**Esimerkki 2.643**

Vaikka tietokonenäköalalla on tapahtunut valtavaa edistystä, koneellista oppimista ei ole yritetty soveltaa hyvin laajoihin lääketieteellisiin kuvatietokantoihin. Esittelemme tekstin ja kuvan välisen syväoppimisjärjestelmän, jonka avulla voidaan poimia ja louhia radiologisten kuvien ja raporttien semanttista vuorovaikutusta kansallisen tutkimussairaalan Picture Archiving and Communication System -järjestelmästä. Luonnollisen kielen prosessoinnin avulla louhimme kokoelman edustavia ∼216K kaksiulotteisia avainkuvia, jotka lääkärit ovat valinneet diagnostisia viitteitä varten, ja sovitamme kuvat ja niiden kuvaukset yhteen automaattisella tavalla. Järjestelmämme toimii vuorovaikutteisesti valvomattoman oppimisen ja valvotun oppimisen välillä dokumentti- ja lausetason tekstikokoelmissa luodaksemme semanttisia merkintöjä ja ennustaaksemme niitä kuvan perusteella. Kun potilaan kuvasta otetaan kuva, radiologian tasojen semanttiset aiheet ennustetaan ja niihin liittyvät avainsanat luodaan. Lisäksi havaitaan useita usein esiintyviä tautityyppejä, jotka ovat läsnä tai puuttuvat, jotta potilaan kuvasta voidaan tehdä tarkempi tulkinta. Tämä osoittaa laajamittaisen oppimisen ja ennustamisen mahdollisuudet sähköisissä potilastiedoissa, jotka ovat saatavilla useimmissa nykyaikaisissa kliinisissä laitoksissa.

**Tulos**

Interleaved Text/Image Deep Mining suuressa radiologisessa tietokannassa automaattista kuvatulkintaa varten.

**Esimerkki 2.644**

Parannamme sellaisten verkko-oppimisalgoritmien laskennallista monimutkaisuutta, jotka vaativat parametrien pienimmän neliösumman regressioestimaattien laskemista usein uudelleen. Ehdotamme kahta stokastista gradienttilaskeutumisjärjestelmää, joissa on satunnaistaminen, jotta regressio-ongelmien todellisia ratkaisuja voidaan seurata tehokkaasti ja saavuttaa O(d) parannus kompleksisuuteen, jossa d on datan ulottuvuus. Ensimmäisessä algoritmissa oletetaan, että regressio-ongelma on vahvasti kupera, ja annamme virheen raja-arvot sekä odotusarvona että suurella todennäköisyydellä (jälkimmäistä tarvitaan usein teoreettisten takuiden antamiseksi korkeamman tason algoritmeille). Toinen algoritmi käsittelee tapauksia, joissa regressio-ongelman vahvaa koveruutta ei voida taata, ja siinä käytetään mukautuvaa regularisointia. Annamme jälleen virherajat sekä odotusarvolle että suurelle todennäköisyydelle. Sovellamme lähestymistapojamme lineaarisiin bandit-algoritmeihin PEGE ja ConfidenceBall ja osoitamme, että molemmissa tapauksissa monimutkaisuus paranee merkittävästi. Koska PEGE-algoritmi takaa vahvan konveksisuuden, menetämme vain logaritmisia kertoimia algoritmin katumissuorituskyvyssä. Toisaalta ConfidenceBall-algoritmissa tehdään mukautuva regularisointi vahvan koveruuden varmistamiseksi, ja tämä johtaa Õ(n1/5)1 :n huononemiseen katumuksessa.

**Tulos**

Online-gradientti laskeutuminen pienimmän neliösumman regressiota varten: Ei-asymptoottiset rajat ja soveltaminen rosvoihin

**Esimerkki 2.645**

Kaksisuuntaiset LSTM-muistit ovat muodostuneet vakiomenetelmäksi, jolla saadaan merkkikohtaisia vektoriedustuksia, joita käytetään syötteenä erilaisissa merkkien merkintätehtävissä (riippumatta siitä, seuraako niitä Viterbi-ennustus vai itsenäinen luokittelu). Tässä artikkelissa ehdotetaan tähän tarkoitukseen vaihtoehtoa Bi-LSTM:ille: iteroituja laajentuneita konvoluutiohermoverkkoja (ID-CNN), joilla on perinteisiä CNN:iä parempi kapasiteetti laajassa kontekstissa ja strukturoidussa ennustamisessa. Kuvaamme verkon rakenteen, parametrien jakamisen ja koulutusmenetelmien erillisen yhdistelmän, joka on paitsi tarkempi kuin Bi-LSTM-CRF:t, myös 8 kertaa nopeampi testiaikana pitkillä sekvensseillä. Lisäksi ID-CNN:t, joissa on itsenäinen luokittelu, mahdollistavat 14-kertaisen nopeutuksen testiajassa, mutta silti Bi-LSTM-CRF:n tarkkuus on verrattavissa Bi-LSTM-CRF:n tarkkuuteen. Osoitamme lisäksi ID-CNN:ien kyvyn yhdistää todisteita pitkien sekvenssien aikana osoittamalla niiden paremman tarkkuuden koko dokumentin (eikä lausekohtaisen) päättelyn yhteydessä. Toisin kuin LSTM-mallit, joiden peräkkäinen käsittely N:n pituisissa lauseissa vaatii O(N)-aikaa jopa rinnakkaisuuden vallitessa, IDCNN-mallit mahdollistavat kiinteäsyvyisten konvoluutioiden suorittamisen rinnakkain koko asiakirjojen yli. Nykyään, kun monet yritykset suorittavat perustason NLP-toimintoja koko verkossa ja suurissa liikennemäärissä, nopeammat menetelmät ovat ensiarvoisen tärkeitä ajan ja energiakustannusten säästämiseksi.

**Tulos**

Nopea ja tarkka sekvenssien merkitseminen iteroitujen dilatoivien konvoluutioiden avulla

**Esimerkki 2.646**

Luonnollisen kielen tuottaminen (NLG) on puhutun dialogin kriittinen osa, ja sillä on merkittävä vaikutus sekä käytettävyyteen että koettuun laatuun. Useimmat yleisesti käytössä olevat NLG-järjestelmät käyttävät sääntöjä ja heuristiikkoja, ja niillä on taipumus tuottaa jäykkiä ja tyyliteltyjä vastauksia ilman ihmiskielen luonnollista vaihtelua. Niitä ei myöskään ole helppo skaalata järjestelmiin, jotka kattavat useita aloja ja kieliä. Tässä artikkelissa esitellään tilastollinen kieligeneraattori, joka perustuu semanttisesti ohjattuun LSTM-rakenteeseen (Long Short-term Memory). LSTM-generaattori voi oppia epäsymmetrisestä datasta optimoimalla yhdessä lauseiden suunnittelun ja pinnan toteutuksen yksinkertaisella ristiinentropian koulutuskriteerillä, ja kielen vaihtelu voidaan helposti saavuttaa ottamalla näytteitä ulostulokandidaateista. Kahdella eri testialueella suoritettu objektiivinen arviointi osoitti, että ehdotettu menetelmä parantaa suorituskykyä aiempiin menetelmiin verrattuna, kun heuristiikkoja on vähemmän. Ihmistuomarit antoivat LSTM-järjestelmälle paremmat pisteet informatiivisuudesta ja luonnollisuudesta ja pitivät sitä yleisesti ottaen parempana kuin muita järjestelmiä.

**Tulos**

Semanttisesti ehdollistettu LSTM-pohjainen luonnollisen kielen tuottaminen puhuttuja dialogijärjestelmiä varten.

**Esimerkki 2.647**

Ehdotamme menetelmää, jolla synteettistä dataa voidaan käyttää luokittelijoiden oppimisen apuna. Synteettinen data, vaikka se luotaisiinkin todellisen datan perusteella, johtaa yleensä siihen, että todellisen datan jakauma muuttuu ominaisuusavaruudessa. Tässä asiakirjassa ehdotetaan monikanavaista automaattista kooderia (Multichannel Autoencoder, MCAE), jolla voidaan kuroa umpeen kuilu todellisen ja synteettisen datan välillä ja oppia yhdessä synteettisestä ja todellisesta datasta. Osoitamme, että MCAE:n avulla on mahdollista oppia parempi ominaisuuksien esitys luokittelua varten. Arvioidaksemme ehdotettua lähestymistapaa teemme kokeita kahdentyyppisillä tietokokonaisuuksilla. Kokeelliset tulokset kahdella tietokokonaisuudella vahvistavat MCAE-mallimme tehokkuuden ja synteettisen datan tuottamismenetelmämme.

**Tulos**

Luokittelijoiden oppiminen synteettisestä datasta monikanavaisen autokooderin avulla

**Esimerkki 2.648**

Valvomaton riippuvuuksien jäsentely, jossa yritetään löytää kielellisiä riippuvuusrakenteita kommentoimattomasta aineistosta, on hyvin haastava tehtävä. Lähes kaikki aiemmat tätä tehtävää koskevat työt keskittyvät generatiivisten mallien oppimiseen. Tässä artikkelissa kehitämme CRF-autokooderiin perustuvan valvomattoman riippuvuuksien jäsentelymallin. Mallimme kooderiosa on diskriminoiva ja globaalisti normalisoitu, minkä ansiosta voimme käyttää monipuolisia piirteitä sekä universaaleja kielellisiä prioreja. Ehdotamme täsmällistä algoritmia jäsennystä varten sekä helposti lähestyttävää oppimisalgoritmia. Arvioimme mallimme suorituskykyä kahdeksalla monikielisellä puupankilla ja havaitsimme, että mallimme saavutti vertailukelpoisen suorituskyvyn uusimpien lähestymistapojen kanssa.

**Tulos**

CRF-autokooderi valvomatonta riippuvuuksien jäsennystä varten∗ \*.

**Esimerkki 2.649**

Ehdotamme säännöllistettyä lineaarista oppimisalgoritmia ominaisuuksien ryhmien järjestämiseen, jossa jokaisesta ryhmästä aiheutuu testiaikaisia kustannuksia tai laskentakustannuksia. Kehitämme yksinkertaisen laajennuksen Orthogonal Matching Pursuit (OMP) -menetelmälle, joka ottaa huomioon muuttuvilla kustannuksilla varustettujen ominaisuuksien ryhmien rakenteen, ja todistamme, että sillä saavutetaan lähes optimaalinen lineaarinen ennuste jokaisella budjettikynnyksellä, jossa valitaan uusi ryhmä. Algoritmimme ja analyysimme on laajennettavissa yleistettyihin lineaarisiin malleihin, joissa on moniulotteisia vastauksia. Osoitamme tuloksena syntyvän lähestymistavan skaalautuvuuden suurilla reaalimaailman datajoukoilla, joissa on monia ominaisuusryhmiä, joihin liittyy testiajan laskentakustannuksia. Menetelmämme parantaa Ryhmä-Lassoa ja Ryhmä-OMP:tä lineaaristen ennusteiden suorituskykyä milloin tahansa mitattuna ajantasaisuudella[7], joka on ennusteiden suorituskyvyn mittari milloin tahansa, ja tarjoaa samalla tiukat suorituskykytakeet.

**Tulos**

Tehokas ominaisuusryhmien sekvensointi milloin tahansa tapahtuvaa lineaarista ennustamista varten

**Esimerkki 2.650**

Seuraavassa kuvataan lyhyesti todennäköisyysgraafimalli, jonka avulla voidaan esittää kausaalirakenteita, päättellä niitä ja oppia niistä Bayes-verkkojen avulla. Tämän jälkeen väitetään, että tämä malli liittyy läheisesti siihen, miten ihmiset päättelevät kausaalirakenteista ja oppivat niistä. Osoitetaan, että psykologian tutkimukset diskonttaamisesta (päättely, joka koskee sitä, miten yhden vaikutuksen syyn olemassaolo tekee toisen syyn vähemmän todennäköiseksi) tukevat hypoteesia, jonka mukaan ihmiset päätyvät samoihin arvioihin kuin algoritmit, jotka tekevät Bayes-verkoissa päättelyä. Seuraavaksi osoitetaan, miten Piaget'n tutkimukset osoittavat, että ihminen oppii kausaalirakenteen havainnoimalla samoja riippumattomuuksia ja riippuvuuksia kuin mitä tietyt algoritmit käyttävät oppiessaan Bayes-verkon rakennetta. Tämän perusteella esitetään kausaalisuuden subjektiivinen määritelmä. Lopuksi käsitellään menetelmiä, joilla näiden väitteiden paikkansapitävyyttä voidaan testata.

**Tulos**

SYYLLISEN TIEDON YHTEISKUNNALLINEN KÄSITTELEMINEN

**Esimerkki 2.651**

Ketjugraafit yhdistävät suunnattuja ja suuntaamattomia graafeja, ja niiden taustalla oleva matematiikka yhdistää näiden kahden ominaisuuksia. Tässä artikkelissa esitetään ketjuharavainten yksinkertaistettu määritelmä, joka perustuu Bayesin (suunnattujen) ja Markovin (suuntaamattomien) verkkojen hierarkkiseen yhdistelmään. Esimerkkejä ketjugraafeista ovat monimuuttujaiset feed-forward-verkot, klusterit, joissa muuttujien välillä on ehdollinen vuorovaikutus, ja Bayes-luokittelijoiden muodot. Ketjugraafeja laajennetaan sitten levyjen notaation avulla, jotta myös näytteet ja data-analyysin ongelmat voidaan esittää graafisessa mallissa. Johtopäätöksessä käsitellään vaikutuksia oppimiseen.

**Tulos**

Ketjugraafit oppimista varten

**Esimerkki 2.652**

Lähestymme haastavaa ongelmaa, joka liittyy urheilulähetysten kohokohtien tuottamiseen pelkkää äänitietoa hyödyntäen. Keskeisten akustisten tapahtumien havaitsemiseen käytetään kielestä riippumatonta, monivaiheista luokittelua, joka toimii pohjana kohokohtien tiivistämiselle. Objektiiviset tulokset ja ihmisten kokemukset osoittavat, että järjestelmämme on erittäin tehokas.

**Tulos**

Urheilun kohokohtien tuottaminen akustisten tapahtumien havaitsemisen perusteella: Rugbyn tapaustutkimus

**Esimerkki 2.653**

Ehdotamme mustan laatikon variationaalista päättelymenetelmää, jolla voidaan approksimoida vaikeasti lähestyttäviä jakaumia yhä monipuolisemmalla approksimointiluokalla. Menetelmämme, jota kutsutaan nimellä variational boosting, tarkentaa iteratiivisesti olemassa olevaa variationaalista approksimaatiota ratkaisemalla sarjan optimointiongelmia, jolloin ammattilainen voi vaihtaa laskenta-ajan tarkkuuteen. Näytämme, miten variatiivista approksimointiluokkaa voidaan laajentaa sisällyttämällä siihen lisää kovarianssirakennetta ja ottamalla käyttöön uusia komponentteja seoksen muodostamiseksi. Sovellamme variationaalista tehostamista synteettisiin ja todellisiin tilastollisiin malleihin ja osoitamme, että tuloksena saadut posterioriset johtopäätökset ovat sekä tarkkuudeltaan että tehokkuudeltaan edullisia verrattuna olemassa oleviin posteriorisiin approksimaatioalgoritmeihin.

**Tulos**

Variational Boosting: Iteratiivisesti tarkentavat posterioriset lähestymiset

**Esimerkki 2.654**

Tässä artikkelissa tutkitaan konvoluutio-neuraaliverkkojen soveltamista still-kuvissa olevien kohteiden havaitsemiseen ja niiden 3D-asennon arviointiin. Tunnistamme suunnattujen objektien erilaiset ominaisuuksien representaatiot ja energiat, jotka johtavat verkon oppimaan nämä representaatiot. Edustuksen valinta on ratkaisevan tärkeää, koska objektin asento on luonteva, jatkuva rakenne, kun taas sen luokka on diskreetti muuttuja. Arvioimme eri lähestymistapoja Pascal3D+-vertailulaskennan yhteisessä objektien havaitsemis- ja asennonarviointitehtävässä käyttäen keskimääräistä Viewpoint Precision -tarkkuutta. Osoitamme, että diskretoituja näkymäpisteitä käyttävä luokitusmenetelmä saavuttaa huippuluokan suorituskyvyn yhteisessä objektin havaitsemisessa ja asennon arvioinnissa ja että se on huomattavasti parempi kuin nykyiset perusratkaisut tässä vertailuarvossa. Osoitamme myös, että näiden kahden tehtävän suorittaminen yhdessä voi parantaa havaintosuorituksia merkittävästi.

**Tulos**

VERTAILEVA TUTKIMUS

**Esimerkki 2.655**

Tarkastelemme tilastollisia ja optimointivaikutuksia, joita on klassisen luonnoksen ja Hessin luonnoksen käyttämisellä matriisiharjuregressio-ongelman (MRR) ratkaisemiseen. Aikaisemmassa tutkimuksessa on tarkasteltu klassisen luonnoksen vaikutuksia pienimmän neliösumman regressioon (LSR), joka on huomattavasti yksinkertaisempi ongelma. Todetaan, että klassisella luonnostelulla on samanlainen vaikutus MRR:n optimointiominaisuuksiin kuin LSR:n ominaisuuksiin - nimittäin se palauttaa lähes optimaaliset ratkaisut. Sen sijaan Hessin luonnoksella ei ole tätä takuuta, vaan approksimointivirheeseen vaikuttaa vastausten "massan" ja optimaalisen tavoitearvon välinen hienovarainen vuorovaikutus. Molempien approksimaatiotyyppien osalta luonnostellun MRR-ongelman regularisointi antaa sille huomattavasti erilaisia tilastollisia ominaisuuksia kuin luonnostellulle LSR-ongelmalle. Luonnostellussa MRR-ongelmassa on erityisesti harha-varianssin kompromissi, jota ei ole luonnostellussa LSR-ongelmassa. Annamme ylä- ja alarajat luonnostellun MRR-ongelman harhojen ja varianssien arvoille; näiden arvojen mukaan varianssi kasvaa merkittävästi, kun käytetään klassisia luonnoksia, kun taas harha kasvaa merkittävästi, kun käytetään Hessin luonnoksia. Empiirisesti voidaan todeta, että luonnosteltujen MRR-ratkaisujen riskit voivat olla kertaluokkaa suuremmat kuin optimaalisten MRR-ratkaisujen riskit. Todistamme teoreettisesti ja empiirisesti, että mallien keskiarvoistaminen pienentää tätä eroa huomattavasti. Näin ollen hajautetussa ympäristössä luonnostelu yhdistettynä mallin keskiarvoistamiseen on tehokas tekniikka, jolla saadaan nopeasti lähes optimaalisia ratkaisuja MRR-ongelmaan ja samalla vähennetään huomattavasti luonnostelusta aiheutuvia tilastollisia riskejä.

**Tulos**

Luonnosteltu harjuregressio: Optimointinäkökulma, tilastollinen näkökulma ja mallin keskiarvoistaminen: Optimointinäkökulma, tilastollinen näkökulma ja mallin keskiarvoistaminen

**Esimerkki 2.656**

Tässä artikkelissa esitellään Bayesin optimointimenetelmä, jonka konvergenssi on eksponentiaalinen ilman apuoptimointia ja ilman δ-peittävää näytteenottoa. Useimmat bayesiläiset optimointimenetelmät vaativat apuoptimointia: ylimääräisen ei-konveksaalisen globaalin optimointiongelman, joka voi olla aikaa vievä ja käytännössä vaikea toteuttaa. Myös olemassa oleva eksponentiaalisen konvergenssin omaava bayesilainen optimointimenetelmä [1] edellyttää pääsyä δ-kattavuusnäytteenottoon, jota pidettiin epäkäytännöllisenä [1, 2]. Meidän lähestymistapamme poistaa molemmat vaatimukset ja saavuttaa eksponentiaalisen konvergenssinopeuden.

**Tulos**

Bayesilainen optimointi eksponentiaalisella konvergenssilla

**Esimerkki 2.657**

Elokuva-arvosteluilla on tärkeä rooli sekä sen määrittämisessä, kuinka todennäköisesti potentiaalinen katsoja katsoo elokuvan, että sen määrittämisessä, kuinka tyytyväisiä katsojat ovat elokuvaan tällä hetkellä. Niitä on saatavilla useista lähteistä, kuten televisio-oppaasta, myydyimmistä tietokirjoista, sanomalehtien kolumneista ja televisio-ohjelmista. Lisäksi elokuva-arviot ovat ratkaisevan tärkeitä suosittelumoottoreille, jotka seuraavat kaikkien käyttäjien käyttäytymistä ja hyödyntävät tietoja ehdottaakseen kohteita, joista he saattaisivat pitää. Elokuva-arviot tarjoavat siis useimmissa tapauksissa tietoa, joka saattaa olla tärkeämpää kuin elokuvan ominaisuuksiin perustuva tieto. On intuitiivisesti houkuttelevaa, että tieto elokuvagenreihin liittyvistä katselumieltymyksistä riittää ennustamaan leimaamattoman elokuvan genren. Elokuvan genrejen ennustamiseksi käsittelemme katselukertoja ominaisuusvektorina, sovellamme Bernoulli-tapahtumamallia elokuvan tietyn genren todennäköisyyden arvioimiseksi ja arvioimme tietyn elokuvan genren jälkitodennäköisyyttä Bayesin säännön avulla. Ehdotetun tekniikan tavoitteena on käyttää tehokkaasti elokuvien arvosanoja elokuvan genren ennustamiseen. Lähestymistavassamme pyrimme vastaamaan kysymykseen: "Kun otetaan huomioon elokuvan katsoneiden käyttäjien joukko, onko mahdollista ennustaa elokuvan genre sen arvosanojen perusteella?"." Simulointituloksemme MovieLens 100k -aineistolla osoittivat ehdotetun tekniikkamme tehokkuuden ja tarkkuuden, sillä täsmällisessä ennusteessa saavutettiin 59 prosentin ennusteprosentti ja 69 prosentin ennusteprosentti, kun mukaan otettiin myös korreloivat genret.

**Tulos**

Monimuuttuja-Bernoulli-malliin ja genrekorrelaatioihin perustuva elokuvan genren ennuste.

**Esimerkki 2.658**

Tutkimme hermosekvenssimallien tehokkuutta premissien valinnassa automaattisessa lauseiden todistamisessa, joka on yksi tärkeimmistä pullonkauloista matematiikan formalisoinnissa. Ehdotamme tähän tehtävään kaksivaiheista lähestymistapaa, joka tuottaa hyviä tuloksia premissien valintatehtävässä Mizar-korpuksessa ja välttää samalla nykyisten huippumallien käsin kehitettyjä ominaisuuksia. Tietojemme mukaan tämä on ensimmäinen kerta, kun syväoppimista on sovellettu teoreemojen todistamiseen.

**Tulos**

DeepMath - Syvät sekvenssimallit lähtökohtien valintaa varten

**Esimerkki 2.659**

Tässä raportissa esitellään Giraffe, shakkimoottori, joka käyttää itsepeliä löytääkseen kaiken aluespesifisen tietämyksensä ja jossa ohjelmoijan käsin antaman tiedon määrä on minimaalinen. Toisin kuin aiemmissa yrityksissä, joissa koneoppimista on käytetty vain parametrien virittämiseen käsin tehdyille arviointifunktioille, Giraffen oppimisjärjestelmä suorittaa myös automaattisen ominaisuuksien louhinnan ja hahmontunnistuksen. Koulutettu arviointitoiminto toimii verrattavissa huipputason shakkimoottoreiden arviointitoimintoihin, jotka kaikki sisältävät tuhansia rivejä huolellisesti käsin kehitettyjä hahmontunnistimia, joita sekä tietokoneen shakkiasiantuntijat että ihmisen shakkimestarit ovat virittäneet useiden vuosien ajan. Giraffe on tähän mennessä onnistunein yritys käyttää kokonaisvaltaista koneoppimista shakin pelaamiseen. Tutkimme myös mahdollisuutta käyttää todennäköisyyskynnyksiä syvyyden sijasta hakupuiden muotoiluun. Syvyyteen perustuvat haut muodostavat käytännössä kaikkien nykyään käytössä olevien shakkimoottoreiden selkärangan, ja se on algoritmi, joka on vakiintunut viimeisen puolen vuosisadan aikana. Alustavat vertailut todennäköisyyspohjaisen haun perustoteutuksen ja syvyyspohjaisen haun perustoteutuksen välillä osoittivat, että uusi todennäköisyyspohjainen lähestymistapamme toimii hieman paremmin kuin vakiintunut lähestymistapa. On myös todisteita siitä, että monet onnistuneet ad hoc -lisäykset syvyyspohjaisiin hakuihin yleistyvät siirtymällä todennäköisyyspohjaiseen hakuun. Uskomme, että todennäköisyyspohjainen haku on perustaltaan oikeampi tapa suorittaa minimax. Lopuksi suunnittelimme toisen koneoppimisjärjestelmän hakupuiden muokkaamiseen todennäköisyyspohjaisen haun puitteissa. Tämä järjestelmä arvioi minkä tahansa aseman ollessa kyseessä kunkin siirron todennäköisyyden olla paras siirto ilman etukäteisarviointia. Järjestelmä on erittäin tehokas, sillä todellinen paras siirto on kolmen parhaan joukossa 70 prosenttia ajasta, kun kustakin asemasta tehdään keskimäärin noin 35 laillista siirtoa. Tämä johti myös pelivoiman merkittävään kasvuun. Kun siirtojen arvioija ohjaa todennäköisyyteen perustuvaa hakua opitun arvioijan avulla, Giraffe pelaa suunnilleen FIDE:n kansainvälisen mestarin tasolla (2,2 % parhaista turnausshakinpelaajista, joilla on virallinen luokitus)12 . Fédération Internationale des Échecs eli Maailman shakkiliitto on kansainvälinen järjestö, joka hallinnoi kaikkia merkittäviä kansainvälisiä shakkikilpailuja. Elo-luokitusjärjestelmän kuvaus on liitteessä A.

**Tulos**

Kirahvi: Deep Reinforcement Learning -oppimisen käyttö shakin pelaamisessa

**Esimerkki 2.660**

Tutkimme optimointialgoritmeja, jotka perustuvat varianssin vähentämiseen stokastista gradienttilaskeutumista (SGD) varten. Tällä alalla on viime aikoina edistytty huomattavasti kehittämällä algoritmeja, kuten SAG, SVRG ja SAGA. Näiden algoritmien on osoitettu päihittävän SGD:n sekä teoreettisesti että empiirisesti. Näiden algoritmien asynkronisia versioita, jotka ovat nykyaikaisten laajamittaisten sovellusten keskeinen vaatimus, ei ole kuitenkaan tutkittu. Me kuromme tämän aukon umpeen esittelemällä yhdistävän kehyksen monille varianssin pienentämistekniikoille. Tämän jälkeen ehdotamme asynkronista algoritmia, joka perustuu kehykseemme, ja todistamme sen nopean konvergenssin. Yleisen lähestymistapamme tärkeä seuraus on, että sen sivutuotteena saadaan aikaan asynkronisia versioita varianssin pienentämisalgoritmeista, kuten SVRG ja SAGA. Menetelmämme saavuttaa lähes lineaarisen nopeutuksen koneoppimisessa yleisesti käytetyissä harvoissa asetuksissa. Osoitamme menetelmämme empiirisen suorituskyvyn asynkronisen SVRG:n konkreettisen toteutuksen avulla.

**Tulos**

Varianssin pienentämisestä stokastisessa gradienttilaskeutumisessa ja sen asynkronisissa muunnelmissa

**Esimerkki 2.661**

Monimutkaisten tehtäväkokonaisuuksien ratkaisemisen oppiminen niin, että samalla hyödynnetään siirtoa ja vältetään katastrofaalinen unohtaminen, on edelleen keskeinen este ihmisen älykkyyden saavuttamiselle. Progressiiviset verkot edustavat edistysaskelta tähän suuntaan: ne ovat immuuneja unohtamiselle ja voivat hyödyntää aiempaa tietämystä aiemmin opittujen ominaisuuksien sivuttaisten yhteyksien kautta. Arvioimme tätä arkkitehtuuria laajasti erilaisissa vahvistusoppimistehtävissä (Atari- ja 3D-silmäkulmapelit) ja osoitamme, että se on parempi kuin tavalliset esivalmennukseen ja hienosäätöön perustuvat perusratkaisut. Käyttämällä uutta herkkyysmittaria osoitamme, että siirtoa tapahtuu sekä matalan tason aisti- että korkean tason ohjauskerroksissa opitun politiikan osalta.

**Tulos**

Progressiiviset neuroverkot

**Esimerkki 2.662**

Etävalvotun relaatioiden louhinnan ydin on se, että se on epätäydellinen monimerkkiluokitusongelma, jossa ominaisuudet ovat harvalukuisia ja kohinaisia. Harvaan ja kohinaan liittyvien haasteiden ratkaisemiseksi ehdotamme luokitusongelman ratkaisemista käyttämällä matriisitäydennystä minimoidun sijan faktoroidulla matriisilla. Muotoilemme relaatioiden luokittelun siten, että testattavien kohteiden (olioparien) tuntemattomat merkinnät täydennetään harvalukuiseen matriisiin, joka yhdistää koulutus- ja testattavat tekstiominaisuudet sekä koulutustunnisteet. Algoritminen kehyksemme perustuu oletukseen, että elementtikohtaisen ominaisuuden ja elementtikohtaisen merkinnän yhteisen matriisin arvo on alhainen. Sovellamme kahta optimointimallia matalan matriisin palauttamiseksi hyödyntäen ominaisuustietomatriisin harvuutta. Matriisin täydentämisongelma ratkaistaan sitten FPC-algoritmilla (fixed point continuation), jolla voidaan löytää globaali optimi. Kokeet kahdella laajalti käytetyllä tietokokonaisuudella, joissa on tekstiominaisuuksien eri ulottuvuuksia, osoittavat, että matalarivinen matriisitäydennysmenetelmämme on huomattavasti parempi kuin perusmenetelmä ja uusimmat menetelmät.

**Tulos**

Errata: Distant Supervision for Relation Extraction with Matrix Completion (Etävalvonta suhteiden poimimisessa matriisien täydentämisellä).

**Esimerkki 2.663**

Webillä on tärkeä rooli ihmisten sosiaalisessa elämässä Web 2.0:n syntymisen jälkeen. Se helpottaa käyttäjien välistä vuorovaikutusta ja antaa heille mahdollisuuden olla vapaasti vuorovaikutuksessa, jakaa ja tehdä yhteistyötä sosiaalisten verkostojen, verkkoyhteisöjen foorumeiden, blogien, wikien ja muiden verkossa toimivien yhteistyövälineiden kautta. Verkkoon liittyy kuitenkin myös kielteisiä piirteitä, kuten kiihottavien viestien lähettäminen. Näin ollen verkkoyhteisöjen foorumeja käsitellessään johtajat pyrkivät aina parantamaan tällaisten foorumien suorituskykyä. Säilyttääkseen rauhallisuuden ja estääkseen normaalin ilmapiirin häiritsemisen johtajat yrittävät aina suojella käyttäjiä pahansuovilta henkilöiltä lähettämällä tällaisen viestin (DO NOT FEED TROLLS). Tällainen varoitus ei kuitenkaan riitä vähentämään tätä ilmiötä. Tässä yhteydessä ehdotamme uutta lähestymistapaa, jonka avulla voidaan havaita pahantahtoiset henkilöt, joita kutsutaan myös "trolleiksi", jotta yhteisön johtajat voivat ottaa näiltä henkilöiltä mahdollisuuden lähettää viestejä verkossa. Jotta ehdotus olisi realistisempi, se on määritelty epävarmoissa puitteissa. Perustuen oletukseen, joka perustuu trollien integroitumiseen onnistuneisiin keskusteluketjuihin, yritämme havaita tällaisten pahansuopaisten käyttäjien läsnäolon. Menetelmä perustuu uskomusfunktioteorian mukaiseen konfliktimittaukseen, jota sovelletaan keskusteluketjun eri viestien välillä. Jotta voimme osoittaa lähestymistapamme toteutettavuuden ja tulokset, testaamme sitä erilaisilla simuloiduilla aineistoilla. Avainsanat: Q&AC, trollit, uskomusfunktioteoria, konfliktimitta.

**Tulos**

Peikkojen tunnistaminen epävarmoissa puitteissa

**Esimerkki 2.664**

Mallintarkastaja voi tuottaa virheelliselle ohjelmalle vastakappaleen, joka on usein pitkä ja vaikeasti ymmärrettävä. Yleensä silmukoita koskeva osa on tämän jäljen ohjeista suurin. Tämä tekee silmukoissa olevien virheiden paikantamisesta kriittisen tärkeää koko ohjelman virheiden analysoimiseksi. Tässä artikkelissa tutkimme LocFaultsin skaalautuvuusominaisuuksia, virheiden paikannusmenetelmäämme, jossa hyödynnetään CFG:n (Control Flow Graph) polkuja vastaesimerkistä MCD:n (Minimal Correction Deviations) ja MCS:n (Minimal Correction Subsets) laskemiseksi kustakin löydetystä MCD:stä. Esitämme lähestymistapamme ajat ohjelmille, joissa While-silmukoita on käännetty b kertaa ja joissa poikkeavien ehtojen määrä vaihtelee 0:sta n:ään. Alustavat tuloksemme osoittavat, että lähestymistapamme rajoituksiin perustuvan ja virtauslähtöisen lähestymistavan ajat ovat paremmat verrattuna BugAssistiin, joka perustuu SAT-menetelmään ja muuttaa koko ohjelman Boolen kaavaksi, vaikka LocFaultsin tarjoamat tiedot ovatkin käyttäjän kannalta ilmaisuvoimaisempia.

**Tulos**

LocFaultsin skaalautuvuuden tutkiminen

**Esimerkki 2.665**

Luottorikkojien ennustaminen on pankkien kaltaisten rahoitusalojen vaarallinen tehtävä. Maksukyvyttömyyden toteaminen ennen lainan myöntämistä on pankkiirille merkittävä ja ristiriitainen tehtävä. Luokittelutekniikat ovat parempi valinta ennakoivaan analyysiin, kuten sen selvittämiseen, onko hakija vaatimaton asiakas vai huijari. Erinomaisen luokittelijan määrittäminen on riskialtis tehtävä mille tahansa teollisuusyrittäjälle, kuten pankkiirille. Tämän ansiosta tietojenkäsittelytieteen tutkijat voivat tehdä tehokkaita tutkimustöitä arvioimalla erilaisia luokittelijoita ja löytämällä parhaan luokittelijan tällaisiin ennakointiongelmiin. Tässä tutkimustyössä tutkitaan LADTree- ja REPTree-luokittelijoiden tuottavuutta luottoriskin ennustamisessa ja verrataan niiden soveltuvuutta eri toimenpiteiden avulla. Saksan luottotietoaineisto on otettu ja sitä on käytetty luottoriskin ennustamiseen avoimen lähdekoodin koneoppimistyökalun avulla.

**Tulos**

LADTREE- JA REPTREE-LUOKITTELIJOIDEN TEHOKKUUSVERTAILU LUOTTORISKIN ENNUSTAMISESSA

**Esimerkki 2.666**

Automaattinen kirjoittamisen arviointi (AWE) on osoittautunut tehokkaaksi mekanismiksi, jolla opiskelijoille voidaan antaa nopeasti palautetta. Se on jo laajalti otettu käyttöön yritystason sovelluksissa, ja sitä aletaan ottaa käyttöön laajamittaisissa yhteyksissä. AWE-mallin kouluttaminen on perinteisesti edellyttänyt useiden satojen kirjoitusesimerkkien erää ja ihmisen antamia pisteitä jokaiselle esimerkille. Tämä vaatimus rajoittaa AWE:n laajamittaista käyttöönottoa, koska esseiden pisteyttäminen ihmisen toimesta on kallista. Tässä arvioidaan algoritmeja, joilla varmistetaan, että AWE-mallit koulutetaan johdonmukaisesti käyttäen kaikkein informatiivisimpia esseitä. Tuloksemme osoittavat, miten harjoitusjoukkojen kokoa voidaan minimoida ja samalla maksimoida ennustuskyky, jolloin kustannukset pienenevät ilman, että tarkkuus kärsii kohtuuttomasti. Lopuksi keskustelemme siitä, miten tämä lähestymistapa voidaan integroida laajamittaisiin AWE-järjestelmiin.

**Tulos**

Tehokas otanta laajamittaisia automaattisia kirjoitusten arviointijärjestelmiä varten

**Esimerkki 2.667**

Muunnettu politiikan iterointi (MPI) on dynaamisen ohjelmoinnin (DP) algoritmi, joka sisältää kaksi tunnettua politiikan ja arvon iterointimenetelmää. Yleisyydestään huolimatta MPI:tä ei ole tutkittu perusteellisesti, varsinkaan sen approksimaatiomuotoa, jota käytetään, kun tila- ja/tai toiminta-avaruudet ovat suuria tai äärettömiä. Tässä artikkelissa ehdotamme kolmea likimääräisen MPI:n (AMPI) toteutusta, jotka ovat tunnettujen likimääräisten DP-algoritmien laajennuksia: fitted-value-iteraatio, fitted-Q-iteraatio ja luokitteluun perustuva politiikka-iteraatio. Tarjoamme virheiden etenemisanalyysejä, jotka yhdistävät likimääräisen politiikan ja arvo-iteraation analyysit. Viimeisen luokitteluun perustuvan toteutuksen osalta kehitämme äärellisten otosten analyysin, joka osoittaa, että MPI:n pääparametrin avulla voidaan hallita tasapainoa luokittelijan estimointivirheen ja yleisen arvofunktion approksimaation välillä.

**Tulos**

Likimääräinen modifioidun politiikan iterointi

**Esimerkki 2.668**

Feature squeezing on hiljattain käyttöön otettu kehys, jonka avulla voidaan lieventää ja havaita haitallisia esimerkkejä. Aiemmassa työssä osoitimme, että se on tehokas useita aiempia menetelmiä vastaan, joilla luodaan haitallisia esimerkkejä. Tässä lyhyessä muistiossa raportoimme tuoreista tuloksista, jotka osoittavat, että yksinkertaiset feature squeezing -tekniikat tekevät syväoppimismalleista huomattavasti kestävämpiä myös Carlini/Wagner-hyökkäyksiä vastaan, jotka ovat tähän mennessä tunnetuimpia löydettyjä vastakohtaisia menetelmiä.

**Tulos**

Ominaisuuksien puristaminen lieventää ja havaitsee Carlini/Wagnerin vastakkaisia esimerkkejä.

**Esimerkki 2.669**

Kyky esittää lauseita tarkasti on keskeinen osa kielen ymmärtämistä. Kuvaamme konvoluutioarkkitehtuurin, jota kutsumme dynaamiseksi konvoluutiohermoverkoksi (DCNN) ja jota käytämme lauseiden semanttiseen mallintamiseen. Verkko käyttää dynaamista k-Max Poolingia, joka on lineaaristen sekvenssien globaali pooling-operaatio. Verkko käsittelee eri pituisia lauseita ja luo lauseen päälle ominaisuuksien kuvaajan, joka pystyy kuvaamaan eksplisiittisesti lyhyen ja pitkän kantaman suhteita. Verkko ei ole riippuvainen jäsennyspuusta, ja sitä voidaan helposti soveltaa mihin tahansa kieleen. Testaamme DCNN:ää neljässä kokeessa: pienen mittakaavan binäärinen ja moniluokkainen tunteiden ennustaminen, kuusisuuntainen kysymysten luokittelu ja Twitter-tunnelmien ennustaminen etävalvonnan avulla. Verkko saavuttaa erinomaisen suorituskyvyn kolmessa ensimmäisessä tehtävässä ja yli 25 prosentin virheen vähennyksen viimeisessä tehtävässä verrattuna vahvimpaan perustasoon.

**Tulos**

Konvolutiivinen neuroverkko lauseiden mallintamiseen

**Esimerkki 2.670**

Esittelemme ohjelmistotyökalun, jossa käytetään uusinta luonnollisen kielen käsittelyä (NLP) ja koneoppimistekniikoita auttamaan sanomalehtien toimittajia laatimaan tehokkaita otsikoita verkkojulkaisua varten. Järjestelmä tunnistaa uutisartikkelin merkittävimmät avainsanat ja asettaa ne paremmuusjärjestykseen sekä yleisen suosion että artikkelin kannalta suoran merkityksen perusteella. Järjestelmä käyttää myös valvottua regressiomallia tunnistamaan otsikot, joita todennäköisesti jaetaan laajasti sosiaalisessa mediassa. Käyttöliittymä on suunniteltu yksinkertaistamaan ja nopeuttamaan toimittajan päätöksentekoprosessia otsikon koostamisesta. Näin ollen työkalu tarjoaa tehokkaan tavan yhdistää automaattisten sitoutumista ja hakukoneoptimointia (SEO) ennustavien tekijöiden hyödyt ihmisen tekemiin arvioihin otsikon kokonaislaadusta.

**Tulos**

Auttaa uutistoimittajia kirjoittamaan parempia otsikoita: Uutisotsikoiden avainsanasisällön ja jaettavuuden parantamiseen tarkoitettu suosittelija.

**Esimerkki 2.671**

Nopeiden verkkojen jatkuva kasvu on haaste reaaliaikaisille verkkoanalyysijärjestelmille. Liikenteen reaaliaikainen luokittelu on yritysten ja Internet-palveluntarjoajien (ISP) ongelma. Tässä työssä esitellään reaaliaikaisen virtausperusteisen verkkoliikenteen luokittelujärjestelmän suunnittelu ja toteutus. Luokitusmonitori toimii putkistona, joka koostuu kolmesta moduulista: pakettien kaappaus ja esikäsittely, virran uudelleen kokoaminen ja luokittelu koneoppimisen (ML) avulla. Moduulit on rakennettu rinnakkaisiksi prosesseiksi, joiden väliset tietorajapinnat on määritelty tarkasti, jotta mitä tahansa moduulia voidaan parantaa ja päivittää itsenäisesti. Tässä putkistossa virran uudelleen kokoamisesta tulee suorituskyvyn pullonkaula. Tässä toteutuksessa käytettiin tehokasta uudelleenkokoamismenetelmää, jonka tuloksena keskimääräinen toimitusviive on noin 0,49 sekuntia. Luokittelumoduulin osalta K-Nearest Neighbor (KNN), C4.5 Decision Tree, Naive Bayes (NB), Flexible Naive Bayes (FNB) ja AdaBoost Ensemble Learning Algorithm -algoritmien suorituskykyä verrataan lähestymistavan validoimiseksi.

**Tulos**

ITCM: REAALIAIKAINEN INTERNET-LIIKENNE

**Esimerkki 2.672**

Tasoitettu analyysi on kehys algoritmin monimutkaisuuden analysoimiseksi, ja se toimii siltana keskimääräisen ja pahimman tapauksen käyttäytymisen välillä. Esimerkiksi Quicksort-algoritmia ja Simplex-algoritmia käytetään laajalti käytännön sovelluksissa, vaikka niiden pahimman tapauksen monimutkaisuus on suuri. Tasoitetun kompleksisuuden avulla pyritään kuvaamaan paremmin tällaisia algoritmeja. Nykyiset teoreettiset rajat lajittelualgoritmien tasoitetulle monimutkaisuudelle ovat vielä melko heikkoja. Lisäksi tasoitetun kompleksisuuden laskeminen empiirisesti sen alkuperäisen määritelmän avulla on laskennallisesti mahdoton tehtävä, jopa vaatimattomilla syötemäärillä. Tässä artikkelissa keskitymme lajittelualgoritmien tasoitetun monimutkaisuuden tarkkaan ennustamiseen koneoppimistekniikoiden avulla. Ehdotamme kahta regressiomallia, joissa otetaan huomioon lajittelualgoritmien erilaisia ominaisuuksia ja joitakin tasoitetun analyysin tunnettuja teoreettisia tuloksia ennusteen laadun parantamiseksi. Esitämme kokeellisia tuloksia Quicksort-, Mergesort- ja optimoidun Bubblesort-lajittelun tasoitetun monimutkaisuuden ennustamisesta suurille syötekokoluokille ja täytämme siten tunnettujen teoreettisten ja empiiristen tulosten välisen aukon.

**Tulos**

Koneoppimisen lähestymistapa lajittelualgoritmien tasoitetun monimutkaisuuden ennustamiseen

**Esimerkki 2.673**

Esittelemme alustavia ajatuksia simulointiin perustuvasta ohjelmointiparadigmasta, joka on suunnattu tekoälysovelluksiin. Lähestymistavassa pyritään integroimaan tekoälyn eri alojen tekniikoita, ja se perustuu ajatukseen, että simuloidut oliot voivat vapaasti vaihtaa tietoja ja käyttäytymismalleja. Määrittelemme simulointiin perustuvan ohjelmointiparadigman peruskäsitteet ja osoitamme, miten sitä voidaan käyttää tekoälysovellusten toteuttamiseen.

**Tulos**

Kohti simulointipohjaista ohjelmointiparadigmaa tekoälysovelluksia varten

**Esimerkki 2.674**

Älykkään sähköverkon hyökkäysten havaitsemisongelmat esitetään tilastollisina oppimisongelmina erilaisissa hyökkäysskenaarioissa, joissa mittauksia havainnoidaan erä- tai online-olosuhteissa. Tässä lähestymistavassa käytetään koneoppimisalgoritmeja mittausten luokittelemiseksi joko turvallisiksi tai hyökkäyksen kohteiksi. Hyökkäysten havaitsemiseen tarjotaan kehys, jonka avulla voidaan hyödyntää kaikkea saatavilla olevaa ennakkotietoa järjestelmästä ja voittaa ongelman harvasta rakenteesta johtuvat rajoitteet ehdotetussa lähestymistavassa. Tunnettuja erä- ja verkko-oppimisalgoritmeja (valvottuja ja puolivalvottuja) käytetään päätös- ja ominaisuustason fuusioinnin kanssa hyökkäysten havaitsemisongelman mallintamiseen. Hyökkäysskenaarioissa käytettyjen hyökkäysvektorien tilastollisten ja geometristen ominaisuuksien ja oppimisalgoritmien välisiä suhteita analysoidaan, jotta havaitsemattomat hyökkäykset voidaan havaita tilastollisten oppimismenetelmien avulla. Ehdotettuja algoritmeja tutkitaan erilaisissa IEEE:n testijärjestelmissä. Kokeelliset analyysit osoittavat, että koneoppimisalgoritmeilla voidaan havaita hyökkäyksiä paremmalla suorituskyvyllä kuin hyökkäysten havaitsemisalgoritmeilla, jotka käyttävät ehdotetussa hyökkäysten havaitsemiskehyksessä tilavektorien estimointimenetelmiä.

**Tulos**

Koneoppimismenetelmät hyökkäysten havaitsemiseksi älykkäässä sähköverkossa

**Esimerkki 2.675**

Tässä artikkelissa tutkitaan uutta ja yleisempää aksiomatisaatiota kuin [6] esitelty aksiomatisaatio todennäköisyyspeleihin liittyvälle preferenssille. Likelihood-vedonlyönnit kuvaavat toimintaa tilanteessa, jossa päätöksentekijä tuntee useita todennäköisyysmalleja ja yhdestä näistä malleista generoidun satunnaisotoksen, mutta ei tunne mallien ennakkotodennäköisyyttä. Tämä uusi aksioomajärjestelmä on saanut inspiraationsa Jensenin aksiomatisoinnista todennäköisyyspelien osalta. Lähestymistapamme tarjoaa uuden näkökulman tiedon rooliin päätöksenteossa monitulkintaisuuden vallitessa. 1 Todennäköisyyspelit Todennäköisyyspelit, jotka esiteltiin [5, 6], kuvaavat toimintaa mallien epäselvyyden tilanteessa, jolle on ominaista, että (1) on olemassa useita todennäköisyysmalleja, (2) on olemassa dataa, joka antaa malleille todennäköisyyksiä, ja (3) ei ole olemassa ennakkotodennäköisyyttä malleista. Muodollisesti tarkastelemme yleistä ongelmaa, jota kuvataan tuplalla (X,Y,Θ,A,x). X,Y ovat muuttujia, jotka kuvaavat kiinnostavaa ilmiötä. X on kokeilumuuttuja, jonka arvoja voidaan havainnoida kokeiden tai tiedonkeruun avulla (esim. laboratoriotestien tulokset, kliiniset havainnot). Y on hyödyllisyysmuuttuja, jonka arvot määrittävät toimien hyödyllisyyden (esim. taudin vaiheet, kasvaimen suhteellinen koko). Θ on joukko malleja, jotka koodaavat tietoa ilmiöstä. Tarkemmin sanottuna Θ on joukko indeksejä ja tieto on koodattu todennäköisyysfunktioihin Prθ(X,Y ) for θ ∈ Θ. A on joukko vaihtoehtoisia toimia (esim. leikkaus, sädehoito, solunsalpaajahoito), jotka ovat funktioita hyötymuuttujasta Y yksikköväliin [0, 1], joka edustaa hyötyjä. Fi∗I kiitän Bharat Raota kannustuksesta ja tuesta sekä UAI-2006-tuomareita rakentavista kommenteista. On syytä huomata, että hyötyarvon käyttäminen sen sijaan, että kokeilumuuttujaan X = x. Peruskysymys, johon on vastattava, on, mikä vaihtoehtoisista toimista on paras valinta, kun otetaan huomioon informaatio. Esittelemme todennäköisyyspelien käsitteen ja johdamme hinnoittelukaavan, joka mahdollistaa niiden vertailun. Huomaa, että kun malli θ ∈ Θ ja havainto x on annettu, hyötymuuttujan Y jakauma on Prθ(y|x). Klassisen Bayesin päätösteorian mukaan toiminnot a ∈ A ovat arvoja niiden odotetun hyödyn mukaan.

**Tulos**

Uusi aksiomatisointi todennäköisyyspeleille.

**Esimerkki 2.676**

Ongelmanratkaisu vastausjoukko-ohjelmoinnissa koostuu kahdesta vaiheesta: ensimmäisestä maadoitusvaiheesta, jossa kaikki muuttujat korvataan järjestelmällisesti termeillä, ja toisesta ratkaisuvaiheesta, jossa lasketaan saadun maadoitusohjelman vakaat mallit. Molempien vaiheiden monimutkainen osa on aggregaattien käsittely, jotka ovat suosittuja kielikonstruktioita, joiden avulla voidaan ilmaista ominaisuuksia joukkojen yli. Tässä artikkelissa käsittelemme aggregaattien käsittelyä gringo-sarjan 4 maadoituksessa. Näin ollen lähestymistapamme on sovellettavissa puoliksi naiiveihin tietokantojen arviointitekniikoihin perustuvaan pohjustukseen. Tarjoamme erityisesti sarjan algoritmeja, joissa käsitellään yksityiskohtaisesti rekursiivisten aggregaattien käsittelyä, ja havainnollistamme tätä toimivalla esimerkillä.

**Tulos**

Rekursiivisten aggregaattien maadoitus: Alustava raportti

**Esimerkki 2.677**

Katastrofaalinen unohtuminen on ongelma, joka tarkoittaa ensimmäisen tehtävän tietojen menettämistä sen jälkeen, kun on harjoiteltu toista tehtävää neuroverkkojen jatkuvassa oppimisessa. Tämän ongelman ratkaisemiseksi ehdotamme inkrementaalista momenttisovitusta (IMM), jossa käytetään Bayesin neuroverkkokehystä. IMM:ssä oletetaan, että neuroverkkojen parametrien jälkijakaumaa approksimoidaan Gaussin jakaumalla, ja se vastaa inkrementaalisesti jälkijakaumien momentteja, jotka on koulutettu ensimmäistä ja toista tehtävää varten. Jotta Gaussin olettamuksemme olisi järkevä, IMM-menettelyssä käytetään erilaisia siirto-oppimistekniikoita, kuten painonsiirtoa, vanhojen ja uusien parametrien L2-normia ja uutta ehdotettua pudotusmenetelmää, jossa käytetään vanhoja parametreja. Analysoimme menetelmiämme MNIST- ja CIFAR-10-tietokannoissa ja arvioimme niitä sitten Google Glass -lasilla kerätyssä todellisessa elämänpäiväkirjatietokannassa. Kokeelliset tulokset osoittavat, että IMM tuottaa parhaan mahdollisen suorituskyvyn eri tietokokonaisuuksissa.

**Tulos**

Katastrofaalisen unohduksen voittaminen inkrementaalisella hetken yhteensovittamisella

**Esimerkki 2.678**

Ehdotamme menetelmää, jolla voidaan rakentaa äärellisen tilaisia reaktiivisia ohjaimia järjestelmille, joiden vuorovaikutus vastustajan ympäristön kanssa mallinnetaan äärettömän pitkäkestoisilla kaksinpeleillä (mahdollisesti) äärettömillä graafeilla. Ehdotettu menetelmä kohdistuu turvallisuuspeleihin, joissa on äärettömän monta tilaa tai niin suuri määrä tiloja, että olisi epäkäytännöllistä - ellei mahdotonta - käyttää tavanomaisia synteesitekniikoita, jotka toimivat koko tilaavaruudessa. Rakennamme tällaisten järjestelmien äärellisten tilojen ohjaimet automaattioppimismenetelmällä, jossa hyödynnetään äärellisiin automaatteihin perustuvaa pelin symbolista esitystä. Koko oppimisprosessin ajan oppija ylläpitää voittoalueen approksimaatiota (joka on esitetty äärellisenä automatiikkana) ja tarkentaa sitä opettajan antamien erityyppisten vastaesimerkkien avulla, kunnes voidaan johtaa tyydyttävä ohjain (jos sellainen on olemassa). Esitämme turvallisuuspelien symbolisen esityksen (joka perustuu säännölliseen mallintarkistukseen), ehdotamme oppijan ja opettajan toteutuksia ja arvioimme niiden suorituskykyä esimerkkien avulla, jotka perustuvat robottien liikesuunnitteluun dynaamisissa ympäristöissä.

**Tulos**

Automaattioppimisen lähestymistapa turvallisuuspelien ratkaisemiseen äärettömien graafien yli.

**Esimerkki 2.679**

Tässä työssä tutkimme tärkeää ongelmaa: ohjelmien oppimista input-output<lb>esimerkkien perusteella. Ehdotamme uudenlaista menetelmää neuraalisen ohjelman oppimiseen, joka käyttää<lb>aluekohtaista ei-differentiaalista konetta, ja osoitamme, että tätä menetelmää<lb>voidaan soveltaa sellaisten ohjelmien oppimiseen, jotka ovat huomattavasti monimutkaisempia kuin aiemmin syntetisoidut<lb>ohjelmat: ohjelmointikielen jäsentäjät tulo-lähtö -pareista<lb> tuntematta taustalla olevaa kielioppia. Päähaasteena on kouluttaa neuraalinen<lb>ohjelma ilman valvontaa suoritusjäljillä. Sen ratkaisemiseksi ehdotamme: (1) <lb>LL-koneita ja niitä käyttäviä neuraalisia ohjelmia, joilla voidaan tehokkaasti säännellä opittujen ohjelmien<lb>avaruutta; ja (2) kaksivaiheista vahvistusoppimiseen perustuvaa<lb>hakutekniikkaa mallin kouluttamiseksi. Arviointimme osoittaa, että lähestymistapamme<lb>voi menestyksekkäästi oppia jäsentämään ohjelmia sekä imperatiivisella kielellä että<lb>funktiokielellä ja saavuttaa 100 prosentin testitarkkuuden, kun taas nykyisten lähestymistapojen<lb>tarkkuudet ovat lähes 0 prosenttia. Tämä on ensimmäinen onnistunut osoitus<lb>vahvistusoppimisen soveltamisesta ei-differentiaalista<lb>konetta käyttävän neuraalisen ohjelman kouluttamiseen, joka voi täysin yleistää testijoukkoihin ei-triviaalissa tehtävässä.

**Tulos**

Neuraalisten ohjelmien oppiminen analysoimaan ohjelmia

**Esimerkki 2.680**

Yrityksiä kouluttaa kattavaa tekoälyä, joka kykenee ratkaisemaan useita tehtäviä, on haitannut krooninen ongelma, jota kutsutaan katastrofaaliseksi unohtamiseksi. Vaikka kaikkien aiempien tietojen pelkkä toistaminen lievittää ongelmaa, se vaatii paljon muistia, ja mikä vielä pahempaa, se on usein mahdoton toteuttaa reaalimaailman sovelluksissa, joissa aiempien tietojen saatavuus on rajallista. Hippokampuksen generatiivisen luonteen innoittamana kädellisten aivojen lyhytkestoisen muistin järjestelmänä ehdotamme Deep Generative Replay -menetelmää, joka on uusi kehys, jossa on yhteistyöhön perustuva kaksoismalliarkkitehtuuri, joka koostuu syvästä generatiivisesta mallista ("generaattori") ja tehtävänratkaisumallista ("ratkaisija"). Ainoastaan näiden kahden mallin avulla aiempien tehtävien harjoittelutiedot voidaan helposti ottaa näytteeksi ja lomittaa uuden tehtävän harjoitustietojen kanssa. Testaamme menetelmiämme useissa peräkkäisissä oppimisympäristöissä, joihin liittyy kuvien luokittelutehtäviä.

**Tulos**

Jatkuva oppiminen syvän generatiivisen toiston avulla

**Esimerkki 2.681**

Perinteisissä latentin käyttäytymisen estimointimenetelmissä käytetään yleensä asenteellisia kysymyksiä, jotka ovat subjektiivisia, ja näitä kysymyksiä ei välttämättä ole aina saatavilla. Oletamme, että latenttien muuttujien estimointiin voidaan käyttää vaihtoehtoista lähestymistapaa suuntaamattomien graafisten mallien avulla. Esimerkiksi ei-parametriset keinotekoiset neuroverkot. Tässä tutkimuksessa tarkastelemme generatiivisten ei-parametristen mallinnusmenetelmien käyttöä latenttien muuttujien estimoinnissa ennakkovalintajakaumasta ilman perinteistä mittausindikaattoreiden käyttöä. Rajoitettua Boltzmannin konetta käytetään latenttien käyttäytymistekijöiden esittämiseen analysoimalla havaittujen valintojen ja selittävien muuttujien välistä suhdetietoa. Algoritmi on mukautettu latentin käyttäytymisen analysointiin diskreetin valinnan skenaariossa, ja käytämme graafista lähestymistapaa arvioidaksemme ja ymmärtäksemme semanttisen merkityksen estimoitujen parametrivektorien arvoista. Havainnollistamme menetelmäämme rahoitusvälineiden valintaa koskevalla tietokokonaisuudella ja teemme tilastollisen analyysin parametrien herkkyydestä ja vakaudesta. Tuloksemme osoittavat, että ei-parametristen tilastollisten testien avulla voimme poimia hyödyllistä latenttia tietoa latenttien konstruktioiden käyttäytymisestä koneoppimismenetelmien avulla ja että niillä on vahva ja merkittävä vaikutus valintaprosessiin. Lisäksi mallinnuskehyksemme osoittaa kestävyyttä syötteen vaihtelun suhteen otannan ja validoinnin avulla. ∗Paper presented at International Choice Modelling Conference 2017 †Laboratory of Innovations in Transportation (LITrans), Department of Civil Engineering, Ryerson University, Toronto, Canada, Email: melvin.wong@ryerson.ca ‡Laboratory of Innovations in Transportation (LITrans), Department of Civil Engineering, Ryerson University, Toronto, Canada, Email: bilal.farooq@ryerson.ca Laboratoire d'Interprétation et de Traitement d'Images et Vidéo (LITIV), Department of Computer and Software Engineering, Polytechnique Montréal, Montréal, Kanada, Email: guillaume-alexandre.bilodeau@polymtl.ca ar X iv :1 70 6. 00 50 5v 1 [ cs .L G ] 1 J un 2 01 7

**Tulos**

Diskriminoiva ehdollinen rajoitettu Boltzmannin kone diskreettia valintaa ja latenttien muuttujien mallintamista varten.

**Esimerkki 2.682**

Variationaaliset autokooderit (VAE) ovat suosittu ja joustava syvän generatiivisen mallin muoto, joka voidaan sovittaa stokastisesti tietyn satunnaisprosessin näytteisiin käyttämällä tietoteoreettista varianssisidettä todelliselle jakaumalle. Kun malli on näin saatu, sitä voidaan oletettavasti käyttää uusien näytteiden tuottamiseen tästä jakaumasta tai matalaulotteisen latentin esityksen tuottamiseen olemassa olevista näytteistä. Vaikka malli on varsin tehokas monilla sovellusalueilla, tietyt tärkeät VAE:n käyttäytymistä säätelevät mekanismit jäävät vaikeasti ratkaistavien integraalien ja niistä johtuvien stokastisten approksimaatioiden vuoksi hämärän peittoon. Koska kyseessä on erittäin epäkonveksaalinen malli, on lisäksi epäselvää, miten taustalla olevan energian minimit liittyvät alkuperäisiin suunnittelutarkoituksiin. Pyrimme määrittelemään nämä kysymykset paremmin analysoimalla joukon helposti lähestyttäviä erikoistapauksia, joiden monimutkaisuus kasvaa. Näin tehdessämme paljastamme mielenkiintoisia yhteyksiä perinteisempiin mittasuhteiden pienentämismalleihin sekä luontaisen, mutta aliarvostetun taipumuksen hylätä poikkeavat mallit, kun estimoidaan latentteja moninaisuuksia. Jälkimmäisen osalta osoitamme, että VAE:tä voidaan pitää viimeaikaisten robustien PCA-mallien luonnollisena evoluutiona, joka kykenee oppimaan epälineaarisia moninaisuuksia, jotka on peitetty karkeilla vääristymillä. Tämän aiemmin tutkimattoman ominaisuuden hintana on kuitenkin mallin mahdollinen romahtaminen rappeutuneeseen jakaumaan, joka saattaa olla vähemmän sopiva uusien näytteiden luomisen perustaksi.

**Tulos**

Variationaalisen automaattisen koodaajan salatut ominaisuudet

**Esimerkki 2.683**

Relationaaliset Markovin satunnaiskentät ovat yleinen ja joustava kehys, jonka avulla voidaan päätellä suuren määrän vuorovaikutuksessa olevien entiteettien attribuuttien yhteistä jakaumaa. Tällaisten mallien oppimisen suurin laskennallinen ongelma on päättely. Jopa silloin, kun käsitellään täydellistä dataa, jossa voidaan tiivistää suuri alue riittävillä tilastoilla, oppiminen edellyttää riittävien tilastojen odotusarvojen laskemista eri parametrivalinnoilla. Tyypillinen ratkaisu tähän ongelmaan on käyttää likimääräisiä päättelymenetelmiä, kuten loopy belief propagation -menetelmää. Vaikka nämä menettelyt ovat varsin tehokkaita, ne vaativat silti laskentaa, joka on mallin vuorovaikutusten (tai ominaisuuksien) määrän suuruusluokkaa. Kun opetellaan suurta relaatiomallia monimutkaisella alueella, jopa tällaiset approksimaatiot vaativat epärealistisen pitkän suoritusajan. Tässä artikkelissa osoitamme, että tietylle relaatiomallien MRF-luokalle, joilla on luontainen symmetria, voimme suorittaa oppimismenetelmissä tarvittavan päättelyn mallitason uskomuspropagandan avulla. Tämän menettelyn suoritusaika on verrannollinen relaatiomallin kokoon eikä niinkään toimialueen kokoon. Lisäksi osoitamme, että tämä laskentamenettely vastaa synkronista loopy belief propagation -menetelmää. Tämä nopeuttaa päättelyä ja oppimisaikaa huomattavasti. Käytämme tätä menettelyä oppiaksemme relationaalisia MRF-malleja, joilla voidaan kuvata suurten proteiini-proteiini-vuorovaikutusverkostojen yhteistä jakaumaa.

**Tulos**

Mallipohjainen päättely symmetrisissä suhteellisissa Markovin satunnaiskentissä (Template Based Inference in Symmetric Relational Markov Random Fields)

**Esimerkki 2.684**

Osoitamme, että [12, 4] esittämä keskimääräinen stabiilisuuden käsite on muuttumaton datan esikäsittelyn suhteen laajalle yleistettyjen lineaaristen mallien luokalle, joka sisältää suurimman osan tunnetuista eksp-konkaavista häviöistä. Toisin sanoen, kun analysoimme tietyn algoritmin stabiilisuusastetta, voimme olettaa, että datan esikonditionointi on optimaalinen. Tämä merkitsee sitä, että ainakaan tilastollisesta näkökulmasta eksplisiittistä regularisointia ei tarvita huonosti ehdollistetun datan kompensoimiseksi, mikä on ristiriidassa laajalti käytetyn lähestymistavan kanssa, joka sisältää regularisoinnin yleistettyjen lineaaristen mallien otoskompleksisuuden analysoimiseksi. Löydöksiemme useita tärkeitä seurauksia ovat seuraavat: a) Osoitamme, että empiirisen riskin minimoinnin (ERM) ylimääräistä riskiä kontrolloidaan esikonditionoidulla vakausasteella. Tämä tuottaa välittömästi suhteellisen lyhyen ja tyylikkään todisteen ERM:n saavuttamille nopeille nopeuksille kontekstissamme. b) Vahvistamme [9] hiljattain esitettyjä rajoja stokastisen Gradient Descent-algoritmin stabiilisuusnopeudelle.

**Tulos**

Keskimääräinen stabiilisuus ei muutu tietojen esikäsittelystä. Vaikutukset eksp-konkaavaan empiiriseen riskin minimointiin

**Esimerkki 2.685**

imbalanced-learn on avoimen lähdekoodin python-työkalupakki, jonka tarkoituksena on tarjota laaja valikoima menetelmiä, joilla voidaan käsitellä koneoppimisessa ja hahmontunnistuksessa usein esiintyvää epätasapainoisen aineiston ongelmaa. Toteutetut uusimmat menetelmät voidaan luokitella neljään ryhmään: i) alimman näytteenoton menetelmät, ii) ylimman näytteenoton menetelmät, iii) ylimman ja alimman näytteenoton yhdistelmä ja iv) ensemble-oppimismenetelmät. Ehdotettu työkalupakki on riippuvainen ainoastaan numpy-, scipy- ja scikit-learn-ohjelmista, ja se jaetaan MIT-lisenssillä. Lisäksi se on täysin yhteensopiva scikit-learnin kanssa ja on osa scikit-learn-contrib-projektia. Dokumentaatio, yksikkötestit ja integraatiotestit ovat mukana käytön ja osallistumisen helpottamiseksi. Työkalupakki on julkisesti saatavilla GitHubissa https://github.com/scikit-learn-contrib/imbalanced-learn.

**Tulos**

Epätasapainoinen oppiminen: Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning (Python-työkalupakki epätasapainoisten aineistojen kiroukseen koneoppimisessa).

**Esimerkki 2.686**

Tässä tutkimuksessa käytetään vektoriavaruusmallia sanojen tunnesuuntausten mittaamiseen. Kaksi edustavaa vektoria positiivista/negatiivista polariteettia varten rakennetaan käyttämällä korkea-ulotteista vektoriavaruutta sekä valvomattomalla että puolivalvotulla tavalla. Sanakohtainen sentimenttisuuntautuneisuuden arvo määritetään ottamalla kosinusetäisyyksien erotus kahteen vertailuvektoriin nähden. Näitä kahta tilaa (valvomaton ja puolivalvottu) verrataan olemassa olevaan valvomattomaan menetelmään (Turney, 2002). Kokeilumme tuloksena osoitamme, että tämä uusi lähestymistapa on huomattavasti parempi kuin aiempi valvomaton lähestymistapa ja että se on myös käytännöllisempi ja datatehokkaampi.

**Tulos**

Uusi lähestymistapa tunteiden suuntautumisen mittaamiseen moniulotteisen vektoriavaruuden avulla.

**Esimerkki 2.687**

Tutkimme yllättävää ilmiötä, joka liittyy datapisteiden pilven esittämiseen polynomien avulla. Aloitamme aiemmin huomaamattomasta empiirisestä havainnosta, jonka mukaan tietopisteiden kokoelman (pilven) tapauksessa tietyn erotetun polynomin alitasojoukot kuvaavat pilven muodon hyvin tarkasti. Tämä erotettu polynomi on neliösumma (SOS), joka on johdettu yksinkertaisella tavalla empiirisen momenttimatriisin käänteisluvusta. Itse asiassa tämä SOS-polynomi liittyy suoraan ortogonaalisiin polynomeihin ja Christoffelin funktioon. Näin voidaan yleistää ja tulkita ortogonaalisten polynomien äärellisyysominaisuuksia ja antaa matemaattinen perustelu havaitulle ilmiölle. Monenlaisten mahdollisten sovellusten joukossa havainnollistamme tulostemme merkitystä verkon tunkeutumisen havaitsemistehtävässä, jossa saavutamme samankaltaisia tuloksia kuin kirjallisuudessa esitetyt nykyiset erityiset menetelmät.

**Tulos**

Tyypillisyyden selvittäminen käänteismomenttimatriisin SOS-polynomin avulla

**Esimerkki 2.688**

Ehdotamme paikallista lähestymistapaa usean ytimen oppimiseen, joka voidaan muotoilla konveksiseksi optimointiongelmaksi tietyn klusterirakenteen yli. Jolle saamme yleistysvirhetakeita ja johdamme Fenchelin duaaliesitykseen perustuvan optimointialgoritmin. Kokeilut reaalimaailman tietokokonaisuuksilla laskennallisen biologian ja tietokonenäön sovellusalueilta osoittavat, että konveksoituneella lokalisoidulla moniydinoppimisella voidaan saavuttaa korkeampi ennustustarkkuus kuin sen globaalilla ja ei-konveksoituneella paikallisella vastineella.

**Tulos**

Paikallinen monikenno-oppiminen-konveksimainen lähestymistapa

**Esimerkki 2.689**

Esittelemme uuden näkökulman graafipohjaisiin menetelmiin, joilla suosittelujärjestelmiä varten voidaan tehdä yhteistoiminnallinen luokittelu. Unlike user-based or item-based methods that compute a weighted average of ratings given by the nearest neighbors, or low-rank approximation methods using convex optimization and the nuclear norm, we formulate matrix completion as a series of semi-supervised learning problems, and propagate the known ratings to the missing ones on the user-user or item-item graph globally. Puolivalvotut oppimisongelmat ilmaistaan Laplace-Beltrami-yhtälöinä moninaisuudella eli harmonisella laajennuksella, ja ne voidaan diskretisoida pisteintegraalimenetelmällä. Osoitamme, että lähestymistapamme ei aseta datapisteille matalarivistä euklidista aliavaruutta, vaan sen sijaan minimoi taustalla olevan moninaisuuden ulottuvuuden. Menetelmämme, jonka nimi on LDM (low dimensional manifold), osoittautuu erityisen tehokkaaksi kohteiden ranking-luokitusten tuottamisessa, sillä se on laskennallisesti erittäin tehokas ja ranking-luokituksen laatu on vankka verrattuna uusimpiin menetelmiin.

**Tulos**

Harmoninen laajentamismenetelmä yhteistoiminnallista luokittelua varten

**Esimerkki 2.690**

Aiemmissa töissä on mallinnettu sanojen kompositionaalisuutta luomalla merkkitason merkitysmalleja, mikä on vähentänyt harvinaisten sanojen harvinaisuuteen liittyviä ongelmia. Monissa kirjoitusjärjestelmissä kompositionaalisuus vaikuttaa kuitenkin myös merkkitasolla: merkin merkitys saadaan sen osien summasta. Tässä artikkelissa mallinnamme tätä vaikutusta luomalla merkkien upotuksia niiden visuaalisten ominaisuuksien perusteella, luomalla merkistä kuvan ja ajamalla se konvoluutiohermoverkon läpi visuaalisen merkin upotuksen tuottamiseksi. Kokeet tekstinluokittelutehtävässä osoittavat, että tällainen malli mahdollistaa harvinaisia merkkejä sisältävien tapausten paremman käsittelyn esimerkiksi kiinan, japanin ja korean kielissä. Lisäksi laadulliset analyysit osoittavat, että ehdotettu malli oppii keskittymään merkkien osiin, jotka sisältävät kategorista sisältöä, mikä johtaa visuaalisessa tilassa yhtenäisiin merkkikuvioihin.

**Tulos**

Hahmotason kompositionaalisuuden oppiminen visuaalisten ominaisuuksien avulla

**Esimerkki 2.691**

Ajallinen abstraktio on avainasemassa oppimisen ja suunnittelun skaalaamisessa vahvistusoppimisessa. Vaikka ajallisesti laajennettujen toimintojen suunnittelu ymmärretään hyvin, tällaisten abstraktioiden luominen itsenäisesti datasta on ollut haastavaa. Me käsittelemme tätä ongelmaa vaihtoehtojen puitteissa [Sutton, Precup & Singh, 1999; Precup, 2000]. Johdamme politiikkagradienttiteoreemoja optioille ja ehdotamme uutta optiokriittistä arkkitehtuuria, joka kykenee oppimaan sekä optioiden sisäiset politiikat että lopetusehdot samanaikaisesti optioita koskevan politiikan kanssa ja ilman tarvetta tarjota mitään ylimääräisiä palkkioita tai alatavoitteita. Kokeelliset tulokset sekä diskreeteissä että jatkuvissa ympäristöissä osoittavat kehyksen joustavuuden ja tehokkuuden. Johdanto Ajallinen abstraktio mahdollistaa tiedon esittämisen eri aikaskaaloilla tapahtuvien toimien kulusta. Vahvistusoppimisessa optiot (Sutton, Precup ja Singh 1999; Precup 2000) tarjoavat kehyksen tällaisten toimintakurssien määrittelyyn ja saumattomaan oppimiseen ja suunnitteluun niiden avulla. Ajallisten abstraktioiden itsenäinen löytäminen on ollut laajan tutkimustyön kohteena viimeisten 15 vuoden aikana (McGovern ja Barto 2001; Stolle ja Precup 2002; Menache, Mannor ja Shimkin 2002; Şimşek ja Barto 2009; Silver ja Ciosek 2012), mutta lähestymistavat, joita voidaan käyttää luontevasti jatkuvien tila- ja/tai toiminta-avaruuksien kanssa, ovat vasta viime aikoina alkaneet tulla mahdollisiksi (Konidaris et al. 2011; Niekum ja Barto 2011; Mann, Mannor ja Precup ; Mankowitz, Mann ja Mannor 2016; Kulkarni et al. 2016; Vezhnevets et al. 2016; Daniel et al. 2016). Suurin osa olemassa olevasta työstä on keskittynyt alatavoitteiden (hyödylliset tilat, jotka agentin tulisi saavuttaa) löytämiseen ja sen jälkeen politiikkojen oppimiseen niiden saavuttamiseksi. Tämä ajatus on johtanut mielenkiintoisiin menetelmiin, mutta niitä on myös vaikea skaalata, koska ne ovat "kombinatorisia". Lisäksi osatavoitteisiin liittyvien toimintatapojen oppiminen voi olla kallista datan ja laskenta-ajan kannalta; pahimmassa tapauksessa se voi olla yhtä kallista kuin koko tehtävän ratkaiseminen. Esitämme vaihtoehtoisen näkemyksen, joka hämärtää rajan vaihtoehtojen löytämisen ongelman ja vaihtoehtojen oppimisen ongelman välillä. Politiikan gradienttiteorian (Sutton et al. 2000) perusteella johdamme uusia tuloksia, jotka mahdollistavat vaihtoehtojen sisäisten politiikkojen ja lopetusfunktioiden asteittaisen oppimisprosessin samanaikaisesti niitä koskevan politiikan kanssa. Tämä lähestymistapa toimii luonnollisesti sekä lineaaristen että epälineaaristen funktioiden approksimaattoreiden kanssa diskreeteissä tai jatkuvissa tila- ja toimintaavaruuksissa. Nykyiset vaihtoehtojen oppimismenetelmät ovat huomattavasti hitaampia, kun ne oppivat yhdestä tehtävästä: suuri osa hyödystä saadaan käyttämällä opittuja vaihtoehtoja uudelleen samankaltaisissa tehtävissä. Sitä vastoin osoitamme, että lähestymistapamme pystyy menestyksekkäästi oppimaan vaihtoehtoja yksittäisen tehtävän sisällä ilman hidastumista ja samalla nopeuttamaan uudelleenkäyttöä. Aloitamme tarkastelemalla taustaa, joka liittyy työhömme kahteen tärkeimpään osatekijään: politiikan gradienttimenetelmiin ja vaihtoehtoihin. Sen jälkeen kuvaamme lähestymistapamme keskeiset ideat: vaihtoehtojen sisäisen politiikan ja lopetusgradientin lauseet. Lisätietoa teknisistä yksityiskohdista on liitteessä. Esitämme kokeellisia tuloksia, jotka osoittavat, että lähestymistapamme oppii mielekkäitä ajallisesti laajennettuja käyttäytymismalleja tehokkaalla tavalla. Toisin kuin muissa menetelmissä, meidän tarvitsee määrittää vain haluttujen vaihtoehtojen määrä; ei tarvita alatavoitteita, ylimääräisiä palkkioita, demonstraatioita, useita ongelmia tai muita erityisiä mukautuksia (lähestymistapa voi kuitenkin haluttaessa toimia pseudopalkintofunktioiden kanssa). Tietojemme mukaan tämä on ensimmäinen kokonaisvaltainen lähestymistapa vaihtoehtojen oppimiseen, joka skaalautuu hyvin suurille alueille vertailukelpoisella tehokkuudella. Esitiedot ja merkinnät Markovin päätösprosessi koostuu tilojen joukosta S, toimintojen joukostaA, siirtymäfunktiosta P : S × A → (S → [0, 1]) ja palkitsemisfunktiosta r : S × A → R. Yksinkertaisuuden vuoksi kehitämme ajatuksiamme olettaen, että tilojen ja toimintojen joukot ovat diskreettejä. Tuloksemme voidaan kuitenkin ulottaa jatkuviin avaruuksiin tavanomaisilla mittateoreettisilla oletuksilla (osa empiirisistä tuloksistamme koskee jatkuvia tehtäviä). (Markovin stationaarinen) toimintapolitiikka on todennäköisyysjakauma tilojen ehdollistamien toimien yli, π : S × A → [0, 1]. Diskontattavissa ongelmissa politiikan π arvofunktio määritellään odotetuksi tuotoksi: Vπ(s) = Eπ [ ∑∞ t=0 γ rt+1 | s0 = s] ja sen toiminta-arvofunktio Qπ(s, a) = Eπ [ ∑∞ t=0 γ rt+1 | s0 = s, a0 = a], missä γ ∈ [0, 1) on diskonttauskerroin. Politiikka π on ahne tietyn toiminta-arvofunktion Q suhteen, jos π(s, a) > 0 iff a = argmaxa′ Q(s, a ′). Diskreetissä MDP:ssä on olemassa ainakin yksi optimaalinen politiikka, joka on ahne taka- X iv :1 60 9. 05 14 0v 1 [ cs .A I] 1 6 Se p 20 16 spect oman toiminta-arvofunktionsa suhteen. Politiikkagradienttimenetelmät (Sutton et al. 2000; Konda ja Tsitsiklis 2000) käsittelevät hyvän politiikan löytämisen ongelmaa suorittamalla stokastista gradienttilaskeutumista suorituskykytavoitteen optimoimiseksi tietyn parametrisoitujen stokastisten politiikkojen πθ-perheen yli. Politiikkagradienttiteoriassa (Sutton et al. 2000) annetaan lausekkeet keskimääräisen palkkion ja diskontatun palkkion tavoitteiden gradientille θ:n suhteen. Diskontatun palkkion tapauksessa tavoite määritellään määritellyn alkutilan (tai jakauman) s0 suhteen: ρ(θ, s0) = Eπθ [ ∑ t=0 γ rt+1 | s0]. Politiikan gradienttiteoria osoittaa, että: ∂ρ(θ,s0) ∂θ = ∑ s μπθ (s | s0) ∑ a ∂πθ(a|s) ∂θ Qπθ (s, a), missä μπθ (s | s0) = ∑∞ t=0 γ t P (st = s | s0) on s0:sta alkavien liikeratojen varrella olevien tilojen diskontattu painotus. Käytännössä politiikan gradientti estimoidaan näytteistä, jotka ovat pitkin politiikan pysyvää jakaumaa. (Thomas 2014) osoitti, että diskonttauskertoimen laiminlyönti tässä stationaarisessa jakaumassa tekee tavanomaisesta politiikkagradientin estimaattorista vinoutuneen. Tämän poikkeaman korjaaminen vähentää kuitenkin myös datan tehokkuutta. Yksinkertaisuuden vuoksi käytämme (Sutton et al. 2000) kehystä ja keskustelemme siitä, miten tuloksiamme voidaan laajentaa (Thomas 2014) mukaisesti. Vaihtoehtojen kehys (Sutton, Precup ja Singh 1999; Precup 2000) virallistaa ajallisesti laajennettujen toimien ajatuksen. Markovilainen optio ω ∈ Ω on kolmikko (Iω, πω, βω), jossa Iω ⊆ S on aloitusjoukko, πω on option sisäinen toimintatapa ja βω : S → [0, 1] on lopetusfunktio. Oletamme myös, että ∀s ∈ S,∀ω ∈ Ω : s ∈ Iω (eli kaikki vaihtoehdot ovat saatavilla kaikkialla), mikä on oletus, joka tehdään useimmissa vaihtoehtojen löytämisalgoritmeissa. Keskustelemme siitä, miten tästä oletuksesta voidaan luopua viimeisessä jaksossa. (Sutton, Precup ja Singh 1999; Precup 2000) osoittavat, että MDP:stä, joka on varustettu joukolla vaihtoehtoja, tulee Semi-Markov-päätösprosessi (Puterman 1994, luku 11), jolla on vastaava optimaalinen arvofunktio vaihtoehtojen yli VΩ(s) ja vaihtoehtoarvofunktio QΩ(s, ω). MDP:n oppimis- ja suunnittelualgoritmeilla on vastineensa tässä ympäristössä. Taustalla olevan MDP:n olemassaolo tarjoaa kuitenkin mahdollisuuden oppia monista eri vaihtoehdoista rinnakkain: tämä on ajatus vaihtoehtojen sisäisestä oppimisesta, jota hyödynnämme työssämme. Vaihtoehtojen oppiminen Otamme käyttöön jatkuvan näkökulman vaihtoehtojen oppimisen ongelmaan. Haluaisimme milloin tahansa tislata kaiken saatavilla olevan kokemuksen järjestelmämme jokaiseen komponenttiin: arvofunktioon ja vaihtoehtoihin liittyviin toimintatapoihin, vaihtoehtojen sisäisiin toimintatapoihin ja lopetusfunktioihin. Tämän tavoitteen saavuttamiseksi keskitymme optiokäytäntöjen ja lopetusfunktioiden oppimiseen olettaen, että ne esitetään differentioituvien parametrisoitujen funktioiden approksimaattoreiden avulla. Tarkastelemme call-and-return -vaihtoehdon toteutusmallia, jossa agentti valitsee vaihtoehdon ω sen vaihtoehtoja koskevan politiikan πΩ mukaisesti ja noudattaa sitten vaihtoehdon sisäistä politiikkaa πω lopettamiseen asti (βω:n sanelemana), jolloin menettely toistetaan. Merkitään πω,θ vaihtoehdon ω sisäistä politiikkaa, jota parametrisoidaan θ:llä, ja βω,θ vaihtoehdon ω lopettamisfunktiota, jota parametrisoidaan θ:llä. Esitämme kaksi uutta tulosta vaihtoehtojen oppimisesta, jotka saadaan käyttämällä apuna politiikka-gradienttiteoriaa (Sutton et al. 2000). Molemmat tulokset on johdettu olettaen, että tavoitteena on oppia vaihtoehtoja, jotka maksimoivat odotetun tuoton nykyisessä tehtävässä. Jos kuitenkin halutaan lisätä lisäinformaatiota tavoitefunktioon, se voidaan helposti tehdä, kunhan se on additiivisesti erilaistuvan funktion muodossa. Oletetaan, että pyrimme optimoimaan suoraan diskontattua tuottoa, joka odotetaan kaikkien sellaisten liikeratojen aikana, jotka alkavat määrätystä tilasta s0 ja vaihtoehdosta ω0, jolloin: ρ(Ω, θ, θ, θ, s0, ω0) = EΩ,θ,ω [ ∑∞ t=0 γ rt+1 | s0, ω0]. Huomaa, että tämä tuotto riippuu optiopolitiikasta sekä optiopolitiikkojen ja lopetusfunktioiden parametreista. Otamme tämän tavoitteen gradientit θ:n ja θ:n suhteen. Tätä varten manipuloimme yhtälöitä, jotka ovat samanlaisia kuin ne, joita käytetään optioiden sisäisessä oppimisessa (Sutton, Precup ja Singh 1999, luku 8). Optioarvofunktion määritelmä voidaan kirjoittaa seuraavasti: QΩ(s, ω) = E Ω,θ,θ [ ∞ ∑ t=0 γrt+1 ∣∣∣∣ s0 = s, ω0 = ω ]

**Tulos**

Vaihtoehtokriittinen arkkitehtuuri

**Esimerkki 2.692**

Oikeusalan ammattilaisten nykyiset tietojen analysointivalmiudet ovat edelleen jäljessä oikeudellisten asiakirjojen saatavuuden räjähdysmäisestä kasvusta digitaalisessa muodossa, mikä lisää tarvetta tehokkaammille oikeudellisen tiedonhaun (IR) ja kysymyksiin vastaamisen (QA) menetelmille. Etenkin tiedonhakutehtäviin liittyy useita ainutlaatuisia haasteita, jotka edellyttävät semanttisesti motivoitujen NLP-tekniikoiden käyttöä. Tässä työssä ehdotetaan kaksivaiheista menetelmää oikeudellista tiedonhakua varten, jossa yhdistetään leksikaalisia tilastoja ja distributiivisia lause-esityksiä Competition on Legal Information Extraction/Entailment (COLIEE) -kilpailun yhteydessä. Yhdistäminen tapahtuu disambiguointisääntöjen avulla, joita sovelletaan leksikaalisiin luokituksiin, kun niitä pidetään epäluotettavina tietyn kyselyn osalta. Kilpailu- ja koetulokset osoittavat, että ehdotettu lähestymistapa parantaa hieman hakutulosten yleistä suorituskykyä. Lisäksi esitetään virhe- ja parannustapausten analyysi, jonka avulla voidaan paremmin ymmärtää tuloksia.

**Tulos**

Oikeustiedon haun parantaminen termien järjestyksen todennäköisyyksien avulla tapahtuvalla jakaumakomposition avulla

**Esimerkki 2.693**

Monimuuttujaisia aikasarjoja esiintyy luonnollisesti monilla aloilla, kuten energia-alalla, bioinformatiikassa, signaalinkäsittelyssä ja rahoituksessa. Useimmissa näistä sovelluksista on pystyttävä vertailemaan näitä strukturoituja tietoja. Tässä yhteydessä dynaaminen aikasarjojen loimennus (DTW) on luultavasti yleisin vertailumenetelmä. Sen parantamiseen oppimalla ei kuitenkaan ole panostettu kovinkaan paljon tutkimustyötä. Tässä asiakirjassa ehdotamme uutta menetelmää DTW:hen perustuvan samankaltaisuuksien oppimiseksi aikasarjojen luokittelun parantamiseksi. Hyödynnämme yhtenäistä stabiilisuuskehystä ja tarjoamme ensimmäiset teoreettiset takeet lineaarisen luokittelun yleistysrajan muodossa. Kokeellinen tutkimus osoittaa, että ehdotettu lähestymistapa on tehokas ja tuottaa samalla harvoja luokittelijoita.

**Tulos**

Samankaltaisuusoppiminen aikasarjojen luokittelussa

**Esimerkki 2.694**

Tässä artikkelissa kuvaamme järjestelmän, jolla luodaan kolmiulotteisia visuaalisia simulaatioita luonnollisen kielen liikeilmaisuista. Käytämme rikasta muodollista mallia tapahtumista ja niiden osallistujista luodaksemme simulaatioita, jotka täyttävät niihin liittyvien lausumien minimirajoitukset, ja tukeudumme fyysisiä esineitä ja liiketapahtumia koskevaan semanttiseen tietämykseen. Tässä asiakirjassa esitetään teknisiä näkökohtia ja keskustellaan edellä mainittujen semanttisten mallien toteuttamisesta tällaisessa järjestelmässä.

**Tulos**

Multimodaaliset semanttiset simulaatiot kielellisesti alimääritellyistä liiketapahtumista

**Esimerkki 2.695**

ADABOOST.MH:n puitteissa ehdotamme vektoriarvoisten päätöspuiden kouluttamista moniluokkaisen reunan optimoimiseksi vähentämättä moniluokkaista ongelmaa K binääriseksi yhden ajonaikaisen luokittelun ongelmaksi. Menetelmän keskeinen elementti on vektoriarvoinen päätöspuu, joka on faktoroitu syötteestä riippumattomaksi vektoriksi, jonka pituus on K, ja etiketistä riippumattomaksi skalaariluokittelijaksi. Puun sisemmissä solmuissa etiketistä riippuvainen vektori hylätään ja binääriluokittimen avulla voidaan tuloavaruus jakaa kahteen alueeseen. Algoritmi säilyttää binäärisen ADABOOSTin käsitteellisen tyylikkyyden, tehon ja laskentatehokkuuden. Kokeissa se on samaa tasoa tukivektorikoneiden ja parhaan olemassa olevan moniluokkaisen tehostamisalgoritmin AOSOLOGITBOOSTin kanssa, ja se on huomattavasti parempi kuin muut tunnetut ADABOOSTin toteutukset.

**Tulos**

ADABOOST.MH:n paluu: moniluokkaiset Hamming-puut.

**Esimerkki 2.696**

Neuroverkkopohjaiset mallit ovat erittäin tehokas työkalu sanojen upotusten luomiseen, ja näiden mallien tavoitteena on ryhmitellä samankaltaiset sanat yhteen. Näitä upotuksia on käytetty ominaisuuksina tulosten parantamiseksi eri sovelluksissa, kuten asiakirjojen luokittelussa ja nimettyjen entiteettien tunnistamisessa. Neuraaliset kielimallit pystyvät oppimaan sanaedustuksia, joita on käytetty semanttisten muutosten tallentamiseen ajassa ja maantieteellisesti. Tämän artikkelin tavoitteena on ensin tunnistaa ja sitten visualisoida, miten sanat muuttavat merkitystään eri tekstikorpuksissa. Koulutamme neuraalisen kielimallin teksteille, jotka ovat peräisin monilta eri tieteenaloilta - filosofiasta, uskonnosta, kaunokirjallisuudesta jne. Jokainen teksti muuttaa sanojen upotuksia, jotta ne edustaisivat sanan merkitystä kyseisessä tekstissä. Esittelemme laskennallisen tekniikan, jonka avulla voidaan havaita sanat, joiden merkitys ja käyttö muuttuvat merkittävästi kielellisesti. Sen jälkeen käytämme tehostettuja hajontakuvioita ja tarinan visualisointia kielellisen muutoksen visualisoimiseksi.

**Tulos**

Kielellisen muutoksen visualisointi

**Esimerkki 2.697**

Esitetään analyysi eräiden käyttäjän ja koneen vuorovaikutuksen muotojen ja käyttötapojen nykyisistä integrointiyrityksistä. Käyttäjälähtöisen älykkään käyttöliittymän uutta käsitettä ehdotetaan multimodaalisen lisätyn todellisuuden ja aivo-tietokone-vuorovaikutuksen pohjalta erilaisiin sovelluksiin: vammaistutkimuksiin, koulutukseen, kotihoitoon, terveydenhuoltoon jne. Esitetään useita multimodaalisen lisäyksen käyttötapauksia. Hahmotellaan näkymiä, jotka liittyvät ihmisen ymmärtämisen parantamiseen välittömän palautteen avulla neurofyysisten kanavien kautta aivo-tietokone-vuorovaikutuksen avulla. Osoitetaan, että aivo-tietokoneliitäntäteknologia (BCI) tarjoaa uusia strategioita, joiden avulla voidaan ylittää nykyisin käytettävissä olevien käyttöliittymien rajoitukset erityisesti toimintarajoitteisten ihmisten kohdalla. Aiempien tutkimusten tulosten perusteella, jotka on tehty edullisista kuluttajille tarkoitetuista ja avoimen lähdekoodin BCI-laitteista, voidaan päätellä, että koneoppimisen (ML), multimodaalisen vuorovaikutuksen (visuaalinen, ääni, tuntoaisti) ja BCI:n yhdistäminen hyötyy välittömästä palautteesta, joka saadaan ML-menetelmillä luokitelluista todellisista neurofyysisistä reaktioista. Yleisesti ottaen BCI yhdistettynä muihin AR-vuorovaikutustapoihin voi tuottaa paljon enemmän tietoa kuin nämä vuorovaikutustyypit itsessään. Jo nykytilanteessa yhdistetyt AR-BCI-käyttöliittymät voisivat tarjota erittäin mukautuvia ja henkilökohtaisia palveluita erityisesti toimintarajoitteisille henkilöille. Asiasanat- lisätty todellisuus, esteettömyyskäyttöliittymät, multimodaalinen käyttöliittymä, aivo-tietokone-käyttöliittymä, sähköinen terveydenhuolto, koneoppiminen, koneen ja koneen välinen vuorovaikutus, ihmisen ja ihmisen välinen vuorovaikutus, ihmisen ja koneen välinen vuorovaikutus.

**Tulos**

Käyttäjälähtöinen älykäs käyttöliittymä multimodaalisen lisätyn todellisuuden ja aivo-tietokone-vuorovaikutuksen pohjalta toimintarajoitteisille henkilöille.

**Esimerkki 2.698**

Kokemuksemme maailmasta on multimodaalinen: näemme esineitä, kuulemme ääniä, tunnemme tekstuurin, haistamme hajuja ja maistamme makuja. Modaliteetti viittaa tapaan, jolla jokin asia tapahtuu tai koetaan, ja tutkimusongelmaa luonnehditaan multimodaaliseksi, kun se sisältää useita tällaisia modaliteetteja. Jotta tekoäly voisi edistyä meitä ympäröivän maailman ymmärtämisessä, sen on kyettävä tulkitsemaan tällaisia multimodaalisia signaaleja yhdessä. Multimodaalisen koneoppimisen tavoitteena on rakentaa malleja, jotka pystyvät käsittelemään ja suhteuttamaan useista eri modaliteeteista peräisin olevaa tietoa. Kyseessä on elinvoimainen monitieteinen ala, jonka merkitys kasvaa jatkuvasti ja jolla on poikkeuksellisia mahdollisuuksia. Sen sijaan, että keskityttäisiin tiettyihin multimodaalisiin sovelluksiin, tässä asiakirjassa tarkastellaan multimodaalisen koneoppimisen viimeaikaisia edistysaskeleita ja esitellään ne yhteisessä taksonomiassa. Menemme tyypillistä varhaisen ja myöhäisen fuusion luokittelua pidemmälle ja tunnistamme laajempia haasteita, joita multimodaalinen koneoppiminen kohtaa: representaatio, kääntäminen, kohdistus, fuusio ja yhteisoppiminen. Tämän uuden taksonomian avulla tutkijat voivat paremmin ymmärtää alan tilannetta ja tunnistaa tulevaisuuden tutkimussuuntia.

**Tulos**

Multimodaalinen koneoppiminen: Taksonomia: Katsaus ja taksonomia

**Esimerkki 2.699**

Monimuotoisuuden puute geneettisen algoritmin populaatiossa voi johtaa geneettisten operaattoreiden huonoon suorituskykyyn, koska etsinnän ja hyödyntämisen välillä ei ole tasapainoa. Tällaisissa tapauksissa geneettiset algoritmit konvergoituvat nopeasti ja epätarkoituksenmukaisesti. Tässä artikkelissa kehitämme uudenlaisen hybridi-geneettisen algoritmin, jolla pyritään saavuttamaan tasapaino etsinnän ja hyödyntämisen välille. Se kohtaa monimuotoisuusongelman käyttämällä nimettyä ahnetta monipuolistamisoperaattoria. Lisäksi ehdotetussa algoritmissa käytetään vanhempien ja lasten välistä kilpailua, jotta voidaan hyödyntää laadukkaita vierailtuja ratkaisuja. Näitä operaattoreita täydennetään yksinkertaisella valintamekanismilla, joka on suunniteltu säilyttämään populaation monimuotoisuus ja hyödyntämään sitä. Lisäksi laajennamme ehdotustamme memeettisten algoritmien alalle ja saamme parannetun mallin, jolla on erinomaisia käytännön tuloksia. Kokeellinen tutkimus osoittaa lähestymistavan pätevyyden sekä sen, kuinka tärkeää on ottaa huomioon etsintä- ja hyödyntämiskäsitteet, kun suunnitellaan evoluutio- ja algoritmimallia.

**Tulos**

GENEETTINEN JA MEMEETTINEN ALGORITMI, JONKA MONIMUOTOISUUDEN TASAPAINO PERUSTUU AHNEESEEN MONIPUOLISTUMISEEN.

**Esimerkki 2.700**

Syvälliset generatiiviset mallit tarjoavat tehokkaan ja joustavan keinon oppia monimutkaisia jakaumia datasta sisällyttämällä neuroverkkoja latenttimuuttujamalleihin. Tällaisten mallien kouluttamiseen käytettävät variatiiviset lähestymistavat ottavat käyttöön todennäköisyyskoodaajan, joka siirtää datan, joka on tyypillisesti valvomatonta, kietoutuneeseen esitysavaruuteen. Vaikka valvomaton oppiminen on usein toivottavaa, joskus jopa välttämätöntä, kun meillä ei ole ennakkotietoa siitä, mitä edustaa, voi tiedon tiettyjen vaihtelun osa-alueiden kuvaaminen toimialatietämyksen avulla usein auttaa oppimaan parempia epäyhtenäisiä representaatioita. Tässä esitellään uusi muotoilu puolivalvotusta oppimisesta varioivissa autokoodereissa, joka mahdollistaa juuri tämän. Se mahdollistaa todennäköisyyskoodereiden joustavan määrittelyn suunnattuina graafisina malleina stokastisen laskentagraafin kautta, joka sisältää sekä jatkuvia että diskreettejä latentteja muuttujia, joiden ehdolliset jakaumat on parametrisoitu neuroverkoilla. Osoitamme, miten riippuvuusrakenteiden ja muutaman merkityn esimerkin antaminen, jotka osoittavat uskottavia arvoja joillekin latenttiavaruuden komponenteille, voi auttaa oppimaan nopeasti eriytettyjä representaatioita. Arvioimme sitten sen kykyä tehdä näin sekä kvalitatiivisesti tutkimalla sen generatiivista kapasiteettia että kvantitatiivisesti käyttämällä epäyhtenäistä esitystä luokitteluun erilaisilla malleilla ja tietokokonaisuuksilla.

**Tulos**

SYVISSÄ GENERATIIVISISSA MALLEISSA

**Esimerkki 2.701**

Optimointi stokastisella gradienttilaskeutumisella on tärkeä osa monia laajamittaisia koneoppimisalgoritmeja. On kuitenkin epäselvää, ovatko nämä algoritmit vankkoja ja laajalti sovellettavissa monissa erilaisissa optimointimaisemissa. Tässä artikkelissa kehitämme kokoelman yksikkötestejä stokastiselle optimoinnille. Kukin yksikkötesti arvioi nopeasti optimointialgoritmin pienessä mittakaavassa, eristetyssä ja hyvin ymmärretyssä ongelmassa, eikä todellisissa skenaarioissa, joissa monet tällaiset ongelmat kietoutuvat toisiinsa. Näiden yksikkötestien läpäiseminen ei ole riittävää, mutta ehdottoman välttämätöntä kaikille algoritmeille, joilla on väitteitä yleisyydestä tai kestävyydestä. Esitämme alustavia kvantitatiivisia ja kvalitatiivisia tuloksia kymmenestä vakiintuneesta algoritmista. Testauskehys on avoimen lähdekoodin, laajennettavissa ja helposti sovellettavissa uusiin algoritmeihin.

**Tulos**

Stokastisen optimoinnin yksikkötestit

**Esimerkki 2.702**

Langattomien sensoriverkkojen yleistyessä teollisuudessa on syntymässä monia tutkimusmahdollisuuksia. Yksi tällainen olennainen ja haastava sovellus on solmujen paikannus. Tässä asiakirjassa käytetään feed-forward-neuraaliverkkoon perustuvaa menetelmää. Käytetään ankkurisolmun majakoiden RSSI-arvoja (Received Signal Strength Indicator). Ankkurisolmujen määrä ja niiden konfiguraatiot vaikuttavat paikannusjärjestelmän tarkkuuteen, mitä käsitellään myös tässä asiakirjassa. Viittä eri koulutusalgoritmia arvioidaan parhaan tuloksen antavan koulutusalgoritmin löytämiseksi. Monikerroksisen perceptronin (MLP) neuroverkkomalli koulutettiin Matlabilla. Ehdotetun menetelmän suorituskyvyn arvioimiseksi reaaliajassa saatu malli toteutettiin sitten Arduino-mikrokontrollerilla. Neljän ankkurisolmun avulla saavutettiin keskimääräinen 2D-paikannusvirhe 0,2953 m 12-12-2 neuroverkon rakenteella. Ehdotettu menetelmä voidaan toteuttaa myös missä tahansa muussa sulautetussa mikrokontrollerijärjestelmässä.

**Tulos**

LANGATTOMIEN SENSORIVERKKOJEN PAIKANNUS: NEUROVERKKO-LÄHESTYMISTAPA

**Esimerkki 2.703**

Monissa sulautetuissa järjestelmissä, kuten kuvantamisjärjestelmissä, järjestelmällä on vain yksi käyttötarkoitus, ja samoja säikeitä suoritetaan toistuvasti. Säikeiden käyttäytymisen profiloinnin avulla järjestelmä voi jakaa kunkin säikeen resurssit siten, että järjestelmän kokonaissuorituskyky paranee. Tutkimme online-resurssienjako-ongelmaa, jossa resurssienhallinta jakaa resursseja samanaikaisesti (tutkimus), oppii vaikutuksen eri kuluttajiin (oppiminen) ja parantaa jakoa kohti optimaalista suorituskykyä (hyödyntäminen). Rakennamme monikätisten bandiittien monipuoliseen kehykseen ja esittelemme online- ja offline-algoritmeja. Näytämme algoritmiemme ansiot ja ominaisuudet laajoilla kokeilla, joissa käytetään sekä synteettistä dataa että todellista välimuistin jakamista säikeille.

**Tulos**

Rosvot kohtaavat tietokonearkkitehtuurin: Cache: Älykkäästi jaetun välimuistin suunnittelu

**Esimerkki 2.704**

Vahvistusoppimisen arvofunktioiden estimoinnin innoittamana tutkimme tilastollisia lineaarisia käänteisongelmia eli ongelmia, joissa ratkaistavan lineaarisen järjestelmän kertoimet havaitaan kohinassa. Tarkastelemme rangaistuja estimaattoreita, joiden suorituskykyä arvioidaan käyttämällä estimaattorin virheen matriisipainotettua kaksinormaalinormia mitattuna suhteessa todellisiin, tuntemattomiin kertoimiin. Tarkasteltavana on kaksi tavoitefunktiota riippuen siitä, onko kohinaisten kertoimien suhteen mitattu virhe neliöllinen vai neliötön. Ehdotamme yksinkertaisia, mutta uusia ja teoreettisesti hyvin perusteltuja datasta riippuvia valintoja regularisointiparametreille molemmissa tapauksissa, joilla vältetään datan pilkkominen. Analyysimme erottava piirre on, että johdamme deterministiset virherajat kertoimien virheen suhteen, jolloin näiden virheiden stokastisten ominaisuuksien analyysi voidaan täysin erottaa. Osoitamme, että tuloksemme johtavat uusiin oivalluksiin ja rajoihin lineaarisen arvofunktion estimoinnissa vahvistusoppimisessa.

**Tulos**

Tilastollinen lineaarinen estimointi rangaistujen estimaattoreiden avulla: sovellus vahvistusoppimiseen.

**Esimerkki 2.705**

Keskustelemme agentin toimien ajasta riippuvaista hyödynnettävyyttä koskevan tiedon esittämisestä ja päättelystä. Ajasta riippuva hyöty on ratkaisevassa asemassa laskennan ja toiminnan välisessä vuorovaikutuksessa rajallisilla resursseilla. Esitämme semantiikan ajasta riippuvalle hyödylle ja kuvaamme ajasta riippuvan tiedon käyttöä päätöksentekokonteksteissa. Havainnollistamme keskusteluamme esimerkeillä ajasta riippuvaisesta päättelystä Protos-järjestelmässä, joka on rakennettu tutkimaan päättelyn ihanteellista hallintaa päättelijöillä, joilla on rajalliset kyvyt.

**Tulos**

Ajasta riippuva hyöty ja toiminta epävarmuuden vallitessa

**Esimerkki 2.706**

Tässä artikkelissa ehdotamme korkea-ulotteisten tietojen universaalimallia, jota kutsutaan nimellä Hybrid Orthogonal Projection and Estimation (HOPE) -malli, jossa yhdistyvät lineaarinen ortogonaalinen projektio ja äärellinen seosmalli yhtenäisessä generatiivisessa mallinnuskehyksessä. Itse HOPE-malli voidaan oppia ilman valvontaa merkitsemättömästä datasta maksimiluotettavuusestimaatioon perustuen sekä kouluttaa erottelevasti merkityn datan perusteella. Kiinnostavampaa on, että olemme osoittaneet ehdotettujen HOPE-mallien olevan läheisessä yhteydessä neuroverkkoihin siinä mielessä, että jokainen piilokerros voidaan muotoilla uudelleen HOPE-malliksi. Tämän seurauksena HOPE-kehystä voidaan käyttää uutena välineenä, jolla voidaan tutkia, miksi ja miten NN:t toimivat, ja mikä tärkeämpää, se tarjoaa myös useita uusia oppimisalgoritmeja NN:ien oppimiseen joko valvotusti tai valvomattomasti. Tässä työssä olemme tutkineet HOPE-kehystä NN:ien oppimisessa useissa vakiotehtävissä, kuten kuvantunnistuksessa MNIST-tietokannassa ja puheentunnistuksessa TIMIT-tietokannassa. Kokeelliset tulokset osoittavat, että HOPE-kehys tuottaa huomattavaa suorituskykyparannusta nykyisiin uusimpiin menetelmiin verrattuna erityyppisissä NN-oppimisongelmissa, kuten valvomattomassa ominaisuuksien oppimisessa, valvotussa tai puolivalvotussa oppimisessa.

**Tulos**

Hybrid Orthogonal Projection and Estimation (HOPE): Uudet puitteet neuroverkkojen koetteluun ja oppimiseen

**Esimerkki 2.707**

Aiemmissa töissä [BGHK92, BGHK93] olemme tutkineet satunnaismaailmojen lähestymistapaa� erityistä (ja varsin tehokasta) menetelmää uskomusasteiden (eli subjektiivisten todennäköisyyksien) tuottamiseksi objektiivisesta (ensimmäisen asteen, slalisti kal ja defaull) tiedosta koostuvasta tietopohjasta. Mutta se, että tietopohjan annetaan sisältää vain objektiivista tietoa, on joskus rajoittavaa. Haluamme toisinaan sisällyttää tietopohjaan myös tietoa uskomusasteista, koska on tilanteita, joissa vanhat uskomukset edustavat tärkeää tietoa, jonka pitäisi vaikuttaa uusiin uskomuksiin. Tässä artikkelissa kuvaamme kolme melko yleistä tekniikkaa, joilla voidaan laajentaa menetelmä, joka tuottaa uskomusasteita objektiivisesta tiedosta, menetelmäksi, joka voi hyödyntää myös uskomusasteita. Kaikki tekniikkamme perustuvat tunnettuihin lähestymistapoihin, kuten a5 ristiemropiaan. Keskustelemme tekniikoiden välisistä yleisistä yhteyksistä ja erityisesti osoitamme, että vaikka tekniikat ovat käsitteellisesti ja teknisesti varsin erilaisia, kaikki tekniikat antavat saman vastauksen, kun niitä sovelletaan satunnaismaailman menetelmään.

**Tulos**

Uusien uskomusten luominen vanhoista\*

**Esimerkki 2.708**

Tutustuminen on ollut keskeinen osa vahvistusoppimista, mutta nykyiset analyyttiset kehykset eivät vieläkään anna tyydyttäviä vastauksia useisiin tärkeisiin kysymyksiin, jotka koskevat tutustumisen tehokkuutta. Näihin kysymyksiin kuuluvat eksploraatioparametrien asettaminen, tilanneanalyysi ja MDP:iden kovuus, jotka kaikki ovat käytännön toimijoiden kannalta väistämättömiä. Teorian ja käytännön välisen kuilun kuromiseksi umpeen ehdotamme uutta analyyttistä kehystä, jota kutsutaan etsinnän onnistumistodennäköisyydeksi. Osoitamme, että kaikkiin edellä mainittuihin tärkeisiin etsintäkysymyksiin voidaan vastata kehyksellämme, ja kehyksemme antamat vastaukset vastaavat käytännön toimijoiden tarpeita paremmin kuin nykyiset vastaukset. Vielä tärkeämpää on, että esittelemme konkreettisen ja käytännöllisen lähestymistavan menestystodennäköisyyksien arvioimiseksi tietyissä MDP:issä ilman, että oppimisalgoritmia tarvitsee itse asiassa ajaa. Sen jälkeen esitämme empiirisiä tuloksia lähestymistapamme todentamiseksi ja osoitamme, miten etsinnän onnistumistodennäköisyyttä voidaan käyttää analysoimaan ja ennustamaan etsinnän käyttäytymistä ja mahdollisia tuloksia, jotka ovat avaimia etsinnän tärkeisiin kysymyksiin vastaamiseen.

**Tulos**

Tutkimuksen onnistumisen todennäköisyys: oppimisen tehokkuuden konkreettinen analyysi.

**Esimerkki 2.709**

Tilastollisten aihepiirimallien käyttäjille on aina ollut rasitteena aihepiirien oikean lukumäärän määrittäminen, joka on useimpien aihepiirimallien keskeinen parametri. Tavallisesti tämän parametrin automaattinen valinta tehdään joko tilastollisen mallivalinnan (esim. ristiinvalidointi, AIC tai BIC) tai Bayesin ei-parametristen mallien (esim. hierarkkinen Dirichlet-prosessi) avulla. Nämä menetelmät perustuvat joko päättelyalgoritmin toistuviin ajoihin, joilla etsitään suuren parametrin arvojen vaihteluväliä, mikä ei sovi suurten tietojen louhintaan, tai ne korvaavat kyseisen parametrin vaihtoehtoisilla parametreilla, jotka ovat vähemmän intuitiivisia ja edelleen vaikeasti määritettävissä. Tässä asiakirjassa tutkimme tämän parametrin "poistamista" uudesta näkökulmasta. Esittelemme ensin PLSA-mallin ei-parametrisen käsittelyn, joka on nimeltään ei-parametrinen todennäköisyyspohjainen latentti semanttinen analyysi (nPLSA). NPLSA:n päättelymenettely mahdollistaa eri aiheiden lukumäärän tutkimisen ja vertailun yhden suorituksen aikana, mutta pysyy silti yhtä yksinkertaisena kuin PLSA:n päättelymenettely. Tämä saavutetaan korvaamalla aiheiden lukumäärää koskeva parametri vaihtoehtoisella parametrilla, joka on asiakirjan minimaalinen hyvyys. Osoitamme, että uusi parametri voidaan edelleen poistaa kahdella parametrittomalla käsittelyllä: joko valvomalla löydettyjen aiheiden monimuotoisuutta tai heikolla käyttäjiltä saatavalla valvonnalla esimerkkiaiheen muodossa. Parametrivapaa aihealuemalli löytää sopivan määrän aihealueita, kun löydettyjen aihealueiden monimuotoisuus on maksimoitu tai kun löydettyjen aihealueiden rakeisuus vastaa esimerkkialuetta. Sekä synteettisillä että todellisilla aineistoilla tehdyt kokeet osoittavat, että parametrivapaa aihealuemalli poimii aihealueita vastaavalla laadulla kuin klassiset aihealuemallit, joissa käytetään "manuaalista siirtoa". Aiheiden laatu on parempi kuin klassisten Bayesin ei-parametristen mallien avulla poimittujen aiheiden laatu. ∗Tämä tutkimus on tehty ensimmäisen kirjoittajan ollessa vierailulla Michiganin yliopistossa. Lupa tehdä digitaalisia tai paperikopioita tästä teoksesta tai sen osasta henkilökohtaiseen tai luokkahuonekäyttöön myönnetään maksutta edellyttäen, että kopioita ei valmisteta tai levitetä voittoa tai kaupallista hyötyä varten ja että kopioissa on tämä ilmoitus ja täydellinen viittaus ensimmäisellä sivulla. Muunlainen kopiointi, uudelleenjulkaiseminen, palvelimille lähettäminen tai luetteloihin jakaminen edellyttää etukäteen annettua erityislupaa ja/tai maksua. Copyright 20XX ACM X-XXXXX-XX-X/XX/XX/XX ...10,00 dollaria.

**Tulos**

"Katso äiti, ei käsiä!" Parametriton aihepiirimalli

**Esimerkki 2.710**

Tietämyksen kokoaminen on lähestymistapa, jolla pyritään ratkaisemaan yleisten päättelyongelmien laskennallinen vaikeasti ratkaistavissa oleva ongelma. Tämän lähestymistavan mukaan tietämyskannat muunnetaan offline-toimintana kohdekieleksi, jota voidaan käyttää on-line-kyselyissä. Reduced ordered binary decision diagram (ROBDD) on yksi vaikutusvaltaisimmista kohdekielistä. Yleistämme ROBDD:tä liittämällä jokaiseen solmuun joitakin epäsuoria kirjaimia, ja uutta kieltä kutsutaan nimellä Reduced Ordered Binary Decision Diagram with Implied Literals (ROBDD-L). Sitten keskustelemme eräänlaisista ROBDD-L:n osajoukoista, joita kutsutaan nimellä ROBDD-i ja joissa on täsmälleen i implikoitua kirjainta (0 ≤ i ≤ ∞). Erityisesti ROBDD-0 on isomorfinen ROBDD:n kanssa; ROBDD-∞ edellyttää, että jokaiseen solmuun liittyy mahdollisimman monta implikoitua kirjainta. Osoitamme, että ROBDD-i on yksikäsitteinen tietyssä muuttujien järjestyksessä, ja ROBDD-∞ on ROBDD-L:n ytimekkäin osajoukko, ja se pystyy täyttämään useimmat tietämyskokoelmakarttaan liittyvät kyselyvaatimukset. Lopuksi ehdotamme ROBDD-i-kokoamisalgoritmia mille tahansa i:lle ja ROBDD-∞-kokoamisalgoritmia. Niiden perusteella toteutamme ROBDD-L-paketin nimeltä BDDjLu ja teemme joitakin johtopäätöksiä alustavista kokeellisista tuloksista: ROBDD-∞ on selvästi pienempi kuin ROBDD kaikkien vertailuarvojen osalta; ROBDD-∞ on pienempi kuin d-DNNF niiden vertailuarvojen osalta, joiden kääntämistulokset ovat suhteellisen pieniä; näyttää siltä, että on parempi muuntaa ROBDDs-∞ FBDD:ksi ja ROBDD:ksi kuin kääntää vertailuarvot suoraan.

**Tulos**

Vähennetty järjestetty binäärinen päätöskaavio, jossa on epäsuoria kirjaimia: Uuden tiedon kokoaminen lähestymistapa

**Esimerkki 2.711**

Maksatulehduksen diagnosointi alkuvaiheessa on tärkeää hoidon parantamiseksi. Nykyään infektioiden havaitsemiseen käytetään sensorien kaltaisia laitteita. Sairausnäytteiden automaattiseen tunnistamiseen tarvitaan tarkkoja luokittelutekniikoita. Tässä tutkimuksessa hyödynnetään tiedonlouhintamenetelmiä maksapotilaiden ja terveiden henkilöiden luokittelussa. Luokittelua varten toteutettiin neljä algoritmia (Naïve Bayes, Bagging, Random Forest ja SVM) käyttäen R-alustaa. Lisäksi luokittelun tarkkuuden parantamiseksi kehitettiin hybridi NeuroSVM-malli, jossa käytettiin SVM:ää ja feed-forward-keinohermoverkkoa (ANN). Hybridimallin suorituskykyä testattiin käyttämällä tilastollisia parametreja, kuten keskimääräistä neliövirheen juurta (RMSE) ja keskimääräistä absoluuttista prosentuaalista virhettä (MAPE). Mallin ennustustarkkuus oli 98,83 prosenttia. Tulosten mukaan hybridimallin kehittäminen paransi ennustustarkkuutta. Lääketieteellisen yhteisön palvelemiseksi potilaiden maksasairauksien ennustamisessa on kehitetty graafinen käyttöliittymä (GUI) käyttäen R:ää. GUI on otettu käyttöön pakettina R-alustan paikallisessa arkistossa, jotta käyttäjät voivat tehdä ennusteita.

**Tulos**

NeuroSVM: Graafinen käyttöliittymä maksapotilaiden tunnistamiseen.

**Esimerkki 2.712**

Lääketieteelliseen kuva-analyysiin ja tietokoneavusteisiin toimenpiteisiin liittyviin ongelmiin käytetään yhä useammin syväoppimiseen perustuvia ratkaisuja. Tämä on johtanut huomattavaan päällekkäiseen työhön ja yhteensopimattomaan infrastruktuuriin monissa tutkimusryhmissä. Tässä työssä esitellään avoimen lähdekoodin NiftyNet-alusta lääketieteellisen kuvantamisen syväoppimista varten. NiftyNetin tavoitteena on nopeuttaa ja yksinkertaistaa tällaisten ratkaisujen kehittämistä ja tarjota yhteinen mekanismi tutkimustulosten levittämiseksi yhteisön käyttöön, mukauttamiseksi ja hyödyntämiseksi. Tämä TensorFlow-pohjainen infrastruktuuri tarjoaa täydellisen modulaarisen syväoppimisputken erilaisia lääketieteellisen kuvantamisen sovelluksia varten, mukaan lukien segmentointi-, regressio-, kuvantuotanto- ja esitysoppimissovellukset, joissa on datan lataus, datan lisääminen, verkkoarkkitehtuurit, häviöfunktiot ja arviointimittarit, jotka on räätälöity lääketieteellisen kuva-analyysin ja tietokoneavusteisten toimenpiteiden erityispiirteisiin ja joissa niitä hyödynnetään.

**Tulos**

NiftyNet: syväoppimisalusta lääketieteelliseen kuvantamiseen

**Esimerkki 2.713**

Yhteistoimintapeleissä mallinnetaan yhteisistä toimista saatavan voiton jakamista esimerkiksi vakauden ja oikeudenmukaisuuden näkökulmasta. Ehdotamme tällaisten pelien luotettavuuslaajennusta, jossa toimijat voivat jättää osallistumatta peliin. Luotettavuuslaajennuksessa kukin agentti "selviytyy" vain tietyllä todennäköisyydellä, ja koalition arvo on todennäköisyys, että sen eloonjääneet jäsenet olisivat peruspelissä voittava koalitio. Tutkimme tällaisten pelien merkittäviä ratkaisukonsepteja ja osoitamme, miten Shapleyn arvoa voidaan approksimoida ja miten ydin voidaan laskea peleissä, joissa on vain vähän agenttityyppejä. Osoitamme myös, että luotettavuuslaajennuksen soveltaminen voi vakauttaa pelin, jolloin ydin ei ole tyhjä, vaikka peruspelissä olisi tyhjä ydin.

**Tulos**

Yhteistoiminnallisten luotettavuuspelien ratkaiseminen

**Esimerkki 2.714**

Aiemmissa tutkimuksissa on ehdotettu kuvapohjaisia epäjärjestysmittareita, jotka korreloivat ihmisen hakuaikojen ja/tai silmänliikkeiden kanssa. Useimmissa malleissa ei kuitenkaan oteta huomioon sitä, että epäjärjestyksen vaikutukset ovat vuorovaikutuksessa ihmisen näköjärjestelmän foveaalisen luonteen kanssa: kauempana foveasta oleva visuaalinen epäjärjestys vaikuttaa yhä haitallisemmin havaitsemiseen. Tässä esitellään uusi foveated clutter -malli, jolla ennustetaan haitallisia vaikutuksia kohteen etsinnässä käyttämällä pakotettua fiksaatiohakutehtävää. Käytämme Feature Congestion -mallia (Rosenholtz et al.) muuna kuin foveoidun häiriön mallina, ja foveoidun häiriön mallissa käytämme perifeeristä arkkitehtuuria Feature Congestionin päälle. Otamme käyttöön PIFC-kertoimen (Peripheral Integration Feature Congestion), joka on mallimme peruskomponentti ja joka moduloi häiriötä epälineaarisena vahvistuksena, joka riippuu eksentrisyydestä. Osoitamme lopuksi, että Foveated Feature Congestion (FFC) -mallin häiriöpisteet (r(44) = -0,82 ± 0,04, p < 0,0001) korreloivat paremmin kohteen havaitsemisen (osumaprosentti) kanssa kuin tavallinen Feature Congestion (r(44) = -0,19 ± 0,13, p = 0,0774) pakotetussa fiksaatiohaussa. Näin ollen mallimme avulla voimme rikastuttaa häiriön havaitsemisen tutkimusta laskemalla fiksaatiokohtaisia häiriökarttoja. Työkalupakki perifeeristen arkkitehtuurien luomiseen: Piranhat: Perifeeriset arkkitehtuurit luonnollisille, hybrideille ja keinotekoisille järjestelmille1.

**Tulos**

Voivatko perifeeriset representaatiot parantaa epäselvyysmittareita monimutkaisissa maisemissa?

**Esimerkki 2.715**

Web 2.0:n tulo on lisännyt verkossa saatavilla olevan tunteikkaan sisällön määrää. Tällaista sisältöä löytyy usein sosiaalisen median verkkosivuilta elokuva- tai tuotearvostelujen, käyttäjäkommenttien, suosittelujen, keskustelupalstojen viestien jne. muodossa. Tunteita herättävän tai mielipiteitä ilmaisevan verkkosisällön oikea-aikaisella löytämisellä on useita etuja, joista tärkein on rahaksi muuttaminen. Ihmisjoukkojen tunteiden ymmärtäminen eri tahoja ja tuotteita kohtaan mahdollistaa paremmat palvelut kontekstisidonnaisia mainoksia, suositusjärjestelmiä ja markkinasuuntausten analysointia varten. Hankkeessamme keskitytään sentimentteihin keskittyvään verkkokierrätyspuitteistoon, joka helpottaa elokuva- ja hotelliarvostelujen tunteisiin perustuvan sisällön nopeaa löytämistä ja niiden analysointia. Käytämme tilastollisia menetelmiä subjektiivisen tyylin ja lauseen polariteetin elementtien tallentamiseen. Artikkelissa käsitellään yksityiskohtaisesti kahta valvottua koneoppimisalgoritmia: K-Nearest Neighbour (K-NN) ja Naïve Bayes' ja verrataan niiden kokonaistarkkuutta, tarkkuutta ja palautusarvoja. Todettiin, että elokuva-arvostelujen tapauksessa Naïve Bayes antoi huomattavasti paremmat tulokset kuin K-NN, mutta hotelliarvostelujen tapauksessa molemmat algoritmit antoivat vähemmän, lähes samat tulokset.

**Tulos**

Arvostelutietokantojen tunneanalyysi Naïve Bayesin ja K-NN-luokittimen avulla

**Esimerkki 2.716**

Nykyään syväarkkitehtuurien käyttö koneoppimisessa on hyvin suosittua. Deep Belief Networks (DBN) ovat syväarkkitehtuureja, jotka käyttävät Restricted Boltzmann Machines (RBM) -pinoja tehokkaan generatiivisen mallin luomiseen harjoitusdatan avulla. Tässä artikkelissa esitellään parannus yleiseen menetelmään, jota yleensä käytetään RBM:ien koulutuksessa. Uudessa menetelmässä käytetään vapaata energiaa kriteerinä, jonka avulla generatiivisesta mallista saadaan eliittinäytteitä. Väitämme, että näillä näytteillä voidaan laskea tarkemmin harjoitusdatan log-todennäköisyyden gradientti. Tulosten mukaan MNIST-testijoukossa saavutettiin 0,99 prosentin virhetaso. Tämä tulos osoittaa, että ehdotettu menetelmä on parempi kuin ensimmäisessä artikkelissa esitelty menetelmä, jossa esiteltiin DBN (1,25 prosentin virhetaso) ja yleiset luokittelumenetelmät, kuten SVM (1,4 prosentin virhetaso) ja KNN (1,6 prosentin virhetaso). Toisessa testissä, jossa käytettiin ISOLET-tietokokonaisuutta, kirjeiden luokitteluvirhe laski 3,59 prosenttiin verrattuna 5,59 prosentin virhetasoon, joka saavutettiin kyseistä tietokokonaisuutta käyttävissä julkaisuissa. Toteutettu menetelmä on saatavilla verkossa osoitteessa "http://ceit.aut.ac.ir/~keyvanrad/DeeBNet Toolbox.html".

**Tulos**

Deep Belief Network -koulutuksen parantaminen käyttämällä Elite-näytteitä vapaan energian minimoimiseksi

**Esimerkki 2.717**

Monissa sovelluksissa sanoilla kuvatut ajatukset liikkuvat usein eri ryhmien välillä. Helpottaaksemme käyttäjiä virtauksen analysoinnissa esitämme virtauskäyttäytymisen mallintamiseen menetelmän, jonka tavoitteena on tunnistaa eri käyttäjäryhmien sanaryhmien väliset lead-lag-suhteet. Dynaamiseen aikavääristykseen perustuvaa parannettua Bayesin ehdollista kointegraatiota käytetään erityisesti eri ryhmien sanojen välisten yhteyksien oppimiseen. Kehitetään tensoripohjainen tekniikka, jolla nämä linkitetyt sanat voidaan ryhmitellä eri klustereihin (ideoihin) ja seurata ideoiden virtausta. Tensoriesityksen tärkein piirre on se, että otamme käyttöön kaksi lisäulottuvuutta sekä ajan että viiveen ja etumatkan suhteiden esittämiseksi. Sekä synteettisillä että todellisilla tietokokonaisuuksilla tehdyt kokeet osoittavat, että menetelmämme on tehokkaampi kuin perinteisiin klusterointitekniikoihin perustuvat menetelmät ja että sillä saavutetaan parempi tarkkuus. Tapaustutkimuksen avulla osoitettiin menetelmämme hyödyllisyys, kun se auttaa käyttäjiä ymmärtämään ideoiden virtausta eri käyttäjäryhmien välillä sosiaalisessa mediassa.

**Tulos**

Sosiaalisten ryhmien välisten ideavirtojen seuranta

**Esimerkki 2.718**

Yhdistämme Riemannin geometriaa ja korkea-ulotteisen kaaoksen keskikenttäteoriaa tutkiaksemme signaalin etenemisen luonnetta geneerisissä syvissä neuroverkoissa, joissa on satunnaispainot. Tuloksemme paljastavat järjestyksen ja kaaoksen välisen ekspressiivisyyden välisen vaiheen siirtymän, jossa kaoottisessa vaiheessa olevat verkot laskevat epälineaarisia funktioita, joiden globaali kaarevuus kasvaa eksponentiaalisesti syvyyden mutta ei leveyden myötä. Osoitamme, että tätä syvien satunnaisfunktioiden yleistä luokkaa ei voida laskea tehokkaasti millään matalalla verkolla, mikä ylittää aiemmat yksittäisten funktioiden analyysiin rajoittuvat työt. Lisäksi formalisoimme ja osoitamme kvantitatiivisesti pitkään arvellun ajatuksen siitä, että syvät verkot voivat purkaa erittäin kaarevat moninaisuudet syöttöavaruudessa tasaisiksi moninaisuuksiksi piilotetussa avaruudessa. Teoreettinen analyysimme syväverkkojen ilmaisuvoimasta soveltuu laajasti mielivaltaisiin epälineaarisuuksiin ja tarjoaa kvantitatiivisen perustan aiemmin abstrakteille käsityksille syväfunktioiden geometriasta.

**Tulos**

Eksponentiaalinen ekspressiivisyys syvissä neuroverkoissa transienttisen kaaoksen avulla

**Esimerkki 2.719**

Tutkitaan assosiatiivisen muistin mallia, joka tallentaa ja hakee luotettavasti paljon enemmän kuvioita kuin verkossa on neuroneita. Ehdotamme yksinkertaista dualismia tämän tiheän assosiatiivisen muistin ja syväoppimisessa yleisesti käytettyjen neuroverkkojen välille. Tämän dualiteetin assosiatiivisen muistin puolella voidaan rakentaa malliperhe, joka interpoloi sujuvasti kahden rajatapauksen välillä. Toista raja-arvoa kutsutaan hahmontunnistuksen piirteiden yhteensovitustilaksi ja toista prototyyppijärjestelmäksi. Dualiteetin syväoppimisen puolella tämä perhe vastaa feedforward-neuraaliverkkoja, joissa on yksi piilotettu kerros ja erilaisia aktivointifunktioita, jotka välittävät näkyvien neuronien toiminnan piilotettuun kerrokseen. Tähän aktivointifunktioiden perheeseen kuuluvat logaritmiat, suoristetut lineaariset yksiköt ja korkeamman asteen suoristetut polynomit. Ehdotetun dualiteetin avulla voidaan soveltaa assosiatiivisesta muistista peräisin olevaa energiapohjaista intuitiota analysoitaessa sellaisten neuroverkkojen laskennallisia ominaisuuksia, joissa on epätavallisia aktivointifunktioita - korkeammat suoristetut polynomit, joita ei tähän asti ole käytetty neuroverkkojen kouluttamiseen. Tiheiden muistien hyödyllisyyttä havainnollistetaan kahdessa testitapauksessa: looginen portti XOR ja käsinkirjoitettujen numeroiden tunnistaminen MNIST-tietoaineistosta.

**Tulos**

Tiheä assosiatiivinen muisti kuvantunnistusta varten

**Esimerkki 2.720**

Havainnointitutkimukset perustuvat ihmisen tilan tarkkaan arviointiin. Käyttäytymisen tunnistusjärjestelmä, joka mallintaa keskustelijoiden tilaa reaaliaikaisesti, voi auttaa merkittävästi mielenterveysalaa. Käyttäytymisen tunnistaminen puheesta on kuitenkin edelleen haastava tehtävä, koska yleistettävien ja edustavien piirteiden löytäminen on vaikeaa kohinan ja korkea-ulotteisen datan vuoksi, varsinkin kun data on rajallista ja kommentoitu karkeasti ja subjektiivisesti. Syvät neuroverkot (Deep Neural Networks, DNN) ovat osoittautuneet lupaaviksi monissa koneoppimistehtävissä, mutta käyttäytymissignaalien käsittelyyn (Behavioral Signal Processing, BSP) niiden käyttöä on rajoitettu rajallisen datamäärän vuoksi. Ehdotamme harvaan kytkettyä ja epäjohdonmukaisesti koulutettua DNN:ää (SD-DNN), jolla voidaan käsitellä rajallista dataa. Ensin jaamme akustiset ominaisuudet osajoukkoihin ja koulutamme useita erillisiä luokittelijoita. Sitten näiden luokittelijoiden piilokerrokset tulevat osaksi syvempää verkkoa, joka integroi kaikki piirrevirrat. Kokonaisjärjestelmä mahdollistaa täydellisen kytkeytyvyyden ja rajoittaa samalla kerrallaan koulutettavien parametrien lukumäärää, ja konvergenssi on mahdollista jopa rajallisella datalla. Esitämme tuloksia useista käyttäytymiskoodeista pariterapia-alalla ja osoitamme edut käyttäytymisen luokittelutarkkuudessa. Osoitamme myös tämän järjestelmän elinkelpoisuuden live-käyttäytymisen merkintöjä varten.

**Tulos**

Harvaan kytketyt ja epäjohdonmukaisesti koulutetut syvät neuroverkot käyttäytymismerkintöjen merkitsemiseen vähäisillä resursseilla: Akustinen luokittelu pariterapiassa: Akustinen luokittelu pariterapiassa

**Esimerkki 2.721**

Verrattuna sana- ja lausetason konvoluutiohermoverkkoihin (ConvNets) merkkitason ConvNets soveltuu paremmin kirjoitus- ja kirjoitusvirheiden käsittelyyn. Tämän vuoksi viimeaikaisissa tekstiluokittelua koskevissa tutkimuksissa on keskitytty pääasiassa merkkitason konvenkronoverkkoihin. Suurin osa näistä tutkimuksista käyttää kuitenkin englanninkielistä korpusta merkkitason tekstiluokitteluun, mutta vain muutama tutkimus on tehty käyttäen kiinalaista korpusta. Tässä tutkimuksessa pyritään kuromaan umpeen tämä aukko tutkimalla merkkitason ConvNets-verkkoja kiinalaisen korpuksen testiluokittelua varten. Olemme rakentaneet laajamittaisen kiinalaisen tietokokonaisuuden, ja tulos osoittaa, että merkkitason ConvNets toimii paremmin kiinalaisessa merkkitietokokonaisuudessa kuin vastaavassa pinyin-muotoisessa tietokokonaisuudessa, joka on aiemmissa tutkimuksissa käytetty yleinen ratkaisu. Tämä on ensimmäinen kerta, kun merkkitason ConvNets-verkkoja on sovellettu kiinalaisten merkkien tietokokonaisuuteen tekstin luokittelussa.

**Tulos**

Merkkitason konvoluutioverkko tekstin luokitteluun sovellettuna kiinalaiseen korpukseen

**Esimerkki 2.722**

Ehdotamme StochAstic Fault diagnosis AlgoRIthmia, nimeltään Safari, jossa minimidiagnoosien laskentatakuita vaihdetaan laskennalliseen tehokkuuteen. Osoitamme empiirisesti käyttämällä 74XXX- ja ISCAS85-vertailupiirien yhdistelmäpiirisarjoja, että Safari nopeuttaa useiden suuruusluokkien verran kahden tunnetun deterministisen algoritmin, CDA∗:n ja HA∗:n, nopeutta monivikadiagnoosien tekemisessä; lisäksi Safari pystyy laskemaan useita monivikadiagnooseja, joita CDA∗ ja HA∗ eivät pysty laskemaan. Todistamme myös, että Safari on optimaalinen useille propositionaalisille vikamalleille, kuten laajalti käytetyille heikkojen vikojen malleille (mallit, joissa ei tiedetä epänormaalista käyttäytymisestä). Keskustelemme Safarin optimaalisuudesta sellaisten vahvan vian piirimallien luokassa, joissa on juuttuneita vikatiloja. Mallintamalla itse algoritmi Markovin ketjuna annamme tarkat rajat lasketun diagnoosin minimille. Safari käyttäytyy myös vahvasti milloin tahansa ja palauttaa diagnoosin minkä tahansa ei-triviaalin päättelyajan jälkeen.

**Tulos**

Likimääräinen mallipohjainen diagnoosi ahneen stokastisen haun avulla

**Esimerkki 2.723**

Tässä artikkelissa käsittelemme uutta lähestymistapaa todennäköisyyspohjaiseen päättelyyn loogisessa kehyksessä. Lähes kaikki kirjallisuudessa ehdotetut todennäköisyyslogiikat perustuvat klassiseen kaksiarvologiikkaan. Kun sumean logiikan ja todennäköisyysteorian väliset erot on ensin selvitetty, ehdotamme tässä sumeaa todennäköisyyslogiikkaa, jolle esitetään täydellisyystulokset (todennäköisyysteoreettisessa mielessä). Lähestymistavan pääajatus on, että terävien lauseiden todennäköisyysarvot voidaan ymmärtää joidenkin sopivien sumeiden lauseiden totuusarvoina teräviin lauseisiin nähden. Lisäksi annetaan sug gestiotlS ja esimerkkejä siitä, miten formalismia voidaan laajentaa niin, että se selviytyy ehdollisista todennäköisyyksistä ja muista epävarmuusformalismeista.

**Tulos**

Sumea logiikka ja todennäköisyys

**Esimerkki 2.724**

Chilessä ei ole riippumatonta tahoa, joka julkaisisi määrällisiä tai laadullisia tutkimuksia perinteisen mediaympäristön ja sen mukauttamisen ymmärtämiseksi sosiaalisessa verkossa. Nykyään chileläiset uutistenlukijat käyttävät yhä useammin sosiaalisia verkkoalustoja ensisijaisena tietolähteenään, joista Twitter on keskeisessä asemassa. Historialliset tiedotusvälineet ja puhtaat toimijat kehittävät erilaisia strategioita lisätäkseen yleisöään ja vaikutusvaltaansa tällä alustalla. Tässä artikkelissa ehdotamme tiedonlouhintatekniikoihin perustuvaa menetelmää, jonka avulla voidaan analysoida ensimmäistä kertaa Chilen uutta mediaympäristöä. Käytämme ryömintätekniikkaa 37 eri chileläisen tiedotusvälineen uutisvirtojen louhimiseen, jotka ovat aktiivisesti esillä Twitterissä, ja ehdotamme useita indikaattoreita niiden vertailemiseksi. Analysoimme niiden tuotantomääriä, potentiaalista yleisöä ja tutkimme NLP-tekniikoiden avulla niiden tuotannon sisältöä: toimituksellista linjaa ja maantieteellistä kattavuutta.

**Tulos**

Chileläisten tiedotusvälineiden toimituksellisten strategioiden diagnosointi Twitterissä automaattisen uutisluokittelijan avulla

**Esimerkki 2.725**

Käytännössä kaikki ohjelmointikielet antavat ohjelmoijalle mahdollisuuden jakaa ohjelma useisiin moduuleihin, mikä tuo mukanaan useita etuja ohjelmistokehityksessä. Tässä artikkelissa olemme kiinnostuneita vastausjoukkojen ohjelmoinnin alueesta, jossa käytetään täysin deklaratiivisia ja ei-monotonisia kieliä. Tässä yhteydessä modulaarisen rakenteen saaminen ohjelmille ei ole suinkaan yksinkertaista, koska kokonaisen ohjelman tulosta ei yleensä voida koostaa sen osien tuloksista. Ymmärtääksemme paremmin disjunktiivisen informaation vaikutuksia modulaarisuuteen rajoitamme analyysin koskemaan disjunktiivisia logiikkaohjelmia (DLP), joihin sovelletaan stabiilin mallin semantiikkaa. Määrittelemme DLP-funktion käsitteen, jossa on hyvin määritelty tulo/lähtö-rajapinta, ja laadimme uuden moduuliopin, joka osoittaa DLP-funktioiden stabiilimallisemantiikan kompositionaalisuuden. Moduuliteoreemi laajentaa tunnettua jakojoukkoteoriaa ja mahdollistaa DLP-funktioiden hajottamisen niiden vahvasti kytkeytyneiden komponenttien perusteella, jotka perustuvat sääntöjen aiheuttamiin positiivisiin riippuvuussuhteisiin. Tässä ympäristössä on myös mahdollista jakaa jaetut disjunktiiviset säännöt komponenttien kesken käyttämällä yleistettyä siirtotekniikkaa. Modulaarisen ekvivalenssin käsite otetaan käyttöön DLP-funktioiden keskinäistä vertailua varten käyttämällä käännöspohjaisen verifiointimenetelmän yleistystä.

**Tulos**

Disjunktiivisten stabiilien mallien modulaarisuusnäkökohdat

**Esimerkki 2.726**

Tässä työssä tutkimme taattua toimitusmallia, jota käytetään laajalti verkkonäytösten mainonnassa. Taatun toimituksen skenaariossa mainosnäytöt (joita joissakin teoksissa kutsutaan myös näyttökerroiksi) käyttäjille taataan mainostajien ja julkaisijoiden etukäteen tekemillä sopimuksilla. Keskeinen ongelma mainosalustalle on se, miten arvokas käyttäjäliikenne saadaan hyödynnettyä täysimääräisesti, jotta saadaan mahdollisimman paljon tuloja. Aiemmista töistä poiketen, joissa yleensä minimoidaan täyttämättä jääneiden sopimusten aiheuttama rangaistus ja jokin muu kustannus (esim. edustavuus), ehdotamme uudenlaista kulutuksen minimointimallia, jossa ensisijaisena tavoitteena on minimoida kaikkien sopimusten täyttämiseksi kulutettu käyttäjäliikenne. Tämän mallin mukaisesti kehitämme lähes optimaalisen menetelmän mainosten toimittamiseksi käyttäjille. Menetelmämme tärkein etu on se, että se kuluttaa lähes mahdollisimman vähän käyttäjäliikennettä kaikkien sopimusten täyttämiseksi, minkä vuoksi voidaan hyväksyä enemmän sopimuksia, jotka tuottavat enemmän tuloja. Sen avulla kustantajat voivat myös arvioida, kuinka suuri osa käyttäjäliikenteestä on tarpeetonta tai lyhyttä, jotta ne voivat myydä tai ostaa tämän osan liikenteestä irtotavarana vaihtomarkkinoilla. Lisäksi se on vankka, kun otetaan huomioon ennakkotieto käyttäjätyyppijakaumasta. Lopuksi simulointi osoittaa, että menetelmämme on parempi kuin perinteiset uusimmat menetelmät.

**Tulos**

Tehokas jakelukäytäntö käyttäjän liikennekulutuksen minimoimiseksi taatussa mainonnassa∗

**Esimerkki 2.727**

CSA-ES on evoluutiostrategia, jossa on kumulatiivinen askelkoon mukauttaminen ja jossa askelkokoa mukautetaan mittaamalla niin sanotun kumulatiivisen polun pituutta. Kumulatiivinen polku on algoritmin toteuttamien edellisten vaiheiden yhdistelmä, jossa kunkin vaiheen merkitys vähenee ajan myötä. Tässä artikkelissa tutkitaan CSA-ES:ää tiukasti kasvavien ja affiinisten lineaaristen funktioiden komposiiteissa tutkimalla sen taustalla olevia Markov-ketjuja. Askelkoon muutoksesta ja vaihtelusta johdetaan tarkkoja tuloksia kumulaation kanssa ja ilman kumulaatiota. Askelkoko divergoi geometrisesti nopeasti useimmissa tapauksissa. Lisäksi tutkitaan kumulaatioparametrin vaikutusta.

**Tulos**

Kumulatiivinen askelkoon mukauttaminen lineaarisille funktioille

**Esimerkki 2.728**

MiRNA- ja geeniekspressioprofiilit ovat osoittautuneet hyödyllisiksi syöpänäytteiden luokittelussa. Tehokkaita luokittelijoita on viime aikoina etsitty ja kehitetty. Kirjallisuudessa tunnetaan useita yrityksiä luokitella syöpänäytteitä miRNA- ja geeniekspressioprofiilien avulla. Viime aikoina bioinformatiikassa on kuitenkin käytetty puolivalvottuja oppimismalleja, joiden avulla on voitu hyödyntää valtavia julkisesti saatavilla olevia aineistoja. Sekä merkittyjen että merkitsemättömien joukkojen käyttämistä näyteluokittelijoiden kouluttamiseen ei ole aiemmin otettu huomioon, kun käytetään geeni- ja miRNA-ekspressiokokonaisuuksia. Lisäksi on olemassa motivaatio integroida sekä miRNA- että geeniekspressio puolivalvottua syöpäluokitusta varten, koska se antaa enemmän tietoa syöpänäytteiden ominaisuuksista. Tässä asiakirjassa kahta puolivalvottua koneoppimismenetelmää, nimittäin itseoppimista ja yhteiskoulutusta, mukautetaan syöpänäytteiden luokittelun laadun parantamiseksi. Näissä lähestymistavoissa hyödynnetään valtavia julkisia korpuksia koulutusdatan rikastuttamiseksi. Itseoppimisessa parannetaan miRNA- ja geenipohjaisia luokittelijoita itsenäisesti. Yhteiskoulutuksessa taas käytetään sekä miRNA- että geeniekspressioprofiileja samanaikaisesti, jotta saadaan erilaisia näkemyksiä syöpänäytteistä. Tietojemme mukaan tämä on ensimmäinen yritys soveltaa näitä oppimismenetelmiä syövän luokitteluun. Lähestymistapoja arvioitiin käyttämällä rintasyövän, hepatosellulaarisen karsinooman (HCC) ja keuhkosyövän ekspressiojoukkoja. Tulokset osoittavat jopa 20 prosentin parannusta F1-mitassa satunnaismetsä- ja SVM-luokittelijoihin verrattuna. Yhteiskoulutus päihittää myös Low Density Separation (LDS) -lähestymistavan noin 25 prosentin parannuksella F1-mitassa rintasyövässä. Avainsanat - miRNA- ja geeniekspressioanalyysi; puolivalvotut lähestymistavat; itseoppiminen; yhteiskoulutus; syöpänäytteiden luokittelijat.

**Tulos**

miRNA:han ja geeniekspressioon perustuva syövän luokittelu itseoppimisen ja yhteiskoulutuksen avulla

**Esimerkki 2.729**

Neuroverkkojen rakentaminen tietopohjan (taulukon) luonnollisen kielen avulla tapahtuvaa kyselyä varten on uusi tutkimusaihe NLP:ssä. Neuroverkkokysely vaatii tyypillisesti useita suoritusvaiheita, koska kyselyt ovat kompositionaalisia. Aiemmissa tutkimuksissa tutkijat ovat kehittäneet joko hajautettuja tai symbolisia kyselyohjelmia taulukkokyselyjä varten. Hajautettu tiedustelija on päästä päähän opittavissa, mutta se on heikko suoritustehokkuuden ja eksplisiittisen tulkittavuuden suhteen. Symbolinen kysyjä on sitä vastoin tehokas suorituksen aikana, mutta sitä on hyvin vaikea kouluttaa erityisesti alkuvaiheessa. Tässä artikkelissa ehdotamme hajautetun ja symbolisen suorituksen yhdistämistä luonnollisen kielen kyselyihin. Havainto on, että täysin hajautettu toteuttaja osoittaa myös mielekästä, vaikkakin epätäydellistä tulkintaa. Voimme siis esivalmentaa symbolisen suorittajan hajautetun suorittajan välitulosten avulla vaiheittain. Kokeet osoittavat, että lähestymistapamme on huomattavasti parempi kuin joko hajautettu tai symbolinen suoritin; lisäksi olemme löytäneet yli 80 prosenttia suoritussekvensseistä, joissa on harjoituksen aikana käytetty vain perustotuuden mukaisia denotaatioita. Yhteenvetona voidaan todeta, että kytketty neuraalinen tiedustelija hyödyntää sekä hajautetun että symbolisen toteuttajan etuja, ja sen suorituskyky, oppimistehokkuus, suoritustehokkuus ja tulkittavuus ovat korkeatasoisia.

**Tulos**

Hajautetun ja symbolisen suorituksen yhdistäminen luonnollisen kielen kyselyissä

**Esimerkki 2.730**

Tässä työssä esittelemme uuden lähestymistavan, jolla voidaan oppia pakattavia representaatioita syväarkkitehtuureissa, joissa käytetään päästä päähän -koulutusstrategiaa. Menetelmämme perustuu kvantisoinnin ja entropian pehmeään (jatkuvaan) relaksaatioon, jonka hehkutamme diskreeteiksi vastineiksi koko koulutuksen ajan. Esittelemme menetelmän kahdessa haastavassa sovelluksessa: Kuvien pakkaaminen ja neuroverkkojen pakkaaminen. Vaikka näitä tehtäviä on tavallisesti lähestytty eri menetelmillä, pehmeästä kovaan kvantisointiin perustuva lähestymistapamme antaa huippuluokan tuloksia molemmissa tapauksissa.

**Tulos**

Pehmeästä kovaan vektorikvantisointi kuvien ja hermoverkkojen oppimaan oppineeseen pakkaukseen alusta loppuun asti

**Esimerkki 2.731**

Luonnollisen kielen kysymyksiin vastaamisen automatisointi on mielenkiintoinen ja hyödyllinen ongelma. Kysymyksiin vastaamisen (QA) järjestelmät suorittavat usein tiedonhakua alkuvaiheessa. Lucenen kaltaisten hakukoneiden tarjoama tiedonhaun suorituskyky rajoittaa järjestelmän kokonaissuorituskykyä. Esimerkiksi lähes 40 prosentissa kysymyksistä haetaan asiakirjoja, joihin ei ole vastausta, alhaisilla sijoilla. Tässä asiakirjassa vastaustekstejä, jotka on saatu aikaisemmista laadunvarmistusarvioinneista, jotka on pidetty osana Text REtrieval Conferences (TREC) -konferenssia, on yhdistetty kyselyihin ja analysoitu, jotta voitaisiin tunnistaa suorituskykyä parantavia sanoja. Näitä sanoja käytetään sitten arvioitaessa kyselyn laajentamismenetelmän suorituskykyä. Tietoon perustuvien laajennussanojen havaittiin auttavan yli 70 prosentissa vaikeista kysymyksistä. Näitä sanoja voidaan käyttää kyselyjen laajentamismenetelmien parantamiseen ja arviointiin. Yksinkertaisen sokean relevanssipalautteen (RF) ennustettiin oikein, että se ei todennäköisesti auta kokonaissuorituskykyä, ja annetaan mahdollinen selitys sen vähäiselle arvolle IR:ssä laadunvarmistuksessa.

**Tulos**

Tietoon perustuva lähestymistapa kyselyiden laajentamiseen kysymyksiin vastaamisessa

**Esimerkki 2.732**

Tässä artikkelissa esitellään kevyt dynaaminen episteeminen looginen kehys automaattista suunnittelua varten epävarmuuden alkaessa. Pelkistämme suunnitelman verifioinnin ja konformisen suunnittelun logiikkamme mallintarkastusongelmiksi. Osoitamme, että iteraatiovapaan fragmentin mallintarkastusongelma on PSPACE-täydellinen. Käyttämällä kahta epätyypillistä (mutta ekvivalenttia) semantiikkaa annamme uusia mallintarkastusalgoritmeja koko kielelle ja iteraatiovapaalle kielelle.

**Tulos**

Dynaaminen episteeminen viitekehys yhdenmukaiselle suunnittelulle

**Esimerkki 2.733**

Rekursiivinen neuroverkko (Recurrent Neural Network, RNN) on yksi suosituimmista luonnollisen kielen prosessointitehtävissä (Natural Language Processing, NLP) käytetyistä arkkitehtuureista, koska sen rekursiivinen rakenne soveltuu hyvin vaihtelevan pituisen tekstin käsittelyyn. RNN voi hyödyntää sanojen hajautettuja esityksiä muuntamalla ensin kunkin tekstin sisältämät merkit vektoreiksi, jotka muodostavat matriisin. Tämä matriisi sisältää kaksi ulottuvuutta: aika-askeleen ulottuvuuden ja ominaisuusvektorin ulottuvuuden. Useimmat nykyiset mallit käyttävät yleensä yksiulotteista (1D) maksimipoolausoperaatiota tai huomiopohjaista operaatiota vain aika-askel-ulottuvuuteen saadakseen kiinteän pituisen vektorin. Ominaisuusvektoriulottuvuuden piirteet eivät kuitenkaan ole toisistaan riippumattomia, ja pelkkä 1D-pooling-operaation soveltaminen itsenäisesti aika-askelulottuvuuteen voi tuhota ominaisuuksien esityksen rakenteen. Toisaalta kaksiulotteisen (2D) pooling-operaation soveltaminen molempiin ulottuvuuksiin voi poimia merkityksellisempiä piirteitä sekvenssimallinnustehtäviä varten. Matriisin molempien ulottuvuuksien piirteiden integroimiseksi tässä asiakirjassa tutkitaan 2D-max-pooling-operaation soveltamista tekstin kiinteäpituisen esityksen saamiseksi. Tässä asiakirjassa hyödynnetään myös 2D-konvoluutiota matriisin mielekkäämpien tietojen poimimiseksi. Kokeet tehdään kuudella tekstiluokittelutehtävällä, mukaan lukien sentimenttianalyysi, kysymysten luokittelu, subjektiivisuuden luokittelu ja uutisryhmien luokittelu. Verrattuna uusimpiin malleihin ehdotetut mallit saavuttavat erinomaisen suorituskyvyn neljässä tehtävässä kuudesta. Erityisesti yksi ehdotetuista malleista saavuttaa parhaan tarkkuuden Stanford Sentiment Treebankin binäärisessä luokittelussa ja hienojakoisessa luokittelussa.

**Tulos**

Tekstien luokittelua parannetaan integroimalla kaksisuuntainen LSTM ja kaksiulotteinen Max Pooling -menetelmä.

**Esimerkki 2.734**

Kun käytössä on jo koulutettu neuroverkko, on usein toivottavaa pystyä lisäämään uusia ominaisuuksia ilman, että jo opittujen tehtävien suorittaminen vaikeutuu. Nykyiset lähestymistavat joko oppivat epäoptimaalisia ratkaisuja, vaativat yhteistä harjoittelua tai lisäävät huomattavasti parametrien määrää kutakin lisättyä tehtävää varten, yleensä yhtä paljon kuin alkuperäinen verkko. Ehdotamme menetelmää, joka säilyttää täysin suorituskyvyn alkuperäisessä tehtävässä vain pienellä lisäyksellä (noin 20 %) tarvittavien parametrien määrässä ja joka toimii samalla tavalla kuin kalliimmat hienosäätömenetelmät, jotka tyypillisesti kaksinkertaistavat parametrien määrän. Opittua arkkitehtuuria voidaan ohjata vaihtamaan eri opittujen representaatioiden välillä, jolloin yksi verkko voi ratkaista tehtävän useilta eri alueilta. Teemme laajoja kokeita, jotka osoittavat menetelmämme tehokkuuden, ja tutkimme sen käyttäytymisen eri näkökohtia.

**Tulos**

Inkrementaalinen oppiminen syväadaptaation avulla

**Esimerkki 2.735**

Tutkimme yhdistelmäkategorisen kieliopin (Combinatory Categorial Grammar, CCG) jäsennyksen monimutkaisuutta Vijay-Shankerin ja Weirin (1994) formalismilla. Päätuloksemme on osoittaa, että mikä tahansa jäsentelyalgoritmi tätä formalismia varten vie välttämättä eksponentiaalisen ajan, kun analyysissä otetaan huomioon kieliopin koko eikä vain syötelauseen pituus. Tämä tulos erottaa Vijay-Shankerin ja Weirin (1994) formalismin heikosti vastaavista formalismeista, kuten Tree-Adjoining Grammarista (TAG), jonka jäsennys voidaan suorittaa ajassa, joka on polynomiaalinen kieliopin ja syöttölauseen yhteenlasketun koon suhteen. Todistuksemme korostaa tärkeitä eroja Vijay-Shankerin ja Weirin (1994) formalismin ja CCG:n nykyisten versioiden välillä.

**Tulos**

CCG:n jäsennyksen monimutkaisuudesta

**Esimerkki 2.736**

Kaikissa tiedonhakuprosesseissa poimitun tiedon arvo on suoraan yhteydessä käytetyn tiedon laatuun. Myös Big Data -ongelmat, jotka johtuvat viime vuosina havaitusta datan laajuuden massiivisesta kasvusta, noudattavat samaa periaatetta. Yleinen ongelma, joka vaikuttaa datan laatuun, on kohina, erityisesti luokitteluongelmissa, joissa kohinalla tarkoitetaan harjoitustapausten virheellistä merkitsemistä, ja sen tiedetään olevan hyvin häiritsevä ominaisuus datassa. Big datan aikakaudella datan massiivinen kasvu asettaa kuitenkin haasteita perinteisille kohinan torjuntaan luoduille ehdotuksille, sillä niillä on vaikeuksia selviytyä näin suuresta datamäärästä. On ehdotettava uusia algoritmeja, joilla voidaan käsitellä Big Data -ongelmissa esiintyvää kohinaa ja tuottaa korkealaatuista ja puhdasta dataa, jota kutsutaan myös älykkääksi dataksi. Tässä asiakirjassa ehdotetaan kahta Big Data -esikäsittelyä, joilla poistetaan kohinaa sisältäviä esimerkkejä: homogeeninen ensemble- ja heterogeeninen ensemble-suodatin, ja kiinnitetään erityistä huomiota niiden skaalautuvuuteen ja suorituskykyyn. Saadut tulokset osoittavat, että näiden ehdotusten avulla voidaan tehokkaasti saada Smart Dataset mistä tahansa Big Data -luokitusongelmasta.

**Tulos**

Älykkään tiedon käyttöönotto: Melun suodatus Big Data -luokittelussa

**Esimerkki 2.737**

Ihmisen visuaalinen järjestelmä pystyy havaitsemaan epänormaalin kuvan ja päättelemään, mikä tekee siitä oudon. Tähän tehtävään ei ole kiinnitetty riittävästi huomiota tietokonenäössä. Tässä artikkelissa tutkimme kuvien erityyppisiä epätyypillisyyksiä kattavammin kuin aiemmin on tehty. Ehdotamme uutta tietokokonaisuutta epänormaaleista kuvista, joissa on laaja valikoima epätyypillisyyksiä. Suunnittelemme ihmiskokeita löytääksemme karkean taksonomian poikkeavuuden syistä. Kokeissamme paljastuu kolme pääluokkaa poikkeavuutta: objektikeskeisyys, kohtauskeskeisyys ja kontekstisidonnaisuus. Tämän taksonomian perusteella ehdotamme kattavaa laskennallista mallia, jolla voidaan ennustaa kaikki kuvien erityyppiset poikkeavuudet ja joka päihittää aiemmat taidot poikkeavuuksien tunnistamisessa.

**Tulos**

Kohti kuvien poikkeavuuksien taksonomiaa ja laskennallisia malleja

**Esimerkki 2.738**

Tässä asiakirjassa esitellään lähestymistapa, jonka avulla voidaan tunnistaa tehokkaita tekniikoita, joita käytetään hakukoneoptimoinnissa (SEO). SEO-tekijöiden ymmärtäminen, jotka voivat vaikuttaa sivun sijoittumiseen hakukoneessa, on merkittävää webmasterille, jotka haluavat houkutella suuren määrän käyttäjiä verkkosivustolleen. Aiemmista asiaa koskevista tutkimuksista poiketen tässä tutkimuksessa kehitimme älykkään Meta-hakukoneen, joka kokoaa yhteen eri hakukoneiden tulokset ja asettaa ne paremmuusjärjestykseen useiden tärkeiden SEO-parametrien perusteella. Tutkimuksessa pyritään osoittamaan, että useampien SEO-parametrien käyttäminen luokittelualgoritmeissa auttaa saamaan parempia hakutuloksia, mikä lisää käyttäjien tyytyväisyyttä. Meta-hakukoneen tuottamat alustavat tulokset olivat nykyisiä hakukoneita parempia, sillä hakutulokset olivat parempia ja tarkempia.

**Tulos**

Innovatiivinen lähestymistapa online Meta Hakukoneoptimointi

**Esimerkki 2.739**

Monitehtäväinen oppiminen voi olla tehokasta, kun yhdessä tehtävässä hyödylliset piirteet ovat hyödyllisiä myös muissa tehtävissä, ja ryhmälasso on vakiomenetelmä yhteisen piirteiden osajoukon valitsemiseksi. Tässä artikkelissa olemme kiinnostuneita monitehtäväoppimisen vähemmän rajoittavasta muodosta, jossa (1) käytettävissä olevat piirteet voidaan järjestää osajoukkoihin samankaltaisuuden käsitteen mukaisesti ja (2) yhdessä tehtävässä hyödylliset piirteet ovat samankaltaisia, mutta eivät välttämättä identtisiä, muihin tehtäviin parhaiten soveltuvien piirteiden kanssa. Tämän artikkelin tärkein panos on uusi menettely nimeltä Sparse Overlapping Sets (SOS) lasso, joka on kovera optimointi, joka valitsee automaattisesti samankaltaiset piirteet toisiinsa liittyviin oppimistehtäviin. SOSlassolle johdetaan virherajat, ja sen johdonmukaisuus osoitetaan neliövirhehäviölle. SOSlasson motivaationa ovat erityisesti usean kohteen fMRI-tutkimukset, joissa toiminnallinen aktiivisuus luokitellaan käyttämällä aivojen vokseleita ominaisuuksina. Kokeet todellisilla ja synteettisillä aineistoilla osoittavat SOSlasson edut verrattuna lassoon ja ryhmälassoon.

**Tulos**

Sparse Overlapping Sets Lasso for Multitask Learning ja sen soveltaminen fMRI-analyysiin

**Esimerkki 2.740**

Toistuvien kuvioiden tai motiivien löytäminen aikasarjoista on tärkeä valvomaton tehtävä, jossa on vielä useita avoimia kysymyksiä, jotka alkavat motiivin määritelmästä. Tässä asiakirjassa tarkistamme motiivin tuen käsitettä ja kuvaamme sitä motiivin määrittelevien kuvioiden tai toistojen määränä. Sen jälkeen ehdotamme GENMOTIFia, geneettistä algoritmia, jolla löydetään motiiveja, joilla on tuki ja joka on samalla riittävän joustava, jotta se voi ottaa huomioon myös muita motiivien määrittelyjä ja tehtävän ominaisuuksia. GENMOTIF on anytime-algoritmi, joka mukautuu helposti moniin tilanteisiin: haku eri segmenttien pituuksien välillä, yhdenmukaisen skaalauksen soveltaminen, useiden ulottuvuuksien käsittely, erilaisten samankaltaisuus- ja ryhmittelykriteerien käyttö jne. GENMOTIF on myös parametriystävällinen: sillä on vain kaksi intuitiivista parametria, jotka eivät vaikuta merkittävästi sen suorituskykyyn, jos ne asetetaan järkeviin rajoihin. Osoitamme lähestymistapamme arvon useissa synteettisissä ja reaalimaailman asetuksissa, joissa tarkastelemme liikennemäärämittauksia, kiihtyvyysanturisignaaleja ja puhelutallenteita.

**Tulos**

Geneettinen algoritmi joustavien motiivien löytämiseksi tuen avulla

**Esimerkki 2.741**

Ehdotamme uutta nopeaa sanojen upottamistekniikkaa, jossa käytetään hash-funktioita. Menetelmä on uudenlaisen satunnaisprojektiotyypin derandomointi: Kun ei oteta huomioon satunnaisprojektioiden suunnittelussa käytettyä klassista rajoitusta (eli pareittaisten etäisyyksien säilyttämistä tietyssä normitetussa avaruudessa), ratkaisumme hyödyntää erittäin harvoja ei-negatiivisia satunnaisprojektioita. Kokeemme osoittavat, että ehdotetulla menetelmällä voidaan saavuttaa kilpailukykyisiä tuloksia, jotka ovat verrattavissa neuraalisiin sulauttamisoppimismenetelmiin, mutta vain murto-osalla näiden menetelmien laskennallisesta monimutkaisuudesta. Vaikka ehdotettu derandomisointi lisää menetelmämme laskennallista ja avaruudellista monimutkaisuutta, mahdollisuus soveltaa painotusmenetelmiä, kuten positiivista pistemäistä keskinäistä informaatiota (PPMI), malleihimme niiden rakentamisen jälkeen (ja pienemmällä ulottuvuudella) antaa tuloksena oleville upotuksille suuren erottelukyvyn. On selvää, että tähän menetelmään liittyy muita satunnaisprojektiopohjaisten tekniikoiden tunnettuja etuja, kuten päivittämisen helppous.

**Tulos**

Sanavektoreiden hahmottaminen häivytyksen avulla

**Esimerkki 2.742**

Tekoälyn pitkäaikainen unelma on ollut rikastuttaa tietokoneohjelmat maalaisjärjellä saatavalla tiedolla, jotta koneet voisivat järkeillä maailmastamme. Tässä artikkelissa esitetään uusi käytännön näkemys, jonka avulla voidaan automatisoida maalaisjärjellä tapahtuvaa päättelyä ensimmäisen kertaluvun logiikan (FOL) ontologioiden avulla. Ehdotamme uutta mustan laatikon testausmenetelmää FOL SUMO -pohjaisille ontologioille hyödyntämällä WordNetiä ja sen kartoitusta SUMOon. Ehdotuksemme sisältää menetelmän, jolla luodaan (puoli)automaattisesti hyvin suuri joukko testejä, ja menettelyn, jolla testit arvioidaan automaattisesti käyttämällä automaattisia teoreemantarkistajia (Automatic Theorem Provers, ATP). Soveltamalla testausehdotustamme pystymme onnistuneesti arvioimaan a) useiden SUMO:n käännösten pätevyyttä FOL:iin ja b) erilaisten automatisoitujen ATP:iden suorituskykyä. Lisäksi pystymme myös arvioimaan tuloksena saatuja testejä eri laatukriteerien mukaan.

**Tulos**

Ensimmäisen kertaluvun loogisten ontologioiden testaus WordNetiä käyttäen

**Esimerkki 2.743**

Sumeiden ohjainten tiedetään toimivan tehokkaina ja tulkittavissa olevina järjestelmän ohjaimina jatkuvissa tila- ja toiminta-alueissa. Tähän mennessä nämä säätimet on rakennettu käsin tai koulutettu automaattisesti joko asiantuntijoiden tuottamien ongelmakohtaisten kustannusfunktioiden perusteella tai sisällyttämällä niihin yksityiskohtaista tietoa optimaalisesta ohjausstrategiasta. Useimmissa todellisissa vahvistusoppimisen (RL) ongelmissa ei ole molempia vaatimuksia automaattiselle koulutusprosessille. Esittelemme uuden hiukkasparven vahvistusoppimismenetelmän (Particle Swarm reinforcement learning, PSRL), joka pystyy rakentamaan sumeat RL-käytännöt pelkästään harjoittelemalla parametreja maailmanmalleihin, jotka on tuotettu satunnaisesti generoiduista näytteistä todellisesta järjestelmästä. Tämä lähestymistapa yhdistää ensimmäistä kertaa itseorganisoituvat sumeat ohjaimet mallipohjaiseen RL:ään. PSRL:ää voidaan käyttää suoraan mihin tahansa RL-ongelmaan, mikä osoitetaan kolmella tavallisella RL-vertailukohteella, vuoristoautolla, kärryn tangon tasapainottamisella ja kärryn tangon heilauttamisella. Kokeissamme saatiin aikaan suorituskykyisiä ja hyvin tulkittavissa olevia sumeaohjauskäytäntöjä.

**Tulos**

Hiukkasparvioptimointi sumeiden vahvistusoppimisperiaatteiden luomiseksi

**Esimerkki 2.744**

Graafimallit ovat tärkeitä monilla aloilla, kuten hajautetussa tietojenkäsittelyssä, älykkäissä opetusjärjestelmissä tai sosiaalisten verkostojen analysoinnissa. Monissa tapauksissa tällaisissa malleissa on otettava huomioon muutokset graafin rakenteessa eli vaihteleva solmujen tai reunojen määrä. Tällaisten muutosten ennustamisen graafien sisällä voidaan odottaa tuottavan tärkeää tietoa taustalla olevasta dynamiikasta, esimerkiksi käyttäjien käyttäytymisestä. Aikaisemmissa ennustetekniikoissa on kuitenkin keskitytty lähes yksinomaan yksittäisiin reunoihin tai solmuihin. Tässä artikkelissa yritämme ennustaa graafin tulevaa tilaa kokonaisuutena. Ehdotamme, että aikasarjojen ennustaminen muotoillaan regressio-ongelmaksi ja että sovellamme dissimilariteetti- tai kernel-pohjaisia regressiotekniikoita, kuten 1-Nearest Neighbour-, kernel-regressio- ja Gaussin prosessiregressiotekniikoita, joita voidaan soveltaa graafien kerneleiden avulla. Regression tuloksena saadaan pseudoeuklidiseen avaruuteen upotettu piste, jota voidaan analysoida käyttämällä myöhempiä dissimilariteetti- tai kernel-pohjaisia käsittelymenetelmiä. Keskustelemme strategioista, joilla Gaussin prosessien regressiota voidaan nopeuttaa kuutiosta lineaariseen aikaan, ja arvioimme lähestymistapaamme kahdella vakiintuneella teoreettisella graafien evoluutiomallilla sekä kahdella todellisella tietokokonaisuudella älykkäiden opetusjärjestelmien alalta. Huomaamme, että yksinkertaiset regressiomenetelmät, kuten kernel-regressio, riittävät kuvaamaan teoreettisten mallien dynamiikkaa, mutta Gaussin prosessiregressio parantaa ennustevirhettä merkittävästi reaalimaailman datassa.

**Tulos**

Aikasarjojen ennustaminen graafeille ydin- ja eroavaisuusavaruuksissa∗††

**Esimerkki 2.745**

Kuvaamme neuroverkkomallin, joka oppii yhdessä tekstien ja tietopohjan (KB) entiteettien hajautettuja representaatioita. Kun otetaan huomioon tietovarastossa oleva teksti, koulutamme ehdotetun mallin ennustamaan tekstin kannalta merkityksellisiä entiteettejä. Mallimme on suunniteltu geneeriseksi, ja sen avulla voidaan helposti ratkaista erilaisia NLP-tehtäviä. Koulutamme mallin käyttämällä suurta tekstikorpusta ja sen Wikipediasta poimittuja oliomerkintöjä. Arvioimme mallia kolmessa tärkeässä NLP-tehtävässä (lauseiden tekstin samankaltaisuus, entiteettien linkittäminen ja faktatietokysymyksiin vastaaminen), joissa käytetään sekä valvomattomia että valvottuja asetuksia. Saavutimme huipputuloksia kaikissa kolmessa näistä tehtävistä.

**Tulos**

Tekstien ja entiteettien hajautettujen esitysten oppiminen tietopohjasta

**Esimerkki 2.746**

Epätäydellisten agenttien tarkoituksenmukaisen käyttäytymisen mallintaminen pienestä määrästä havaintoja on haastava tehtävä. Kun rajoitutaan yhden toimijan päätöksentekoteoreettiseen asetelmaan, käänteisen optimaalisen ohjauksen tekniikoissa oletetaan, että havaittu käyttäytyminen on likimain optimaalinen ratkaisu tuntemattomaan päätösongelmaan. Näillä tekniikoilla opitaan hyötyfunktio, joka selittää esimerkkikäyttäytymisen, ja sitä voidaan sitten käyttää ennustamaan tai jäljittelemään tarkasti tulevaa käyttäytymistä samanlaisissa havaituissa tai havaitsemattomissa tilanteissa. Tässä työssä tarkastelemme samankaltaisia tehtäviä kilpailevilla ja yhteistyöhön perustuvilla moniagenttialueilla. Toisin kuin yksiagenttiympäristöissä, pelaaja ei voi tässä tapauksessa maksimoida palkkiotaan lyhytnäköisesti, vaan sen on spekuloitava, miten muut agentit voivat vaikuttaa pelin lopputulokseen. 1 ar X iv :1 30 8. 35 06 v1 [ cs .G T ] 1 5 A ug 2 01 3 peliteoreettisen katumuksen käsitteen ja maksimaalisen entropian periaatteen avulla esitämme tekniikan käyttäytymisen ennustamiseen ja yleistämiseen.

**Tulos**

Laskennallinen järkeistäminen: Käänteinen tasapaino-ongelma

**Esimerkki 2.747**

Monimerkkisten tuotosten ennustamisen oppiminen on haastavaa, mutta monissa ongelmissa tuotoksille on olemassa luonnollinen metriikka, jota voidaan käyttää ennusteiden parantamiseen. Tässä artikkelissa kehitämme häviöfunktion usean merkin oppimista varten, joka perustuu Wassersteinin etäisyyteen. Wassersteinin etäisyys tarjoaa luonnollisen käsitteen todennäköisyysmittojen erilaisuudelle. Vaikka tarkan Wassersteinin etäisyyden optimointi on kallista, viimeaikaisessa työssä on kuvattu säännelty approksimaatio, joka voidaan laskea tehokkaasti. Kuvaamme tehokkaita oppimisalgoritmeja, jotka perustuvat tähän regularisointiin ja laajentavat Wassersteinin häviön todennäköisyysmitoista normalisoimattomiin mittoihin. Kuvaamme myös tilastollisen oppimisrajan häviölle ja osoitamme yhteydet kokonaisvaihtelunormiin ja Jaccard-indeksiin. Wassersteinin häviö voi kannustaa ennusteiden tasaisuuteen suhteessa lähtöavaruuden valittuun metriikkaan. Osoitamme tämän ominaisuuden todellisen datan tunnisteiden ennustamisongelmassa käyttäen Yahoo Flickr Creative Commons -tietokokonaisuutta, ja saavutamme paremman suorituskyvyn kuin perustasolla, joka ei käytä metriikkaa.

**Tulos**

Oppiminen Wasserstein-tappion kanssa

**Esimerkki 2.748**

Yhteisöpohjaiset kysymysten vastausalustat ovat houkutelleet runsaasti käyttäjiä jakamaan tietoa ja oppimaan toisiltaan. CQA-alustojen nopean laajentumisen myötä syntyy paljon päällekkäisiä kysymyksiä, mikä saa käyttäjät hämmentymään oikean viitteen valinnassa. Meidän on kiireesti otettava käyttöön tehokkaita automaattisia algoritmeja, joilla voidaan käyttää uudelleen historiallisia kysymyksiä ja vastauksia niihin. Tässä asiakirjassa keskitymme ongelman ratkaisemiseen kysymysten haun avulla, jonka tavoitteena on löytää historiallisia kysymyksiä, jotka ovat merkityksellisiä tai semanttisesti vastaavia, jotta kysely voidaan ratkaista suoraan. Tämän tehtävän haasteita ovat kysymysten väliset leksikaaliset aukot sanojen monitulkintaisuuden ja sanojen yhteensopimattomuuden ongelmaan. Lisäksi kysyttyjen lauseiden rajallinen sanamäärä aiheuttaa sanaominaisuuksien vähäisyyttä. Näiden haasteiden lieventämiseksi ehdotamme uutta HNIL-nimistä kehystä, joka koodaa kysymyksen sisällön lisäksi myös kysyjän sosiaaliset vuorovaikutussuhteet parantaakseen kysymyksen upottamisen suorituskykyä. Tarkemmin sanottuna sovellamme satunnaiskävelyyn perustuvaa oppimismenetelmää ja rekursiivista neuroverkkoa sovittamaan yhteen kysyjän kysymyksen ja muiden käyttäjien esittämien historiallisten kysymysten samankaltaisuudet. Laajat kokeet laajamittaisella tietokokonaisuudella, joka on peräisin reaalimaailman CQA-sivustolta Quora, osoittavat, että heterogeenisten sosiaalisten verkostotietojen käyttö on tässä tehtävässä muita huipputason ratkaisuja parempi.

**Tulos**

Kysymysten haku yhteisöpohjaista kysymysten vastaamista varten heterogeenisen verkoston integrointioppimisen avulla

**Esimerkki 2.749**

Tässä asiakirjassa tutkitaan valintamenetelmien vaikutusta hajautettujen evoluutioalgoritmien on-line-ohjauksessa. Ehdotamme mEDEA-algoritmin muunnelmaa, johon lisäämme valintaoperaattorin, ja sovellamme sitä tehtävälähtöisessä skenaariossa. Arvioimme neljää valintamenetelmää, jotka aiheuttavat erilaista valintapaineen voimakkuutta monirobottinavigointitehtävässä, jossa on esteiden välttämistehtävä, ja kollektiivisessa metsästystehtävässä. Kokeet osoittavat, että pieni valintapaineen intensiteetti riittää nopeasti hyviin suorituksiin kyseisissä tehtävissä. Otamme käyttöön erilaisia mittareita valintamenetelmien vertailemiseksi ja osoitamme, että mitä suurempi valintapaine on, sitä parempia suorituksia saavutetaan, erityisesti haastavammassa ruoanhakutehtävässä.

**Tulos**

Valintamenetelmien vertailu on-line hajautetussa evolutiivisessa robotiikassa

**Esimerkki 2.750**

Termi "affordanssi" tarkoittaa esineiden käyttäytymiseen liittyvää merkitystä. Ehdotamme kognitiivista arkkitehtuuria affordanssien havaitsemiseksi visuaalisessa modaliteetissa. Tämä malli perustuu liikesekvenssien sisäiseen simulointiin. Jokaisen liikkeen vaiheen tuloksena syntyvä aistitila ennustetaan eteenpäin suuntautuvalla mallilla, joka puolestaan käynnistää uuden (simuloidun) motorisen käskyn tuottamisen käänteisellä mallilla. Näin syntyy sarja mentaalisia mielikuvia aistien ja moottorin alueella. Alkaen todellisesta aistitilasta simuloidaan rinnakkain suuri määrä tällaisia sarjoja. Lopullinen affordanssin havaitseminen perustuu tuotettuihin motorisiin komentoihin. Sovellamme tätä mallia reaalimaailman liikkuvaan robottiin, joka joutuu kohtaamaan estejärjestelyjä, joista osa on ohitettavissa (käytävä) ja osa ei (umpikuja). Robotin tehtävänä on havaita oikea este ("läpipääsykelpoinen" tai "ei-läpipääsykelpoinen"). Tarvittavat sisäiset mallit hankitaan hierarkkisessa koulutusprosessissa. Tämän jälkeen robottiagentti pystyy erottamaan käytävät ja umpikujat luotettavasti toisistaan. Tämä reaalimaailman tulos vahvistaa ehdotetun mentaalisen simulointimenetelmän pätevyyttä. Lisäksi vertailemme useita simulointiprosessin avaintekijöitä suorituskyvyn ja tehokkuuden osalta. Rahoitusselvitys: Tämä tutkimus ei ole saanut erityistä avustusta miltään julkiselta, kaupalliselta tai voittoa tavoittelemattomalta rahoittajalta. 1 ar X iv :1 61 1. 00 27 4v 1 [ cs .A I] 1 N ov 2 01 6

**Tulos**

Affordanssien havaitseminen visuomotorisen simulaation avulla

**Esimerkki 2.751**

Monivaiheinen temporaalisen eron (TD) oppiminen, jossa päivityskohteet sisältävät tietoa useista aika-askeleista eteenpäin, on yksi suosituimmista TD-oppimisen muodoista lineaaristen funktioiden approksimoinnissa. Syynä tähän on se, että monivaiheiset menetelmät tuottavat usein huomattavasti parempaa suorituskykyä kuin niiden yksiportaiset vastineet, koska päivityskohteiden harha on pienempi. Epälineaaristen funktioiden approksimoinnissa yksivaiheiset menetelmät näyttävät kuitenkin olevan normi. Osasyynä voi olla se, että monilla aloilla suositut monivaiheiset menetelmät TD(λ) ja Sarsa(λ) eivät toimi hyvin, kun ne yhdistetään epälineaariseen funktion approksimointiin. Erityisesti ne ovat hyvin alttiita arvoestimaattien poikkeavuudelle. Tässä artikkelissa selvitämme syyn tähän. Lisäksi ehdotamme analyysimme perusteella uutta monivaiheista TD-menetelmää epälineaaristen funktioiden approksimaatiota varten, jolla tämä ongelma voidaan ratkaista. Vahvistamme menetelmämme tehokkuuden käyttämällä kahta vertailutehtävää, joissa funktioiden approksimointina käytetään neuroverkkoja.

**Tulos**

Tehokas monivaiheinen temporaalisen eron oppiminen epälineaarisen funktion approksimointiin

**Esimerkki 2.752**

Kun todennäköisyyspohjaiset järjestelmät yleistyvät ja tulevat laajempaan käyttöön, järjestelmän havaintoja ja suosituksia selittävän mekanismin tarve on entistä tärkeämpi. Järjestelmä tarvitsee myös mekanismin kilpailevien selitysten järjestämistä varten. Tarkastelemme kahta kirjallisuudessa esiintyvää edustavaa selitysmallia, joista toinen on Gardenforsin ja toinen Pearlin ansiota, ja osoitamme, että molemmissa on merkittäviä ongelmia. Ehdotamme lähestymistapaa "paremman selityksen" käsitteen määrittelyyn, jossa yhdistyvät jotkin molempien piirteet sekä Pearlin ja muiden kausaalisuutta käsittelevä uudempi työ.

**Tulos**

Selityksen määrittäminen todennäköisyysjärjestelmissä

**Esimerkki 2.753**

Käsittelemme uutta ongelmaa, joka liittyy tietokilpailutyyppisten tietokysymysten automaattiseen tuottamiseen DBpedian kaltaisesta tietograafista. Tällaisilla kysymyksillä on runsaasti sovelluksia esimerkiksi käyttäjien kouluttamiseen tai heidän tietämyksensä arviointiin tietyllä alalla. Ongelman ratkaisemiseksi ehdotamme kokonaisvaltaista lähestymistapaa. Lähestymistavassa valitaan ensin nimetty entiteetti tietograafista vastaukseksi. Sen jälkeen se luo strukturoidun kolmoiskuviokyselyn, jonka ainoana tuloksena on vastaus. Jos halutaan monivalintakysymys, lähestymistapa valitsee vaihtoehtoisia vastausvaihtoehtoja. Lopuksi lähestymistapamme käyttää mallipohjaista menetelmää strukturoidun kyselyn sanallistamiseksi ja luonnollisen kielen kysymyksen tuottamiseksi. Keskeinen haaste on arvioida, kuinka vaikea tuotettu kysymys on ihmiskäyttäjille. Tätä varten hyödynnämme Jeopardy! -tietokilpailusta saatuja historiallisia tietoja ja semanttisesti kommentoitua web-asiakirjakokoelmaa, kehitämme sopivia piirteitä ja koulutamme logistisen regressioluokittimen kysymyksen vaikeuden ennustamiseksi. Kokeet osoittavat yleisen lähestymistapamme toimivuuden.

**Tulos**

Tietämyskysymykset tietämyskaavioista

**Esimerkki 2.754**

Ratkaiseva osa tietopohjajärjestelmää, joka poimii uusia faktoja tekstikorpuksista, on harjoitusdatan tuottaminen relaatioiden poimijoita varten. Tässä artikkelissa esitellään menetelmä, jolla maksimoidaan juuri koulutettujen relaatio-uutinten tehokkuus mahdollisimman pienin annotaatiokustannuksin. Manuaalista merkitsemistä voidaan vähentää merkittävästi etäkohtien valvonnalla, joka on menetelmä, jolla harjoitusaineisto rakennetaan automaattisesti sovittamalla suuri tekstikorpus olemassa olevaan tietopohjaan, jossa on tunnettuja faktoja. Esimerkiksi kaikki lauseet, joissa mainitaan sekä "Barack Obama" että "Yhdysvallat", voivat toimia positiivisina harjoitustapauksina relaatiolle born in(subject,object). Etävalvonta johtaa kuitenkin tyypillisesti erittäin häiriintyneeseen harjoitusjoukkoon: monet harjoituslauseet eivät todellisuudessa ilmaise aiottua relaatiota. Ehdotamme, että etävalvonta yhdistetään minimaaliseen manuaaliseen valvontaan tekniikalla, jota kutsutaan ominaisuuksien merkitsemiseksi, jotta voidaan poistaa kohinaa suuresta ja meluisesta alkuperäisestä harjoitusjoukosta, jolloin tarkkuus paranee merkittävästi. Parannamme tätä lähestymistapaa edelleen ottamalla käyttöön Semantic Label Propagation -menetelmän, jossa käytetään ehdokkaiden harjoitustapausten matalaulotteisten esitysten samankaltaisuutta koulutusjoukon laajentamiseksi, jotta voidaan lisätä muistamista ja säilyttää samalla korkea tarkkuus. Ehdotettua strategiaa harjoitusaineiston tuottamiseksi tutkitaan ja arvioidaan vakiintuneella testikokoelmalla, joka on suunniteltu tietopohjien muodostamistehtäviä varten. Kokeelliset tulokset osoittavat, että Semantic Label Propagation -strategia johtaa huomattavaan suorituskyvyn paranemiseen verrattuna nykyisiin lähestymistapoihin ja vaatii samalla lähes mitättömän vähän manuaalista merkintätyötä.

**Tulos**

Tietopohjan kartuttaminen semanttisen etiketin levittämisen avulla

**Esimerkki 2.755**

On ehdotettu, että konvolutiivisia neuroverkkoja (Convolutional Neural Networks, CNN), joissa on konvoluutio- ja pooling-operaatioita taajuusakselilla, käytettäisiin ominaisuuksien taajuussiirtymien muuttumattomuuden saavuttamiseksi. Tämä ei kuitenkaan sovellu siihen, että akustiset piirteet vaihtelevat taajuuden mukaan. Tässä artikkelissa väitämme, että aika-akselin suuntainen konvoluutio on tehokkaampi. Ehdotamme myös, että syviin CNN:iin lisätään IMP-kerros (intermap pooling). Tässä kerroksessa kunkin ryhmän suodattimet poimivat yhteisiä mutta spektrisesti erilaisia piirteitä, minkä jälkeen kerros yhdistää kunkin ryhmän piirrekartat. Tämän seurauksena ehdotettu IMP CNN voi saavuttaa epäherkkyyden eri puhujille ja lausumille ominaisten spektristen vaihteluiden suhteen. IMP CNN -arkkitehtuurin tehokkuus osoitetaan useissa LVCSR-tehtävissä. Jopa ilman puhujan mukauttamistekniikoita arkkitehtuuri saavutti 12,7 prosentin WER-arvon Hub5'2000-arviointitestisarjan SWB-osiossa, mikä on kilpailukykyinen muiden huipputason menetelmien kanssa.

**Tulos**

Syvät CNN:t aika-akselilla ja Intermap Pooling spektrivaihteluiden kestävyyden varmistamiseksi.

**Esimerkki 2.756**

Tässä artikkelissa raportoidaan kahden riippuvuuksien jäsentäjän, Stanfordin ja Miniparin, suorituskyvystä biolääketieteellisissä teksteissä. Näiden jäsentäjien suorituskyky riippuvuuksien osoittamisessa kahden biolääketieteellisen käsitteen välille, joiden on jo osoitettu olevan yhteydessä toisiinsa, ei ole tyydyttävä. Sekä Stanford että Minipar, jotka ovat tilastollisia jäsentäjiä, eivät pysty määrittämään riippuvuussuhdetta kahden toisiinsa kytkeytyneen käsitteen välille, jos ne ovat vähintään yhden lausekkeen päässä toisistaan. Miniparin suorituskykyä biomedisiinisen tekstin jäsentämisessä mitataan myös tarkkuuden, palautuksen ja liitospistemäärän F-tuloksen (esim. oikein tunnistettu pää riippuvuussuhteessa) avulla käyttäen Stanfordia kultaisena standardina. Tulokset viittaavat siihen, että Minipar ei vielä sovellu biolääketieteellisten tekstien jäsentämiseen. Lisäksi laadullinen tutkimus paljastaa, että Miniparin heikentyneeseen suorituskykyyn vaikuttavat olennaisesti myös jäsentäjien toimintaperiaatteiden erot.

**Tulos**

Stanfordin ja Miniparin jäsentimen suorituskyky biolääketieteellisissä teksteissä

**Esimerkki 2.757**

Luokkien epätasapaino on yksi koneoppimisen haastavista ongelmista monissa reaalimaailman sovelluksissa, kuten hiili- ja kaasupurkausten onnettomuuksien seurannassa: purkausten ennakointitiedot ovat kuitenkin äärimmäisen pienempiä kuin normaalit tiedot, joihin me todella keskitymme. Kustannusherkkä mukautusmenetelmä on tyypillinen algoritmitason menetelmä, jolla vastustetaan datajoukon epätasapainoa. SVM-luokittimelle, jota on muutettu siten, että se sisältää vaihtelevan rangaistusparametrin (C) kullekin tarkastellulle esimerkkiryhmälle. C-arvo määritetään kuitenkin empiirisesti tai lasketaan arviointimittarin mukaan, joka on laskettava iteratiivisesti ja aikaa vievästi. Tässä asiakirjassa esitellään uusi kustannustietoinen SVM-menetelmä, jonka rangaistusparametri C optimoidaan klusterin todennäköisyystiheysfunktion (PDF) perusteella, ja klusterin PDF-arvo arvioidaan vain samankaltaisuusmatriisin ja joidenkin ennalta määritettyjen hyperparametrien perusteella. Kokeelliset tulokset erilaisilla vakiomuotoisilla vertailutietoaineistoilla ja reaalimaailman aineistoilla, joissa on erilaisia epätasapainon suhteita, osoittavat, että ehdotettu menetelmä on tehokas verrattuna yleisesti käytettyihin kustannusherkkiin tekniikoihin.

**Tulos**

Kustannustietoisen SVM:n optimointi epätasapainoiselle aineistolle: Klusterin yhdistäminen luokitteluun.

**Esimerkki 2.758**

Variationaalisissa autokoodereissa (VAE) käytetään usein Gaussin tai kategorian jakaumaa päättelyprosessin mallintamiseen. Tämä rajoittaa variationaalista oppimista, koska tämä yksinkertaistettu oletus ei vastaa todellista jälkijakaumaa, joka on yleensä paljon kehittyneempi. Tämän rajoituksen purkamiseksi ja mielivaltaisen parametrisen jakauman soveltamiseksi päättelyn aikana tässä artikkelissa johdetaan puolijatkuva latentti esitys, joka approksimoi jatkuvaa tiheyttä määrättyyn tarkkuuteen asti ja jota on paljon helpompi analysoida kuin sen jatkuvaa vastinetta, koska se on pohjimmiltaan diskreetti. Esittelemme ehdotuksen soveltamalla jälkijakaumana polynomi-eksponenttiperheen jakaumia, jotka ovat universaaleja todennäköisyystiheysfunktion generaattoreita. Kokeelliset tuloksemme osoittavat johdonmukaisia parannuksia verrattuna yleisesti käytettyihin VAE-malleihin.

**Tulos**

Karkearakeiset eksponentiaaliset muunnelma-autokooderit

**Esimerkki 2.759**

Lehtisuonisto muodostaa perustan lehtien luonnehdinnalle ja luokittelulle. Eri lajeilla on erilaisia lehtisuonten kuvioita. Lehtisuonten segmentointi auttaa pitämään kirjaa kaikista lehdistä niiden tietyn suonikuvion mukaan ja tarjoaa siten tehokkaan tavan hakea ja tallentaa eri kasvilajeja koskevia tietoja tietokantaan sekä tehokkaan keinon luonnehtia kasveja lehtisuonirakenteen perusteella, joka on ainutlaatuinen jokaiselle lajille. Algoritmissa ehdotetaan uutta tapaa lehtien suonien segmentointiin Odd Gabor -suotimien ja morfologisten operaatioiden avulla paremman tuloksen aikaansaamiseksi. Odd Gabor -suodatin antaa tehokkaan tuloksen ja on vankka ja skaalautuva verrattuna nykyisiin tekniikoihin, koska se havaitsee lehdissä olevat hienot kuitumaiset suonet paljon tehokkaammin.

**Tulos**

Lehtisuonten segmentointi Odd Gabor -suodattimien ja morfologisten operaatioiden avulla

**Esimerkki 2.760**

OA-mallit (Or's of And's) koostuvat pienestä määrästä konjunktioiden disjunktioita, joita kutsutaan myös disjunktiiviseksi normaalimuodoksi. Esimerkki OA-mallista on seuraava: Jos (x1 = 'sininen' JA x2 = 'keskimmäinen') TAI (x1 = 'keltainen'), niin ennuste Y = 1, muuten ennuste Y = 0. Ja-mallien etuna on, että ne ovat tulkittavissa ihmisasiantuntijoille, koska ne ovat joukko ehtoja, jotka kuvaavat tiiviisti tietyn osajoukon ominaisuuksia. Esittelemme kaksi optimointiin perustuvaa koneoppimiskehystä OA-mallien rakentamiseen, Optimoitu OA (OOA) ja sen nopeampi versio, Optimoitu OA approksimaatioiden kanssa (OOAx). Todistamme teoreettisia rajoja OA-mallin mallien ominaisuuksille. Rakennamme OA-malleja obstruktiivisen uniapnean diagnostiseksi seulontatyökaluksi, jolla saavutetaan korkea tarkkuus ja huomattava parannus tulkittavuuteen muihin menetelmiin verrattuna.

**Tulos**

Optimoitujen tai- ja-parien oppiminen

**Esimerkki 2.761**

Tässä artikkelissa esitellään uusi deterministinen approksimointitekniikka Bayes-verkoissa. Menetelmä, "Expectation Propagation", yhdistää kaksi aiempaa tekniikkaa: oletetun tiheyden suodatuksen, joka on Kalmanin suodattimen laajennus, ja loopy be lief propagationin, joka on uskomuspropagaation laajennus Bayesin verkoissa. Loopy belief propa gation, koska se levittää täsmällisiä uskomustiloja, on hyödyllinen rajoitetulle uskomusverkkojen luokalle, kuten niille, jotka ovat puhtaasti diskreettejä. Expec tation Propagation approksimoi uskomustiloja säilyttämällä vain odotukset, kuten keskiarvon ja varianssin, ja iteroi, kunnes nämä odotukset ovat johdonmukaisia koko verkossa. Tämän vuoksi sitä voidaan soveltaa hybridiverkkoihin, joissa on diskreettejä ja jatkuvia solmuja. Gaussin sekoitusmalleilla tehdyt kokeet osoittavat, että Expectation Propagation on vakuuttavasti parempi kuin menetelmät, joilla on samanlaiset laskentakustannukset: Laplace-menetelmä, vari ational Bayes ja Monte Carlo. Expectation Propagation tarjoaa myös tehokkaan algoritmin Bayes-pistekone-luokittelijoiden kouluttamiseen.

**Tulos**

Odotusten eteneminen likimääräistä Bayesin päättelyä varten

**Esimerkki 2.762**

Sairauden etenemisen ja hoidon seurannan kannalta merkityksellisiä kuvantamismarkkereita kuvaavien mallien saaminen on haastavaa. Mallit perustuvat tyypillisesti suuriin tietomääriin, joissa on kommentoituja esimerkkejä<lb>tunnetuista merkkiaineista, joiden havaitsemista pyritään automatisoimaan. Suuri annotointitarve ja rajoittuminen tunnettujen merkkiaineiden sanastoon rajoittavat tällaisten lähestymistapojen<lb>tehoa. Tässä tutkimuksessa suoritetaan valvomatonta oppimista, jotta<lb>kuvantamisdatan poikkeavuudet voidaan tunnistaa merkkiainekandidaateiksi. Käytämme AnoGANia, syvää konvolutiivista generatiivista vastakkaisverkkoa, joka<lb>oppii normaalin anatomisen vaihtelun moninaisuuden, ja sen mukana on uusi<lb>anomalian pisteytysjärjestelmä, joka perustuu kuva-avaruuden ja la-<lb>tenttiavaruuden väliseen kartoitukseen. Sovellettuna uusiin tietoihin malli merkitsee poikkeavuudet ja pisteyttää<lb>kuvapatsaat, jotka osoittavat niiden sopivuuden opittuun jakaumaan. Verkkokalvon optisen koherenssitomografian kuvista saadut tulokset osoittavat, että<lb>lähestymistapa tunnistaa oikein poikkeavat kuvat, kuten kuvat, joissa on verkkokalvon nestettä tai hyperheijastavia polttopisteitä.

**Tulos**

Valvomaton poikkeamien havaitseminen generatiivisten vastakkaisverkkojen avulla merkkien löytämisen ohjaamiseksi

**Esimerkki 2.763**

Teleoperaattorin suuri investointi kohdistuu infrastruktuuriin ja sen ylläpitoon, kun taas liiketoiminnan tulot ovat suhteessa siihen, kuinka suuri ja hyvä asiakaskunta on. Esittelemme dataan perustuvan analyysistrategian, joka perustuu kombinatoriseen optimointiin ja historiatietojen analysointiin. Tiedot kattavat käyttäjien historiallisen liikkuvuuden yhdellä Ruotsin alueella viikon aikana. Soveltamalla ehdotettua menetelmää tapaustutkimukseen olemme tunnistaneet maantieteellis-demografisten segmenttien optimaalisen osuuden asiakaskannasta, kehittäneet toiminnon suunnitellun markkinointikampanjan potentiaalin arvioimiseksi ja tutkineet ongelmaa, joka koskee markkinointikampanjoiden avulla tavoiteltavien maantieteellis-demografisten segmenttien optimaalista määrää ja tyyppejä. Sumean logiikan avulla data-analyysin päätelmät muunnetaan automaattisesti ymmärrettäviksi suosituksiksi luonnollisella kielellä.

**Tulos**

Suositukset televiestintäalan markkinointikampanjoiksi jalanjälki-analyysin perusteella.

**Esimerkki 2.764**

Tässä työssä esitellään ja analysoidaan kolmea konvoluutiohermoverkkomallia (convolutional neural network, CNN) kuvien tehokasta pikselikohtaista luokittelua varten. Kun käytetään konvoluutiohermoverkkoja luokittelemaan yksittäisiä pikseleitä koko kuvan laikuissa, liukuikkunaverkkoja käytettäessä suoritetaan paljon turhia laskutoimituksia. Nämä uudet arkkitehtuurit ratkaisevat tämän ongelman joko poistamalla tarpeettomat laskutoimitukset tai käyttämällä täysin konvoluutioarkkitehtuureja, jotka luonnostaan ennustavat monia pikseleitä kerralla. Näiden kolmen mallin toteutukset ovat käytettävissä Caffe-kirjaston päällä olevan uuden apuohjelman avulla. Apuohjelma tukee monia erilaisia kuvan syöttö- ja tulostusformaatteja, esikäsittelyparametreja ja menetelmiä, joilla tasataan merkintähistogrammi harjoittelun aikana. Caffe-kirjastoa on laajennettu uusilla kerroksilla ja uudella backendillä, jotta se on käytettävissä laajemmalla laitteistovalikoimalla, kuten suorittimilla ja GPU:lla OpenCL:n avulla. AMD:n GPU:lla nopeutuminen on ollut 54-kertainen (SK-verkko), 437-kertainen (U-verkko) ja 320-kertainen (USK-verkko), kun lähtötasona on käytetty SK-verkkoa vastaavaa SW-verkkoa (liukuva ikkuna). Tarran läpäisykyky on jopa yksi megapikseli sekunnissa. Analysoiduilla neuroverkoilla on erityispiirteitä, joita sovelletaan harjoittelun tai käsittelyn aikana, eivätkä kaikki aineistot sovellu kaikille arkkitehtuureille. Ennusteiden laatua arvioidaan kahdella neuroverkkoaineistolla, joista toinen on ISBI 2012 -haasteaineisto. Koulutuksessa käytettiin kahta eri häviöfunktiota, Malis-häviötä ja Softmax-häviötä. Koko putki, joka koostuu malleista, käyttöliittymästä ja modifioidusta Caffe-kirjastosta, on saatavilla avoimen lähdekoodin ohjelmistona työnimellä Project Greentea.

**Tulos**

Tehokkaat konvoluutio-neuraaliverkot pikseleittäin tapahtuvaan luokitteluun heterogeenisissä laitteistojärjestelmissä

**Esimerkki 2.765**

Vastausjoukko-ohjelmoinnissa on tutkittu erilaisia ekvivalenssin käsitteitä, kuten vahvasta ja yhtenäisestä ekvivalenssista tunnettuja käsitteitä, pääasiassa sellaisten ohjelmien tunnistamiseksi, jotka voivat toimia korvaavina ohjelmina muuttamatta semantiikkaa, esimerkiksi ohjelmien optimoinnissa. Tällaisia semanttisia vertailuja luonnehditaan yleensä erilaisilla mallivalinnoilla Hereand-There (HT) -logiikassa. Yhtenäistä ekvivalenssia varten voidaan kuitenkin saada HT-mallien avulla oikeat luonnehdinnat vain äärellisille teorioille, vastaavasti ohjelmille. Tässä artikkelissa osoitamme, että HT:n vastamallien valinta kuvaa yhtenäistä ekvivalenssia myös äärettömille teorioille. Tämä tulos muunnetaan eri ekvivalenssikäsitteiden johdonmukaisiksi luonnehdinnoiksi vastamalleilla sekä HT-mallien ja vastamallien sekoituksella (ns. ekvivalenssitulkinnoilla). Lisäksi yleistämme ohjelmien niin sanotun relativoidun hyperekvivalenssin käsitteen propositioteorioihin ja sovellamme samaa metodologiaa saadaksemme semanttisen luonnehdinnan, joka soveltuu äärettömiin asetelmiin. Tämä mahdollistaa tulosten siirtämisen ensimmäisen kertaluvun teorioihin hyvin yleisessä semantiikassa, joka annetaan HT:n kvantifioidun version avulla. Näin saamme yleiset puitteet erilaisten ekvivalenssikäsitteiden tutkimiselle teorioiden osalta vastausjoukkojen semantiikassa. Lisäksi todistamme tarkoituksenmukaisen ominaisuuden, joka mahdollistaa laajennettujen allekirjoitusten yksinkertaistetun käsittelyn, ja annamme lisätuloksia ei-maantieteellisille logiikkaohjelmille. Erityisesti yhtenäinen ekvivalenssi on yhteneväinen avoimessa ja tavallisessa vastausjoukkojen semantiikassa, ja äärellisille ei-maantieteellisille ohjelmille näissä semantiikoissa myös yhtenäisen ekvivalenssin tavanomainen luonnehdinta maadoituksen maksimaalisten ja totaalisten HT-mallien suhteen on oikea jopa äärettömille alueille, kun vastaavat maadoitusohjelmat ovat äärettömiä. Ilmestyy julkaisussa Theory and Practice of Logic Programming (TPLP).

**Tulos**

Yleinen kehys vastausjoukkojen ohjelmoinnin ekvivalensseille vastamallien avulla tässä ja tässä ∗ -logiikassa.

**Esimerkki 2.766**

Kieli on sosiaalinen ilmiö, ja sen sosiaaliseen luonteeseen kuuluu, että se muuttuu jatkuvasti. Laskennallisen kielitieteen (CL) piirissä on viime aikoina havaittavissa kiinnostuksen kasvua kielen sosiaalista ulottuvuutta kohtaan. Tässä artikkelissa esitämme katsauksen "laskennallisen sosiolingvistiikan" kehittyvään alaan, joka kuvastaa tätä lisääntynyttä kiinnostusta. Tavoitteena on antaa kattava katsaus sosiolingvistisiä aiheita käsittelevään CL-tutkimukseen, jossa käsitellään muun muassa kielen ja sosiaalisen identiteetin välistä suhdetta, kielenkäyttöä sosiaalisessa vuorovaikutuksessa ja monikielistä viestintää. Lisäksi osoitamme, että tutkimukseen osallistuvien tutkimusyhteisöjen välillä on mahdollista saavuttaa synergiaa osoittamalla, miten CL:ssä laajalti käytetyt laajamittaiset aineistolähtöiset menetelmät voivat täydentää nykyisiä sosiolingvistisiä tutkimuksia ja miten sosiolingvistiikka voi antaa tietoa ja kyseenalaistaa CL-tutkimuksissa käytettyjä menetelmiä ja oletuksia. Toivomme voivamme välittää näiden kahden yhteisön välisen tiiviimmän yhteistyön mahdolliset hyödyt ja lopuksi keskustelemme avoimista haasteista.

**Tulos**

Laskennallinen sosiolingvistiikka: A Survey

**Esimerkki 2.767**

Suurten tekstien tiivistäminen on edelleen avoin ongelma kielenkäsittelyssä. Tässä työssä kehitämme täysimittaisen putken uutisartikkeleiden tiivistelmien tuottamiseksi käyttäen abstraktia merkityskuvausta (AMR). Ensin luodaan tarinoiden AMR-graafit, sitten poimitaan tarinagraafeista tiivistelmägraafit ja lopuksi luodaan lauseet tiivistelmägraafista. Yhteenveto-AMR:ien poimimiseen tarina-AMR:istä käytetään kaksivaiheista prosessia. Ensin tekstistä etsitään tärkeät lauseet ja sitten poimitaan yhteenveto-AMR:t näistä valituista lauseista. Suoritamme aiempia AMR:ää tiivistämiseen käyttäviä menetelmiä yli 3 ROGUE-1-pistettä paremmalla tuloksella. CNN-Dailymail-korpuksessa saavutamme tuloksia, jotka ovat kilpailukykyisiä vahvan lyijy-3-perustason kanssa, kunnes tiivistelmäkuvaajan poimintavaihe on suoritettu.

**Tulos**

Tekstin tiivistäminen abstraktin merkityksen esittämisen avulla

**Esimerkki 2.768**

Käsittelemme ongelmaa, joka liittyy kahden kuvan välisen vastaavuuden määrittämiseen geometrisen mallin, kuten affiinisen tai ohutlevy-spline-muunnoksen, mukaisesti ja sen parametrien estimointiin. Tämän työn panos on kolmitahoinen. Ensinnäkin ehdotamme konvoluutiohermoverkkoarkkitehtuuria geometrista yhteensovittamista varten. Arkkitehtuuri perustuu kolmeen pääkomponenttiin, jotka jäljittelevät tavanomaisia vaiheita, joita ovat piirteiden louhinta, yhteensovittaminen ja samanaikainen virheiden havaitseminen sekä mallin parametrien estimointi, ja jotka ovat koulutettavissa alusta loppuun. Toiseksi osoitamme, että verkkoparametrit voidaan kouluttaa synteettisesti tuotetuista kuvista ilman manuaalista merkintää ja että sovituskerroksemme lisää merkittävästi yleistämiskykyä ennen näkemättömiin kuviin. Lopuksi osoitamme, että sama malli pystyy suorittamaan sekä instanssi- että kategoriatason täsmäytyksen, mikä antaa huippuluokan tuloksia haastavassa Proposal Flow -tietokannassa.

**Tulos**

Geometrisen yhteensovittamisen konvolutiivinen neuroverkkoarkkitehtuuri

**Esimerkki 2.769**

Syväoppimisvälineet ovat viime aikoina saaneet paljon huomiota sovelletussa koneoppimisessa. Tällaiset regressio- ja luokittelutyökalut eivät kuitenkaan mahdollista mallin epävarmuuden huomioon ottamista. Bayesin mallit tarjoavat mahdollisuuden pohtia mallin epävarmuutta, mutta niiden laskentakustannukset ovat yleensä liian korkeat. Osoitamme, että monikerroksisissa perceptronimalleissa (MLP) esiintyvä pudotus voidaan tulkita Bayesin approksimaatioksi. Saadaan tuloksia epävarmuuden mallintamisesta dropout MLP-malleille - poimitaan olemassa olevista malleista tietoa, joka on tähän asti heitetty pois. Tämä lieventää epävarmuuden esittämisen ongelmaa syväoppimisessa uhraamatta laskentatehoa tai testitarkkuutta. Suoritamme eksploratiivisen tutkimuksen pudotetun epävarmuuden ominaisuuksista. Erilaisia verkkoarkkitehtuureja ja epälineaarisuuksia arvioidaan ekstrapolointi-, interpolointi- ja luokittelutehtävissä. Osoitamme, että mallin epävarmuus on tärkeää luokittelutehtävissä käyttäen MNISTiä esimerkkinä, ja käytämme mallin epävarmuutta Bayesin putkistossa, jossa syvä vahvistusoppiminen on konkreettinen esimerkki.

**Tulos**

Pudottaminen Bayesin approksimaationa: Mallin epävarmuuden esittäminen syväoppimisessa

**Esimerkki 2.770**

Tässä asiakirjassa on kaksi osaa. Ensimmäisessä osassa käsitellään sanojen upotuksia. Keskustelemme niiden tarpeesta, joistakin menetelmistä niiden luomiseksi ja joistakin niiden mielenkiintoisista ominaisuuksista. Vertaamme niitä myös kuvan upotuksiin ja tarkastelemme, miten sanojen upottamista ja kuvan upottamista voidaan yhdistää eri tehtävien suorittamiseksi. Toisessa osassa toteutamme konvoluutiohermoverkon, joka on koulutettu valmiiksi koulutettujen sanavektoreiden päälle. Verkkoa käytetään useisiin lausetason luokittelutehtäviin, ja sillä saavutetaan huipputason (tai vertailukelpoisia) tuloksia, mikä osoittaa ennalta koulutettujen sanojen upotusten suuren tehon satunnaisiin upotuksiin verrattuna.

**Tulos**

Sanojen sulautukset ja niiden käyttö lauseiden luokittelutehtävissä

**Esimerkki 2.771**

Stokastisen optimoinnin asynkroniset rinnakkaistoteutukset ovat viime aikoina saaneet valtavan suosion teoriassa ja käytännössä. Asynkroniset toteutukset, joissa ei ole lukitusta, ovat tehokkaampia kuin sellaiset, joissa on kirjoitus- tai lukituslukko. Tässä artikkelissa keskitymme yhdistettyyn tavoitefunktioon, joka koostuu sileästä koverasta funktiosta f ja lohko-erotettavasta koverasta funktiosta, joita esiintyy laajalti koneoppimisessa ja tietokonenäössä. Ehdotamme asynkronista stokastista lohkokoordinaattista laskeutumisalgoritmia, jossa on nopeutettu varianssin vähentämistekniikka (AsySBCDVR), jotka ovat lukituksettomia toteutuksessa ja analyysissä. AsySBCDVR on erityisen tärkeä, koska se voi skaalautua hyvin otoskoon ja ulottuvuuden kanssa samanaikaisesti. Osoitamme, että AsySBCDVR saavuttaa lineaarisen konvergenssinopeuden, kun funktio f on optimaalisen vahvan konveksisuuden ominaisuuden kanssa, ja sublineaarisen nopeuden, kun f on yleisen konveksisuuden kanssa. Vielä tärkeämpää on, että jaetulla muistilla varustetussa rinnakkaisessa järjestelmässä voidaan saavuttaa lähes lineaarinen nopeus.

**Tulos**

Asynkroninen stokastinen lohkokoordinaattien laskeutuminen varianssin vähentämisen avulla

**Esimerkki 2.772**

Tarkastelemme yleistä verkko-oppimista asiantuntijoiden neuvojen avulla, jossa katumus määritellään suhteessa kilpailijaluokkaan, joka on määritelty painotetulla automaatilla, joka kattaa asiantuntijoiden sarjat. Kehyksemme kattaa useita aiemmin tutkittuja ongelmia, erityisesti ongelman, jossa kilpaillaan k-siirtyviä asiantuntijoita vastaan. Esitämme useita algoritmeja tälle ongelmalle, mukaan lukien automaattipohjainen algoritmi, joka laajentaa painotettua enemmistöä, ja tehokkaampia algoritmeja, jotka perustuvat epäonnistumisen siirtymien käsitteeseen. Lisäksi esitämme tehokkaita algoritmeja, jotka perustuvat kilpailija-automaatin kompaktiin approksimaatioon, erityisesti tehokkaisiin n-grammimalleihin, jotka saadaan minimoimalla Rényin divergenssi, ja tutkimme laajasti tällaisten mallien approksimaatio-ominaisuuksia. Laajennamme algoritmeja ja tuloksia myös nukkuvien asiantuntijoiden kehykseen. Lopuksi kuvaamme approksimaatiomenetelmiemme laajentamista verkossa tapahtuvaan koveraan optimointiin ja yleiseen peililähtöasetelmaan.

**Tulos**

Verkko-oppiminen asiantuntija-automaatteja vastaan

**Esimerkki 2.773**

Tässä artikkelissa esitellään kahden kerroksen kernel-koneiden kehys, joka yleistää klassisia kernel-menetelmiä. Uusi oppimismenetelmä tarjoaa muodollisen yhteyden useilla kerroksilla varustettujen laskenta-arkkitehtuurien ja tavallisissa regularisointimenetelmissä käytetyn kerneloppimisen teeman välille. Ensin esitetään kaksikerroksisille verkoille edustajateoreema, joka osoittaa, että kullakin kerroksella olevien ytimien äärelliset lineaariset yhdistelmät ovat optimaalisia arkkitehtuureja aina, kun vastaavat funktiot ratkaisevat sopivia variaatio-ongelmia toistuvissa kernel-Hilbert-avaruuksissa (RKHS). Näiden arkkitehtuurien ilmaisema tulo-lähtökartta osoittautuu vastaavaksi kuin sopivat yksikerroksiset kernelikoneet, joissa myös kernelifunktio opitaan datasta. Viime aikoina niin sanotut usean ytimen oppimismenetelmät ovat saaneet paljon huomiota koneoppimiskirjallisuudessa. Tässä artikkelissa osoitetaan, että moninkertaisen ytimen oppimismenetelmät ovat erityistapauksia kaksikerroksisista kernel-koneista, joissa toinen kerros on lineaarinen. Lopuksi esitellään yksinkertainen ja tehokas usean ytimen oppimismenetelmä nimeltä RLS2 (regularized least squares with two layers), ja sen suorituskykyä useissa oppimisongelmissa analysoidaan laajasti. Käytettävissä on avoimen lähdekoodin MATLAB-työkaluryhmä RLS2-mallien kouluttamiseen ja validointiin graafisella käyttöliittymällä.

**Tulos**

Kernel-koneet, joissa on kaksi kerrosta ja usean kernelin oppiminen

**Esimerkki 2.774**

Tutkimme kahta sekalaista robustia/keskimääräistapauksen submodulaarista jako-ongelmaa, joita kutsumme yhteisnimellä Submodulaarinen jako. Nämä ongelmat yleistävät sekä ongelman puhtaasti robustit tapaukset (nimittäin max-min submodulaarinen oikeudenmukainen jako (SFA) Golovin (2005) ja min-max submodulaarinen kuormanjako (SLB) Svitkina ja Fleischer (2008)) että myös keskimääräiset tapaukset (eli submodulaarinen hyvinvointiongelma (SWP) Vondrák (2008) ja submodulaarinen monitieosio (SMP) Chekuri ja Ene (2011a)). Vaikka robusteja versioita on tutkittu teoriayhteisössä Goemans et al. (2009); Golovin (2005); Khot ja Ponnuswami (2007); Svitkina ja Fleischer (2008); Vondrák (2008), nykyisessä työssä on keskitytty tiukkoihin approksimointitakuisiin, eivätkä tuloksena saadut algoritmit yleensä ole skaalautuvia hyvin suuriin reaalimaailman sovelluksiin. Tämä on ristiriidassa keskimääräisen tapauksen kanssa, jossa useimmat algoritmit ovat skaalautuvia. Tässä artikkelissa kuromme umpeen tämän kuilun ehdottamalla useita uusia algoritmeja (mukaan lukien ahneisiin, majorisaatio-minimisaatio-, minorisaatio-maximisaatio- ja relaksaatioalgoritmeihin perustuvat algoritmit), jotka skaalautuvat suuriin kokoluokkiin ja joilla saavutetaan teoreettiset approksimaatiotakuut, jotka ovat lähellä huipputason tuloksia, ja joissain tapauksissa saavutetaan uusia tiukkoja rajoja. Tarjoamme myös uusia skaalautuvia algoritmeja, joita voidaan soveltaa robustien ja keskimääräisten ääritavoitteiden additiivisiin yhdistelmiin. Osoitamme, että näillä ongelmilla on monia sovelluksia koneoppimisessa (ML). Näitä ovat mm: 1) data 1 ar X iv :1 51 0. 08 86 5v 2 [ cs .D S] 1 6 A ug 2 01 6 Wei, Iyer, Wang, Bai, Bilmes osiointi ja kuorman tasapainottaminen hajautetuille konealgoritmeille rinnakkaiskoneissa; 2) datan klusterointi; ja 3) monimerkkisten kuvien segmentointi (vain) Boolen submodulaaristen funktioiden avulla pikselin osioinnin avulla. Osoitamme empiirisesti algoritmiemme tehokkuuden reaalimaailman ongelmissa, joihin liittyy datan osiointi tavanomaisten koneoppimistavoitteiden hajautettua optimointia varten (mukaan lukien sekä konveksiset että syvät neuroverkkotavoitteet), sekä myös puhtaasti valvomattomassa (eli ei valvottua tai puolivalvottua oppimista eikä vuorovaikutteista segmentointia) kuvan segmentoinnissa.

**Tulos**

Yhdistetty robusti/keskimääräinen submodulaarinen jako (Mixed Robust/Average Submodular Partitioning): Nopeat algoritmit, takuut ja sovellukset rinnakkaiseen koneoppimiseen ja monileimakuvasegmentointiin.

**Esimerkki 2.775**

<Osoitamme, että minkä tahansa stokastisella gradienttimenetelmällä koulutetun mallin yleistysvirhe on häviävä. Todistamme tämän osoittamalla, että menetelmä on algoritmisesti stabiili<lb>Bousquet'n ja Elisseeffin merkityksessä. Analyysimme käyttää vain konveksisen<lb>ja jatkuvan optimoinnin alkeistyökaluja. Tuloksemme pätevät sekä konveksiseen että ei-konveksiseen optimointiin<lb>vakioiduilla Lipschitz- ja tasaisuusoletuksilla.<lb>Soveltamalla tuloksiamme konveksiseen tapaukseen tarjoamme uusia selityksiä sille, miksi stokastisen gradienttilaskeutumisen useat<lb>epochit yleistyvät hyvin käytännössä. Epäkonveksisessa tapauksessa tarjoamme<lb>uuden tulkinnan neuroverkkojen yleisistä käytännöistä ja muodollisen<lb>perustelun vakautta edistäville mekanismeille suurten, syvien mallien koulutuksessa. Käsitteellisesti<lb>löydöksemme korostavat koulutusajan lyhentämisen tärkeyttä sen ilmeisen hyödyn lisäksi.

**Tulos**

Stokastisen gradienttilaskeutumisen stabiilius

**Esimerkki 2.776**

Tietueiden yhdistämisen ja usean kohteen seurannan kaltaiset tehtävät, joissa on kyse havaittujen tietojen taustalla olevien kohteiden joukon rekonstruoinnista, ovat erityisen haastavia probabilistisen päättelyn kannalta. Viimeaikaisessa työssä on saavutettu tehokkaita ja tarkkoja päätelmiä tällaisissa ongelmissa käyttämällä Markovin ketju Monte Carlo (MCMC) -tekniikoita, joissa on räätälöityjä ehdotusjakaumia. Tällä hetkellä tällaisen järjestelmän toteuttaminen edellyttää MCMC-tilojen esitysten ja hyväksymistodennäköisyyslaskentojen koodaamista, jotka ovat sovelluskohtaisia. Vaihtoehtoisena lähestymistapana, jota tässä asiakirjassa noudatetaan, on käyttää yleiskäyttöistä todennäköisyysmallinnuskieltä (kuten BLOG) ja yleistä Metropolis-Hastings MCMC-algoritmia, joka tukee käyttäjän antamia ehdotusjakaumia. Algoritmimme saa joustavuutta käyttämällä MCMC-tiloja, jotka ovat vain osittaisia kuvauksia mahdollisista maailmoista; annamme ehdot, joiden mukaan MCMC osittaisten maailmojen yli tuottaa oikeita vastauksia kyselyihin. Näytämme myös, miten kontekstikohtaista Bayes-verkkoa käytetään tunnistamaan hyväksymistodennäköisyyden tekijät, jotka on laskettava tietylle ehdotetulle siirrolle. Kokeelliset tulokset sitaattien täsmäytystehtävässä osoittavat, että yleiskäyttöinen MCMC-moottorimme on hyvässä vertailussa sovelluskohtaisen järjestelmän kanssa.

**Tulos**

Yleiskäyttöinen MCMC-ulottuvuus relaatiorakenteiden yli

**Esimerkki 2.777**

Suosittelujärjestelmissä käytetään usein latentteja ominaisuuksia selittämään käyttäjien käyttäytymistä ja kuvaamaan kohteiden ominaisuuksia. Kun käyttäjät ovat vuorovaikutuksessa eri kohteiden kanssa ajan mittaan, käyttäjä- ja kohdeominaisuudet voivat vaikuttaa toisiinsa, kehittyä ja kehittyä yhdessä ajan mittaan. Näiden ominaisuuksien hienojakoisen epälineaarisen yhteisevoluution kuvaamiseksi täsmällisesti ehdotamme rekursiivista yhteisevoluutio-ominaisuuksien sulauttamisprosessimallia, jossa yhdistyvät rekursiivinen neuroverkko (RNN) ja moniulotteinen pisteprosessimalli. RNN oppii epälineaarisen esityksen käyttäjän ja kohteen ominaisuuksista, joissa otetaan huomioon käyttäjän ja kohteen ominaisuuksien keskinäinen vaikutus ja ominaisuuksien kehittyminen ajan myötä. Kehitämme myös tehokkaan stokastisen gradienttialgoritmin malliparametrien oppimista varten, joka voidaan helposti skaalata miljooniin tapahtumiin. Kokeet erilaisilla reaalimaailman tietokokonaisuuksilla osoittavat, että käyttäjien käyttäytymisen ennustaminen on parantunut merkittävästi verrattuna nykyiseen tekniikkaan.

**Tulos**

Toistuvat yhteisevolutionaariset ominaisuuksien sulauttamisprosessit suosituksia varten

**Esimerkki 2.778**

Useilla autonvalmistajilla on kehitteillä tai tuotannossa automatisoituja ajojärjestelmiä (ADS), jotka tarjoavat moottoritiepilottitoimintoja. Tämäntyyppiset ADS-järjestelmät rajoittuvat yleensä vain rajoitetusti liikennöitäviin moottoriteihin, toisin sanoen siirtyminen manuaalisesta tilasta automaattiseen tilaan tapahtuu vasta sen jälkeen, kun ramppien yhdistämisprosessi on suoritettu manuaalisesti. Yksi suuri haaste automaation laajentamisessa ramppien sulautumiseen on se, että automaattisen ajoneuvon on otettava huomioon ja optimoitava pitkän aikavälin tavoitteet (esim. onnistunut ja sujuva sulautuminen), kun lyhyen aikavälin toimet on suoritettava turvallisesti. Lisäksi sulautumisprosessiin liittyy vuorovaikutusta muiden ajoneuvojen kanssa, joiden käyttäytymistä on joskus vaikea ennustaa mutta jotka voivat vaikuttaa sulautuvan ajoneuvon optimaalisiin toimiin. Tällaisen monimutkaisen ohjausongelman ratkaisemiseksi ehdotamme, että sovellamme syvävahvistusoppimisen (Deep Reinforcement Learning, DRL) tekniikoita optimaalisen ajopolitiikan löytämiseksi maksimoimalla pitkän aikavälin palkkio vuorovaikutteisessa ympäristössä. Käytämme vuorovaikutteisen ympäristön mallintamiseen LSTM-arkkitehtuuria (Long Short-Term Memory), josta sisäinen tila, joka sisältää historiallisia ajotietoja, välitetään Deep Q-Networkille (DQN). DQN:ää käytetään Q-funktion approksimointiin, joka ottaa sisäisen tilan syötteenä ja tuottaa Q-arvoja tuotoksena toiminnan valintaa varten. Tämän DRL-arkkitehtuurin avulla vuorovaikutteisen ympäristön historiallinen vaikutus pitkän aikavälin palkkioon voidaan ottaa huomioon optimaalista ohjauskäytäntöä määritettäessä. Ehdotettua arkkitehtuuria on mahdollista laajentaa ja soveltaa muihin autonomisen ajamisen skenaarioihin, kuten monimutkaisen risteyksen läpi ajamiseen tai kaistanvaihtoon vaihtuvissa liikennevirtaolosuhteissa. Avainsanat- Autonominen ajaminen; moottoritien rampin sulautuminen; syvä vahvistusoppiminen; pitkä lyhytkestoinen muisti; syvä Q-verkko; ohjauspolitiikka.

**Tulos**

Syvän vahvistusoppimisarkkitehtuurin muotoilu kohti autonomista ajamista rampin sulkeutumista varten

**Esimerkki 2.779**

Stokastisen variationaalisen päättelyn (SVI) avulla voimme skaalata Bayes-laskentaa massiivisille aineistoille. Se käyttää stokastista optimointia variaatiojakauman sovittamiseen helposti laskettavien luonnollisten, meluisten gradienttien mukaan. Kuten useimmissa perinteisissä stokastisissa optimointimenetelmissä, SVI:ssä käytetään varotoimia, jotta voidaan käyttää puolueettomia stokastisia gradientteja, joiden odotukset ovat yhtä suuret kuin todelliset gradientit. Tässä artikkelissa tutkimme ajatusta puolueellisten stokastisten gradienttien seuraamisesta SVI:ssä. Menetelmämme korvaa luonnollisen gradientin vastaavalla tavalla rakennetulla vektorilla, joka käyttää kiinteän ikkunan liukuvaa keskiarvoa joistakin edellisistä termeistään. Osoitamme tämän tekniikan monet edut. Ensinnäkin sen laskentakustannukset ovat samat kuin SVI:llä, ja tallennustarpeet lisääntyvät vain vakiokertoimella. Toiseksi se pienentää huomattavasti varianssia verrattuna harhattomiin estimaatteihin, pienentää harhaa verrattuna keskiarvogradientteihin ja johtaa pienempään keskineliövirheeseen verrattuna koko gradienttiin. Testaamme menetelmäämme latentilla Dirichlet-allokaatiolla kolmen suuren korpuksen avulla.

**Tulos**

Tasoitetut gradientit stokastista variatiivista päättelyä varten

**Esimerkki 2.780**

ADMM-menetelmä (Alternating Direction Method of Multippliers) on tunnustettu monipuoliseksi lähestymistavaksi nykyaikaisten laajamittaisten koneoppimis- ja signaalinkäsittelyongelmien tehokkaaseen ratkaisemiseen. Kun datan koko ja/tai ongelman ulottuvuus on suuri, voidaan käyttää ADMM:n hajautettua versiota, joka pystyy jakamaan laskentakuorman ja datajoukon laskentasolmujen verkolle. Valitettavasti tällaisen algoritmin suora synkroninen toteutus ei skaalautu hyvin ongelman koon kasvaessa, sillä algoritmin nopeutta rajoittavat hitaimmat laskentasolmut. Tämän ongelman ratkaisemiseksi ehdotimme eräässä oheisjulkaisussa asynkronista hajautettua ADMM-algoritmia (AD-ADMM) ja tutkimme sen pahimman tapauksen konvergenssiehtoja. Tässä asiakirjassa jatkamme tutkimusta kuvaamalla ehtoja, joilla AD-ADMM saavuttaa lineaarisen konvergenssin. Olosuhteemme sekä tuloksena saadut lineaariset nopeudet paljastavat eri algoritmiparametrien, verkon viiveen ja verkon koon vaikutuksen algoritmin suorituskykyyn. Osoittaaksemme ehdotetun AD-ADMM:n ylivoimaisen aikatehokkuuden, testaamme AD-ADMM:ää suuritehoisella tietokoneklusterilla ratkaisemalla suuren mittakaavan logistisen regression ongelman. Avainsanat- Hajautettu optimointi, ADMM, Asynkroninen, Konsensusoptimointi ⋆Tsung-Hui Chang on vastaava kirjoittaja. Osoite: School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Kiina 518172, E-mail: tsunghui.chang@ieee.org. Wei-Cheng Liao on sähkö- ja tietotekniikan laitokselta, Minnesotan yliopisto, Minneapolis, MN 55455, Yhdysvallat, Sähköposti: mhong@umn.edu Mingyi Hong on teollisuus- ja tuotantojärjestelmätekniikan laitokselta, Iowa State University, Ames, 50011, Yhdysvallat, Sähköposti: mingyi@iastate.edu Xiangfeng Wang on Shanghai Key Lab for Trustworthy Computing, Software Engineering Institute, East China Normal University, Shanghai, 200062, Kiina, Sähköposti: xfwang@sei.ecnu.edu.cn 10. syyskuuta 2015 LUONNOS.

**Tulos**

Asynkroninen hajautettu ADMM suuren mittakaavan optimointiin - Osa II: Lineaarinen konvergenssianalyysi ja numeerinen suorituskyky

**Esimerkki 2.781**

Monissa peräkkäisissä päätöksenteko-ongelmissa saatamme haluta hallita riskiä minimoimalla kustannusten vaihtelun jonkin mittarin sen lisäksi, että minimoimme vakiokriteerin. Ehdollinen riskiarvo (CVaR, Conditional Value-at-Risk) on suhteellisen uusi riskimittari, joka korjaa joitakin tunnettujen varianssiin liittyvien riskimittareiden puutteita, ja laskennallisen tehokkuutensa vuoksi se on saavuttanut suosiota rahoituksessa ja operaatiotutkimuksessa. Tässä artikkelissa tarkastelemme keskiarvo-CVaR-optimoinnin ongelmaa MDP:ssä. Johdamme ensin kaavan tämän riskisensitiivisen kohdefunktion gradientin laskemiseksi. Sen jälkeen kehitämme politiikkagradientti- ja toimijakriittiset algoritmit, jotka kumpikin käyttävät tiettyä menetelmää tämän gradientin estimointiin ja päivittävät politiikkaparametrit laskeutumissuunnassa. Todistamme algoritmiemme konvergenssin paikallisesti riskiherkkiin optimaalisiin politiikkoihin. Lopuksi osoitamme algoritmiemme hyödyllisyyden optimaalisen pysäytysongelman yhteydessä.

**Tulos**

Algoritmit CVaR-optimoinnissa MDP:ssä

**Esimerkki 2.782**

Huolimatta todistusteorian merkittävästä kehityksestä on todistuksen todentajan käsitteeseen kiinnitetty yllättävän vähän huomiota. Erityisesti matemaattinen yhteisö voi olla kiinnostunut tutkimaan erityyppisiä todisteiden todentajia (ihmisiä, ohjelmia, oraakkeleita, yhteisöjä, superälyjä jne.) matemaattisina objekteina, niiden ominaisuuksia, niiden valtuuksia ja rajoituksia (erityisesti ihmismatemaatikoilla), minimi- ja maksimikompleksisuutta sekä todentajien itsevarmennukseen ja itseviittaukseen liittyviä kysymyksiä. Ehdotamme alustavaa luokittelujärjestelmää todentajille ja esitämme alkeellisen analyysin tämän tärkeän alan ratkaistuista ja avoimista ongelmista. Tärkein panoksemme on epäkorjattavuuden käsitteen muodollinen esittely, jonka osalta artikkeli voisi toimia yleisenä siteerauksena teoreemojen todistamisen, ohjelmistojen ja tekoälyn verifioinnin aloilla.

**Tulos**

Todentajateoria aksioomeista matemaattisten todisteiden todentamattomuuteen, ohjelmistot ja tekoälytekniikka

**Esimerkki 2.783**

Artikkelissa tutkitaan koneoppimisongelmia, joissa jokainen esimerkki kuvataan joukon Boolen piirteiden avulla ja joissa hypoteesit esitetään lineaarisilla kynnyselementeillä. Yksi tapa lisätä opittujen hypoteesien ilmaisuvoimaa tässä yhteydessä on laajentaa ominaisuusjoukkoa siten, että se sisältää perusominaisuuksien yhdistelmiä. Tämä voidaan tehdä eksplisiittisesti tai mahdollisuuksien mukaan käyttämällä kernel-funktiota. Keskitytään tunnettuihin Perceptron- ja Winnow-algoritmeihin ja osoitetaan kompromissi sen laskennallisen tehokkuuden, jolla algoritmi voidaan ajaa laajennetussa ominaisuusavaruudessa, ja vastaavan oppimisalgoritmin yleistämiskyvyn välillä. Aluksi kuvataan useita kernelfunktioita, jotka kuvaavat joko rajoitettuja konjunktiomuotoja tai kaikkia konjunktioita. Osoitamme, että näitä ytimiä voidaan käyttää Perceptron-algoritmin tehokkaaseen suorittamiseen eksponentiaalisen monien konjunktioiden piirreavaruudessa; osoitamme kuitenkin myös, että tällaisia ytimiä käyttäen Perceptron-algoritmi voi todistettavasti tehdä eksponentiaalisen määrän virheitä jopa yksinkertaisia funktioita opittaessa. Tämän jälkeen tarkastelemme kysymystä siitä, voidaanko kernel-funktioita käyttää vastaavasti multiplikatiivisen päivityksen Winnow-algoritmin suorittamiseen laajennetussa funktioavaruudessa, jossa on eksponentiaalisen monta konjunktiota. Tunnetut ylärajat viittaavat siihen, että Winnow-algoritmi voi oppia DNF-kaavoja (Disjunctive Normal Form) polynomiaalisella virherajalla tässä ympäristössä. Osoitamme kuitenkin, että on laskennallisesti vaikeaa simuloida Winnowin käyttäytymistä DNF:n oppimisessa tällaisessa ominaisuusavaruudessa. Tämä tarkoittaa, että kernel-funktiot, jotka vastaavat Winnow'n ajamista tämän ongelman osalta, eivät ole tehokkaasti laskettavissa ja että ei ole olemassa yleistä konstruktiota, jolla Winnow'ta voitaisiin ajaa kerneleiden avulla.

**Tulos**

Tehokkuus ja Boolen ytimien konvergenssi online-oppimisalgoritmeja varten

**Esimerkki 2.784**

Tässä artikkelissa tarkastelemme monitehtäväoppimisen ongelmaa, jossa oppijalle annetaan joukko ennustustehtäviä, jotka on ratkaistava. Aiemmista töistä poiketen luovumme olettamuksesta, että kaikille tehtäville on saatavilla merkittyä harjoitusdataa. Sen sijaan ehdotamme aktiivista tehtävänvalintaa, jossa oppija voi valita vain merkitsemättömän datan perusteella tyypillisesti pienen osajoukon tehtäviä, joista hän saa merkityt esimerkit. Jäljelle jääviin tehtäviin, joihin ei ole saatavilla merkintöjä, etsitään ratkaisuja siirtämällä tietoa valituista tehtävistä. Analysoimme kahta siirtostrategiaa ja kehitämme yleistysrajat kummallekin. Tämän teoreettisen analyysin perusteella ehdotamme kahta algoritmia, joilla leimattujen tehtävien valinta tehdään periaatteellisella tavalla, ja osoitamme niiden tehokkuuden synteettisellä ja todellisella datalla.

**Tulos**

Aktiivinen tehtävän valinta monitehtäväoppimista varten

**Esimerkki 2.785**

Sosiaalisten verkostojen tutkimus on kasvava tutkimusalue. Suurin osa nykyisestä tutkimuksesta käsittelee kuitenkin verkostoja, jotka yksinkertaisesti koodaavat, onko suhteita olemassa vai ei. Sen sijaan allekirjoitetuissa verkostoissa suhteet voivat olla positiivisia ("tykkää", "luottaa") tai negatiivisia ("ei pidä", "epäluottamus"). Sosiaalisen tasapainon teoria osoittaa, että allekirjoitetuilla verkoilla on taipumus noudattaa joitakin paikallisia malleja, jotka puolestaan saavat aikaan tiettyjä globaaleja piirteitä. Tässä artikkelissa hyödynnämme sosiaalisen tasapainon teorian sekä paikallisia että globaaleja näkökohtia kahdessa allekirjoitettujen verkkojen analyysin perusongelmassa: merkkien ennustamisessa ja klusteroinnissa. Sosiaalisen tasapainon paikallisten mallien perusteella ehdotamme ensin kahta merkkien ennustamismenetelmäperhettä: sosiaalisen epätasapainon mittareita (MOIs) ja valvottua oppimista, jossa käytetään korkean järjestyksen syklejä (HOC). Nämä menetelmät ennustavat reunojen merkkejä kolmioiden ja l-syklien perusteella suhteellisen pienillä l:n arvoilla. Mielenkiintoista on, että tarkastelemalla sosiaalisen epätasapainon mittoja osoitamme, että klassisella Katzin mittarilla, jota käytetään laajalti merkitsemättömien linkkien ennustamisessa, on itse asiassa tasapainoteoreettinen tulkinta, kun sitä sovelletaan merkityille verkoille. Tasapainotettujen verkkojen globaalin rakenteen perusteella ehdotamme lisäksi tehokasta matalan luokan mallinnusmenetelmää sekä merkkien ennustamista että klusterointia varten. Tarjoamme matalan luokan mallinnusmenetelmälle teoreettiset suorituskykytakeet konveksisten relaksaatioiden avulla, skaalaamme sen suuriin ongelmakokoihin käyttämällä matriisifaktorointiin perustuvaa algoritmia ja tarjoamme laajaa kokeellista validointia, mukaan lukien vertailut paikallisten lähestymistapojen kanssa. Kokeelliset tuloksemme osoittavat, että omaksumalla globaalimman näkökulman tasapainorakenteeseen saamme merkittäviä suorituskyky- ja laskennallisia parannuksia ennustus- ja klusterointitehtävissä allekirjoitetuissa verkoissa. Työssämme korostuu siis tasapainoteorian globaalin näkökulman hyödyllisyys allekirjoitettujen verkkojen analysoinnissa.

**Tulos**

Ennustaminen ja klusterointi allekirjoitetuissa verkoissa: Paikallisesta globaaliin näkökulmaan

**Esimerkki 2.786**

Generatiivisia vastakkaisverkkoja on ehdotettu keinoksi kouluttaa tehokkaasti syviä generatiivisia neuroverkkoja. Ehdotamme generatiivista vastakkaisverkkomallia, joka toimii jatkuvassa sekventiaalisessa datassa, ja sovellamme sitä kouluttamalla sitä klassisen musiikin kokoelmaan. Toteamme, että se tuottaa musiikkia, joka kuulostaa yhä paremmalta, kun mallia koulutetaan, raportoimme tilastoja tuotetusta musiikista ja annamme lukijan arvioida laatua lataamalla tuotettuja kappaleita.

**Tulos**

C-RNN-GAN: Jatkuvat toistuvat neuroverkot, joissa on vastakohtainen koulutus.

**Esimerkki 2.787**

Koulutamme generaattorin suurimman todennäköisyyden avulla ja koulutamme saman generaattoriarkkitehtuurin myös Wasserstein GAN:lla. Sen jälkeen vertaamme luotuja näytteitä, tarkkoja log-todennäköisyystiheyksiä ja likimääräisiä Wassersteinin etäisyyksiä. Osoitamme, että riippumaton kriitikko, joka on koulutettu lähentämään Wassersteinin etäisyyttä validointijoukon ja generaattorijakauman välillä, auttaa havaitsemaan ylisovittamisen. Lopuksi kehitämme uudenlaisen nopean oppimisen kriitikon, joka perustuu ideoihin, jotka ovat peräisin yhden otoksen oppimista käsittelevästä kirjallisuudesta.

**Tulos**

Todellisten NVP:iden Maximum Likelihood- ja GAN-pohjaisen koulutuksen vertailu

**Esimerkki 2.788**

Neuraaliset Turingin koneet (NTM) [2] sisältävät muistikomponentin, joka simuloi aivojen "työmuistia" tietojen tallentamiseksi ja hakemiseksi, mikä helpottaa yksinkertaisten algoritmien oppimista. Toistaiseksi on ehdotettu vain lineaarisesti organisoitua muistia, ja kokeilujen aikana havaitsimme, että malli ei aina konvergoi ja että se ylisovittuu helposti tiettyjä tehtäviä käsiteltäessä. Uskomme, että muistikomponentti on avain joihinkin NTM:n virheellisiin käyttäytymismalleihin, ja muistikomponentin parempi organisointi voisi auttaa torjumaan näitä ongelmia. Tässä artikkelissa ehdotamme NTM:lle useita erilaisia muistirakenteita, ja todistimme kokeissa, että kaksi ehdottamaamme NTM-mallia, joka sisältää rakenteellisen muistin, voi johtaa parempaan konvergenssiin nopeuden ja ennustustarkkuuden suhteen kopiointitehtävässä ja assosiatiivisessa muistitehtävässä kuin artikkelissa [2].

**Tulos**

Neuraalisten Turingin koneiden strukturoitu muisti

**Esimerkki 2.789**

Label distribution learning (LDL) on yleinen oppimiskehys, jossa instanssille annetaan yhden tai useamman labelin sijasta jakauma useiden labelien joukosta. Nykyisillä LDL-menetelmillä on joko rajoitettuja oletuksia merkkijakauman ilmaisumuodosta tai rajoituksia representaatio-oppimisessa. Tässä artikkelissa esitellään etikettijakauman oppimismetsät (label distribution learning forests, LDLFs), joka on uusi differentioituviin päätöspuihin perustuva etikettijakauman oppimisalgoritmi, jolla on useita etuja: 1) Päätöspuilla on mahdollisuus mallintaa mitä tahansa yleistä etikettijakaumien muotoa lehtisolmujen ennusteiden sekoituksella. 2) Differentioituvien päätöspuiden oppiminen voidaan yhdistää representaatio-oppimiseen, esimerkiksi syvien ominaisuuksien oppimiseen päästä päähän -periaatteella. Määrittelemme metsille jakaumapohjaisen tappiofunktion, joka mahdollistaa kaikkien puiden oppimisen yhdessä, ja osoitamme, että lehtisolmujen ennusteiden päivitysfunktio, joka takaa tappiofunktion tiukan pienenemisen, voidaan johtaa variationaalisen rajoittamisen avulla. Ehdotettujen LDLF:ien tehokkuus todennetaan kahdessa LDL-ongelmassa, kuten iän arvioinnissa ja elokuvien yleisön mielipiteiden ennustamisessa, ja ne osoittavat merkittäviä parannuksia uusimpiin LDL-menetelmiin verrattuna.

**Tulos**

Etiketin jakauman oppimismetsät

**Esimerkki 2.790**

Tässä artikkelissa tutkitaan parametrisoituja likimääräisiä viestien välitysjärjestelmiä, jotka perustuvat rajoitettuun päättelyyn ja jotka ovat saaneet vaikutteita Pearlin uskomusten etenemisalgoritmista (BP). Aloitamme rajoitetun päättelyn miniklusterointialgoritmilla ja siirrymme sitten iteratiiviseen järjestelmään nimeltä Iterative Join-Graph Propagation (IJGP), jossa yhdistyvät sekä iterointi että rajoitettu päättely. Algoritmi IJGP kuuluu Generalized Belief Propagation -algoritmien luokkaan, joka on kehys, joka salli yhteydet tilastollisen fysiikan approksimatiivisiin algoritmeihin, ja sen on empiirisesti osoitettu ylittävän miniklusterointi- ja uskomuspropagointialgoritmien suorituskyvyn sekä useiden muiden nykyaikaisten algoritmien suorituskyvyn useissa verkkoluokissa. Tarjoamme myös tietoa iteratiivisen BP:n ja IJGP:n tarkkuudesta suhteuttamalla nämä algoritmit tunnettuihin rajoitusten etenemisjärjestelmien luokkiin.

**Tulos**

Join-Graph Propagation -algoritmit

**Esimerkki 2.791**

Esitämme tässä artikkelissa tutkimuksen, jossa tarkastellaan näppäilydynamiikan todennusmenetelmän käyttömahdollisuuksia ja hyötyjä yhteistoimintajärjestelmissä. Todentaminen on haastava kysymys, kun halutaan taata yhteistoiminnallisten järjestelmien käytön turvallisuus pääsynvalvontavaiheessa. Nykyisessä tekniikassa on monia ratkaisuja, kuten kertakäyttösalasanojen tai älykorttien käyttö. Tässä asiakirjassa keskitytään biometrisiin ratkaisuihin, jotka eivät edellytä mitään ylimääräistä tunnistinta. Näppäilydynamiikka on mielenkiintoinen ratkaisu, koska se käyttää vain näppäimistöä ja on käyttäjille näkymätön. Tällä alalla on julkaistu monia menetelmiä. Teemme vertailevan tutkimuksen monista niistä ottaen huomioon yhteistoiminnallisissa järjestelmissä käytettävät toiminnalliset rajoitukset.

**Tulos**

Näppäilydynamiikan todennus yhteistoimintajärjestelmiä varten

**Esimerkki 2.792**

Esittelemme globaalisti konvergentin algoritmin puupainotetun (TRW) variaatiotavoitteen optimoimiseksi marginaalisen polytoopin yli. Algoritmi perustuu ehdolliseen gradienttimenetelmään (Frank-Wolfe) ja siirtää pseudomarginaalit marginaalipolytoopin sisällä toistuvien MAP-kutsujen (Maximum a posteriori) avulla. Tämän modulaarisen rakenteen ansiosta voimme hyödyntää mustan laatikon MAP-ratkaisuja (sekä tarkkoja että likimääräisiä) variatiiviseen päättelyyn, ja saamme tarkempia tuloksia kuin puupainotteiset algoritmit, jotka optimoivat paikallisen johdonmukaisuusrelaksaation yli. Teoreettisesti rajoitamme ehdotetun algoritmin alioptimaalisuuden huolimatta siitä, että TRW-tavoitteella on rajoittamattomat gradientit marginaalisen polytoopin rajalla. Empiirisesti osoitamme, että tulosten laatu paranee, kun tiukennamme relaksaatiota marginaalisen polytoopin ja jännityspuupolytoopin yli synteettisillä ja todellisilla tapauksilla.

**Tulos**

Frank-Wolfe-este marginaalista päättelyä varten

**Esimerkki 2.793**

Kun tällaisten tapahtumapaikkojen tyyppi ja määrä kasvavat, tekstivarantojen automaattisesta tunteiden analysoinnista on tullut olennainen tiedonlouhintatehtävä. Tässä artikkelissa tutkimme ongelmaa, joka liittyy mielipiteiden louhintaan epävirallisten lyhyiden tekstien kokoelmasta. Teksteistä havaitaan sekä positiivisen että negatiivisen tunteen voimakkuus. Keskitymme muuhun kuin englannin kieleen, jolla on vain vähän resursseja tekstinlouhintaa varten. Tämä lähestymistapa auttaisi parantamaan sentimenttianalyysia kielissä, joissa ei ole luetteloa mielipiteitä sisältävistä sanoista. Ehdotamme uutta menetelmää, jossa teksti projisoidaan tiheisiin ja matalaulotteisiin ominaisvektoreihin sanojen tunteen voimakkuuden mukaan. Tunnistamme positiivisten ja negatiivisten tunteiden sekoituksen monimuuttujaisella asteikolla. Ehdotetun kehyksen empiirinen arviointi turkkilaisilla twiiteillä osoittaa, että lähestymistapamme tuottaa hyviä tuloksia mielipiteiden louhinnassa.

**Tulos**

Mielipiteiden louhinta ei-englanninkielisestä lyhyestä tekstistä

**Esimerkki 2.794**

Arabit käyttävät päivittäisessä viestinnässään paikallisia murteita, joita on vaikea tunnistaa automaattisesti tavanomaisilla luokitusmenetelmillä. Murteiden tunnistamisen haasteellinen tehtävä vaikeutuu entisestään, kun kyseessä ovat samaan maakuntaan/alueeseen kuuluvat vähävaraiset murteet. Tässä artikkelissa aluksi analysoidaan tilastollisesti Algerian murteita, jotta voidaan määritellä niiden erityispiirteet, jotka liittyvät prosodiatietoihin, jotka erotetaan lausetasolla karkearakeisen konsonantti/vokaali-segmentoinnin jälkeen. Näiden analyysitulosten perusteella ehdotamme hierarkkista luokittelua puhutun arabian algerialaisen murteen tunnistamiseksi (HADID). Siinä hyödynnetään sitä tosiasiaa, että murteilla on luontainen ominaisuus jäsentyä luonnostaan hierarkkiseksi. HADID-menetelmässä sovelletaan ylhäältä alaspäin suuntautuvaa hierarkkista luokittelua, jossa käytetään syviä neuroverkkomenetelmiä (Deep Neural Networks, DNN) paikallisen luokittelijan rakentamiseksi jokaiselle hierarkkisen murrerakenteen vanhemmalle solmulle. Järjestelmämme on toteutettu ja arvioitu Algerian arabian murteiden korpuksella. Hierarkkinen murrerakenne johdetaan historiallisista ja kielitieteellisistä tiedoista. Tulokset paljastavat, että HADIDissa paras luokittelija on DNN verrattuna tukivektorikoneeseen. Lisäksi HADID-järjestelmämme antaa 63,5 prosentin parannuksen tarkkuudessa verrattuna perusluokitusjärjestelmään Flat. Lisäksi kokonaistulokset osoittavat, että prosodiaan perustuva HADID-luokituksemme soveltuu puhujasta riippumattomaan murteen tunnistamiseen, ja se vaatii alle kuuden sekunnin testilauseita. Sähköpostiosoitteet: sm.bougrine@lagh-univ.dz (Soumia Bougrine), hadda\_cherroun@mail.lagh-univ.dz (Hadda Cherroun), djelloul.ziadi@univ-rouen.fr (Djelloul Ziadi ) Preprint submitted to Elsevier March 30, 2017 ar X iv :1 70 3. 10 06 5v 1 [ cs .C L ] 2 9 M ar 2 01 7

**Tulos**

Hierarkkinen luokittelu puhutun arabiankielisen murteen tunnistamiseksi prosodian avulla: Algerian murteet

**Esimerkki 2.795**

Toistuvat neuroverkot (Recurrent Neural Networks, RNN) ja erityisesti niiden LSTM-muistilla (Long ShortTerm Memory) varustetut muunnelmat ovat herättäneet uutta kiinnostusta, koska niitä on sovellettu menestyksekkäästi monenlaisiin koneoppimisongelmiin, joihin liittyy peräkkäisiä tietoja. Vaikka LSTM-verkot tuottavat käytännössä poikkeuksellisia tuloksia, niiden suorituskyvyn lähde ja rajoitukset tunnetaan edelleen melko huonosti. Käyttämällä merkkitason kielimalleja tulkinnanvaraisena testialustana pyrimme kuromaan umpeen tätä aukkoa tarjoamalla kattavan analyysin niiden representaatioista, ennusteista ja virhetyypeistä. Kokeilumme paljastavat erityisesti sellaisten tulkittavien solujen olemassaolon, jotka pitävät kirjaa pitkän kantaman riippuvuuksista, kuten rivien pituuksista, lainausmerkeistä ja sulkeista. Lisäksi laaja analyysi äärellisen horisontin n-grammimalleilla viittaa siihen, että verkot löytävät ja hyödyntävät näitä riippuvuuksia aktiivisesti. Lopuksi esitämme yksityiskohtaisen virheanalyysin, jossa ehdotetaan jatkotutkimusalueita.

**Tulos**

Toistuvien verkkojen visualisointi ja ymmärtäminen

**Esimerkki 2.796**

Tarkastelemme Bayesin aktiivisen oppimisen ja koesuunnittelun ongelmaa, jossa tavoitteena on oppia jonkin tuntemattoman kohdemuuttujan arvo informatiivisten, meluisten testien sarjan avulla. Toisin kuin aiemmissa töissä, keskitymme haastavaan, mutta käytännössä merkitykselliseen tilanteeseen, jossa testitulokset voivat olla ehdollisesti riippuvaisia piilotetusta kohdemuuttujasta. Tällaisissa olettamuksissa yleiset heuristiikat, kuten ahnaasti suoritettavat testit, jotka maksimoivat kohteen epävarmuuden vähenemisen, toimivat usein huonosti. Tässä artikkelissa ehdotamme ECED:tä, uutta, laskennallisesti tehokasta aktiivista oppimisalgoritmia, ja todistamme vahvat teoreettiset takuut, jotka pätevät korreloituneilla, kohinaisilla testeillä. Sen sijaan, että ECED optimoisi suoraan ennustusvirheen, se valitsee jokaisessa vaiheessa testin, joka maksimoi voiton sijaistavoitteessa, jossa otetaan huomioon testien väliset riippuvuudet. Analyysimme perustuu informaatioteoreettiseen apufunktioon, jolla seurataan ECED:n edistymistä, ja käyttää mukautuvaa submodulaarisuutta lähes optimaalisen rajan saavuttamiseen. Osoitamme ECED:n vahvan empiirisen suorituskyvyn kahdessa ongelmatapauksessa, mukaan lukien Bayesin koesuunnittelutehtävä, jonka tarkoituksena on erottaa toisistaan taloudelliset teoriat siitä, miten ihmiset tekevät riskipäätöksiä, ja aktiivinen preferenssien oppimistehtävä pareittaisten vertailujen avulla.

**Tulos**

Lähes optimaalinen Bayesin aktiivinen oppiminen korreloituneilla ja kohinaisilla testeillä.

**Esimerkki 2.797**

Rekursiiviset neuroverkot (Recurrent Neural Networks, RNN) pystyvät oppimaan koodaamaan ja hyödyntämään aktivoitumishistoriaa mielivaltaisella aikaskaalalla. Käytännössä kuitenkin tiedetään, että nykyaikaisilla gradienttilaskeutumiseen perustuvilla koulutusmenetelmillä on vaikeuksia pitkän aikavälin riippuvuuksien oppimisessa. Tässä kuvataan uusi koulutusmenetelmä, jossa käytetään rinnakkaisia rinnakkaisia kloonattuja verkkoja, joilla kaikilla on samat painot, jotka on koulutettu eri ärsykevaiheessa ja jotka säilyttävät itsenäiset aktivaatiohistoriat. Koulutus etenee suorittamalla rekursiivisesti eräpäivityksiä rinnakkaisten kloonien yli, kun aktivointihistoriaa kasvatetaan asteittain. Tämä mahdollistaa ristiriitojen etenemisen hierarkkisesti lyhytaikaisista konteksteista kohti pidempiaikaisia konteksteja, kunnes ne on ratkaistu. Havainnollistamme rinnakkaisten kloonien menetelmää ja hierarkkista konfliktin etenemistä hahmotason syvällä RNN:llä, jonka tehtävänä on muistaa Herman Melvillen Moby Dick -teoksen kappale.

**Tulos**

Hierarkkinen konfliktien leviäminen: Jakson oppiminen toistuvassa syvässä neuroverkossa: Jakson oppiminen toistuvassa syvässä neuroverkossa

**Esimerkki 2.798**

Reservoir computing on uusi, tehokas ja joustava koneoppimistekniikka, joka on helppo toteuttaa laitteistossa. Aika-multipleksattua arkkitehtuuria käyttämällä laitteistovaraajatietokoneet ovat viime aikoina saavuttaneet digitaalisiin toteutuksiin verrattavan suorituskyvyn. Optoelektronisten järjestelmien avulla on saavutettu reaaliaikaisen tiedonkäsittelyn mahdollistavat toimintanopeudet. Tällä hetkellä suurin suorituskyvyn pullonkaula on lukukerros, jossa käytetään hidasta digitaalista jälkikäsittelyä. Olemme suunnitelleet analogisen lukulaitteen, joka soveltuu aikamultipleksoituihin optoelektronisiin säiliötietokoneisiin ja joka pystyy toimimaan reaaliajassa. Lukuohjelma on rakennettu ja testattu kokeellisesti vakiomuotoisella vertailutehtävällä. Sen suorituskyky on parempi kuin muilla kuin reservoir-menetelmillä, ja siinä on runsaasti parantamisen varaa. Näin ollen tässä työssä ratkaistaan yksi tärkeimmistä rajoituksista laitteistovarastotietokoneiden tulevan kehityksen kannalta.

**Tulos**

Optisten säiliötietokoneiden analoginen lukulaite

**Esimerkki 2.799**

Tässä kuvataan työtä, jonka tarkoituksena on oppia verbien alaluokkia morfologisesti rikkaassa kielessä käyttäen vain minimaalisia kielellisiä resursseja. Tavoitteenamme on oppia verbien alaluokkia quechua-kielelle, joka on morfologisesti rikas kieli, jonka resurssit ovat vähäiset, annotoimattomasta korpuksesta. Vertailemme tuloksia, jotka on saatu soveltamalla tätä lähestymistapaa noteeraamattomaan arabian kielen korpukseen, niihin tuloksiin, jotka on saatu käsittelemällä samaa tekstiä puupankkimuodossa. Alkuperäisen suunnitelman mukaan oli tarkoitus käyttää vain morfologista analysaattoria ja annotoimatonta korpusta, mutta kokeet osoittavat, että tämä lähestymistapa ei yksinään ole tehokas arabian verbien yhdistelmäpotentiaalin oppimisessa yleensä. Tämän tiedon hankkimiseen tarvittavien resurssien alaraja on jonkin verran korkeampi, ja se edellyttää ilmeisesti useimpien kielten osalta puheosien merkitsintälaitetta ja chunkeria ja arabian osalta morfologista disambiguateria.

**Tulos**

Resurssikevyt lähestymistapa verbien valenssien oppimiseen

**Esimerkki 2.800**

Sähköiset potilastiedot sisältävät tärkeitä kliinisiä tietoja potilaista. Näiden tietojen tehokkaalla ja tuloksellisella käytöllä voitaisiin täydentää tai jopa korvata manuaalinen potilaskertomusten tarkastelu keinona tutkia ja parantaa terveydenhuollon laatua ja turvallisuutta. Osa näistä kliinisistä tiedoista on kuitenkin vapaata tekstiä, ja ne on esikäsiteltävä ennen käyttöä automaattisissa järjestelmissä. Yleinen vapaan tekstin tietolähde ovat radiologian raportit, jotka radiologit yleensä sanelevat selittääkseen tulkintansa. Pyrimme osoittamaan tietokoneoppimisen avulla tietokonetomografiakuvantamisraporttien luokittelun binäärisiin tuloksiin eli murtuman suhteen positiivisiin ja negatiivisiin tuloksiin käyttämällä tavallista tekstiluokittelua ja aiheen mallintamiseen perustuvia luokittelijoita. Aihepiirien mallintaminen tuottaa raportteihin tulkittavia teemoja (aihepiirijakaumia), jotka ovat tiiviimpi esitys kuin yleisesti käytetty sanasäkki-esitys ja joita voidaan käsitellä raakatekstiä nopeammin myöhemmissä automatisoiduissa prosesseissa. Esittelemme uusia luokittelijoita, jotka perustuvat tähän raporttien aihepiirien mallintamiseen perustuvaan esitystapaan. Aggregate topic classifier (ATC) ja confidence-based topic classifier (CTC) käyttävät yhtä ainoaa aihetta, joka määritetään koulutustietokannasta eri mittareiden perusteella, raporttien luokitteluun testitietokannassa. Vaihtoehtoisesti samankaltaisuuteen perustuva aihealueluokittelija (STC) mittaa raporttien aihejakaumien samankaltaisuutta ennustetun luokan määrittämiseksi. Ehdotetut aiheen mallintamiseen perustuvat luokitinjärjestelmät ovat kilpailukykyisiä nykyisten tekstiluokittelutekniikoiden kanssa, ja ne tarjoavat tehokkaan ja tulkittavissa olevan esityksen.

**Tulos**

Aihepiirien mallintaminen kliinisten raporttien luokittelua varten

**Esimerkki 2.801**

Verkkotietojen louhinnasta on tullut tärkeä tutkimusalue, koska sitä voidaan soveltaa lukuisiin ongelmiin. Tässä artikkelissa esitellään NOESIS, avoimen lähdekoodin kehys verkostotiedon louhintaa varten, joka tarjoaa laajan kokoelman verkkoanalyysitekniikoita, mukaan lukien verkkojen rakenteellisten ominaisuuksien analysointi, yhteisöjen havaitsemismenetelmät, linkkien pisteytys ja linkkien ennustaminen, sekä verkkojen visualisointialgoritmeja. Siinä on myös täydellinen itsenäinen graafinen käyttöliittymä, joka helpottaa kaikkien näiden tekniikoiden käyttöä. NOESIS-kehys on suunniteltu käyttäen vankkoja oliosuuntautuneita suunnitteluperiaatteita ja jäsenneltyä rinnakkaisohjelmointia. Koska NOESIS on kevyt kirjasto, jolla on minimaalinen määrä ulkoisia riippuvuuksia ja jolla on salliva ohjelmistolisenssi, se voidaan sisällyttää muihin ohjelmistoprojekteihin. Se on julkaistu BSD-lisenssillä, ja se on saatavilla osoitteesta http://noesis.ikor.org.

**Tulos**

NOESIS-verkkopohjainen tutkimus-, simulointi- ja induktiojärjestelmä (Network-Oriented Exploration, Simulation and Induction System)

**Esimerkki 2.802**

Sanojen sulauttaminen, erityisesti sen viimeaikainen kehitys, lupaa termien välisen samankaltaisuuden kvantifioinnin. Ei ole kuitenkaan selvää, missä määrin tämä samankaltaisuusarvo voi olla aidosti merkityksellinen ja hyödyllinen myöhemmissä tehtävissä. Tutkimme, miten malleista saatu samankaltaisuuspistemäärä todella kertoo termien sukulaisuudesta. Tarkastelemme ja kvantifioimme ensin sanojen upotusmallien epävarmuustekijää samankaltaisuusarvon suhteen. Tämän tekijän perusteella otamme käyttöön eri ulottuvuuksilla yleisen kynnysarvon, joka suodattaa tehokkaasti hyvin samankaltaiset termit. Arviointimme neljällä tiedonhakukokoelmalla tukee lähestymistapamme tehokkuutta, sillä käyttöönotetun kynnysarvon tulokset ovat huomattavasti parempia kuin perustason tulokset, mutta ovat yhtä hyviä tai tilastollisesti erottamattomia optimaalisista tuloksista.

**Tulos**

Epävarmuus neuroverkon sanojen upottamisessa Samankaltaisuuden kynnysarvon tutkiminen

**Esimerkki 2.803**

Tässä tutkimuksessa selvitetään ongelmaa, joka liittyy suurten tekstikokonaisuuksien merkintöjen ennustamiseen, kun jokaiselle tekstille voidaan antaa useita merkintöjä. Ongelma saattaa vaikuttaa triviaalilta, kun merkintöjen määrä on pieni, ja se voidaan helposti ratkaista käyttämällä sarjaa yksi-vsall-luokittelijoita. Kun merkintöjen määrä kuitenkin kasvaa useisiin tuhansiin, parametriavaruudesta tulee erittäin suuri, eikä yksi vastaan kaikki -tekniikkaa ole enää mahdollista käyttää. Tässä ehdotamme mallia, joka perustuu korkeamman kertaluvun sanavektorin momenttien faktorointiin sekä etikettien ja sanojen välisiin ristikkäismomentteihin usean etiketin ennustamista varten. Mallimme tarjoaa taatut konvergenssirajat uutetuille parametreille. Lisäksi mallimme tarvitsee vain kolme läpikäyntiä koulutustietokannan läpi parametrien poimimiseksi, mikä johtaa erittäin skaalautuvaan algoritmiin, joka voi kouluttaa miljoonia asiakirjoja ja satoja tuhansia merkintöjä sisältäviä gigatavun kokoisia tietoja käyttäen nimellisesti yhtä prosessoria, jossa on 16 Gt RAM-muistia. Mallimme saavuttaa 10x-15x-luokan nopeuden suurissa tietokokonaisuuksissa ja tuottaa samalla kilpailukykyisen suorituskyvyn verrattuna olemassa oleviin vertailualgoritmeihin.

**Tulos**

Suuren mittakaavan etiketin ennustaminen harvoille tiedoille todennäköisillä takuilla

**Esimerkki 2.804**

Esitämme käytännöllisen ja tilastollisesti johdonmukaisen järjestelmän binääriluokittelijoiden aktiiviseen oppimiseen yleisillä häviöfunktioilla. Algoritmimme käyttää tärkeyspainotusta näytteenottovirheiden korjaamiseen, ja hallitsemalla varianssia pystymme antamaan tiukat etikettien monimutkaisuuden rajat oppimisprosessille. Passiivisesti merkityillä tiedoilla tehdyt kokeet osoittavat, että tämä lähestymistapa vähentää merkintäkompleksisuutta, jota tarvitaan hyvän ennustuskyvyn saavuttamiseksi monissa oppimisongelmissa.

**Tulos**

Tärkeyspainotettu aktiivinen oppiminen

**Esimerkki 2.805**

Tässä artikkelissa käsitellään mallien valintaa piilotetuille Markov-malleille (HMM). Yleistämme faktorisoidun asymptoottisen Bayesin päättelyn (FAB), joka on hiljattain kehitetty riippumattomien piilomuuttujien (eli sekamallien) mallinvalintaa varten, ajasta riippuville piilomuuttujille. Kuten FAB seosmalleissa, FAB HMM:lle johdetaan iteratiivisena alarajan maksimointialgoritmina faktoroidulle informaatiokriteerille (FIC). Se perii seosmallien FAB:stä useita toivottuja ominaisuuksia HMM:ien oppimiselle, kuten FIC:n asymptoottisen johdonmukaisuuden marginaalisen log-likelihoodin kanssa, kutistumisvaikutuksen piilotettujen tilojen valinnassa, FIC:n alemman rajan monotonisen kasvun iteratiivisen optimoinnin kautta. Lisäksi siinä ei ole viritettävää hyperparametria, joten sen mallinvalintaprosessi voidaan täysin automatisoida. Kokeelliset tulokset osoittavat, että FAB päihittää nykyaikaisen variationaalisen Bayesin HMM:n ja ei-parametrisen Bayesin HMM:n mallinvalintatarkkuuden ja laskennallisen tehokkuuden osalta.

**Tulos**

Faktorisoidut asymptoottiset Bayesin piilomarkov-mallit

**Esimerkki 2.806**

Lausuman voimakkuudella voi olla merkittävä vaikutus yleisöön. Esimerkiksi kansainvälisiä suhteita voi rasittaa se, miten yhden maan tiedotusvälineet kuvaavat tapahtumaa toisessa maassa, ja artikkelit voidaan hylätä, koska ne liioittelevat tai vähättelevät tuloksiaan. Siksi on tärkeää ymmärtää lausuman voimakkuuden vaikutukset. Ensimmäinen askel on pystyä erottamaan vahvat ja heikot lausumat toisistaan. Jopa tätä ongelmaa on kuitenkin tutkittu liian vähän, mikä johtuu osittain tietojen puutteesta. Koska vahvuus on luonnostaan suhteellista, väitteitä sisältävien tekstien tarkistukset ovat luonnollinen tietolähde vahvuuseroja koskeville tiedoille. Tässä artikkelissa esitellään akateemisen kirjallisuuden lausetason tarkistuksia sisältävä korpus. Kuvaamme myös näkemyksiä, joita olemme saaneet tätä tehtävää varten tehdyistä annotaatiopyrkimyksistä.

**Tulos**

A Corpus of Sentence-level Revisions in Academic Writing: Askel kohti lausumien vahvuuden ymmärtämistä viestinnässä

**Esimerkki 2.807**

Säännöstely on hyvin tutkittu ongelma neuroverkkojen yhteydessä. Sitä käytetään yleensä parantamaan yleistyksen suorituskykyä, kun syöttönäytteiden määrä on suhteellisen pieni tai ne ovat voimakkaasti kohinan saastuttamia. Parametrisen mallin regularisointi voidaan toteuttaa eri tavoin, joista osa on varhaisia pysäytyksiä (Morgan ja Bourlard, 1990), painojen hajoamista ja tuotoksen tasoittamista, joilla vältetään ylisovittaminen tarkasteltavan mallin harjoittelun aikana. Bayesiläisestä näkökulmasta monet regularisointitekniikat vastaavat tiettyjen ennakkojakaumien asettamista mallin parametreille (Krogh ja Hertz, 1991). Käyttämällä Bishopin approksimaatiota (Bishop, 1995) tavoitefunktiolle, kun parametrisen funktion tuloon lisätään rajoitetun tyyppistä kohinaa, johdamme Taylorin laajennuksen ylemmän kertaluvun termit ja analysoimme kohinan aiheuttamien regularisointitermien kertoimia. Tutkimme erityisesti sitä, miten kartoitusfunktion Hessianin rankaiseminen syötteen suhteen vaikuttaa yleistystehoon. Osoitamme myös, miten voimme kontrolloida itsenäisesti tätä kerrointa rankaisemalla eksplisiittisesti kartoitusfunktion Jacobiania korruptoituneilla syötteillä.

**Tulos**

Melun lisääminen regularisoidulla tavoitteella koulutetun mallin syötteeseen.

**Esimerkki 2.808**

Huolimatta kiinnostuksesta käyttää monikielistä tietoa sanojen upotusten oppimiseen eri tehtävissä, kirjallisuudesta puuttuu systemaattinen vertailu mahdollisista lähestymistavoista. Arvioimme laajasti neljää suosittua lähestymistapaa kieltenvälisten upotusten tuottamiseen, joista kukin vaatii erilaista valvontaa, neljällä typografisesti erilaisella kieliparilla. Arviointiasetelmamme kattaa neljä erilaista tehtävää, mukaan lukien yksikielisen ja kieltenvälisen samankaltaisuuden sisäinen arviointi sekä ekstrinsinen arviointi semanttisten ja syntaktisten sovellusten osalta. Osoitamme, että kalliita kieltenvälisiä tietoja vaativat mallit suoriutuvat lähes aina paremmin, mutta halvalla valvotut mallit osoittautuvat usein kilpailukykyisiksi tietyissä tehtävissä.

**Tulos**

Sanojen sulautusten monikieliset mallit: Empiirinen vertailu

**Esimerkki 2.809**

Syväoppimisesta on viime vuosina tullut suosittu menetelmä, jolla ratkaistaan ongelmia esineiden ja puheen tunnistamisesta robottien havaitsemiseen ja ihmisten sairauksien ennustamiseen. Tässä artikkelissa esitellään konvoluutiohermoverkkojen (convolutional neural networks, CNN) ja pinottujen autokoodereiden (stacked autoencoders, SAE) hybridiarkkitehtuuri, jonka avulla opitaan toimintojen sarja, joka epälineaarisesti muuttaa syötemuodon tai -jakauman kohdemuodoksi tai -jakaumaksi, jolla on sama tuki. Vaikka tällainen kehys voi olla hyödyllinen useissa eri ongelmissa, kuten robottireittien suunnittelussa, peräkkäisessä päätöksenteossa peleissä ja materiaalin käsittelyreittien tunnistamisessa haluttujen mikrorakenteiden saavuttamiseksi, tässä artikkelissa keskitytään nesteen muodonmuutosten hallintaan mikrofluidisessa kanavassa sijoittamalla tarkoituksellisesti peräkkäisiä pylväitä, millä on merkittävä vaikutus valmistukseen biolääketieteellisissä ja tekstiilisovelluksissa, joissa halutaan erittäin kohdennettuja muotoja. Ehdotamme arkkitehtuuria, joka ennustaa samanaikaisesti epämuodostumattoman ja halutun virtausmuodon välisellä epälineaarisella muunnosreitillä sijaitsevan välimuodon ja oppii sitten kausaalisen toiminnan - yksittäisen pilarin, joka johtaa virtauksen muodonmuutokseen - yksi kerrallaan. Vaiheittaisten muodonmuutosten oppiminen tarjoaa syvällisen näkemyksen virtauksen fysikaalisesta muodonmuutoksesta. Tulokset osoittavat, että nykyisessä kehyksessä mallimme pystyy ennustamaan pilarien sarjan, joka rekonstruoi virtauksen muodon, joka muistuttaa suuresti haluttua muotoa.

**Tulos**

Syvä toimintajaksojen oppiminen kausaalista muodonmuodostusta varten

**Esimerkki 2.810**

Värikielen tuottaminen on olennaisen tärkeää perustellun kielen tuottamisen kannalta. Värikuvauksilla on monia haastavia ominaisuuksia: ne voivat olla epämääräisiä, koostumuksellisesti monimutkaisia ja denotaatiorikkaita. Esittelemme tehokkaan lähestymistavan värikuvausten tuottamiseen käyttämällä rekurrenssia neuroverkkoja ja Fouriertransformoitua värikuvausta. Mallimme on parempi kuin aiemmat työt ehdollisen kielen mallinnustehtävässä, jossa käytetään suurta korpusta naturalistisia värikuvauksia. Lisäksi mallin tuotoksen tutkiminen paljastaa, että se pystyy tuottamaan tarkasti paitsi perusväritermit myös kuvaukset, joilla on ei-konveksisia denotaatioita ("vihertävä"), paljaita modifikaattoreita ("kirkas", "tylsä") ja kompositionaalisia fraaseja ("haalistunut teal"), joita ei ole nähty koulutuksessa.

**Tulos**

Kompositionaalisten värikuvausten luomisen opettelu

**Esimerkki 2.811**

Toimialariippumaton suunnittelu on yksi tekoälyn perustavanlaatuisista osa-alueista. Suunnittelutehtävän kuvaus koostuu maailman alkutilasta, tavoitteesta ja joukosta toimintoja maailman tilan muuttamiseksi. Tavoitteena on löytää toimintojen sarja eli suunnitelma, joka muuttaa maailman alkutilan tavoitetilaksi. Optimaalisessa suunnittelussa emme ole kiinnostuneita löytämään vain yhtä suunnitelmaa, vaan yhden halvimmista suunnitelmista. Nykyään merkittävä lähestymistapa optimaaliseen suunnitteluun on heuristinen tila-avaruushaku, jota ohjaavat sallitut heuristiset funktiot. On kehitetty lukuisia sallittuja heuristiikkoja, joilla kullakin on omat vahvuutensa ja heikkoutensa, ja on hyvin tiedossa, että optimaaliseen suunnitteluun ei yleisesti ottaen ole olemassa yhtä "parasta" heuristiikkaa. Näin ollen on vaikea kysymys, mikä heuristiikka valitaan tiettyä suunnittelutehtävää varten. Tämä vaikeus voidaan välttää yhdistämällä useita heuristiikkoja, mutta tämä edellyttää lukuisten heurististen estimaattien laskemista jokaisessa tilassa, ja kompromissi tähän käytetyn ajan ja eri heuristiikkojen yhdistettyjen etujen ansiosta säästetyn ajan välillä voi olla suuri. Esittelemme uuden menetelmän, joka vähentää optimaalisen suunnittelun kannalta hyväksyttävien heuristiikkojen yhdistämisestä aiheutuvia kustannuksia säilyttäen samalla sen edut. Idealisoidun hakuavaruusmallin avulla muotoilemme päätöksentekosäännön, jolla valitaan paras heuristiikka, joka lasketaan kussakin tilassa. Tämän jälkeen esitellään aktiivinen online-oppimismenetelmä, jolla opitaan luokittelija, jonka kohdekäsitteenä on kyseinen päätössääntö, ja käytetään opittua luokittelijaa päättämään, mikä heuristiikka lasketaan kussakin tilassa. Arvioimme tätä tekniikkaa empiirisesti ja osoitamme, että se on huomattavasti parempi kuin tavanomainen menetelmä, jossa yhdistetään useita heuristiikkoja niiden pistemäisen maksimin avulla.

**Tulos**

Online Speedup Learning optimaalista suunnittelua varten

**Esimerkki 2.812**

Tilastollisia käsikirjoituksia, tapahtumasekvenssien malleja, jotka mahdollistavat todennäköisyyspohjaisen päättelyn implisiittisistä tapahtumista asiakirjoista, on tutkittu pienessä mutta kasvavassa määrin. Nämä järjestelmät toimivat NLP-putken tuottamien strukturoitujen verbi-argumenttitapahtumien pohjalta. Vertailemme näitä järjestelmiä uusimpiin rekurrenssia neuroverkkomalleihin, jotka toimivat suoraan raakamerkkien perusteella lauseiden ennustamiseksi, ja havaitsimme, että jälkimmäiset ovat suunnilleen verrattavissa edellisiin, kun kyse on asiakirjoista puuttuvien tapahtumien ennustamisesta.

**Tulos**

Lausetason LSTM-kielimallin käyttö käsikirjoitusten päättelyyn

**Esimerkki 2.813**

Esittelemme uudenlaisen generatiivisten todennäköisyysmallien koulutusperiaatteen, joka on vaihtoehto maksimiluotettavuudelle. Ehdotettu GSN-kehys (Generative Stochastic Networks) yleistää Denoising Auto-Encoders (DAE) -menetelmää ja perustuu sellaisen Markov-ketjun siirtymäoperaattorin oppimiseen, jonka stationaarijakauma arvioi datajakauman. Siirtymäjakauma on ehdollinen jakauma, joka yleensä sisältää pienen siirron, joten sillä on vähemmän hallitsevia moodeja ja se on unimodaalinen pienten siirtojen rajalla. Tämä yksinkertaistaa oppimisongelmaa, jolloin se ei muistuta niinkään tiheyden estimointia vaan enemmänkin valvottua funktion approksimointia, jonka gradientit saadaan backprop-menetelmällä. Tässä esitetyt teoreemat tarjoavat todennäköisyysperusteisen tulkinnan äänenvaimennuksen automaattikoodereille ja yleistävät niitä; tämän kehyksen yhteydessä tarkasteltuna injektoidun kohinan avulla oppivat automaattikooderit ovat GSN:ien erikoistapaus ja ne voidaan tulkita generatiivisiksi malleiksi. Teoriat tarjoavat myös mielenkiintoisen perustelun riippuvuusverkoille ja yleistetyille pseudolikelihoodeille ja määrittelevät sopivan yhteisen jakauman ja näytteenottomekanismin, vaikka ehdolliset eivät olisikaan johdonmukaisia. GSN:iä voidaan käyttää puuttuvien syötteiden kanssa, ja niitä voidaan käyttää muuttujien osajoukkojen näytteenottoon loput muuttujat huomioon ottaen. Kokeet, joilla nämä teoreettiset tulokset validoidaan, tehdään sekä synteettisillä tietokokonaisuuksilla että kuvatietokokonaisuuksilla. Kokeissa käytetään erityistä arkkitehtuuria, joka jäljittelee Deep Boltzmann Machine Gibbs sampleria, mutta jonka avulla koulutus voidaan suorittaa backprop-menetelmällä rekursiivisen neuroverkon kautta, jonka sisälle syötetään kohinaa, ja ilman kerroksittaista esivalmennusta.

**Tulos**

GSN:t: Generatiiviset stokastiset verkot

**Esimerkki 2.814**

Tässä tutkielmassa esitellään uusi algoritmi ajoneuvojen reititysongelmaan nimeltä Enhanced Bees Algorithm. Se on mukautettu melko tuoreesta algoritmista, Bees-algoritmista, joka on kehitetty jatkuviin optimointiongelmiin. Osoitamme, että parannetun Bees-algoritmin tulokset ovat kilpailukykyisiä parhaiden ajoneuvojen reititysongelmaa varten saatavilla olevien metaheuristiikkojen kanssa - se pystyy saavuttamaan tuloksia, jotka ovat 0,5 %:n sisällä optimaalisesta ratkaisusta yleisesti käytetyllä testitapausten joukolla. Osoitamme, että algoritmi toimii hyvin suoritusaikana, sillä se tuottaa tuloksia, jotka ovat alle 2 % optimaalisesta ratkaisusta 60 sekunnissa, joten se soveltuu käytettäväksi reaalimaailman lähetysskenaarioissa. Lisäksi esitämme lyhyen historian kirjallisuudesta tunnetuista tuloksista sekä yksityiskohtaisen kuvauksen ajoneuvojen reititysongelman ratkaisemiseksi kehitetyistä perusmenetelmistä.

**Tulos**

Bees-algoritmi ajoneuvojen reititysongelmaa varten

**Esimerkki 2.815**

Analysoimme tässä artikkelissa satunnaista ominaisuuskarttaa, joka perustuu [1] esiteltyyn invarianssiteoriaan (I-teoria). Tarkemmin sanottuna ryhmäinvariantti signaalisignaali saadaan ryhmämuunnettujen satunnaisprojektioiden kumulatiivisten jakaumien avulla. Analyysimme yhdistää invarianttien ominaisuuksien oppimisen kernelimenetelmiin, sillä osoitamme, että tämä ominaisuuskartta määrittelee odotetun Haar-integraatiokernelin, joka on invariantti määritellylle ryhmätoiminnalle. Osoitamme, miten tämä epälineaarinen satunnainen ominaisuuskartta approksimoi tätä ryhmäinvarianttiytymätöntä ydintä tasaisesti N pisteen joukossa. Lisäksi osoitamme, että se määrittelee funktioavaruuden, joka on tiheä vastaavassa Invariantti Reproducing Kernel Hilbert Space -avaruudessa. Lopuksi kvantifioimme empiirisen riskin minimoinnin konvergenssin virhetasot sekä tällaista invariantti-edustusta signaalien luokitteluun käyttävän oppimisalgoritmin näytekompleksisuuden vähenemisen klassisessa valvotun oppimisen ympäristössä.

**Tulos**

Oppiminen ryhmäinvarianttien ominaisuuksien avulla: Kernel-näkökulma

**Esimerkki 2.816**

Ajallisten mallien, kuten dynaamisten Bayes-verkkojen, parametrit voidaan mallintaa Bayesin kontekstissa staattisina tai atemporaalisina muuttujina, jotka vaikuttavat siirtymätodennäköisyyksiin jokaisella aika-askeleella. Hiukkassuodattimet epäonnistuvat malleissa, jotka sisältävät tällaisia muuttujia, kun taas Gibbsin näytteenottoa parametrimuuttujista käyttävät menetelmät voivat aiheuttaa näytekohtaisen kustannuksen, joka kasvaa lineaarisesti havaintosarjan pituuden myötä. Storvik (2002) kehitti menetelmän tarkkojen riittävien tilastojen inkrementaaliseen laskentaan, joka joissakin tapauksissa vähentää otoskohtaiset kustannukset vakioksi. Tässä artikkelissa osoitamme yhteyden Storvikin suodattimen ja Kalmanin suodattimen välillä parametriavaruudessa ja asetamme yleisempiä ehtoja, joiden vallitessa Storvikin suodatin toimii. Laajennetun Kalman-suodattimen analogiaan nojautuen kehitämme ja analysoimme sekä teoreettisesti että kokeellisesti Taylorin approksimaation parametrien posteriorille, jonka avulla Storvikin menetelmää voidaan soveltaa laajempaan luokkaan malleja. Sekä synteettisillä esimerkeillä että todellisilla sovelluksilla tekemämme kokeet osoittavat parannusta nykyisiin menetelmiin verrattuna.

**Tulos**

Laajennettu parametrisuodatin

**Esimerkki 2.817**

Monia todellisia järjestelmiä on mallinnettu verkostokäsitteiden avulla, ja kirjoitetut tekstit ovat erityinen esimerkki tietoverkoista. Viime vuosina verkostomenetelmien käyttö kielen analysoinnissa on mahdollistanut useiden mielenkiintoisten havaintojen tekemisen, mukaan luettuna uusien mallien esittäminen perustavanlaatuisten universaalien mallien syntymisen selittämiseksi. Vaikka syntaktiset verkot, yksi yleisimmistä kirjoitettujen tekstien verkostomalleista, osoittavat sekä mittakaavattomia että pienimaailman ominaisuuksia, tällainen esitys ei pysty kuvaamaan muita tekstin piirteitä, kuten aiheiden tai aihepiirien järjestäytymistä. Tässä yhteydessä ehdotamme uutta verkkoesitystä, jonka päätarkoituksena on kaapata sanojen semanttiset suhteet yksinkertaisella tavalla. Tätä varten yhdistämme kaikki sanat, jotka esiintyvät samassa semanttisessa kontekstissa, joka määritellään kolmella tavalla. Osoitamme, että ehdotettu esitys suosii semanttisesti toisiinsa liittyvien sanojen yhteisöjen syntymistä, ja tätä ominaisuutta voidaan käyttää merkityksellisten aiheiden tunnistamiseen. Ehdotettua menetelmää aiheiden havaitsemiseksi sovellettiin valikoituihin Wikipedia-artikkeleihin. Olemme havainneet, että yleisesti ottaen menetelmämme ovat parempia kuin perinteiset bag-of-words -representaatiot, mikä viittaa siihen, että korkean tason tekstiedustus voi olla hyödyllinen tekstien semanttisten piirteiden tutkimisessa.

**Tulos**

Aihealueiden segmentointi yhteisön havaitsemisen avulla monimutkaisissa verkoissa

**Esimerkki 2.818**

Koneoppimiskilpailuissa, kuten ImageNet Large Scale Visual Recognition Challenge [RDS15] ja KDD Cup, kilpailijat voivat lähettää ratkaisukandidaatteja ja saada oraakkelilta (yleensä kilpailun järjestäjiltä) arvionsa tarkkuuden verrattuna perustotuusmerkintöihin. Yksi yleisimmin käytetyistä tarkkuusmittareista binäärisissä luokittelutehtävissä on Area Under the Receiver Operating Characteristics Curve (AUC). Tässä artikkelissa esitämme konseptitodistuksia siitä, miten arvausjoukon AUC:n tuntemusta voidaan käyttää kahdessa erityyppisessä hyökkäyksessä parantamaan kyseisten arvausten tarkkuutta. Toisaalta osoitamme myös erään AUC-hyökkäyksen vaikeasti lähestyttävyyden osoittamalla, että niiden n esimerkin binääristen merkintöjen määrä, joiden osalta ratkaisukandidaatti saa AUC-pistemäärän c, kasvaa eksponentiaalisesti n:ssä, kun jokainen c ∈ (0, 1).

**Tulos**

Hyödyntämällä oraakkelia, joka ilmoittaa AUC-pisteet koneoppimiskilpailuissa.

**Esimerkki 2.819**

Tutkimme tekniikoita, joiden avulla voidaan maksimoida diskurssitiedon tehokkuus tekijyyden määrittelyssä. Esittelemme uudenlaisen menetelmän, jolla diskurssiominaisuudet upotetaan konvolutiiviseen neuroverkkotekstiluokittimeen ja jolla saavutetaan huomattavasti parempi tulos kuin nykyisellä tekniikalla. Tutkimme empiirisesti useita featurisointimenetelmiä ymmärtääkseen olosuhteet, joissa diskurssipiirteet edistävät ei-triviaalia suorituskyvyn kasvua, ja analysoimme diskurssin upotuksia.

**Tulos**

Keskustelutietojen tehokas hyödyntäminen tekijyyden määrittelyssä∗ \*.

**Esimerkki 2.820**

Mieltymyksillä on tärkeä rooli jokapäiväisessä elämässämme. CP-verkot, tai lyhyesti CP-verkot, ovat graafisia malleja ehdollisten kvalitatiivisten mieltymysten esittämiseksi ceteris paribus -oletusten ("kaikki muut ovat samat") mukaisesti. Intuitiivisesta luonteestaan ja monipuolisesta esitystavastaan huolimatta dominanssitestaus CP-verkoilla on laskennallisesti monimutkaista, vaikka CP-verkot rajattaisiinkin binääriarvoisiin preferensseihin. Binäärisille CP-verkoille on olemassa toteuttamiskelpoisia algoritmeja, mutta nämä algoritmit ovat epätäydellisiä moniarvoisille CP-verkoille. Tässä artikkelissa tunnistamme moniarvoisten CP-verkkojen luokan, jota kutsumme enemmän tai vähemmän CP-verkoiksi ja jolla on sama laskennallinen monimutkaisuus kuin binäärisillä CP-verkoilla. More-or-less CP-verkoissa hyödynnetään attribuuttiarvojen monotonisuutta ja käytetään välejä sellaisten arvojen yhdistämiseen, jotka aiheuttavat samankaltaisia preferenssejä. Tämän jälkeen esitellään dominanssitestausta varten etsinnänohjaussääntö, joka karsii tehokkaasti hakuavaruuden säilyttäen samalla kattavuuden.

**Tulos**

Enemmän tai vähemmän CP-verkot

**Esimerkki 2.821**

Arvioimme koneoppimistekniikoita vihreän energian (tuuli-, aurinko- ja biomassa) ennustamiseen sääennusteiden perusteella. Sää muodostuu useista ominaisuuksista: lämpötila, pilvipeite, tuulen nopeus ja suunta, jotka ovat diskreettejä satunnaismuuttujia. Yksi tavoitteistamme on ennustaa sää aikaisempien säätietojen perusteella. Lisäksi olemme kiinnostuneita löytämään korrelaatioita (riippuvuussuhteita, jotta voimme vähentää tietokokonaisuuden ulottuvuutta) näiden muuttujien välillä, ennustamaan puuttuvia tietoja ennustamaan poikkeamia sääennusteissa (työnsuunnittelua varten vihreässä valvontakeskuksessa), löytämään klustereita tiedoista (jotka muodostuvat läheisesti toisiinsa liittyvistä muuttujista, esim. PCA:n avulla, jota voidaan käyttää turhien muuttujien poistamiseen), luokittelemaan, löytämään (ei-lineaarisia SVM:n avulla) regressiomalleja, kouluttamaan keinotekoisia neuroverkkoja aiempien tietojen perusteella, jotta niitä voidaan käyttää ennustamiseen tulevaisuudessa.

**Tulos**

Koneoppimistekniikoiden arviointi vihreän energian ennustamisessa

**Esimerkki 2.822**

Esittelemme uudenlaisen kehyksen multimodaalisten syväoppimismallien arvioimiseksi niiden kielen ymmärtämis- ja yleistämiskyvyn kannalta. Tässä lähestymistavassa keinotekoinen data luodaan automaattisesti kokeilijan määritysten mukaisesti. Datan sisältöä sekä harjoittelun että arvioinnin aikana voidaan hallita yksityiskohtaisesti, mikä mahdollistaa sellaisten tehtävien luomisen, jotka edellyttävät todellisia yleistämiskykyjä, erityisesti aiemmin esiteltyjen käsitteiden yhdistämistä uudella tavalla. Osoitamme menetelmämme potentiaalin arvioimalla erilaisia visuaalisten kysymysten vastausmalleja neljässä eri tehtävässä ja osoitamme, miten kehyksemme antaa meille yksityiskohtaisen näkemyksen niiden kyvyistä ja rajoituksista. Avoimella kehyksellämme toivomme voivamme edistää edistystä multimodaalisen kielen ymmärtämisen alalla.

**Tulos**

SHAPEWORLD: Uusi testimenetelmä multimodaalisen kielen ymmärtämistä varten.

**Esimerkki 2.823**

Hiukkassuodatin (PF) on yleisimmin käytetty Bayesin peräkkäinen estimointimenetelmä epälineaaristen dynaamisten järjestelmien piilotilojen määrittämiseksi. Se kärsii kuitenkin edelleen tietyistä ongelmista, kuten hiukkasten monimuotoisuuden menetyksestä, suuresta hiukkasmäärän tarpeesta ja tärkeystiheysfunktioiden kalliista valinnasta. Tässä asiakirjassa esitellään uusi PF nimeltä eksponentiaalinen luonnollinen hiukkassuodatin (xNPF) edellä mainittujen ongelmien ratkaisemiseksi. Tässä lähestymistavassa ehdotetaan tilan siirtymätodennäköisyyttä, jossa käytetään luonnollista gradienttioppimista, joka tasapainottaa etsintää ja hyödyntämistä kestävämmin. Ehdotetulla tiheysfunktiolla varustettu PF ei tarvitse suurta määrää hiukkasia ja se säilyttää hiukkasten monimuotoisuuden ajon aikana. Ehdotettua järjestelmää arvioidaan ajassa muuttuvan parametrin estimointiongelmassa HIV-viruksen immuunivasteen dynaamisessa mallissa. Tätä mallia käytetään osoittamaan xNPF:n suorituskyky verrattuna useisiin nykyaikaisiin hiukkassuodatinmalleihin, kuten Annealed PF, Bootstrap PF, iteratiivinen PF, ekvivalenttipainoinen PF ja älykäs PF. Tulokset osoittavat, että xNPF konvergoi paljon lähemmäs todellisia tavoitetiloja kuin muut menetelmät.

**Tulos**

Eksponentiaalinen luonnollinen hiukkassuodatin

**Esimerkki 2.824**

Esittelemme NMDR-mallin (nonparametric metadata dependent relational), joka on bayesiläinen ei-parametrinen stokastinen lohkomalli verkkotietoja varten. NMDR-mallin mukaan kuhunkin solmuun liittyvät entiteetit voivat kuulua eri tavoin rajattomaan joukkoon piileviä yhteisöjä. Opittujen regressiomallien avulla nämä jäsenyydet voivat riippua mielivaltaisista solmun metatiedoista ja ne voidaan ennustaa niistä. Kehitämme tehokkaita MCMC-algoritmeja NMDR-mallien oppimiseen osittain havaituista solmujen suhteista. Retrospektiiviset MCMC-menetelmät mahdollistavat sen, että näytteenottajamme voi työskennellä suoraan NMDR:n äärettömän tarranmurtoedustuksen kanssa, jolloin vältytään äärellisiltä typistyksiltä. Tuloksemme osoittavat hyödyllisten piilevien yhteisöjen hyödyntämisen todellisista sosiaalisista ja ekologisista verkostoista sekä metatietojen hyödyllisyyden linkkien ennustamistehtävissä.

**Tulos**

Ei-parametrinen metatiedoista riippuvainen relaatiomalli (Nonparametric Metadata Dependent Relational Model)

**Esimerkki 2.825**

Tarkastelemme verkko-oppimisen ongelmaa epäspesifisissä lineaarisissa stokastisissa monikätisissä bandit-ongelmissa. Uusimpien lineaaristen bandit-algoritmien, kuten Optimism in the Face of Uncertainty Linear bandit (OFUL) -algoritmin, katoamistakuu pätee olettaen, että käsien odotetut palkkiot ovat täysin lineaarisia niiden ominaisuuksien suhteen. On kuitenkin kiinnostavaa tutkia mahdollisten virheellisten spesifikaatioiden vaikutusta lineaarisissa bandit-malleissa, joissa odotetut palkkiot poikkeavat lineaarisesta aliavaruudesta, joka määräytyy aseiden ominaisuuksien perusteella. Vaikka OFUL:n on hiljattain osoitettu olevan kestävä suhteellisen pienille poikkeamille lineaarisuudesta, osoitamme, että minkä tahansa lineaarisen bandit-algoritmin, joka nauttii optimaalisesta katumussuorituskyvystä täydellisen lineaarisessa ympäristössä (esim. OFUL), on kärsittävä lineaarisesta katumuksesta lineaarisen mallin harvan additiivisen häiriön yhteydessä. Tämän negatiivisen tuloksen voittamiseksi määrittelemme luonnollisen luokan bandit-malleja, joille on ominaista ei-harvinaiset poikkeamat lineaarisuudesta. Väitämme, että OFUL-algoritmi voi epäonnistua epälineaarisen katumuksen saavuttamisessa jopa sellaisissa malleissa, joissa on ei-harvinaiset poikkeamat. Lopuksi kehitämme uudenlaisen bandit-algoritmin, joka koostuu lineaarisuuden hypoteesitestistä, jonka jälkeen päätetään käyttää joko OFUL-algoritmia tai UCB-algoritmia (Upper Confidence Bound). Täysin lineaarisille bandit-malleille algoritmi osoittaa todistettavasti OFUL:n suotuisan katumussuorituskyvyn, kun taas väärin määritellyille malleille, jotka täyttävät ei-harvinaisen poikkeaman ominaisuuden, algoritmi välttää lineaarisen katumuksen ilmiön ja turvautuu UCB:n epälineaariseen katumuksen skaalaukseen. Numeeriset kokeet synteettisellä datalla ja suositustiedoilla Yahoo! Learning to Rank Challenge -tietokannasta tukevat empiirisesti havaintojamme.

**Tulos**

Epätäsmälliset lineaariset rosvot

**Esimerkki 2.826**

Tarkastelemme ongelmaa, jossa on tarkoitus palauttaa tarkasti matriisi B, jonka koko on M × M ja joka edustaa todennäköisyysjakaumaa M tuloksen yli, kun käytettävissä on havaittu matriisi "lukumääriä", jotka on tuotettu ottamalla riippumattomat näytteet jakaumasta B. Miten taustalla olevan matriisin B rakenteellisia ominaisuuksia voidaan hyödyntää, jotta saadaan aikaan laskennallisesti tehokkaita ja tietoteoreettisesti optimaalisia rekonstruktio-algoritmeja? Milloin tarkka rekonstruktio voidaan toteuttaa harvan datan järjestelmässä? Tämä perusongelma on keskeisellä sijalla useissa kysymyksissä, joita eri yhteisöt parhaillaan pohtivat. Tällaisia ovat esimerkiksi yhteisöjen havaitseminen harvoissa satunnaisgrafiineissa, strukturoitujen mallien, kuten aihepiirimallien tai piilomarkovin mallien, oppiminen ja luonnollisen kielen käsittelyyhteisön pyrkimykset laskea "sanojen upotuksia". Monet tämän ongelman näkökohdat - sekä oppimisen että ominaisuuksien testaamisen/arvioinnin kannalta ja sekä algoritmisen että tietoteoreettisen puolen osalta - ovat edelleen avoimia. Tuloksemme koskevat tilannetta, jossa B:llä on tietty rank 2 -rakenne. Tätä asetelmaa varten ehdotamme tehokasta (ja käytännössä toteuttamiskelpoista) algoritmia, joka palauttaa tarkasti taustalla olevan M × M -matriisin käyttäen Θ(M)-näytteitä. Tämä tulos on helposti siirrettävissä Θ(M)-näytteenottoalgoritmeihin, joilla opitaan aihepiirimalleja, joissa on kaksi aihepiiriä M-kokoisten sanakirjojen yli, ja opitaan piilomarkov-malleja, joissa on kaksi piilotettua tilaa ja M-alkioihin tuetut havaintojakaumat. Nämä lineaariset näytteiden monimutkaisuudet ovat optimaalisia vakiokertoimia myöten erittäin vahvassa mielessä: jopa taustalla olevan matriisin perusominaisuuksien testaaminen (kuten onko matriisin arvo 1 vai 2) vaatii Ω(M) näytettä. Lisäksi esitämme vielä vahvemman alarajan, jossa sen erottaminen, onko havaintosarja poimittu M:n havaintojen tasaisesta jakaumasta vai onko se tuotettu HMM:llä, jossa on kaksi piilotettua tilaa, vaatii Ω(M) havaintoa. Tämä estää sublineaarisen otoksen hypoteesitestit perusominaisuuksille, kuten identiteetille tai yhdenmukaisuudelle, sekä sublineaarisen otoksen estimaattorit suureille, kuten HMM:ien entropianopeudelle. Sublineaarisen otoksen ominaisuuksien testaamisen mahdottomuus näissä asetuksissa on kiehtovaa ja korostaa merkittäviä eroja näiden strukturoitujen asetusten ja standardiasetusten välillä, joissa i.i.d. otokset otetaan strukturoimattomasta jakaumasta, jonka tukikoko on M . ∗MIT. Sähköposti: qqh@mit.edu. †Washingtonin yliopisto. Sähköposti: sham@cs.washington.edu ‡Stanfordin yliopisto. Sähköposti: kweihao@gmail.com §Stanford University. Sähköposti: valiant@stanford.edu. Gregoryn ja Weihaon panosta on tuettu NSF CAREER Award CCF-1351108 -palkinnolla ja Okawa-säätiön tutkimusapurahalla. ar X iv :1 60 2. 06 58 6v 1 [ cs .L G ] 2 1 Fe b 20 16

**Tulos**

Strukturoitujen todennäköisyysmatriisien palauttaminen

**Esimerkki 2.827**

Pyrimme valottamaan äskettäin käyttöön otetun neuraalisen konekäännösparadigman vahvuuksia ja heikkouksia. Tätä varten suoritamme monitahoisen arvioinnin, jossa vertailemme uusimpien neuraalisten konekäännösjärjestelmien ja fraasipohjaisten konekäännösjärjestelmien tuottamia tuotoksia yhdeksän kielisuuntaa varten useilla eri ulottuvuuksilla. Mittaamme erityisesti tuotosten samankaltaisuutta, niiden sujuvuutta ja uudelleenjärjestelyn määrää, lauseen pituuden vaikutusta ja suorituskykyä eri virheluokissa. Huomaamme, että neuraalisten konekäännösjärjestelmien tuottamat käännökset eroavat huomattavasti toisistaan, ovat sujuvampia ja tarkempia sanajärjestyksen osalta verrattuna fraasipohjaisten järjestelmien tuottamiin käännöksiin. Neuraaliset konekäännösjärjestelmät ovat myös tarkempia tuottamaan taivutusmuotoja, mutta ne suoriutuvat huonosti käännettäessä hyvin pitkiä lauseita.

**Tulos**

Monipuolinen arviointi neuraalisen ja lausepohjaisen konekääntämisen välillä 9 kielisuunnassa.

**Esimerkki 2.828**

Esittelemme merkkitason rekursiivisen neuroverkon, joka tuottaa relevanttia ja johdonmukaista tekstiä, kun sille annetaan aputietoa, kuten tunne tai aihe. Käyttämällä yksinkertaista syötteen monistamisstrategiaa säilytämme apusyötteen signaalin laajemmilla sekvenssiväleillä kuin mitä on mahdollista kouluttaa ajassa tapahtuvalla takaisinkulkeutumisella. Tärkeimmät tuloksemme keskittyvät suureen 1,5 miljoonan olutarvostelun korpukseen BeerAdvocatesta. Generatiivisessa tilassa verkkomme tuottaa käskystä arvosteluja, jotka on räätälöity tähtiluokituksen tai tuotekategorian mukaan. Generatiivinen malli voi toimia myös päinvastoin, jolloin luokittelu onnistuu yllättävän tarkasti. Käänteisen mallin suorituskyky tarjoaa suoraviivaisen tavan määrittää, mitä generatiivinen malli tietää ilman, että tukeudutaan liikaa subjektiiviseen analyysiin. Arvostelun perusteella malli voi määrittää tarkasti vastaavan arvosanan ja päätellä oluen luokan (IPA, Stout jne.). Hyödynnämme tätä kykyä ja seuraamme havaittua tunnetta ja luokkaan kuulumista, kun arvostelun jokaista merkkiä käsitellään. Kvantitatiiviset ja kvalitatiiviset empiiriset arvioinnit osoittavat, että malli ymmärtää merkityksen ja oppii tekstin epälineaarisen dynamiikan, kuten negaation vaikutuksen mielialaan, vaikka sillä ei ole a priori käsitystä sanoista. Koska malli toimii merkkitasolla, se pystyy käsittelemään kirjoitusvirheitä, slangia ja suuria sanastoja ilman, että siihen tarvitaan mitään nimenomaisesti tätä tarkoitusta varten suunniteltua konetta.

**Tulos**

MERKKITASON GENERATIIVISET TEKSTIMALLIT

**Esimerkki 2.829**

Tässä asiakirjassa esitetään yhteenveto viimeaikaisesta edistyksestä, jota olemme saavuttaneet fysioterapian tietokonenäköteknologian alalla helppokäyttöisten ja kohtuuhintaisten laitteiden avulla. Esittelemme ensin etäterveysvalmennusjärjestelmän, jonka olemme rakentaneet Microsoft Kinectin avulla. Koska Kinectin kaappaama liiketieto on kohinaista, tutkimme Kinectin tietotarkkuutta suhteessa korkean tarkkuuden liikkeenkaappausjärjestelmään. Ehdotamme myös datan jakaumaan perustuvaa poikkeavien tietojen poistoalgoritmia. Kinectin kaappaamasta meluisesta datasta kinemaattisen parametrin tuottamiseksi ehdotamme kinemaattista suodatusalgoritmia, joka perustuu unscented Kalman Filteriin ja ihmisen luurangon kinemaattiseen malliin. Ehdotetulla algoritmilla saadaan tasainen kinemaattinen parametri, jonka kohina on vähäisempi kuin Kinectin raa'an liikedatan perusteella tuotetun kinemaattisen parametrin.

**Tulos**

Etävalmennusjärjestelmä ja ihmisen liiketietojen analysointi fysioterapiaa varten Microsoft Kinectin avulla

**Esimerkki 2.830**

Monet koneoppimisessa esiintyvät kombinatoriset ongelmat voidaan pelkistää submodulaarisen funktion minimointiongelmaksi. Submodulaariset funktiot ovat koverien funktioiden luonnollinen diskreetti analogi, ja ne voidaan minimoida vahvasti polynomisessa ajassa. Valitettavasti nykyaikaiset algoritmit yleistä submodulaarista minimointia varten ovat vaikeasti ratkaistavissa suuremmissa ongelmissa. Tässä artikkelissa esittelemme uuden alaluokan submodulaarisia minimointiongelmia, joita kutsumme hajotettaviksi. Hajotettavat submodulaariset funktiot ovat sellaisia, jotka voidaan esittää modulaarisiin funktioihin sovellettujen koverien funktioiden summina. Kehitämme algoritmin, SLG, jolla voidaan tehokkaasti minimoida hajotettavia submodulaarisia funktioita, joissa on kymmeniä tuhansia muuttujia. Algoritmimme hyödyntää viimeaikaisia tuloksia tasoitetusta koverasta minimoinnista. Sovellamme SLG:tä synteettisiin vertailuarvoihin ja yhteiseen luokitus- ja segmentointitehtävään ja osoitamme, että se päihittää nykyaikaiset yleiskäyttöiset submodulaariset minimointialgoritmit useilla kertaluokilla.

**Tulos**

Hajotettavien submodulaaristen funktioiden tehokas minimointi

**Esimerkki 2.831**

Tässä työssä esitellään menetelmä, jolla voidaan virittää sekvenssipohjainen generatiivinen malli molekyylien de novo -suunnittelua varten, joka voi lisääntyneen episodisen todennäköisyyden avulla oppia tuottamaan rakenteita, joilla on tietyt halutut ominaisuudet. Osoitamme, miten tämä malli voi suorittaa erilaisia tehtäviä, kuten tuottaa analogeja kyselyrakenteelle ja tuottaa yhdisteitä, joiden ennustetaan olevan aktiivisia biologista kohdetta vastaan. Periaatteen osoittamiseksi malli koulutetaan ensin tuottamaan molekyylejä, jotka eivät sisällä rikkiä. Toisena esimerkkinä malli koulutetaan tuottamaan analogeja lääkkeelle Celecoxib, tekniikka, jota voitaisiin käyttää scaffold hoppingiin tai kirjaston laajentamiseen yhdestä molekyylistä lähtien. Kun malli viritetään tuottamaan yhdisteitä, joiden ennustetaan olevan aktiivisia dopamiinireseptoria D2 vastaan, malli tuottaa rakenteita, joista yli 95 prosenttia ennustetaan aktiivisiksi, mukaan lukien kokeellisesti vahvistetut aktiiviset aineet, joita ei ole sisällytetty generatiiviseen malliin eikä aktiivisuuden ennustemalliin.

**Tulos**

Molekyylien De-Novo-suunnittelu syvän vahvistusoppimisen avulla

**Esimerkki 2.832**

Semanttista jäsennystä koskevissa nykyisissä tutkimuksissa keskitytään pääasiassa toimialueen sisäiseen tilanteeseen. Muotoilemme monialaisen semanttisen jäsentelyn toimialueiden välisenä sopeutumisongelmana: koulutamme semanttisen jäsentäjän joillakin lähdealueilla ja sopeutamme sen sitten kohdealueelle. Koska loogiset muodot ovat hyvin erilaisia eri aloilla, tämä ongelma asettaa ainutlaatuisia ja kiehtovia haasteita. Muuntelemalla loogiset muodot luonnollisen kielen kanonisiksi lausumiksi pelkistämme semanttisen jäsennyksen parafraasiksi ja kehitämme tarkkaavaisen sekvenssistä sekvenssiin -parafraasimallin, joka on yleinen ja joustava mukautettavaksi eri aloille. Havaitsemme kaksi ongelmaa, pienet mikro- ja suuret makrovarianssit, jotka estävät esivalmisteltujen sanasulkeumien suoran käytön neuroverkoissa, ja ehdotamme standardointitekniikoita ratkaisuksi. Suositussa OVERNIGHT-tietokannassa, joka sisältää kahdeksan toimialuetta, osoitamme, että sekä toimialueiden välinen harjoittelu että standardoitu esivalmennettu sanojen upotus voivat parantaa tuloksia merkittävästi.

**Tulos**

Toimialarajat ylittävä semanttinen jäsennys parafraasien avulla

**Esimerkki 2.833**

Tässä artikkelissa esittelemme nmtpy:n, Theanoon perustuvan joustavan Python-työkalupaketin neuraalisen konekäännöksen ja muiden neuraalisten sekvenssistä sekvenssiin -arkkitehtuurien kouluttamiseen. nmtpy irrottaa verkon määrittelyn koulutus- ja päättelyapuohjelmista, mikä helpottaa uuden arkkitehtuurin lisäämistä ja vähentää kirjoitettavan koodin määrää. nmtpy:tä on käytetty LIUM:n parhaana pidetyissä esityksissä WMT:n Multimodaalisen konekääntämisen ja Uutiskääntämisen tehtäviin vuosina 2016 ja 2017. 1 KATSAUS nmtpy on refaktoroitu, laajennettu ja vain Python 3 -versio dl4mt-tutorial 1:stä, joka on Theanon (Theano Development Team, 2016) toteutus attenttisesta neuraalisesta konekääntämisestä (Neural Machine Translation, NMT) (Bahdanau et al., 2014). Maaliskuussa 2017 MIT-lisenssillä MIT-lisenssillä avoimen lähdekoodin2 saaneen nmtpy-projektin kehitys alkoi maaliskuussa 2016 pyrkimyksenä sovittaa dl4mt-tutorial multimodaalisiin käännösmalleihin. nmtpy:stä on nyt tullut tehokas työkalupakki, jossa uuden mallin lisääminen on yhtä yksinkertaista kuin johdannainen abstraktista perusluokasta, jolla täytetään joukko perusmetodeja ja (valinnaisesti) toteutetaan mukautettu data-iteraattori. Koulutus- ja päättelyapuohjelmat ovat mahdollisimman malli-agnostisia, joten niitä voidaan käyttää erilaisiin sekvenssigeneraattoriverkkoihin, kuten multimodaaliseen NMT:hen ja kuvatekstitykseen, muutamia mainitakseni. Tämä joustavuus ja tarjottujen arkkitehtuurien runsas määrä (jakso 3) erottaa nmtpy:n Nematusista (Sennrich et al., 2017), toisesta dl4mt-tutorialista johdetusta NMT-ohjelmistosta.

**Tulos**

NMTPY: JOUSTAVA TYÖKALUPAKKI EDISTYNEILLE NEURAALISILLE KONEKÄÄNNÖSJÄRJESTELMILLE.

**Esimerkki 2.834**

Tässä artikkelissa käsittelemme ongelmaa, joka liittyy suhteen q p estimointiin, jossa p on tiheysfunktio ja q on toinen tiheys tai yleisemmin mielivaltainen funktio. Tämän suhteen tuntemista tai approksimointia tarvitaan useissa päättelyyn ja integrointiin liittyvissä ongelmissa, erityisesti silloin, kun on tarpeen keskiarvoistaa funktio yhden todennäköisyysjakauman suhteen, kun otetaan näyte toisesta jakaumasta. Sitä kutsutaan usein tilastollisessa päättelyssä tärkeysotannaksi, ja se liittyy läheisesti myös siirtooppimisen kovariaattisiirtymäongelmaan sekä erilaisiin MCMC-menetelmiin. Se voi olla hyödyllinen myös silloin, kun halutaan erottaa avaruuden, vaikkapa moninaisuuden, taustalla oleva geometria sille määritellystä tiheysfunktiosta. Lähestymistapamme perustuu siihen, että q p:n estimointiongelma muotoillaan uudelleen käänteisongelmaksi kerneliä vastaavan integraalioperaattorin suhteen ja näin se pelkistetään integraaliyhtälöksi, jota kutsutaan ensimmäisen lajin Fredholmin ongelmaksi. Tämä muotoilu yhdistettynä regularisointitekniikoihin ja kernelimenetelmiin johtaa periaatteelliseen kerneliin perustuvaan kehykseen algoritmien rakentamiseksi ja niiden teoreettiseksi analysoimiseksi. Tuloksena syntynyt algoritmiperhe (FIRE, Fredholm Inverse Regularized Estimator) on joustava, yksinkertainen ja helppo toteuttaa. Esitämme yksityiskohtaisen teoreettisen analyysin, mukaan lukien Gaussin ytimen keskittymisrajat ja konvergenssinopeudet, kun kyseessä ovat R:ssä määritellyt tiheydet, kompaktit alueet R:ssä ja sileät d-ulotteiset euklidisen avaruuden alamannekiinit. Esitämme myös kokeellisia tuloksia, mukaan lukien sovelluksia luokitteluun ja puolivalvottuun oppimiseen kovariaattisiirtymän puitteissa, ja osoitamme joitakin rohkaisevia kokeellisia vertailuja. Näytämme myös, miten algoritmiemme parametrit voidaan valita täysin valvomattomasti.

**Tulos**

Käänteinen tiheys käänteisenä ongelmana: Fredholmin yhtälön lähestymistapa.

**Esimerkki 2.835**

Harkitaan tehokkaan joukkoistamisjärjestelmän suunnittelua M -aarista luokittelutehtävää varten. Joukkojen työntekijät suorittavat yksinkertaisia binäärisiä mikrotehtäviä, joiden tulokset yhdistetään lopputuloksen saamiseksi. Tarkastelemme uutta skenaariota, jossa työntekijöillä on hylkäämisvaihtoehto, jolloin he voivat jättää mikrotehtävät väliin, jos he eivät pysty tai halua vastata. Esimerkiksi puheen transkriptiossa työntekijät, jotka eivät tunne kieltä, eivät välttämättä pysty vastaamaan mikrotehtäviin, jotka keskittyvät fonologisiin ulottuvuuksiin, jotka eivät kuulu heidän kategoriseen käsitykseensä. Esittelemme aggregointimenetelmän, jossa käytetään painotettua enemmistöäänestyssääntöä, jossa jokaisen työntekijän vastaukselle annetaan optimoitu painoarvo, jotta joukon luokittelusuorituskyky maksimoitaisiin. Arvioimme järjestelmän suorituskykyä sekä eksaktissa että asymptoottisessa muodossa. Lisäksi tarkastelemme asetelmaa, jossa voi olla joukko ahneita työntekijöitä, jotka suorittavat mikrotehtäviä, vaikka eivät pystyisi suorittamaan niitä luotettavasti. Harkitsemme ahneiden työntekijöiden käsittelyyn unohdus- ja karkotusstrategiaa ja kehitämme algoritmin, joka vaihtaa adaptiivisesti näiden kahden välillä sen perusteella, kuinka suuri osuus ahneista työntekijöistä on arvioitu anonyymissä joukossa. Simulaatiotulokset osoittavat, että suorituskyky on parempi kuin perinteisessä enemmistöpäätöksessä.

**Tulos**

Usean kohteen luokittelu joukkoistamisen avulla hylkäämisvaihtoehdolla varustettuna

**Esimerkki 2.836**

Erilaiset reaalimaailman prosessit (verkoissa) tuottavat tietosarjoja, joiden monimutkaista ajallista dynamiikkaa on tutkittava. Erityisesti tapahtumien aikaleimat voivat sisältää tärkeää tietoa taustalla olevasta verkkodynamiikasta, jota ei muuten saada jatkuvista signaaleista tasaisesti poimituista aikasarjoista. Lisäksi useimmissa monimutkaisissa prosesseissa tapahtumasarjat ja tasaisesti näytteistetyt aikasarjatiedot voivat olla vuorovaikutuksessa keskenään, mikä tekee näiden kahden tietolähteen yhteisestä mallintamisesta välttämätöntä. Edellä mainittujen ongelmien ratkaisemiseksi hyödynnämme tässä asiakirjassa (ajallisten) pisteprosessien rikkaita puitteita tapahtumatietojen mallintamiseen ja sen intensiteettifunktion ajantasaiseen päivittämiseen synergisen kaksoiskaksikon Recurrent Neural Networks (RNN) avulla. Ehdotetussa arkkitehtuurissa intensiteettifunktiota moduloidaan synergisesti yhdellä RNN:llä, jonka syötteenä on asynkronisia tapahtumia, ja toisella RNN:llä, jonka syötteenä on aikasarja. Lisäksi mallin tulkittavuuden parantamiseksi otetaan käyttöön hermopisteprosessin huomiomekanismi. Koko malli, jossa on tapahtumatyypin ja aikaleiman ennusteiden lähtökerrokset, voidaan kouluttaa päästä päähän, ja se mahdollistaa mustan laatikon käsittelyn intensiteetin mallintamiseksi. Todistamme mallimme paremmuuden synteettisessä datassa ja kolmessa reaalimaailman vertailutietoaineistossa.

**Tulos**

Tapahtumasarjojen ja aikasarjojen yhteinen mallintaminen huomiokaksoisilla toistuvilla neuroverkoilla.

**Esimerkki 2.837**

Mahdollisuus tallentaa tietoja liiketoimintaprosessien suorittamisesta niin sanottuihin tapahtumalokeihin on johtanut prosessien suorittamisen analysointiin ja prosessimallin toimivuuden arviointiin tarkoitettujen työkalujen yleistymiseen. Nämä työkalut ovat kuitenkin usein hyvin jäykkiä käsitellessään tapahtumalokeja, jotka sisältävät epätäydellistä tietoa prosessin suorittamisesta. Vaikka kyky käsitellä epätäydellisiä tapahtumatietoja on yksi prosessinlouhinnan manifestissa mainituista haasteista, suorituksen jäljityksen vaatimustenmukaisuuden arviointi edellyttää silti, että suoritetaan täydellinen, alusta loppuun ulottuva jäljitys. Tässä asiakirjassa hyödynnetään abduktiota, jotta voidaan tarjota joustava mutta laskennallisesti tehokas kehys tapahtumalokin epätäydellisyyden eri muotojen käsittelyyn. Lisäksi siinä ehdotetaan klassisen vaatimustenmukaisuuden käsitteen jalostamista vahvaksi ja ehdolliseksi vaatimustenmukaisuudeksi, jotta epätäydelliset lokit voidaan ottaa huomioon. Lopuksi kokeellinen arviointi osoittaa esitetyn lähestymistavan toteutettavuuden.

**Tulos**

Puutteellisten tapahtumalokien vaatimustenmukaisuuden vähentäminen

**Esimerkki 2.838**

Universaali induktio on keskeinen kysymys AGI:ssä. Sen käytännön sovellettavuus voidaan saavuttaa valitsemalla vertailukone tai algoritmien esitys, joka on sovittu ympäristön kanssa. Tämän koneen pitäisi olla päivitettävissä, jotta seuraavat tehtävät voitaisiin ratkaista tehokkaammin. Tutkimme tätä ongelmaa kombinaatiologiikan esimerkillä, joka on hyvin yksinkertainen Turingin täydellinen referenssikone, joka mahdollistaa ohjelmien esitysten muuttamisen ottamalla käyttöön erilaisia primitiivisten kombinaattoreiden joukkoja. Geneettistä ohjelmointijärjestelmää käytetään etsimään kombinaattori-ilmaisuja, jotka on helppo hajottaa osailmaisuiksi, jotka yhdistetään uudelleen ristiinkytkennässä. Kokeemme osoittavat, että kehitetty järjestelmä voi ratkaista matalan kompleksisuuden induktio- tai ennustustehtäviä (paljon tehokkaammin kuin raa'an voiman avulla); hyödyllisiä kombinaattoreita voidaan paljastaa ja sisällyttää esitykseen, mikä yksinkertaistaa vaikeampia tehtäviä. Optimaaliset yhdistelmäjoukot riippuvat kuitenkin erityisestä tehtävästä, joten vertailukone olisi valittava adaptiivisesti koordinoidusti hakukoneen kanssa.

**Tulos**

Universaali induktio vaihtelevilla yhdistelmäjoukoilla

**Esimerkki 2.839**

Tarkastelemme ongelmaa, jossa saman dynaamisen ohjelmointiongelman muunnelma ratkaistaan toistuvasti peräkkäisissä kokeissa. Esimerkki tarkastelemastamme ongelmatyypistä on optimaalisen binäärihakupuun löytäminen. Jokaisen kokeilun alussa oppija valitsee todennäköisyyspohjaisesti puun, jonka sisäisissä solmuissa on n avainta ja lehdissä n+ 1 avainten välisiä aukkoja. Sen jälkeen sille ilmoitetaan avainten ja aukkojen frekvenssit ja se veloitetaan valitun puun keskimääräisillä hakukustannuksilla. Ongelma on online, koska taajuudet voivat muuttua kokeiden välillä. Tavoitteena on kehittää algoritmeja, joilla on sellainen ominaisuus, että niiden keskimääräiset kokonaishakukustannukset (tappio) kaikissa kokeissa ovat lähellä jälkikäteen valitun parhaan puun kokonaishäviötä kaikissa kokeissa. Haasteena on tietenkin se, että algoritmin on käsiteltävä eksponentiaalista määrää puita. Kehitämme menetelmän tällaisten ongelmien ratkaisemiseksi laajalle luokalle dynaamisen ohjelmoinnin algoritmeja. Kehyksemme avulla voimme laajentaa online-oppimisalgoritmeja, kuten Hedge [9] ja Component Hedge [15], huomattavasti laajempaan luokkaan kombinatorisia kohteita kuin aiemmin oli mahdollista.

**Tulos**

Dynaaminen online-ohjelmointi

**Esimerkki 2.840**

Esittelemme ensimmäisen globaalin rekursiivisen neuraalisen jäsennysmallin, jolla on optimaalisuustakuu dekoodauksen aikana. Globaalien ominaisuuksien tukemiseksi luovumme dynaamisista ohjelmista ja etsimme sen sijaan suoraan kaikkien mahdollisten alipuiden avaruudesta. Vaikka tämä avaruus on eksponentiaalisesti suuri lauseen pituuden suhteen, osoitamme, että on mahdollista oppia tehokas A\*-jäsennysohjelma. Täydennämme olemassa olevia jäsennysmalleja, joilla on informatiiviset rajat ulkopuolisille pisteille, globaalilla mallilla, jolla on väljät rajat mutta jonka tarvitsee mallintaa vain ei-lokaaleja ilmiöitä. Globaali malli koulutetaan uudella tavoitteella, joka kannustaa jäsentäjää etsimään sekä tehokkaasti että tarkasti. Lähestymistapaa sovelletaan CCG-jäsennykseen, ja se parantaa uusinta tarkkuutta 0,4 F1:llä. Jäsennin löytää optimaalisen jäsennyksen 99,9 %:lle pidättäytyneistä lauseista ja tutkii keskimäärin vain 190 alipuuta.

**Tulos**

Globaali neuraalinen CCG-jäsennys optimaalisuustakuilla varustettuna

**Esimerkki 2.841**

Ehdotamme mallia, jolla opitaan visuaalisesti perusteltuja sanojen upotuksia (vis-w2v) kuvaamaan visuaalisia käsityksiä semanttisesta sukulaisuudesta. Vaikka tekstin avulla koulutetut sanojen upotukset ovat olleet erittäin menestyksekkäitä, ne eivät pysty paljastamaan visuaaliseen maailmaan sisältyviä semanttisen yhteyden käsitteitä. Visuaalinen pohjustus voi esimerkiksi auttaa meitä ymmärtämään, että sellaiset käsitteet kuin syöminen ja tuijottaminen liittyvät toisiinsa, sillä kun ihmiset syövät jotain, heillä on tapana myös tuijottaa ruokaa. Syömisen ja tuijottamisen kaltaisten moninaisten suhteiden perusteleminen visuaalisesti on haastava tehtävä, vaikka visuaalisen näkemisen alalla on viime aikoina tapahtunut edistystä. Ymmärrämme, että sanojen visuaalinen perustaminen riippuu visuaalisen maailmamme semantiikasta eikä kirjaimellisista pikseleistä. Käytämme siis leikekuvista luotuja abstrakteja kohtauksia visuaalisen perustan tarjoamiseen. Huomaamme, että oppimamme upotukset kuvaavat hienojakoisia visuaalisesti perusteltuja käsityksiä semanttisesta yhteydestä. Näemme parannuksia pelkkään tekstiin perustuviin sanojen upotuksiin (word2vec) verrattuna kolmessa tehtävässä: common-sense assertion -luokittelussa, visuaalisessa parafraasoinnissa ja tekstipohjaisessa kuvahaussa. Koodimme ja tietokokonaisuutemme ovat saatavilla verkossa.

**Tulos**

Visual Word2Vec (vis-w2v): Visuaalisesti perusteltujen sanojen sulautusten oppiminen abstraktien maisemien avulla

**Esimerkki 2.842**

Dialogitekstien tunnistamisella on tärkeä rooli keskustelujen ymmärtämisessä. Sitä on sovellettu laajalti monilla aloilla, kuten vuoropuhelujärjestelmissä, automaattisessa konekääntämisessä ja automaattisessa puheentunnistuksessa, ja se on erityisen hyödyllinen järjestelmissä, joissa on ihmisen ja tietokoneen välisiä luonnollisen kielen vuoropuhelurajapintoja, kuten virtuaaliavustajat ja chatbotit. Ensimmäinen vaihe dialogitoiminnan tunnistamisessa on dialogitoiminnan rajan tunnistaminen lausumissa. Tässä artikkelissa keskitytään lausuman segmentointiin vietnaminkielisten lausumien dialogitoimintojen rajojen mukaan eli funktionaalisten segmenttien tunnistamiseen. Tutkimme huolellisesti funktionaalisten segmenttien tunnistamista kahdella lähestymistavalla: (1) koneoppimismenetelmällä, jossa käytetään maksimaalista entropiaa (ME) ja ehdollisia satunnaiskenttiä (CRF); (2) syväoppimismenetelmällä, jossa käytetään kaksisuuntaista pitkää lyhytkestoista muistia (LSTM), jossa on CRF-kerros (Bi-LSTM-CRF), kahdella erilaisella keskusteluaineistolla: (1) Facebook-viesteillä (Message-data); (2) puhelinkeskustelujen transkriptiolla (Phone-data). Tietojemme mukaan tämä on ensimmäinen työ, jossa sovelletaan syväoppimiseen perustuvaa lähestymistapaa dialogitekstien segmentointiin. Kuten tulokset osoittavat, syväoppimismenetelmä toimii huomattavasti paremmin kuin perinteiset koneoppimismenetelmät. Lisäksi tämä on myös ensimmäinen tutkimus, jossa käsitellään vietnaminkielisen dialogin tekojen ja toiminnallisten segmenttien tunnistamista. Avainsanat: dialogitoimintojen segmentointi, toiminnallinen segmentti, vietnaminkielinen keskustelu.

**Tulos**

Dialogitekstien segmentointi vietnaminkielisissä ihmisen ja ihmisen välisissä keskusteluteksteissä

**Esimerkki 2.843**

Ihmisillä on taipumus yleistää uusi substantiivi hierarkkisen taksonomian perustasolle - kognitiivisesti erottuvaan kategoriaan, kuten "koira" - ja yleistämisen aste riippuu esimerkkien määrästä ja tyypistä. Äskettäin on myös osoitettu, että esimerkkien esitysajankohdan muutoksella on vaikutusta, mikä on yllättäen kääntänyt aiemmin havaitun perustason yleistämisen mallin päinvastaiseksi. Tutkimme täsmällisiä mekanismeja, jotka voisivat johtaa tällaiseen käyttäytymiseen, laajentamalla sanojen oppimisen ja sanojen yleistämisen laskennallista mallia niin, että se integroi muistin ja tarkkaavaisuuden kognitiiviset prosessit. Tuloksemme osoittavat, että unohtamisen ja uutuuden huomioimisen vuorovaikutus sekä herkkyys sekä esimerkkien tyypille että merkkitiheydelle mahdollistavat sen, että malli voi toistaa empiiriset tulokset eri esitysajankohdista. Tuloksemme vahvistavat tarvetta sisällyttää sanojen oppimismalleihin yleisiä kognitiivisia prosesseja, jotta sanaston omaksumisessa havaittujen käyttäytymismallien kirjoa voitaisiin ymmärtää paremmin.

**Tulos**

Muistin ja tarkkaavaisuuden vuorovaikutus uuden sanan yleistämisessä: A Computational Investigation

**Esimerkki 2.844**

Nykyiset syväoppimismenetelmät ovat olleet erittäin menestyksekkäitä, kun on käytetty suurilla GPU-pohjaisilla tietokoneilla koulutettuja konvoluutiohermoverkkoja (CNN). Tämän lähestymistavan kolme rajoitusta ovat: 2) verkot konfiguroidaan manuaalisesti optimaalisten tulosten saavuttamiseksi ja 3) neuronimallin toteuttaminen on kallista sekä kustannuksiltaan että teholtaan. Tässä artikkelissa arvioimme syväoppimisen malleja käyttämällä kolmea eri laskentaarkkitehtuuria näiden ongelmien ratkaisemiseksi: kvanttilaskentaa monimutkaisten topologioiden kouluttamiseen, suurteholaskentaa (HPC) verkkotopologian automaattiseen määrittämiseen ja neuromorfista laskentaa pienitehoiseen laitteistototeutukseen. Käytämme kokeessamme MNIST-tietokokonaisuutta nykyisten kvanttitietokoneiden syöttökokorajoitusten vuoksi. Tuloksemme osoittavat, että näiden kolmen arkkitehtuurin käyttäminen yhdessä on mahdollista edellä mainittujen syväoppimisen rajoitusten poistamiseksi. Osoitamme, että kvanttitietokone voi löytää laadukkaita arvoja kerroksen sisäisten yhteyksien painoille kohtuullisessa ajassa verkon monimutkaisuuden kasvaessa, että korkean suorituskyvyn tietokone voi löytää optimaaliset kerrospohjaiset topologiat ja että neuromorfinen tietokone voi esittää muista arkkitehtuureista johdetun monimutkaisen topologian ja painotukset pienitehoisessa memristisessä laitteistossa. Huom: Tämän käsikirjoituksen on laatinut UT-Battelle, LLC Yhdysvaltain energiaministeriön kanssa tehdyn sopimuksen nro DE-AC05-00OR22725 nojalla. Hyväksymällä artikkelin julkaistavaksi Yhdysvaltain hallitus pidättää itsellään ja kustantaja tunnustaa, että Yhdysvaltain hallitus pidättää itsellään ei-yksinomaisen, maksullisen, peruuttamattoman, maailmanlaajuisen lisenssin julkaista tai jäljentää tämän käsikirjoituksen julkaistua muotoa tai antaa muiden tehdä niin Yhdysvaltain hallituksen tarkoituksiin. Energiaministeriö tarjoaa yleisön saataville nämä liittovaltion rahoittaman tutkimuksen tulokset DOE:n julkista saatavuutta koskevan suunnitelman mukaisesti (http://energy.gov/downloads/doe-public-access-plan). 1 ar X iv :1 70 3. 05 36 4v 1 [ cs .N E ] 1 5 M ar 2 01 7

**Tulos**

Tutkimus monimutkaisista syväoppimisverkoista suuritehoisilla, neuromorfisilla ja kvanttitietokoneilla.

**Esimerkki 2.845**

Konekielinen translitterointi on prosessi, jossa sanan kirjoitusasu muutetaan automaattisesti lähdekielestä kohdekieleen säilyttäen samalla ääntäminen. Järjestysjärjestyksestä järjestykseen -oppiminen on viime aikoina noussut esiin uutena paradigmana valvotussa oppimisessa. Tässä asiakirjassa on ehdotettu merkkipohjaista koodaaja-dekooderimallia, joka koostuu kahdesta toistuvasta neuroverkosta. Enkooderi on kaksisuuntainen rekursiivinen neuroverkko, joka koodaa merkkijonon kiinteän pituiseen vektoriesitykseen, ja dekooderi tuottaa kohdejonon käyttämällä huomiopohjaista rekursiivista neuroverkkoa. Kooderi, dekooderi ja huomiomekanismi koulutetaan yhdessä maksimoimaan tavoitesekvenssin ehdollinen todennäköisyys lähdesekvenssin perusteella. Kokeemme eri tietokokonaisuuksilla osoittavat, että ehdotetulla kooderin ja dekooderin mallilla voidaan saavuttaa huomattavasti parempi translitteroinnin laatu kuin perinteisillä tilastollisilla malleilla.

**Tulos**

Neuraalinen koneellinen translitterointi: Alustavat tulokset

**Esimerkki 2.846**

Kernel-approksimaatio satunnaistettujen ominaisuuskarttojen avulla on viime aikoina herättänyt paljon kiinnostusta. Tässä työssä havaitsemme, että aiemmat lähestymistavat polynomiaalisen ytimen approksimointiin luovat karttoja, jotka ovat rankkipuutteellisia, eivätkä näin ollen hyödynnä projisoidun ominaisuusavaruuden kapasiteettia tehokkaasti. Tämän haasteen ratkaisemiseksi ehdotamme kompakteja satunnaisia ominaisuuskarttoja (CRAFTMaps), joiden avulla polynomiytimiä voidaan approksimoida tiiviimmin ja tarkemmin. Todistamme CRAFTMapsin virherajat ja osoitamme sen paremman suorituskyvyn ytimen rekonstruoinnissa verrattuna aiempiin approksimaatiojärjestelmiin. Näytämme, miten strukturoituja satunnaismatriiseja voidaan käyttää CRAFTMapsien tehokkaaseen tuottamiseen, ja esittelemme CRAFTMapsien avulla yhden läpikäynnin algoritmin epälineaaristen moniluokkaisten luokittelijoiden oppimiseen. Esitämme kokeita useilla vakiomuotoisilla aineistoilla, joiden suorituskyky on kilpailukykyinen uusimpien tulosten kanssa.

**Tulos**

Kompaktit satunnaiset ominaisuuskartat

**Esimerkki 2.847**

Moniagenttinen polunetsintäongelma (MAPF) on viime aikoina saanut paljon huomiota. Se ei kuitenkaan kuvaa monien reaalimaailman alojen, kuten automaattisten varastojen, tärkeitä ominaisuuksia, sillä niissä agentit ovat jatkuvasti tekemisissä uusien tehtävien kanssa. Siksi tutkimme tässä artikkelissa MAPF-ongelman elinikäistä versiota, jota kutsutaan moniagenttiseksi nouto- ja toimitusongelmaksi (multiagent pickup and delivery, MAPD). MAPD-ongelmassa agenttien on huolehdittava toimitustehtävien virrasta verkkoympäristössä. Kullekin jakelutehtävälle on osoitettava yksi agentti. Tämän agentin on ensin siirryttävä tiettyyn noutopaikkaan ja sitten tiettyyn toimituspaikkaan välttäen törmäyksiä muiden agenttien kanssa. Esittelemme kaksi toisistaan riippumatonta MAPD-algoritmia, Token Passing (TP) ja Token Passing with Task Swaps (TPTS). Teoreettisesti osoitamme, että ne ratkaisevat kaikki hyvin muotoillut MAPD-tapaukset, jotka ovat realistinen MAPD-tapausten alaluokka. Kokeellisesti vertaamme niitä keskitettyyn MAPD-algoritmiin, jolla ei ole tätä takuuta, simuloidussa varastojärjestelmässä. TP voidaan helposti laajentaa täysin hajautetuksi MAPD-algoritmiksi, ja se on paras valinta, kun reaaliaikainen laskenta on ensisijainen huolenaihe, koska se on tehokas MAPD-tapauksissa, joissa on satoja agentteja ja tehtäviä. TPTS vaatii rajoitettua viestintää agenttien välillä, ja se tasapainoilee hyvin TP:n ja keskitetyn MAPD-algoritmin välillä.

**Tulos**

Elinikäinen moniagenttinen polunetsintä online-nouto- ja jakelutehtäviä varten∗\*

**Esimerkki 2.848**

Suosittelujärjestelmät ovat nousemassa tärkeäksi liiketoimintasovellukseksi, jolla on merkittäviä taloudellisia vaikutuksia. Tällä hetkellä suosittuja järjestelmiä ovat esimerkiksi Amazonin kirjasuositukset, Netflixin elokuvasuositukset ja Pandoran musiikkisuositukset. Tässä artikkelissa käsittelemme ongelmaa, joka liittyy suosittelujärjestelmien tietoihin liittyvien todennäköisyyksien estimointiin käyttämällä ei-parametrista kernel-sujutusta. Estimoinnissa tulkitsemme puuttuvat kohteet satunnaisesti sensuroiduiksi havainnoiksi ja saamme aikaan tehokkaita laskentajärjestelmiä, joissa hyödynnetään generoivien funktioiden kombinatorisia ominaisuuksia. Demonstraamme lähestymistapaamme useilla tapaustutkimuksilla, jotka koskevat todellista elokuvasuositustietoa. Tulokset ovat vertailukelpoisia uusimpien tekniikoiden kanssa, mutta samalla ne tarjoavat myös todennäköisyyteen perustuvia preferenssiestimaatteja, jotka eivät kuulu perinteisten suosittelujärjestelmien piiriin.

**Tulos**

Todennäköisyyksien arviointi suositusjärjestelmissä

**Esimerkki 2.849**

Sanaston kokonaisrakenteen ja sen dynamiikan tutkiminen tuli mahdolliseksi diakronisten tekstikorpusten, erityisesti Google Books Ngramin, luomisen ansiosta. Tässä artikkelissa käsitellään kysymystä ydinsanaston muutosnopeudesta ja siitä, missä määrin ydinsanat kattavat tekstit. Vertaillaan eri aikakausia kolmen viime vuosisadan ajalta ja kuutta tärkeintä Google Books Ngramissa esitettyä eurooppalaista kieltä. Tärkein tulos on ydinsanojen muutosnopeuden suuri vakaus, joka on analoginen Swadesh-luettelon vakauden kanssa.

**Tulos**

Kielen ydinsanaston dynamiikka

**Esimerkki 2.850**

Stokastisia gradienttikorotettuja päätöspuita käytetään laajalti monimuuttujaluokitus- ja regressiotehtävissä. Tässä artikkelissa esitellään nopeusoptimoitu ja välimuistiin tallentava toteutus monimuuttujaluokittelua varten nimellä FastBDT. FastBDT on kertaluokkaa nopeampi sovitus- ja sovellusvaiheessa verrattuna suosittuihin toteutuksiin ohjelmistokehyksissä, kuten TMVA, scikit-learn ja XGBoost. Suoritusajan optimointiin käytettyjä konsepteja ja suorituskykytutkimuksia käsitellään yksityiskohtaisesti tässä asiakirjassa. Keskeisiä ideoita ovat mm: Syöttötietojen tasataajuinen jako, jonka avulla kalliit liukulukuoperaatiot voidaan korvata kokonaislukuoperaatioilla ja samalla parantaa luokittelun laatua; välimuistiin tallentuva lineaarinen pääsy syöttötietoihin, toisin kuin tavanomaisissa toteutuksissa, joissa pääsy on satunnaista. FastBDT tarjoaa rajapinnat C/C++:lle, Pythonille ja TMVA:lle. Sitä käytetään laajalti korkean energian fysiikan alalla Belle II -kokeilussa.

**Tulos**

Nopeusoptimoitu ja välimuistiystävällinen toteutus stokastisille gradienttikorotetuille päätöspuille monimuuttujaluokittelua varten.

**Esimerkki 2.851**

Esittelemme LTLS:n, tekniikan moniluokkaiseen ja monimerkkiseen ennustamiseen, joka voi suorittaa koulutuksen ja päättelyn logaritmisessa ajassa ja tilassa. LTLS sulauttaa suuret luokitusongelmat yksinkertaisiksi strukturoiduiksi ennustamisongelmiksi ja perustuu tehokkaisiin dynaamisen ohjelmoinnin algoritmeihin päättelyssä. Koulutamme LTLS:ää stokastisella gradienttilaskeutumisella useilla moniluokkaisilla ja monimerkkisillä tietokokonaisuuksilla ja osoitamme, että pienestä muistijalanjäljestä huolimatta se on usein kilpailukykyinen nykyisten lähestymistapojen kanssa.

**Tulos**

Log-aika ja log-avaruus Extreme-luokittelu

**Esimerkki 2.852**

Tässä artikkelissa tutkimme ongelmaa, joka liittyy mobiililaitteen paikantamiseen sen sulautettujen antureiden lukemien perusteella koneoppimismenetelmiä hyödyntäen. Tarkastelemme todellista ympäristöä, keräämme suuren 3110 datapisteen tietokokonaisuuden ja tutkimme useiden koneoppimisalgoritmien suorituskykyä mobiililaitteen paikantamisessa. Olemme löytäneet algoritmeja, jotka antavat jopa 0,76 metrin keskivirheen, mikä on parempi kuin muiden kirjallisuudessa raportoitujen sisäilman paikannusjärjestelmien tarkkuus. Ehdotamme myös hybridi-instanssipohjaista lähestymistapaa, jonka avulla nopeus kasvaa kymmenkertaiseksi ilman tarkkuuden menetystä suorassa käyttöönotossa verrattuna tavallisiin instanssipohjaisiin menetelmiin, mikä mahdollistaa nopean ja tarkan paikannuksen. Lisäksi selvitämme, miten pienemmällä tiheydellä kerätyt pienemmät tietokokonaisuudet vaikuttavat paikannuksen tarkkuuteen, mikä on tärkeää reaalimaailman ympäristöissä käytettäväksi. Lopuksi osoitamme, että nämä lähestymistavat soveltuvat reaalimaailman käyttöön arvioimalla niiden suorituskykyä liikkeessä tapahtuvassa online-kokeessa.

**Tulos**

Koneoppiminen sisätilojen paikannukseen matkapuhelinpohjaisten antureiden avulla

**Esimerkki 2.853**

Likimääräisen jälkijakauman valinta on yksi variationaalisen päättelyn keskeisistä ongelmista. Useimmissa variationaalisen päättelyn sovelluksissa käytetään yksinkertaisia posterioristen approksimaatioiden perheitä tehokkaan päättelyn mahdollistamiseksi, ja niissä keskitytään keskikenttään tai muihin yksinkertaisiin strukturoituihin approksimaatioihin. Tällä rajoituksella on merkittävä vaikutus variaatiomenetelmillä tehtyjen päätelmien laatuun. Esittelemme uuden lähestymistavan joustavien, mielivaltaisen monimutkaisten ja skaalautuvien likimääräisten posteriorijakaumien määrittämiseksi. Approksimaatiomme ovat jakaumia, jotka on rakennettu normalisointivirran avulla, jolloin yksinkertainen alkutiheys muunnetaan monimutkaisemmaksi soveltamalla käänteismuunnosten sarjaa, kunnes haluttu monimutkaisuustaso on saavutettu. Käytämme tätä näkemystä normalisoivista virroista kehittääksemme äärellisten ja äärettömän pienten virtojen luokkia ja tarjotaksemme yhtenäisen näkemyksen lähestymistavoista rikkaiden posterioristen approksimaatioiden rakentamiseen. Osoitamme, että teoreettiset edut, joita saadaan, kun posteriorit vastaavat paremmin todellista posterioria, yhdistettynä amortisoitujen variaatiolähestymistapojen skaalautuvuuteen, parantavat selvästi variaatiolaskennan suorituskykyä ja sovellettavuutta.

**Tulos**

Variatiivinen päättely normalisoivilla virtauksilla

**Esimerkki 2.854**

Lauseet ovat luonnollisen kielen tärkeitä semanttisia yksiköitä. Lauseiden yleinen, distributiivinen esitys, joka pystyy vangitsemaan latentin semantiikan, on hyödyllinen monille jatkokehityssovelluksille. Havaitsemme lauseiden yksinkertaisen geometrian - tietyn lauseen sanarepresentaatiot (keskimäärin 10,23 sanaa kaikissa SemEval-tietoaineistoissa ja keskihajonta 4,84) sijaitsevat karkeasti ottaen matalan sijan aliavaruudessa (karkeasti ottaen sijalla 4). Tämän havainnon perusteella esitämme lauseen sen sanavektoreiden kattaman matalan sijan aliavaruuden avulla. Tällainen valvomaton esitys on empiirisesti validoitu semanttisilla tekstin samankaltaisuustehtävillä 19 eri tietokokonaisuudessa, joissa se päihittää kehittyneet neuroverkkomallit, mukaan lukien ohitetut ajatusvektorit, keskimäärin 15 prosentilla.

**Tulos**

Lauseiden esittäminen matalarakeisina aliavaruuksina

**Esimerkki 2.855**

Yksityinen joukkojen leikkaus (PSI) toteutetaan yleensä useiden käyttäjäparien välisten salauskierrosten sarjana, mutta tässä työssä PSI toteutetaan yksinkertaisemmin: jokainen joukko on salattava vain kerran, minkä jälkeen kunkin käyttäjäparin on suoritettava vain yksi tavallinen joukkojen vertailu. Tämä on tyypillisesti suuruusluokkia nopeampaa kuin tavallinen PSI, mutta vertailussa on jonkin verran "epätarkkuutta", joka voi kuitenkin olla siedettävää tai jopa toivottavaa. Tämä osoitetaan tapauksessa, jossa joukot koostuvat WordNetin avulla käsitellyistä englanninkielisistä sanoista. Sähköposti: 1054h34@gmail.com

**Tulos**

Nopea ja sumea yksityisten joukkojen risteytys

**Esimerkki 2.856**

Submodulaariset funktiot kuvaavat erilaisia diskreettejä ongelmia koneoppimisessa, signaalinkäsittelyssä ja tietokonenäössä. Submodulaaristen<lb>funktioiden minimointiin liittyy kuitenkin useita algoritmisia haasteita. Viimeaikaisessa työssä esiteltiin<lb>helppokäyttöinen, rinnakkaistettavissa oleva algoritmi sellaisten submodulaaristen funktioiden minimoimiseksi, jotka<lb>hajoavat "yksinkertaisten" submodulaaristen funktioiden summaksi. Empiirisesti tämä al-<lb>goritmi toimii erittäin hyvin, mutta teoreettista analyysia ei esitetty. Tässä<lb>paperissa osoitamme, että algoritmi konvergoi lineaarisesti, ja annamme ylä- ja<lb>alarajat konvergenssinopeudelle. Todistuksemme perustuu<lb>submodulaaristen polyedereiden geometriaan ja nojautuu spektraalisen graafiteorian tuloksiin.

**Tulos**

Hajotettavan submodulaarisen funktion minimoinnin konvergenssinopeudesta (On the Convergence Rate of Decomposable Submodular Function Minimization)

**Esimerkki 2.857**

Koraani on muslimien pyhä kirja. Se sisältää tietoa monilta aloilta. Usein ihmiset etsivät tiettyjä käsitteitä pyhästä Koraanista käsitteiden välisten suhteiden perusteella. Pyhän Koraanin ontologinen mallintaminen voi olla hyödyllistä tällaisessa tilanteessa. Tässä asiakirjassa olemme mallintaneet pyhän Koraanin luontoon liittyviä käsitteitä käyttäen OWL:ää (Web Ontology Language) / RDF:ää (Resource Description Framework). Menetelmämme sisältää pyhässä Koraanissa mainittujen luontoon liittyvien käsitteiden tunnistamisen ja näiden käsitteiden välisten suhteiden tunnistamisen. Nämä käsitteet ja suhteet esitetään OWL-ontologian luokkina/instansseina ja ominaisuuksina. Myöhemmin tulososiossa osoitetaan, että ontologisen mallin avulla SPARQL-kyselyillä voidaan hakea kiinnostavia jakeita ja käsitteitä. Näin ollen tämä mallinnus auttaa pyhän Koraanin semanttisessa haussa ja kyselyissä. Tässä työssä olemme käyttäneet Sahih Internationalin englanninkielistä Koraanin käännöstä, Protege OWL Editoria ja kyselyihin SPARQL:ää. Avainsanat - Koraanin ontologia; semanttinen Koraani; koraanitiedon esittäminen.

**Tulos**

Ontologisen mallintamisen soveltaminen Koraanin "luonto"-alueeseen

**Esimerkki 2.858**

Go on haastavampi peli kuin muut lautapelit, koska aseman tai siirron arviointifunktion rakentaminen on vaikeaa. Tässä artikkelissa tutkimme, voidaanko syviä konvoluutioverkkoja käyttää suoraan tämän tiedon esittämiseen ja oppimiseen. Koulutamme suuren 12-kerroksisen konvoluutiohermoverkon valvotun oppimisen avulla tietokannasta, joka sisältää ihmisten ammattilaispelejä. Verkko ennustaa oikein asiantuntijan siirron 55 prosentissa asemista, mikä vastaa 6 danin ihmispelaajan tarkkuutta. Kun koulutettua konvoluutioverkkoa käytettiin suoraan Go-pelien pelaamiseen ilman hakua, se voitti perinteisen hakuohjelman GnuGo:n 97 prosentissa peleistä, ja se vastasi huipputason Monte-Carlo-puuhakua, joka simuloi kahta miljoonaa asemaa siirtoa kohden.

**Tulos**

LIIKKEEN ARVIOINTI GO:SSA SYVIEN KONVOLUUTIO-NEURAALIVERKKOJEN AVULLA

**Esimerkki 2.859**

Klusterianalyysillä on tärkeä rooli monien tietoon perustuvien järjestelmien päätöksentekoprosessissa. Klusterointisovelluksia varten on olemassa lukuisia erilaisia lähestymistapoja, kuten heuristisia tekniikoita, todennäköisyysmalleja ja perinteisiä hierarkkisia algoritmeja. Tässä asiakirjassa ehdotetaan uutta heuristista lähestymistapaa, joka perustuu big bang-big crunch -algoritmiin klusterointiongelmia varten. Ehdotetussa menetelmässä ei ainoastaan hyödynnetä heuristista luonnetta tyypillisten klusterointialgoritmien, kuten k-meansin, lieventämiseksi, vaan se hyötyy myös muistiin perustuvasta järjestelmästä verrattuna sen kaltaisiin heuristisiin tekniikoihin. Lisäksi ehdotetun algoritmin suorituskykyä tutkitaan useiden vertailutestifunktioiden ja tunnettujen tietokokonaisuuksien perusteella. Kokeelliset tulokset osoittavat, että ehdotettu menetelmä on huomattavasti parempi kuin vastaavat algoritmit.

**Tulos**

Muistirikastettu Big Bang Big Crunch -optimointialgoritmi tietojen klusterointiin

**Esimerkki 2.860**

Herkkyysanalyysin tekeminen vaikutuskaavioille päätöspiirikehyksen avulla on erityisen kätevää, koska osittaisderivaatat jokaisen parametrin suhteen ovat helposti saatavilla [Bhattacharjya ja Shachter, 2007; 2008]. Tässä artikkelissa esitellään kolme epälineaarista herkkyysanalyysimenetelmää, jotka hyödyntävät tätä osittaisjohdannaistietoa ja jotka eivät näin ollen vaadi päätöstilanteen uudelleenarviointia useita kertoja. Näytämme erityisesti, miten voidaan tehokkaasti vertailla strategioita päätöksentekotilanteissa, suorittaa herkkyysanalyysi riskin välttämiselle ja laskea täydellisen suojauksen arvo [Seyller, 2008].

**Tulos**

Kolme uutta herkkyysanalyysimenetelmää vaikutuskaavioille

**Esimerkki 2.861**

Esittelemme kaksi toisiinsa liittyvää menetelmää MasterPrint-tunnisteiden, synteettisten sormenjälkien, luomiseksi, jotka sormenjälkien todentamisjärjestelmä tunnistaa moneksi eri ihmiseksi. Molemmat menetelmät aloitetaan kouluttamalla generatiivista vastakkaisverkkoa (Generative Adversarial Network, GAN) joukolla oikeita sormenjälkikuvia. Tämän jälkeen generaattoriverkkoa käytetään etsimään kuvia, jotka voidaan tunnistaa useiksi henkilöiksi. Ensimmäisessä menetelmässä käytetään evolutiivista optimointia latenttien muuttujien avaruudessa ja toisessa gradienttipohjaista hakua. Menetelmämme avulla pystytään suunnittelemaan MasterPrint, jota kaupallinen sormenjälkijärjestelmä vastaa 22 prosenttia kaikista käyttäjistä tiukassa turvallisuusympäristössä ja 75 prosenttia kaikista käyttäjistä löysemmällä turvallisuusympäristöllä.

**Tulos**

DeepMasterPrint: Sormenjälkien luominen esityshyökkäyksiä varten

**Esimerkki 2.862**

Tässä artikkelissa kehitetään ajatus jäsenyysfunktion osoittamisesta OWL-ontologian (Web Ontology Language) elementeille, jotta ontologiasta voidaan luoda sumeat säännöt. OWL-ontologian elementtien jäsenyysfunktioiden määrittäminen on jo osittain kuvattu, mutta se koski tapausta, jossa oli käytettävissä useita saman alan OWL-ontologioita ja ne yhdistettiin yhdeksi ontologiaksi. Tämän artikkelin tarkoituksena on esitellä tapa, jolla OWL-ontologian elementtien jäsenyysfunktiot määritetään siinä tapauksessa, että käytettävissä on vain yksi ontologia. OWL-ontologiasta generoidut sumeat säännöt ovat välttämättömiä SWES-tietokannan (Semantic Web Expert System) täydentämiseksi. SWES on asiantuntijajärjestelmä, joka pystyy poimimaan tietoa verkosta löytyvistä OWL-ontologioista ja toimii käyttäjän yleisenä asiantuntijana.

**Tulos**

Yhden OWL-ontologian elementtien jäsenyysfunktion määrittäminen

**Esimerkki 2.863**

Kirjoittajia kannustetaan lähettämään uusia artikkeleita INFORMS-lehtiin käyttämällä tyylitiedostomallia, joka sisältää lehden otsikon. Mallin käyttö ei kuitenkaan tarkoita, että artikkeli on hyväksytty julkaistavaksi kyseisessä lehdessä. INFORMS-lehden mallit on tarkoitettu yksinomaan INFORMS-lehteen lähettämistä varten, eikä niitä saa käyttää papereiden levittämiseen painettuna tai verkossa tai papereiden lähettämiseen muuhun julkaisuun.

**Tulos**

Dynaaminen lähes optimaalinen algoritmi lineaariseen online-ohjelmointiin

**Esimerkki 2.864**

Jotta agenteista ja roboteista tulisi entistä hyödyllisempiä, niiden on kyettävä oppimaan nopeasti muilta kuin teknisiltä käyttäjiltä. Tässä artikkelissa tutkitaan ongelmaa, joka liittyy ihmisopettajan välittämien käyttäytymismallien vuorovaikutteiseen oppimiseen positiivisen ja negatiivisen palautteen avulla. Suuressa osassa aiempaa tätä ongelmaa käsittelevää työtä on oletettu, että ihmiset antavat päätöksistä palautetta, joka on riippuvainen opetettavasta käyttäytymisestä ja riippumaton oppijan nykyisestä toimintatavasta. Esitämme empiirisiä tuloksia, jotka osoittavat, että tämä oletus on väärä - siihen, antavatko ihmisopettajat myönteistä vai kielteistä palautetta päätöksestä, vaikuttaa oppijan nykyinen politiikka. Väitämme, että politiikasta riippuvainen palaute on tavallista ja mahdollistaa lisäksi hyödyllisiä koulutusstrategioita, joista agenttien tulisi hyötyä. Tämän oivalluksen pohjalta esittelemme COACH-algoritmin (Convergent Actor-Critic by Humans), joka on algoritmi politiikasta riippuvan palautteen perusteella tapahtuvaan oppimiseen ja joka konvergoi paikalliseen optimiin. Lopuksi osoitamme, että COACH voi menestyksekkäästi oppia useita käyttäytymismalleja fyysisellä robotilla, vaikka kuvien ominaisuudet olisivat meluisia.

**Tulos**

Vuorovaikutteinen oppiminen politiikasta riippuvaisesta ihmispalautteesta

**Esimerkki 2.865**

Tässä artikkelissa tarkastellaan yleistä datan sovittamisongelmaa verkkojärjestelmässä, jossa monet laskentasolmut on yhdistetty toisiinsa suuntaamattomalla graafilla. Tällaisella ongelmalla on monia reaalimaailman sovelluksia, ja sitä on tutkittu laajasti kirjallisuudessa. Olemassa olevat ratkaisut kuitenkin joko tarvitsevat keskitetyn ohjaimen tiedon jakamiseen tai vaativat eri solmujen välistä aukkosynkronointia, mikä lisää käytännön toteutusten vaikeutta erityisesti hyvin suuressa ja heterogeenisessa järjestelmässä. Sen vastakohtana tässä artikkelissa käsittelemme verkon yli tapahtuvaa tietojen sovittamisongelmaa stokastisena ohjelmointiongelmana, johon liittyy monia rajoituksia. Suunnittelemme täysin hajautetun ja asynkronoidun stokastisen gradienttilaskeutumisalgoritmin (SGD-algoritmi) mukauttamalla hiljattain julkaistun artikkelin [18] tuloksia. Osoitamme, että algoritmimme voi saavuttaa globaalin optimaalisuuden ja konsensuksen asymptoottisesti vain paikallisilla laskutoimituksilla ja viestinnällä. Lisäksi annamme terävän alarajan konvergenssinopeudelle säännöllisen graafin tapauksessa. Tämä tulos vastaa intuitiota ja antaa ohjeita "hyvän" verkkotopologian suunnitteluun konvergenssin nopeuttamiseksi. Suunnitelmamme ansiot vahvistetaan myös kokeilla sekä synteettisillä että todellisilla tietokokonaisuuksilla.

**Tulos**

Täysin hajautettu ja asynkronoitu stokastinen gradientin laskeutuminen verkottuneille järjestelmille

**Esimerkki 2.866**

Data-teksti-järjestelmät ovat tehokkaita tuottamaan automaattisesti raportteja tiedoista, ja ne yksinkertaistavat siten monimutkaisten tietojen esittämistä. Sen sijaan, että datatekstijärjestelmät esittäisivät tietoja visualisointitekniikoiden avulla, ne käyttävät luonnollista (ihmis)kieltä, joka on yleisin tapa kommunikoida ihmisten välillä. Lisäksi data-teksti-järjestelmät voivat mukauttaa tulostussisältöä käyttäjien mieltymysten, taustan tai kiinnostuksen kohteiden mukaan, joten niiden kanssa on miellyttävää olla vuorovaikutuksessa. Sisällön valinta on tärkeä osa jokaista data-teksti-järjestelmää, koska se on moduuli, joka määrittää, mitä saatavilla olevista tiedoista käyttäjälle välitetään. Tässä katsauksessa esitellään aluksi datasta tekstiin -tuotantoalaa, kuvataan yleinen datasta tekstiin -järjestelmän arkkitehtuuri ja tarkastellaan sen jälkeen uusimpia sisällönvalintamenetelmiä. Lopuksi annetaan suosituksia lähestymistavan valintaa varten ja käsitellään tulevan tutkimuksen mahdollisuuksia.

**Tulos**

Sisällön valinta data-teksti-järjestelmissä: Katsaus

**Esimerkki 2.867**

Osittaisvalvottua oppimista, joka perustuu matalan tiheyden erotteluperiaatteeseen, kuten klusteri- ja moninaisuusoletuksiin, on tutkittu laajasti viime vuosikymmeninä. Tällaiset puolivalvotut oppimismenetelmät eivät kuitenkaan aina toimi<lb>hyvin, koska klusteri- ja moninaisuusoletuksia rikotaan. Tässä artikkelissa<lb>ehdotamme uudenlaista lähestymistapaa puolivalvottuun oppimiseen, joka ei vaadi<lb> tällaisia rajoittavia oletuksia. Keskeinen ideamme on yhdistää oppiminen positiivisesta<lb>ja negatiivisesta datasta (tavallinen valvottu oppiminen) ja oppiminen positiivisesta ja<lb>merkitsemättömästä datasta (PU-oppiminen), jälkimmäisen taataan pystyvän hyödyntämään merkitsemätöntä<lb>dataa ilman klusteri- ja moninaisuusoletuksia. Osoitamme teoreettisesti ja<lb>kokeellisesti lähestymistapamme hyödyllisyyden.

**Tulos**

Matalan tiheyden erotteluperiaatteen ohi: Uusi lähestymistapa puolivalvottuun oppimiseen

**Esimerkki 2.868**

Tässä artikkelissa ehdotamme neuroverkkomallia, jossa on uusi Sequential Attention -kerros, joka laajentaa pehmeää huomiokykyä antamalla painotuksia syötesarjan sanoille siten, että otetaan huomioon paitsi se, miten hyvin sana vastaa kyselyä, myös se, miten hyvin ympäröivät sanat vastaavat kyselyä. Arvioimme tätä lähestymistapaa luetun ymmärtämistehtävässä (Who did What ja CNN) ja osoitamme, että se parantaa huomattavasti Stanford Reader -ohjelman kaltaista vahvaa perustasoa. Tuloksena syntyvä malli on kilpailukykyinen nykyiseen tekniikkaan nähden.

**Tulos**

Peräkkäinen huomio

**Esimerkki 2.869**

Ehdotamme kahta uutta tekniikkaa - pullonkaulaominaisuuksien pinoamista ja pienimmän generointivirheen koulutuskriteeriä - syvään neuroverkkoon (DNN) perustuvan puhesynteesin suorituskyvyn parantamiseksi. Tekniikoilla ratkaistaan nykyisissä tyypillisissä DNN-pohjaisissa synteesikehyksissä esiintyvät ongelmat, jotka liittyvät kehyskohtaiseen riippumattomuuteen ja staattisten ja dynaamisten piirteiden välisen suhteen tuntemattomuuteen. Pullonkaulaominaisuuksien pinoaminen, joka on akustisesti informoitu kielellinen esitys, tarjoaa tehokkaan tavan sisällyttää yksityiskohtaisempi kielellinen konteksti syötteeseen. Pienimmän generointivirheen koulutuskriteerillä minimoidaan ulostulon kokonaisvirhe koko lausuman aikana sen sijaan, että minimoitaisiin virhe kehyskohtaisesti erikseen, ja näin otetaan huomioon staattisten ja dynaamisten piirteiden välinen vuorovaikutus. Näitä kahta tekniikkaa voidaan helposti yhdistää suorituskyvyn parantamiseksi entisestään. Esitämme sekä objektiivisia että subjektiivisia tuloksia, jotka osoittavat ehdotettujen tekniikoiden tehokkuuden. Subjektiiviset tulokset osoittavat, että näiden kahden tekniikan yhdistäminen johtaa huomattavasti luonnollisempaan synteettiseen puheeseen kuin perinteiset DNN- tai LSTM-järjestelmät (Long Short Short Memory).

**Tulos**

Trajektorimallinnuksen parantaminen DNN-pohjaista puhesynteesiä varten käyttämällä pinottuja pullonkaulaominaisuuksia ja minimaalisen generointivirheen harjoittelua.

**Esimerkki 2.870**

Syviä neuroverkkoja koulutettaessa oletetaan yleensä, että harjoitusesimerkit ovat tasaisen vaikeita oppia. Tai toisin sanoen oletetaan, että koulutusvirhe jakautuu tasaisesti harjoitusesimerkkien kesken. Näiden oletusten perusteella kutakin harjoitusesimerkkiä käytetään yhtä monta kertaa. Tämä oletus ei kuitenkaan välttämättä päde monissa tapauksissa. Hiljattain otettiin käyttöön "Oddball SGD" (novelty-driven stochastic gradient descent), jolla koulutusta ohjataan todennäköisyyteen perustuvalla tavalla virhejakauman mukaan - koulutustiheys on verrannollinen koulutusvirheen suuruuteen. Tässä artikkelissa osoitetaan videon koodaamiseen syvän neuroverkon avulla, että oddball SGD:tä voidaan käyttää tasaisen virheen aikaansaamiseksi koko harjoitusjoukossa.

**Tulos**

Yhtenäinen oppiminen syvässä neuroverkossa "Oddball" Stochastic Gradient Descent -menetelmän avulla

**Esimerkki 2.871**

Tässä projektissa on siirretty melko täydellinen todistusteoreettinen formalisointi Lambek-laskennasta (ei-assosiatiivinen ja mielivaltaiset laajennukset) Coq-todistusavustajasta HOL4-teoremankirjoittajaan, jossa on joitakin parannuksia ja uusia teoreemoja. Lambek-laskennan kolme päättelyjärjestelmää (syntaktinen laskenta, luonnollinen päättely ja sekvenssilaskenta) on määritelty ja monia niihin liittyviä teoreemoja on todistettu. Näiden järjestelmien välinen vastaavuus todistetaan muodollisesti. Lopuksi on suunniteltu ja toteutettu sekvenssilaskennan todisteiden formalisointi (jossa Coqilla on sisäänrakennetut tuet) HOL4:llä. Eräät perustulokset, mukaan lukien niin sanottujen "leikkausvapaiden" todisteiden aliformuloiden ominaisuudet, todistetaan muodollisesti. Tätä työtä voidaan pitää alustavana työnä kohti kategoriakielioppeihin perustuvaa kielen jäsentäjää, joka ei ole multimodaalinen, mutta jolla on silti kyky tukea kontekstisidonnaisia kieliä mukautettujen laajennusten avulla.

**Tulos**

Formalisoitu Lambek-lasku korkeamman asteen logiikassa (HOL4)

**Esimerkki 2.872**

Tässä artikkelissa käsitellään ongelmaa, joka koskee seuraavaksi todennäköisimmin esiintyvien k tapahtuman ennustamista historiallisista reaaliaikaisista tapahtumavirroista. Nykyisissä kausaaliennustekyselyihin liittyvissä lähestymistavoissa on useita rajoituksia. Ensinnäkin niissä etsitään tyhjentävästi asyklisen kausaaliverkon kautta todennäköisimmät k tapahtumaa, mutta todellisten tapahtumavirtojen tiedot heijastavat usein syklistä kausaalisuutta. Toiseksi ne sisältävät konservatiivisia oletuksia, joiden tarkoituksena on sulkea pois kaikki mahdolliset ei-kausaaliset linkit kausaaliverkosta; tämä johtaa siihen, että monet harvinaisemmat mutta tärkeät kausaalilinkit jätetään pois. Korjaamme nämä rajoitukset ehdottamalla uudenlaista tapahtumien etusijajärjestysmallia ja ajonaikaista kausaalisen päättelyn mekanismia. Tapahtumien etusijajärjestysmalli rakentaa ensimmäisen kertaluvun absorboivan Markovin ketjun inkrementaalisesti tapahtumavirtojen yli, jossa kahden tapahtuman välinen reuna merkitsee niiden välistä ajallista etusijajärjestyssuhdetta, joka on välttämätön edellytys kausaalisuudelle. Tämän jälkeen ajonaikainen kausaalinen päättelymekanismi oppii kausaalisuhteet dynaamisesti kyselyn käsittelyn aikana. Tämä tapahtuu poistamalla osa ajallisista etusijajärjestyssuhteista, jotka eivät ole kausaalisia muiden tapahtumien läsnä ollessa tapahtumien etusijajärjestysmallissa. Tässä asiakirjassa esitellään kaksi kyselyjen käsittelyalgoritmia - toinen suorittaa mallin tyhjentävän haun ja toinen tehokkaamman suppeamman haun, joka päättyy aikaisin. Kokeet, joissa käytetään kahta todellista tietokokonaisuutta (kaskadoituvat sähkökatkokset sähköjärjestelmissä ja verkkosivujen katselu), todentavat todennäköisyyteen perustuvien top-k-ennustuskyselyjen tehokkuuden ja algoritmien tehokkuuden. Vähennetty hakualgoritmi vähensi suoritusaikaa suhteessa tyhjentävään hakuun 25-80 prosenttia (sovelluksesta riippuen), mutta tarkkuus heikkeni vain vähän.

**Tulos**

Reaaliaikainen Top-K-ennustekyselyjen käsittely tapahtumavirtojen yli

**Esimerkki 2.873**

Kehitämme todennäköisyysperusteisen latenttimuuttujamallin, jonka avulla voimme löytää korpuksista semanttisia kehyksiä - tapahtumatyyppejä ja niiden osallistujia. Esitämme Dirichlet-multinomiaalisen mallin, jossa kehykset ovat latentteja kategorioita, jotka selittävät verbi-subjekti-objekti-kolmioiden yhdistämisen, kun otetaan huomioon dokumenttitason harvinaisuus. Analysoimme, mitä malli oppii, ja vertaamme sitä FrameNetiin, ja havaitsemme, että se oppii joitakin uusia ja mielenkiintoisia kehyksiä. Tämä asiakirja sisältää myös keskustelua päättelyyn liittyvistä kysymyksistä, kuten keskittymäparametrien oppimisesta, sekä pienimuotoisen virheanalyysin syntaktisen jäsennyksen tarkkuudesta. Huomautus: tämä työ julkaistiin alun perin verkossa lokakuussa 2012 osana CMU MLD:n Data Analysis Project -vaatimusta. Tässä versiossa ei ole uusia kokeita tai tuloksia, mutta siihen on lisätty jonkin verran keskustelua uusista asiaan liittyvistä töistä.

**Tulos**

Kehysten oppiminen tekstistä valvomattomalla latenttimuuttujamallilla

**Esimerkki 2.874**

Potilaille määrätään usein lääkkeitä, joiden tarkoituksena on parantaa potilaan terveydentilaa, mutta useimpien reseptilääkkeiden valitettava seuraus on ei-toivottujen sivuvaikutusten esiintyminen. Nykyiset lääkevalvontamenetelmät todennäköisesti ilmoittavat tehokkaasti haittavaikutuksista, joita esiintyy useammalla kuin yhdellä potilaalla tuhannesta, mutta harvinaisemmista haittavaikutuksista voi kestää vuosikymmeniä ennen kuin nämä samat menetelmät antavat merkkejä, jolloin lääkettä määräävien potilaiden sairastavuus tai kuolleisuus on vaarassa, kun harvinaista haittavaikutusta ei havaita. Tässä asiakirjassa ehdotamme harvinaisista haittavaikutuksista ilmoittamiseen uutta laskennallista meta-analyysikehystä, jossa yhdistyvät nykyiset menetelmät, verkosta saatu tieto, metrioppiminen ja puolivalvottu klusterointi. Uudenlainen kehys pystyi ilmaisemaan monia tunnettuja harvinaisia ja vakavia sivuvaikutuksia tutkittujen lääkkeiden osalta, kuten jänteen repeämä, kun lääkkeeksi määrättiin siprofloksasiinia tai levofloksasiinia, munuaisten vajaatoiminta naprokseenin yhteydessä ja masennus, joka liittyy Rimonabanttiin. Lisäksi se tuotti useimpien tutkittujen lääkkeiden osalta harvinaisia haittavaikutuksia koskevia signaaleja nykyisiä menetelmiä tiukemmalla signaalikynnyksellä, ja sillä on potentiaalia tulla olennaiseksi osaksi markkinoille saattamisen jälkeistä valvontaa harvinaisten haittavaikutusten havaitsemiseksi.

**Tulos**

Uusi puolivalvottu algoritmi harvinaisten lääkemääräysten sivuvaikutusten löytämiseen

**Esimerkki 2.875**

Osoitamme vahvan yhteyden frequentististen PAC-Bayesin rajojen ja Bayesin marginaalisen todennäköisyyden välillä. Toisin sanoen negatiivisen log-likelihoodin tappiofunktion osalta osoitamme, että PAC-Bayesin yleistysrajojen minimointi maksimoi Bayesin marginaalisen todennäköisyyden. Tämä tarjoaa vaihtoehtoisen selityksen Bayesin Occamin partaveitsen kriteereille olettaen, että data on i.i.d.-jakauman tuottamaa. Koska negatiivinen log-likelihood on rajoittamaton tappiofunktio, perustelemme ja ehdotamme PAC-Bayesin teoreemaa, joka on räätälöity sub-Gamma-häviöperheelle, ja osoitamme, että lähestymistapamme on järkevä klassisissa Bayesin lineaarisissa regressiotehtävissä.

**Tulos**

PAC-Bayesin teoria kohtaa Bayesin päättelyn.

**Esimerkki 2.876**

Koulutusteknologiassa ja oppimistieteissä on useita käyttötarkoituksia ennustemallille, joka kertoo, suoriutuuko oppilas tehtävästä oikein vai ei. Esimerkiksi älykäs opetusjärjestelmä voi käyttää tällaista mallia arvioidakseen, hallitseeko oppilas jonkin taidon vai ei. Analysoimme, mikä merkitys tietojen ajantasaisuudella on tällaisten ennusteiden tekemisessä, eli kysymme, onko opiskelijan suorituksesta tehdyillä suhteellisen tuoreilla havainnoilla enemmän merkitystä kuin suhteellisen vanhoilla havainnoilla. Kehitämme uuden Recent-Performance Factors Analysis -mallin, jossa otetaan huomioon tietojen uusiminen. Uusi malli parantaa merkittävästi ennustetarkkuutta sekä nykyisiin logistisen regression suorituskykymalleihin että uusiin perusmalleihin verrattuna arvioinneissa, jotka on tehty reaalimaailman ja synteettisillä tietokokonaisuuksilla. Toissijaisena panoksena osoitamme, miten laajalti käytetty ristiinvalidointi 0-1-menetyksen tappiolla on huonompi kuin AIC ja ristiinvalidointi L1-ennustusvirheiden tappiolla mallin suorituskyvyn mittarina.

**Tulos**

Suorituksen ennustaminen ohjauksen aikana viimeaikaisen suorituksen mallien avulla

**Esimerkki 2.877**

Ehdotamme valvottua koneoppimismenetelmää nykyisten signaalin ja kuvan talteenottomenetelmien tehostamiseksi ja osoitamme sen tehokkuuden esimerkkinä tietokonetomografiakuvan rekonstruktiosta. Tekniikkamme perustuu useiden kuvaestimaattien paikalliseen epälineaariseen fuusioon, jotka kaikki on saatu soveltamalla valittua rekonstruktio-algoritmia, jonka ohjausparametrien arvot vaihtelevat. Yleensä tällaisilla tuloskuvilla on erilainen bias/varianssin kompromissi. Kuvien fuusio suoritetaan feed-forward-neuraaliverkolla, joka on koulutettu joukon tunnettujen esimerkkien perusteella. Numeeriset kokeet osoittavat, että rekonstruktion laatu paranee verrattuna nykyisiin suoriin ja iteratiivisiin rekonstruktiomenetelmiin.

**Tulos**

Paikkatietoon mukautuva rekonstruktio tietokonetomografiassa neuroverkkoja käyttäen

**Esimerkki 2.878**

Esittelemme uuden tulkinnan kahdelle toisiinsa liittyvälle käsitteelle: ehdollinen hyöty ja hyötyjen riippumattomuus. Perinteisestä tulkinnasta poiketen uusi tulkinta tekee käsitteistä niiden probabilististen vastineiden suoria analogeja. Näiden käsitteiden kuvaamiseksi muodollisesti vetoamme edellisessä artikkelissa esiteltyyn hyötyjakauman käsitteeseen. Osoitamme, että hyötyjakaumat, joiden rakenne on identtinen todennäköisyysjakaumien rakenteen kanssa, voidaan nähdä additiivisten moniattribuuttihyötyfunktioiden erikoistapauksena, ja osoitamme, miten tämän erikoistapauksen avulla voimme vangita ehdollisen hyötyarvon ja hyötyriippumattomuuden uudet merkitykset. Lopuksi esitellään hyötyverkkojen käsite, joka tekee hyötyjen osalta saman kuin Bayesin verkot todennäköisyyksien osalta. Hyötyverkot hyödyntävät erityisesti ehdollisen hyödyn ja hyödyn riippumattomuuden uutta tulkintaa hyötyjakauman kompaktin esittämisen mahdollistamiseksi.

**Tulos**

Ehdollinen hyöty, hyötyjen riippumattomuus ja hyötyverkot.

**Esimerkki 2.879**

Esittelemme ensimmäiset vahvistusoppimisen differentiaalisesti yksityiset algoritmit, joita sovelletaan kiinteän politiikan arviointiin. Luomme kaksi lähestymistapaa eriytyneen yksityisyyden saavuttamiseksi, teemme teoreettisen analyysin näiden kahden algoritmin yksityisyydestä ja hyödyllisyydestä ja osoitamme lupaavia tuloksia yksinkertaisilla empiirisillä esimerkeillä.

**Tulos**

Eriytetysti yksityisen politiikan arviointi∗

**Esimerkki 2.880**

Laitteistovikojen suuri todennäköisyys estää monien kehittyneiden robottien (esim. jalkaisten robottien) luotettavan käyttöönoton todellisissa tilanteissa (esim. katastrofin jälkeinen pelastustoiminta). Sen sijaan, että robotit yrittäisivät diagnosoida vika(t), ne voisivat sopeutua kokeilemalla ja erehtymällä, jotta ne pystyisivät suorittamaan tehtävänsä. Parhaat robotiikan koe-error-algoritmit ovat kuitenkin kaikki episodisia: jokaisen kokeilun välillä robotti on palautettava samaan tilaan, eli robotti ei opi itsenäisesti. Tässä artikkelissa esittelemme uudenlaisen oppimisalgoritmin nimeltä "Reset-free Trial-and-Error" (RTE), jonka avulla robotit voivat toipua vaurioista suorittaessaan tehtäviään. Arvioimme sitä kuusijalkaisella robotilla, joka on vaurioitunut usealla eri tavalla (esim. puuttuva jalka, lyhennetty jalka jne.) ja jonka tavoitteena on saavuttaa sarja tavoitteita areenalla. Kokeemme osoittavat, että robotti pystyy palauttamaan suurimman osan liikkumiskyvystään muutamassa minuutissa ympäristössä, jossa on esteitä, ja ilman ihmisen puuttumista asiaan. Kaiken kaikkiaan tämä uusi algoritmi antaa mahdollisuuden harkita robottien lähettämistä paikkoihin, jotka ovat todella liian vaarallisia ihmisille ja joissa robotteja ei voida pelastaa.

**Tulos**

Resetoimaton koe-ja-virhe-oppiminen datatehokasta robotin vaurioiden korjausta varten

**Esimerkki 2.881**

Koalition muodostaminen on keskeinen aihe moniagenttijärjestelmissä. Koalitioiden avulla agentit pystyvät saavuttamaan tavoitteita, joita ne eivät ehkä pystyisi saavuttamaan yksin. Aikaisempi työ on osoittanut koalitiopelien ongelmien olevan laskennallisesti vaikeita. Wooldridge ja Dunne (Artificial Intelligence 2006) tutkivat useiden luonnollisten päätöksenteko-ongelmien klassista laskennallista monimutkaisuutta koalitionaalisissa resurssipeleissä (Coalitional Resource Games, CRG), joissa jokaisella agentilla on joukko resursseja ja joissa koalitiot voivat saada aikaan joukon tavoitteita, jos niillä on kollektiivisesti tarvittava määrä resursseja. Koalitionaalisten resurssipelien syöttö niputtaa yhteen useita elementtejä, esim. agenttijoukon Ag, tavoitejoukon G, resurssijoukon R jne. Shrot, Aumann ja Kraus (AAMAS 2009) tarkastelevat koalitionmuodostusongelmia CRG-mallissa käyttäen parametrisoidun kompleksisuuden teoriaa. Heidän hienostunut analyysinsä osoittaa, että kaikki syötteen osat eivät toimi yhtä lailla jotkin ongelman tapaukset ovat todellakin käsiteltävissä, kun taas toiset ovat edelleen vaikeasti ratkaistavissa. Vastaamme Shrotin, Aumannin ja Krausin avoimeksi jättämään tärkeään kysymykseen osoittamalla, että SC-ongelma (koalition onnistumisen tarkistaminen) on W[1]-vaikea, kun se parametrisoidaan koalition koon mukaan. Tämän jälkeen pystymme osoittamaan SC:stä yhden ainoan reduktioteeman avulla, että erilaiset Wooldridgen et al. esittelemät resursseihin, resurssirajoihin ja resurssikonflikteihin liittyvät ongelmat ovat 1. W[1]-kovia tai co-W[1]-kovia, kun ne on parametrisoitu koalition koon mukaan. 2. para-NP-vaikea tai co-para-NP-vaikea, kun parametrina on |R|. 3. FPT, kun parametrina on joko |G| tai |Ag|+ |R|. ∗Tuki osittain Googlen tiedekunnan tutkimuspalkinnolla, ONR Young Investigator Awardilla ja NSF CAREER -palkinnolla. †Department of Computer Science , University of Maryland at College Park, USA, email: rchitnis@cs.umd.edu ‡Department of Computer Science , University of Maryland at College Park, USA. email: hajiagha@cs.umd.edu §Department of Computer Science , University of Maryland at College Park, USA. email: vliaghat@cs.umd.edu.

**Tulos**

Koalitionaalisten resurssipelien ongelmien parametrisoitu monimutkaisuus

**Esimerkki 2.882**

Sanallinen ruumiinavaus on kirjaus haastattelusta, joka koskee vahvistamattoman kuolemantapauksen olosuhteita. Kehitysmaissa, jos kuolema sattuu muualla kuin terveydenhuoltolaitoksissa, kenttätyöntekijä haastattelee vainajan sukulaista kuoleman olosuhteista; tämä suullinen ruumiinavaus voidaan tarkistaa muualla. Raportoimme vertailevasta tutkimuksesta, joka koskee kuolemansyyn luokitteluun sovellettavan tekstiluokittelun prosesseja: ominaisuuksien arvojen esittämistä, koneoppimisen luokittelualgoritmeja ja ominaisuuksien vähentämisstrategioita, jotta voidaan tunnistaa sopivimmat lähestymistavat, joita voidaan soveltaa suullisen ruumiinavauksen tekstin luokitteluun. Osoitamme, että normalisoidulla termifrekvenssillä ja tavanomaisella TFiDF-menetelmällä saavutetaan vertailukelpoinen suorituskyky useilla luokittelijoilla. Tulokset osoittavat myös, että tukivektorikone on parempi kuin muut tässä tutkimuksessa käytetyt luokittelualgoritmit. Lopuksi osoitamme, että paikallisesti puolivalvotun ominaisuuksien vähentämisstrategian käyttö on tehokasta suorituskyvyn tarkkuuden lisäämiseksi.

**Tulos**

Vertaileva tutkimus koneoppimismenetelmistä verbaalisen ruumiinavaustekstin luokittelua varten.

**Esimerkki 2.883**

Tarkastelemme pääkomponenttianalyysia saastuneille aineistoille, jotka ovat erittäin moniulotteisia, jolloin kunkin havainnon moniulotteisuus on verrattavissa havaintojen lukumäärään tai jopa suurempi kuin havaintojen lukumäärä. Ehdotamme determinististä korkea-ulotteista robustia PCA-algoritmia, joka perii kaikki satunnaistetun vastineensa teoreettiset ominaisuudet, eli se on helposti lähestyttävä, robusti saastuneille pisteille, helposti kerneloitavissa, asymptoottisesti johdonmukainen ja saavuttaa maksimaalisen robustisuuden - 50 %:n hajoamispisteen. Vielä tärkeämpää on, että ehdotetulla menetelmällä on huomattavasti parempi laskennallinen tehokkuus, minkä vuoksi se soveltuu laajamittaisiin reaalisovelluksiin.

**Tulos**

Vankka PCA korkea-ulotteisissa aineistoissa: Deterministinen lähestymistapa

**Esimerkki 2.884**

Taustaa: Keuhkosyöpä tunnettiin ensisijaisena syöpänä, ja syövän eloonjäämisaste on noin 15 prosenttia. Keuhkosyövän varhainen toteaminen on johtava tekijä eloonjäämisasteessa. Kaikki keuhkosyövän oireet (piirteet) ilmenevät vasta, kun syöpä leviää muille alueille. Keuhkosyövän tarkka varhainen havaitseminen on välttämätöntä, jotta eloonjäämisprosentti kasvaisi. Tarkkaa havaitsemista varten on määriteltävä tehokkaat piirteet ja poistettava tarpeettomat piirteet kaikkien piirteiden joukosta.Piirteiden valinta on ongelma, jossa valitaan informatiiviset piirteet kaikkien piirteiden joukosta. Materiaalit ja menetelmät: Keuhkosyöpätietokanta koostuu 32 potilastiedostosta, joissa on 57 ominaisuutta. Tämä tietokanta on kerätty Hongin ja Youngin toimesta ja indeksoitu Kalifornian yliopiston Irvinen arkistoon. Kokeellinen sisältö sisältää kliinisistä tiedoista ja röntgentiedoista jne. poimitut tiedot. Tiedot kuvaavat 3 erilaista patologista keuhkosyöpää, ja kaikki piirteet ottavat kokonaislukuarvon 0-3. Tutkimuksessamme ehdotetaan uutta menetelmää keuhkosyövän tehokkaiden piirteiden tunnistamiseksi. Se perustuu hyperheuristiseen menetelmään. Tulokset: Saavutimme 80,63 %:n tarkkuuden käyttämällä supistettua 11 piirteen joukkoa.Ehdotettua menetelmää verrataan viiden koneoppimisen piirteiden valinnan tarkkuuteen.Näiden viiden menetelmän tarkkuudet ovat 60,94, 57,81, 68,75, 60,94 ja 68,75. Johtopäätökset: Ehdotetun menetelmän suorituskyky on parempi ja tarkkuus on korkein. Siksi ehdotettua mallia suositellaan tehokkaan taudin oireen tunnistamiseen. Nämä havainnot ovat erittäin tärkeitä terveystutkimuksessa, erityisesti lääketieteellisten resurssien kohdentamisessa potilaille, jotka on ennustettu korkean riskin potilaiksi.

**Tulos**

Hyper-heuristinen algoritmi tehokkaiden ominaisuuksien löytämiseksi keuhkosyöpäsairauden diagnosoimiseksi

**Esimerkki 2.885**

Tarkastelemme kysymystä evoluutioalgoritmien stabiilisuudesta kohdekäsitteen asteittaisille muutoksille, <lb>tai ajelehtimiselle. Määrittelemme algoritmin olevan kestävä ajautumiselle, jos<lb>jollakin käänteisellä polynomiaalisella ajautumisnopeudella kohdefunktiossa se konvergoi tarkkuuteen 1 - ǫ<lb>polynomiaalisilla resursseilla ja pysyy sitten tuon tarkkuuden sisällä loputtomiin, paitsi<lb>todennäköisyydellä ǫ milloinkin. Osoitamme, että jokainen evoluutioalgoritmi<lb>Valiantin [19] merkityksessä voidaan muuntaa Feldmanin [9] Correlational Query -tekniikan avulla<lb> tällaiseksi ajelehtimisen kestäväksi algoritmiksi. Tietyille evoluutioalgoritmeille, kuten Boolen<lb>yhdisteille, annamme rajat sille, kuinka paljon ne kestävät ajautumista. Kehitämme joitakin<lb>uusia evoluutioalgoritmeja, jotka kestävät merkittävää ajautumista. Erityisesti annamme<lb>algoritmin lineaaristen erottimien evoluutiolle pallosymmetrisen jakauman yli, joka on<lb>kestävä O(ǫ/n):n ajelehtimisnopeudelle, ja toisen algoritmin yleisemmille tuotos<lb>normaalijakaumille, joka kestää pienemmän ajelehtimisnopeuden. Ylläoleva käänteistulos voidaan tulkita myös tulokseksi, joka koskee itse<lb>evolvoituvuuden käsitteen kestävyyttä määritelmän muutoksissa. Toisena tämänsuuntaisena tuloksena osoitamme<lb>että jokainen evoluutioalgoritmi voidaan muuntaa kvasimonotoniseksi algoritmiksi, joka voi kehittyä<lb> mistä tahansa lähtöpisteestä ilman, että suorituskyky koskaan laskee merkittävästi<lb>lähtöpisteen suorituskyvyn alle. Tämä mahdollistaa sen, että useista tunnetuista robustisuuskäännöksistä voidaan poistaa suorituskyvyn mielivaltaisen<lb>alenemisen hieman luonnoton piirre.

**Tulos**

Kehitys ajelehtivien kohteiden kanssa

**Esimerkki 2.886**

Kahden otoksen testien tavoitteena on arvioida, onko kaksi otosta, SP ∼ P ja SQ ∼ Q, poimittu samasta jakaumasta. Ehkäpä mielenkiintoista on, että yksi suhteellisen tutkimaton menetelmä kahden otoksen testien rakentamiseksi on binääriluokittelijoiden käyttö. Rakennetaan tietokokonaisuus yhdistämällä SP:n n esimerkkiä positiiviseen merkintään ja SQ:n m esimerkkiä negatiiviseen merkintään. Jos nollahypoteesi "P = Q" on tosi, binääriluokittimen luokittelutarkkuuden tämän tietokokonaisuuden pidättämällä osajoukolla pitäisi pysyä lähellä sattuman tasoa. Kuten osoitamme, tällaiset luokittimen kahden otoksen testit (C2ST) oppivat lennossa sopivan esityksen datasta, palauttavat testitilastot tulkittavissa olevissa yksiköissä, niillä on yksinkertainen nollajakauma, ja niiden ennustava epävarmuus mahdollistaa tulkinnan siitä, missä P ja Q eroavat toisistaan. Tämän artikkelin tavoitteena on selvittää C2ST:n ominaisuudet, suorituskyky ja käyttötarkoitukset. Ensin analysoidaan niiden tärkeimmät teoreettiset ominaisuudet. Toiseksi vertaamme niiden suorituskykyä useisiin uusimpiin vaihtoehtoihin. Kolmanneksi ehdotamme niiden käyttöä sellaisten generatiivisten mallien otoslaadun arviointiin, joiden todennäköisyys on vaikeasti määriteltävissä, kuten generatiiviset vastakkaisverkot (Generative Adversarial Networks, GAN). Neljänneksi esitellään GAN:ien ja C2ST:n uutta sovellusta kausaalisyyksien löytämiseen.

**Tulos**

LUOKITTELIJAN KAHDEN OTOKSEN TESTIEN TARKISTAMINEN

**Esimerkki 2.887**

Yhä useammat tahot, erityisesti ihmiset, voivat hankkia, jakaa ja käsitellä tietoja. Tämän ilmiön hajautettu luonne on vaikuttanut monien joukkoistamishankkeiden kehittymiseen. Tämä skenaario on yleinen useimmissa asiantuntijoiden ja ei-asiantuntijoiden ryhmämielipide- ja arviointitehtävissä (mukaan lukien monet internet- tai verkkokäyttäjien käyttäytymisen muodot), joissa keskeisenä elementtinä on useista lähteistä peräisin olevien havaintojen ja mielipiteiden yhdistäminen.

**Tulos**

Joukkoistamiseen osallistujien arviointi, kun ei ole olemassa pohjatietoa

**Esimerkki 2.888**

Tässä artikkelissa esitellään uusi tekniikka (osittain) sidottujen kyselyjen optimoimiseksi disjunktiivisten Datalog-ohjelmien yli, joissa on ositettu negaatio. Tekniikka hyödyntää kyselyn sidonnaisuuksien etenemistä ja laajentaa Magic Set -optimointitekniikkaa (joka on alun perin määritelty ei-disjunktiivisille ohjelmille). Disjunktiivisten Datalog-ohjelmien tärkeä piirre on epämonotonisuus, joka edellyttää epädeterministisiä toteutuksia, kuten takaperin etsimistä. Uuden menetelmän erityispiirre on, että optimointia voidaan hyödyntää myös ei-deterministisessä vaiheessa. Erityisesti sen jälkeen, kun laskennan aikana on tehty joitakin oletuksia, ohjelman osat voivat muuttua epäolennaisiksi näiden oletusten mukaisen kyselyn kannalta. Tämä mahdollistaa hakuavaruuden dynaamisen karsimisen. Sitä vastoin aiemmin määriteltyjen Magic Set -menetelmien vaikutus disjunktiiviselle Datalogille rajoittuu prosessin deterministiseen osaan. Näin ollen ehdotetun menetelmän käytön potentiaalinen suorituskykyhyöty voi olla eksponentiaalinen, kuten empiirisesti voitiin havaita. Menetelmän oikeellisuus vahvistetaan ja todistetaan muodollisella tavalla Magic Setien ja perustelemattomien joukkojen välisen vahvan suhteen ansiosta, jota ei ole aiemmin tutkittu kirjallisuudessa. Tämä tieto mahdollistaa menetelmän ja oikeellisuustodistuksen laajentamisen luonnollisella tavalla myös ohjelmiin, joissa on ositettu negaatio. Ehdotettu menetelmä on toteutettu DLV-järjestelmässä, ja erilaisia kokeita on tehty sekä synteettisellä että todellisella datalla. Synteettisellä datalla tehdyt kokeelliset tulokset vahvistavat Magic Set -menetelmän käyttökelpoisuuden disjunktiivisille Datalog-ohjelmille, ja ne korostavat laskennallista hyötyä, joka uudella menetelmällä voidaan saavuttaa verrattuna aiemmin ehdotettuun Magic Set -menetelmään disjunktiivisille Datalog-ohjelmille. Todellisesta sovelluksesta otetuilla tiedoilla tehdyt lisäkokeet osoittavat Magic Set -menetelmän edut sovellusskenaariossa, joka on saanut paljon huomiota viime vuosina, eli ongelmassa, joka koskee käyttäjien kyselyihin vastaamista mahdollisesti epäjohdonmukaisissa tietokannoissa, jotka ovat peräisin itsenäisten tietolähteiden integroinnista.

**Tulos**

Taika-joukot disjunktiivisille datalog-ohjelmille

**Esimerkki 2.889**

Nollasummapelit on helppo ratkaista, koska ne voidaan esittää yksinkertaisina Markovin päätösprosesseina. Näin ei kuitenkaan ole yleissummaisten stokastisten pelien kohdalla. Filar ja Vrieze [2004] ovat laatineet melko yleisen optimointiongelman muotoilun yleisen summan stokastisille peleille. Siinä optimointiongelmalla on kuitenkin epälineaarinen tavoite ja epälineaarisia rajoitteita, joilla on erityinen rakenne. Koska tämän optimointiongelman sekä tavoitteen että rajoitusten gradientit ovat hyvin määriteltyjä, gradienttipohjaiset järjestelmät näyttävät olevan luonnollinen valinta. Käsittelemme gradienttijärjestelmää, joka on viritetty kahden pelaajan stokastisia pelejä varten. Osoitamme simuloinneilla, että tämä järjestelmä todellakin konvergoi Nashin tasapainoon yksinkertaisessa maastonetsintäongelmassa, joka on mallinnettu yleisen summan stokastisena pelinä. Osoittautuu kuitenkin, että vain optimointiongelman globaalit minimit vastaavat taustalla olevan yleisen summan stokastisen pelin Nash-tasapainoja, kun taas gradienttijärjestelmät takaavat vain konvergenssin paikallisiin minimiin. Tämän jälkeen esitämme tärkeitä välttämättömiä ehtoja, jotta gradienttijärjestelmät konvergoituvat Nashin tasapainotiloihin yleissummaisissa stokastisissa peleissä.

**Tulos**

Tutkimus asteittaisesta laskeutumisesta yleissummaisten stokastisten pelien osalta (Gradient Descent Schemes for General-Sum Stochastic Games)

**Esimerkki 2.890**

Generatiivisten vastakkaisverkkojen kouluttaminen on epävakaata suurissa ulottuvuuksissa, kun todellinen datajakauma sijaitsee alempiulotteisella moninaisuudella. Tällöin erottelija pystyy helposti erottamaan lähes kaikki generoidut näytteet, jolloin generaattori jää ilman mielekkäitä gradientteja. Ehdotamme, että yhtä generaattoria koulutetaan samanaikaisesti useita diskriminaattoreita vastaan, joista kukin tarkastelee eri satunnaista matalaulotteista datan projektiota. Osoitamme, että yksittäiset erottelijat antavat tällöin vakaita gradientteja generaattorille ja että generaattori oppii tuottamaan näytteitä, jotka ovat johdonmukaisia koko datan jakauman kanssa, jotta ne tyydyttävät kaikki erottelijat. Osoitamme kokeellisesti tämän lähestymistavan käytännön hyödyn ja osoitamme, että sen avulla voidaan tuottaa kuvanäytteitä, jotka ovat laadukkaampia kuin perinteinen harjoittelu yhdellä diskriminaattorilla.

**Tulos**

GAN-koulutuksen vakauttaminen useilla satunnaisprojektioilla

**Esimerkki 2.891**

Yleisradiouutisten transkriptio on mielenkiintoinen ja haastava sovellus suuren sanavaraston jatkuvalle puheentunnistukselle (LVCSR). Esittelemme yksityiskohtaisesti manuaalisesti segmentoidun ja annotoidun korpuksen rakenteen, joka sisältää yli 160 tuntia saksalaisia uutislähetyksiä, ja ehdotamme sitä LVCSR-järjestelmien arviointikehykseksi. Esitämme omat kokeelliset tuloksemme korpuksesta, jotka on saatu huippuluokan LVCSR-dekooderilla, mittaamme eri ominaisuusjoukkojen ja dekoodausparametrien vaikutusta ja osoitamme näin, että testijoukon reaaliaikainen dekoodaus on toteutettavissa pöytätietokoneella 9,2 prosentin sanavirheprosentilla.

**Tulos**

Yleisradion uutiskorpus saksalaisten LVCSR-järjestelmien arviointia ja virittämistä varten.

**Esimerkki 2.892**

Spektraalimenetelmät ovat edistäneet huomattavasti latenttien muuttujien mallien estimointia, ja ne ovat tuottaneet joukon uusia ja tehokkaita algoritmeja, joilla on vahvat teoreettiset takeet. Nykyiset spektrialgoritmit rajoittuvat kuitenkin suurelta osin diskreettien tai Gaussin jakaumien sekoituksiin. Tässä artikkelissa ehdotamme kernelmenetelmää moninäkökulmaisten latenttien muuttujien mallien oppimiseen, jolloin jokainen seoskomponentti voi olla epäparametrinen. Menetelmän keskeinen ajatus on upottaa moninäkökulmaisen latentin muuttujan yhteinen jakauma toistavaan kernel-hilbert-avaruuteen, minkä jälkeen latentit parametrit otetaan talteen käyttämällä robustia tensoripower-menetelmää. Todetaan, että ehdotetun menetelmän näytteenottokompleksisuus on kvadraattinen latenttien komponenttien lukumäärän suhteen ja matalan polynomin suhteen muiden relevanttien parametrien suhteen. Näin ollen ei-parametrinen tensorimenetelmämme latenttimuuttujamallien oppimiseen on otos- ja laskentatehokas. Lisäksi ei-parametrinen tensorivoimakkuusmenetelmä vertautuu kokeissamme hyvin EM-algoritmiin ja muihin olemassa oleviin spektrialgoritmeihin.

**Tulos**

Moninäkökulmaisten latenttimuuttujamallien ei-parametrinen estimointi

**Esimerkki 2.893**

Tähän mennessä ennustusvälin estimointiin ehdotetut menetelmät keskittyvät tapauksiin, joissa syötemuuttujat ovat numeerisia. Tietoaineistoissa, joissa on pelkästään nimellisiä syötemuuttujia, havaitaan tietueita, joissa on täsmälleen sama syötemuuttuja x, mutta erilaiset reaaliarvoiset tuotokset, jotka johtuvat järjestelmälle ominaisesta kohinasta. Nykyiset ennustusvälien estimointimenetelmät eivät käytä esitystapoja, joilla voitaisiin tarkasti mallintaa tällaista luontaista kohinaa nimellisillä syötteillä. Ehdotamme uutta ennustusvälien estimointimenetelmää, joka on räätälöity tämäntyyppisille tiedoille, jotka ovat yleisiä biologiassa ja lääketieteessä. Kutsumme menetelmää nimellä Distribution Adaptive Prediction Interval Estimation given Nominal inputs (DAPIEN), ja siinä on neljä päävaihetta. Ensin valitaan jakaumafunktio, joka edustaa parhaiten järjestelmän luontaista kohinaa kaikille ainutlaatuisille syötteille. Sitten päätellään valitun jakaumafunktion parametrit θi (esim. θi = [meani, variancei]) kaikille yksilöllisille tulovektoreille xi ja luodaan uusi vastaava harjoitusjoukko käyttäen pareja xi , θi. III). Tämän jälkeen koulutetaan malli ennustamaan θ uuden xu:n perusteella. Lopuksi laskemme ennustusvälin uudelle näytteelle käyttäen kumulatiivisen jakaumafunktion käänteislukua, kun koulutettu malli on ennustanut parametrit θ. Vertasimme DAPIEN-menetelmää yleisesti käytettyyn Bootstrap-menetelmään kolmella synteettisellä tietokokonaisuudella. Tuloksemme osoittavat, että DAPIEN tuottaa Bootstrapiin verrattuna tiukemmat ennustusvälit säilyttäen samalla halutun kattavuuden. Tämä työ voi helpottaa regressiomenetelmien laajempaa käyttöä lääketieteessä ja biologiassa, joissa on tarpeen antaa tiukkoja ennustusvälejä ja säilyttää kattavuus, kun syötemuuttujat ovat nimellisiä.

**Tulos**

DICTION INTERVAL ESTIMATION USING NOMINAL VARIABLES

**Esimerkki 2.894**

Tässä asiakirjassa tarkastellaan kahta erillistä kysymystä: Pystymmekö suorittamaan luonnollisen kielen käsittelytehtäviä ilman sanastoa? ja Pitäisikö meidän tehdä niin? Nykyiset luonnollisen kielen käsittelytekniikat perustuvat joko sanoihin yksikköinä tai käyttävät grammojen kaltaisia yksiköitä vain perusluokittelutehtävissä. Kuinka lähelle kone voi päästä päättelemään korpuksen sanojen ja lauseiden merkityksiä ilman leksikonia, vain grammoihin perustuen? Oma motivaatiomme tämän kysymyksen esittämiseen perustuu pyrkimyksiimme löytää sanojen ja lausekkeiden suosittuja trendejä kiinalaisesta sosiaalisesta verkkomediasta. Tämä kirjoitetun kiinan muoto käyttää niin paljon neologismeja, luovia merkkien sijoitteluja ja kirjoitusjärjestelmien yhdistelmiä, että sitä on kutsuttu "Marskin kieleksi". Lukijoiden on usein käytettävä visuaalisia jonoja, ääneen lukemisesta saatuja äänijonoja sekä tietämystään ja ymmärrystään ajankohtaisista tapahtumista ymmärtääkseen viestin. Suosittujen suuntausten analyysin kannalta erityinen ongelma on se, että sanaston rakentaminen on vaikeaa, kun uusien tapojen keksiminen viittaamaan sanaan tai käsitteeseen on helppoa ja yleistä. Luonnollisen kielen prosessoinnin osalta väitämme tässä asiakirjassa, että uudet kielenkäytön muodot sosiaalisessa mediassa haastavat koneiden kyvyt toimia sanojen kanssa ymmärtämisen perusyksikkönä, ei ainoastaan kiinan kielellä vaan mahdollisesti myös muilla kielillä.

**Tulos**

Kieli ilman sanoja: Pointillist-malli luonnollisen kielen käsittelyyn

**Esimerkki 2.895**

Puuttuvien attribuuttiarvojen imputointi lääketieteellisissä tietokokonaisuuksissa piilotiedon poimimiseksi lääketieteellisistä tietokokonaisuuksista on mielenkiintoinen tutkimusaihe, joka on hyvin haastava. Puuttuvia arvoja ei voida poistaa lääketieteellisistä tiedoista. Syynä voi olla se, että joitakin testejä ei ole tehty, koska ne ovat kustannustehokkaita, arvoja on jäänyt puuttumaan kliinisten tutkimusten yhteydessä, arvoja ei ole ehkä kirjattu, vain muutamia syitä mainitakseni. Tiedonlouhinnan tutkijat ovat ehdottaneet erilaisia lähestymistapoja puuttuvien arvojen löytämiseksi ja korvaamiseksi luokittelutarkkuuden lisäämiseksi, jotta sairauksia voidaan ennustaa tarkasti. Tässä asiakirjassa ehdotetaan uutta imputointimenetelmää puuttuvien arvojen imputoimiseksi ja luokittelun suorittamiseksi puuttuvien arvojen korjaamisen jälkeen. Lähestymistapa perustuu klusterointikäsitteeseen, ja sen tavoitteena on tietueiden ulottuvuuden vähentäminen. Käsitelty tapaustutkimus osoittaa, että puuttuvat arvot voidaan korjata ja imputoida tehokkaasti pienentämällä dimensioita. Ehdotetun lähestymistavan merkitys luokittelussa näkyy tapaustutkimuksessa, jossa annetaan yksi luokkatunnus toisin kuin usean tunnisteen antaminen, jos dimensioiden pienentämistä ei tehdä. Avainsanat- imputointi; puuttuvat arvot; ennuste; lähin naapuri, klusteri, sairauskertomukset, dimensioiden pienentäminen.

**Tulos**

Innovatiivinen imputointi- ja luokittelumenetelmä tautien tarkkaa ennustamista varten.

**Esimerkki 2.896**

Bilattice-pohjainen kolmio tarjoaa tyylikkään algebrallisen rakenteen epämääräisen ja epävarman tiedon kanssa päättelyyn. Bilattice-pohjaisen kolmion intervallien totuus- ja tietämysjärjestys ei kuitenkaan pysty käsittelemään toistuvia uskomusten tarkistuksia, mikä on ei-monotonisen päättelyn olennainen piirre. Lisäksi bilattice-pohjaisen kolmion intervalleille antama järjestys ei ole joskus intuitiivinen. Tässä työssä rakennamme vaihtoehtoisen algebrallisen rakenteen, nimittäin esijärjestykseen perustuvan kolmion, ja muotoilemme sille asianmukaiset loogiset konnektiivit. Se on bilattice-pohjaisen kolmion parannus, jolla voidaan käsitellä uskomusten tarkistamista ei-monotonisessa päättelyssä.

**Tulos**

Ennakkotilaukseen perustuva kolmio: Bilattice-Based Triangle for Belief Revision in Nonmonotonic Reasoning: A Modified Version of Bilattice-Based Triangle for Belief Revision in Nonmonotonic Reasoning.

**Esimerkki 2.897**

Jakautumissemanttiset mallit oppivat sanojen vektoriedustukset niiden esiintymisyhteyksien avulla. Vaikka kontekstin valinta (joka usein tapahtuu liukuvan ikkunan muodossa) vaikuttaa suoraan tuloksena saataviin upotuksiin, tämän mallin komponentin tarkkaa roolia ei vielä täysin ymmärretä. Tässä asiakirjassa esitetään konteksti-ikkunoiden järjestelmällinen analyysi, joka perustuu neljään eri hyperparametriin. Koulutamme jatkuvia SkipGram-malleja kahdella englanninkielisellä korpuksella näiden hyperparametrien eri yhdistelmillä ja arvioimme niitä sekä leksikaalisen samankaltaisuuden että analogian tehtävissä. Merkittäviä kokeellisia tuloksia ovat ristikkäisten lauseyhteyksien myönteinen vaikutus ja oikean kontekstin ikkunoiden yllättävän hyvä suorituskyky.

**Tulos**

Konteksti-ikkunoiden uudelleenmäärittely sanojen upotusmalleja varten: Kokeellinen tutkimus

**Esimerkki 2.898**

Esittelemme menetelmän, jolla voidaan kouluttaa kvantittuneita neuroverkkoja (QNN) - neuroverkkoja, joiden painot ja aktivoinnit ovat erittäin epätarkkoja (esim. 1-bittisiä), ajonaikaisesti. Harjoitteluaikana kvantisoituja painoja ja aktivaatioita käytetään parametrien gradienttien laskemiseen. Eteenpäin suuntautuvan syötön aikana QNN:t vähentävät huomattavasti muistin kokoa ja käyttöoikeuksia ja korvaavat useimmat aritmeettiset operaatiot bittiviisailla operaatioilla. Tämän seurauksena virrankulutuksen odotetaan vähenevän huomattavasti. Koulutimme QNN:t MNIST-, CIFAR-10-, SVHN- ja ImageNet-tietokantoihin. Tuloksena saadut QNN:t saavuttavat ennustustarkkuuden, joka on verrattavissa niiden 32-bittisiin vastineisiin. Esimerkiksi AlexNetin kvantifioitu versio, jossa on 1-bittiset painot ja 2-bittiset aktivoinnit, saavuttaa 51 prosentin top-1-tarkkuuden. Lisäksi kvantisoimme myös parametrien gradientit 6-bittisiksi, mikä mahdollistaa gradienttien laskennan vain bittiviisaalla operaatiolla. Kvantisoituja rekursiivisia neuroverkkoja testattiin Penn Treebank -tietokannassa, ja ne saavuttivat vastaavan tarkkuuden kuin niiden 32-bittiset vastineet, jotka käyttivät vain 4 bittiä. Lopuksi ohjelmoimme binäärisen matriisikertolaskennan GPU-ytimen, jonka avulla MNIST QNN:ää voidaan ajaa 7 kertaa nopeammin kuin optimoimattomalla GPU-ytimellä ilman, että luokittelutarkkuus heikkenee. QNN-koodi on saatavilla verkossa. 1 ar X iv :1 60 9. 07 06 1v 1 [ cs .N E ] 2 2 Se p 20 16 Hubara, Courbariaux, Soudry, El-Yaniv ja Bengio.

**Tulos**

Kvantisoidut neuroverkot Kvantisoidut neuroverkot: Neuraaliverkkojen kouluttaminen matalan tarkkuuden painoilla ja aktivoinneilla

**Esimerkki 2.899**

Nykyään Twitteriin liittyvät maantieteelliset tiedot ovat ratkaisevan tärkeitä hienojakoisille sovelluksille. Twitteristä on kuitenkin saatavilla vain vähän maantieteellistä tietoa, mikä tekee monien sovellusten toteuttamisesta haastavaa. Tällaisessa tilanteessa twiitin sijainnin arviointi on tutkimuksen tärkeä tavoite. Toisin kuin useimmissa aiemmissa tutkimuksissa, joissa luokittelutehtävänä on ennalta määritellyn piirin estimointi, tässä tutkimuksessa käytetään todennäköisyysjakaumaa twiitin rikkaamman informaation esittämiseksi, ei ainoastaan sijainnin vaan myös sen moniselitteisyyden osalta. Tämän mallinnuksen toteuttamiseksi ehdotamme konvoluutioseoksen tiheysverkkoa (CMDN), joka käyttää tekstidataa seosmallin parametrien estimointiin. Kokeellisesti saadut tulokset paljastavat, että CMDN saavutti tarkkojen koordinaattien ennustamisessa käytettävistä menetelmistä parhaan ennustuskyvyn. Se tarjoaa myös kvantitatiivisen esityksen sijainnin epäselvyydestä kunkin twiitin osalta, joka toimii asianmukaisesti luotettavien sijaintiarvioiden poimimisessa.

**Tulos**

Tiheyden estimointi geopaikannusta varten konvolutiivisen sekoitustiheysverkon avulla

**Esimerkki 2.900**

Kernel-pohjaiset klusterointialgoritmit pystyvät kuvaamaan reaalimaailman tietojen epälineaarisen rakenteen. Erilaisista kernel-pohjaisista klusterointialgoritmeista kernel k -means on saavuttanut suosiota sen yksinkertaisen iteratiivisen luonteen ja helpon toteutuksen vuoksi. Sen ajonaikainen monimutkaisuus ja muistijalanjälki kasvavat kuitenkin neliöllisesti aineiston koon kasvaessa, joten suuria aineistoja ei voida klusteroida tehokkaasti. Tässä artikkelissa ehdotamme satunnaistamiseen perustuvaa approksimointijärjestelmää, jota kutsutaan approksimatiiviseksi k-means-ytimen k-meansiksi. Lähestymme klusterikeskuksia käyttämällä kernelin samankaltaisuutta muutaman otospisteen ja kaikkien datajoukon pisteiden välillä. Osoitamme, että ehdotetulla menetelmällä saavutetaan parempi klusterointisuorituskyky kuin perinteisillä matalan kernelin approksimaatioon perustuvilla klusterointijärjestelmillä. Osoitamme myös, että sen ajoaika ja muistivaatimukset ovat huomattavasti pienemmät kuin kernel k -means -menetelmän, ja klusterointilaatu heikkenee vain vähän useilla julkisilla suurilla aineistoilla. Sitten käytämme ensemble-klusterointitekniikoita algoritmimme suorituskyvyn parantamiseksi entisestään.

**Tulos**

Skaalautuva ytimen klusterointi: Ydinkeskiarvo: Approximate Kernel k -means Clustering: Approximate Kernel k -means

**Esimerkki 2.901**

MOOC-kurssien (massiivisten avoimien verkkokurssien) suosion myötä joukkoistamisen avulla tapahtuvasta luokittelusta on tullut yleinen lähestymistapa laajamittaisissa kursseissa. Joukkoistaminen kuitenkin rajoittaa sellaisten monimutkaisten tehtävien arvosanojen saamista, jotka vaativat erityisiä taitoja ja ponnisteluja, sillä joukkoresursointi ei ole riittävän tietämyksellistä, koska joukosta tulevat työntekijät eivät tunne niitä riittävästi. Joukkoarvioijien rajallisen tietämyksen vuoksi osittaisiin näkökulmiin perustuvasta arvioinnista tulee suuri haaste arvioitaessa monimutkaisia tehtäviä joukkoistamisen avulla. Tämä koskee erityisesti sellaisia tehtäviä, joiden arvioinnissa ei tarvita vain erityistä tietämystä, vaan jotka olisi myös arvioitava kokonaisuutena sen sijaan, että ne olisi pilkottu pienempiin ja yksinkertaisempiin osatehtäviin. Ehdotamme kehystä monimutkaisten tehtävien arvioimiseksi useiden eri näkökulmien avulla, jotka ovat asiantuntijoiden määrittelemiä erilaisia arviointinäkökulmia, jotta saadaan aikaan yhdenmukaisuus. Kunkin näkökulman arvosanojen yhdistämiseen käytetään luokittelijoiden variansseihin perustuvaa aggregointialgoritmia. Tunnistamme myös arvioijien ennakkoluulot ja poistamme ne kunkin tehtävän näkökulman osalta. Ennakkoluokitusmalli määrittää, miten käyttäytyminen on ennakkoluokittajien kesken vinoutunutta, ja se havaitaan tilastollisella tekniikalla. Ehdotettua lähestymistapaa analysoidaan synteettisellä aineistolla. Osoitamme, että mallimme antaa tarkempia tuloksia verrattuna luokitusmenetelmiin, joissa ei käytetä eri näkymiä ja joissa ei käytetä vääristymiä poistavaa algoritmia. Avainsanat - monimutkainen tehtävä; joukkoistaminen; näkymä; puolueellinen malli; debias; Vancouver-algoritmi

**Tulos**

Monimutkaisen tehtävän arviointi joukkoistamisen avulla: Lähestymistapa useita näkökulmia

**Esimerkki 2.902**

Tässä asiakirjassa esitellään Ontology Web Search Engine -hanke. Tämän artikkelin päätarkoituksena on kehittää tällainen projekti, joka voidaan toteuttaa helposti. Ontology Web Search Engine on ohjelmisto, jolla etsitään ja indeksoidaan ontologioita verkossa. Tarkoitetaan OWL-ontologioita (Web Ontology Languages), jotka ovat välttämättömiä SWES:n (Semantic Web Expert System) toiminnalle. SWES on asiantuntijajärjestelmä, joka käyttää verkosta löytämiään ontologioita, luo niistä sääntöjä ja täydentää tietopohjaansa näillä tuotetuilla säännöillä. SWES:n odotetaan toimivan yleisenä asiantuntijajärjestelmänä keskivertokäyttäjälle.

**Tulos**

KOHTI ONTOLOGIAN VERKKOHAKUKONETTA

**Esimerkki 2.903**

Suurten tietokokonaisuuksien lisääntyessä kernelmenetelmien suosio laski ja neuroverkot ottivat jälleen vallan. Ydinmenetelmien suurin ongelma on se, että ydinmatriisi kasvaa kvadraattisesti datapisteiden määrän myötä. Useimmat yritykset skaalata kernelmenetelmiä ratkaisevat tämän ongelman hylkäämällä datapisteitä tai jonkin kernelikartan approksimaation perusfunktioita. Tässä esitämme yksinkertaisen mutta tehokkaan vaihtoehdon kernelmenetelmien skaalaamiseen, jossa otetaan huomioon koko datajoukko emprikaalisen kernelikartan kaksinkertaisen stokastisen optimoinnin avulla. Algoritmi on suoraviivainen toteuttaa erityisesti rinnakkaistoteutusympäristöissä; se hyödyntää klassisten kernelfunktioiden koko tehoa ja monipuolisuutta ilman tarvetta muotoilla eksplisiittisesti kernelikartan approksimaatiota. Esitämme empiiristä näyttöä siitä, että algoritmi toimii suurilla datajoukoilla.

**Tulos**

Kaksinkertaisesti stokastinen suuren mittakaavan kerneloppiminen empiirisen kernelikartan avulla

**Esimerkki 2.904**

Monitehtäväoppimisen (multi-task learning, MTL) toimintatapana on alusta lähtien ollut minimoida empiiristen riskien tehtäväkohtainen keskiarvo. Esittelemme MTL:lle yleistetyn häviökompositio-paradigman, joka sisältää alaperheenä muotoilujen kirjon. Tämän spektrin yksi päätepiste on minimax MTL: uusi MTL-muotoilu, joka minimoi tehtävien empiiristen riskien maksimin. Erään minimax MTL:n tietyn relaksaation avulla saadaan minimax MTL:n ja klassisen MTL:n välille jatkumo MTL-formulointeja. Täydellinen paradigma itsessään on häviökompositionaalinen, ja se toimii empiiristen riskien vektorilla. Se sisältää minimax MTL:n, sen lievennykset ja monia uusia MTL-formulointeja erikoistapauksina. Osoitamme teoreettisesti, että minimax MTL pyrkii välttämään huonoimpia tuloksia äskettäin laadituissa testitehtävissä oppimaan oppimisen (LTL) testiympäristössä. Useiden MTL-formulointien tulokset synteettisissä ja todellisissa ongelmissa MTL- ja LTL-testiympäristöissä ovat rohkaisevia.

**Tulos**

Minimax-monitehtäväoppiminen ja yleistetty häviö-koostumusparadigma MTL:lle

**Esimerkki 2.905**

Vuorovaikutteiset aihepiirimallit ovat tehokkaita välineitä, kun halutaan ymmärtää suuria tekstikokoelmia. Nykyiset otantaan perustuvat interaktiiviset aihepiirimallinnusmenetelmät skaalautuvat kuitenkin huonosti suuriin aineistoihin. Ankkurimenetelmät, jotka käyttävät yksittäistä sanaa aiheen yksilöimiseen, tarjoavat vuorovaikutteisessa työssä tarvittavan nopeuden, mutta niistä puuttuu mekanismi, jolla voidaan syöttää aiempaa tietoa, ja niistä puuttuu intuitiivinen semantiikka, jota tarvitaan käyttäjälle suunnatuissa sovelluksissa. Ehdotamme sanojen yhdistelmiä ankkureiksi, jotka menevät nykyisiä yhden sanan ankkurialgoritmeja pidemmälle - lähestymistapa, jota kutsumme "Tandem-ankkureiksi". Aloitamme tämän lähestymistavan synteettisellä tutkimuksella, sitten sovellamme lähestymistapaa interaktiiviseen aiheiden mallintamiseen käyttäjätutkimuksessa ja vertaamme sitä interaktiivisiin ja ei-interaktiivisiin lähestymistapoihin. Tandem-ankkurit ovat nopeampia ja intuitiivisempia kuin nykyiset interaktiiviset lähestymistavat. Aihepiirimallit tislaavat suuret tekstikokoelmat aihepiireiksi, jolloin saadaan korkean tason yhteenveto aineiston temaattisesta rakenteesta ilman manuaalista merkintää. Sen lisäksi, että aihepiirien mallintamista käytetään ajankohtaisten suuntausten löytämisen helpottamiseksi (Gardner et al., 2010), sitä käytetään monenlaisissa ongelmissa, kuten asiakirjojen luokittelussa (Rubin et al., 2012), tiedonhaussa (Wei ja Croft, 2006), kirjoittajan tunnistamisessa (Rosen-Zvi et al., 2004) ja tunneanalyysissä (Titov ja McDonald, 2008). Aihepiirimallien vakuuttavin käyttö on kuitenkin auttaa käyttäjiä ymmärtämään suuria tietokokonaisuuksia (Chuang et al., 2012). Vuorovaikutteinen aihepiirimallinnus (Hu et al., 2014) antaa ei-asiantuntijoille mahdollisuuden tarkentaa automaattisesti luotuja aihepiirejä, jolloin aihepiirimallit eivät ole enää niinkään "ota tai jätä" -ehdotus. Ihmisten panos harjoittelun aikana parantaa mallin laatua ja antaa käyttäjille mahdollisuuden ohjata aiheita tietyllä tavalla, jolloin malli voidaan räätälöidä tiettyä jatkokäsittelytehtävää tai analyysia varten. Haittapuolena on, että interaktiivinen aihepiirien mallintaminen on hidasta - algoritmit skaalautuvat yleensä korpuksen koon mukaan - ja vaatii käyttäjältä epäintuitiivista tietoa must-linkki- ja cannot-linkki-rajoitusten muodossa (Andrzejewski et al., 2009). Korjaamme nämä interaktiivisen aihepiirimallinnuksen puutteet käyttämällä interaktiivista versiota ankkurisanojen algoritmista aihepiirimalleille. Ankkurialgoritmi (Arora et al., 2013) on vaihtoehtoinen aihepiirien mallinnusalgoritmi, joka skaalautuu aineistossa olevien uniikkien sanatyyppien lukumäärän eikä dokumenttien tai tokenien lukumäärän mukaan (luku 1). Tämä tekee ankkurialgoritmista riittävän nopean vuorovaikutteiseen käyttöön myös web-mittakaavan dokumenttikokoelmissa. Ankkurimenetelmän haittapuolena on se, että ankkurisanat - sanat, joilla on suuri todennäköisyys kuulua yhteen aihepiiriin - eivät ole intuitiivisia. Laajennamme ankkurialgoritmia siten, että siinä voidaan käyttää useita ankkurisanoja yhdessä (jakso 2). Tandem-ankkurit eivät ainoastaan paranna vuorovaikutteista tarkennusta, vaan tekevät myös taustalla olevasta ankkuripohjaisesta menetelmästä intuitiivisemman. Vuorovaikutteisessa aihepiirien mallintamisessa tandem-ankkurit tuottavat laadukkaampia aihepiirejä kuin yhden sanan ankkurit (jakso 3). Tandem-ankkurit tarjoavat puitteet nopealle vuorovaikutteiselle aihepiirien mallintamiselle: käyttäjät parantavat ja tarkentavat olemassa olevaa mallia monisana-ankkureiden avulla (jakso 4). Verrattuna olemassa oleviin menetelmiin, kuten Interactive Topic Models (Hu et al., 2014), menetelmämme on paljon nopeampi.

**Tulos**

Tandem-ankkurointi: usean sanan ankkurimenetelmä interaktiiviseen aihepiirien mallintamiseen.

**Esimerkki 2.906**

Tutkimme suoran summan ongelmaa eri tavoin yksityisen PAC-oppimisen yhteydessä: Mikä<lb>on k oppimistehtävän samanaikaisen ratkaisemisen näytekompleksisuus differentiaalisen yksityisyyden vallitessa, ja miten<lb>verrataan tätä kustannusta k oppimistehtävän ratkaisemiseen ilman yksityisyyttä? Asetelmassamme yksittäinen<lb>näyte koostuu toimialueen elementistä x, joka on merkitty k:lla tuntemattomalla käsitteellä (c1, . . . . ,<lb>ck). <lb>monilukuoppijan tavoitteena on tuottaa k hypoteesia (h1, . . . , hk), jotka yleistävät syötetyt esimerkit.<lb>Epätietoisuudesta piittaamatta näytteenoton monimutkaisuus, joka tarvitaan k käsitteen samanaikaiseen oppimiseen, on<lb>olennaisesti sama kuin yhden käsitteen oppimiseen. Differentiaalisen yksityisyyden suojan vallitessa perusstrategia<lb>kunkin hypoteesin oppimisesta itsenäisesti tuottaa näytekompleksisuuden, joka kasvaa polynomisesti k:n kanssa.<lb>Joillekin käsiteluokille annamme monioppimismenetelmiä, jotka vaativat vähemmän näytteitä kuin perusstrategia.<lb>Kahdingossa annamme kuitenkin myös alempia rajoja, jotka osoittavat, että hyvin yksinkertaisillekin käsiteluokille<lb>yksityisen monioppimisen näytekustannusten on kasvettava polynomisesti k:n suhteen.

**Tulos**

Useiden käsitteiden samanaikainen yksityinen oppiminen

**Esimerkki 2.907**

Syy-seuraussuhteiden päättely havaintoaineistosta on aktiivisen tutkimus- ja kehitystyön kohteena tilastotieteessä ja tietojenkäsittelytieteessä. Tätä tarkoitusta varten on kehitetty monia tilasto-ohjelmistoista riippuvaisia työkalupaketteja. Nämä työkalupaketit eivät kuitenkaan sovellu suuriin tietokokonaisuuksiin. Tässä artikkelissa kuvaamme joukon tekniikoita, joilla voidaan ilmaista kausaalista päättelyä koskevat tehtävät havaintoaineistosta SQL:llä. Tämä paketti tukee uusimpia kausaaliseen päättelyyn käytettäviä menetelmiä, ja se toimii mittakaavassa tietokantamoottorissa. Lisäksi esittelemme useita optimointitekniikoita, jotka nopeuttavat kausaalista päättelyä merkittävästi sekä online- että offline-ympäristössä. Arvioimme tekniikkojemme laatua ja suorituskykyä kokeilemalla todellisia tietokokonaisuuksia.

**Tulos**

ZaliQL: SQL-pohjainen kehys kausaalisten johtopäätösten tekemiseen suurista tiedoista.

**Esimerkki 2.908**

Inkluusio-ongelma käsittelee sitä, miten voidaan (graafisesti) kuvailla, ovatko kaikki DAG:n K tuottaman mallin riippumattomuuslausumat toisen DAG:n L tuottamassa mallissa. Meek (1997) päätteli, että tämä inkluusio pätee, jos on olemassa sellainen DAG:ien sarja L:stä K:hon, että vain tietyt "lailliset" nuolten kääntämisoperaatiot ja "lailliset" nuolten lisäysoperaatiot suoritetaan, jotta saadaan sarjan seuraava DAG. Tässä artikkelissa esitämme useita DAG-mallien inkluusion karakterisointeja ja tarkistamme Meekin arvelun siinä tapauksessa, että DAG:t K ja L eroavat toisistaan korkeintaan yhden vierekkäisyyden osalta. Lämmittelynä annetaan DAG:ien vastaavuuden graafisten karakterisointien tiukka todistus.

**Tulos**

Bayes-verkkojen sulkeutumisen luonnehtiminen

**Esimerkki 2.909**

1 Rivit, prosessit ja litteät kieliopit ....................................................................................... 2 2 Rank-tulkinta-arkkitehtuuri monilineaarisille kieliopeille ......................................... 6 2.1 Hahmotelma Rank-tulkinta-arkkitehtuurin kehyksestä .............................................. 6 2.2 Alustava huomautus lineaarisuudesta ja hierarkiasta fraseaalisella asteikolla ..................................... 7 2.3 Rank-tulkinta-arkkitehtuurin karakterisointi .................................................. 9 2.3.1 Taustaa ..................................................................................................................... 9 2.3.2 Muodollinen yhteenveto ........................................................................................................... 10 2.3.3 Kontrasti perinteisiin näkemyksiin kielen arkkitehtuurista .............................................. 11 2.3.4 Rankkien sui generis -ominaisuuksien etsiminen ..................................................................... 11 2.4 Proseduraaliset näkökulmat rank-hierarkiaan ................................................................. 12 3 Diskurssijärjestys .............................................................................................................. 15 3.1 Diskurssin kuvioinnin ensisijaisuus ................................................................................ 15 3.2 Vierusparin intonointi ........................................................................................ 16 3.3 Chanted 'call' intonaatio .................................................................................................. 18 4 Lausuman tai tekstin arvojärjestys ................................................................................................... 20 5 Lauseen arvojärjestys ................................................................................................................... 22 5.1 Lauseen rakenteen ominaispiirteet .................................................................................. 22 5.2 Lineaariset sekvenssit, iteraatio: säännölliset ja epäsäännölliset kieliopit ........................................ 23 5.3 Huomautus pitkän matkan ja ristikkäisriippuvuuksista ................................................... 26 5.4 Prosodis-foneettinen tulkinta lauseen arvojärjestyksessä ........................................................ 27 6 Sanan arvojärjestys ..................................................................................................................... 30 6.1 Litteät sanat ......................................................................................................................... 30 6.2 Litteät johdokset ................................................................................................................. 30 6.3 Litteät yhdyssanat ................................................................................................................ 31 6.4 Prosodis-foneettinen tulkinta sana-asteella ............................................................ 31 7 Yhteenveto ja päätelmät .................................................................................................... 33 7.1 Dualiteetista monilinjaiseen kielioppiin ja sijatulkinta-arkkitehtuuriin ................. 33 7.2 Yleistäminen stokastisiin litteisiin lineaarisiin malleihin ................................................................. 34 7.3 Tuleva työ ...................................................................................................................... 35 8 Viitteet ........................................................................................................................... 36

**Tulos**

Multilineaarinen kielioppi: sijat ja tulkinnat

**Esimerkki 2.910**

Biologiset hermoverkot ovat järjestelmiä, joilla on poikkeuksellisia laskentakykyjä, joita evoluutio, kehitys ja elinikäinen oppiminen ovat muokanneet. Näiden tekijöiden vuorovaikutus johtaa mukautuvan käyttäytymisen ja älykkyyden syntymiseen, mutta koko vuorovaikutusjärjestelmän monimutkaisuus estää keskeisten tekijöiden ymmärtämisen. Tällaisten monimutkaisten luonnonilmiöiden innoittamana Evolved Plastic Artificial Neural Networks (EPANN) käyttää simuloitua evoluutiota in-silico kasvattaakseen muovisia neuroverkkoja, keinotekoisia järjestelmiä, jotka koostuvat antureista, ulostuloista ja muovisista komponenteista, jotka muuttuvat vastauksena ympäristössä saatuihin aisti-ulostulokokemuksiin. Nämä järjestelmät voivat paljastaa sopeutumisen keskeiset algoritmiset osatekijät, löytää itsenäisesti uusia sopeutumisalgoritmeja ja johtaa hypoteeseihin biologisen sopeutumisen synnystä. EPANNit ovat edistyneet huomattavasti kahden viime vuosikymmenen aikana. Nykyiset tieteelliset ja teknologiset edistysaskeleet keinotekoisten neuroverkkojen alalla luovat nyt edellytykset radikaalisti uusille lähestymistavoille ja tuloksille. Erityisesti useimmissa syvissä neuroverkoissa nykyisin käytettävien käsin suunniteltujen rakenteiden ja algoritmien rajoitukset voitaisiin voittaa joustavammilla ja innovatiivisemmilla ratkaisuilla. Tämä asiakirja kokoaa yhteen erilaisia inspiroivia ideoita, jotka määrittelevät EPANN-alan. Tärkeimmät laskennalliset menetelmät ja tulokset käydään läpi. Lopuksi esitellään uusia mahdollisuuksia ja kehitystä.

**Tulos**

Syntynyt oppimaan: Kehittyneiden muovisten keinotekoisten neuroverkkojen inspiraatio, kehitys ja tulevaisuus.

**Esimerkki 2.911**

Kuvaamme yksinkertaisen järjestelmän, jonka avulla agentti voi tutkia ympäristöään valvomattomasti. Järjestelmässämme on kaksi versiota samasta agentista, Alice ja Bob, vastakkain. Alice ehdottaa Bobille tehtävää, jonka Bob yrittää suorittaa, minkä jälkeen Bob yrittää suorittaa tehtävän. Tässä työssä keskitymme (lähes) palautuviin ympäristöihin tai ympäristöihin, jotka voidaan palauttaa, ja Alice "ehdottaa" tehtävää suorittamalla joukon toimintoja, minkä jälkeen Bobin on osittain peruutettava tai toistettava ne. Sopivan palkitsemisrakenteen avulla Alice ja Bob luovat automaattisesti tutkimussuunnitelman, mikä mahdollistaa agentin valvomattoman harjoittelun. Kun agentti käytetään RL-tehtävässä ympäristössä, tämä valvomaton harjoittelu vähentää oppimiseen tarvittavien jaksojen määrää.

**Tulos**

Sisäinen motivaatio ja automaattiset opetussuunnitelmat epäsymmetrisen itsepelin avulla

**Esimerkki 2.912**

Esittelemme lähestymistavan, jolla havaitaan Java-luokkiin viittaavien ohjelmistoalan entiteettien väliset koordinaattisuhteet. Yleensä suhteet löydetään tarkastelemalla tekstiobjekteihin liittyviä korpustilastoja. Joillakin teknisillä aloilla meillä on kuitenkin käytettävissämme lisätietoa olioiden nimeämistä reaalimaailman objekteista, mikä viittaa siihen, että "maadoitettuja" olioita koskevan tiedon yhdistäminen korpustilastoihin voisi johtaa parempiin menetelmiin suhteiden löytämiseksi. Tätä varten kehitämme Java-luokille samankaltaisuusmittarin, jossa käytetään ohjelmistojen käyttötapoja koskevaa jakaumatietoa, jonka yhdistämme korpustilastoihin, jotka koskevat niiden kontekstien jakaumaa, joissa luokat esiintyvät tekstissä. Lähestymistapamme avulla ristiinvalidointitarkkuutta voidaan parantaa huomattavasti, noin 60 prosentista 88 prosenttiin. Ihmisten tekemien merkintöjen tulokset osoittavat, että luokittimemme F1-tulos on 86 prosenttia 1000 parhaasta ennustetusta parista.

**Tulos**

Ohjelmistokokonaisuuksien välisten koordinaattitermisuhteiden perusteltu löytäminen

**Esimerkki 2.913**

Monikätisen rosvon ongelma (multi-armed bandit problem, MBP) on ongelma, jossa pyritään löytämään mahdollisimman tarkasti ja nopeasti kannattavin vaihtoehto joukosta vaihtoehtoja, jotka antavat stokastisia palkkioita aiempien kokemusten perusteella. Jäykän kappaleen vaihtelevien liikkeiden innoittamana hinauspelissä muotoilimme ainutlaatuisen hakualgoritmin, jota kutsumme "hinauksen dynamiikaksi (TOW)", MBP:n ratkaisemiseksi tehokkaasti [1-5]. Kognitiivinen media access, jolla tarkoitetaan monikäyttäjäkanavien jakamista kognitiivisessa radiossa, voidaan tulkita kilpailulliseksi moniaseisen banditin ongelmaksi (competitive multi-armed bandit problem, CMBP); ongelmana on määrittää optimaalinen strategia kanavien jakamiseksi käyttäjille, joka tuottaa maksimaalisen kokonaishyödyn kaikille käyttäjille [6]. Tässä osoitamme, että CMBP:n ratkaisemiseen on mahdollista rakentaa fyysinen laite, jota kutsumme nimellä "TOW Bombe", hyödyntämällä TOW-dynamiikkaa, joka esiintyy kytketyissä kokoonpuristumattomissa nestesylintereissä. Tällä analogisella laskentalaitteella saavutetaan "sosiaalisesti maksimaalinen" resurssien jako, joka maksimoi kognitiivisen väliaineen käytön kokonaispalkkiot maksamatta valtavia laskentakustannuksia, jotka kasvavat eksponentiaalisesti ongelman koon funktiona.

**Tulos**

Päätöksentekijä, joka käyttää yhdistettyjä kokoonpuristumattomien nesteiden sylintereitä.

**Esimerkki 2.914**

Tiedonlouhinnan alalla on ehdotettu monia klusterointimenetelmiä, mutta vakioversioissa ei oteta huomioon epävarmoja tietokantoja. Tässä artikkelissa käsitellään uutta lähestymistapaa epävarmojen tietojen klusterointiin käyttämällä hierarkkista klusterointia, joka on määritelty uskomusfunktiokehyksessä. Uskomushierarkkisen klusteroinnin päätavoitteena on sallia kohteen kuuluminen yhteen tai useampaan klusteriin. Jokaiseen kuulumiseen liitetään uskomusaste, ja klusterit yhdistetään pignististen ominaisuuksien perusteella. Kokeet todellisilla epävarmoilla tiedoilla osoittavat, että ehdotettua menetelmää voidaan pitää suotuisana työkaluna.

**Tulos**

Uskomushierarkkinen klusterointi

**Esimerkki 2.915**

Käsittelemme alueen suojaamiseen liittyvää ongelmaa graafipohjaisissa skenaarioissa, joissa on useita liikkuvia agentteja ja joissa agenttien välillä ylläpidetään yhteyksiä sen varmistamiseksi, että ne voivat kommunikoida keskenään. Ongelma koostuu kahdesta vihamielisestä agenttiryhmästä, jotka liikkuvat suuntaamattomassa graafissa, jonka molemmat ryhmät jakavat. Agentit sijoitetaan graafin kärkipisteisiin; enintään yksi agentti voi olla kärkipisteessä, ja ne voivat siirtyä viereisiin kärkipisteisiin konfliktivapaasti. Joukkueilla on epäsymmetriset tavoitteet: yhden joukkueen hyökkääjien tavoitteena on tunkeutua tietylle alueelle, kun taas vastajoukkueen puolustajien tavoitteena on suojella aluetta hyökkääjien tunkeutumiselta miehittämällä valitut kärkipisteet. Puolustajien joukkueen on säilytettävä omien agenttiensa miehittämien kärkipisteiden yhteydet näkyvyysgraafissa. Näkyvyysgraafi mallintaa mahdollisuutta kommunikaatioon kärkipisteiden välillä. Tutkimme strategioita, joiden avulla puolustajaryhmän miehittämät kärkipisteet voidaan jakaa hyökkäävien agenttien estämiseksi siten, että yhteydet säilyvät samalla. Tätä varten varataan osajoukko puolustavista agenteista, jotka eivät yritä estää hyökkääjiä vaan sijoitetaan tukemaan joukkueen yhteyksiä. Strategioiden suorituskykyä testataan useilla vertailuarvoilla. Strategian onnistuminen riippuu suuresti instanssin tyypistä, joten yksi tämän työn panoksista on se, että tunnistamme sopivia strategioita erilaisille instanssityypeille.

**Tulos**

Ad-Hoc-viestintäverkon ylläpito alueellisissa suojausskenaarioissa, joissa on vastapuolen agentteja.

**Esimerkki 2.916**

Tutkimme yhteisten häviöpintojen ominaisuuksia niiden Hessin matriisin avulla. Erityisesti syväoppimisen yhteydessä osoitamme empiirisesti, että Hessin matriisin spektri koostuu kahdesta osasta: (1) nollan lähelle keskitetystä massasta ja (2) massasta poikkeavista arvoista. Esitämme numeerisia todisteita ja matemaattisia perusteluja seuraaville Sagun et al. [2016] esittämille arvauksille: Kiinnittämällä dataa ja lisäämällä parametrien lukumäärää ainoastaan skaalautuu spektrin bulkki; ulottuvuuden kiinnittäminen ja datan muuttaminen (esimerkiksi lisäämällä lisää klustereita tai tekemällä datasta vähemmän erotettavaa) vaikuttaa vain outlieriin. Uskomme, että havainnoillamme on silmiinpistäviä vaikutuksia ei-konveksaaliseen optimointiin suurissa ulottuvuuksissa. Ensinnäkin tällaisten maisemien litteys (jota voidaan mitata Hessianin singulaarisuudella) merkitsee sitä, että klassiset käsitykset vetovoima-alueista voivat olla varsin harhaanjohtavia. Ja että keskustelu leveistä/kapeista altaista saattaa tarvita uutta näkökulmaa yliparametrisoinnin ja redundanssin ympärille, jotka kykenevät luomaan suuria kytkeytyneitä komponentteja maiseman alaosaan. Toiseksi, pienen määrän suurten ominaisarvojen riippuvuus aineiston jakaumasta voidaan liittää mallitulosten gradienttien kovarianssimatriisin spektriin. Tätä silmällä pitäen voimme arvioida uudelleen yhteyksiä mallin data-arkkitehtuuri-algoritmikehyksessä toivoen, että se valaisisi nykyaikaisissa sovelluksissa esiintyvien korkea-ulotteisten ja ei-konvekseja sisältävien tilojen geometriaa. Esitämme erityisesti tapauksen, joka yhdistää nämä kaksi havaintoa: gradienttiin perustuva menetelmä näyttää kiipeävän ensin ylämäkeen ja sitten laskevan alas kahden pisteen välillä; vaikka itse asiassa ne sijaitsevat samassa altaassa.

**Tulos**

Yliparametrisoitujen neuroverkkojen Hessianin empiirinen analyysi

**Esimerkki 2.917**

Menetelmien kehittäminen tekstiyksiköiden informatiivisen sisällön käsittelemiseksi yhteensovitusprosessissa on suuri haaste automaattisissa yhteenvetojen arviointijärjestelmissä, jotka käyttävät kiinteää n-grammin yhteensovitusta. Rajoitus aiheuttaa epätarkkuutta vertailu- ja viiteyhteenvetojen yksiköiden välillä. Tässä tutkimuksessa esitellään uusi avainsanapohjainen tiivistelmien arviointilaite (KpEval) automaattisten tiivistelmien arviointia varten. KpEval luottaa avainsanoihin, koska ne ilmaisevat tekstin tärkeimmät käsitteet. Arviointiprosessissa avainsanoja käytetään niiden lemma-muodossa tekstin vastaavana yksikkönä. Järjestelmää sovellettiin TAC2011-tapahtumassa esitetyn arabialaisen monidokumenttidatan eri tiivistelmien arviointiin. Tulokset osoittivat, että uusi arviointitekniikka korreloi hyvin tunnettujen arviointijärjestelmien kanssa: Rouge-1, Rouge-2, Rouge-SU4 ja AutoSummENG-MeMoG. KpEval korreloi vahvimmin AutoSummENG-MeMoG:n kanssa, Pearsonin ja Spearmanin korrelaatiokertoimet ovat 0,8840 ja 0,9667. Yleiset termit Automaattinen tiivistelmän arviointi, automaattinen tiivistäminen, avainsanojen poiminta, luonnollisen kielen käsittely, laskennallinen kielitiede, tiedonhaku.

**Tulos**

Avainlauseisiin perustuva automaattisen tekstin tiivistämisen arviointi

**Esimerkki 2.918**

Laajennamme regressio-ongelmien tehostamisen teoriaa verkko-oppimisympäristöön. Boostingin eräkohtaisesta asetelmasta yleistettynä heikon oppimisalgoritmin käsite mallinnetaan lineaarisilla tappiofunktioilla toimivaksi online-oppimisalgoritmiksi, joka kilpailee regressiofunktioiden perusluokan kanssa, kun taas vahvan oppimisalgoritmin käsite on sileillä konveksilla tappiofunktioilla toimiva online-oppimisalgoritmi, joka kilpailee laajemman regressiofunktioiden luokan kanssa. Tärkein tuloksemme on online gradientin tehostamisalgoritmi, joka muuntaa heikon onlineoppimisalgoritmin vahvaksi algoritmiksi, jossa funktioiden suurempi luokka on perusluokan lineaarinen alue. Esitämme myös yksinkertaisemman tehostamisalgoritmin, joka muuntaa heikon online-oppimisalgoritmin vahvaksi algoritmiksi, kun suurempi funktioluokka on perusluokan kupera runko, ja todistamme sen optimaalisuuden.

**Tulos**

Online Gradientin tehostaminen

**Esimerkki 2.919**

Gaussin prosessit ovat monipuolisia funktioiden jakaumia, jotka tarjoavat Bayesin ei-parametrisen lähestymistavan tasoittamiseen ja interpolointiin. Esittelemme yksinkertaisia suljetussa muodossa olevia ytimiä, joita voidaan käyttää Gaussin prosessien kanssa kuvioiden löytämiseksi ja ekstrapoloinnin mahdollistamiseksi. Nämä ytimet johdetaan mallintamalla spektritiheys - ytimen Fourier-muunnos - Gaussin seoksella. Ehdotetut ytimet tukevat laajaa luokkaa stationaarisia kovariansseja, mutta Gaussin prosessien päättely pysyy yksinkertaisena ja analyyttisenä. Esittelemme ehdotettuja ytimiä löytämällä kuvioita ja suorittamalla pitkän aikavälin ekstrapolointia synteettisillä esimerkeillä sekä ilmakehän hiilidioksiditrendeillä ja lentomatkustajatiedoilla. Osoitamme myös, että voimme rekonstruoida vakiokovarianssit kehyksessämme.

**Tulos**

Gaussin prosessin kovarianssiytimet kuvioiden löytämiseen ja ekstrapolointiin

**Esimerkki 2.920**

Tarkka ohjelmistokehitystyön arviointi on ratkaisevan tärkeää ohjelmistoprojektien onnistumisen kannalta. Vaikka alan ammattilaiset ovat kehittäneet ja toteuttaneet monia tekniikoita ja algoritmisia malleja, ohjelmistokehitystyön tarkka ennustaminen on edelleen haastava tehtävä ohjelmistosuunnittelun alalla, erityisesti epävarmojen ja epätäsmällisten syötteiden ja rinnakkaisten ominaisuuksien käsittelyssä. Tässä artikkelissa on käytetty älykästä hybridimallia, jossa yhdistetään neuroverkkomalli, joka on integroitu sumeaan malliin (neuro-sumea malli), parantamaan ohjelmistokustannusten arvioinnin tarkkuutta. Ehdotetun mallin suorituskykyä arvioidaan suunnittelemalla ja suorittamalla arviointi julkaistujen hanke- ja teollisuustietojen avulla. Tulokset ovat osoittaneet, että ehdotettu malli pystyy parantamaan arviointitarkkuutta 18 prosentilla MMRE-kriteerin (Mean Magnitude of Relative Error) perusteella.

**Tulos**

ÄLYKÄS HYBRIDIMALLI OHJELMISTOKUSTANNUSTEN ARVIOINTIA VARTEN

**Esimerkki 2.921**

Havaitsemme ensin symbolien jatkuvien vektoriesitysten mahdollisen heikkouden neuraalisessa konekääntämisessä. Toisin sanoen symbolin jatkuva vektoriedustus tai sanan upotusvektori koodaa useita samankaltaisuusulottuvuuksia, mikä vastaa sanan useamman kuin yhden merkityksen koodaamista. Tästä seuraa, että neuraalisen konekääntämisen koodaus- ja dekooderiverkkojen on käytettävä huomattava osa kapasiteetistaan lähde- ja kohdesanojen erotteluun lähdelauseen määrittelemän kontekstin perusteella. Tämän havainnon perusteella ehdotamme tässä artikkelissa sanojen upotusvektoreiden kontekstualisointia käyttämällä lähdelauseen epälineaarista sanapussiesitystä. Lisäksi ehdotamme, että erityismerkit (kuten numerot, varsinaiset substantiivit ja lyhenteet) esitetään kirjoitetuilla symboleilla, jotta voidaan helpottaa sellaisten sanojen kääntämistä, jotka eivät sovellu hyvin käännettäviksi jatkuvien vektorien avulla. En-Fr- ja En-De-kokeet osoittavat, että ehdotetut kontekstualisointi- ja symbolisointimenetelmät parantavat neuraalisten konekäännösjärjestelmien käännöslaatua merkittävästi.

**Tulos**

Kontekstiriippuvainen sanojen esittäminen neuraalista konekääntämistä varten

**Esimerkki 2.922**

Esittelemme useita uusia arkkitehtuurin ominaisuuksia ja koulutusmenetelmiä, joita sovellamme generatiivisiin vastakkaisverkkoihin (Generative Adversarial Networks, GAN). Keskitymme kahteen GAN:ien sovellukseen: puolivalvottuun oppimiseen ja sellaisten kuvien tuottamiseen, joita ihmiset pitävät visuaalisesti realistisina. Toisin kuin useimmissa generatiivisia malleja koskevissa töissä, ensisijaisena tavoitteenamme ei ole kouluttaa mallia, joka antaa suuren todennäköisyyden testidatalle, emmekä vaadi, että malli pystyy oppimaan hyvin ilman merkintöjä. Uusia tekniikoitamme käyttämällä saavutamme huipputuloksia puolivalvotussa luokittelussa MNIST-, CIFAR-10- ja SVHN-luokituksissa. Luodut kuvat ovat korkealaatuisia, kuten visuaalinen Turing-testi vahvistaa: mallimme tuottaa MNIST-näytteitä, joita ihmiset eivät pysty erottamaan todellisesta datasta, ja CIFAR-10-näytteitä, joiden inhimillinen virhetaso on 21,3 prosenttia. Esitämme myös ImageNet-näytteitä, joiden tarkkuus on ennennäkemätön, ja osoitamme, että menetelmiemme avulla malli oppii tunnistettavia ImageNet-luokkien piirteitä.

**Tulos**

Parannetut tekniikat GANien kouluttamiseen

**Esimerkki 2.923**

Sarkasmia, joka johtuu numeeristen osien esiintymisestä tekstissä, on aiemmin mainittu automaattisten sarkasmintunnistusmenetelmien tekemänä virheenä. Esitämme ensimmäisen tutkimuksen, jossa havaitaan sarkasmia numeroissa, kuten lauseessa "Love waking up at 4 am". Analysoimme ongelman haasteita ja esittelemme sääntöpohjaisia, koneoppimisen ja syväoppimisen lähestymistapoja sarkasmin havaitsemiseksi tekstin numeerisissa osissa. Deep Learning -lähestymistapamme päihittää neljä aiempaa sarkasmin havaitsemiseen liittyvää työtä sekä sääntöpohjaiset ja koneoppimisen lähestymistavat twiittien tietokokonaisuudessa, ja F1-tulos on 0,93. Tämä osoittaa, että numeroita sisältävän tekstin erityinen huomioiminen voi olla hyödyllistä sarkasmin havaitsemisen huipputason parantamiseksi.

**Tulos**

"2 tuntia aikaa kirjoittaa essee on hauskaa!": Sarkasmin havaitseminen tekstin numeerisista osista

**Esimerkki 2.924**

Toby Walsh esittää teoksessaan "The Singularity May Never Be Near" kuusi argumenttia tukeakseen näkemystään, jonka mukaan teknologinen singulariteetti voi tapahtua, mutta se on epätodennäköistä. Tässä artikkelissa analysoimme kutakin hänen perusteluistaan ja päädymme samankaltaisiin päätelmiin, mutta annamme enemmän painoarvoa todennäköisyydelle "todennäköisesti tapahtuu".

**Tulos**

Singulariteetti voi olla lähellä

**Esimerkki 2.925**

Nopea kriisitoiminta edellyttää viestien reaaliaikaista analysointia. Katastrofin jälkeen vapaaehtoiset yrittävät luokitella twiittejä määrittääkseen tarpeet, esim. tarvikkeet, infrastruktuurin vahingot jne. Kun käytössä on merkittyjä tietoja, valvottu koneoppiminen voi auttaa viestien luokittelussa. Merkityn datan vähäisyys aiheuttaa huonoa suorituskykyä koneen koulutuksessa. Voidaanko vanhoja twiittejä käyttää uudelleen luokittelijoiden kouluttamiseen? Miten voimme valita merkittyjä twiittejä koulutukseen? Tutkimme erityisesti aiempien tapahtumien merkityn datan hyödyllisyyttä. Auttavatko eri kieliset merkityt twiitit? Tarkkailemme menneistä katastrofeista saatujen koulutusjoukkojen eri yhdistelmillä koulutettujen luokittelijoidemme suorituskykyä. Teemme laajoja kokeiluja todellisilla kriisitietoaineistoilla ja osoitamme, että aiemmat leimatut tiedot ovat hyödyllisiä, kun sekä lähde- että kohdetapahtumat ovat samantyyppisiä (esim. molemmat maanjäristyksiä). Samankaltaisilla kielillä (esim. italia ja espanja) kielialueiden välinen mukauttaminen oli hyödyllistä, mutta eri kielillä (esim. italia ja englanti) suorituskyky heikkeni.

**Tulos**

Kielialueiden välinen mukauttaminen kriiseihin liittyvien lyhyiden viestien luokittelua varten

**Esimerkki 2.926**

Recurrent Neural Networks (RNN) on viime aikoina saavuttanut parhaan suorituskyvyn offline-käsialan tekstintunnistuksessa. Samaan aikaan RNN:n oppiminen gradienttilaskeutumisen avulla johtaa hitaaseen konvergenssiin, ja harjoitteluaika on erityisen pitkä, kun harjoittelutietokanta koostuu kokonaisista tekstiriveistä. Tässä artikkelissa ehdotamme helppoa tapaa nopeuttaa stokastista gradienttilaskeutumista tässä asetelmassa ja yleisemmin sekvenssien tunnistamisen oppimisessa. Periaate on nimeltään opetussuunnitelmaoppiminen eli muotoilu. Ideana on oppia ensin tunnistamaan lyhyitä sekvenssejä ennen kuin harjoitellaan kaikkia käytettävissä olevia harjoitussekvenssejä. Kokeet kolmella eri käsinkirjoitustekstitietokannalla (Rimes, IAM, OpenHaRT) osoittavat, että tämän strategian yksinkertainen toteutus voi nopeuttaa tekstintunnistukseen tarkoitetun RNN:n koulutusta merkittävästi ja jopa parantaa suorituskykyä merkittävästi joissakin tapauksissa.

**Tulos**

Opetussuunnitelman mukainen oppiminen käsinkirjoitetun tekstin rivien tunnistamiseksi

**Esimerkki 2.927**

Syvällisen koneoppimisen menetelmien avulla matalan tason representaatioita voidaan käyttää tehokkaasti uudelleen abstraktimpien korkean tason representaatioiden tuottamiseen. Alun perin syväoppimista on sovellettu passiivisesti (esim. luokittelutarkoituksiin). Viime aikoina sitä on laajennettu arvioimaan autonomisten agenttien toimien arvoa vahvistusoppimisen (RL) puitteissa. Tällaisen arvofunktion täydentämiseksi voidaan oppia eksplisiittisiä malleja ympäristöstä. Vaikka mallipohjaisessa RL:ssä on jo käytetty "litteitä" konnektionistisia menetelmiä, tähän mennessä vain RL:n modevapaita muunnelmia on varustettu syväoppimisen menetelmillä. Ehdotamme syvämallipohjaisen RL:n muunnelmaa, jonka avulla agentti voi oppia mielivaltaisesti abstrakteja hierarkkisia esityksiä ympäristöstään. Tässä artikkelissa esittelemme tutkimusta siitä, miten tällaiset hierarkkiset representaatiot voivat perustua agentin ja sen ympäristön väliseen sensomotoriseen vuorovaikutukseen.

**Tulos**

Hierarkkisten vahvistusoppimismallien perustaminen tiedonsiirtoa varten

**Esimerkki 2.928**

Ehdotamme Edwardia, Turing-täydellistä todennäköisyysohjelmointikieltä. Edward määrittelee kaksi kompositionaalista esitystä - satunnaismuuttujat ja päättely. Käsitellessämme päättelyä ensimmäisen luokan kansalaisena mallintamisen rinnalla osoitamme, että todennäköisyysohjelmointi voi olla yhtä joustavaa ja laskennallisesti tehokasta kuin perinteinen syväoppiminen. Joustavuuden osalta Edwardin avulla on helppo sovittaa sama malli käyttäen erilaisia yhdistettävissä olevia päättelymenetelmiä, jotka vaihtelevat piste-estimaatiosta variationaaliseen päättelyyn ja MCMC:hen. Lisäksi Edward voi käyttää mallinnusrepresentaatiota uudelleen osana päättelyä, mikä helpottaa rikkaiden variaatiomallien ja generatiivisten vastakkaisverkkojen suunnittelua. Tehokkuuden vuoksi Edward on integroitu TensorFlow'hun, mikä tarjoaa huomattavaa nopeutusta nykyisiin todennäköisyysjärjestelmiin verrattuna. Esimerkiksi logistisen regression vertailutehtävässä osoitamme, että Edward on vähintään 35 kertaa nopeampi kuin Stan ja 6 kertaa nopeampi kuin PyMC3. Edward ei myöskään aiheuta mitään yleiskustannuksia: se on yhtä nopea kuin käsinkirjoitettu TensorFlow.

**Tulos**

SYVÄ TODENNÄKÖISYYSOHJELMOINTI

**Esimerkki 2.929**

Etukäteispolariteettileksikoiden johtaminen sentimenttianalyysia varten - jossa positiiviset tai negatiiviset pisteet liitetään sanoihin asiayhteydestä irrotettuna - on haastava tehtävä. Yleensä on vaikea löytää kompromissia tarkkuuden ja kattavuuden välillä, ja se riippuu leksikon rakentamisessa käytetystä menetelmästä. Manuaalisesti annotoidut sanastot tarjoavat korkean tarkkuuden, mutta niiden kattavuus on heikko, kun taas automaattinen johdanto olemassa olevasta tiedosta takaa korkean kattavuuden, mutta tarkkuus on heikompi. Koska ennakkopolariteettien automaattinen johtaminen on vähemmän aikaa vievää kuin manuaalinen annotointi, tällaiset lähestymistavat, erityisesti SentiWordNet-resurssiin perustuvat lähestymistavat, ovat puhjenneet kovaa vauhtia. Tässä artikkelissa vertaamme SentiWordNetiin perustuvia yleisimmin käytettyjä tekniikoita uudempiin tekniikoihin ja yhdistämme ne oppimiskehykseen (ns. ensemble-menetelmä). Hyödyntämällä manuaalisesti rakennettuja ennakkopolariteettileksikoita ensemble-menetelmämme pystyy paremmin ennustamaan tuntemattomien sanojen ennakkoarvon ja päihittää kaikki muut SentiWordNet-menetelmät. Tämän tekniikan avulla olemme rakentaneet noin 155 000 sanan SentiWords-sanakirjan, joka sisältää noin 155 000 sanaa ja jonka tarkkuus ja kattavuus ovat korkeat. Lopuksi osoitamme, että tunneanalyysitehtävissä leksikkomme avulla voimme päihittää sekä SentiWordNetistä johdetut yksittäiset mittarit että suositut manuaalisesti annotoidut tunneanalyysileksiköt.

**Tulos**

SentiWords: Suuri tarkkuus ja kattavuus sanastossa tunneanalyysiä varten: Deriving a High Precision and High Coverage Lexicon for Sentiment Analysis

**Esimerkki 2.930**

Logistista häviöfunktiota suositellaan usein koneoppimisessa ja tilastotieteessä sileäksi ja tiukasti kupera-alaiseksi korvikkeeksi 0-1-häviölle. Tässä artikkelissa tutkimme, onko logistinen häviö näiden sileys- ja kuperaominaisuuksien ansiosta parempi kuin muut yleisesti pidetyt vaihtoehdot, kuten saranahäviö. Osoitamme, että tunnetuista asymptoottisista rajoista poiketen logistinen häviö ei paranna yleistä ei-sileää häviöfunktiota, kuten saranahäviötä, niin kauan kuin ennustus-/optimointi-iteraatioiden määrä on alle eksponentiaalinen. Osoitamme erityisesti, että stokastisen logistisen optimoinnin konvergenssinopeus rajoittuu alhaaltapäin polynomilla päätösjoukon halkaisijan ja ennustamisiteraatioiden lukumäärän suhteen, ja annamme vastaavan tiukan ylärajan. Tämä ratkaisee McMahanin ja Streeterin (2012) avoimen COLT-ongelman.

**Tulos**

Logistinen regressio: Tiukat rajat stokastiselle ja online-optimoinnille∗\*

**Esimerkki 2.931**

Kaikkien aikojen ensimmäinen ihminen vastaan tietokone no-limit Texas hold 'em -kilpailu järjestettiin 24. huhtikuuta - 8. toukokuuta 2015 River's Casinolla Pittsburghissa, PA:ssa. Tässä artikkelissa esittelen ajatuksiani kilpailun suunnittelusta, agenttiarkkitehtuurista ja saaduista kokemuksista.

**Tulos**

Pohdintani ensimmäisestä mies vastaan kone No-Limit Texas Hold 'em -kilpailusta∗\*

**Esimerkki 2.932**

Tutkimme useita kysymyksiä Kalain et al. (2009) luotettavan agnostisen oppimiskehyksen avulla, joka kattaa oppimistehtävät, joissa tietynlainen virhe on kalliimpi kuin toisenlainen. Positiivisesti luotettava luokittelija on sellainen, joka ei tee vääriä positiivisia virheitä. Positiivisen luotettavan agnostisen kehyksen tavoitteena on tuottaa hypoteesi, jolla on seuraavat ominaisuudet: (i) sen väärien positiivisten virheiden määrä on korkeintaan ǫ, (ii) sen väärien negatiivisten virheiden määrä on korkeintaan ǫ suurempi kuin luokan parhaan positiivisen luotettavan luokittelijan. Tähän läheisesti liittyvä käsite on täysin luotettava agnostinen oppiminen, jossa tarkastellaan osittaisluokittelijoita, joiden sallitaan ennustaa "tuntematon" joillekin syötteille. Paras täysin luotettava osittaisluokittelija on sellainen, joka ei tee virheitä ja minimoi todennäköisyyden ennustaa "tuntematon", ja täysin luotettavan oppimisen tavoitteena on tuottaa hypoteesi, joka on lähes yhtä hyvä kuin luokan paras täysin luotettava osittaisluokittelija. Jakaumariippumattomassa oppimisessa tunnetuimmat PAC-oppimisalgoritmit käyttävät tyypillisesti polynomikynnyksisiä esityksiä, kun taas uusimmat agnostiset oppimisalgoritmit käyttävät pistemäisiä polynomiapproksimaatioita. Osoitamme, että yksipuoliset polynomi-approksimaatiot, joka on polynomikynnysrepresentaatioiden ja pistemäisten polynomi-approksimaatioiden välimuoto, riittävät oppimiseen luotettavissa agnostisissa asetuksissa. Tämän jälkeen osoitamme, että enemmistöt voidaan oppia täysin luotettavasti ja enemmistöjen disjunktiot voidaan oppia positiivisesti luotettavasti rakentamalla sopivia yksipuolisia polynomiapproksimaatioita. Täysin luotettava algoritmimme enemmistöjä varten tarjoaa ensimmäisen todisteen siitä, että täysin luotettava oppiminen voi olla ehdottomasti helpompaa kuin agnostinen oppiminen. Algoritmimme täyttävät myös vahvat attribuuttitehokkuusominaisuudet, ja monissa tapauksissa ne tarjoavat sujuvaa kompromissia näytteen monimutkaisuuden ja suoritusajan välillä. Kalifornian yliopisto, Berkeley. Sähköposti: vkanade@eecs.berkeley.edu The Simons Institute for the Theory of Computing at UC Berkeley. Sähköposti: jthaler@seas.harvard.edu

**Tulos**

Jakelusta riippumaton luotettava oppiminen

**Esimerkki 2.933**

Voidaanko monimutkaisia tutkimus- ja hyödyntämisvaihtoehtoja rinnakkaistaa? Esimerkkinä mainittakoon optimaalisen suuritehoisen koesuunnittelun ongelma, jossa haluamme suunnitella peräkkäin koe-eriä, jotta voimme samanaikaisesti oppia korvikefunktion, joka yhdistää ärsykkeen vasteeseen, ja tunnistaa funktion maksimin. Formalisoimme tehtävän monihaarainen bandiitti -ongelmaksi, jossa tuntematon voittofunktio poimitaan Gaussin prosessista (GP), ja yksittäisen haaran sijasta vedämme jokaisella kierroksella usean haaran erän rinnakkain. Kehitämme GP-BUCB:n, periaatteellisen algoritmin erien valintaan, joka perustuu GP-UCB-algoritmiin peräkkäistä GP-optimointia varten. Todistamme yllättävän tuloksen; verrattuna peräkkäiseen lähestymistapaan rinnakkaisen algoritmin kumulatiivinen katumus kasvaa vain vakiokertoimella, joka on riippumaton erän koosta B. Tuloksemme tarjoavat tiukkaa teoreettista tukea rinnakkaisuuden hyödyntämiselle Bayesin globaalissa optimoinnissa. Osoitamme lähestymistapamme tehokkuuden kahdessa reaalimaailman sovelluksessa.

**Tulos**

Exploration-Exploitation Tradeoffs -vaihtoehtojen rinnakkaistaminen Gaussin prosessin Bandit-optimoinnilla

**Esimerkki 2.934**

Hyödyllisen tiedon oppiminen pitkien aikaviiveiden yli on kriittinen ja vaikea ongelma ajallisille neuromalleille tehtävissä, kuten kielen mallintamisessa. Nykyiset arkkitehtuurit, jotka käsittelevät tätä ongelmaa, ovat usein monimutkaisia ja kalliita kouluttaa. Delta-RNN (Delta Recurrent Neural Network) on yksinkertainen ja suorituskykyinen rakenne, joka yhdistää aiemmin ehdotetut ajalliset neuromallit. DeltaRNN-mallit ylläpitävät pidempiaikaista muistia oppimalla interpoloimaan nopeasti muuttuvan datapohjaisen esityksen ja hitaasti muuttuvan, implisiittisesti vakaan tilan välillä. Tämä vaatii tuskin yhtään enempää parametreja kuin klassinen yksinkertainen rekursiivinen verkko. Mallit päihittävät suositut monimutkaiset arkkitehtuurit, kuten LSTM (Long Short Term Memory) ja GRU (Gated Recurrent Unit), ja ne saavuttavat huippuluokan suorituskyvyn kielen mallintamisessa merkki- ja sanatasolla ja tuottavat vertailukelpoisen suorituskyvyn sanan alitason mallintamisessa.

**Tulos**

Yksinkertaisempien kielimallien oppiminen Delta-rekurrenssia neuroverkkokehyksen avulla

**Esimerkki 2.935**

Vastausjoukkojen ohjelmointi (ASP) on vakiintunut formalismi ei-monotoniseen päättelyyn. ASP-ohjelmalla ei voi olla vastausjoukkoa syklisen oletusnegaation vuoksi. Tällöin ei ole mahdollista tehdä mitään johtopäätöstä, vaikka sitä ei olisikaan tarkoitettu. Viime aikoina on ehdotettu useita parakoherentteja semantiikoita, jotka käsittelevät tätä ongelmaa, ja näiden semantiikoiden mahdollisia sovelluksia on tunnistettu useita. Parakoherentteja semantiikoita ei kuitenkaan ole voitu soveltaa käytännössä tehokkaiden algoritmien ja toteutusten puutteen vuoksi. Tässä artikkelissa puututaan tähän puutteeseen, ja ehdotetaan useita erilaisia algoritmeja puolistabiilien ja puolitasapainomallien laskemiseen ja toteutetaan ne vastausjoukon ratkaisukehykseen. Uusien algoritmien suorituskykyä verrataan empiirisesti ASP-kilpailujen vertailuarvoilla.

**Tulos**

Parakoherenttien vastausjoukkojen laskemisesta

**Esimerkki 2.936**

Syvät neuroverkot ovat aiheuttaneet vallankumouksen monissa luokittelutehtävissä. Tähän liittyvä meneillään oleva vallankumous - jota ei myöskään teoreettisesti ymmärretä - koskee niiden kykyä toimia geneerisinä malleina monimutkaisille tietotyypeille, kuten kuville ja teksteille. Näitä malleja koulutetaan käyttämällä sellaisia ideoita kuin varioivia autokoodereita ja generatiivisia adversaaliverkkoja (Generative Adversarial Networks). Selitämme ensimmäisen kerran monikerroksisten verkkojen ilmaisukykyä antamalla riittävän kriteerin sille, että funktio on approksimoitavissa neuroverkolla, jossa on n piilokerrosta. Keskeinen tekijä on Barronin teoreema [Bar93], joka antaa Fourierin kriteerin funktion approksimoitavuudelle neuroverkolla, jossa on yksi piilokerros. Osoitamme, että n funktion, jotka täyttävät tietyt Fourier-ehdot ("Barronin funktiot"), koostumus voidaan approksimoida n+ 1-kerroksisella neuroverkolla. Todennäköisyysjakaumien osalta tämä tarkoittaa kriteeriä, jonka mukaan todennäköisyysjakauma on lähestyttävä Wassersteinin etäisyydellä - joka on luonnollinen metriikka todennäköisyysjakaumille - neuroverkolla, jota sovelletaan kiinteään perusjakaumaan (esim. monimuuttujaiseen gaussiseen). Viimeaikaisen alarajoja koskevan työn pohjalta annamme myös esimerkkifunktion, joka osoittaa, että Barronin funktioiden koostumus on ilmaisuvoimaisempi kuin pelkät Barronin funktiot.

**Tulos**

Neuroverkkojen kyvystä ilmaista jakaumia

**Esimerkki 2.937**

Tässä artikkelissa esitellään laskennallisen logiikan perusteet toimijuuden mallille, jota kutsutaan KGP-malliksi (Knowledge, Goals and Plan). Tämä malli mahdollistaa heterogeenisten agenttien määrittelyn, jotka voivat olla vuorovaikutuksessa toistensa kanssa ja jotka voivat käyttäytyä sekä ennakoivasti että reaktiivisesti, jolloin ne voivat toimia dynaamisissa ympäristöissä mukauttamalla tavoitteitaan ja suunnitelmiaan, kun tällaisissa ympäristöissä tapahtuu muutoksia. KGP tarjoaa erittäin modulaarisen agenttiarkkitehtuurin, joka integroi kokoelman päättely- ja fyysisiä kykyjä, jotka on koottu siirtymiin, jotka päivittävät agentin tilaa päättelyn, havaitsemisen ja toiminnan seurauksena. Siirtymiä organisoidaan sykliteorioilla, jotka määrittelevät siirtymien suoritusjärjestyksen ottaen huomioon dynaamisen kontekstin ja agentin mieltymykset, sekä valintaoperaattoreilla, jotka antavat syötteitä siirtymille.

**Tulos**

KGP-agenttien laskennallisen logiikan perusteet

**Esimerkki 2.938**

Submodulaaristen funktioiden konveksisten relaksaatioiden avulla saatujen strukturoitujen rangaistusten proksimaaliongelman tiedetään vastaavan separoituvien konveksisten funktioiden minimointia vastaavilla submodulaarisilla polyedereillä. Tässä artikkelissa paljastamme kattavan luokan strukturoituja seuraamuksia, joiden seuraamuksille tämä ongelma voidaan ratkaista tehokkaasti ratkaistavan parametrisen maksimivirtaoptimoinnin luokan avulla. Tämän jälkeen osoitamme, että Gallon et al. [17] ehdottama parametrinen maxflow-algoritmi ja sen muunnokset, joka toimii pahimmassa tapauksessa vain vakiokertoimen kustannuksella vastaavan maxflow-optimoinnin yhdestä laskutoimituksesta, voidaan sovittaa ratkaisemaan näiden rangaistusten proksimaaliset ongelmat. Useat olemassa olevat strukturoidut rangaistukset täyttävät nämä ehdot; näin ollen regularisoitu oppiminen näillä rangaistuksilla on ratkaistavissa nopeasti käyttämällä parametrista maxflow-algoritmia. Tutkimme myös ehdotetun kehyksen empiiristä suoritusajan suorituskykyä.

**Tulos**

Parametriset maksimivirrat strukturoituun harvaan oppimiseen submodulaaristen funktioiden konvekseilla relaksaatioilla

**Esimerkki 2.939**

Cloze-tyylinen luetun ymmärtäminen on edustava ongelma asiakirjan ja kyselyn välisen suhteen etsimisessä. Tässä artikkelissa esitellään yksinkertainen mutta uudenlainen malli nimeltä attention-over-attention reader, jonka avulla voidaan ratkaista paremmin cloze-tyylinen luetun ymmärtämistehtävä. Mallimme tarkoituksena on asettaa toinen huomiomekanismi dokumenttitason huomion päälle ja saada aikaan "tarkkaavaisuus" lopullisia vastausennusteita varten. Mallimme etuna on se, että se on yksinkertaisempi kuin vastaavat teokset ja antaa samalla erinomaisen suorituskyvyn. Ehdotamme myös N-best-uudelleenjärjestämisstrategiaa, jolla ehdokkaiden paikkansapitävyys voidaan tarkistaa ja suorituskykyä parantaa entisestään. Kokeelliset tulokset osoittavat, että ehdotetut menetelmät päihittävät huomattavasti eri huipputason järjestelmät julkisissa tietokokonaisuuksissa, kuten CNN- ja Children's Book Test -tietokannoissa.

**Tulos**

Tarkkaavaisuus yli tarkkaavaisuuden - Neuraaliverkot lukemisen ymmärtämisessä

**Esimerkki 2.940**

Tässä artikkelissa astutaan perinteisten NLP-tehtävien, kuten automaattisen puheentunnistuksen (ASR) ja konekääntämisen (MT), mukavuusalueen ulkopuolelle ja käsitellään kahta uutta ongelmaa, jotka liittyvät automaattiseen monikieliseen uutisseurantaan: TV- ja radio-ohjelmien ASR-äänikirjoitusten segmentointi yksittäisiksi tarinoiksi ja eri lähteistä ja eri kielistä tulevien yksittäisten tarinoiden ryhmittely tarinalinjoiksi. Samoja tapahtumia käsittelevien tarinoiden klusterointi on olennainen tehtävä tiedotusvälineiden seurannassa. Ratkaisemme nämä kaksi ongelmaa yhdessä hyödyntämällä sekvenssistä sekvenssiin -neuraalisten käännösmallien matalaulotteisia semanttisia esityskykyjä. Morfologisesti rikkaiden kielten monikielisen neuraalisen käännöksen yhteisen monitehtäväoppimisen mahdollistamiseksi korvaamme huomiomekanismin liukuikkunamekanismilla ja käytämme sekvenssistä sekvenssiin -neuraalista käännösmallia merkkitasolla sanatason sijaan. Tarinan segmentointi- ja tarinan klusterointiongelma ratkaistaan tarkastelemalla neuraalisen käännösprosessin sivutuotteena syntyviä matalaulotteisia vektoreita. Tämän artikkelin tulokset kuvaavat uudenlaista lähestymistapaa automaattiseen tarinan segmentointi- ja juoniklusterointiongelmaan.

**Tulos**

Merkkitason neuraalinen kääntäminen monikielisen median seurannassa SUMMA-hankkeessa

**Esimerkki 2.941**

Tässä asiakirjassa kuvataan USTC:n NELSLIP-järjestelmiä, jotka lähetettiin vuoden 2016 TAC:n Knowledge Base Population (KBP) -kilpailujen kolmikielisen entiteettien havaitsemisen ja linkittämisen (EDL) raidalle. Olemme rakentaneet kaksi järjestelmää entiteettien löytämiseen ja mainintojen havaitsemiseen (MD): toisessa käytetään ehdollista RNNLM:ää ja toisessa huomiopohjaista enkooderi-dekooderi-kehystä. Entiteettien yhdistämisjärjestelmä (EL) koostuu kahdesta moduulista: sääntöpohjaisesta ehdokkaiden luomisesta ja neuroverkkojen todennäköisyysjärjestysmallista. Lisäksi NIL-klusteroinnissa käytetään joitakin yksinkertaisia merkkijonojen täsmäytyssääntöjä. Paras järjestelmämme saavutti lopulta F1-pistemäärän 0,624 end-to-end typed mention ceaf plus -mittarilla.

**Tulos**

USTC:n NELSLIP-järjestelmät kolmikielisten entiteettien havaitsemiseen ja linkitystehtäviin TAC KBP 2016 -tapahtumassa.

**Esimerkki 2.942**

Sanat voidaan esittää koostamalla sanan alayksiköiden, kuten sanasegmenttien, merkkien ja/tai merkkien n-grammien, esityksiä. Vaikka tällaiset representaatiot ovat tehokkaita ja voivat kuvata sanojen morfologisia säännönmukaisuuksia, niitä ei ole vertailtu järjestelmällisesti, eikä tiedetä, miten ne ovat vuorovaikutuksessa eri morfologisten typologioiden kanssa. Esitämme kielen mallinnustehtävässä kokeita, joissa systemaattisesti vaihdellaan (1) esityksen perusyksikköä, (2) näiden esitysten koostumusta ja (3) mallinnettavan kielen morfologista typologiaa. Tuloksemme vahvistavat suurelta osin aiemmat havainnot siitä, että merkkiedustukset ovat tehokkaita monissa kielissä, vaikka havaitsimmekin, että bi-LSTM:ien avulla koostettujen merkkien trigramiedustusten aiemmin tutkimaton yhdistelmä on parempi kuin useimmat muut asetukset. Löydämme kuitenkin myös parantamisen varaa: hahmomallit eivät yllä ennustustarkkuuteen, joka on sama kuin mallilla, jolla on pääsy eksplisiittisiin morfologisiin analyyseihin.

**Tulos**

Hahmoista sanoihin ja niiden välissä: Saammeko morfologiaa talteen?

**Esimerkki 2.943**

Ehdotetaan uus-sumean neuronin muunnosta (laajennettu uus-sumea neuroni (ENFN)), jolle on ominaista paremmat approksimointiominaisuudet. Ehdotetaan adaptiivista oppimisalgoritmia, jolla on sekä seuranta- että tasoitusominaisuuksia ja joka ratkaisee epästationaaristen, "meluisten" stokastisten ja kaoottisten signaalien ennustus-, suodatus- ja tasoitustehtäviä. ENFN:n erityispiirteenä on sen laskennallinen yksinkertaisuus verrattuna muihin keinotekoisiin neuroverkkoihin ja neurofuzzy-järjestelmiin.

**Tulos**

Laajennettu neo-sumea neuroni ja sen mukautuva oppimisalgoritmi

**Esimerkki 2.944**

Luokittelu on laajalti käytetty tekniikka tiedonlouhinnan alalla, jossa suurten tietokantojen luokittelualgoritmeihin liittyvät välittömät ongelmat ovat skaalautuvuus ja tehokkuus. Ehdotamme parannuksia nykyiseen C4.5-päätöspuualgoritmiin. Tässä asiakirjassa attribuuttipohjainen induktio (AOI) ja relevanssianalyysi yhdistetään käsitehierarkian tietämykseen ja HeightBalancePriority-algoritmiin päätöspuun rakentamiseksi yhdessä monitasoisen louhinnan kanssa. Prioriteetit attribuuteille määritetään arvioimalla informaatioentropiaa eri abstraktiotasoilla päätöspuun rakentamiseksi HeightBalancePriority-algoritmia käyttäen. Muunneltuja DMQL-kyselyjä käytetään ymmärtämään ja tutkimaan C4.5-luokittimen tuottamien päätöspuiden puutteita koulutustietokannassa, ja tuloksia verrataan ehdotettuun lähestymistapaan.

**Tulos**

HYÖDYLLISTEN SÄÄNTÖJEN POIMIMINEN PARANTAMALLA PÄÄTÖSPUUINDUKTIOTA KÄYTTÄMÄLLÄ INFORMAATIOENTROPIAA.

**Esimerkki 2.945**

Koska päivittäin syntyy paljon uutta tietoa, tietämysgraafissa olevan tiedon paikkansapitävyyden määrittäminen ja puuttuvien osien täyttäminen ovat ratkaisevan tärkeitä tehtäviä monille tutkijoille ja käytännön toimijoille. Tähän haasteeseen vastaamiseksi on kehitetty useita tietämysgraafien täydennysmenetelmiä, joissa käytetään matalaulotteisia graafisulautumia. Vaikka tutkijat jatkavat näiden mallien parantamista yhä monimutkaisemman ominaisuusavaruuden avulla, osoitamme, että yksinkertaisilla muutoksilla taustalla olevan mallin arkkitehtuurissa voidaan päihittää uusimmat mallit ilman monimutkaista ominaisuuksien suunnittelua. Tässä työssä esittelemme ProjE-nimisen jaetun muuttujan neuroverkkomallin, joka täyttää tietämysgraafin puuttuvan tiedon oppimalla tietämysgraafin entiteettien ja särmien yhteisiä upotuksia ja tekemällä hienovaraisia mutta tärkeitä muutoksia tavalliseen häviöfunktioon. ProjE:n parametrien koko on pienempi kuin 11:llä 15:stä olemassa olevasta menetelmästä, ja sen suorituskyky on 37 prosenttia parempi kuin nykyisen parhaan menetelmän suorituskyky vakiotietoaineistoissa. Osoitamme myös uuden faktantarkistustehtävän avulla, että ProjE pystyy määrittämään tarkasti monien deklaratiivisten väitteiden todenperäisyyden. Tietograafeista (Knowledge Graphs, KG) on tullut keskeinen resurssi monissa koneoppimisen, tiedonlouhinnan ja tekoälysovellusten tehtävissä, kuten kysymysten vastaamisessa [34], entiteettien disambiguoinnissa [7], nimettyjen entiteettien yhdistämisessä [14], faktojen tarkistamisessa [32] ja linkkien ennustamisessa [28], muutamia mainitakseni. Näkemyksemme mukaan KG:t ovat esimerkki heterogeenisesta tietoverkosta, joka sisältää entiteettisolmuja ja suhdesolmuja, jotka vastaavat RDF-tyylisiä kolmioita 〈h, r, t〉, joissa h edustaa pääentiteettiä ja r on suhde, joka yhdistää h:n loppuentiteettiin t. KG:itä käytetään laajalti monissa käytännön tehtävissä, mutta niiden virheettömyyttä ja täydellisyyttä ei kuitenkaan taata. Sen vuoksi on tarpeen kehittää tietämysgraafien täydentämismenetelmiä (KGC) puuttuvien tai virheellisten suhteiden löytämiseksi, jotta KG:iden yleistä laatua voitaisiin parantaa, minkä avulla voidaan puolestaan parantaa tai luoda mielenkiintoisia sovelluksia. KGC-tehtävä voidaan jakaa kahteen osatehtävään, jotka eivät sulje toisiaan pois: i) olioiden ennustaminen ja ii) suhteiden ennustaminen. Entiteettien ennustamistehtävä ottaa syötteenä osittaiskolmikon 〈h, r, ?〉 ja tuottaa tulosteena järjestetyn luettelon ehdolla olevista entiteeteistä: Määritelmä 1. (Entiteettien ranking-ongelma) Kun tietämysgraafi G = {E,R} ja syötekolmio 〈h, r, ?〉, entiteettien ranking-ongelma yrittää löytää optimaalisen järjestetyn listan siten, että ∀ej∀ei ((ej ∈ E- ∧ ei ∈ E+)→ ei ≺ ej), missä E+ = {e ∈ {e1, e2, . . . . . , el}|〈h, r, e〉 ∈ G} ja E- = {e ∈ {el+1, el+2, . . . . , e|E|}|〈h, r, e〉 / ∈ G}. Pää- ja häntä-entiteettien erottaminen toisistaan on yleensä mielivaltaista, joten voimme helposti korvata 〈h, r, ?〉:n 〈?, r, t〉:lla. Suhteiden ennustamistehtävässä pyritään löytämään järjestetty luettelo suhteista, jotka yhdistävät pää-entiteetin ja häntä-entiteetin, eli 〈h, ?, t〉. Keskustellessamme tämän työn yksityiskohdista keskitymme erityisesti entiteettien ennustamistehtävään; on kuitenkin helppoa mukauttaa menetelmä suhteiden ennustamistehtävään muuttamalla syötettä. Viime vuosina on kehitetty useita KGC-algoritmeja, ja menestyneimmillä malleilla on yksi yhteinen piirre: ne käyttävät matalaulotteisia upotusvektoreita entiteettien ja suhteiden esittämiseen. Monet upotusmallit, esimerkiksi Unstructured [3], TransE [4], TransH [35] ja TransR [25], käyttävät marginaaliin perustuvaa pareittaista ranking-menetysfunktiota, joka mittaa kunkin mahdollisen tuloksen pistemäärää h+ r:n ja t:n välisenä Ln-etäisyytenä. Näissä malleissa menetysfunktiot ovat kaikki samoja, joten mallit eroavat toisistaan sen suhteen, miten ne muuntavat 1 ar X iv :1 61 1. 05 42 5v 1 [ cs .A I] 1 6 N ov 2 01 6 entiteettien upotukset h ja t suhteessa suhteiden upotuksiin r. Pelkän h + r:n lisäämisen sijaan Knowledge Vault [8] ja HolE [29] oppivat ilmaisuvoimaisempia yhdistelmäoperaattoreita ennustamaan 〈h, r, t〉 olemassaoloa KG:ssä. Muut mallit, kuten Neural Tensor Network (NTN) [33] ja Compositional Vector Space Model (CVSM) [27], sisällyttävät monikerroksisen neuroverkkoratkaisun olemassa oleviin malleihin. Valitettavasti nämä mallit joko (i) eivät skaalautu hyvin tai (2) tarkastelevat vain yhtä suhdetta kerrallaan [10], mikä rajoittaa niiden käyttökelpoisuutta suurissa reaalimaailman KG:ssä. Huolimatta mallin suuresta koosta edellä mainitut menetelmät käyttävät vain singleton-kolmioita, eli pituudeltaan 1-pituisia polkuja KG:ssä. PTransE [24] ja RTransE [10] käyttävät laajennettua polkutietoa 2- ja 3-hop poluista tietämysgraafin yli. Näillä laajennetuilla malleilla saavutetaan erinomainen suorituskyky syöttötietojen runsauden ansiosta; valitettavasti niiden mallien koko kasvaa eksponentiaalisesti polkujen pituuden kasvaessa, mikä pahentaa entisestään skaalautuvuusongelmia, jotka liittyvät taustalla olevien mallien parametrien jo ennestään suureen määrään. Toinen mielenkiintoinen havainto on, että jotkin nykyisistä malleista eivät ole itsenäisiä malleja, eli ne vaativat valmiiksi koulutettuja KG-sulkeumia (RTransE, CVSM), valmiiksi valittuja polkuja (PTransE, RTransE) tai kunkin solmun valmiiksi laskettuja sisällön sulkeumia (DKRL [36]), ennen kuin mallin kouluttaminen voidaan edes aloittaa. TransR ja TransH ovat itsenäisiä malleja, mutta niiden kokeiluissa raportoidaan vain tuloksia, joissa käytetään syötteenä valmiiksi koulutettuja TransE:n upotuksia. Nämä seikat huomioon ottaen tässä työssä mietimme uudelleen joitakin aiempien mallien tekemiä peruspäätöksiä luodaksemme KGC:lle projektion upotusmallin (ProjE). ProjE-mallissa on neljä osaa, jotka erottavat sen aiheeseen liittyvistä töistä: 1. Sen sijaan, että mittaisimme syötekolmikon 〈h, r, ?〉 ja entiteettiehdokkaiden etäisyyttä yhtenäisellä tai suhteeseen liittyvällä tasolla, päätämme projisoida entiteettiehdokkaat kohdevektoriin, joka edustaa syötetietoa. 2. Toisin kuin nykyiset mallit, jotka käyttävät muunnosmatriiseja, me yhdistämme tulotietoja edustavat upotusvektorit kohdevektoriksi käyttämällä opittavissa olevaa yhdistelyoperaattoria. Näin vältetään suuren määrän muunnosmatriisien lisääminen käyttämällä entiteetti-enkeddaatioita uudelleen. 3. Sen sijaan, että optimoisimme marginaaliin perustuvaa pareittaista luokitteluhäviötä, optimoimme ehdokas-entiteettien (tai suhteiden) luettelon luokitteluhäviötä kollektiivisesti. Lisäksi käytämme ehdokkaiden otantaa hyvin suurten tietokokonaisuuksien käsittelyyn. 4. Toisin kuin monet vastaavat mallit, jotka vaativat esivalmisteltuja tietoja edeltävistä malleista tai tutkivat kalliita monipolkuisia polkuja tietämysgraafin läpi, ProjE on itsenäinen malli, joka kattaa pituus-1-särmiä.

**Tulos**

ProjE: Upotusprojektio tietämysgraafin täydentämistä varten

**Esimerkki 2.946**

Generatiiviset mallit, jotka määrittelevät yhteisiä jakaumia jäsennyspuiden ja lauseiden yli, ovat hyödyllisiä jäsennyksessä ja kielen mallintamisessa, mutta ne asettavat rajoituksia ominaisuuksien laajuudelle ja ovat usein syrjiviä malleja suorituskykyisempiä. Ehdotamme jäsennystä ja kielen mallintamista varten kehystä, jossa yhdistetään generatiivinen malli ja diskriminatiivinen tunnistusmalli koodaaja-dekooderiympäristössä. Esitämme odotusten maksimointiin ja variationaaliseen päättelyyn perustuvia tulkintoja kehyksestä ja osoitamme, että se mahdollistaa jäsennyksen ja kielen mallintamisen yhdessä toteutuksessa. Englanninkielisellä Penn Treenbank -kielipankilla kehyksemme saavuttaa kilpailukykyisen suorituskyvyn konstituenttien jäsennyksessä ja vastaa samalla yhden mallin kielimallinnuksen huipputason tuloksia.

**Tulos**

Generatiivinen jäsennin ja diskriminoiva tunnistusalgoritmi

**Esimerkki 2.947**

Skaalautuva ja tehokas tutkimus on edelleen keskeinen haaste vahvistusoppimisessa (RL). Vaikka on olemassa menetelmiä, joilla on optimaalisuustakuu diskreeteissä tila- ja toiminta-avaruuksissa, näitä menetelmiä ei voida soveltaa korkea-ulotteisissa syvissä RL-skenaarioissa. Näin ollen useimmat nykyaikaiset RL-menetelmät perustuvat yksinkertaisiin heuristiikkoihin, kuten -greedy-etsintään tai Gaussin kohinan lisäämiseen ohjaukseen. Tässä artikkelissa esitellään Variational Information Maximizing Exploration (VIME), eksploraatiostrategia, joka perustuu agentin uskomusta ympäristön dynamiikasta koskevan informaatiovoiton maksimointiin. Ehdotamme käytännön toteutusta, jossa käytetään Bayesin neuroverkoissa variationaalista päättelyä, joka käsittelee tehokkaasti jatkuvia tila- ja toiminta-avaruuksia. VIME muuttaa MDP:n palkitsemisfunktiota, ja sitä voidaan soveltaa useiden eri RL-algoritmien kanssa. Osoitamme, että VIME:llä saavutetaan huomattavasti parempi suorituskyky kuin heuristisilla tutkimusmenetelmillä useissa jatkuvissa ohjaustehtävissä ja algoritmeissa, myös tehtävissä, joissa palkkiot ovat hyvin harvat.

**Tulos**

Uteliaisuuteen perustuva tutkiminen syvässä vahvistusoppimisessa Bayesin neuroverkkojen avulla

**Esimerkki 2.948**

Latentit voimamallit (LFM) ovat joustavia malleja, joissa yhdistyvät mekanistisen mallintamisen periaatteet (eli fysikaaliset mallit) ja ei-parametriset tietoon perustuvat komponentit. Useissa LFM-mallien keskeisissä sovelluksissa tarvitaan epälineaarisuutta, mikä johtaa analyyttisesti vaikeasti ratkaistavissa olevaan päättelyyn. Tässä työssä osoitamme, miten epälineaariset LFM:t voidaan esittää epälineaarisina valkean kohinan ohjaamina tila-avaruusmalleina, ja esittelemme tehokkaan epälineaariseen Kalman-suodatukseen ja tasoitukseen perustuvan menetelmän likimääräistä tila- ja parametri-inferenssiä varten. Havainnollistamme ehdotetun menetelmän suorituskykyä kahden simuloidun esimerkin avulla ja sovellamme sitä reaalimaailman ongelmaan, joka koskee GPS-satelliittien kiertoratojen pitkän aikavälin ennustamista.

**Tulos**

Tila-avaruuden päättely ei-lineaarisille latenttivoimamalleille ja soveltaminen satelliitin kiertoradan ennustamiseen

**Esimerkki 2.949**

Tietojärjestelmän attribuuttihahmojen tutkimus, joka sisältää keskeiset, välttämättömät ja tarpeettomat attribuutit, on attribuuttien vähentämisen perusasia ja tärkeä kysymys. Monet attribuuttihahmojen arviointimenetelmät perustuvat kohteiden ja attribuuttien väliseen suhteeseen. Tässä asiakirjassa ehdotetaan attribuuttihahmojen arviointia varten uudenlaisia arviointiteorioita, jotka perustuvat täysin attribuuttien väliseen suhteeseen. Menetelmässä vertaillaan kahta uutta attribuuttijoukkoa E(a) ja N(a) tässä työssä ehdotetun määritellyn attribuutin a suhteen. Päättelemme, että se, mihin attribuutin tyyppiin a kuuluu, määräytyy pohjimmiltaan E(a):n ja N(a):n välisen suhteen perusteella. Toiseksi annetaan tiiviimpiä ja selkeämpiä tuloksia attribuuttihahmojen arvostelusta analysoimalla tarkennuksen ja täsmällisen tarkennuksen ominaisuuksia E(a):n ja N(a):n välillä topologiassa. Lisäksi keskustellaan attribuuttien välisistä suhteista, jotka ovat hyödyllisiä reduktioiden rakentamisessa tämän asiakirjan viimeisessä jaksossa. Viimeisessä osassa ehdotetaan E(a):een perustuvaa reduktioalgoritmia, ja tämä algoritmi on edellä esitetyn attribuuttihahmojen analyysin laajennettu sovellus.

**Tulos**

Tietojärjestelmän attribuuttihahmoja koskevat uudenlaiset arviointiteoriat.

**Esimerkki 2.950**

Avoimen tiedonlouhinnan (Open Information Extraction, OIE) tavoitteena on poimia pintasuhteita ja niiden argumentteja luonnollisesta tekstistä valvomattomasti ja toimialariippumattomasti. Tässä artikkelissa tutkitaan, miten liian spesifisiä uuttoja voidaan vähentää OIE-järjestelmissä tuottamatta epäinformatiivisia tai epätarkkoja tuloksia. Ehdotamme MinIE:tä, OIE-järjestelmää, joka tuottaa minimoituja, kommentoituja otteita. MinIE-järjestelmän ytimessä on OIE-uutteiden uudelleenkirjoittaminen (1) tunnistamalla ja poistamalla osat, joita pidetään liian spesifisinä; (2) esittämällä polariteettia, modaliteettia, attribuutiota ja määriä koskevat tiedot sopivilla annotaatioilla varsinaisen uutteen sijasta. Suoritimme kokeellisen tutkimuksen useilla todellisilla tietokokonaisuuksilla ja havaitsimme, että MinIE saavuttaa kilpailukykyisen tai korkeamman tarkkuuden ja palautuksen kuin useimmat aiemmat järjestelmät ja tuottaa samalla paljon lyhyempiä otteita.

**Tulos**

MinIE: Faktojen minimointi avoimessa tiedonlouhinnassa (MinIE: Minimizing Facts in Open Information Extraction)

**Esimerkki 2.951**

Joissakin luonnollisissa kielissä sanoilla voi olla koosterakenne. Tämän rakenteen elementtejä ovat juuri (joka voi olla myös komposiitti), etuliitteet ja suffiksit, joilla voidaan ilmaista erilaisia vivahteita ja suhteita muihin sanoihin. Näin ollen sanan sisäinen rakenne on otettava huomioon, jotta sanasta voidaan muodostaa asianmukainen esitys. Tekstikorpuksesta poimitaan joukko usein esiintyviä osasanoja ja näistä valitaan mallit eli osasanat, jotka sisältävät tietoa merkkien n-grammien säännönmukaisuuksista. Valinta tehdään käyttämällä kuvioihin perustuvaa Conditional Random Field -mallia [23,19], jossa on l1-regularisointi. Lisäksi jokaisesta sanasta muodostetaan uusi sarja kuvioiden aakkosista. Uuden aakkoston symbolit rajoittavat paikallisen tilastollisen kontekstin vahvemmin kuin merkit, joten ne mahdollistavat paremmat esitykset R:ssä ja ovat parempia rakennuspalikoita sanan esittämistä varten. Alasanatietoisen kielen mallintamisen tehtävässä kuvioihin perustuvat mallit ovat 2-20 perpleksisyyspistettä parempia kuin merkkeihin perustuvat analogit. Myös rekursiivinen neuroverkko, jossa sana esitetään sen kuvioiden upotusten summana, on kilpailukykyisen ja huomattavasti kehittyneemmän merkkipohjaisen konvoluutioarkkitehtuurin veroinen.

**Tulos**

Kuviot verrattuna merkkeihin sanatietoisessa neuraalisessa kielimallinnuksessa (Subword-aware Neural Language Modeling)

**Esimerkki 2.952**

Suosittelu- ja yhteissuodatusjärjestelmät ovat tärkeitä nykyaikaisissa tieto- ja sähköisen kaupankäynnin sovelluksissa. Koska nämä järjestelmät ovat yhä suositumpia alalla, niiden tuotokset voivat vaikuttaa liiketoimintaa koskevaan päätöksentekoon, mikä kannustaa vastapuolia vaarantamaan tällaisten järjestelmien saatavuuden tai eheyden. Esittelemme yhteissuodatusjärjestelmiin kohdistuvan tietomyrkytyshyökkäyksen. Osoitamme, miten tehokas hyökkääjä, jolla on täydet tiedot oppijasta, voi tuottaa haitallisia tietoja maksimoidakseen haitalliset tavoitteensa ja samalla jäljitellä normaalia käyttäjäkäyttäytymistä välttääkseen havaitsemisen. Vaikka oletus täydellisestä tietämyksestä vaikuttaa äärimmäiseltä, se mahdollistaa vankan arvion yhteissuodatusjärjestelmien haavoittuvuudesta erittäin motivoituneille hyökkäyksille. Esitämme tehokkaita ratkaisuja kahdelle suositulle faktorointiin perustuvalle yhteissuodatusalgoritmille: vaihtoehtoiselle minimointimuodostelmalle ja ydinnormin minimointimenetelmälle. Lopuksi testaamme ehdotettujen algoritmien tehokkuutta reaalimaailman tiedoilla ja keskustelemme mahdollisista puolustusstrategioista.

**Tulos**

Tietojen myrkytyshyökkäykset tekijöihin perustuvaan yhteissuodatukseen

**Esimerkki 2.953**

Esitämme tehokkaan menetelmän löysästi skaalattujen rakenteellisten SVM:ien kouluttamiseen. Vaikka eniten rikkovan etiketin löytäminen marginaalin suhteen skaalautuvassa muodossa on usein helppoa, koska kohdefunktio hajoaa rakenteen suhteen, tämä ei päde slack-skaalautuvassa muodossa, ja eniten rikkovan etiketin löytäminen voi olla hyvin vaikeaa. Keskeinen panoksemme on tehokas menetelmä, jolla löydetään eniten rikkova merkintä slack-rescaloidussa formulaatiossa, kun on annettu oraakkeli, joka palauttaa eniten rikkovan merkinnän (hieman muunnetussa) marginaalin rescaloidussa formulaatiossa. Osoitamme, että menetelmämme mahdollistaa tarkan ja skaalautuvan koulutuksen slack-rescaled SVM:lle, mikä lyhentää suoritusaikaa kertaluokkaa verrattuna aiempiin slack-rescaled SVM:n lähestymistapoihin.

**Tulos**

Nopea ja skaalautuva rakenteellinen SVM löysällä skaalautumisella (Slack Rescaling)

**Esimerkki 2.954**

Tekstianalyysi sisältää tekstin leksikaalisen analyysin, ja sitä on tutkittu ja käytetty laajasti erilaisissa sovelluksissa. Viime vuosikymmenen aikana tutkijat ovat ehdottaneet monia tehokkaita ratkaisuja suurten tekstitietokokonaisuuksien analysoimiseksi/luokittelemiseksi, mutta lyhyiden tekstien analysointi/luokittelu on edelleen haasteellista, koska 1) tiedot ovat hyvin harvalukuisia, 2) ne sisältävät kohinasanoja ja 3) tekstin syntaktista rakennetta on vaikea ymmärtää. Lyhytsanomapalvelu (SMS) on matkapuhelimen/älypuhelimen tekstiviestipalvelu, jota kaikki matkapuhelimen käyttäjät käyttävät usein. SMS-palvelun suosion vuoksi markkinointiyritykset käyttävät tätä palvelua nykyään myös suoramarkkinointiin, joka tunnetaan myös nimellä SMS-markkinointi.Tässä asiakirjassa on ehdotettu ontologiapohjaista SMS-ohjainta, joka analysoi tekstiviestin ja luokittelee sen ontologian avulla lailliseksi tai roskapostiksi. Ehdotettua järjestelmää on testattu eri skenaarioissa, ja kokeelliset tulokset osoittavat, että ehdotettu ratkaisu on tehokas sekä tehokkuuden että ajan suhteen. Avainsanat-Lyhyt tekstiluokitus; tekstiviestien roskaposti; tekstianalyysi; ontologiapohjainen tekstiviestien roskaposti; tekstianalyysi ja ontologia.

**Tulos**

Ontologiaan perustuva tekstiviestien ohjain älypuhelimille

**Esimerkki 2.955**

Tässä artikkelissa tarkastelemme verkko-oppimisen kehyksen muunnelmaa, jossa käytetään asiantuntijan neuvoja ja rajoitettua/kaistapalautetta. Pidämme kutakin asiantuntijaa oppivana yksikkönä, joka pyrkii heijastamaan tarkemmin tiettyjä reaalimaailman sovelluksia. Asetelmassamme palaute millä tahansa hetkellä t on rajoitettua siinä mielessä, että se on vain sen asiantuntijan i käytettävissä, jonka keskusalgoritmi (ennustaja) on valinnut, ts, Tarkastelemme yleistä mustan laatikon lähestymistapaa, jossa ennustaja ei kontrolloi tai tiedä asiantuntijoiden oppimisdynamiikkaa lukuun ottamatta seuraavaa ei-katumuksen oppimisominaisuutta: minkä tahansa asiantuntijan j keskimääräinen katumus häviää vähintäänO(t j ) nopeudella tj oppimisaskeleella, jossa β ∈ [0, 1] on parametri. Monikätisten rosvojen ongelman jälkikäteen parhaan toiminnan kanssa kilpailemisen hengessä tavoitteemme on olla kilpailukykyinen suhteessa kumulatiivisiin tappioihin, joita algoritmi voisi saada, jos se noudattaisi politiikkaa, jossa valitaan aina yksi asiantuntija. Todistamme seuraavan vaikeustuloksen: ilman minkäänlaista koordinointia ennustajan ja asiantuntijoiden välillä on mahdotonta suunnitella ennustaja, jolla saavutetaan katumattomat takuut. Tämän kovuustuloksen kiertämiseksi tarkastelemme käytännön oletusta, jonka mukaan ennustaja voi "ohjata" asiantuntijoiden oppimisprosessia suodattamalla/estämällä osan heidän ympäristöstä havaitsemistaan palautteista, eli antamalla valitun asiantuntijan i olla oppimatta ajanhetkellä t joidenkin aika-askelten ajan. Tämän jälkeen suunnittelemme uudenlaisen katumattoman oppimisalgoritmin LEARNEXP tähän ongelmaympäristöön ohjaamalla huolellisesti asiantuntijoiden havaitsemia palautteita. Todistamme, että LEARNEXP saavuttaa pahimmassa tapauksessa odotetun kumulatiivisen katumuksenO(T 1 2-β ) T aika-askeleen jälkeen ja vastaa katumuksen Θ(T 1 2 ) -rajaa monikätisten bandiittien erikoistapauksessa.

**Tulos**

Oppiminen käyttämään oppijoiden neuvoja

**Esimerkki 2.956**

Yleisesti käytetty oppimissääntö on minimoida keskimääräinen häviö harjoitusjoukossa. Muilla oppimisalgoritmeilla, kuten AdaBoostilla ja hard-SVM:llä, pyritään minimoimaan maksimaalinen tappio harjoitusjoukossa. Keskimääräinen tappio on suositumpi erityisesti syväoppimisessa kolmesta syystä. Ensinnäkin se voidaan kätevästi minimoida käyttämällä online-algoritmeja, jotka käsittelevät muutamia esimerkkejä jokaisella iteraatiolla. Toiseksi usein väitetään, että ei ole järkevää minimoida tappiota harjoitusjoukossa liikaa, koska se ei heijastu yleistyshäviöön. Viimeiseksi, maksimihäviö ei ole kestävä poikkeaville arvoille. Tässä artikkelissa kuvaamme ja analysoimme algoritmin, jolla mikä tahansa online-algoritmi voidaan muuntaa maksimihäviön minimoijaksi. Osoitamme, että joissakin tilanteissa parempi tarkkuus harjoitusjoukossa on ratkaisevan tärkeää, jotta saadaan hyvä suorituskyky näkymättömissä esimerkeissä. Lopuksi ehdotamme lähestymistavan vankkoja versioita, joilla voidaan käsitellä poikkeamia.

**Tulos**

Maksimaalisen tappion minimointi: Miten ja miksi?

**Esimerkki 2.957**

Tässä artikkelissa esitellään hybridi COAW-algoritmi, jossa käytetään Cuckoo-optimointialgoritmia ja yksinkertaista additiivista painotusmenetelmää monitavoiteongelmien ratkaisemiseksi. Cuckoo-algoritmi on tehokas ja jäsennelty menetelmä epälineaaristen jatkuvien ongelmien ratkaisemiseen. Ehdotetun COAW-algoritmin luomat Paretorajat ovat tarkkoja ja niillä on hyvä hajonta. Menetelmä on nopea Paretorajojen löytämisessä ja tunnistaa Paretorajojen alku- ja loppupisteet oikein. Ehdotetun algoritmin validointia varten analysoitiin useita kokeellisia ongelmia. Tulokset osoittavat COAW-algoritmin tehokkuuden monitavoiteongelmien ratkaisemisessa.

**Tulos**

UUSI HYBRIDI-COAW-MENETELMÄ MONITAVOITEONGELMIEN RATKAISEMISEKSI

**Esimerkki 2.958**

Tässä artikkelissa rakennetaan Turingin kone, joka voi ratkaista β ′:n, joka on REcomplete. Tällainen kone on mahdollinen vain, jos tietojenkäsittelytieteen ja matematiikan perusteissa on jotain vikaa. Siksi tarkistamme työmme tarkastelemalla hyvin tarkkaan Cantorin diagonaalisointia ja konstruoimme uudenlaisen formaalin kielen Abelin ryhmänä, jonka avulla voimme ekvivalenssisuhteiden avulla tarjota ei-triviaalin vastaesimerkin Cantorin väitteelle. Aivan kuin tämä ei olisi vielä riittänyt, havaitsemme sitten, että Gödelin diagonalisaation lemman impredikatiivinen luonne johtaa loogiseen tautologiaan, mikä mitätöi kaiken menetelmän taustalla olevan merkityksen, eikä jätä epäilystäkään siitä, että diagonalisaatio on virheellinen. Näitä perustavia argumentteja koskeva löytömme avaa oven P vs. NP -ongelman ratkaisemiseen. 1 Turingin todistus Entscheidungsproblemasta on ollut

**Tulos**

Vahvempi perusta tietojenkäsittelytieteelle ja P=NP:lle

**Esimerkki 2.959**

Tarkastelen tässä monien tekoälyn osa-alueiden nykytilannetta ja arvioin, milloin on kohtuullista odottaa ihmisen tasoista tekoälyn kehitystä. Tunnettujen tekoälytutkijoiden ennusteet vaihtelevat Andrew Ng:n hyvin pessimistisistä ennusteista Geoffrey Hintonin paljon maltillisempiin ennusteisiin ja DeepMindin perustajan Shane Leggin optimistisiin ennusteisiin. Kun otetaan huomioon viime vuosien valtava kehitysvauhti ja tekoälyasiantuntijoiden ennusteiden laaja kirjo, keskustellaan myös tekoälyn turvallisuuskysymyksistä.

**Tulos**

Katsaus tekoälyn uusimpaan kehitystasoon ja sen soveltaminen tekoälyn turvallisuusongelmaan.

**Esimerkki 2.960**

Viime aikoina on ehdotettu useita lähestymistapoja hajautettujen syvien moniagenttikäytäntöjen oppimiseen, jotka koordinoivat toimintaansa erilaistuvan viestintäkanavan kautta. Vaikka nämä politiikat ovat tehokkaita monissa tehtävissä, niiden aiheuttamien viestintästrategioiden tulkinta on edelleen haaste. Tässä ehdotamme, että agenttien viestejä tulkitaan kääntämällä ne. Toisin kuin tyypillisissä konekäännösongelmissa, meillä ei ole rinnakkaista dataa, josta oppia. Sen sijaan kehitämme käännösmallin, joka perustuu näkemykseen, että agenttien viestit ja luonnollisen kielen merkkijonot tarkoittavat samaa asiaa, jos ne saavat kuulijan uskomaan samaan uskomukseen maailmasta. Esitämme teoreettisia takeita ja empiirisiä todisteita siitä, että lähestymistapamme säilyttää sekä viestien semantiikan että pragmatiikan varmistamalla, että käännöstason kautta kommunikoivat pelaajat eivät kärsi merkittävää palkkiotappiota suhteessa pelaajiin, joilla on yhteinen kieli.1

**Tulos**

Kääntäminen Neuralese

**Esimerkki 2.961**

Ympyröinti on edustava esimerkki ei-monotonisesta päättelytekniikasta. Ympyröintiä on usein tutkittu ensimmäisen kertaluvun teorioiden osalta, mutta myös sen propositionaalinen versio on ollut laajan tutkimuksen kohteena, sillä sen on osoitettu vastaavan laajennettua suljetun maailman oletusta (extended closed world assumption, ECWA). Lisäksi propositionaalisen sirkumskription päättely on tunnettu esimerkki päätösongelmasta polynomihierarkian toisella tasolla. Tässä artikkelissa ehdotetaan uutta Boolean Satisfiability (SAT)-pohjaista algoritmia propositionaalisen kiertokirjoituksen sisältämälle päättelylle, jossa tutkitaan propositionaalisen kiertokirjoituksen suhdetta minimimalleihin. Uusi algoritmi on saanut vaikutteita SAT-pohjaisessa mallintarkistuksessa yleisesti käytetyistä ideoista, nimittäin vastaesimerkkien ohjaamasta abstraktioiden tarkentamisesta. Lisäksi uutta algoritmia tarkennetaan laskemaan teorian sulkeutuminen GCWA-oletukselle (generalized close world assumption). Kokeelliset tulokset osoittavat, että uusi algoritmi voi ratkaista ongelmatapauksia, joita muut ratkaisut eivät pysty ratkaisemaan.

**Tulos**

Vastaesimerkin ohjaama abstraktioiden tarkennusalgoritmi propositionaalista kiertokirjoitusta varten.

**Esimerkki 2.962**

Tiivistelmä Stokastista gradienttilaskeutumista (Stochastic gradient descent, SGD) matalan sijan faktorointiin [9] käytetään yleisesti matriisiongelmien nopeuttamiseen, mukaan lukien matriisien täydentäminen, aliavaruusseuranta ja SDP-relaksaatio. Tässä artikkelissa esitämme askelkokojärjestelmän SGD:lle matalan rankin pienimmän neliösumman ongelmassa ja todistamme, että laajoissa näytteenotto-olosuhteissa menetelmämme konvergoi globaalisti satunnaisesta alkupisteestä O(ǫn logn) askelissa vakiotodennäköisyydellä vakiorangin ongelmille. SGD:n modifikaatiomme liittää sen stokastiseen tehoitteluun (stochastic power iteration). Esitämme myös kokeita, joilla havainnollistamme algoritmin suoritusaikaa ja konvergenssia.

**Tulos**

Stokastisen gradienttilaskennan globaali konvergenssi joissakin ei-konveksisissa matriisiongelmissa

**Esimerkki 2.963**

Tässä artikkelissa kuvataan syväoppimiseen perustuva lähestymistapamme Twitterin tunteiden analysointiin osana SemEval-2016 Task 4 -tehtävää. Käytämme konvoluutiohermoverkkoa sentimentin määrittämiseen ja osallistumme kaikkiin osatehtäviin eli kaksi-, kolmi- ja viisipisteasteiseen sentimenttiluokitukseen sekä kaksi- ja viisipisteasteiseen sentimentin kvantifiointiin. Saavutamme kilpailukykyisiä tuloksia kahden pisteen asteikolla tapahtuvassa tunteiden luokittelussa ja kvantifioinnissa, ja sijoitumme viidenneksi ja neljänneksi (kolmanneksi ja toiseksi vaihtoehtoisilla mittareilla), vaikka käytämme vain valmiiksi koulutettuja sulautumia, jotka eivät sisällä mitään tunteita koskevaa tietoa. Saavutamme hyvän tuloksen kolmen pisteen asteikolla tapahtuvassa tunteiden luokittelussa, jossa sijoitumme kahdeksanneksi 35:stä, mutta huonosti viiden pisteen asteikolla tapahtuvassa tunteiden luokittelussa ja kvantifioinnissa. Virheanalyysi paljastaa, että tämä johtuu mallin vähäisestä ilmaisukyvystä negatiivisten tunteiden kuvaamiseen sekä kyvyttömyydestä ottaa huomioon ordinaalitietoa. Ehdotamme parannuksia näiden ja muiden ongelmien ratkaisemiseksi.

**Tulos**

INSIGHT-1 SemEval-2016 -tapahtumassa Tehtävä 4: Convolutional Neural Networks for Sentiment Classification and Quantification (Konvoluutio-neuraaliverkot tunteiden luokitteluun ja kvantifiointiin)

**Esimerkki 2.964**

Valvotun luokittelun vakiolähestymistapa sisältää log-lossin minimoinnin luokitteluvirheen ylärajana. Vaikka tämä on tiukka raja optimoinnin alkuvaiheessa, siinä ylikorostetaan virheellisesti luokiteltujen esimerkkien vaikutusta kaukana päätösrajasta. Ylärajan päivittäminen optimoinnin aikana johtaa parempiin luokitusasteisiin samalla kun oppiminen muutetaan minimointiongelmien sarjaksi. Lisäksi, kun luokittelija on osa laajempaa järjestelmää, tämä muutos mahdollistaa luokittelijan suorituskyvyn yhdistämisen koko järjestelmän suorituskykyyn, mikä mahdollistaa ulkoisten rajoitusten saumattoman käyttöönoton.

**Tulos**

TIUKEMMAT RAJAT JOHTAVAT PAREMPIIN LUOKITTELIJOIHIN.

**Esimerkki 2.965**

Upotukset ovat yleisiä representaatioita, jotka ovat hyödyllisiä monissa NLP-tehtävissä. Tässä artikkelissa esitellään DENSIFIER, menetelmä, joka oppii sulautusavaruuden ortogonaalisen muunnoksen, joka keskittää tehtävän kannalta olennaisen tiedon ultradensiiviseen aliavaruuteen, jonka dimensiomäärä on 100-kertaisesti pienempi kuin alkuperäisen avaruuden. Osoitamme, että DENSIFIERin tuottamat ultradense embedding-avaruudet saavuttavat huipputason leksikon luomistehtävässä, jossa sanat merkitään kolmenlaisella leksikaalisella informaatiolla - tunnetilalla, konkreettisuudella ja frekvenssillä. SemEval2015 10B -tunnelma-analyysitehtävässä osoitamme, että ultradense-avaruutta käytettäessä ei menetetä tietoa, mutta harjoittelu on kertaluokkaa tehokkaampaa ultradense-avaruuden tiiviyden ansiosta.

**Tulos**

Ultradense Word Embeddings by Orthogonal Transformation (sanojen upottaminen ortogonaalisella muunnoksella)

**Esimerkki 2.966**

Tässä artikkelissa esitellään useiden syvän konvoluutioautokooderin mallien kehittäminen Caffe-syväoppimiskehyksessä ja niiden kokeellinen arviointi MNIST-tietokannan avulla. Olemme luoneet viisi konvoluutioautokooderin mallia, jotka eroavat arkkitehtuuriltaan toisistaan sen mukaan, onko autokooderin kooderi- ja dekooderiosissa pooling- ja unpooling-kerroksia vai ei. Tuloksemme osoittavat, että kehitetyt mallit tuottavat erittäin hyviä tuloksia dimensioiden pienentämisessä ja valvomattomassa klusterointitehtävässä sekä pieniä luokitteluvirheitä, kun käytimme opittua sisäistä koodia valvotun lineaarisen luokittelijan ja monikerroksisen perceptronin syötteenä. Parhaat tulokset saatiin mallilla, jossa kooderiosa sisältää konvoluutio- ja poolointikerrokset, ja sen jälkeen analoginen dekooderiosa, jossa on dekonvoluutio- ja unpoolointikerrokset ilman, että dekooderiosassa käytetään kytkentämuuttujia. Artikkelissa käsitellään myös käytännön yksityiskohtia syvän konvoluutioautokooderin luomisesta erittäin suositussa Caffe-syväoppimiskehyksessä. Uskomme, että tässä paperissa esitelty lähestymistapamme ja tuloksemme voivat auttaa muita tutkijoita rakentamaan tehokkaita syviä neuroverkkoarkkitehtuureja tulevaisuudessa.

**Tulos**

Syvä konvolutiivinen automaattinen kooderi, jossa on pooling- ja unpooling-kerrokset Caffe-ohjelmassa.

**Esimerkki 2.967**

Ehdotamme online-konveksuaalista optimointialgoritmia (RESCALEDEXP), joka saavuttaa<lb>optimaalisen katumuksen rajoittamattomassa ympäristössä ilman ennakkotietoa tappiofunktioiden rajoista<lb>. Todistamme alemman rajan, joka osoittaa eksponentiaalisen eron sellaisten olemassa olevien algoritmien, jotka vaativat tunnetun rajan<lb> häviöfunktioille, ja minkä tahansa algoritmin, joka ei vaadi tällaista tietoa, katumuksen välillä.<lb>RESCALEDEXP vastaa tätä alempaa rajaa asymptoottisesti itera-<lb>tionien määrässä. RESCALEDEXP on luonnollisesti hyperparametrivapaa, ja osoitamme empiirisesti, että se vastaa aiempia optimointialgoritmeja, jotka vaativat hyperparametrin<lb>optimointia. 1 Online Convex Optimization Online Convex Optimization (OCO) [1, 2] tarjoaa tyylikkään kehyksen meluisten,<lb>vastakkaisten tai muuttuvien ympäristöjen mallintamiseen. Ongelma voidaan esittää muodollisesti<lb>seuraavien määritelmien avulla: <lb>Konvex joukko: JoukkoW on kovera, josW sisältyy johonkin reaaliseen vektoriavaruuteen ja tw+(1- t)w′ ∈W<lb>kaikille w,w′ ∈W ja t ∈ [0, 1].<lb>Konveksinen funktio: f :W → R on konveksinen funktio, jos f(tw + (1- t)w′) ≤ tf(w) + (1- t)f(w′)<lb>kaikille w,w′ ∈W ja t ∈ [0, 1]. OCO-ongelma on toistuvien kierrosten peli, jossa kierroksella t oppija ensin valitsee elementin<lb>wt jostain koverasta avaruudesta W , sitten saa koveran tappiofunktion `t ja kärsii tappion `t(wt). Oppijan<lb>regret jonkun muun u ∈W suhteen määritellään seuraavasti

**Tulos**

Online-konveksinen optimointi rajoittamattomilla alueilla ja tappioilla

**Esimerkki 2.968**

Formalisoimme yksinkertaistettujen lautapelien kielen, joka kuvaa mielivaltaisten lautapelien alaluokkaa. Kielen rakenne perustuu säännöllisiin lausekkeisiin, mikä tekee säännöistä helposti konekäsiteltäviä ja pitää ne samalla lyhyinä ja melko helppolukuisina.

**Tulos**

YKSINKERTAISTETUT LAUTAPELIT

**Esimerkki 2.969**

Ehdotamme uutta tensorifaktorointimenetelmää, Sparse Hierarchical Tucker (Sparse H-Tucker), harvoille ja korkea-asteisille tensoreille. Sparse H-Tucker -menetelmä on saanut innoituksensa nimensä mukaisesta klassisesta hierarkkisesta Tucker-menetelmästä, jonka tavoitteena on laskea syötetystä datajoukosta puurakenteinen faktorisointi, jota alan asiantuntija voi helposti tulkita. Sparse H-Tucker -menetelmässä käytetään kuitenkin sisäkkäistä näytteenottotekniikkaa, jolla ratkaistaan hierarkkisen Tucker-menetelmän keskeinen skaalautuvuusongelma, joka on hankalan tiheän ydintensorin luominen; lähestymistapamme tuloksena on nopeampi, tilatehokkaampi ja tarkempi menetelmä. Testaamme menetelmäämme laajasti todellisella terveydenhuollon tietokokonaisuudella, joka on kerätty 30 000 potilaasta ja jonka tuloksena on 18. kertaluvun harva datatensor. Toisin kuin kilpailevat menetelmät, Sparse H-Tucker voi analysoida koko datajoukon yhdellä monisäikeisellä koneella. Se pystyy myös tekemään sen tarkemmin ja nopeammin kuin uusimmat menetelmät: 12. kertaluvun osajoukolla syötetystä datasta Sparse H-Tucker on 18-kertaisesti tarkempi ja 7,5-kertaisesti nopeampi kuin aiempi uusin menetelmä. Jopa matalan kertaluvun tensoreiden (esim. 4 kertaluvun) analysointiin menetelmämme tarvitsee lähes kertaluokkaa vähemmän aikaa ja yli kaksi kertaluokkaa vähemmän muistia verrattuna perinteisiin tensorifaktorointimenetelmiin, kuten CP:hen ja Tuckeriin. Lisäksi havaitsemme, että Sparse H-Tucker skaalautuu lähes lineaarisesti nollasta poikkeavien tensorielementtien määrän mukaan. Tuloksena saatava malli tarjoaa myös tulkittavissa olevan tautihierarkian, jonka kliininen asiantuntija on vahvistanut.

**Tulos**

Harva hierarkkinen Tucker-faktorointi ja sen soveltaminen terveydenhuoltoon

**Esimerkki 2.970**

Tekstin kohdistaminen ja tekstin laatu ovat ratkaisevia tekijöitä konekäännösjärjestelmien, joidenkin NLP-työkalujen ja muiden kaksikielistä dataa vaativien tekstinkäsittelytehtävien tarkkuuden kannalta. Tässä tutkimuksessa ehdotetaan kielestä riippumatonta kahden lauseen suodatusmenetelmää, joka perustuu puolan (joka ei ole asentoherkkä kieli) ja englannin kokeisiin. Tämä puhdistusmenetelmä kehitettiin TED Talks -korpuksella, ja sitä testattiin aluksi myös Wikipedian vertailukelpoisella korpuksella, mutta sitä voidaan käyttää millä tahansa tekstialueella tai kieliparilla. Ehdotetussa lähestymistavassa käytetään erilaisia heuristiikkoja lauseiden vertailua varten. Joissakin niistä hyödynnetään synonyymejä ja tekstin semanttista ja rakenteellista analyysia lisätietona. Tietohävikin minimointi varmistettiin. Työkalun avulla käsitellyn tekstin avulla MT-järjestelmän pistemäärän parantumisesta keskustellaan.

**Tulos**

Melua sisältävien rinnakkaisten ja vertailukelpoisten korpusten suodatusmenetelmä kaksikielisten vastaavien tietojen poimimiseksi lausetasolla.

**Esimerkki 2.971**

Paikannukseen kykenevien laitteiden räjähdysmäinen kasvu yhdessä Internet-palvelujen käytön lisääntymisen kanssa on johtanut siihen, että tietoisuus paikkatiedon merkityksestä ja käytöstä monissa sovelluksissa on lisääntynyt. Navigointisovellukset (joita usein kutsutaan kartoiksi) käyttävät erilaisia saatavilla olevia tietolähteitä matka-ajan laskemiseen ja ennustamiseen sekä useiden vaihtoehtojen laskemiseen julkisen liikenteen, auton tai jalankulkijan reittiä varten. Tässä asiakirjassa arvioidaan Maps-sovellusten jalankulkutapaa kolmessa tärkeimmässä älypuhelinten käyttöjärjestelmässä (Android, iOS ja Windows Phone). Artikkelissa osoitetaan, että iOS:n, Androidin ja Windows Phonen Maps-sovellukset ennustavat jalankulkutilassa matka-aikaa oppimatta yksilön liikkumisprofiilista. Lisäksi osoitamme, että näissä sovelluksissa on erityinen tiedon laatuongelma, joka liittyy siihen, että niistä puuttuu tietoa jalankulkijoiden risteysten sijainnista ja tyypistä. Lopuksi havainnollistamme, miten yksilöiden liikkumisprofiilista oppiminen eri ennakoivan analytiikan malleilla parantaa matka-ajan arvioinnin tarkkuutta.

**Tulos**

Ennustava analytiikka matka-ajan arvioinnin parantamiseksi Applen, Googlen ja Microsoftin navigointisovelluksissa

**Esimerkki 2.972**

Monte carlo -tutkimuksia tehtiin joidenkin vaihtoehtoisten menettelyjen käyttäytymisen vertaamiseksi epävarmuuden vallitessa. Useiden Bayesin mallin, lineaarisen mallin ja oletusarvoisen päättelyn menettelyjen käyttäytymistä tutkittiin kalibrointivirheen lisääntyvän tason yhteydessä. Mielenkiintoisin tulos on, että Bayesin menettelyillä oli taipumus tuottaa enemmän äärimmäisiä posteriorisia uskomusarvoja (posterioriset uskomukset lähellä 0,0 tai 1,0) kuin muilla tekniikoilla, mutta lineaariset mallit tuottivat suhteellisesti harvemmin vahvaa tukea virheelliselle päätelmälle. Myös todisteiden välisten todennäköisyysriippuvuuksien huomioon ottaminen oli tärkeää sekä bayesiläisissä että lineaarisissa päivitysmenetelmissä.

**Tulos**

Päättely epävarmuuden vallitessa:

**Esimerkki 2.973**

Tässä asiakirjassa kuvataan ontologian graafien luomisjärjestelmän esikäsittelyvaihe eri alojen Punjabi-tekstiasiakirjoista. Tässä tutkimuksessa keskitytään Punjabi-tekstiasiakirjojen esikäsittelyyn. Esikäsittely on syötetyn tekstin strukturoitu esitys. Ontologian graafin luomisen esikäsittely sisältää tekstin rajoitusten sallimisen, erikoismerkkien ja välimerkkien poistamisen, päällekkäisten termien poistamisen, pysäytyssanojen poistamisen, termien poimimisen vertaamalla syötettyjä termejä sanakirjan ja hakemistoluettelon termeihin. AvainsanatOntologia, esikäsittelyvaihe, ontologiagraafi, tietämyksen esittäminen, luonnollisen kielen prosessointi.

**Tulos**

Punjabi-kielisen toimialueontologian graafin luomisjärjestelmän esikäsittelyn suorittaminen

**Esimerkki 2.974**

Tutkimme konvoluutiohermoverkkojen (convolutional neural networks, CNN) käyttöä aukkojen täyttämistehtävässä puhutun kielen ymmärtämisessä. Ehdotamme uutta CNN-arkkitehtuuria sekvenssien merkitsemistä varten, joka ottaa huomioon aiemmat kontekstissa olevat sanat, joiden järjestystieto on säilynyt, ja kiinnittää erityistä huomiota nykyiseen sanaan ja sen ympäröivään kontekstiin. Lisäksi se yhdistää luokittelua varten aiempien ja tulevien sanojen tiedot. Ehdotettu CNN-arkkitehtuuri päihittää jopa aiemmin parhaan yhdistävän rekursiivisen neuroverkkomallin, ja sen avulla saavutetaan ATIS-vertailutietokannassa huippuluokan tulokset, joiden F1-tulos on 95,61 %, ilman kielellistä lisätietoa ja -resursseja.

**Tulos**

Sekventiaaliset konvolutiiviset neuroverkot aukkojen täyttämiseen puhutun kielen ymmärtämisessä

**Esimerkki 2.975**

Tässä työssä ehdotetaan matalan kompleksisuuden epälineaarisuusmallia ja kehitetään siihen mukautuvia algoritmeja. Malli perustuu hajotettaviin - tai tensorikielellä sanottuna rank-one - Volterran ytimiin. Se voidaan kuvata myös FIR-suodattimien tuotteena, mikä selittää sen vähäisen monimutkaisuuden. Rank-one-malli on mielenkiintoinen myös siksi, että se on peräisin approksimaatioteorian hyvin ratkaistusta ongelmasta. Artikkelissa käytetään tällaista mallia estimointiteorian yhteydessä tarkan gradienttityyppisen algoritmin kehittämiseen, josta johdetaan adaptiivisia algoritmeja, kuten pienimmän keskiarvon neliösuodatin (LMS) ja sen datan uudelleenkäyttöön perustuva versio TRUE-LMS. Käsitellään vakaus- ja konvergenssikysymyksiä. Tämän jälkeen algoritmeja testataan simulaatioissa, jotka osoittavat sen hyvän suorituskyvyn verrattuna muihin kirjallisuudessa esiintyviin epälineaarisiin käsittelyalgoritmeihin.

**Tulos**

Epälineaariset mukautuvat algoritmit Rank-One-tensorimalleille

**Esimerkki 2.976**

Tässä asiakirjassa käytetään kotitalouksien sähkökuormitusprofiilien klusterointiin lähestymistapaa, jota on käytetty menestyksekkäästi Portugalin tietojen kanssa, ja sovelletaan sitä Yhdistyneen kuningaskunnan tietoihin. Klusterointitekniikoita sovelletaan, ja havaitaan, että portugalilaisessa työssä käytetty tekniikka (kaksivaiheinen prosessi, jossa yhdistetään Self Organised Maps ja Kmeans) ei sovellu Yhdistyneen kuningaskunnan tietoihin. Työ osoittaa, että jopa yhdeksän kotitalousklusteria voidaan tunnistaa, ja käyttöprofiilien erot ovat visuaalisesti silmiinpistäviä. Tämä osoittaa, että on tarkoituksenmukaista eritellä sähkönkäyttömallit yksityiskohtaisemmin kuin sähköalan nykyisin julkaisemat kaksi kuormitusprofiilia. Artikkelissa esitetään yksityiskohtaisesti alustavat tulokset, jotka on saatu Milton Keynesissä noin vuonna 1990 kerätyistä tiedoista. Jatkotyössä keskitytään tarkkojen ja merkityksellisten ryhmien muodostamiseen samankaltaisista sähkönkäyttäjistä, jotta kysynnänhallinta-aloitteet voidaan kohdentaa paremmin tärkeimmille kohdeasiakkaille.

**Tulos**

Klusterointikehyksen soveltaminen Yhdistyneen kuningaskunnan kotitaloussähkötietoihin

**Esimerkki 2.977**

Bayes-verkoissa todennäköisin selitys (Most Probable Explanation, MPE) on täydellinen muuttujan instantiointi, jolla on suurin todennäköisyys nykyisen todistusaineiston perusteella. Tässä artikkelissa käsitellään ongelmaa, joka liittyy MPE:n kestävyysehtojen löytämiseen yksittäisten parametrimuutosten yhteydessä. Erityisesti kysymme seuraavaa: Kuinka paljon yksittäisen verkkoparametrin muutosta on varaa tehdä, kun MPE pysyy muuttumattomana? Kuvaamme menettelyn, joka on ensimmäinen laatuaan ja joka laskee tämän vastauksen kaikille Bayes-verkon parametreille ajassa O(n exp(w)), jossa n on verkon muuttujien lukumäärä ja w sen puunleveys.

**Tulos**

Todennäköisimpien selitysten kestävyydestä

**Esimerkki 2.978**

Opimme rikkaita luonnollisia äänikuvauksia hyödyntämällä suuria määriä luonnosta kerättyä merkitsemätöntä äänidataa. Hyödynnämme näön ja äänen luonnollista synkronointia akustisen edustuksen oppimisessa kahden miljoonan merkitsemättömän videon avulla. Merkitsemättömien videoiden etuna on se, että niitä voidaan hankkia taloudellisesti valtavassa mittakaavassa, mutta ne sisältävät hyödyllisiä signaaleja luonnollisesta äänestä. Ehdotamme oppilas-opettaja -koulutusmenettelyä, jolla siirretään hyvin vakiintuneista visuaalisista tunnistusmalleista peräisin olevaa visuaalista tietämystä äänimodaaliin käyttämällä merkitsemätöntä videota siltana. Äänikuvauksemme parantaa suorituskykyä merkittävästi verrattuna uusimpiin tuloksiin akustisen kohtauksen ja kohteen luokittelun vakioesimerkkien avulla. Visualisoinnit viittaavat siihen, että ääniverkossa syntyy automaattisesti jonkin verran korkean tason semantiikkaa, vaikka se on koulutettu ilman perustotuusmerkintöjä.

**Tulos**

SoundNet: Oppiminen ääniedustuksista merkitsemättömästä videosta

**Esimerkki 2.979**

Kuvaamme MITREn osallistumista SemEval-2016 Task 6, Detecting Stance in Tweets -tehtävään. Tämä ponnistus saavutti tehtävän A parhaat tulokset valvotussa kannan havaitsemisessa, ja sen keskimääräinen F1-pistemäärä oli 67,8, kun arvioitiin, oliko twiitin kirjoittaja jonkin aiheen puolesta vai vastaan. Käytimme rekurrenssia neuroverkkoa, joka oli alustettu kaukovalvonnalla opituilla piirteillä kahdella suurella merkitsemättömällä tietokokonaisuudella. Koulutimme sanojen ja lauseiden upotukset word2vec-menetelmällä, jonka jälkeen käytimme näitä ominaisuuksia oppiaksemme lause-esityksiä hashtagien ennustamisen aputehtävän avulla. Nämä lauseenvektorit hienosäädettiin sitten asennon havaitsemista varten useilla sadoilla leimatuilla esimerkeillä. Tuloksena oli suorituskykyinen järjestelmä, joka käytti siirto-oppimista maksimoidakseen käytettävissä olevan harjoitusdatan arvon.

**Tulos**

MITRE SemEval-2016 -tapahtumassa Tehtävä 6: Transfer Learning for Stance Detection (Siirto-oppiminen asennon havaitsemiseksi)

**Esimerkki 2.980**

Nykyisten vahvistusoppimismenetelmien avulla on viime aikoina ollut mahdollista oppia pelaamaan tuntemattomia 3D-pelejä raakapikseleistä. Tässä työssä tutkimme haasteita, joita tällaisissa monimutkaisissa ympäristöissä syntyy, ja teemme yhteenvedon nykyisistä menetelmistä, joilla niitä voidaan lähestyä. Valitsemme Doom-pelissä tehtävän, jota ei ole vielä lähestytty. Agentin tavoitteena on taistella vihollisia vastaan viidestä huoneesta koostuvassa 3D-maailmassa. Harjoittelemme DQN- ja LSTMA3C-algoritmeja tässä tehtävässä. Tulokset osoittavat, että molemmat algoritmit oppivat järkeviä toimintatapoja, mutta eivät saavuta korkeita pistemääriä, kun otetaan huomioon harjoittelun määrä. Tarjoamme näkemyksiä opitusta käyttäytymisestä, mikä voi toimia arvokkaana lähtökohtana Doom-alan jatkotutkimukselle.

**Tulos**

Syvä vahvistusoppiminen raakapikseleistä Doomissa

**Esimerkki 2.981**

Tässä työssä keskitytään vastaamaan Freebase-tietokantaan kohdistuviin yksittäisiin faktatietokysymyksiin. Kunkin kysymyksen vastaus voidaan saada yhdestä Freebase-tietokannassa olevasta muodosta (subjekti, predikaatti, objekti). Tähän tehtävään, yksinkertaisiin kysymyksiin vastaamiseen (SimpleQA), voidaan vastata kaksivaiheisella putkella: entiteettien linkittäminen ja faktojen valinta. Faktojen valinnassa sovitamme subjektin ja kysymyksessä mainitun entiteetin yhteen merkkitason konvolutiivisen neuroverkon (char-CNN) avulla ja predikaatin ja kysymyksen sanatason CNN:n (wordCNN) avulla. Tässä työssä on kaksi keskeistä panosta. (i) Ehdotetaan yksinkertaista ja tehokasta entiteettilinkitystä Freebasen yli. Entiteettilinkkaajamme on parempi kuin SimpleQA-tehtävän huipputason entiteettilinkkaaja. (ii) Word-CNN:n päälle on pinottu uudenlainen tarkkaavainen maxpooling, jotta predikaattiesitys voidaan sovittaa tehokkaammin yhteen predikaattikeskeisen kysymysesityksen kanssa. Kokeet osoittavat, että järjestelmämme asettaa uuden huippuluokan tässä tehtävässä.

**Tulos**

Yksinkertainen kysymyksiin vastaaminen huomioivan konvolutiivisen neuroverkon avulla

**Esimerkki 2.982**

Tekoälymenetelmiä on usein sovellettu tiettyjen toimintojen tai tehtävien suorittamiseen kyberpuolustuksen alalla. Koska vastustajan menetelmät ovat kuitenkin yhä monimutkaisempia ja vaikeammin hahmotettavia, hajanaiset pyrkimykset ymmärtää kyberhyökkäyksiä ja erityisesti haittaohjelmiin perustuvia hyökkäyksiä eivät tarjoa haittaohjelmien analyytikoille riittäviä keinoja ymmärtää haittaohjelmien aiempia, nykyisiä ja tulevia ominaisuuksia. Tässä asiakirjassa esittelemme MAAGI-järjestelmän (Malware Analysis and Attributed using Genetic Information). MAAGI-järjestelmän perusajatuksena on, että haittaohjelmien käyttäytymisen ja biologisten organismien käyttäytymisen välillä on vahvoja yhtäläisyyksiä, ja biologisesti inspiroituneiden menetelmien soveltaminen haittaohjelmien korporaatioihin voi auttaa analyytikkoja ymmärtämään paremmin haittaohjelmien hyökkäysten ekosysteemiä. Haittaohjelmien ja analyysin monimutkaisuuden vuoksi MAAGI-järjestelmä tukeutuu vahvasti tekoälytekniikoihin, jotta se voisi tarjota tämän mahdollisuuden. Järjestelmä on jo tuottanut lupaavia tuloksia kehityskaarensa aikana, ja toivottavasti se innostaa tekoäly- ja kyberpuolustusyhteisöjen integroitumista.

**Tulos**

Tekoälyyn perustuva haittaohjelmien analysointi

**Esimerkki 2.983**

Osoitamme, miten silmäseuranta-korpuksia voidaan käyttää lauseiden pakkausmallien parantamiseen, ja esittelemme uuden monitehtäväoppimisalgoritmin, joka perustuu monikerroksisiin LSTM-malleihin. Saavutamme suorituskyvyn, joka on kilpailukykyinen tai parempi kuin uusimpien lähestymistapojen suorituskyky.

**Tulos**

Lauseiden tiivistämisen parantaminen oppimalla ennustamaan katseita

**Esimerkki 2.984**

Esittelemme diskriminoivan rekursiivisen harvahkon autokooderimallin, joka koostuu rekursiivisesta koodaajasta, joka koostuu rekursiivisista lineaarisista yksiköistä, jotka on purettu kiinteän iteraatiomäärän ajan ja yhdistetty kahteen lineaariseen dekooderiin, jotka rekonstruoivat syötteen ja ennustavat sen valvotun luokituksen. Koulutus backpropagation-through-time-menetelmällä minimoi aluksi valvomattoman harvan rekonstruktiovirheen; sen jälkeen tappiofunktiota täydennetään valvotun luokittelun diskriminoivalla termillä. Ajallisesti rullaamattomaan muotoon sisältyvä syvyys antaa järjestelmälle paljon enemmän edustuskykyä, mutta koulutettavien parametrien määrä pysyy kiinteänä. Alun perin jäsentymättömästä verkosta piilotetut yksiköt jakautuvat kategorisiin yksiköihin, joista kukin edustaa syötteen prototyyppiä, jolla on hyvin määritelty luokka, ja osayksiköihin, jotka edustavat näiden prototyyppien muodonmuutoksia. Rekurrenssikoodaajan opittu organisaatio on hierarkkinen: part-yksiköt toimivat suoraan syötteen perusteella, kun taas kategoristen yksiköiden toiminta rakentuu ajan myötä vuorovaikutuksessa part-yksiköiden kanssa. Jopa käyttämällä pientä määrää piilotettuja yksiköitä kerrosta kohti, diskriminoivat rekursiiviset harvahkot autokooderit saavuttavat erinomaisen suorituskyvyn MNIST-testissä.

**Tulos**

Diskriminoivat rekursiiviset sparkeat autokooderit (Sparse Auto-Encoders)

**Esimerkki 2.985**

Monissa luonnollisen kielen käsittelytehtävissä (NLP) asiakirja mallinnetaan yleisesti sanapussina käyttäen termifrekvenssi-inversi-dokumenttifrekvenssi (TF-IDF) -vektoria. TF-IDF-ominaisuusvektorin yksi merkittävä puute on se, että se jättää huomiotta sanajärjestykset, jotka sisältävät syntaktisia ja semanttisia suhteita asiakirjan sanojen välillä, ja ne voivat olla tärkeitä joissakin NLP-tehtävissä, kuten genreluokittelussa. Tässä asiakirjassa ehdotetaan uutta hajautettua vektoriedustusta asiakirjasta: ensin luodaan yksinkertainen rekurrendi-neuraalisen verkon kielimalli (RNN-LM) tai pitkän lyhytkestoisen muistin RNN-kielimalli (LSTM-LM) kaikista tehtävän asiakirjoista; joitakin LM-parametreja mukautetaan sitten kunkin asiakirjan mukaan, ja mukautetut parametrit vektoroidaan asiakirjan edustamiseksi. Uudet asiakirjavektorit merkitään DV-RNN:ksi ja DV-LSTM:ksi. Uskomme, että uudet asiakirjavektorimme pystyvät tallentamaan asiakirjoihin jonkin verran korkean tason peräkkäistä tietoa, jota muut nykyiset asiakirjojen esitystavat eivät pysty tallentamaan. Uusia asiakirjavektoreita arvioitiin dokumenttien genreluokittelussa kolmessa korpuksessa: Brown Corpus, BNC Baby Corpus ja keinotekoisesti luotu Penn Treebank -tietokanta. Niiden luokittelusuorituksia verrataan TF-IDF-vektorin suorituskykyyn ja uusimpaan hajautetun muistin malliin kappale-vektorista (PV-DM). Tulokset osoittavat, että DV-LSTM päihittää TF-IDF:n ja PV-DM:n useimmissa tapauksissa merkittävästi, ja ehdotettujen asiakirjavektorien yhdistäminen TF-IDF:n tai PVDM:n kanssa voi parantaa suorituskykyä entisestään.

**Tulos**

Rekursiivinen neuroverkko Kielimallin mukauttaminen Johdettu dokumenttivektori

**Esimerkki 2.986**

Laajennamme Stochastic Gradient Variational Bayes -menetelmää, jotta voimme tehdä posteriorisen päättelyn Stick-Breaking-prosessien painoille. Tämän kehityksen ansiosta voimme määritellä Stick-Breaking Variational Autoencoder (SB-VAE), Bayesin ei-parametrinen versio variational autoencoderista, jolla on latentti representaatio, jolla on stokastinen ulottuvuus. Kokeellisesti osoitamme, että SB-VAE ja sen puolivalvottu muunnos oppivat erittäin erottelevia latentteja representaatioita, jotka usein päihittävät Gaussin VAE:n.

**Tulos**

STICK-BREAKING VARIATIONAL AUTOENCODERS (MUUNNELMA-AUTOKOODERIT)

**Esimerkki 2.987**

Formalisoimme jaettujen valvontaprotokollien synteesin, jolla on oikeellisuustakuu temporaalilogiikan määrittelyille. Tarkemmin sanottuna esittelemme mallinnusformalismin, jossa sekä ihminen että autonomiaprotokolla voivat antaa robotille käskyjä tietyn tehtävän suorittamiseksi. Nämä komennot yhdistetään robotin yhteiseksi syötteeksi. Autonomiaprotokolla syntetisoidaan käyttämällä abstraktiota ihmisen mahdollisista käskyistä, jossa otetaan huomioon päätöksissä esiintyvä satunnaisuus, joka johtuu esimerkiksi väsymyksestä tai käsillä olevan ongelman ymmärtämättömyydestä. Synteesi on suunniteltu siten, että tuloksena oleva robotin käyttäytyminen täyttää annetut turvallisuus- ja suorituskykymäärittelyt, esim. ajallisessa logiikassa. Ratkaisumme perustuu epälineaariseen ohjelmointiin, ja käsittelemme luontaista skaalautuvuusongelmaa esittämällä vaihtoehtoisia menetelmiä. Arvioimme lähestymistavan toteutettavuutta ja skaalautuvuutta kokeellisella arvioinnilla.

**Tulos**

Yhteisten valvontaprotokollien synteesi, jonka turvallisuus ja suorituskyky voidaan taata.

**Esimerkki 2.988**

Tässä esiteltävässä työssä suunnitellaan monikerroksiseen perceptroniin (MLP) perustuva kuvioluokitin käsinkirjoitettujen Bangla-numeroiden tunnistamiseen 76 elementin ominaisvektorin avulla. Bangla on Intian toiseksi suosituin kirjoitusasu ja kieli ja maailman viidenneksi suosituin kieli. Tässä käsinkirjoitettujen Bangla-numeroiden esittämiseen kehitetty ominaisuusjoukko sisältää 24 varjo-ominaisuutta, 16 keskipiste-ominaisuutta ja 36 pisimmän juoksun ominaisuutta. Kokeilemalla 6000 näytteen tietokannalla tekniikka tuottaa keskimäärin 96,67 prosentin tunnistusasteen, kun tulokset on validoitu kolminkertaisesti ristiin. Se on hyödyllinen sovelluksissa, jotka liittyvät käsinkirjoitettujen Bangla Digit -merkkien OCR-tunnistukseen, ja sitä voidaan laajentaa myös käsinkirjoitettujen Bangla-aakkosten merkkien OCR-tunnistukseen.

**Tulos**

MLP-pohjainen lähestymistapa käsinkirjoitettujen "Bangla"-numeroiden tunnistamiseen

**Esimerkki 2.989**

LSTM-malleista on tullut monien syvien NLP-mallien perusrakenne. Viime vuosina on ehdotettu monia parannuksia ja muunnelmia syville sekvenssimalleille yleensä ja erityisesti LSTM-malleille. Ehdotamme ja analysoimme useita LSTM-verkkojen lisäyksiä ja muutoksia, joiden tuloksena on parempi suorituskyky tekstiluokitustietokannoissa. Havaitsemme perinteisiin LSTM-verkkoihin nähden huomattavia parannuksia Monte Carlo -testiaikaisen mallin keskiarvoistamisen, keskiarvopoolauksen ja jäännösyhteyksien sekä neljän muun ehdotetun muutoksen avulla. Analyysimme tarjoaa yksinkertaisen, luotettavan ja laadukkaan perusmallin.

**Tulos**

LSTMS:ÄÄ KOSKEVIEN VIIMEAIKAISTEN OIVALLUSTEN YHDISTÄMINEN

**Esimerkki 2.990**

Sosiaalista mediaa ja tiedonlouhintaa käytetään yhä useammin poliittisten ja yhteiskunnallisten kysymysten analysointiin. Tässä tutkimuksessa luokittelemme sosiaalisen median käyttäjät alueellaan käynnissä olevia itsenäisyysliikkeitä tukeviksi tai vastustaviksi. Itsenäisyysliikkeet syntyvät alueilla, joiden kansalaisilla on ristiriitaisia kansallisia identiteettejä. Käyttäjät, joilla on vastakkaiset kansalliset identiteetit, tukevat tai vastustavat tunnetta siitä, että he ovat osa itsenäistä kansakuntaa, joka eroaa virallisesti tunnustetusta maasta. Kuvaamme menetelmän, joka perustuu käyttäjien itse ilmoittamaan sijaintiin ja jonka avulla muodostamme tietokokonaisuuksia kolmesta alueesta - Kataloniasta, Baskimaasta ja Skotlannista - ja testaamme kielestä riippumattomia luokittelijoita, jotka käyttävät neljää erityyppistä piirrettä. Osoitamme, että lähestymistapa on tehokas suurten annotoitujen tietokokonaisuuksien rakentamisessa ja että sen avulla voidaan saavuttaa tarkkoja, kielestä riippumattomia luokittelusuorituksia, jotka vaihtelevat 85 prosentista 97 prosenttiin kolmella tutkittavalla alueella. Data-analyysi osoittaa, että on olemassa kaikukammioita, jotka eristävät vastakkaiset kansalliset identiteetit toisistaan.

**Tulos**

Sosiaalisen median käyttäjien kannanottojen luokittelu itsenäisyysliikkeissä

**Esimerkki 2.991**

Tekstien luokittelu on nykyään haastava ja kuuma ala, ja sillä on suuri merkitys tekstien luokittelusovelluksissa. Tällä alalla on tehty paljon tutkimustyötä, mutta tekstidokumenttien kokoelma on luokiteltava toisensa poissulkeviin luokkiin poimimalla käsitteitä tai piirteitä valvotun oppimisen paradigman ja erilaisten luokittelualgoritmien avulla. Tässä asiakirjassa ehdotetaan uutta sumeaan samankaltaisuuteen perustuvaa käsitteiden louhintamallia (FSCMM), jolla joukko tekstidokumentteja voidaan luokitella ennalta määriteltyihin luokkaryhmiin (CG) tarjoamalla niille koulutusta ja valmistelua lause-, asiakirja- ja integroitujen korpusten tasoilla sekä ominaisuuksien vähentämistä ja epäselvyyksien poistamista kullakin tasolla järjestelmän korkean suorituskyvyn saavuttamiseksi. Fuzzy Feature Category Similarity Analyzer (FFCSA) -analysaattoria (FFCSA) käytetään analysoimaan kutakin integroidun korporaation ominaisuusvektorin (ICFV) uutettua ominaisuutta vastaavien luokkien tai luokkien kanssa. Tässä mallissa käytetään tukivektorikone-luokittelijaa (SVMC) luokittelemaan harjoitustietokuviot oikein kahteen ryhmään, + 1 ja - 1, ja näin saadaan tarkkoja ja oikeita tuloksia. Ehdotettu malli toimii tehokkaasti ja tuloksellisesti, ja sen suorituskyky ja tarkkuus ovat korkeat. Avainsanat-tekstien luokittelu; luonnollisen kielen käsittely; ominaisuuksien louhinta; käsitteiden louhinta; sumea samankaltaisuusanalysaattori; ulottuvuuden vähentäminen; lausetaso; asiakirjataso; integroitu korporaatiotason käsittely.

**Tulos**

Sumeaan samankaltaisuuteen perustuva käsitteiden louhintamalli tekstin luokittelua varten

**Esimerkki 2.992**

Taustalla olevan puheen vääristyminen on yleinen ongelma yksikanavaisissa puheenparannusalgoritmeissa, ja se estää tällaisten menetelmien laajemman käytön. Ehdotetaan sanakirjaan perustuvaa puheenparannusmenetelmää, jossa korostetaan taustalla olevan puheen säilyttämistä. Puhtaasta puheesta otetaan näytteitä ja klusteroidaan spektrilaikkuja sanakirjan kouluttamiseksi. Kun kyseessä on meluisa puheen spektrilaatikko, valitaan sanakirjaan parhaiten sopiva merkintä, jota käytetään arvioimaan kohinan voimakkuutta kussakin aikataajuusbinissä. Melun estimointivaihe muotoillaan outlierien havaitsemisongelmaksi, jossa melun oletetaan esiintyvän kussakin binäärissä vain, jos se on outlieri parhaaseen vastaavaan sanakirjamerkintään nähden vastaavassa binäärissä. Tässä kehyksessä asetetaan etusijalle sellaisten spektrielementtien poistaminen, jotka poikkeavat voimakkaasti koulutettuun sanakirjaan tallennetusta tyypillisestä puhutusta yksiköstä. Jopa ilman erillistä kohinamallia tällä menetelmällä voidaan vähentää merkittävästi kohinaa erilaisissa epästationäärisissä kohinatyypeissä ja säilyttää samalla tehokkaasti taustalla oleva puhe haastavammissa meluisissa ympäristöissä.

**Tulos**

YKSIKANAVAISEN PUHEEN PARANTAMINEN POIKKEAMIEN HAVAITSEMISEN AVULLA

**Esimerkki 2.993**

Päätösteoreettisessa ongelmanratkaisussa on kyse tietyn ongelman, kuten monimutkaisen ihmisen tekemän laitteen korjaamisen, odotettujen kustannusten minimoimisesta. Tässä artikkelissa tarkastelemme tilanteita, joissa joudutaan purkamaan osa laitteesta, jotta päästään käsiksi tiettyihin klustereihin ja toimenpiteisiin. Tarkemmin sanottuna tutkimme vianetsintää, jossa klustereista koostuvassa puussa on riippumattomia toimintoja, joissa klusterin sisällä olevia toimintoja ei voi suorittaa ennen kuin klusteri on avattu. Ongelma ei ole triviaali, koska klusterin avaamiseen ja sulkemiseen liittyy kustannuksia. Vianmääritys, jossa on riippumattomia toimintoja ja jossa ei ole klustereita, voidaan ratkaista O(n - lg n)-ajassa (n on toimintojen lukumäärä) Kadanen ja Simonin tunnetulla P-over-C-algoritmilla, mutta puuklusterimallille ei ole vielä löydetty tehokasta ja optimaalista algoritmia. Tässä artikkelissa kuvaamme "bottom-up P-over-C" -algoritmin, joka on O(n - lg n)-aikainen, ja osoitamme, että se on optimaalinen silloin, kun klustereita ei tarvitse sulkea, jotta voidaan testata, ovatko toimet ratkaisseet ongelman.

**Tulos**

Kustannusklustereiden vianmäärityksen kustannukset sisäpiiritiedon avulla

**Esimerkki 2.994**

Vakaan mallin (SM) semantiikasta puuttuvat olemassaolon, relevanssin ja kumulatiivisuuden ominaisuudet. Jos tarkastelemme ennakoivasti SM-semantiikan konservatiivisten laajennusten luokkaa (eli semantiikoita, joissa jokaiselle normaalille logiikkaohjelmalle P haetaan P:n stabiilien mallien joukon yläjoukko), voidaan miettiä, miten tämän luokan semantiikat käyttäytyvät edellä mainittujen ominaisuuksien suhteen. Tällaista kysymystä käsitellään tässä artikkelissa. Määrittelemme suuren luokan SM-semantiikan konservatiivisia laajennuksia, joita kutsutaan nimellä affix stable model semantics (ASM), ja tutkimme edellä mainittuja ominaisuuksia luokan ASM kahdessa ei-disjointisessa alaryhmässä, joita kutsutaan tässä nimillä ASMh ja ASMm. Tästä tutkimuksesta seuraa joukko tuloksia, jotka helpottavat semantiikan arviointia luokassa ASMh ∪ASMm olemassaolon, relevanssin ja kumulatiivisuuden ominaisuuksien suhteen ja paljastavat samalla näiden ominaisuuksien välisiä suhteita. Työmme lähestymistavan tuloksena valotetaan SM-semantiikan karakterisointia, sillä osoitamme, että ominaisuudet (olemassaolon puuttuminen) ja (varovaisen monotonisuuden puuttuminen) ovat ekvivalentteja, mikä on vastoin kirjallisuudesta löytyviä tätä asiaa koskevia väitteitä. Lisäksi luonnehdimme SM-semantiikan relevanssihäiriötä selkeämmin kuin kirjallisuudessa yleensä esitetään.

**Tulos**

Vakaan mallin semantiikan laajennusten ominaisuudet

**Esimerkki 2.995**

Tässä artikkelissa käsitellään menetelmää, jolla laajennettuun relaatiotietomalliin perustuva todennäköisyyspohjainen päättelyjärjestelmä toteutetaan. Tämä malli tarjoaa yhtenäisen lähestymistavan erilaisiin sovelluksiin, kuten dynaamiseen ohjelmointiin, harvojen lineaaristen yhtälöiden ratkaisemiseen ja rajoitusten etenemiseen. Tässä kehyksessä todennäköisyysmalli esitetään yleistettynä relaatiotietokantana. Myöhemmin esitettävät todennäköisyyskyselyt voidaan käsitellä tavanomaisina relaatiokyselyinä. Tavanomaiset tietokantojen hallintajärjestelmät voidaan helposti ottaa käyttöön tällaisen approksimatiivisen päättelyjärjestelmän toteuttamiseksi.

**Tulos**

Menetelmä todennäköisyysmallin toteuttamiseksi relaatiotietokantana.

**Esimerkki 2.996**

Monikielisessä kysymysten vastaamisessa kysymys on joko käännettävä asiakirjan kielelle tai päinvastoin. Suunnan lisäksi käännöksen suorittamiseen on olemassa useita menetelmiä, joista neljää tarkastelemme tässä artikkelissa: sanapohjainen, 10-best-, kontekstipohjainen ja kielioppipohjainen. Rakennamme ominaisuuden kullekin käännössuunnan ja menetelmän yhdistelmälle ja koulutamme mallin, joka oppii optimaaliset ominaisuuksien painotukset. Suuressa foorumitietokannassa, joka koostuu englannin-, arabian- ja kiinankielisistä viesteistä, uudenlainen kääntämään oppiva lähestymistapamme oli tehokkaampi kuin vahva perusmalli (p < 0,05): käännetään kaikki teksti englanniksi ja koulutetaan sitten luokittelija vain englanninkielisen (alkuperäisen tai käännetyn) tekstin perusteella.

**Tulos**

Kääntämisen oppiminen monikielistä kysymyksiin vastaamista varten

**Esimerkki 2.997**

Tilastollisen relationaalisen oppimisen (SRL) menetelmät ovat osoittaneet, että luokittelutarkkuutta voidaan parantaa integroimalla näytteiden välisiä suhteita. Iteratiivisen luokittelun tai relaksoivan merkitsemisen kaltaisilla tekniikoilla tämä saavutetaan levittämällä tietoa toisiinsa liittyvien näytteiden välillä päättelyprosessin aikana. Kun vain muutama näyte on merkitty ja näytteiden väliset yhteydet ovat harvoja, kollektiiviset päättelymenetelmät ovat parantaneet huomattavasti tuloksia verrattuna tavanomaisiin ominaisuuksiin perustuviin ML-menetelmiin. Toisin kuin ominaisuuksiin perustuva ML, kollektiiviset päättelymenetelmät vaativat kuitenkin monimutkaisia päättelymenettelyjä ja ovat usein riippuvaisia vahvasta oletuksesta, jonka mukaan toisiinsa liittyvien näytteiden merkinnät ovat johdonmukaisia. Tässä artikkelissa esitellään uusia relaatio-ominaisuuksia vakiomuotoisia ML-menetelmiä varten ottamalla tietoa suorista ja epäsuorista suhteista. Osoitamme empiirisesti kolmella vakiomuotoisella vertailutietoaineistolla, että relaatio-ominaisuuksiemme avulla saadaan kollektiivisiin päättelymenetelmiin verrattavia tuloksia. Lopuksi osoitamme, että ehdotuksemme päihittää nämä menetelmät, kun lisätietoa on saatavilla.

**Tulos**

Graafipohjaiset relaatio-ominaisuudet kollektiivista luokittelua varten

**Esimerkki 2.998**

Tässä asiakirjassa kuvataan uusi lähestymistapa, jonka avulla voidaan luoda yksinkertaistettu kaksijalkaisen kävely. Lähestymistavassa yhdistetään klassinen dynaaminen mallinnus ja käänteisen kinematiikan ratkaisija, joka perustuu hiukkasparvioptimointiin (Particle Swarm Optimization, PSO). Ensin käytetään käänteistä heiluria, IP, robotin yksinkertaistetun dynaamisen mallin saamiseksi ja kaksoiskärjen liikkumisen kannalta keskeisen pisteen, massakeskipisteen, COM, tavoiteasennon laskemiseksi. Ehdotettu algoritmi, jota kutsutaan IK-PSO:ksi (Inverse Kinematics PSO, käänteinen kinemaattinen PSO), palauttaa käänteisen kinemaattisen ratkaisun, joka vastaa kyseistä COM-arvoa nivelrajoituksia kunnioittaen. Tässä asiakirjassa käytetään inertiapainon PSO-muunnosta tuottamaan mahdollinen ratkaisu stabiilisuuteen perustuvan kuntoarvofunktion ja nivelten liikerajoitusten joukon mukaisesti. Menetelmää sovelletaan menestyksekkäästi jalkojen liikkeen tuottamiseen. Koska ehdotus perustuu ennalta laskettuun COM:iin, joka täyttää kaksijalkaisen vakauden vaatimukset, sen avulla voidaan suunnitella myös kävelyä, jota sovelletaan pienikokoiseen kaksijalkaiseen robottiin. Yleiset termit Robotiikka, robottimallinnus, laskennallinen älykkyys.

**Tulos**

IK-PSO, PSO Inverse Kinematics Solver ja sen soveltaminen biped Gait Generation -menetelmään.

**Esimerkki 2.999**

Tarkastelemme manipulointiongelmia, joissa manipuloijalla on vain osittaista tietoa muiden kuin manipuloijien äänistä. Tällaista osittaista tietoa kuvataan informaatiojoukolla, joka on joukko manipuloimattomien profiilien profiileja, jotka ovat manipuloijalle erottamattomia. Kun otetaan huomioon tällainen informaatiojoukko, hallitseva manipulaatio on manipuloijan antama ei-totuudenmukainen ääni, joka tekee voittajasta vähintään yhtä mieluisan (ja joskus jopa mieluisamman) kuin voittaja, kun manipulaattori äänestää totuudenmukaisesti. Kun manipulaattorilla on täydet tiedot, hallitsevan manipulaation olemassaolon laskeminen on P:ssä monille yleisille äänestyssäännöille (tunnettujen tulosten mukaan). Osoitamme, että kun manipulaattorilla ei ole tietoa, hallitsevaa manipulaatiota ei ole olemassa monille yleisille äänestyssäännöille. Kun manipulaattorin tiedot esitetään osittaisjärjestyksillä ja vain pieni osa preferensseistä on tuntemattomia, hallitsevan manipulaation laskeminen on NP-vaikeaa monille yleisille äänestyssäännöille. Tuloksemme valaisevat siis sitä, voidaanko strategista käyttäytymistä estää rajoittamalla tietoa muiden äänestäjien äänistä.

**Tulos**

Hallitsevat manipulaatiot äänestyksissä, joissa on osittaista tietoa

**Esimerkki 2.1000**

Laajamittaisen koneoppimisen hajautetut optimointialgoritmit kärsivät tiedonsiirron pullonkaulasta. Viestinnän vähentäminen tekee eri koneiden osatöiden tehokkaasta yhdistämisestä haastavampaa. Tässä artikkelissa esittelemme uudenlaisen yleistyksen hiljattain kehitetystä kommunikaatiotehokkaasta primaari-duaalikoordinaattitason nousukehyksestä (COCOA). Kehyksemme, COCOA+, sallii paikallisten päivitysten additiivisen yhdistämisen globaaleihin parametreihin jokaisella iteraatiokerralla, kun taas aiemmat järjestelmät sallivat vain konservatiivisen keskiarvon muodostamisen. Annamme vahvemmat (primaali-duaali-) konvergenssinopeustakuut sekä COCOA:lle että uusille muunnoksillemme ja yleistämme molempien menetelmien teorian kattamaan myös ei-sileät konveksit tappiofunktiot. Esitämme laajan kokeellisen vertailun useilla todellisilla hajautetuilla tietokokonaisuuksilla, jotka osoittavat huomattavasti parempaa suorituskykyä, erityisesti kun koneiden määrää kasvatetaan.

**Tulos**

Lisääminen vs. keskiarvoistaminen hajautetussa primääri-divisio-optimoinnissa

**Esimerkki 2.1001**

Kielen ymmärtäminen kulkee käsi kädessä kyvyn kanssa integroida monimutkaista kontekstuaalista tietoa, joka on saatu havaitsemalla. Tässä työssä esittelemme uudenlaisen tehtävän perusteltua kielellistä ymmärtämistä varten: lauseen erottaminen, kun on annettu visuaalinen kohtaus, joka kuvaa yhtä lauseen mahdollisista tulkinnoista. Tätä varten esittelemme uuden multimodaalisen korpuksen, joka sisältää monitulkintaisia lauseita, jotka edustavat monenlaisia syntaktisia, semanttisia ja diskursiivisia monitulkintaisuuksia, sekä videoita, jotka visualisoivat kunkin lauseen eri tulkintoja. Ratkaisemme tämän tehtävän laajentamalla visuaalista mallia, joka määrittää, onko lause kuvattu videolla. Osoitamme, miten tällaista mallia voidaan mukauttaa tunnistamaan saman lauseen eri tulkintoja, jolloin lauseet voidaan erottaa yhtenäisellä tavalla eri moniselitteisyystyyppien välillä.

**Tulos**

Ymmärrätkö mitä tarkoitan? Kielellisten epäselvyyksien visuaalinen ratkaiseminen

**Esimerkki 2.1002**

Multimodaalinen tunneanalyysi on yhä suositumpi tutkimusalue, joka laajentaa perinteistä kielipohjaista tunneanalyysin määritelmää multimodaaliseen asetelmaan, jossa kielen rinnalla on muitakin merkityksellisiä modaliteetteja. Tässä artikkelissa esitämme multimodaalisen sentimenttianalyysin ongelman modaliteetin sisäisen ja modaliteettien välisen dynamiikan mallintamisena. Esittelemme uudenlaisen mallin, nimeltään Tensor Fusion Network, joka oppii molemmat dynamiikat alusta loppuun. Ehdotettu lähestymistapa on räätälöity verkkovideoiden puhutun kielen sekä siihen liittyvien eleiden ja äänen epävakaaseen luonteeseen. Kokeissa mallimme päihittää nykyiset lähestymistavat sekä multimodaalisessa että unimodaalisessa tunneanalyysissä.

**Tulos**

Tensor-fuusioverkko multimodaaliseen tunneanalyysiin

**Esimerkki 2.1003**

Ehdotamme moniagenttijärjestelmää, jonka osajoukkona on syöttöverkot ja joka on vapaa kerrosrakenteesta matriisi-vektorijärjestelmän avulla. Syviä verkkoja verrataan usein aivojen aivokuoreen tai visuaaliseen havaintojärjestelmään. Yksi suurimmista eroista ihmisaivoihin on matriisi-vektorikertolaskennan käyttö kerrosrakenteeseen perustuen. Se auttaisi ymmärtämään ihmisaivojen toimintatapaa, jos onnistuisimme kehittämään hyvän syväverkkomallin ilman kerrosarkkitehtuuria säilyttäen samalla niiden suorituskyvyn. Aivojen neokorteksi toimii neuronien paikallisten vuorovaikutusten aggregointina, joka on melko samanlainen kuin moniagenttijärjestelmä, joka koostuu autonomisista, osittain havainnoivista agenteista kuin yksiköistä, jotka on kohdistettu pylväsvektoreihin ja joita manipuloidaan globaalin tason algoritmilla. Siksi oletamme, että se on tehokas lähestymistapa biologisesti uskottavamman mallin kehittämiseksi säilyttäen samalla yhteensopivuuden syvien verkkojen kanssa, jotta yksiköt voidaan vuorotella useiden agenttien kanssa. Menetelmämme etuna on myös skaalautuvuus ja muistin tehokkuus. Toteutimme uudelleen Stacked Denoising Autoencoder(SDAE) -menetelmän konkreettisena esimerkkinä moniagenttijärjestelmämme kanssa ja tarkistimme sen vastaavuuden vakiomuotoisen SDAE:n kanssa sekä teoreettisesta että empiirisestä näkökulmasta. Lisäksi ehdotimme myös moniagenttisen SDAE:n muunnelmaa nimeltä "Sparse Connect SDAE" ja osoitimme sen laskennallisen edun MNIST-tietokannan avulla.

**Tulos**

MONIAGENTTIJÄRJESTELMÄ KERROKSETTOMASSA VERKOSSA

**Esimerkki 2.1004**

Rekursiiviset neuromallit, jotka käyttävät syntaktisia jäsennyspuita tuottaakseen rekursiivisesti representaatioita alhaalta ylöspäin jäsennyslapsista, ovat suosittuja uusia arkkitehtuureja, jotka lupaavat vangita rakenteellisia ominaisuuksia, kuten negaation laajuuden tai pitkien etäisyyksien semanttiset riippuvuudet. Avoimena on kuitenkin edelleen kysymys siitä, mihin tehtäviin tämä jäsennykseen perustuva menetelmä soveltuu. Tässä artikkelissa vertailemme rekursiivisia neuromalleja sekventiaalisiin rekursiivisiin neuromalleihin, jotka rakentuvat pelkästään sanasekvenssien perusteella. Tutkimme viittä tehtävää: (1) lauseiden ja (2) syntaktisten lausekkeiden tunnearvojen luokittelu, (3) kysymyksiin vastaaminen, (4) diskurssien jäsentäminen, (5) semanttiset suhteet (esim. substantiivien väliset komponentti-kokonaisuus-suhteet); Huomaamme, että rekursiiviset mallit suoriutuvat rekursiivisista malleista yhtä hyvin tai paremmin kuin rekursiiviset mallit kaikissa tehtävissä yhtä lukuun ottamatta yhtä tehtävää: nominien väliset semanttiset suhteet. Analyysimme osoittaa, että sekventiaaliset mallit selviytyvät hyvin tehtävistä, jotka perustuvat negaation laajuuteen (kuten sentimentit). Rekursiiviset mallit auttavat vain tehtävissä, jotka edellyttävät sanojen välisten pitkien etäisyyksien suhteiden esittämistä. Tuloksemme tarjoavat näkemyksiä representaatioiden oppimiseen tarkoitettujen neuraalisten arkkitehtuurien suunnittelusta.

**Tulos**

Milloin puurakenteet ovat välttämättömiä representaatioiden syväoppimisessa?

**Esimerkki 2.1005**

Tässä artikkelissa ehdotamme ja tutkimme neuroverkkojen uutta muistirakennetta nimeltä Hierarchical Attentive Memory (HAM). Se perustuu binääripuuhun, jonka lehdet vastaavat muistisoluja. Tämän ansiosta HAM pystyy suorittamaan muistin käytön Θ(logn)-kompleksisuudella, mikä on merkittävä parannus verrattuna vakiomuotoiseen tarkkaavaisuusmekanismiin, joka vaatii Θ(n)-operaatioita, jossa n on muistin koko. Osoitamme, että HAM:lla täydennetty LSTM-verkko voi oppia algoritmeja ongelmiin, kuten yhdistämiseen, lajitteluun tai binäärihakuun, puhtaista input-output-esimerkeistä. Erityisesti se oppii lajittelemaan n lukua ajassa Θ(n logn) ja yleistyy hyvin syötesarjoihin, jotka ovat paljon pidempiä kuin koulutuksen aikana nähdyt. Osoitamme myös, että HAM voidaan kouluttaa toimimaan klassisten tietorakenteiden tavoin: pino, FIFO-jono ja prioriteetti-jono.

**Tulos**

Tehokkaiden algoritmien oppiminen hierarkkisen tarkkaavaisen muistin avulla

**Esimerkki 2.1006**

Ehdotamme uutta algoritmia regularisoidun empiirisen tappion minimoimiseksi: Stochastic Dual Newton Ascent (SDNA). Menetelmämme on luonteeltaan dualistinen: jokaisessa iteraatiossa päivitämme satunnaisen osajoukon dualismuuttujia. Toisin kuin nykyiset menetelmät, kuten stokastinen dual-koordinaattinousu, SDNA pystyy kuitenkin hyödyntämään kaiken esimerkkien sisältämän kaarevuusinformaation, mikä johtaa silmiinpistäviin parannuksiin sekä teoriassa että käytännössä - joskus jopa suuruusluokkia. Erikoistapauksessa, jossa primäärissä käytetään L2-säännöstelijää, duaaliongelma on kovera kvadraattinen maksimointiongelma, johon on lisätty separoituva termi. Tässä järjestelmässä SDNA ratkaisee jokaisessa vaiheessa proksimaalisen osaongelman, johon liittyy kvadraattisen funktion Hessin satunnainen pääomatriisi; tästä johtuu menetelmän nimi. Jos lisäksi häviöfunktiot ovat kvadraattisia, menetelmämme voidaan tulkita uudeksi muunnelmaksi hiljattain esitellystä Iterative Hessian Sketch -menetelmästä.

**Tulos**

SDNA: Stokastinen kaksois-Newtonin nousu empiiristä riskin minimointia varten

**Esimerkki 2.1007**

Klusterointi on tehokas tiedonlouhintatekniikka, jolla luodaan kiinnostuksen kohteena olevia ryhmiä. Erilaisista klusterointimenetelmistä k-means-algoritmit ja min-cut-algoritmit ovat suosituimpia yksinkertaisuutensa ja tehokkuutensa vuoksi. Klassinen k-means-algoritmi jakaa tietopisteet useisiin osajoukkoihin päivittämällä iteratiivisesti klusterointikeskukset ja niihin liittyvät tietopisteet. Sitä vastoin min-cut-algoritmeissa rakennetaan painotettu suuntaamaton graafi, joka jakaa graafin kärjet kahteen joukkoon. Nykyisillä klusterointialgoritmeilla on kuitenkin taipumus klusteroida datapisteiden vähemmistö osajoukkoihin, mitä on vältettävä, kun kohdetietokanta on tasapainoinen. Tasapainotetun tietokokonaisuuden tarkemman klusteroinnin saavuttamiseksi ehdotamme, että k-means- ja min-cut-menetelmissä käytetään eksklusiivista lasso-menetelmää klusterointitulosten tasapainoasteen säätämiseksi. Optimoimalla eksklusiivisen lasson päälle rakennettuja kohdefunktioitamme voimme tehdä klusterointituloksesta mahdollisimman tasapainoisen. Laajat kokeet useilla laajamittaisilla tietokokonaisuuksilla vahvistavat ehdotettujen algoritmien edut verrattuna uusimpiin klusterointialgoritmeihin.

**Tulos**

Tasapainotettu k-Means ja Min-Cut-klusterointi

**Esimerkki 2.1008**

Tässä raportissa tarkastelemme dynaamista Bayes-verkkoa (DBN) mallina, joka pyrkii sisällyttämään ajallisen ulottuvuuden ja epävarmuuden. Aloitamme DBN:n perusteista, joissa keskitymme erityisesti päättelyn ja oppimisen käsitteisiin ja algoritmeihin. Sen jälkeen esittelemme eri tasoja ja menetelmiä DBN:ien luomiseksi sekä lähestymistapoja ajallisen ulottuvuuden sisällyttämiseksi staattiseen Bayes-verkkoon. AvainsanatDBN, DAG, päättely, oppiminen, HMM, EM-algoritmi, SEM, MLE, kytketyt HMM:t.

**Tulos**

Dynaamisen Bayes-verkon luonnehdinta Dynaaminen Bayes-verkko on ajallinen verkko.

**Esimerkki 2.1009**

Aggregaatteja sisältävät loogiset ohjelmat (LP) ovat yksi tärkeimmistä loogisen ohjelmoinnin (LP) kielellisistä laajennuksista. Tässä työssä ehdotamme perustelemattoman joukon ja perustelemattoman semantiikan käsitteiden yleistämistä ohjelmille, joilla on monotonisia ja antimonotonisia aggregaatteja (LPm,a-ohjelmat). Erityisesti esitämme LPm,a-ohjelmille uuden perusteettoman joukon käsitteen, joka on järkevä yleistys alkuperäisestä määritelmästä tavalliselle (aggregaatittomalle) LP:lle. Tältä pohjalta määrittelemme LPm,a-ohjelmille hyvin perustellun operaattorin, jonka kiintopistettä kutsutaan hyvin perustelluksi malliksi (tai hyvin perustelluksi semantiikaksi) LPm,a-ohjelmille. Tässä yleistyksessä säilytetään perustelemattomien joukkojen tärkeimmät ominaisuudet ja tavallisen LP:n perustelemattoman semantiikan tärkeimmät ominaisuudet, erityisesti perustelemattoman mallin olemassaolo ja yksikäsitteisyys sekä vahva suhde LPm,a-ohjelmien vastausjoukkojen semantiikkaan. Osoitamme, että yksi D̃-well-founded-semantiikka, jonka Pelov, Denecker ja Bruynooghe ovat määritelleet laajemmalle aggregaattiluokalle, jossa käytetään approksimointioperaattoreita, on yhteneväinen tässä työssä LPm,a-ohjelmille määritellyn well-founded-mallin kanssa. Käsittelemme myös joitakin kompleksisuuskysymyksiä, ja ennen kaikkea annamme muodollisen todisteen LPm,a-ohjelmien hyvin perustellun mallin laskettavuudesta. Lisäksi todistamme, että yleisille LP-ohjelmille, jotka voivat sisältää aggregaatteja, jotka eivät ole monotonisia eivätkä antimonotonisia, aggregaatti-ilmaisujen tyydyttämisen päättäminen osittaisten tulkintojen suhteen on coNP-täydellinen. Tästä seuraa, että yleisille LP-ohjelmille ei todennäköisesti ole olemassa hyvin perusteltua semantiikkaa, joka mahdollistaisi helposti laskettavan laskennan, mikä oikeuttaa rajoituksen LPm,a-ohjelmiin. Lopuksi esittelemme DLV:tä laajentavan prototyyppijärjestelmän, joka tukee hyvin perusteltua semantiikkaa LPm,a-ohjelmille, ja on tätä kirjoitettaessa ainoa toteutettu järjestelmä, joka tekee niin. Tällä prototyypillä tehdyt kokeet osoittavat, että aggregaattikonstruktioilla on merkittäviä laskennallisia etuja verrattuna vastaaviin aggregaatittomiin koodauksiin.

**Tulos**

Aggregaatteja sisältävien vastausjoukko-ohjelmien rajoittamattomat joukot ja hyvin perusteltu semantiikka

**Esimerkki 2.1010**

Stokastinen gradientti-MCMC (Stochastic gradient MCMC, SG-MCMC) on ollut tärkeässä asemassa laajamittaisessa Bayes-oppimisessa, ja sen teoreettiset konvergenssiominaisuudet ovat hyvin kehittyneet. Tällaisissa SG-MCMC-sovelluksissa on yhä suositumpaa käyttää hajautettuja järjestelmiä, joissa stokastiset gradientit lasketaan joidenkin vanhentuneiden parametrien perusteella, jolloin saadaan niin sanottuja vanhoja gradientteja. Vaikka vanhentuneita gradientteja voidaan käyttää suoraan SG-MCMC:ssä, niiden vaikutusta konvergenssiominaisuuksiin ei ole tutkittu hyvin. Tässä artikkelissa kehitämme teoriaa osoittaaksemme, että vaikka SG-MCMC-algoritmin harha ja MSE riippuvat stokastisten gradienttien vanhentuneisuudesta, sen estimointivarianssi (suhteessa odotettuun estimaattiin, joka perustuu määrättyyn määrään näytteitä) on siitä riippumaton. Yksinkertaisessa Bayesin hajautetussa järjestelmässä, jossa SG-MCMC:n avulla joukko työntekijöitä laskee vanhoja gradientteja asynkronisesti, teoriamme osoittaa lineaarisen nopeutumisen estimointivarianssin pienenemisessä työntekijöiden lukumäärän mukaan. Synteettisellä datalla ja syvillä neuroverkoilla tehdyt kokeet vahvistavat teoriamme ja osoittavat SG-MCMC:n tehokkuuden ja skaalautuvuuden vanhentuneiden gradienttien kanssa.

**Tulos**

Stokastinen gradientti MCMC ja vanhentuneet gradientit (Stale Gradients)

**Esimerkki 2.1011**

Laajennamme karkeisiin joukkoihin perustuvaa standardimenetelmää käsittelemään valtavia määriä numeerisia attribuutteja verrattuna pieneen määrään käytettävissä olevia objekteja. Tässä ehdotetaan uudenlaista lähestymistapaa klusterointiin yhdessä ulottuvuuden vähentämisen kanssa; hybridi Fuzzy C Means-Quick Reduct (FCMQR) -algoritmia yhden geenin valintaa varten. Geenivalinta on prosessi, jossa valitaan informatiivisempia geenejä. Se on yksi tärkeistä vaiheista tiedon löytämisessä. Ongelmana on, että kaikki geenit eivät ole tärkeitä geeniekspressiotiedoissa. Osa geeneistä voi olla tarpeettomia ja osa voi olla merkityksettömiä ja kohinaisia. Tässä tutkimuksessa koko tietokokonaisuus jaetaan samankaltaisten geenien asianmukaisiin ryhmiin soveltamalla Fuzzy C Means (FCM) -algoritmia. Korkean luokan erottelukykyiset geenit on valittu niiden riippuvuusasteen perusteella soveltamalla kaikkiin tuloksena syntyneisiin klustereihin karkeaan joukko-oppiin perustuvaa Quick Reduct -algoritmia. Keskimääräinen korrelaatioarvo (Average Correlation Value, ACV) lasketaan korkean luokan erotteleville geeneille. Klusterit, joiden ACV-arvo on 1, määritetään merkittäviksi klustereiksi, joiden luokittelutarkkuus on yhtä suuri tai suuri verrattuna koko aineiston tarkkuuteen. Ehdotettua algoritmia arvioidaan WEKA-luokittimien avulla ja verrataan niitä keskenään. Lopuksi leukemiasyöpäaineistoon liittyvät kokeelliset tulokset vahvistavat, että lähestymistapamme on varsin lupaava, vaikka se vaatii varmasti lisätutkimuksia.

**Tulos**

Uudenlainen lähestymistapa yksittäisen geenin valintaan klusterointia ja ulottuvuuden vähentämistä käyttäen

**Esimerkki 2.1012**

Sanan tunnetila riippuu siitä, millä alalla sitä käytetään. Yhteiskuntatieteiden laskennallisessa tutkimuksessa tarvitaan siis sentimenttisanakirjaa, joka on ominaista tutkittaville aloille. Yhdistämme alakohtaiset sanojen upotukset ja etiketin etenemismenetelmän, jotta voimme tuottaa tarkkoja alakohtaisia tunnesanastoja käyttämällä pieniä siemensanajoukkoja. Osoitamme, että lähestymistapamme on huippuluokkaa tunnesanastojen tuottamisessa aluespesifisistä korpuksista ja että puhtaasti korpuspohjainen lähestymistapamme on parempi kuin menetelmät, jotka perustuvat käsin kuratoituihin resursseihin (esim. WordNet). Kehyksemme avulla tuotamme ja julkaisemme historiallisia sentimenttisaksikoita 150 vuoden englannin kielelle ja yhteisökohtaisia sentimenttisaksikoita 250 verkkoyhteisölle sosiaalisen median Reddit-foorumilta. Indusoimamme historialliset sanakirjat osoittavat, että yli 5 prosenttia sentimenttiä sisältävistä (ei-neutraaleista) englanninkielisistä sanoista on vaihtanut täysin napaisuutta viimeisten 150 vuoden aikana, ja yhteisökohtaiset sanakirjat korostavat, miten sentimentti vaihtelee huomattavasti eri yhteisöjen välillä.

**Tulos**

Aluekohtaisten aatesanastojen tuottaminen merkitsemättömistä korpuksista

**Esimerkki 2.1013**

Tiedonhaussa on yritetty useaan otteeseen ratkaista erilaisia taivutuksen yhteensovittamiseen liittyviä ongelmia. Stemming on yleinen lähestymistapa tähän tarkoitukseen. Monista stemming-tekniikoista tilastollinen stemming on osoittautunut tehokkaaksi useissa kielissä, erityisesti voimakkaasti taivutetuissa kielissä. Tässä artikkelissa ehdotamme menetelmää, jolla löydetään affiksit sanan eri paikoissa. Yleiset tilastolliset tekniikat perustuvat pitkälti merkkijonojen samankaltaisuuteen etuliitteiden ja suffiksien yhteensovittamisessa. Koska infiksit ovat yleisiä epäsäännöllisissä/epämuodollisissa taivutuksissa morfologisesti monimutkaisissa teksteissä, on tarpeen löytää infiksit stemmingia varten. Tässä artikkelissa ehdotamme menetelmää, jonka tavoitteena on löytää tilastollisia taivutussääntöjä, jotka perustuvat sanaparien pienimmän muokkausetäisyyden taulukkoon ja sääntöjen todennäköisyyksiin kielessä. Näitä sääntöjä käytetään sanojen tilastolliseen kantamiseen, ja niitä voidaan käyttää erilaisissa tekstinlouhintatehtävissä. Kokeelliset tulokset CLEF 2008 ja CLEF 2009 English-Persian CLIR -tehtävissä osoittavat, että ehdotettu menetelmä on MAP-arvoltaan huomattavasti parempi kuin kaikki perusmenetelmät.

**Tulos**

SS4MCT: Morfologisesti monimutkaisten tekstien tilastollinen stemmeri

**Esimerkki 2.1014**

Viimeaikainen lisääntynyt kiinnostus tekoälyn etiikkaan saattaa jättää monet opettajat miettimään, miten käsitellä moraalisia, eettisiä ja filosofisia kysymyksiä tekoälykursseillaan. Kouluttajina haluamme kehittää opetussuunnitelmia, jotka valmistavat opiskelijoita paitsi tekoälyn ammattilaisiksi myös ymmärtämään niitä moraalisia, eettisiä ja filosofisia vaikutuksia, joita tekoälyllä tulee olemaan yhteiskuntaan. Tässä artikkelissa tarjoamme tekoälyn opettajille käytännön tapausesimerkkejä ja linkkejä resursseihin. Annamme myös konkreettisia ehdotuksia siitä, miten tekoälyn etiikka voidaan sisällyttää yleiseen tekoälykurssiin ja miten voidaan opettaa itsenäistä tekoälyn etiikan kurssia.

**Tulos**

Tekoälykurssien eettiset näkökohdat

**Esimerkki 2.1015**

Kulttuurin vaikutus visuaalisten tunteiden havaitsemiseen on viime aikoina herättänyt huomiota multimediatutkimuksessa. Tässä tutkimuksessa tarjoamme tehokkaita laskennallisen kielitieteen työkaluja, joiden avulla voidaan tutkia, hakea ja selata 16 000 monikielisen affektiivisen visuaalisen käsitteen ja 7,3 miljoonan Flickr-kuvan muodostamaa tietokokonaisuutta. Ensin suunnitellaan tehokas joukkoistamiskokeilu, jolla kerätään visuaalisiin käsitteisiin liittyviä ihmisten tunnearvioita. Sen jälkeen käytämme sanojen upotuksia näiden käsitteiden esittämiseen matalaulotteisessa vektoriavaruudessa, jolloin voimme laajentaa käsitteiden ympärillä olevia merkityksiä ja saada näin tietoa eri kielten välisistä yhtäläisyyksistä ja eroista. Vertailemme erilaisia käsite-esityksiä uuden arviointitehtävän avulla, joka perustuu visuaalisen semanttisen sukulaisuuden käsitteeseen. Näiden representaatioiden perusteella suunnittelemme klusterointijärjestelmiä monikielisten visuaalisten käsitteiden ryhmittelyyn ja arvioimme niitä uusilla metriikoilla, jotka perustuvat joukkoistettuihin tunnemerkintöihin ja visuaaliseen semanttiseen sukulaisuuteen. Ehdotetun klusterointikehyksen avulla voimme analysoida koko monikielisen tietokokonaisuuden perusteellisesti, ja esitämme myös sovelluksen kasvotietojen osajoukkoon, jossa tutkitaan muotokuviin liittyvien affektiivisten visuaalisten käsitteiden kulttuurisia oivalluksia.

**Tulos**

Monikielinen visuaalinen tunnekäsitteiden vertailu

**Esimerkki 2.1016**

Tässä artikkelissa esitellään koneoppimisratkaisuja käytännön ongelmaan, joka liittyy luonnollisen kielen tuottamiseen (Natural Language Generation, NLG), erityisesti sananmuodostukseen tamilin kaltaisissa agglutinatiivisissa kielissä, valvotulla tavalla. Morfologinen generaattori on tärkeä osa tekoälyn luonnollisen kielen käsittelyä. Se tuottaa sanamuotoja, kun juuren ja liitetiedostojen perusteella. Morfofoneemiset muutokset, kuten lisäys, poisto, vuorottelu jne., tapahtuvat, kun kaksi tai useampi morfeemi tai sana liitetään yhteen. Sandhi-säännöt olisi määriteltävä eksplisiittisesti sääntöpohjaisissa morfologisissa analysaattoreissa ja generaattoreissa. Koneoppimisessa järjestelmä voi oppia nämä säännöt automaattisesti harjoitusnäytteistä ja soveltaa niitä sitten uusiin syötteisiin. Tässä artikkelissa ehdotimme koneoppimismalleja, jotka oppivat substantiivien deklinaatioiden morfofoneemiset säännöt annetusta harjoitusaineistosta. Nämä mallit koulutetaan oppimaan sandhi-sääntöjä käyttämällä erilaisia oppimisalgoritmeja, ja näiden algoritmien suorituskyky esitellään. Tämän perusteella voimme päätellä, että morfologisen prosessoinnin, kuten sanamuotojen muodostamisen, koneoppiminen voidaan menestyksekkäästi oppia valvotulla tavalla ilman sääntöjen eksplisiittistä kuvausta. Keskustellaan päätöspuiden ja Bayesin koneoppimisalgoritmien suorituskyvystä substantiivien deklinaatioissa.

**Tulos**

KONEOPPIMINEN FONOLOGISESTI EHDOLLISTETTUJEN SUBSTANTIIVIEN DEKLINAATIOIDEN OPPIMISESTA TAMILIN MORFOLOGISIA GENERAATTOREITA VARTEN.

**Esimerkki 2.1017**

Kuuluisaa biologisesti inspiroitunutta hierarkkista mallia, jota Riesenhuber ja Poggio ehdottivat ensimmäisen kerran, on sovellettu menestyksekkäästi useisiin visuaalisiin tunnistustehtäviin. Malli kykenee saavuttamaan joukon sijainti- ja mittakaavasietoista tunnistusta, joka on keskeinen ongelma hahmontunnistuksessa. Tässä asiakirjassa esitellään eräiden muiden biologisten koetulosten perusteella muisti- ja assosiaatiomekanismit edellä mainittuun biologisesti inspiroituun malliin. Työn tärkeimmät motiivit ovat a) jäljitellä aktiivista muisti- ja assosiaatiomekanismia ja lisätä "ylhäältä alaspäin" tapahtuva säätö edellä mainittuun biologisesti inspiroituun hierarkkiseen malliin ja b) rakentaa algoritmi, joka voi säästää tilaa ja säilyttää hyvän tunnistussuorituskyvyn. Työn yksityiskohdat voidaan selittää seuraavasti: (1) Kohteiden muistamisprosessissa: Ehdotettu malli jäljittelee joitakin ihmisen muistimekanismin ominaisuuksia seuraavasti: (a) Mallissamme yksi objekti muistetaan semanttisten ominaisuuksien ja erityisten kuvalaikkujen avulla (mikä vastaa episodista muistia). Semanttiset attribuutit kuvaavat kohteen jokaista osaa selkeällä fyysisellä merkityksellä, esimerkiksi ovatko kasvojen silmät ja suut "suuret" vai "pienet" ja niin edelleen. Yksi erityinen laikku valitaan, jos vastaavan semanttisen ominaisuuden arvo on kaukana keskimääräisestä. Laikan tulisi olla kohteen näkyvin osa. (b) Mallissamme yhden objektin eri piirteet (semanttiset ominaisuudet ja erityiset laikut) tallennetaan hajautettuihin paikkoihin ja eri objektien yhteinen piirre tallennetaan aggregoituna, jolloin voidaan oppia luokittelemaan eri objektien samankaltaisten piirteiden erot. Kunkin objektin samankaltaisuuskynnykset voidaan oppia, kun uusia objekteja opitaan. (2) Objektien tunnistamisprosessissa: Biologisessa prosessissa siihen liittyvä tunnistaminen, mukaan lukien tuttuuden erottelu ja muistettava yhteensovittaminen. Ehdotetussa mallissamme, ensin jäljitellään tuttuuden erottelua ("tietäen" vaikka episodi), vertaamme ehdokkaiden erikoislaikkuja tallennettujen kohteiden erikoislaikkujen kanssa käyttämällä edellä mainittua biologisesti inspiroitunutta hierarkkista mallia, jossa ehdokkailla ja tallennetuilla kohteilla on samat näkyvät semanttiset piirteet. Sen jälkeen jäljitellään muistinvaraista yhteensovittamista, jolloin erityisten laikkujen vertailutulokset yhdistetään semanttisten piirteiden vertailuun. Uutta mallia sovelletaan myös esineiden tunnistamisprosesseihin. Ensisijaiset kokeelliset tulokset osoittavat, että menetelmämme on tehokas ja vaatii paljon vähemmän muistia.

**Tulos**

Muistin ja assosiaatiomekanismin tuominen biologisesti inspiroituun visuaaliseen malliin

**Esimerkki 2.1018**

Ehdotamme tehokasta tekniikkaa monimerkkiluokittelua varten, joka perustuu kalibrointiin, jolla tarkoitamme sellaisen linkkifunktion oppimista, joka kuvaa riippumattomat ennusteet yhteisiksi ennusteiksi. Vaikka ehdotuksemme naiivi toteutus edellyttäisi erillisten luokittelijoiden kouluttamista jokaiselle etiketille, osoitamme, että luonnollisissa tietokokonaisuuksissa ja lineaarisissa luokittelijoissa voimme välttää tämän hyödyntämällä satunnaistetun lineaarialgebran tekniikoita. Lisäksi algoritmimme soveltuu yhtä hyvin moniluokkaiseen luokitteluun. Lopputuloksena on algoritmi, joka skaalautuu hyvin suuriin monimerkkisiin ja moniluokkaisiin ongelmiin ja tarjoaa huipputarkkuutta monissa tietokokonaisuuksissa.

**Tulos**

Multilabel-ennuste kalibroinnin avulla

**Esimerkki 2.1019**

Tässä artikkelissa käsitellään SYNTAGMAa, sääntöpohjaista NLP-järjestelmää, joka käsittelee syntaktisen monitulkintaisuuden vähentämisen ja sanojen merkityksen määrittelyn hankalia kysymyksiä sekä tarjoaa innovatiivisia ja omaperäisiä ratkaisuja konstituenttien tuottamiseen ja rajoitteiden hallintaan. Jotta järjestelmän toiminnasta saataisiin käsitys, kuvataan järjestelmän yleinen rakenne ja komponentit sekä sen leksikaaliset, syntaktiset ja semanttiset resurssit. Tämän jälkeen käsitellään mekanismia, joka suorittaa valikoivan jäsennyksen syntaktisen ja semanttisen tiedon vuorovaikutuksen avulla ja johtaa jäsennyksen suorittajan johdonmukaiseen ja tarkkaan tulkintaan syötetekstistä.

**Tulos**

Syntaksin ja semantiikan vuorovaikutuksen jäsentämisstrategiat. SYNTAGMAn sisällä

**Esimerkki 2.1020**

Algoritmisen käsittelyn historiallinen tausta etymologian ja metodologian osalta käännetään matemaattisen logiikan ja tietojenkäsittelytieteen termeiksi. Muodollinen logiikkarakenne esitellään esimerkkikysymysten avulla, jotka esitetään Fiqh-luvuille loogisen kyselykielen määrittelemiseksi. Perustaksi suunnitellaan yleinen algoritmi Fiqhrulings-päätösten tekemiseksi, jotta voidaan mahdollistaa ja tehostaa oikeusvaltioperiaatetta (vs. oikeusvaltioperiaatetta) täydellisellä avoimuudella ja islamilaisen oikeuden täydellisellä algoritmisella kattavuudella, joka lopulta tarjoaa oikeusvarmuuden, oikeudellisen tasa-arvon ja täyden oikeudellisen vastuun. Tämä toteutetaan erottelemalla ja palauttamalla klassinen Fiqh-metodologia (usul al-Fiqh) ensimmäisen kertaluvun logiikan (FOL) osajoukkojen ilmaisuvoiman avulla ja korvaamalla ad hoc -perustelu kestävästi vääristettävällä rationaalisella argumentoinnilla. Tuloksia käsitellään muodollisten Fiqh-järjestelmien täydellisyyden, ratkaistavuuden ja monimutkaisuuden kannalta. Muodostetaan ja validoidaan päätösongelma formaalille Fiqh-järjestelmälle.

**Tulos**

Islamilaisen oikeuskäytännön (Fiqh) algoritmi ja päätösongelman validointi (Entscheidungsproblem)

**Esimerkki 2.1021**

Syvä vahvistusoppimisen (RL) soveltaminen todellisissa järjestelmissä kärsii hitaasta datanäytteenotosta. Ehdotamme parannettua generatiivista vastakkaisverkkoa (EGAN) RL-agentin alustamiseksi, jotta oppiminen nopeutuisi. EGAN hyödyntää tilojen ja toimintojen välistä suhdetta parantaakseen GAN:n tuottamien tietonäytteiden laatua. Agentin esivalmennus EGAN:lla osoittaa jyrkempää oppimiskäyrää, jossa oppimisen alussa oppimisaika paranee 20 prosenttia verrattuna siihen, että esivalmennusta ei olisi tehty, ja parannus GAN:lla tapahtuvaan harjoitteluun verrattuna on noin 5 prosenttia pienemmillä vaihteluilla. Reaaliaikaisissa järjestelmissä, joissa on harva ja hidas tiedonkeruu, EGANia voitaisiin käyttää nopeuttamaan koulutusprosessin alkuvaiheita.

**Tulos**

Tehostettu kokemusten toistaminen tehokasta vahvistusoppimista varten

**Esimerkki 2.1022**

Monet logistiikkatoiminnot liittyvät yritysten ja prosessien välisten materiaalivirtojen yhdistämiseen. Tällaisissa sovelluksissa yhdistyvät määrälliset päätökset, esimerkiksi lähetettävien tavaroiden määrä (varastonhallinta), ja reitityspäätökset, joita käsitellään ajoneuvojen reitityksen alalla. On selvää, että nämä kaksi aluetta leikkaavat toisiaan huomattavassa määrin, mikä vaikeuttaa tällaisten ongelmien ratkaisemista. Viime aikoina tässä yhteydessä on tehty intensiivistä tutkimusta, jota kutsutaan yleisesti varastojen reititysongelmiksi [2, 3] (IRP). IRP:stä löytyy useita muunnelmia, jotka vaihtelevat deterministisistä kysyntätapauksista stokastisiin malleihin. Yritysten käytännön näkökulmasta todellisuus on paljon monimutkaisempi kuin tunnettu kysyntä ja paljon epävarmempi kuin stokastinen laki. Itse asiassa yrityksillä on usein vain osittainen tieto kysynnästä suunnitteluhorisontin aikana. Havaintomme tästä ilmiöstä voidaan muuntaa uudentyyppisiksi tiedoiksi, joita ehdotamme kokeellisia jatkotutkimuksia varten. Oletamme tässä, että kuluvan kauden kysyntä tunnetaan kauden alussa. Lisäksi meillä on likimääräinen yleiskuva kysynnästä 5 seuraavan jakson, 20 seuraavan jakson ja 60 seuraavan jakson ajalta. Tämä yleiskatsaus on melko hyvä (esim. se ei poikkea todellisuudesta enempää kuin ±10 %), mutta emme tietenkään voi ennustaa varmuudella, mitä seuraavilla kausilla tapahtuu. Tämän työn yleistavoitteena on tarjota käytännön optimointimenetelmiä yrityksille, jotka käsittelevät varaston reititysongelmia, ottaen huomioon tämän uudenlaiset tiedot. Yritykset eivät myöskään toisinaan pysty käsittelemään joka jakso muuttuvia suunnitelmia, vaan ne haluaisivat ottaa käyttöön säännöllisiä rakenteita asiakkaiden palvelemiseksi. Koska työmme on pitkäaikainen hanke, aiomme vähitellen kehittää ratkaisukeinoamme. Ensimmäisessä vaiheessa keskitymme varastojen reititysongelmaan, jossa on yksi tuote ja deterministinen tunnettu kysyntä rajallisen ajanjakson aikana. Toisin kuin [1], oletamme, että reitityskustannukset ja varastointikustannukset eivät ole vertailukelpoisia, ja siksi niitä olisi käsiteltävä kahtena eri tavoitteena. Tietojemme mukaan tämä on ensimmäinen kerta, kun kaksitavoitteista lähestymistapaa tarkastellaan tässä ongelmassa.

**Tulos**

Käytännön inventaarion reititys: Ongelman määrittely ja optimointimenetelmä

**Esimerkki 2.1023**

Esitetään tehokas ja intuitiivinen algoritmi puhujien tunnistamiseen pitkästä aineistosta (kuten YouTuben pitkästä keskustelusta, Cocktail-juhlien nauhoitetusta äänestä tai videosta).Automaattisen puhujan tunnistamisen tavoitteena on tunnistaa erilaisten puhujien määrä ja valmistella malli kyseiselle puhujalle uuttamalla, luonnehtimalla ja puhujakohtaisella informaatiolla, joka sisältyy puhesignaaliin. Sillä on monia erilaisia sovelluksia erityisesti seuraavilla aloilla: valvonta , maahanmuutto lentokentällä , tietoverkkoturvallisuus , transkriptio useissa samankaltaisissa äänilähteissä, joissa transkriptiota on vaikea määrittää mielivaltaisesti. Yleisimmin puheparametrisointi käytetään puhujan todentaminen, K-mean, cepstral analyysi, on yksityiskohtainen. Sen jälkeen selitetään Gaussin seosmallinnus, joka on puhujan mallinnustekniikka. Gaussin seosmallit (GMM), ehkä vankin koneoppimisalgoritmi, on esitelty, jotta voidaan tutkia ja arvioida huolellisesti puhujan tunnistamista tekstistä riippumattomassa tekstissä. Gaussin seosmallien soveltaminen tai käyttö seurantaan & Analysointi puhujan identiteetti kannustetaan tuttuus, tietoisuus tai ymmärrys saatu kokemuksen kautta, että Gaussin spektri kuvaavat ominaisuuksia puhujan spektrinen konformatiivinen malli ja merkittävä kyky GMM rakentaa oikukas tiheydet sen jälkeen havainnollistamme "Expectation maksimointi" iteratiivinen algoritmi, joka ottaa jonkin mielivaltaisen arvon alkuarvion ja jatkaa iteratiivista prosessia, kunnes konvergenssi arvo havaitaan Olemme yrittäneet saada 85 ~ 95 %:n tarkkuus käyttäen puhujan mallintaminen vektorikvantisointi ja Gaussin sekoitusmalli ,joten tekemällä useita kokeita saamme 79 - 82 % tunnistustuloksen käyttämällä vektorikvantisointia ja 85 - 92 % tunnistustuloksen käyttämällä vektorikvantisointia.6 % tunnistusasteesta käyttämällä GMM-mallinnusta odotusmaksimointiparametrin estimoinnilla riippuen parametrin vaihtelusta.

**Tulos**

PUHUJAN TUNNISTAMINEN YOUTUBE-TIEDOISTA

**Esimerkki 2.1024**

Tässä artikkelissa esitellään puheentunnistukseen tarkoitettu yksinkertainen alusta loppuun -malli, jossa yhdistyvät konvoluutioverkkoon perustuva akustinen malli ja graafinen dekoodaus. Se on koulutettu tuottamaan kirjaimia, joissa on transkriptoitua puhetta, ilman, että foneemeja on pakko kohdistaa. Esittelemme automaattisen segmentointikriteerin, jonka avulla voidaan kouluttaa sekvenssiannotaatiosta ilman kohdistusta ja joka on CTC:n [6] tasolla, mutta yksinkertaisempi. Osoitamme kilpailukykyisiä tuloksia sanavirheprosentissa Librispeech-korpuksella [18] MFCC-ominaisuuksien avulla ja lupaavia tuloksia raa'alla aaltomuodolla.

**Tulos**

Wav2Letter: päästä päähän - ConvNet-pohjainen puheentunnistusjärjestelmä

**Esimerkki 2.1025**

Puolivalvotun oppimisen menestys riippuu ratkaisevasti sen skaalautuvuudesta valtavaan määrään merkitsemätöntä dataa, jota tarvitaan taustalla olevan moninaisen rakenteen tallentamiseen paremman luokittelun aikaansaamiseksi. Koska pareittaisen samankaltaisuuden laskeminen harjoitusaineiston välillä on kohtuuttoman kallista useimmille syötetiedoille, tällä hetkellä ei ole olemassa mitään yleistä valmista puolivalvotun oppimisen menetelmää/työkalua, jolla voitaisiin oppia kymmenillä miljoonilla tai useammilla datapisteillä. Tässä asiakirjassa otimme käyttöön kahden matalan rankin etikettien etenemisalgoritmin, GLNP:n (Global Linear Neighborhood Propagation) ja Kernel Nyström Approximationin, idean ja toteutimme näiden kahden algoritmin rinnakkaistetun version, jota kiihdytettiin Nesterovin kiihdytetyllä projisoidulla gradienttilaskeutumisella Big-data Label Propagationia (BigLP) varten. Rinnakkaisia algoritmeja testataan viidellä todellisella tietokokonaisuudella, joiden koko vaihtelee 7000:sta 10 000 000:een, ja 100 000 000 näytteen simulaatiotietokokonaisuudella. Kokeissa toteutus voi skaalautua 100 000 000 näytteen ja satojen ominaisuuksien tietokokonaisuuksiin, ja algoritmit paransivat ennustustarkkuutta merkittävästi myös silloin, kun vain hyvin pieni osa datasta on merkitty. Tulokset osoittavat, että BigLP-toteutus on hyvin skaalautuva suurille datamäärille ja tehokas hyödyntämään merkitsemätöntä dataa puolivalvottuun oppimiseen.

**Tulos**

Low-rank Label Propagation puolivalvottuun oppimiseen 100 miljoonalla näytteellä.

**Esimerkki 2.1026**

Tässä artikkelissa ehdotamme moniydinluokittimen oppimisalgoritmia, jolla optimoidaan tietty epälineaarinen ja ei-monikuukautinen monimuuttujaluokittimen suorituskykymittaus. Lisäksi ydinfunktion valintaan ja ytimen parametrien virittämiseen liittyvän ongelman ratkaisemiseksi ehdotimme optimaalisen ytimen rakentamista painotetulla lineaarisella yhdistelmällä joistakin ehdokasytimistä. Luokittelijan parametrin ja ytimen painon oppiminen yhdistetään yhdeksi ainoaksi tavoitefunktioksi, jossa minimoidaan tietyn monimuuttujaisen suorituskykymittarin yläraja. Tavoitefunktiota optimoidaan luokitteluparametrin ja ytimen painon osalta vuorotellen iteratiivisessa algoritmissa leikkaustasoalgoritmia käyttäen. Kehitettyä algoritmia arvioidaan kahdella eri kuvioluokitusmenetelmällä monimuuttujaisen suorituskykymittarin optimointiongelmien osalta. Kokeiden tulokset osoittavat, että ehdotettu algoritmi päihittää kilpailevat menetelmät.

**Tulos**

Usean ytimen monimuuttujaisen suorituskyvyn oppiminen leikkaavan tason algoritmia käyttäen

**Esimerkki 2.1027**

Useat kirjoittajat ovat hiljattain kehittäneet riskiherkkiä politiikkagradienttimenetelmiä<lb>, jotka täydentävät tavanomaista odotettujen kustannusten minimointiongelmaa kustannusten muuttuvuuden mittarilla<lb>. Näissä tutkimuksissa on keskitytty tiettyihin riskimittoihin, kuten<lb>varianssiin tai ehdolliseen riskiarvoon (CVaR). Tässä työssä laajennamme pol-<lb>icy-gradienttimenetelmän koko koherenttien riskimittojen luokkaan, joka on laajalti<lb>hyväksytty muun muassa rahoituksessa ja operaatiotutkimuksessa. Tarkastelemme<lb>sekä staattisia että aikakonsistentteja dynaamisia riskimittoja. Staattisten riskimittojen osalta<lb>lähestymistapamme on politiikkagradienttialgoritmien hengessä ja yhdistää tavanomaisen<lb>näytteenottomenetelmän ja koveran ohjelmoinnin. Dynaamisten riskimittojen osalta lähestymistapamme on toimijakriittinen ja sisältää eksplisiittisen arvofunktion approksimaation.<lb> Tärkeintä on, että panoksessamme esitämme yhtenäisen lähestymistavan riskisensitiiviseen<lb>vahvistusoppimiseen, joka yleistää ja laajentaa aiempia tuloksia.

**Tulos**

Johdonmukaisten riskinarviointitoimenpiteiden poliittinen porrastus

**Esimerkki 2.1028**

Suunnitelmien tunnistamisen tavoitteena on löytää havaittujen toimintojen taustalla olevat tavoitesuunnitelmat (eli toimintasekvenssit), kun käytössä on historiasuunnitelmakirjastoja tai toimialueen malleja. Aiemmat lähestymistavat joko löytävät suunnitelmia "sovittamalla" havaitut toiminnot mahdollisimman hyvin yhteen suunnitelmakirjastojen kanssa olettaen, että kohdesuunnitelmat ovat suunnitelmakirjastoista, tai päättelevät suunnitelmia suorittamalla toimialuemalleja, jotka selittävät parhaiten havaitut toiminnot, olettaen, että täydelliset toimialuemallit ovat käytettävissä. Todellisen maailman sovelluksissa kohdesuunnitelmat eivät kuitenkaan useinkaan ole suunnitelmakirjastoista ja täydellisiä toimialueen malleja ei useinkaan ole saatavilla, koska täydellisten suunnitelmasarjojen ja täydellisten toimialueen mallien rakentaminen on usein vaikeaa tai kallista. Tässä artikkelissa suunnitelmakirjastoja tarkastellaan korporaatioina, ja korporaatioiden avulla opitaan toimintojen vektorimuotoisia esityksiä; sitten löydetään kohdesuunnitelmia vektorimuotoisten esitysten perusteella. Lähestymistapamme kykenee löytämään suunnitelmakirjastojen ulkopuolisia suunnitelmia ilman, että tarvitaan toimitettuja toimialueen malleja. Osoitamme empiirisesti lähestymistapamme tehokkuuden vertaamalla sen suorituskykyä perinteisiin suunnitelmien tunnistamismenetelmiin kolmella suunnittelualueella.

**Tulos**

Toimintojen hajautettuihin esityksiin perustuvien suunnitelmien löytäminen

**Esimerkki 2.1029**

Tutkimme matalien ja syvien neuroverkkojen ilmaisuvoimaa lineaarisilla aktivointifunktioilla. Määritämme uusia tarkkoja ylä- ja alarajoja verkon monimutkaisuudelle Sobolev-avaruuksien approksimaatioissa. Erityisesti todistamme, että syvät ReLU-verkot approksimoivat sileitä funktioita tehokkaammin kuin matalat verkot. 1D Lipschitz-funktioiden approksimaatioiden tapauksessa kuvaamme adaptiivisia syvyys-6-verkkoarkkitehtuureja, jotka ovat tehokkaampia kuin tavalliset matalat arkkitehtuurit.

**Tulos**

Syvillä ReLU-verkoilla tehtävien approksimaatioiden virherajat

**Esimerkki 2.1030**

Todennäköisyyteen perustuvassa logiikassa jopa kohtalaisen suuret ongelmat voivat tuottaa lineaarisia rajoitusjärjestelmiä, joissa on niin monta muuttujaa, että tarkat menetelmät ovat epäkäytännöllisiä. Tämä vaikeus voidaan korjata monissa kiinnostavissa tapauksissa ottamalla käyttöön kolmiarvoinen logiikka (tosi, epätosi ja "ei kiinnosta"). Kolmiarvoisen lähestymistavan avulla voidaan rakentaa "pakattuja" rajoitusjärjestelmiä, joilla on samat ratkaisujoukot kuin kaksiarvoisilla vastaavilla järjestelmillä, mutta joissa voi olla huomattavasti vähemmän muuttujia. Keskustellaan tekniikoista, joilla voidaan laskea piste-estimaatteja lauseiden jälkitodennäköisyyksille. 1. LAUSEKKEIDEN PROLIFEROINTI Nilssonin (1986) probabilistisen logiikan entailment-ongelmassa johdetaan yhden lauseen (jäljempänä "kohde") ennakkotodennäköisyyden estimaatti muiden ("lähde"-) lauseiden joukon ennakkotodennäköisyyksistä. Lähdelauseiden ennakko-uskomukset asettavat rajoituksia, jotka ovat muotoa P=VW L wi = l Wi <': 0 summa kaikista "maailmoista"

**Tulos**

Todennäköisyyslogiikan pakatut rajoitukset ja niiden tarkistaminen

**Esimerkki 2.1031**

Käsitekarttoja voidaan käyttää tärkeiden tietojen tiiviseen esittämiseen ja rakenteiden luomiseen suuriin asiakirjakokoelmiin. Siksi tutkimme monidokumenttien tiivistämisen muunnelmaa, joka tuottaa tiivistelmiä käsitekarttojen muodossa. Tällä hetkellä sopivia arviointitietokantoja tätä tehtävää varten ei kuitenkaan ole. Tämän puutteen korjaamiseksi esittelemme äskettäin luodun käsitekarttojen korpuksen, jossa tiivistetään heterogeenisiä kokoelmia opetusaiheisia verkkodokumentteja. Se on luotu käyttämällä uutta joukkoistamismenetelmää, jonka avulla voidaan tehokkaasti määrittää tärkeät elementit suurista asiakirjakokoelmista. Julkaisemme korpuksen yhdessä perusjärjestelmän ja ehdotetun arviointiprotokollan kanssa, jotta tätä tiivistämisen muunnelmaa voidaan tutkia lisää.

**Tulos**

Pisteiden yhdistäminen: Suurten asiakirjakokoelmien tiivistäminen ja jäsentäminen käsitekarttojen avulla.

**Esimerkki 2.1032**

Laajennamme monikätisen bandiitin mallia, jossa on yksikkökohtaiset vaihtokustannukset, sisällyttämällä siihen toimintojen välinen metriikka. Tarkastelemme tapausta, jossa toimintojen metriikka voidaan mallintaa täydellisellä binääripuulla, ja kahden lehden välinen etäisyys on niiden pienimmän yhteisen esivanhemman alipuun koko, mikä abstrahoi tapauksen, jossa toiminnot ovat pisteitä jatkuvalla välillä r0, 1s ja vaihtokustannus on niiden etäisyys. Tässä ympäristössä esitämme uuden algoritmin, joka määrittää katumuksen r Op ? kT ` T {kq, jossa k on toimien lukumäärä ja T on aikahorisontti. Kun toimintajoukko vastaa kokonaista r0, 1s-väliä, voimme hyödyntää menetelmäämme Lipschitz-häviöfunktioilla tapahtuvan bandit-oppimisen tehtävässä, jolloin algoritmimme saavuttaa optimaalisen katumuksen r ΘpT 2{3q, joka on sama, joka saadaan, kun liikkeistä ei ole rangaistusta. Pääsovelluksessamme käytämme uutta algoritmiamme adaptiivisen hinnoittelun ongelman ratkaisemiseen. Tarkemmin sanottuna tarkastelemme tapausta, jossa yksi myyjä kohtaa kärsivällisiä ostajia. Jokaisella ostajalla on oma arvo ja aikaikkuna, jonka aikana hän on kiinnostunut ostamaan, ja hän ostaa ikkunan alimpaan hintaan, jos se on hänen arvoaan alhaisempi. Osoitamme, että sopivalla hintojen diskretisoinnilla myyjä voi saavuttaa jälkikäteen parhaaseen kiinteään hintaan verrattuna katumuksen r OpT 2{3q, mikä ylittää ongelman aikaisemman katumuksen rajan r OpT 3{4q.

**Tulos**

Liikkumiskustannuksia ja mukautuvaa hinnoittelua sisältävät rosvot

**Esimerkki 2.1033**

Tässä artikkelissa esitämme kaksi aksiomatisaatiota algebralliselle odotushyödylle, joka on erityinen yleistetty odotushyöty, von Neumann-Morgenstern-asetelmassa, eli epävarmuuden representaation oletetaan olevan annettu, ja sitä kuvataan uskottavuusmittarilla, joka on arvotettu semiringissä, joka voi olla osittain järjestetty. Osoitamme, että odotetun hyödyn aksioomien kanssa identtiset aksioomat johtavat siihen, että preferenssejä edustaa algebrallinen odotettu hyöty. Tämä algebrallinen lähestymistapa mahdollistaa monien aiempien lauseiden (odotettu hyöty, binäärinen possibilistinen hyöty, ...) yhdistämisen samaan yleiseen kehykseen ja osoittaa, että saadulla hyötyarvolla on samat hienot ominaisuudet kuin odotetulla hyötyarvolla: lineaarisuus, dynaaminen johdonmukaisuus, taustalla olevan epävarmuusesityksen autodynaamisuus, päätöksentekokriteerin autodynaamisuus ja mahdollisuus mallintaa päätöksentekijän asennetta epävarmuutta kohtaan.

**Tulos**

Aksiomaattiset perusteet eräälle yleistetyn odotetun hyödyn luokalle: Algebrallinen odotettu hyöty

**Esimerkki 2.1034**

s sisältävät näkyvästi tietoja, jotka ovat merkityksellisiä kiinnostuksen kohteena olevan väestöryhmän, toimenpiteen, vertailun ja sairauden kannalta. Vektoriavaruusmalli mittaa lisäksi kyselyn ja lainauksen välistä samankaltaisuutta käsitteiden perusteella (eikä vain termien tai sanojen perusteella).

**Tulos**

Hybridinen viittausten hakualgoritmi näyttöön perustuvan kliinisen tiedon tiivistämiseen: Vektorin samankaltaisuus ja kyselyjen laajentaminen suuren tarkkuuden saavuttamiseksi.

**Esimerkki 2.1035**

Ehdotamme muunnettuun adsorptioon perustuvaa etikettien etenemismenetelmää maantieteellisen sijainnin ennustamiseen, johon on tehty kaksi parannusta: (1) "julkkis"-solmujen poistaminen sijainnin homofilian lisäämiseksi ja helpommin lähestyttävyyden parantamiseksi ja (2) tekstipohjaisten maantieteellisen sijainnin ennakkoarvioiden sisällyttäminen testikäyttäjille. Kolmella Twitterin vertailutietoaineistolla tehdyillä kokeilla saavutetaan huipputuloksia ja osoitetaan parannusten tehokkuus.

**Tulos**

Twitter-käyttäjän maantieteellinen paikannus yhdistetyn teksti- ja verkkoennustemallin avulla

**Esimerkki 2.1036**

Tässä artikkelissa tutkitaan erilaisten verkko-oppimisalgoritmien käyttäytymistä. Uskomusverkkojen valinta, joilla on tietyt minimaalisuusominaisuudet, osoittautuu NP-vaikeaksi, mikä oikeuttaa hakuheuristiikkojen käytön. Cooperin ja Her skovitsin Bayesin mittaukseen perustuvia hakuheuristiikkoja ja MDL-mittaa (Minimum Description Length) verrataan niiden ominaisuuksien osalta sekä rajallisella että äärellisellä tietokantakoolla. Osoitetaan, että MDL-mitalla on enemmän toivottavia ominaisuuksia kuin Bayesin mittauksella. Kokeelliset tulokset viittaavat siihen, että uskomusverkkojen todennäköisyyksien oppimisessa tasoittaminen on hyödyllistä.

**Tulos**

Bayesin uskomusverkko-oppimisalgoritmien ominaisuudet

**Esimerkki 2.1037**

Gaussin prosessin bandit-optimointi on osoittautunut tehokkaaksi työkaluksi meluisten mustan laatikon funktioiden optimointiin. Yksi esimerkki koneoppimisessa on hyperparametrin optimointi, jossa jokainen kohdefunktion arviointi vaatii mallin kouluttamista, mikä voi vaatia päivien tai jopa viikkojen laskentatyömäärän. Useimmat menetelmät tähän niin sanottuun "Bayesin optimointiin" mahdollistavat vain parametriavaruuden peräkkäisen tutkimisen. Usein on kuitenkin toivottavaa ehdottaa parametrien arvojen eriä tai sarjoja, joita voidaan tutkia samanaikaisesti, erityisesti silloin, kun käytettävissä on suuria rinnakkaisia prosessointilaitteita. Erämenetelmät edellyttävät erän eri arviointien välisen vuorovaikutuksen mallintamista, mikä voi olla kallista monimutkaisissa skenaarioissa. Tässä artikkelissa ehdotamme uutta lähestymistapaa Bayesin optimoinnin rinnakkaistamiseen mallintamalla erän monimuotoisuutta Determinanttipisteprosessien (DPP) avulla, joiden ytimet opitaan automaattisesti. Näin voimme yleistää aiempaa tulosta sekä todistaa paremmat DPP-näytteenottoon perustuvat katumusrajat. Kokeemme erilaisissa synteettisissä ja todellisissa robotiikan ja hyperparametrin optimointitehtävissä osoittavat, että DPP-pohjaiset menetelmämme, erityisesti DPP-näytteenottoon perustuvat, päihittävät nykyaikaiset menetelmät.

**Tulos**

Batched Gaussin prosessin bandit-optimointi determinanttipisteprosessien avulla

**Esimerkki 2.1038**

Rekursiivisten neuroverkkojen kouluttaminen on hankalampaa kuin feedforward-verkkojen, koska Bengio et al. (1994) kuvaavat katoavan ja räjähtävän gradientin ongelmia. Tässä artikkelissa yritämme ymmärtää räjähtävän gradientin ongelman taustalla olevia peruskysymyksiä tutkimalla sitä analyyttisestä, geometrisesta ja dynaamisen järjestelmän näkökulmasta. Analyysimme avulla perustellaan yksinkertaista mutta tehokasta ratkaisua, jossa räjähtävä gradientti leikataan normileikkauksella. Kokeellisessa osiossa tämän heuristisen ratkaisun ja tavallisen SGD:n välinen vertailu antaa empiiristä näyttöä hypoteesillemme ja osoittaa, että tällainen heuristiikka on välttämätön, jotta saavutetaan huipputason tulokset hahmojen ennustamistehtävässä ja moniäänisen musiikin ennustamistehtävässä.

**Tulos**

Räjähtävän gradientin ongelman ymmärtäminen

**Esimerkki 2.1039**

Graafin osioinnilla, joka on hyvin tutkittu rinnakkaislaskennan ongelma, on monia sovelluksia monilla eri aloilla, kuten hajautetussa tietojenkäsittelyssä, sosiaalisten verkostojen analysoinnissa, tiedonlouhinnassa ja monilla muilla aloilla. Tässä artikkelissa esitellään FGPGA, tehokas geneettinen lähestymistapa toteutettavissa olevien graafin osioiden tuottamiseen. Menetelmämme ottaa huomioon osioiden heterogeenisuuden ja kapasiteettirajoitukset tasapainoisen osioinnin varmistamiseksi. Tällaisella lähestymistavalla on useita sovelluksia mobiilipilvilaskennassa, johon kuuluu ohjelmistosovellusten toteuttamiskelpoinen käyttöönotto resursseiltaan paremmassa infrastruktuurissa pilvipalvelussa mobiilin käsivarren sijasta. Ehdotettu lähestymistapa on kevyt ja soveltuu siten käytettäväksi pilviarkkitehtuurissa. Varmistamme luotujen osioiden toteutettavuuden estämällä liian suurten osioiden luomisen alustuksen ja haun aikana. Ehdotettu menetelmämme, jota testattiin vakiomuotoisilla vertailutietoaineistoilla, päihittää merkittävästi uusimmat menetelmät osioiden laadun ja ratkaisujen toteutettavuuden osalta.

**Tulos**

FGPGA: Tehokas geneettinen lähestymistapa toteuttamiskelpoisten graafin osioiden tuottamiseen

**Esimerkki 2.1040**

Verkkosovellusten suunnittelun nykyinen monimutkaisuus tekee ohjelmistoturvallisuudesta vaikeasti saavutettavan tavoitteen. Hyökkääjä voi tutustua verkkopalveluun ja hyökätä omaan tahtiinsa. Moving Target Defense (MTD) web-sovelluksissa on tehokas mekanismi, jolla voidaan mitätöidä tämä tiedustelun etu, mutta kehys vaatii hyvän kytkentästrategian, kun sen web-pinon eri kokoonpanojen välillä vaihdetaan. Tämän ongelman ratkaisemiseksi ehdotamme todellisen MTD-verkkosovelluksen mallintamista toistuvana Bayes-pelinä. Sen jälkeen muotoilemme optimointiongelman, joka tuottaa tehokkaan vaihtostrategian ottaen huomioon eri web-pakettikonfiguraatioiden välillä vaihtamisesta aiheutuvat kustannukset. Jotta tämä malli voidaan sisällyttää kehitettyyn MTD-järjestelmään, kehitämme automaattisen järjestelmän, jolla luodaan hyökkäysjoukkoja yhteisistä haavoittuvuuksista ja altistumisista (Common Vulnerabilities and Exposures, CVE) syötetyille hyökkääjätyypeille, joilla on ennalta määritellyt valmiudet. Kehyksemme saa realistiset palkitsemisarvot tämän pelin pelaajille (puolustajille ja hyökkääjille) käyttämällä kansallisesta haavoittuvuustietokannasta (National Vulnerability Database, NVD) saatuja CVE:itä koskevaa turvallisuusalan asiantuntemusta. Käsittelemme myös kysymystä haavoittuvuuksien priorisoinnista, sillä niiden korjaaminen parantaa monenkeskisen tietoturvajärjestelmän turvallisuutta. Lopuksi osoitamme ehdotetun mallin kestävyyden arvioimalla sen suorituskykyä, kun hyökkääjätiedoista ei ole varmuutta.

**Tulos**

Verkkosovellusten puolustaminen liikkuvan kohteen avulla Bayesin Stackelberg-pelien avulla

**Esimerkki 2.1041**

Kun syvien neuroverkkojen (DNN) monimutkaisuus kasvaa ja ne pystyvät ottamaan vastaan yhä suurempia datamääriä, muistin ja energian kulutukseen on kiinnitetty yhä enemmän huomiota teollisissa sovelluksissa, erityisesti mobiililaitteissa. Tässä asiakirjassa esitellään uusi rakenne, joka perustuu funktionaaliseen hashiin DNN:ien pakkaamiseksi, nimittäin FunHashNN. FunHashNN käyttää useita edullisia hash-funktioita jokaisen syvän verkon merkinnän kohdalla hakemaan arvoja pakkausavaruudesta ja käyttää sitten pientä rekonstruktioverkkoa kyseisen merkinnän palauttamiseksi. Rekonstruktioverkko liitetään koko verkkoon ja koulutetaan yhdessä. FunHashNN sisältää hiljattain ehdotetut HashedNets-verkot [7] degeneroituneena tapauksena, ja se hyötyy suuremmasta arvokapasiteetista ja pienemmistä rekonstruktiohäviöistä. Keskustelemme lisäksi laajennuksista, joissa on kaksoisavaruuden hashaus ja monihyppelyt. Useissa vertailutietoaineistoissa FunHashNN osoittaa korkeaa pakkaussuhdetta ilman, että ennustustarkkuus heikkenee juurikaan.

**Tulos**

Toiminnallinen häivytys neuroverkkojen pakkaamiseen

**Esimerkki 2.1042**

On osoitettu, että laajamittaiset realistiset tietoon perustuvan konekääntämisen (KBMT) sovellukset edellyttävät valtavan määrän tietoa kielestä ja maailmasta. Tämä tieto on koodattu laskennallisiin kielioppiin, leksikoihin ja toimialamalleihin. Toinen lähestymistapa, jossa vältetään tarve kerätä ja analysoida valtavaa tietämystä, on esimerkkipohjainen lähestymistapa, joka on tämän asiakirjan aiheena. Osoitamme tässä asiakirjassa, että esimerkkipohjainen lähestymistapa ei sovellu arabian kielelle kääntämiseen. Sen vuoksi esitellään peruslähestymistavan muutos, jolla parannetaan käännösprosessin tarkkuutta. Uuden lähestymistavan perusajatuksena on parantaa tekniikkaa, jolla mallipohjaiset lähestymistavat valitsevat sopivat mallit. Se perustuu siihen, että rinnakkaisesta kaksikielisestä korpuksesta poimitaan kaikki mahdolliset mallit, jotka voisivat vastata lähdelauseen osia. Nämä mallit valitaan lähdelauseeseen soveltuviksi ehdokasosiksi. Vastaavat arabian kielen mallit poimitaan myös ja esitetään suoraviivaisen graafin avulla. Kukin haara edustaa yhtä mahdollista mallisarjaa, joka on ehdokas edustamaan kohdelausetta. Lyhin jatkuva polku tai todennäköisin puun haara valitaan edustamaan kohdelausetta. Lopuksi luodaan valitun puun haaran arabian käännökset.

**Tulos**

Paras mallien vastaavuus tekniikka esimerkkipohjaista konekäännöstä varten

**Esimerkki 2.1043**

Esittelemme uuden päätöksentekosäännön, maximin safety, joka pyrkii säilyttämään suuren marginaalin huonoimpaan lopputulokseen nähden samalla tavalla kuin minimax regret pyrkii minimoimaan etäisyyden parhaaseen lopputulokseen. Väitämme, että maksimiturvallisuus on arvokas sekä kuvauksellisesti että normatiivisesti. Kuvailevasti maksimiturvallisuus selittää tunnetun houkutuslintuvaikutuksen, jossa hallitsevan vaihtoehdon käyttöönotto muuttaa mieltymyksiä muiden vaihtoehtojen välillä. Normatiivisesti esitämme aksiomatisaation, joka luonnehtii maksimiturvallisuuden aiheuttamia preferenssejä, ja osoitamme, että maksimiturvallisuudella on paljon samaa käyttäytymisperustaa minimax-katumuksen kanssa.

**Tulos**

Maksimiturvallisuus: Kun häviämättä jättäminen on parempi vaihtoehto kuin voittaminen.

**Esimerkki 2.1044**

Automaattisessa kuvankuvauksessa on viime vuosina tapahtunut nopeaa ja merkittävää edistystä, mutta mitkä ovat tämän alan avoimet ongelmat? Useimpia töitä on arvioitu tekstipohjaisilla samankaltaisuusmittareilla, jotka osoittavat vain, että parannuksia on tapahtunut, mutta eivät selitä, mikä on parantunut. Tässä artikkelissa esitellään yksityiskohtainen virheanalyysi kuvauksista, jotka on tuotettu huipputason huomiopohjaisella mallilla. Analyysimme toimii kahdella tasolla: ensin tarkistamme kuvausten tarkkuuden ja sitten luokittelemme epätarkoissa kuvauksissa havaitsemamme virhetyypit. Huomaamme, että vain 20 prosentissa kuvauksista ei ole virheitä, ja yllättäen 26 prosentissa kuvauksista ei ole mitään tekemistä kuvan kanssa. Lopuksi korjaamme manuaalisesti yleisimmin esiintyvät virhetyypit (esim. sukupuolen tunnistaminen), jotta voimme arvioida, kuinka suuri suorituskyky on palkittu näiden virheiden korjaamisesta, ja havaitsemme 0,2-1 BLEU-pisteen parannuksen kutakin virhetyyppiä kohti.

**Tulos**

Automaattisen kuvankuvauksen parantamisen varaa: virheanalyysi

**Esimerkki 2.1045**

Puhutun sisällön hakeminen puhuttujen kyselyjen avulla eli puhuttujen termien havaitseminen (STD) on houkuttelevaa, koska se mahdollistaa signaalien yhteensovittamisen suoraan akustisella tasolla ilman, että niitä tarvitsee kirjoittaa tekstiksi. Tässä ehdotamme päästä päähän -kyselyä esimerkin mukaan - STD-mallia, joka perustuu tarkkaavaisuuteen perustuvaan monihyppelyverkkoon, jonka syötteenä on puhuttu kysely ja useita lauseita sisältävä äänikatkelma; ulostulossa ilmoitetaan, sisältääkö äänikatkelma kyselyn. Malli voidaan kouluttaa joko valvotussa skenaariossa käyttäen leimattua dataa tai valvomatta. Valvotussa skenaariossa havaitsemme, että huomiomekanismi ja useat siirtymät parantavat suorituskykyä ja että huomiopainot ilmaisevat havaittujen termien aikaväliä. Valvomattomassa tilanteessa malli jäljittelee olemassa olevan esimerkkikyselyjen STD-järjestelmän käyttäytymistä, ja sen suorituskyky on verrattavissa olemassa olevaan järjestelmään, mutta hakuaika on lyhyempi.

**Tulos**

ESIMERKKIKYSELYYN PERUSTUVAN PUHUTUN TERMIN HAVAITSEMINEN KÄYTTÄMÄLLÄ HUOMIOPERUSTEISIA MONISATAMAVERKKOJA

**Esimerkki 2.1046**

Monien vähävaraisten tai uhanalaisten kielten puhuttuja kieliaineistoja kommentoidaan todennäköisemmin käännöksillä kuin transkriptioilla. Viimeaikaisessa työssä on hyödynnetty tällaisia merkintöjä puheen ja käännöksen välisen rinnastuksen tuottamiseksi ilman tekstin transkriptioita. Tutkimme, voiko tällaisten tietojen antaminen auttaa tuottamaan parempia (väärin sovitettuja) joukkoresurssien transkriptioita, jotka puolestaan voisivat olla arvokkaita puheentunnistusjärjestelmien harjoittelussa, ja osoitamme, että niistä voi todellakin olla hyötyä pienen mittakaavan tapaustutkimuksen avulla. Esittelemme myös yksinkertaisen, foneettisesti tiedostavan merkkijonojen keskiarvotekniikan, joka tuottaa laadukkaampia transkriptioita.

**Tulos**

Tapaustutkimus puheesta käännökseksi -tasauksen käytöstä kielidokumentoinnissa.

**Esimerkki 2.1047**

Ensembling on tunnettu tekniikka neuraalisessa konekääntämisessä (NMT) järjestelmän suorituskyvyn parantamiseksi. Yhden neuroverkon sijasta useita neuroverkkoja, joilla on sama topologia, koulutetaan erikseen, ja dekooderi tuottaa ennusteita keskiarvoistamalla yksittäiset mallit. Ensemblaaminen parantaa usein huomattavasti tuotettujen käännösten laatua. Se ei kuitenkaan sovellu tuotantojärjestelmiin, koska se on työläs ja hidas. Tässä työssä pyritään lyhentämään suoritusaikaa niin, että se vastaa yhden järjestelmän suoritusaikaa käännösten laadusta tinkimättä. Ensiksi osoitamme, että kokonaisuus voidaan purkaa yhdeksi suureksi neuroverkoksi, joka jäljittelee kokonaisuusjärjestelmän tuotosta. Osoitamme, että purkaminen voi jo käytännössä parantaa suoritusaikaa, koska GPU:lla voidaan tehdä enemmän työtä. Jatkamme kuvaamalla joukon tekniikoita, joilla voidaan pienentää taitettua verkkoa vähentämällä kerrosten dimensiota. Raportoimme, että japaninenglannilla saatu verkko on kooltaan ja dekoodausnopeudeltaan yhden NMT-verkon kokoinen, mutta sen suorituskyky on 3-kokoonpanon järjestelmän tasolla.

**Tulos**

Neuraalisten konekääntäjäkokoonpanojen purkaminen ja kutistaminen

**Esimerkki 2.1048**

Viimeaikaiset tutkimukset osoittavat, että useimmilla brasilialaisilla oppilailla on vakavia ongelmia lukutaidossaan. Tämän taidon täysipainoinen kehittäminen on avainasemassa jokaisen kansalaisen akateemisen ja ammatillisen tulevaisuuden kannalta. Lapsille suunnattujen lukumateriaalien monimutkaisuuden luokitteluvälineillä pyritään parantamaan lukemisen ja tekstin ymmärtämisen opetuksen laatua. Fengin työtä [11] pidetään englannin kielen osalta luokkatasojen ennustamisen huipputasona, ja siinä saavutettiin 74 prosentin tarkkuus luokiteltaessa automaattisesti neljä tekstin monimutkaisuuden tasoa lähikouluasteille. Portugalin kielessä ei ole luokittelijoita tietokirjallisille teksteille lähiluokille. Tässä artikkelissa ehdotamme järjestelmää tekstien manuaalista merkitsemistä varten viidellä luokkatasolla, jota käytetään räätälöityyn lukemiseen, jotta vältetään lukemisessa edistyneempien oppilaiden kiinnostuksen puute ja niiden oppilaiden lukemisen estäminen, joiden on vielä edettävä. Saavutimme 52 prosentin tarkkuuden luokitellessamme tekstejä 5 tasolle ja 74 prosentin tarkkuuden 3 tasolle. Tulokset osoittautuvat lupaaviksi, kun niitä verrataan uusimpiin töihin.

**Tulos**

Portugalinkielisten tietokirjallisten tekstien monimutkaisuuden automaattinen luokittelu varhaiskouluikäisille portugalin kielellä

**Esimerkki 2.1049**

Oletuksena on, että luovuus syntyy sisäisen maailmanmallin eli maailmankuvan itsekorjautumiskyvystä. Luovan yksilön ainutlaatuisesti hiottu maailmankatsomus johtaa omaleimaiseen tyyliin, joka on tunnistettavissa sekä eri aloilla että niiden välillä. Lisäksi oletetaan, että luovuus on yleisluonteista siinä mielessä, että on olemassa useita väyliä, joiden kautta yksilön maailmankatsomuksen omaleimaisuus voidaan ilmaista. Näitä hypoteeseja testattiin taideopiskelijoiden ja luovan kirjoittamisen opiskelijoiden avulla. Taideopiskelijat arvasivat huomattavasti sattumaa paremmin sekä sen, minkä maalauksen oli tehnyt mikä viidestä kuuluisasta taiteilijasta, että sen, minkä taideteoksen oli tehnyt mikä heidän ikätoverinsa. Vastaavasti luovan kirjoittamisen opiskelijat arvasivat huomattavasti sattumaa paremmin sekä sen, minkä viidestä kuuluisasta kirjailijasta kukin oli kirjoittanut, että sen, minkä heidän ikätovereistaan kukin oli kirjoittanut. Nämä tulokset tukevat hypoteesia, jonka mukaan luova tyyli on tunnistettavissa. Lisäksi luovan kirjoittamisen opiskelijat arvasivat huomattavasti sattumaa paremmin, kuka heidän ikätovereistaan oli tuottanut tietyn taideteoksen, mikä tukee hypoteesia, jonka mukaan luova tyyli on tunnistettavissa paitsi eri alojen sisällä myös eri aloilla.

**Tulos**

Yksilöllisen luovan tyylin tunnistettavuus eri osa-alueilla ja niiden välillä: Alustavat tutkimukset

**Esimerkki 2.1050**

GLEU-mittaria ehdotettiin kielioppivirheiden korjausten arvioimiseksi käyttämällä n-grammien päällekkäisyyttä vertailulauseiden joukon kanssa, toisin kuin tiettyjen annotoitujen virheiden tarkkuutta/palautumista (Napoles et al., 2015). Tässä artikkelissa kuvataan GLEU-mittariin tehtyjä parannuksia, joilla puututaan ongelmiin, joita syntyy, kun käytetään yhä useampia viitejoukkoja. Toisin kuin alun perin esitelty metriikka, muutettu metriikka ei vaadi virittämistä. Suosittelemme, että tätä versiota käytetään alkuperäisen version sijasta.1

**Tulos**

GLEU ilman viritystä

**Esimerkki 2.1051**

Esitämme ensimmäisen askeleen kohti kehystä, jonka avulla voidaan määritellä ja käsitellä normatiivisia asiakirjoja tai sopimuksia, jotka on kuvattu sopimuskeskeisinä (ContractOriented (C-O) Diagrams) kaavioina. Nämä kaaviot tarjoavat tällaisille teksteille visuaalisen esityksen, joka antaa mahdollisuuden ilmaista allekirjoittajan velvoitteet, luvat ja kiellot, joko ajoitusrajoituksin tai ilman niitä, sekä seuraamukset, jotka johtuvat sopimuksen täyttämättä jättämisestä. Tässä työssä esitellään C-O-kaavioiden sanallistamiseen tarkoitettu CNL, verkkopohjainen työkalu, joka mahdollistaa muokkaamisen kyseisessä CNL:ssä, sekä toinen työkalu, jolla kaavioita voidaan visualisoida ja käsitellä vuorovaikutteisesti. Tämän jälkeen osoitetaan, miten näitä proof-ofconcept-työkaluja voidaan käyttää soveltamalla niitä pieneen esimerkkiin.

**Tulos**

CNL sopimuskeskeisille kaavioille

**Esimerkki 2.1052**

Gaussin seosmallin sovittamiseen käytettävän EM-algoritmin (Expectation Maximization) konvergenssin nopeuden tiedetään olevan riippuvainen seoskomponenttien päällekkäisyyden määrästä. Tässä artikkelissa tutkimme sekoituskertoimien vaikutusta EM:n konvergenssiin. Osoitamme, että kun sekoituskomponentit ovat jonkin verran päällekkäisiä, EM:n konvergenssi hidastuu, kun sekoituskertoimien dynaaminen vaihteluväli kasvaa. Ehdotamme determinististä antiannelointialgoritmia, joka parantaa merkittävästi EM:n konvergenssin nopeutta tällaisille seoksille, joissa sekoituskertoimet eivät ole tasapainossa. Ehdotettua algoritmia verrataan muihin tavallisiin optimointitekniikoihin, kuten BFGS, Conjugate Gradient ja perinteinen EM-algoritmi. Lopuksi ehdotamme samanlaista determinististä antiannelointiin perustuvaa algoritmia Dirichlet-prosessin seosmallille ja osoitamme sen edut perinteiseen variationaaliseen Bayesin lähestymistapaan verrattuna.

**Tulos**

EM-algoritmin konvergenssi Gaussin sekoituksille, joiden sekoituskertoimet ovat epätasapainossa.

**Esimerkki 2.1053**

Esittelemme "eksponentiaalisen lineaarisen yksikön" (ELU), joka nopeuttaa oppimista<lb>syvissä neuroverkoissa ja johtaa suurempaan luokittelutarkkuuteen. Kuten rektifioidut<lb>lineaariset yksiköt (ReLU), vuotavat ReLU:t (LReLU) ja parametrisoidut ReLU:t (PRe-<lb>LU), myös ELU:t välttävät katoavan gradientin identiteetin kautta positiivisille arvoille.<lb>Ja ELU:illa on kuitenkin paremmat oppimisominaisuudet verrattuna yksiköihin, joilla on<lb>muuta aktivointifunktiota. Toisin kuin ReLU:illa, ELU:illa on negatiivisia arvoja, mikä<lb>antaa niille mahdollisuuden työntää yksikön keskimääräiset aktivaatiot lähemmäs nollaa. Nollakeskiarvot nopeuttavat<lb>oppimista, koska ne tuovat gradientin lähemmäs yksikön luonnollista gradienttia. Me<lb>näytämme, että yksikön luonnollinen gradientti eroaa normaalista gradientista bias shift<lb>termillä, joka on verrannollinen saapuvien yksiköiden keskimääräiseen aktivaatioon. Kuten erä<lb>normalisointi, ELU:tkin työntävät keskiarvoa kohti nollaa, mutta huomattavasti pienemmällä<lb>laskennallisella jalanjäljellä. Vaikka muilla aktivointifunktioilla, kuten LReLU:lla ja PRe-<lb>LU:lla, on myös negatiivisia arvoja, ne eivät takaa kohinaa kestävää deaktivointitilaa.<lb>ELU:t kyllästyvät negatiiviseen arvoon pienemmillä syötteillä ja vähentävät siten<lb>propagoituvaa vaihtelua ja informaatiota. Siksi ELU:t koodaavat tiettyjen ilmiöiden esiintymisastetta<lb>syötteessä, mutta ne eivät mallinna<lb>kvantitatiivisesti niiden puuttumisen astetta. Näin ollen ELU:iden väliset riippuvuudet ovat paljon<lb>helpommin mallinnettavissa ja erilliset käsitteet häiritsevät toisiaan vähemmän.<lb>Havaitsimme, että ELU:t johtavat paitsi nopeampaan oppimiseen myös parempaan yleistymis<lb>suorituskykyyn, kun verkoissa on monta kerrosta (≥ 5). ELU-verkot olivat<lb>kymmenen parhaan raportoidun CIFAR-10-tuloksen joukossa ja tuottivat parhaan julkaistun tuloksen<lb>CIFAR-100:ssa ilman, että turvauduttiin moninäkökulmaiseen arviointiin tai mallien keskiarvottamiseen. Kuvaverkossa<lb> ELU-verkot nopeuttavat oppimista huomattavasti verrattuna ReLU<lb>verkkoon, jolla on sama arkkitehtuuri, ja niiden luokitteluvirhe<lb> on alle 10 % yhden sadon ja yhden mallin verkolla.

**Tulos**

EKSPONENTIAALISET LINEAARISET YKSIKÖT (ELUS)

**Esimerkki 2.1054**

Esitetään menetelmä rytmillistä pars<lb>ing-ongelmaa varten: Kun on olemassa havaittujen<lb>musiikillisten nuottien alkamisajankohtien sarja, arvioimme samanaikaisesti<lb> vastaavaa notoitua rytmi<lb>ja tempoprosessia. On<lb>kehitetty graafinen malli, joka<lb>esittää<lb>tempon ja rytmin kehityksen ja suhteuttaa nämä piil<lb>den suureet havaittavaan esitykseen.<lb>Rytmimuuttujat ovat diskreettejä ja<lb>tempo- ja havainnointimuuttujat ovat jatkuvia<lb>jatkuvia. Näytämme, miten lasketaan<lb>glob<lb>todennäköisin konfiguraatio tempo<lb>ja rytmimuuttujille, kun otetaan huomioon havainto<lb>nuottien alkamisajoista. Alustavia kokeita<lb>esitetään pienellä aineistolla. Hahmotellaan yleiskuvaus MAP-estimaattien laskemiseksi<lb>kertaluonteisille ehdollisille Gaussin jakaumille<lb>.

**Tulos**

Sekamuotoinen graafinen malli rytmistä jäsentelyä varten

**Esimerkki 2.1055**

Paikallisten optimaalisten ratkaisujen ominaisuudet monitavoitteisissa kombinatorisissa optimointiongelmissa ovat ratkaisevan tärkeitä paikallisten hakualgoritmien tehokkuuden kannalta, erityisesti silloin, kun nämä algoritmit perustuvat Pareto-dominanssiin. Tällaiset paikalliset hakualgoritmit palauttavat tyypillisesti joukon toisistaan riippumattomia Pareto-lokaalioptimaalisia (PLO) ratkaisuja eli PLO-joukon. Tässä artikkelissa tutkitaan kahta PLO-joukkojen näkökohtaa Pareto-paikallishakua (PLS) koskevien kokeiden avulla. Ensinnäkin tarkastellaan useiden ongelman ominaisuuksien vaikutusta PLO-joukkojen ominaisuuksiin monitavoitteisissa NK-maastoissa, joissa on korreloituneita tavoitteita. Raportoimme erityisesti, että joko tavoitteiden lukumäärän lisääminen tai tavoitteiden välisen korrelaation vähentäminen johtaa PLO-joukkojen koon eksponentiaaliseen kasvuun, kun taas muuttuvalla korrelaatiolla on vain vähäinen vaikutus. Toiseksi tutkimme suoritusaikaa ja saavutettua laatua, kun käytetään rajoittavia arkistointimenetelmiä, joilla rajoitetaan PLS:n käsittelemän arkiston kokoa ja siten löydetyn PLO-joukon enimmäiskokoa. Väitämme, että PLS:n suoritusajan ja ongelmakappaleen vaikeuden välillä on selvä yhteys.

**Tulos**

Paikalliset optimaaliset joukot ja rajattu arkistointi monitavoitteisissa NK-maastoissa, joissa on korreloituneita tavoitteita.

**Esimerkki 2.1056**

Aihepiirimallit ovat hyödyllinen menetelmä suurten tekstikokonaisuuksien dimensioiden pienentämiseen ja eksploratiiviseen data<lb>analyysiin. Useimmat lähestymistavat aihepiirimallien päättelyyn ovat perustuneet<lb>maksimimallin todennäköisyystavoitteeseen. On olemassa tehokkaita algoritmeja, jotka lähestyvät tätä tavoitetta,<lb>mutta niillä ei ole todistettavia takeita. Viime aikoina on esitelty algoritmeja, jotka tarjoavat<lb>todistettavia rajoja, mutta nämä algoritmit eivät ole käytännöllisiä, koska ne ovat tehottomia eivätkä ro-<lb>bust mallin oletusten rikkomiselle. Tässä artikkelissa esitämme algoritmin aiheen malli<lb>inferenssille, joka on sekä todistettavissa että käytännöllinen. Algoritmi tuottaa tuloksia, jotka ovat verrattavissa<lb>parhaisiin MCMC-toteutuksiin, mutta toimii samalla kertaluokkia nopeammin.

**Tulos**

Käytännöllinen algoritmi aihepiirien mallintamiseen todistettavilla takuilla

**Esimerkki 2.1057**

Vaikka sosiaalisen median alustoja käytetäänkin yhä enemmän tiedon ja uutisten keräämiseen, sen moderoimaton luonne johtaa usein huhujen eli sellaisten tietojen syntymiseen ja leviämiseen, joita ei ole varmistettu julkaisuhetkellä. Samalla sosiaalisen median alustojen avoimuus tarjoaa mahdollisuuksia tutkia, miten käyttäjät jakavat huhuja ja keskustelevat niistä, ja tutkia, miten luonnollisen kielen käsittely- ja tiedonlouhintatekniikoita voidaan käyttää huhujen todenperäisyyden määrittämiseen. Tässä tutkimuksessa esittelemme ja käsittelemme kahdenlaisia huhuja, joita sosiaalisessa mediassa liikkuu: pitkäaikaisia huhuja, jotka kiertävät pitkiä aikoja, ja vasta syntyneitä huhuja, jotka syntyvät nopeatempoisten tapahtumien, kuten uutislähetysten, aikana, jolloin raportteja julkaistaan pätkittäin ja usein niiden alkuvaiheessa vahvistamattomina. Esitämme yleiskatsauksen sosiaalisen median huhuja koskevaan tutkimukseen, jonka perimmäisenä tavoitteena on kehittää huhujen luokittelujärjestelmä, joka koostuu neljästä osatekijästä: huhujen havaitseminen, huhujen seuranta, huhujen kannan luokittelu ja huhujen todenperäisyyden luokittelu. Perehdymme tieteellisessä kirjallisuudessa esiteltyihin lähestymistapoihin, joita on käytetty näiden neljän komponentin kehittämisessä. Teemme yhteenvedon tähänastisista ponnisteluista ja saavutuksista huhujen luokittelujärjestelmien kehittämiseksi ja esitämme lopuksi ehdotuksia tulevaisuuden tutkimuskohteiksi sosiaalisen median louhinnan alalla huhujen havaitsemiseksi ja selvittämiseksi.

**Tulos**

Huhujen havaitseminen ja ratkaiseminen sosiaalisessa mediassa: Tutkimus

**Esimerkki 2.1058**

Nykyään tiedonlouhintatekniikoita hyödynnetään lääketieteessä sairauksien diagnosoimiseksi, voittamiseksi ja hoitamiseksi. Neuroverkko on yksi niistä tekniikoista, joita käytetään laajalti diagnosointiin lääketieteen alalla. Tässä artikkelissa arvioidaan yhdeksän algoritmin tehokkuutta sydän- ja verisuonitautien diagnosoinnissa, jotka ovat neuroverkko-oppimisen perustana. Algoritmeja arvioidaan tarkkuuden, herkkyyden, läpinäkyvyyden, AROC:n ja konvergenssinopeuden suhteen 10-kertaisen ristiinvalidoinnin avulla. Tulokset osoittavat, että harjoitusvaiheessa Lonberg-M-algoritmilla on paras tehokkuus kaikkien mittareiden osalta, algoritmilla OSS on suurin tarkkuus testausvaiheessa, algoritmilla SCG on suurin läpinäkyvyys ja algoritmilla CGB on suurin herkkyys. Avainsanat - sydän- ja verisuonisairaudet; neuroverkko; oppimisalgoritmit.

**Tulos**

Sydän- ja verisuonitautien diagnosoinnissa käytettävien neuroverkon oppimisalgoritmien vertailu

**Esimerkki 2.1059**

Multimedian päättely, joka soveltuu muun muassa multimediasisällön analysointiin ja videokohtausten korkeatasoiseen tulkintaan, perustuu edustetun tietämysalueen muodolliseen ja kattavaan käsitteellistämiseen. Useimmat multimediaontologiat eivät kuitenkaan ole tyhjentäviä roolimäärittelyjen osalta eivätkä sisällä monimutkaisia roolien sisällyttämisiä ja roolien keskinäisiä riippuvuuksia. Useimmissa multimediaontologioissa ei itse asiassa ole lainkaan roolilaatikkoa, ja niissä toteutetaan vain perusjoukko käytettävissä olevista loogisista konstruktoreista. Näin ollen niiden soveltaminen multimedia-ajatteluun on rajallista. Edellä mainittujen ongelmien ratkaisemiseksi on esitelty VidOnt, ensimmäinen multimediaontologia, jossa on SROIQ(D)-ilmaisukyky ja DL-turvallinen sääntökokonaisuus seuraavan sukupolven multimediajärkeilyä varten. Yleisestä käytännöstä poiketen muodollinen perusta on asetettu yhteen kaikkein ilmaisuvoimaisimmista kuvauslogiikoista, ja ontologia on validoitu alan johtavilla päättelijöillä, nimittäin HermiT:llä ja FaCT++:lla. Tässä asiakirjassa esitellään myös parhaita käytäntöjä multimediaontologioiden kehittämiseksi, jotka perustuvat ontologiatekniikan lähestymistapaani.

**Tulos**

Uusi lähestymistapa multimedia-ontologian kehittämiseen audiovisuaalisten LOD-tietoaineistojen automaattista päättelyä varten.

**Esimerkki 2.1060**

Esittelemme uudenlaisen diffuusiojärjestelmän verkossa tapahtuvaa kernel-pohjaista verkko-oppimista varten. Toistaiseksi minkä tahansa verkossa toimivan oppimisalgoritmin, joka toimii toistuvassa kernel-Hilbert-avaruudessa (RKHS), suurin haittapuoli on tarve päivittää kasvava määrä parametreja iteraatioiden edetessä. Monimutkaisuuden lisäksi tämä johtaa hajautetussa ympäristössä lisääntyneeseen viestintäresurssien tarpeeseen. Sen sijaan ehdotetussa menetelmässä ratkaisu approksimoidaan kiinteän kokoisena vektorina (joka on suurempi kuin syöttöavaruus) satunnaisia Fourier-ominaisuuksia käyttäen. Tämä mahdollistaa tavanomaisten lineaaristen yhdistely- ja mukauttamistekniikoiden käytön. Tietojemme mukaan tämä on ensimmäinen kerta, kun esitetään täydellinen protokolla RKHS:n hajautettua online-oppimista varten. Lisäksi esitetään edellytykset asymptoottiselle konvergenssille ja verkkokatumuksen rajoittamiselle. Simuloidut testit havainnollistavat ehdotetun järjestelmän suorituskykyä.

**Tulos**

Verkossa tapahtuva hajautettu verkko-oppiminen RKH-avaruudessa satunnaisia Fourier-ominaisuuksia käyttäen

**Esimerkki 2.1061**

Ehdotamme uutta aktiivisen oppimisen (AL) menetelmää tekstiluokittelua varten, joka perustuu konvoluutiohermoverkkoihin (CNN). AL-menetelmässä valitaan manuaalisesti leimattavat tapaukset, jotta mallin suorituskyky voitaisiin maksimoida mahdollisimman pienellä vaivalla. Neuraaliset mallit hyödyntävät sanojen upotuksia ominaisuuksina ja virittävät ne kulloisenkin tehtävän mukaan. Väitämme, että tekstiluokittelussa käytettävien neuraalisten AL-strategioiden olisi keskityttävä valitsemaan instansseja, jotka vaikuttavat eniten upotusavaruuteen (eli saavat aikaan erottelevia sanaesityksiä). Tämä on vastakohta perinteisille AL-lähestymistavoille (esim. epävarmuusnäytteenotto), joissa määritellään korkeamman tason tavoitteet. Ehdotamme yksinkertaista lähestymistapaa, jossa valitaan tapaukset, jotka sisältävät sanoja, joiden upotukset todennäköisesti päivittyvät suurimmalla todennäköisyydellä, ja siten opitaan nopeasti erottelevia, tehtäväkohtaisia upotuksia. Empiiriset tulokset osoittavat, että menetelmämme on parempi kuin AL-peruslähestymistavat.

**Tulos**

Aktiivinen erotteleva sanojen upottamisen oppiminen

**Esimerkki 2.1062**

Esittelemme Grid Beam Search (GBS) -algoritmin, joka laajentaa palkkihakua siten, että siihen voidaan sisällyttää ennalta määritettyjä leksikaalisia rajoituksia. Algoritmia voidaan käyttää minkä tahansa mallin kanssa, joka tuottaa sekvenssin ŷ = {y0 . . . . . yT } maksimoimalla p(y|x) = ∏ t p(yt|x; {y0 . . . . . yt-1}). Leksikaaliset rajoitteet ovat lauseita tai sanoja, joiden on oltava läsnä tulostussekvenssissä. Tämä on hyvin yleinen tapa sisällyttää lisätietoa mallin tulosteeseen ilman, että malliparametreja tai harjoitusdataa tarvitsee muuttaa. Osoitamme leksikaalisesti rajoitetun dekoodauksen toteuttamiskelpoisuuden ja joustavuuden tekemällä kokeiluja neuraalisella interaktiivisella ennakoivalla käännöksellä sekä neuraalisen konekäännöksen toimialueen mukauttamisella (Domain Adaptation for Neural Machine Translation). Kokeet osoittavat, että GBS:n avulla voidaan parantaa merkittävästi käännöksen laatua vuorovaikutteisissa skenaarioissa ja että GBS:n avulla voidaan saavuttaa huomattavia suorituskykyparannuksia toimialan mukauttamisskenaarioissa jopa ilman käyttäjän panosta.

**Tulos**

Lexically Constrained Decoding for Sequence Generation Using Grid Beam Search (Leksikaalisesti rajoitettu dekoodaus sekvenssin tuottamiseksi käyttäen ruudukkopalkkihakua)

**Esimerkki 2.1063**

Ominaisuuksien valinnalla on tärkeä rooli tiedonlouhintaprosessissa. Sitä tarvitaan, jotta voidaan käsitellä liian suurta määrää ominaisuuksia, jotka voivat muodostua laskennalliseksi taakaksi oppimisalgoritmeille. Se on tarpeen myös silloin, kun laskennallisia resursseja ei ole niukasti, sillä se parantaa koneoppimistehtävien tarkkuutta, kuten tulevissa jaksoissa nähdään. Tässä katsauksessa käsitellään erilaisia ominaisuuksien valintatapoja sekä niiden ja eri koneoppimisalgoritmien välistä suhdetta.

**Tulos**

Tutkimus ominaisuuksien valinnasta

**Esimerkki 2.1064**

Vankat uskomusten tarkistusmenetelmät ovat ratkaisevan tärkeitä virtaavan datan tilanteissa, kun olemassa olevaa tietoa (tai uskomuksia) päivitetään uusien todisteiden perusteella. Bayesin ehdollistaminen on ensisijainen mekanismi, jota käytetään uskomusten tarkistamiseen todennäköisyyspohjaista päättelyä käyttävissä datafuusiojärjestelmissä. Perinteiset ehdollistamismenetelmät kohtaavat kuitenkin useita haasteita, jotka johtuvat datan/lähteiden puutteellisuudesta big data -ympäristöissä, jotka hyödyntävät pehmeiden (eli ihmisten tai ihmisiin perustuvien) lähteiden lisäksi myös kovia (eli fysiikkaan perustuvia) antureita. Tämän artikkelin tavoitteena on tutkia Bayesin ehdollistamisen luonnollisinta laajennusta, joka soveltuu todisteiden päivittämiseen tällaisten epävarmuuksien vallitessa. Tarkastelemalla todisteiden päivitysprosessia ajatuskokeiluna johdetaan elegantti strategia vankkaa todisteiden päivittämistä varten äärimmäisten epävarmuustekijöiden vallitessa, jotka ovat tyypillisiä big data -ympäristöille. Erityisesti Fagin-Halpernin ehdollisten käsitteiden avulla johdetaan luonnollinen laajennus Bayesin ehdollistamiseen todisteille, jotka ovat yleisen uskomusfunktion muodossa. Esitetty työ eroaa olennaisesti ehdollisesta päivitysyhtälöstä (Conditional Update Equation, CUE) ja kirjoittajien omista laajennuksista siihen. Kehityksestä annetaan yleiskatsaus havainnollistavien esimerkkien avulla. Lisäksi esitetään näkemyksiä parametrien valinnasta erilaisissa fuusiokonteksteissa.

**Tulos**

Todisteiden päivittäminen virtatiedon käsittelyä varten suurissa tiedoissa: Pehmeän ja kovan datan fuusioympäristöissä.

**Esimerkki 2.1065**

Kuvien rekisteröinti tarkoittaa, että kuvat, joilla on vaihteleva suuntaus, monimodaaliset tai moniaikaiset kuvat kartoitetaan yhteen koordinaattijärjestelmään. Digitaaliset korkeusmallit (Digital Elevation models, DEM) ovat kuvia, joihin on upotettu maastotietoa. DEM:n ja DEM:n välinen rekisteröinti sisältää sellaisten DEM:ien rekisteröinnin, joilla on erilainen suunta, jotka on saatettu kartoittaa eri aikoina tai jotka on käsitelty eri resoluutioilla. Vaikka DEM-rekisteröinti on hyvin tärkeää, on olemassa vain kourallinen menetelmiä, joista useimmat koskevat DEM:n ja topografisen kartan tai DEM:n ja kaukokartoitetun kuvan välistä rekisteröintiä. Kognitiivisten kartoituskäsitteiden käyttäminen DEM-rekisteröinnissä on kehittynyt tästä perusajatuksesta, jonka mukaan avaruuden ja kohteiden välistä kartoitusta ja niiden välisten suhteiden määrittelyä käytetään perustavanlaatuisten kiintopisteiden muodostamiseen, jotka on merkittävä, tallennettava ja joita on käsiteltävä ympäristössä tai muussa ympäristökandidaattiympäristössä tai sen ympäristössä, eli meidän tapauksessamme DEM:issä. Paikkatietoisen kognition menetelmien asteittainen kaksitasoinen kapselointi sisältää maamerkkitietoa ja asettelutietoa, ja siitä voi olla hyötyä DEM-rekisteröinnissä. Avaruuspohjainen lähestymistapa, jossa korostetaan tarkasteltavana olevan ympäristön eksplisiittistä laajuutta, ja kohdepohjainen lähestymistapa, jossa korostetaan paikallisen ympäristön kohteiden välisiä suhteita, jotka ovat kaksi kognitiivisen kartoituksen paradigmaa, voidaan integroida metodisesti tähän DEM-rekisteröinnin kolmiportaiseen arkkitehtuuriin. Aluksi suoritetaan P-malliin perustuva segmentointi, jota seuraa maamerkkien muodostaminen kontekstuaalista kartoitusta varten, jossa käytetään kontekstuaalisen pyramidin muodostamista. Sen lisäksi, että maamerkkejä käytetään rekisteröinnin avainpisteiden löytämiseen, muunnosten ja muutosten havaitsemiseen on käytetty euklidiseen etäisyyteen perustuvaa muodonmuutoslaskentaa. Aluksi suoritetaan P-malliin perustuva segmentointi ja sen jälkeen maamerkkien muodostaminen kontekstuaalista kartoitusta varten, jossa käytetään kontekstuaalista pyramidin muodostamista. Maamerkit on luokiteltu joko tasaisiksi alueiksi, joilla ei ole suuria vaihteluita maanpinnan korkeuksissa; huippuihin, jotka löytyvät, kun korkeus kasvaa asteittain verrattuna tasaisiin alueisiin; laaksoihin, joissa korkeus vähenee asteittain DEM:ssä; ja lopuksi aaltoileviin alueisiin, joilla on hyvin matalia harjuja ja pohjakorkeuksia. Fraktaalipohjaisen pakkauksen on todettu antavan hyviä tuloksia tila- ja laskentavaatimusten kannalta, kun kyseessä on rinnakkaisrekisteröityjen DEM-kuvien lopullinen tallennus. Tässä asiakirjassa on pyritty toteuttamaan DEM-DEM-rekisteröinti, joka perustuu ihmisen paikkatietomenetelmään, joka perustuu muistamiseen. Menetelmää voidaan laajentaa DEM-topografisen kartan ja DEM-kaukokartoitettujen kuvien rekisteröintiin. Kokeelliset tulokset osoittavat, että DEM-rekisteröinti voidaan toteuttaa tehokkaasti ehdotetulla menetelmällä.

**Tulos**

Kognitiiviseen kartoitukseen ja kontekstuaaliseen pyramidiin perustuva digitaalisen korkeusmallin rekisteröinti ja sen tehokas tallentaminen fraktaalipohjaista pakkausta käyttäen.

**Esimerkki 2.1066**

Tässä työssä käytetään L-järjestelmää tekstisekvenssin puurakenteen muodostamiseen ja sen monimutkaisuuden johtamiseen [1]. Se toimii tekstin rakenteellisen monimutkaisuuden mittarina. Sitä sovelletaan poikkeavuuksien havaitsemiseen tiedonsiirrossa. Avainsanat: tekstin monimutkaisuus, poikkeamien havaitseminen, rakenteellinen monimutkaisuus, uudelleenkirjoitussääntö, kontekstiton kielioppi, L-järjestelmä.

**Tulos**

Symbolittoman sekvenssin syntaktisesti herkkä monimutkaisuus

**Esimerkki 2.1067**

Sanojen matalaulotteiset esitykset mahdollistavat<lb>tarkkojen NLP-mallien kouluttamisen rajoitetulla<lb>annotoidulla datalla. Vaikka useimmat representaatiot eivät huomioi sanojen paikallista kontekstia, luonnollinen tapa tuottaa kontekstiriippuvaisia representaatioita on muodostaa päätelmä todennäköisyysmalliin perustuvassa latenttimuuttujamallissa. Kun otetaan huomioon viimeaikainen menestys<lb>jatkuvien vektoriavaruus-sanarepresentaatioiden<lb>kehitämme tällaisen päättelymenettelyn<lb>jatkuville tiloille, joissa sanojen representaatiot<lb>annetaan lineaarisen dynaamisen<lb>systeemin posteriorisen keskiarvon avulla. Tällöin tehokas päättely voidaan<lb>suorittaa Kalmanin suodatuksen avulla. Oppimis<lb>algoritmimme on erittäin skaalautuva, sillä se toimii<lb>yksinkertaisilla yhteisesiintymisluvuilla sekä param-<lb>eterin alustuksessa momenttimenetelmää käyttäen<lb>että myöhemmissä EM-<lb>iteraatioissa. Kokeiluissamme käytämme pääteltyjä sanojen sulautuksia ominaisuuksina tavanomaisissa merkintätehtävissä, ja havaitsemme merkittäviä tarkkuusparannuksia. Lopuksi Kalman-suodattimen päivitykset voidaan nähdä lineaarisena toistuvana hermoverkkona. Osoitamme,<lb>että käyttämällä mallimme parametreja ini-<lb>tialisoimaan epälineaarisen rekursiivisen neuroverkon lan-<lb>guage-mallia lyhennetään sen koulutusaikaa päivällä<lb>ja saadaan alhaisempi perpleksiteetti.

**Tulos**

Lineaarisen dynaamisen järjestelmän malli tekstille

**Esimerkki 2.1068**

Tiivistelmässä on esitettävä yhteenveto artikkelin sisällöstä vähintään 70 ja enintään 150 sanalla. Se kirjoitetaan 9-kirjaimella, ja se sijoitetaan 1,0 cm:n etäisyydelle oikeasta ja vasemmasta marginaalista. Ennen ja jälkeen tiivistelmän on kaksi tyhjää riviä. . . .

**Tulos**

Diskriminoiva parametrien estimointi satunnaiskävelyjen segmentointia varten: Tekninen raportti

**Esimerkki 2.1069**

Stable Marriage Problem (SMP) on tunnettu yhteensovitusongelma, jonka Gale ja Shapley esittivät ja ratkaisivat ensimmäisen kerran [7]. Sittemmin on tutkittu useita muunnelmia ja laajennuksia tästä ongelmasta, jotta se kattaisi laajemman joukon sovelluksia. Aina kun uutta muunnelmaa tarkastellaan, on kuitenkin kehitettävä ja toteutettava uusi algoritmi. Vaihtoehtoisesti ehdotamme tässä asiakirjassa SMP:n koodausta käyttäen vastausjoukkojen ohjelmointia (ASP). Koodaustamme voidaan helposti laajentaa ja mukauttaa erityissovellusten tarpeisiin. Esimerkkinä näytämme, miten stabiileja vastineita voidaan löytää, kun yksilöt voivat nimetä kumppaneita, joita ei voida hyväksyä, ja kun mieltymysten väliset siteet sallitaan. Tämän jälkeen osoitamme, miten ASP:hen perustuva koodauksemme antaa luonnollisesti mahdollisuuden valita tietyt stabiilit matchingit, jotka ovat optimaalisia tietyn kriteerin mukaan. Joka kerta voimme luottaa yleisiin ja tehokkaisiin valmiisiin vastausjoukkojen ratkaisijoihin löytääksemme (optimaaliset) stabiilit yhteensovitukset.

**Tulos**

Vakaiden täsmäämisongelmien mallintaminen vastausjoukkojen ohjelmoinnilla

**Esimerkki 2.1070**

Ideatiheys (ID) mittaa sitä, kuinka nopeasti ideoita tai alkeispredikaatioita ilmaistaan lausumassa tai tekstissä. Alhaisemman ID:n on todettu olevan yhteydessä lisääntyneeseen riskiin sairastua Alzheimerin tautiin (Snowdon et al., 1996; Engelman et al., 2010). ID:tä on käytetty kahtena eri versiona: propositionaalinen ideatiheys (PID) laskee ilmaistut ideat, ja sitä voidaan soveltaa mihin tahansa tekstiin, kun taas semanttinen ideatiheys (SID) laskee ennalta määritellyt tietosisältöyksiköt, ja se soveltuu luonnollisesti paremmin normatiivisille alueille, kuten kuvien kuvaustehtäviin. Tässä artikkelissa kehitämme DEPID:n, uuden riippuvuussuhteisiin perustuvan menetelmän PID:n laskemiseksi, ja sen version DEPID-R:n, joka mahdollistaa toistuvien ideoiden poissulkemisen - AD-puheelle ominaisen piirteen. Vertailemme ensimmäistä kertaa automaattisesti poimittuja PID- ja SID-arvoja diagnostisessa luokittelutehtävässä kahdella eri AD-tietoaineistolla, jotka kattavat sekä suljetut aihealueet että vapaasti palautettavat alueet. Vaikka SID toimii paremmin normatiivisessa aineistossa, PID:n lisääminen johtaa pieneen mutta merkittävään parannukseen (+1,7 Fscore). Vapaa-aiheisessa tietokokonaisuudessa PID toimii odotetusti paremmin kuin SID (F-tulos 77,6 vs. 72,3), mutta automaattisen SID:n perustana olevasta sanojen sulauttamisen klusteroinnista johdettujen piirteiden lisääminen parantaa tuloksia huomattavasti, jolloin F-tulos on 84,8. Tämä on hyvä tulos.

**Tulos**

Ideatiheys Alzheimerin taudin ennustamiseksi transkriptoidusta puheesta

**Esimerkki 2.1071**

Korppien paradoksista kirjoittavat filosofit toteavat usein, että Nicodin ehto (NC) pätee, kun tietyt taustatiedot ovat olemassa, mutta ei päde, kun muut taustatiedot ovat olemassa, mutta harvoin he menevät pidemmälle. Toisin sanoen ei yleensä tutkita, mikä taustatieto tekee NC:stä todellisen tai väärän. Tämän artikkelin tavoitteena on täyttää tämä aukko. Meille "(objektiivinen) taustatieto" rajoittuu tietoon, joka voidaan ilmaista todennäköisyystapahtumina. Kaikki muu konfiguraatio katsotaan subjektiiviseksi ja a priori todennäköisyysjakauman ominaisuudeksi. Tutkimme NC:tä kahdessa erityisessä tilanteessa. Ensimmäisessä tapauksessa tiedetään täydellinen kuvaus joistakin yksilöistä, esimerkiksi tiedetään jokaisesta yksilöryhmästä, ovatko ne mustia ja ovatko ne korppeja. Toisessa tapauksessa tietyn ominaisuuden omaavien yksilöiden lukumäärä on annettu, esimerkiksi tiedetään, kuinka monta korppia tai kuinka monta mustaa on (kyseisessä populaatiossa). Vaikka jotkin paradoksin tunnetuimmista vastauksista ovat mitta-riippuvaisia, keskustelumme ei rajoitu mihinkään tiettyyn todennäköisyysmittaan. Kiinnostavin tuloksemme on, että jälkimmäisessä tapauksessa NC rikkoo yksinkertaista induktiivista päättelyä (nimittäin projisoitavuutta). Koska jälkimmäinen sääntö liittyy NC:hen verrattuna läheisemmin induktiivisen päättelyn intuitiiviseen käsitykseemme ja on suoremmin perusteltavissa sillä, tämä jännite puhuu NC:n uskottavuutta vastaan. Lopuksi ehdotamme, että NC:n epävirallinen esitys voi vaikuttaa intuitiivisesti uskottavalta, koska se voidaan helposti erehtyä luulemaan analogian avulla tapahtuvaksi päättelyksi.

**Tulos**

Nicodin ehdosta, induktiosäännöistä ja Korppi-paradoksista.

**Esimerkki 2.1072**

Varastoautomaatio on herättänyt viime vuosina suurta kiinnostusta, ja ehkäpä näkyvimmin se on näkynyt Amazon Picking Challenge -kilpailussa (APC) [1]. Täysin autonominen varaston poiminta- ja sijoittelujärjestelmä edellyttää vankkaa visuaalista näkemistä, joka tunnistaa ja paikantaa kohteet luotettavasti sekavissa ympäristöissä, itsestään sulkeutuvien kohtien, anturihäiriöiden ja suuren kohteiden määrän keskellä. Tässä artikkelissa esitellään lähestymistapa, jossa hyödynnetään monikuvamaista RGB-D-dataa ja itseohjautuvaa, tietoon perustuvaa oppimista näiden vaikeuksien voittamiseksi. Lähestymistapa oli osa MITPrinceton-tiimin järjestelmää, joka saavutti APC 2016 -tapahtumassa 3. ja 4. sijan ahtaus- ja poimintatehtävissä. Ehdotetussa lähestymistavassa segmentoidaan ja merkitään kohtauksen useita näkymiä täysin konvolutiivisen neuroverkon avulla ja sovitetaan sitten valmiiksi skannattuja 3D-objektimalleja tuloksena olevaan segmentointiin 6D-objektin asennon saamiseksi. Syvän neuroverkon kouluttaminen segmentointia varten vaatii yleensä suuren määrän harjoitusdataa. Ehdotamme itseohjautuvaa menetelmää, jolla voidaan luoda suuri merkitty tietokokonaisuus ilman työlästä manuaalista segmentointia. Osoitamme, että järjestelmämme pystyy arvioimaan luotettavasti esineiden 6D-asennon erilaisissa tilanteissa. Kaikki koodi, datat ja vertailuarvot ovat saatavilla osoitteessa http://www.cs.princeton.edu/∼andyz/apc2016.

**Tulos**

Moninäkymäinen itseohjautuva syväoppiminen 6D-poseerauksen arvioimiseksi Amazon Picking Challenge -haasteessa

**Esimerkki 2.1073**

Possibilistiset logiikkapohjat ja possibilistiset graafit ovat kaksi erilaista kiinnostavaa kehystä tiedon esittämiseksi. Ensin mainitussa luokitellaan tiedon osat (jotka ilmaistaan loogisilla kaavoilla) niiden varmuuden tason mukaan, kun taas jälkimmäisessä esitetään muuttujien väliset suhteet. Nämä kaksi esitystyyppiä ovat semanttisesti ekvivalentteja, kun ne johtavat samaan mahdollisuuksien jakaumaan (joka asettaa mahdolliset tulkinnat järjestykseen). Mahdollisuusjakauma voidaan purkaa ketjusäännön avulla, joka voi perustua kahteen erityyppiseen ehdollistamiseen, joita on olemassa mahdollisuusteoriassa (toinen perustuu tuloon numeerisessa ympäristössä ja toinen minimioperaatioon laadullisessa ympäristössä). Nämä kaksi ehdollistamistapaa aiheuttavat kahdenlaisia possibilistisia kuvaajia. Molemmissa tapauksissa esitetään näiden graafien kääntäminen possibilistisiksi perusteiksi. Käänteinen käännös possibilistisesta tietopohjasta min-pohjaiseksi graafiksi kuvataan myös.

**Tulos**

Possibilistiset logiikan perusteet ja possibilistiset graafit

**Esimerkki 2.1074**

Tämä on liitetiedote hiljattain tekemäämme tutkimukseen, jossa tarkastelimme rajoitettujen empaattisten ajallisen eron oppimisalgoritmien (ETD) heikkoja konvergenssiominaisuuksia teoreettisesta näkökulmasta. Se täydentää jälkimmäistä analyysia simulointituloksilla ja havainnollistaa joidenkin ETD-algoritmien käyttäytymistä kolmen esimerkkiongelman avulla.

**Tulos**

Joitakin simulaatiotuloksia empaattisista ajallisen eron oppimisalgoritmeista∗∗ .

**Esimerkki 2.1075**

Proteiinien subcellulaarisen lokalisoinnin ennustaminen on tärkeä ja haastava ongelma. Perinteiset biologiset kokeet ovat kalliita ja aikaa vieviä, joten yhä useammat tutkimukset suuntautuvat useisiin koneoppimismenetelmiin proteiinien subcellulaarisen sijainnin ennustamiseksi. Nykyisissä uusimmissa menetelmissä on kaksi pääongelmaa. Ensinnäkin suurin osa nykyisistä tekniikoista on suunniteltu käsittelemään moniluokkaista mutta ei monimerkkistä luokittelua, mikä jättää huomiotta useiden merkkien välisen yhteyden. Todellisuudessa monipaikkaiset proteiinit viittaavat siihen, että niillä on elintärkeitä ja ainutlaatuisia biologisia merkityksiä, joihin on syytä kiinnittää erityistä huomiota, eikä niitä voida jättää huomiotta. Toiseksi, tekniikat epätasapainoisen datan käsittelemiseksi monimerkkiluokitusongelmassa ovat merkittäviä, mutta niitä on vähemmän. Näiden kahden ongelman ratkaisemiseksi olemme kehittäneet HPSLPred-nimisen monimerkkiluokittimen, jota voidaan soveltaa monimerkkiluokitukseen, jossa proteiinilähde on epätasapainossa. HPSLPrediä varten perustettiin käyttäjäystävällinen verkkopalvelin osoitteeseen http://server.malab.cn/HPSLPred. Käyttäjien mukavuuden vuoksi.

**Tulos**

HPSLPred: Ensemble Multi-label Classifier for Human Protein Subcellular Location Prediction with Imbalanced Source: Ensemble Multi-label Classifier for Human Protein Subcellular Location Prediction with Imbalanced Source

**Esimerkki 2.1076**

Pieniulotteisten jakaumien koostaminen, jonka perusteet luotiin UAI'97 -tapahtuman pöytäkirjassa julkaistussa oppaassa (Jirousek 1997), näytti olevan vaihtoehtoinen väline moniulotteisten todennäköisyysmallien kuvaamiseen. Toisin kuin graafiset Markov-mallit, jotka määrittelevät moniulotteiset jakaumat deklaratiivisesti, tämä lähestymistapa on pikemminkin proseduraalinen. Matalaulotteisten jakaumien järjestäminen sopivaan järjestykseen määrittelee täysin vastaavan laskentamenettelyn; siksi erilaisten generoivien järjestysten tutkiminen on yksi alan keskeisistä ongelmista. Näin ollen vaikuttaa siltä, että erityiset, täydellisiksi kutsutut sekvenssit ovat tärkeässä asemassa. Niiden tärkeimmät karakterisointiteoriat esitellään tässä artikkelissa. Tämän artikkelin päätulos on kuitenkin ratkaisu marginalisaatio-ongelmaan yleisille sarjoille. Pääteoreemassa kuvataan tapa saada generoiva sekvenssi, joka määrittelee mielivaltaisen generoivan sekvenssin määrittelemän distription marginaalia vastaavan mallin. Tästä teoreemasta lukija näkee, missä määrin nämä laskutoimitukset ovat lo cal; ts. sarja koostuu marginaalijakaumista, joiden laskeminen on tehtävä summaamalla yli poistettavien muuttujien arvojen (tutkielmassa käsitellään äärellistä mallia) .

**Tulos**

Marginalisointi koostetuissa todennäköisyysmalleissa

**Esimerkki 2.1077**

Monilla tieteenaloilla, kuten tietojärjestelmissä, tietojenkäsittelytieteessä ja toiminnanohjauksessa, aikataulutusongelmat ovat tärkeitä päätöksentekotehtäviä. Koska monet aikataulutusongelmat ovat vahvassa mielessä NP-vaikeita, on tarpeen kehittää ratkaisuheuristiikkoja. Ratkaisumenetelmiä on tutkittu vain vähän sellaisten aikataulutusongelmien osalta, joissa asetusaika on rinnakkaisissa koneissa, jotka eivät ole yhteydessä toisiinsa, eikä rinnakkaisia tietokonearkkitehtuureja ole tietojemme mukaan vielä hyödynnetty. Korjaamme tämän puutteen ehdottamalla ja toteuttamalla uuden ratkaisuheuristiikan ja testaamalla erilaisia rinnakkaistamisstrategioita. Laskennallisissa kokeissamme osoitamme, että heuristiikkamme laskee lähes optimaalisia ratkaisuja myös suurille tapauksille ja että laskenta-aikaa voidaan vähentää huomattavasti rinnakkaistamisella.

**Tulos**

Korkean suorituskyvyn tietojenkäsittely aikataulupäätösten tukena: Rinnakkainen syvyys-ensimmäinen haku -heuristiikka

**Esimerkki 2.1078**

Tässä artikkelissa ehdotamme konvoluutiohermoverkkoja kysymys- ja vastauslauseiden optimaalisen esityksen oppimiseen. Niiden tärkein näkökohta on parin kahden jäsenen sanojen välisten vastaavuuksien antaman relaatiotiedon käyttö. Vastaukset koodataan sulautumina, joilla on lisäparametreja (ulottuvuuksia), jotka verkko virittää. Näiden avulla kysymysten ja vastausten väliset vuorovaikutussuhteet voidaan ottaa paremmin huomioon, mikä parantaa tarkkuutta merkittävästi. Testaamme mallejamme kahdella laajalti käytetyllä vastauslauseiden valintavertailumallilla. Tulokset osoittavat selvästi, miten tehokkaasti relaatiotietomme mahdollistaa sen, että suhteellisen yksinkertainen verkkomme lähestyy alan huippua.

**Tulos**

Kysymys-vastausparien suhteellisten tietojen mallintaminen konvolutiivisten neuroverkkojen avulla

**Esimerkki 2.1079**

Äärellisen automaatin synkronoiva sekvenssi on syötesekvenssi, joka vie kaikki tilat samaan tilaan. Synkronoivan sekvenssin olemassaolon tarkistaminen ja synkronoivan sekvenssin löytäminen, jos sellainen on olemassa, voidaan suorittaa polynomiajassa. Lyhimmän synkronoivan sekvenssin löytämisen ongelman tiedetään kuitenkin olevan NP-vaikea. Tässä työssä tutkitaan vastausjoukko-ohjelmoinnin käyttökelpoisuutta tämän optimointiongelman ratkaisemisessa verrattuna raakavoima-algoritmeihin ja SAT-pohjaisiin lähestymistapoihin.

**Tulos**

Lyhimpien synkronointisekvenssien tuottaminen vastausjoukkojen ohjelmoinnin avulla

**Esimerkki 2.1080**

Tässä artikkelissa ehdotetaan hierarkkista tarkkaavaisuus-neuraalista käännösmallia, jossa keskitytään parantamaan lähdepuolen hierarkkisia representaatioita kattamalla sekä paikallinen että globaali semanttinen informaatio kaksisuuntaisen puupohjaisen koodaimen avulla. Kohdesanojen ennustamistodennäköisyyden maksimoimiseksi käytetään huomiomekanismin painotettua muunnelmaa tasapainottamaan huomiotietoa leksikaalisten ja fraasivektoreiden välillä. Harvinaisten sanojen puupohjaisen koodauksen avulla ehdotettua mallia laajennetaan sananalatasolle, jotta sanaston ulkopuolisen sanan (OOV) ongelmaa voidaan lieventää. Empiiriset tulokset osoittavat, että ehdotettu malli päihittää merkittävästi sekvenssistä sekvenssiin -huomioon perustuvat ja puupohjaiset neuraaliset käännösmallit englannin ja kiinan välisissä käännöstehtävissä.

**Tulos**

Kohti kaksisuuntaisia hierarkkisia representaatioita tarkkaavaisuuteen perustuvassa neuraalisessa konekääntämisessä

**Esimerkki 2.1081**

Tässä artikkelissa esitellään hybridilähestymistapa verkkopalvelujen automaattiseen koostamiseen, joka tuottaa semanttisia inputoutput-pohjaisia koosteita, joilla on optimaalinen päästä päähän - QoS, minimoiden tuloksena syntyvän koosteen palvelujen määrän. Ehdotetussa lähestymistavassa on neljä päävaihetta: 1) koostumuksen graafin luominen pyyntöä varten; 2) optimaalisen koostumuksen laskeminen, joka minimoi yhden objektiivisen QoS-funktion; 3) monivaiheiset optimoinnit hakuavaruuden pienentämiseksi tunnistamalla ekvivalentit ja hallitsevat palvelut; ja 4) hybridi paikallinen ja globaali haku optimaalisen QoS:n saamiseksi mahdollisimman pienellä palvelumäärällä. Laajamittainen validointi Web Service Challenge 2009-2010 -kilpailun tietokokonaisuuksilla ja satunnaisesti luoduilla tietokokonaisuuksilla osoittaa, että: ja globaalin optimoinnin yhdistelmä on yleinen ja tehokas tekniikka optimaalisten koostumusten löytämiseksi erilaisissa skenaarioissa, ja 2) hybridistrategia toimii paremmin kuin uusin tekniikka, sillä se tuottaa ratkaisuja, joissa on vähemmän palveluita ja optimaalinen QoS. Avainsanat Palvelukokoonpano; palveluoptimointi; hybridialgoritmi; QoS-tietoinen; semanttiset verkkopalvelut.

**Tulos**

Hybridioptimointialgoritmi laajamittaiseen QoS-tietoiseen palvelukokoonpanoon

**Esimerkki 2.1082**

Graafinen moniagenttimalli (GMM) edustaa agenttijoukon käyttäytymisen yhteistä jakaumaa. Yksi tietolähde agenttien käyttäytymisestä voi olla peliteoreettinen analyysi, jota kuvaavat useat viime vuosina kehitetyt graafiset peliesitykset. GMM:t yleistävät tätä lähestymistapaa ilmaisemaan mielivaltaisia jakaumia, jotka perustuvat pelikuvauksiin tai muihin tietolähteisiin, jotka vaikuttavat uskomuksiin agenttien käyttäytymisestä. GMM:ien joustavuuden havainnollistamiseksi esitellään pelistä johdettuja malleja, jotka sallivat todennäköisyyteen perustuvan poikkeaman tasapainosta, sekä malleja, jotka perustuvat heuristiseen toimintojen valintaan. Tutkimme kolmea erilaista menetelmää näiden mallien integroimiseksi yhteen malliin, joka edustaa yhdistettyjä tietolähteitä. Arvioidaksemme yhdistetyn mallin ennustuskykyä käsittelemme todellisena tuloksena vahvistusoppimisprosessin tuottamaa käyttäytymistä. Havaitsimme, että kahden tietolähteen yhdistäminen millä tahansa menetelmällä tuottaa parempia ennusteita kuin kumpikaan tietolähde yksinään. Yhdistämismenetelmistä tiedon yhdistäminen päihittää tässä empiirisessä kokeessa tutkitut mielipidepooli- ja suoran päivityksen menetelmät.

**Tulos**

Tietämyksen yhdistäminen graafisissa moniagenttimalleissa

**Esimerkki 2.1083**

Ristretto: Convolutional Neural Networksin laitteistokohtainen approksimointi Convolutional Neural Networks (CNN) on saavuttanut viime vuosina merkittäviä läpimurtoja. Niiden suorituskyky tietokonenäössä on vastannut ihmisen kykyjä ja joillakin aloilla jopa ylittänyt ne. Syvät neuroverkot pystyvät kuvaamaan monimutkaisia epälineaarisia piirteitä, mutta tämän kyvyn hintana ovat suuret laskenta- ja muistivaatimukset. Uusimmat neuroverkot vaativat miljardeja aritmeettisia operaatioita ja miljoonia parametreja. Jotta sulautetut laitteet, kuten älypuhelimet, Google-lasit ja valvontakamerat, saisivat käyttöönsä syväoppimisen hämmästyttävän tehon, voidaan käyttää erityisiä laitteistokiihdyttimiä, joilla voidaan vähentää sekä suoritusaikaa että virrankulutusta. Sovelluksissa, joissa nopeaa yhteyttä pilvipalveluun ei voida taata tai joissa yksityisyys on tärkeää, laskenta on suoritettava paikallisesti. Viime aikoina on ehdotettu monia laitteistokiihdyttimiä syviä neuroverkkoja varten. Kiihdyttimien suunnittelun ensimmäinen tärkeä vaihe on syvien neuroverkkojen laitteistokohtainen approksimointi, joka mahdollistaa energiatehokkaan päättelyn. Esittelemme Ristretton, nopean ja automatisoidun kehyksen CNN-approksimaatiota varten. Ristretto simuloi mukautetun laitteistokiihdyttimen laitteistoaritmetiikkaa. Kehys pienentää verkon parametrien ja resurssi-intensiivisten kerrosten ulostulojen bittinopeutta, mikä pienentää merkittävästi sirun pinta-alaa kertolaskuyksiköiden osalta. Vaihtoehtoisesti Ristretto voi poistaa kertojien tarpeen kokonaan, jolloin saadaan aikaan pelkkä yhteenlaskijaaritmetiikka. Työkalu hienosäätää trimmattuja verkkoja korkean luokittelutarkkuuden saavuttamiseksi. Koska syvien neuroverkkojen kouluttaminen voi olla aikaa vievää, Ristretto käyttää erittäin optimoituja rutiineja, jotka toimivat GPU:lla. Tämä mahdollistaa minkä tahansa verkon nopean tiivistämisen. Kun enimmäistoleranssi on 1 %, Ristretto pystyy onnistuneesti tiivistämään CaffeNet- ja SqueezeNet-verkot 8-bittisiksi. Ristretton koodi on saatavilla.

**Tulos**

Ristretto: Laitteisto-orientoitunut konvoluutio-neuraaliverkkojen approksimointi

**Esimerkki 2.1084**

Robottisovellusten yleistyminen on johtanut valtavan määrän jäsentymättömän datan syntymiseen, kun taas nykyinen pilvipalveluinfrastruktuuri on suunniteltu käsittelemään rajallisia määriä jäsenneltyä dataa. Tämän ongelman ratkaisemiseksi ehdotamme robottipilvipalveluja varten oppimis-muistamis-recall-reduce-paradigmaa (learn-memorize-recall-reduce paradigm). Oppimisvaihe muuntaa saapuvan jäsentymättömän datan jäsennellyksi dataksi, muistinvaihe tarjoaa tehokkaan tallennustilan massiiviselle tietomäärälle, muistinvaihe tarjoaa tehokkaan keinon hakea raakadataa ja pelkistämisvaihe tarjoaa keinot, joilla tämä massiivinen määrä jäsentymätöntä dataa voidaan ymmärtää rajallisilla laskentaresursseilla.

**Tulos**

Opi-muistiin-palauta-pelkistä-pelkistä vähennä: Robottipilvilaskennan paradigma

**Esimerkki 2.1085**

Dialogijärjestelmä on järjestelmä, joka on vuorovaikutuksessa ihmisen kanssa luonnollisella kielellä. Tällä hetkellä monet yliopistot kehittävät vuoropuhelujärjestelmää omalla alueellisella kielellään. Tässä asiakirjassa käsitellään dialogijärjestelmää, sen komponentteja, haasteita ja arviointia. Tämä asiakirja auttaa tutkijoita saamaan tietoa dialogijärjestelmästä.

**Tulos**

Dialogijärjestelmä: Lyhyt katsaus

**Esimerkki 2.1086**

Esitämme Bayes-menetelmän useiden neliömatriisien, jotka eivät välttämättä ole symmetrisiä, likimääräistä diagonalisointia varten. Näiden matriisien yhteisten ominaisvektoreiden ja ominaisarvojen näytteiden simuloimiseksi johdetaan Gibbs-näytteenottaja. Useita synteettisiä esimerkkejä käytetään havainnollistamaan ehdotetun Gibbs-näytteenottimen suorituskykyä, ja sen jälkeen tehdään vertailuja useisiin muihin yhteisdiagonalisointialgoritmeihin, mikä osoittaa, että Gibbs-näytteenottimella saavutetaan uusinta tekniikkaa vastaava suorituskyky tarkastelluissa esimerkeissä. Sivutuotteena Gibbsin näytteenottimen tulosta voitaisiin käyttää log marginaalisen todennäköisyyden estimointiin, mutta käytämme Bayesin informaatiokriteeriin (BIC) perustuvaa approksimaatiota, joka tarkastelluissa synteettisissä esimerkeissä paikallisti oikein yhteisten ominaisvektoreiden määrän. Sovelsimme näytteenottajaa onnistuneesti lähteiden erotteluongelmaan sekä yhteiseen pääkomponenttianalyysiin ja yhteiseen spatiaaliseen kuvioanalyysiin.

**Tulos**

Bayesilainen lähestymistapa neliömatriisien likimääräiseen yhteiseen diagonalisointiin

**Esimerkki 2.1087**

Ketjurakenteinen pitkä lyhytkestoinen muisti (LSTM) on osoittautunut tehokkaaksi monenlaisissa ongelmissa, kuten puheentunnistuksessa ja konekääntämisessä. Tässä artikkelissa ehdotamme sen laajentamista puurakenteisiin, joissa muistisolu voi heijastaa rekursiivisessa prosessissa useiden lapsisolujen tai useiden jälkeläissolujen historiamuistia. Kutsumme mallia S-LSTM:ksi, joka tarjoaa periaatteellisen tavan tarkastella pitkän matkan vuorovaikutusta hierarkioiden, esimerkiksi kielen tai kuvan jäsentelyrakenteiden, yli. Hyödynnämme malleja semanttiseen kompositioon tekstin merkityksen ymmärtämiseksi, mikä on perustavanlaatuinen ongelma luonnollisen kielen ymmärtämisessä, ja osoitamme, että se päihittää uusinta rekursiivista mallia korvaamalla sen kompositiokerrokset S-LSTM-muistilohkoilla. Osoitamme myös, että annettujen rakenteiden hyödyntäminen auttaa saavuttamaan paremman suorituskyvyn kuin ilman rakenteiden huomioimista.

**Tulos**

Pitkä lyhytkestoinen muisti puurakenteiden yli

**Esimerkki 2.1088**

Tässä asiakirjassa esitetään kokonaisnäkemys systeemiseen lupus erythematosukseen (SLE) liittyvistä tieteellisistä julkaisuista, ja lähtökohtana ovat artikkelien tiivistelmät. Tiivistelmät ovat ajan myötä kehittyneet yhä monimutkaisemmiksi käytettyjen termien osalta, mikä tekee välttämättömäksi käyttää kehittyneitä tilastollisia menetelmiä ja vastata muun muassa seuraaviin kysymyksiin: Miten sanasto kehittyy ajan myötä? Mitkä artikkelit ovat vaikutusvaltaisimpia? Mitkä artikkelit ovat ottaneet käyttöön uusia termejä ja sanastoa? Näihin kysymyksiin vastaamiseksi analysoimme tietokokonaisuutta, joka koostuu 506 tiivistelmästä, jotka on ladattu 115 eri lehdestä ja jotka kattavat 18 vuoden ajanjakson.

**Tulos**

Miten tieteellinen kirjallisuus on kehittynyt ajan myötä? Uusi tilastollinen lähestymistapa, jossa käytetään sananseurantaan perustuvia menetelmiä.

**Esimerkki 2.1089**

Palvelutasosopimus (Service Level Agreement, SLA) on olennainen osa pilvipalvelujärjestelmiä, jotta asiakkaille voidaan varmistaa palveluiden maksimaalinen saatavuus. Jos SLA-sopimusta rikotaan, palveluntarjoajan on maksettava seuraamuksia. Näin ollen SLA-sopimusten rikkomisen ennustaminen hyödyttää sekä asiakkaita että palveluntarjoajia. Tässä asiakirjassa tarkastelemme kahta koneoppimismallia: Naive Bayes- ja Random Forest -luokittelijoita SLA-sopimusrikkomusten ennustamiseksi. Koska SLA-sopimusrikkomukset ovat reaalimaailmassa harvinainen tapahtuma (∼ 0,2 %), luokittelutehtävästä tulee haastavampi. Näiden haasteiden voittamiseksi käytämme useita uudelleennäytteenottomenetelmiä, kuten Random Over and Under Sampling, SMOTH, NearMiss (1,2,3), One-sided Selection, Neighborhood Cleaning Rule jne., tasapainottaaksemme tietokokonaisuuden uudelleen. Käytämme Google Cloud Cluster -jälkeä tietokokonaisuutena näiden eri menetelmien tutkimiseen. Huomaamme, että satunnaismetsät, joissa on SMOTE-ENN-uusintanäytteenotto, toimivat muiden menetelmien joukossa parhaiten, sillä niiden tarkkuus on 0,9988 % ja F1-tulos 0,9980.

**Tulos**

SLA-rikkomusten ennustaminen pilvipalveluissa: Koneoppimisen näkökulma

**Esimerkki 2.1090**

Monet kombinatoriset ongelmat käsittelevät preferenssejä ja rikkomuksia, joiden tavoitteena on löytää ratkaisuja, joiden kustannukset ovat mahdollisimman pienet. Painotettu rajoitusten tyydyttäminen on kehys tällaisten ongelmien mallintamiseen, joka koostuu joukosta kustannusfunktioita, joilla mitataan muuttujien määritysten eri yhdistelmien rikkomusastetta tai preferenssejä. Tyypilliset painotettujen rajoitusten tyydyttämisongelmien (WCSP) ratkaisumenetelmät perustuvat haara- ja sidontahakuun, joka tehdään käytännölliseksi käyttämällä tehokkaita johdonmukaisuustekniikoita, kuten AC\*, FDAC\* ja EDAC\*, piilotetun kustannustiedon päättelemiseksi ja arvojen karsimiseksi haun aikana. Nämä tekniikat on kuitenkin suunniteltu niin, että ne ovat tehokkaita vain taulukkomuodossa esitetyille binäärisille ja ternäärisille kustannusfunktioille. Monien tosielämän ongelmien ratkaisemisessa tarvitaan korkean ariteetin (tai globaalin) kustannusfunktioita. Tutkimme tehokasta esitysmallia ja algoritmeja, joiden avulla johdonmukaisuustekniikoiden hyödyt voidaan tuoda myös korkean ariteetin kustannusfunktioihin, jotka usein johdetaan klassisen rajoitusten tyydyttämisen kovista globaaleista rajoitteista. Kirjallisuuden mukaan jotkin globaalit kustannusfunktiot voidaan esittää virtausverkkoina, ja minimikustannusvirtausalgoritmia voidaan käyttää tällaisten verkkojen minimikustannusten laskemiseen polynomiajassa. Osoitamme, että tämän virtaukseen perustuvan algoritmimenetelmän naiivi soveltaminen globaaleihin kustannusfunktioihin voi johtaa vahvempaan ∅-käänteisen johdonmukaisuuden muotoon. Lisäksi osoitamme, miten menetelmää voidaan muokata käsittelemään kustannusprojektioita ja laajennuksia, jotta voidaan ylläpitää yleistettyjä versioita AC\*:sta ja FDAC\*:sta kustannusfunktioille, joissa on enemmän kuin kaksi muuttujaa. Vastaava yleistäminen vahvemmalle EDAC\*:lle ei ole yhtä suoraviivaista. Paljastamme värähtelyongelman, kun EDAC\*:a sovelletaan kustannusfunktioihin, joissa on useampi kuin yksi muuttuja. Oskillaation välttämiseksi ehdotamme heikkoa versiota EDAC\*:sta ja yleistämme sen heikoksi EDGAC\*:ksi ei-binäärisille kustannusfunktioille. Käyttämällä erilaisia vertailuarvoja, joissa käytetään kovien globaalien rajoitusten ALLDIFFERENT, GCC, SAME ja REGULAR pehmeitä muunnoksia, empiiriset tulokset osoittavat, että ehdotuksemme parantaa tilannetta jopa kertaluokkaa perinteiseen rajoitusten optimointimenetelmään verrattuna sekä ajan että karsinnan suhteen.

**Tulos**

Yhdenmukaisuustekniikat virtaukseen perustuville projisointiturvallisille globaaleille kustannusfunktioille painotetussa rajoitusten tyydyttämisessä

**Esimerkki 2.1091**

Jotta voimme kuroa umpeen ihmisten ja koneiden välisen kuilun kuvien ymmärtämisessä ja kuvaamisessa, tarvitsemme lisätietoa siitä, miten ihmiset kuvaavat havaitsemansa kohtauksen. Tässä artikkelissa tutkimme alhaalta ylöspäin suuntautuvan saliaatiopohjaisen visuaalisen huomion ja kohteiden viittausten välistä yhteisymmärrystä kohtauksen kuvauskonstruktioissa. Tutkimme ihmisten kirjoittamien kuvausten ja koneella luotujen kuvausten ominaisuuksia. Sen jälkeen ehdotamme saliaatiopainotteista kuvakuvauskirjoitusmallia, jonka avulla voidaan tutkia matalan tason vihjeistä saatavia hyötyjä kielimalleissa. Opimme, että (1) ihmiset mainitsevat kuvauksissaan korostuneemmat kohteet aikaisemmin kuin vähemmän korostuneet, (2) mitä paremmin kuvatekstimalli suoriutuu, sitä paremmin se vastaa ihmisten kuvauksia, (3) ehdotettu saliencyboostattu malli ei perusmuotoonsa verrattuna paranna merkittävästi MS COCO -tietokannassa, mikä osoittaa, että eksplisiittinen alhaalta ylöspäin tapahtuva tehostaminen ei auta, kun tehtävä on opittu hyvin ja viritetty aineistoon, (4) saliencyboostatulla mallilla havaitaan kuitenkin parempaa yleistämiskykyä aineistoissa, joita ei ole nähty.

**Tulos**

Voiko erottuvuustiedosta olla hyötyä kuvien kuvatekstimalleille?

**Esimerkki 2.1092**

Robusti pääkomponenttianalyysi (RPCA) on tehokas työkalu, jolla voidaan palauttaa matalien sijojen rakenne puhtaista tiedoista, jotka ovat korruptoituneet harvan kohinan tai poikkeavien tekijöiden kanssa. Monissa matalan tason visio-ongelmissa tiedetään, että puhtaan datan taustalla oleva rakenne on matalarivinen, mutta myös puhtaan datan tarkka arvo tunnetaan. Kun näihin ongelmiin sovelletaan tavanomaista rangin minimointia, tavoitefunktio muotoillaan tavalla, joka ei täysin hyödynnä ongelmia koskevaa a priori-tietoa rangista. Tämä havainto motivoi meitä tutkimaan, onko olemassa parempaa vaihtoehtoista ratkaisua, kun käytetään rank-minimointia. Tässä artikkelissa ehdotamme ydinnormien minimoinnin sijasta singulaariarvojen osittaissumman minimointia, mikä implisiittisesti kannustaa tavoitejärjestysrajoituksen noudattamiseen. Kokeelliset analyysimme osoittavat, että kun näytteiden määrä on puutteellinen, lähestymistapamme johtaa suurempaan onnistumisprosenttiin kuin perinteinen rankin minimointi, kun taas näillä kahdella lähestymistavalla saadut ratkaisut ovat lähes identtisiä, kun näytteiden määrä on yli riittävä. Sovellamme lähestymistapaamme erilaisiin matalan tason visio-ongelmiin, kuten suuren dynaamisen alueen kuvantamiseen, liikkeen reunojen havaitsemiseen, fotometriseen stereoon, kuvien kohdistamiseen ja palauttamiseen, ja osoitamme, että tuloksemme ovat parempia kuin perinteisellä ydinnormin rankin minimointimenetelmällä saadut tulokset.

**Tulos**

Singulaaristen arvojen osittaissumman minimointi robustissa PCA:ssa: algoritmi ja sovelluksia

**Esimerkki 2.1093**

Ehdolliset todennäköisyydet ovat keskeinen käsite koneoppimisessa. Esimerkiksi merkin Y optimaalinen ennustaminen syötteen X perusteella vastaa Y:n ehdollisen todennäköisyyden maksimointia X:n perusteella. Yleinen lähestymistapa päättelytehtäviin<lb>on ehdollisten todennäköisyyksien mallin oppiminen. Nämä mallit perustuvat kuitenkin usein<lb>vahvoihin oletuksiin (esim, log-lineaariset mallit), ja siksi niiden estimaatti<lb>ehdollisista todennäköisyyksistä ei ole kestävä ja riippuu suuresti niiden oletusten paikkansapitävyydestä.<lb>Tässä ehdotamme kehystä ehdollisten todennäköisyyksien päättelyyn ilman<lb>olettamusta jakaumien taustalla olevista jakaumista, paitsi tietämystä niiden<lb>toisen kertaluvun marginaaleista, jotka voidaan estimoida datan perusteella. Näytämme, miten tämä<lb>asettelu johtaa ehdollisten todennäköisyyksien taattuihin rajoihin, jotka voidaan laskea<lb>tehokkaasti erilaisissa tilanteissa, kuten strukturoidussa ennustamisessa. Lopuksi<lb>sovellamme niitä puolivalvottuun syväoppimiseen ja saamme tuloksia, jotka ovat kilpailukykyisiä<lb>muuttuja-autokoodereiden kanssa.

**Tulos**

Robustit ehdolliset todennäköisyydet

**Esimerkki 2.1094**

Arabian asiakirjojen klusterointi on tärkeä tehtävä, jotta perinteisillä tiedonhakujärjestelmillä (IR) voidaan saada hyviä tuloksia, etenkin kun arabiankielisten online-asiakirjojen määrä kasvaa nopeasti. Asiakirjojen klusteroinnilla pyritään automaattisesti ryhmittelemään samankaltaisia asiakirjoja yhteen klusteriin käyttäen erilaisia samankaltaisuus- ja etäisyysmittareita. Tähän tehtävään vaikuttaa usein asiakirjojen pituus, ja asiakirjojen hyödylliseen informaatioon liittyy usein suuri määrä kohinaa, minkä vuoksi on tarpeen poistaa tämä kohina ja säilyttää samalla hyödyllinen informaatio, jotta asiakirjojen klusteroinnin suorituskykyä voidaan parantaa. Tässä asiakirjassa ehdotamme, että arvioimme tekstin tiivistämisen vaikutusta latentin semanttisen analyysin mallin avulla arabialaisten asiakirjojen klusterointiin edellä mainittujen ongelmien ratkaisemiseksi käyttämällä viittä samankaltaisuus-/etäisyysmittaa: Euklidinen etäisyys, kosininen samankaltaisuus, Jaccard-kerroin, Pearsonin korrelaatiokerroin ja Kullback-Leiblerin keskiarvoinen divergenssi kahdella eri tavalla: ilman stemmingiä ja stemmingin kanssa. Kokeelliset tulokset osoittavat, että ehdotettu lähestymistapa ratkaisee tehokkaasti kohinaisten tietojen ja asiakirjojen pituuden aiheuttamat ongelmat ja parantaa siten merkittävästi klusterointitulosta.

**Tulos**

SEMANTTINEN ANALYYSI ARABIALAISTEN ASIAKIRJOJEN KLUSTEROINNIN TEHOSTAMISEKSI

**Esimerkki 2.1095**

Tieteellisiin mittauksiin liittyvän tiedon ymmärtämiseksi ja dokumentoimiseksi on tehty merkittäviä ponnisteluja. Monet näistä ponnisteluista johtivat yhteen tai useampaan korkealaatuiseen ontologiaan, jotka kuvaavat joitakin tieteellisten mittausten näkökohtia, mutta eivät kattavasti ja johdonmukaisesti integroituna. Huomaamme esimerkiksi, että monet näistä korkealaatuisista ontologioista eivät ole kunnolla linjassa keskenään, ja mikä vielä haastavampaa, niissä on erilaisia ja usein ristiriitaisia käsitteitä ja lähestymistapoja empiirisiä mittauksia koskevan tiedon koodaamiseksi. Yhtenäisen näkemyksen puuttumisen vuoksi tutkijoiden on usein haastavaa määrittää, onko kaksi tieteellistä mittausta tehty semanttisesti yhteensopivalla tavalla, mikä vaikeuttaa päätöstä siitä, pitäisikö mittauksia analysoida yhdessä vai ei. Tässä artikkelissa esittelemme Human-Aware Sensor Network Ontology -ontologian, joka on anturi-infrastruktuuri-ontologian ja provenance-ontologian kattava linjaus ja integrointi. HASNetOa on kehitetty yli vuoden ajan, ja useat tiedeyhteisöt ovat tarkastelleet, jakaneet ja käyttäneet sitä. Ontologia on ollut käytössä useiden laajamittaisten ekologisten seurantatoimien (havaintojen) ja empiiristen kokeiden tiedonhallinnan tukena.

**Tulos**

Ihmistietoinen sensoriverkon ontologia: Semanttinen tuki empiiriselle tiedonkeruulle

**Esimerkki 2.1096**

Tässä artikkelissa tutkitaan stokastisia ja vastakohtaisia kombinatorisia monikätisiä bandit-ongelmia. Stokastisessa tilanteessa johdamme ensin ongelmakohtaiset katumuksen alarajat ja analysoimme, miten nämä rajat skaalautuvat päätösavaruuden ulottuvuuden mukaan. Sen jälkeen ehdotamme COMBUCB-algoritmeja, jotka hyödyntävät tehokkaasti ongelman kombinatorista rakennetta, ja johdamme niiden katumukselle äärellisen ajan ylärajan. Nämä rajat parantavat olemassa olevien algoritmien katumuksen ylärajoja, ja osoitamme numeerisesti, että COMBUCB on huomattavasti parempi kuin mikään muu algoritmi. Ehdotamme kaksi yksinkertaista algoritmia, nimittäin COMBEXP-1 ja COMBEXP-2 semibandit- ja bandit-palautetta varten. Niiden katumukset skaalautuvat samalla tavalla kuin uusimpien algoritmien, vaikka niiden toteutus on yksinkertainen.

**Tulos**

Stokastiset ja vastapuoliset kombinatoriset rosvot (Stochastic and Adversarial Combinatorial Bandits)

**Esimerkki 2.1097**

Sarkasmia pidetään yhtenä tunneanalyysin vaikeimmista ongelmista. Indonesian sosiaalisessa mediassa tekemämme havainnon mukaan ihmisillä on tietyissä aiheissa tapana kritisoida jotakin asiaa sarkasmilla. Tässä ehdotimme kahta lisäominaisuutta sarkasmin havaitsemiseksi sen jälkeen, kun yleinen tunneanalyysi on suoritettu. Ominaisuudet ovat negatiivisuusinformaatio ja interjektiosanojen määrä. Käytimme myös käännettyä SentiWordNetiä sentimenttiluokittelussa. Kaikki luokittelut tehtiin koneoppimisalgoritmeilla. Kokeelliset tulokset osoittivat, että lisäominaisuudet ovat varsin tehokkaita sarkasmin havaitsemisessa. Avainsanat- Sentimenttianalyysi, sarkasmi, luokittelu,

**Tulos**

Indonesian sosiaalisen median tunteiden analysointi sarkasmin havaitsemisen avulla

**Esimerkki 2.1098**

Syvien neuroverkkojen menestyksen laajentaminen luonnollisen kielen ymmärtämiseen ja symboliseen päättelyyn edellyttää monimutkaisia toimintoja ja ulkoista muistia. Viimeaikaiset neuro-ohjelmien induktiomenetelmät ovat yrittäneet ratkaista tämän ongelman, mutta ne rajoittuvat tyypillisesti erilaistuvaan muistiin, eivätkä näin ollen voi skaalautua pieniä synteettisiä tehtäviä laajemmalle. Tässä työssä ehdotamme Manager-ProgrammerComputer-kehystä, joka integroi neuroverkot ei-differentioitavaan muistiin tukeakseen abstrakteja, skaalautuvia ja tarkkoja operaatioita ystävällisen neurotietokoneen käyttöliittymän avulla. Esittelemme erityisesti neuraalisen symbolisen koneen, joka sisältää sekvenssistä sekvenssiin toimivan neuraalisen "ohjelmoijan" ja ei-eristyvän "tietokoneen", joka on koodia avustava Lisp-tulkki. Jotta REINFORCEa voidaan käyttää menestyksekkäästi harjoitteluun, sitä täydennetään likimääräisillä kultaohjelmilla, jotka löydetään iteratiivisella maksimiluotettavuusharjoitteluprosessilla. NSM kykenee oppimaan semanttisen jäsentäjän heikosta valvonnasta laajassa tietopohjassa. Se saavuttaa uuden huippuluokan suorituskyvyn WEBQUESTIONSSP:ssä, joka on haastava semanttinen jäsennystietokanta. Aiempiin lähestymistapoihin verrattuna NSM on kokonaisvaltainen, joten se ei ole riippuvainen ominaisuuksien suunnittelusta tai toimialakohtaisesta tiedosta.

**Tulos**

Neuraaliset symboliset koneet: Freebase-tietokannan semanttisten jäsentäjien oppiminen heikolla valvonnalla (lyhyt versio).

**Esimerkki 2.1099**

Säädön perustavanlaatuinen ongelma on oppia havainnoista järjestelmän malli, jota voidaan hyödyntää säätimien synteesissä. Jotta nykyiset menetelmät tarjoaisivat hyvät suorituskykytakeet, niiden on oletettava, että todellinen järjestelmä kuuluu oppimisen aikana tarkasteltujen mallien luokkaan. Esittelemme iteratiivisen menetelmän, jolla on vahvat takuut myös agnostisessa tapauksessa, jossa järjestelmä ei kuulu luokkaan. Näytämme erityisesti, että mitä tahansa katumatonta online-oppimisalgoritmia voidaan käyttää lähes optimaalisen politiikan saamiseksi, jos jollakin mallilla saavutetaan alhainen koulutusvirhe ja pääsy hyvään eksploraatiojakaumaan. Lähestymistapamme soveltuu sekä diskreetteihin että jatkuviin alueisiin. Osoitamme sen tehokkuuden ja skaalautuvuuden kirjallisuudessa esitetyllä haastavalla helikopterialueella.

**Tulos**

Agnostinen järjestelmän tunnistaminen mallipohjaista vahvistusoppimista varten

**Esimerkki 2.1100**

Kahden viime vuosikymmenen aikana on ehdotettu useita sumeisiin aikasarjoihin perustuvia ennustemenetelmiä. Suurin osa sumeiden aikasarjojen menetelmistä on esitetty autojen liikenneonnettomuuksien ennustamiseen. Nykyisten menetelmien ennustetarkkuus ei kuitenkaan ole riittävän hyvä. Tässä asiakirjassa verrataan ehdottamaamme uutta sumeiden aikasarjojen ennustamismenetelmää olemassa oleviin menetelmiin. Menetelmämme perustuu auto-onnettomuuksien historiatietojen jakamiseen keskiarvojen perusteella. Ehdotettu menetelmä kuuluu k:nnen kertaluvun ja aikamuuttujien menetelmiin. Ehdotetulla menetelmällä voidaan saavuttaa paras ennustetarkkuusaste auto-onnettomuuksien ennustamisessa kuin nykyisillä menetelmillä. AvainsanatSumeat joukot, sumeat loogiset ryhmät, sumeat tiedot, sumeat aikasarjat.

**Tulos**

Epätarkkuuden minimointi sumeiden tietojoukkojen jakamisella - analyyttisen menetelmän validointi

**Esimerkki 2.1101**

Elävät organismit kietovat toisiinsa pehmeät (esim. lihakset) ja kovat (esim. luut) materiaalit, mikä antaa niille luontaisen joustavuuden ja kimmoisuuden, joka usein puuttuu perinteisistä jäykistä roboteista. Pehmeän robotiikan kehittyvällä alalla pyritään hyödyntämään näitä samoja ominaisuuksia joustavien koneiden luomiseksi. Pehmeiden materiaalien luonne asettaa kuitenkin huomattavia haasteita suunnitteluun, rakentamiseen ja ohjaukseen liittyville näkökohdille, ja tähän asti valtaosa pehmeiden robottien kävelytavoista on suunniteltu käsin empiirisen kokeilun ja erehdyksen avulla. Tässä käsikirjoituksessa kuvataan helposti koottava, kireysmateriaalipohjainen pehmeä robotti, joka kykenee erittäin dynaamisiin liikkeisiin ja joka osoittaa rakenteellista ja käyttäytymiskestävyyttä fyysisten vaurioiden sattuessa. Tämän mahdollistaa koneoppimisalgoritmi, joka pystyy löytämään uusia kävelytapoja minimaalisella määrällä fyysisiä kokeiluja. Nämä tulokset vahvistavat entisestään pehmeitä robottimenetelmiä, joilla pyritään hyödyntämään monimutkaisen materiaalidynamiikan vuorovaikutusta ja luomaan runsaasti dynaamisia käyttäytymismalleja.

**Tulos**

Pehmeät tensegrity-robotit

**Esimerkki 2.1102**

Tekstin ja kuvan välisen tiedon kääntäminen on tekoälyn perusongelma, joka yhdistää luonnollisen kielen käsittelyn ja tietokonenäön. Viime vuosina kuvatekstien tuottamisen suorituskyky on parantunut merkittävästi rekursiivisten neuroverkkojen (RNN) ansiosta. Samaan aikaan tekstistä kuvaan -tekniikan avulla on alettu tuottaa uskottavia kuvia käyttämällä tietosarjoja tietyistä luokista, kuten linnuista ja kukista. On jopa nähty kuvien tuottamista moniluokkaisista tietokokonaisuuksista, kuten Microsoft Common Objects in Context (MSCOCO), generatiivisten vastakkaisverkkojen (GAN) avulla. Monimutkaisen muotoisten objektien syntetisointi on kuitenkin edelleen haastavaa. Esimerkiksi eläimillä ja ihmisillä on monia vapausasteita, mikä tarkoittaa, että ne voivat saada monia monimutkaisia muotoja. Ehdotamme uutta koulutusmenetelmää nimeltä Image-Text-Image (I2T2I), jossa yhdistetään teksti-kuvasta-kuvaan ja kuvasta-tekstiin (kuvateksti) -synteesi ja parannetaan teksti-kuvasta-kuvaan -synteesin suorituskykyä. Osoitamme, että I2T2I pystyy tuottamaan MSCOCO:n avulla parempia moniluokkaisia kuvia kuin nykytekniikka. Osoitamme myös, että I2T2I voi saavuttaa siirto-oppimista käyttämällä esivalmennettua kuvatekstimoduulia tuottamaan ihmiskuvia MPII Human Pose -tietokannassa (MHP) ilman lauseannotaatiota.

**Tulos**

I2T2I: TEKSTIN JA KUVAN SYNTEESIN OPPIMINEN TEKSTIDATAN LISÄÄMISEN AVULLA.

**Esimerkki 2.1103**

Koneoppimisessa kantajoukon tunnistamisongelmana on löytää joukko satunnaismuuttujia, jotka selittävät parhaiten valittua muuttujaa, kun otetaan huomioon tiedot ja jokin ennalta määritetty pisteytysfunktio. Tämä ongelma on kriittinen osa Bayes-verkkojen rakenteiden oppimista ja Markov-huopien löytämistä, ja siksi sillä on monia käytännön sovelluksia petosten havaitsemisesta kliiniseen päätöksenteon tukeen. Tässä artikkelissa esitellään uusi hajautetun muistin lähestymistapa tarkan vanhempien joukkojen määritysongelmaan. Skaalautuvuuden saavuttamiseksi johdamme teoreettisia rajoja hakuavaruuden rajoittamiseksi, kun käytetään MDL-pisteytysfunktiota, ja järjestämme taustalla olevan dynaamisen ohjelmoinnin uudelleen siten, että laskentatiheys kasvaa ja hienojakoinen synkronointi poistuu. Tämän jälkeen suunnittelemme lähestymistapamme tehokkaan toteutuksen Apache Spark -alustalle. Kokeellisten tulosten avulla osoitamme, että menetelmä säilyttää vahvan skaalautuvuuden 500-ytimisessä itsenäisessä Spark-klusterissa, ja sitä voidaan käyttää 70 muuttujaa sisältävien datajoukkojen tehokkaaseen käsittelyyn, mikä on paljon nykyisin saatavilla olevien ratkaisujen ulottumattomissa.

**Tulos**

Skaalautuva tarkka vanhempien joukkojen tunnistaminen Bayes-verkkojen oppimisessa Apache Sparkilla

**Esimerkki 2.1104**

Tehokkaat ja vankat algoritmit hajautettuun estimointiin verkoissa ovat olennaisen tärkeitä monissa hajautetuissa järjestelmissä. Otoskeskiarvojen hajautettu estimointi on ollut paljon esillä, mutta U-tilastojen laskeminen, joka perustuu kalliimpaan keskiarvojen laskemiseen havaintoparien yli, on vähemmän tutkittu alue. Tällaiset datafunktiot ovat kuitenkin olennaisen tärkeitä tilastollisen populaation yleisten ominaisuuksien kuvaamisessa, ja tärkeitä esimerkkejä ovat muun muassa Area Under the Curve, empiirinen varianssi, Gini-keskiarvoero ja klusterin sisäinen pistehajonta. Tässä artikkelissa ehdotetaan uusia synkronisia ja asynkronisia satunnaistettuja juorualgoritmeja, jotka samanaikaisesti levittävät tietoja verkon yli ja ylläpitävät paikallisia estimaatteja kiinnostavasta U -tilastosta. Määritämme konvergenssinopeusrajat O(1/t) ja O(log t/t) synkroniselle ja asynkroniselle tapaukselle, joissa t on iteraatioiden lukumäärä, ja joissa on eksplisiittiset datasta ja verkosta riippuvat termit. Nopeusanalyysin kannalta suotuisten vertailujen lisäksi numeeriset kokeet antavat empiiristä näyttöä siitä, että ehdotetut algoritmit ylittävät aiemmin esitellyn lähestymistavan.

**Tulos**

Gossip-algoritmien laajentaminen U-tilastojen hajautettuun estimointiin.

**Esimerkki 2.1105**

Tässä artikkelissa esitellään luonnollisista dialogeista kerätty tietokokonaisuus, jonka avulla voidaan testata dialogijärjestelmien kykyä oppia uusia faktoja käyttäjän lausumista koko dialogin ajan. Tämä vuorovaikutteinen oppiminen auttaa ratkaisemaan yhden avoimen toimialan dialogijärjestelmien yleisimmistä ongelmista, joka on dialogijärjestelmän päättelyyn soveltuvien faktojen vähäisyys. Ehdotettu tietokokonaisuus, joka koostuu 1900 kerätystä dialogista, mahdollistaa vuorovaikutteisen oppimisen simuloinnin, jossa käyttäjiltä saadaan vuorovaikutteisia selityksiä ja kysymyksiä, joita voidaan käyttää vuorovaikutteiseen oppimiseen.

**Tulos**

Tiedonkeruu vuorovaikutteista oppimista varten dialogin avulla

**Esimerkki 2.1106**

Politiikan arvioinnissa pyritään arvioimaan arvofunktio, joka ennustaa tilojen pitkän aikavälin arvot tietyllä politiikalla. Se on ratkaiseva vaihe monissa vahvistusoppimisalgoritmeissa. Tässä artikkelissa keskitytään politiikan arviointiin lineaarisen funktion approksimaation avulla kiinteän tietokokonaisuuden avulla. Ensin muutamme empiirisen politiikan arviointiongelman (kvadraattiseksi) kongevoidun-konkaavaksi satulapisteongelmaksi ja esittelemme sitten primaari-duaali-erägradienttimenetelmän sekä kaksi stokastista varianssin vähennysmenetelmää ongelman ratkaisemiseksi. Nämä algoritmit skaalautuvat lineaarisesti sekä otoskoon että ominaisuuksien ulottuvuuden suhteen. Lisäksi ne saavuttavat lineaarisen konvergenssin myös silloin, kun satulapisteongelmalla on vain vahva konkaviteetti kaksoismuuttujissa mutta ei vahvaa kuperuutta alkumuuttujissa. Numeeriset kokeet vertailuongelmilla osoittavat menetelmiemme tehokkuuden.

**Tulos**

Stokastiset varianssin vähentämismenetelmät politiikan arvioinnissa

**Esimerkki 2.1107**

Strukturoitu harva optimointi on tärkeä ja haastava ongelma, kun analysoidaan korkea-ulotteista dataa monissa eri sovelluksissa, kuten bioinformatiikassa, lääketieteellisessä kuvantamisessa, sosiaalisissa verkostoissa ja tähtitieteessä. Vaikka useita strukturoituja harvennusmalleja, kuten puita, ryhmiä, klustereita ja polkuja, on tutkittu, kytkettyjä aliraportteja on tutkittu nykyisessä kirjallisuudessa harvoin. Yksi tärkeimmistä teknisistä haasteista on se, että ei ole olemassa mitään strukturoitua harvuutta aiheuttavaa normia, joka voisi suoraan mallintaa kytkettyjen aliraporttien avaruutta, eikä kytketyille aliraporteille ole olemassa täsmällistä projisointiorakkelia sen NP-kovuuden vuoksi. Tässä artikkelissa tutkitaan tehokkaita likimääräisiä projektio-orakkeleita yhdistetyille aliraporteille ja ehdotetaan kahta uutta tehokasta algoritmia, GRAPH-IHT ja GRAPH-GHTP, joilla optimoidaan yleistä epälineaarista kohdefunktiota, jolle on asetettu muuttujien tukeen kohdistuva liitettävyysrajoitus. Ehdotetuilla algoritmeilla on vahvat takuut, jotka vastaavat useita nykyisiä harvuusrajoitteisen optimoinnin menetelmiä, kuten Projected Gradient Descent (PGD), Approximate Model Iterative Hard Thresholding (AM-IHT) ja Gradient Hard Thresholding Pursuit (GHTP), konvergenssinopeuden ja approksimointitarkkuuden suhteen. Sovellamme ehdotettuja algoritmeja optimoidaksemme useita tunnettuja graafien skannausstatistiikkoja useissa sovelluksissa, jotka koskevat kytkettyjen aliraporttien havaitsemista, ja kokeelliset tulokset osoittavat, että ehdotetut algoritmit ylittävät nykyaikaiset menetelmät.

**Tulos**

Tekninen raportti: Graph-Structured Sparse Optimization for Connected Subgraph Detection (Graafirakenteinen harva optimointi yhdistetyn aliraportin havaitsemiseen)

**Esimerkki 2.1108**

Koska reaalimaailman ongelmien ennustamiseen liittyy epätarkkuutta ja epävarmuutta, keinotekoiset neuroverkkotekniikat (ANN) ovat tulleet yhä hyödyllisemmiksi mallintamisessa ja optimoinnissa. Tässä asiakirjassa esitellään keinotekoinen neuroverkko-lähestymistapa sähköenergian kulutuksen ennustamiseen. Sähköjärjestelmien tehokkaan suunnittelun ja toiminnan varmistamiseksi energiaoperaattorit tarvitsevat optimaalisia ennustusvälineitä voiton maksimoimiseksi ja energian kuluttajien mahdollisimman suuren tyytyväisyyden saavuttamiseksi. Gazan kaistalla kulutettua sähköenergiaa koskevat kuukausittaiset tiedot kerättiin vuodesta 1994 vuoteen 2013. Tiedot koulutettiin ja ehdotettu malli validoitiin käyttämällä 2-kertaista ja K-kertaista ristiinvalidointitekniikkaa. Mallia on testattu todellisilla energiankulutustiedoilla, ja sen suorituskyky on tyydyttävä.

**Tulos**

Keinotekoisten neuroverkkomenetelmien käyttö sähköenergian kulutuksen ennustamiseen

**Esimerkki 2.1109**

Ehdotamme siirtosyväoppimisen (TDL) kehystä, jolla voidaan siirtää yhden modaalisen neuroverkon avulla saatu tieto verkkoon, jossa on eri modaliteetti. Erityisesti osoitamme, että voimme hyödyntää puhedataa hienosäätääksemme videotunnistusta varten koulutettua verkkoa, kun käytössä on aluksi samassa semantiikassa oleva rinnakkainen audio-videotietoaineisto. Lähestymistapamme oppii ensin analogiaa säilyttävät sulautumat kunkin verkon välikerroksista opittujen abstraktien representaatioiden välille, mikä mahdollistaa semantiikkatason siirron lähde- ja kohdemodaliteettien välillä. Sitten sovellamme neuroverkko-operaatiotamme, joka hienosäätää kohdeverkkoa lähdeverkosta siirretyn lisätiedon avulla pitäen kohdeverkon topologian muuttumattomana. Vaikka esitämme lähestymistapamme sovelluksena audiovisuaalisen tunnistustehtävän, kehyksemme on joustava, joten se voi toimia minkä tahansa multimodaalisen tietokokonaisuuden kanssa tai minkä tahansa jo olemassa olevan syväverkon kanssa, jolla on yhteinen perussemantiikka. Tässä keskeneräistä työtä koskevassa raportissa pyrimme tarjoamaan kattavia tuloksia ehdotetun lähestymistavan eri kokoonpanoista kahdessa laajalti käytetyssä audiovisuaalisessa tietokokonaisuudessa ja keskustelemme ehdotetun lähestymistavan mahdollisista sovelluksista.

**Tulos**

Multimodaalinen syväoppiminen ja sovellukset audio-visuaaliseen tunnistukseen

**Esimerkki 2.1110**

Tässä artikkelissa esitellään toimintakieleen ja vastausjoukkojen ohjelmointiin perustuva lähestymistapa suunnittelu- ja aikataulutusongelmien ratkaisemiseen hybridialueilla, joilla on sekä diskreettiä että jatkuvaa käyttäytymistä. Käytämme toimintakieltä H toimialueen esittämiseen ja käännämme tuloksena olevan teorian A-Prolog-ohjelmaksi. Tällä tavoin vähennämme suunnittelu- ja aikataulutusongelmien ratkaisujen löytämisen ongelman A-Prolog-ohjelmien vastausjoukkojen laskemiseen. Esitämme kirjallisuudessa olevan suunnittelu- ja aikataulutusesimerkin ja näytämme, miten se mallinnetaan H-kielellä. Näytämme, miten tuloksena oleva H-teoria käännetään vastaavaksi A-Prolog-ohjelmaksi. Laskemme tuloksena saadun ohjelman vastausjoukot käyttämällä hybridiratkaisinta nimeltä EZCSP, joka integroi löyhästi rajoitteiden ratkaisijan ja vastausjoukkojen ratkaisijan. Ratkaisijan avulla voimme päätellä rajoitteita reaalien yli ja laskea ratkaisuja monimutkaisiin suunnittelu- ja aikataulutusongelmiin. Tulokset ovat osoittaneet, että lähestymistapaamme voidaan soveltaa mihin tahansa suunnittelu- ja aikataulutusongelmaan hybridialueilla.

**Tulos**

Suunnittelu ja aikataulutus hybridialueilla vastausjoukkojen ohjelmointia käyttäen

**Esimerkki 2.1111**

Tarkastelemme aluksi k-pariteettien oppimisen ongelmaa on-line virheisiin sidotussa mallissa: kun on annettu piilovektori x ∈ {0, 1}, jossa |x| = k, ja sarja "kysymyksiä" a1, a2, - - - - ∈ {0, 1} , jossa algoritmin on vastattava jokaiseen kysymykseen sanoilla 〈ai, x〉 (mod 2), mikä on paras kompromissi algoritmin tekemien virheiden lukumäärän ja algoritmin aikakompleksisuuden välillä? Parannamme Buhrmanin et al. [BGM10] aiempaa parasta tulosta exp(k)-kerroin aikakompleksisuuden suhteen. Toiseksi tarkastelemme ongelmaa, joka koskee k-pariteettien oppimista luokittelukohinan ollessa läsnä nopeudella η ∈ (0, 1/2). Polynomiaikaisen algoritmin löytäminen tälle ongelmalle (kun η > 0 ja k = ω(1)) on oppimisteorian pitkäaikainen haaste. Grigorescu et al. [GRV11] osoittivat algoritmin, joka toimii ajassa ( n k/2 )1+4η+o(1) . Huomaa, että tämä algoritmi vaatii luonnostaan aikaa ( n k/2 ) myös silloin, kun kohinanopeus η on polynomisesti pieni. Havaitaan, että riittävän pienellä kohinanopeudella on mahdollista rikkoa ( n k/2 ) -raja. Erityisesti, jos jollekin funktiolle f(n) = ω(1) ja α ∈ [1/2, 1), k = n/f(n) ja η = o(f(n)/log n), niin ongelmalle on olemassa algoritmi, jonka suoritusaika on poly(n) - ( n k )1-α - e.

**Tulos**

K-pariteettien oppimisesta kohinan kanssa ja ilman kohinaa

**Esimerkki 2.1112**

Englannista intiankieliseen konekääntämiseen liittyy rakenteellisten ja morfologisten eroavaisuuksien haaste. Tässä asiakirjassa kuvataan englannista intian kieleen tapahtuvaa tilastollista konekääntämistä käyttäen esijärjestystä ja suffiksien erottelua. Esijärjestämisessä käytetään sääntöjä lähdelauseiden rakenteen siirtämiseksi ennen harjoittelua ja kääntämistä. Tämä syntaktinen rakenneuudistus auttaa tilastollista konekäännöstä ratkaisemaan rakenteelliset eroavaisuudet ja siten parantamaan käännöksen laatua. Suffiksien erottelua käytetään englannin kielen ja erittäin agglutinatiivisten intialaisten kielten välisten morfologisten eroavaisuuksien poistamiseen. Osoitamme, että esijärjestelyn ja suffiksien erottelun käyttö auttaa parantamaan englannin ja intian kielen konekääntämisen laatua.

**Tulos**

MTIL17: Tilastollinen konekääntäminen englannista intian kieleen MTIL17: Tilastollinen konekääntäminen englannista intian kieleen

**Esimerkki 2.1113**

Tutkimme binääriluokittelun ongelmaa koneoppimisessa, mutta luokittimen sallitaan pidättäytyä mistä tahansa tiedosta, jolloin se voi olla tietämätön todellisesta luokkamerkinnästä sitoutumatta mihinkään ennusteeseen. Tämä perustuu suoraan sovelluksiin, kuten lääketieteelliseen diagnostiikkaan ja petosriskien arviointiin, joissa virheellisillä ennusteilla voi olla tuhoisia seurauksia. Keskitymme viimeaikaiseen teoreettisesti perusteltuun työhön tällä alalla, jossa kuvataan, miten pidättäytymisen salliminen voi johtaa virheiden vähenemiseen hyvin yleisissä asetuksissa. Kaksi aluetta nostetaan esiin: yllättävä mahdollisuus virheettömään oppimiseen ja perustavanlaatuinen kompromissi riittävän usein tapahtuvan ennustamisen ja virheellisten ennusteiden välttämisen välillä. Tarkastelemme tehokkaita algoritmeja, joilla on todistettavat takuut kummallekin näistä alueista. Käsittelemme myös yhteyksiä muihin skenaarioihin, erityisesti aktiiviseen oppimiseen, sillä ne viittaavat lupaaviin jatkotutkimussuuntiin tällä kehittyvällä alalla.

**Tulos**

Pidättäytymisen hyödyllisyys binääriluokittelussa

**Esimerkki 2.1114**

Evoluutio on johtanut siihen, että monet luonnolliset älykkäät olennot ovat kehittyneet hyvin pitkälle ja pystyvät ennustamaan mekaanisia ilmiöitä nopeasti ja tarkasti. Ihmiset ovat menestyksekkäästi kehittäneet fysiikan lakeja tällaisten mekaanisten ilmiöiden abstrahoimiseksi ja mallintamiseksi. Tekoälyn yhteydessä viimeaikaisessa työssä on keskitytty arvioimaan fysikaalisia parametreja aistitietojen perusteella ja käyttämään niitä fysikaalisissa simulaattoreissa pitkän aikavälin ennusteiden tekemiseen. Sen sijaan me tutkimme yhden neuroverkon tehokkuutta mekaanisten ilmiöiden pitkäaikaisennusteiden loppuun asti tapahtuvassa ennustamisessa. Laajan arvioinnin perusteella osoitamme, että tällaiset verkot voivat päihittää vaihtoehtoiset lähestymistavat, joilla on jopa pääsy ground-truth-fysikaalisiin simulaattoreihin, erityisesti silloin, kun jotkin fysikaaliset parametrit ovat havaitsemattomia tai niitä ei tunneta a priori. Lisäksi verkostomme tuottaa tulosten jakauman, joka kuvaa tietojen luontaista epävarmuutta. Lähestymistapamme osoittaa ensimmäistä kertaa, että anturidatasta voidaan tehdä käyttökelpoisia pitkän aikavälin ennusteita ilman, että taustalla olevia fysikaalisia lakeja on mallinnettava eksplisiittisesti.

**Tulos**

Fyysisen pitkän aikavälin ennustajan oppiminen

**Esimerkki 2.1115**

Semanttinen segmentointi edellyttää kuvapikselien yksityiskohtaista merkitsemistä objektiluokittain. Yksittäisten kohteiden yksityiskohtaisen muodon kuvaamiseen tarvitaan paikallisista kuvapisteistä saatua tietoa. Tämä tieto on kuitenkin monitulkintaista ja voi johtaa kohinaisiin merkintöihin. Kuvan sisällön globaalilla päättelyllä voidaan sen sijaan vangita yleiset semanttiset käsitteet. Kannatamme sitä, että kuvan käsitteiden kokonaisvaltainen päättely tarjoaa arvokasta tietoa yksityiskohtaista pikselimerkintää varten. Ehdotamme yleistä kehystä, jolla voidaan hyödyntää kokonaisvaltaista tietoa LabelBank-pankin muodossa pikselitason segmentoinnissa. Osoitamme, että kehyksemme pystyy parantamaan semanttisen segmentoinnin suorituskykyä erilaisissa ympäristöissä. Opettelemme malleja holistisen LabelBankin poimimiseksi visuaalisista vihjeistä, attribuuteista ja/tai tekstikuvauksista. Osoitamme parannuksia semanttisen segmentoinnin tarkkuudessa vakiotietoaineistoissa useilla uusimmilla segmentointiarkkitehtuureilla ja holistisilla päättelymenetelmillä.

**Tulos**

LabelBank: Semanttisen segmentoinnin globaalien näkökulmien uudelleentarkastelu

**Esimerkki 2.1116**

Vaiheen hakuongelmissa ratkaistaan lineaarisia yhtälöitä, mutta niistä puuttuu merkkitieto (tai vaihe, jos kyseessä ovat<lb>kompleksiset luvut). Kahden viime vuosikymmenen aikana suosittu yleinen empiirinen lähestymistapa<lb> tämän ongelman moniin muunnelmiin on ollut vuorotteleva minimointi eli vuorottelu puuttuvan vaiheinformaation ja ratkaisukandidaatin arvioinnin välillä. Tässä artikkelissa osoitamme, että yksinkertainen vuorotteleva minimointialgoritmi konvergoi geometrisesti yhden tällaisen ongelman ratkaisuun<lb> - vektorin x löytäminen y,A:sta, jossa y = |Ax| ja |z| tarkoittaa vektoria<lb>, joka koostuu z:n elementtiviisaista suuruuksista - olettaen, että A on Gaussinen.<lb>Empiirisesti algoritmimme toimii samankaltaisesti kuin hiljattain ehdotetut konveksiset tekniikat tälle<lb>muuttujalle (jotka perustuvat "nostamiseen" konveksiseen matriisiongelmaan) näytteenottokompleksisuudeltaan ja<lb>kestävyydeltään kohinan suhteen. Algoritmimme on kuitenkin paljon tehokkaampi ja se voi skaalautua suuriin<lb>ongelmiin. Analyyttisesti osoitamme geometrisen konvergenssin ratkaisuun ja näytteenoton monimutkaisuuden<lb>, joka poikkeaa log-kertoimilla ilmeisistä alarajoista. Todistamme myös lähes optimaalisen skaalautumisen<lb>tapauksessa, jossa tuntematon vektori on harva. Työmme edustaa ainoaa tunnettua teoreettista<lb>takuuta vuorottelevalle minimoinnille mille tahansa vaihehakuongelmien muunnokselle ei<lb>konveksisessa ympäristössä.

**Tulos**

Vaiheen haku käyttäen vuorottelevaa minimointia

**Esimerkki 2.1117**

Vuorovaikutteisten toimijoiden dynaamisissa järjestelmissä tapahtuvien toimenpiteiden kausaalivaikutusten arviointi on alikehittynyt. Tässä asiakirjassa tarkastelemme ongelman monimutkaisuutta tavanomaisten lähestymistapojen avulla ja osoitamme, että tarvitaan sopivampia menetelmiä. Neyman-Rubinin kausaalimallin pohjalta kehitämme kausaalisen päättelymenetelmän ja selvitämme vakausoletukset, jotka ovat välttämättömiä pätevän kausaalisen päättelyn kannalta. Menetelmämme koostuu ajallisesta komponentista, joka mallintaa agenttien omaksumien käyttäytymismallien kehittymistä ajan kuluessa, ja käyttäytymiskomponentista, joka mallintaa agenttien toimien jakautumista hyväksyttyjen käyttäytymismallien perusteella. Tämä mahdollistaa kiinnostavien suureiden pitkän aikavälin estimaattien imputoimisen ja siten interventioiden pitkän aikavälin kausaalivaikutusten arvioinnin. Esittelemme menetelmämme käyttäytymispeliteoriaan perustuvalla tietokokonaisuudella ja keskustelemme avoimista ongelmista, jotka kannustavat tulevaan tutkimukseen.

**Tulos**

Neyman-Rubin-mallin mukainen tilastollinen päättely pitkän aikavälin kausaalivaikutuksista monitoimijajärjestelmissä.

**Esimerkki 2.1118**

Tässä artikkelissa esitellään keskustelumalli, joka sisältää sekä kontekstin että osallistujan roolin kahden osapuolen keskusteluissa. Tutkitaan erilaisia arkkitehtuureja osallistujan roolin ja kontekstitiedon integroimiseksi LSTM-kielimalliin (Long Short-term Memory). Keskustelumalli voi toimia kielimallina tai kielen generointimallina. Ubuntu Dialog Corpus -korpuksella tehdyt kokeet osoittavat, että mallimme pystyy kuvaamaan usean vuoron vuorovaikutusta osallistujien välillä. Ehdotettu menetelmä päihittää perinteisen LSTM-mallin mitattuna kielimallin perpleksisyydellä ja vastausjärjestyksellä. Generoidut vastaukset osoittavat tyypillisiä eroja kahden osallistujan roolien välillä.

**Tulos**

LSTM-pohjaiset keskustelumallit

**Esimerkki 2.1119**

Oikeinkirjoitusvirheitä syntyy tekstiin joko kirjoittamisen aikana tai silloin, kun käyttäjä ei tiedä oikeaa foneemia tai grafeemia. Jos kieli sisältää monimutkaisia sanoja, kuten sandhi, joissa kaksi tai useampi morfeemi liittyy toisiinsa joidenkin sääntöjen perusteella, oikeinkirjoituksen tarkistamisesta tulee hyvin työlästä. Tällaisissa tilanteissa on erittäin hyödyllistä, että oikeinkirjoituksen tarkistusohjelma, jossa on sandhi-jakajalaite, ilmoittaa käyttäjälle virheistä ja antaa ehdotuksia. Tässä artikkelissa ehdotetaan uutta sandhien jakoalgoritmia. Sandhi-jakajalla voidaan jakaa noin 7000 yleisintä sandhi-sanaa Kannadan kielessä, joita käytetään testinäytteinä. Sandhi-jakaja integroitiin kannadan kielen oikeinkirjoituksen tarkistusohjelmaan, ja siihen lisättiin ehdotusten luomiseen tarkoitettu mekanismi. Näin kehitettiin kattava, alustasta riippumaton, itsenäinen oikeinkirjoituksen tarkistusohjelma, jossa on sandhi-jakajan sovellusohjelmisto, jota testattiin laajasti sen tehokkuuden ja oikeellisuuden varmistamiseksi. Tämän sandhi-jakajalla varustetun oikeinkirjoituksen tarkistusohjelman vertaileva analyysi tehtiin, ja tulosten mukaan sandhi-jakajalla varustetun Kannada-oikeinkirjoituksen tarkistusohjelman suorituskyky oli parempi. Se on kaksi kertaa nopeampi, 200 kertaa tilatehokkaampi ja 90-prosenttisen tarkka monimutkaisissa substantiiveissa ja 50-prosenttisen tarkka monimutkaisissa verbeissä. Tällaisella sandhi-jakajalla varustetulla oikeinkirjoituksen tarkistusohjelmalla on suuri merkitys muun muassa konekäännösjärjestelmissä ja äänenkäsittelyssä. Tämä on ensimmäinen sandhi-jakajalaite kannadaksi, ja uuden algoritmin etuna on, että se voidaan laajentaa kaikkiin intialaisiin kieliin. Avainsanat- Luonnollisen kielen prosessointi; morfologia; laskennallinen kielitiede; Sandhi splitter; oikeinkirjoituksen tarkistus.

**Tulos**

Kannadan oikeinkirjoituksen tarkistusohjelma Sandhi Splitterin kanssa

**Esimerkki 2.1120**

Artikkelin tavoitteena on tarjota tarkka lähestymistapa hierarkkisesta CRM:stä poimitun Poisson-prosessin generoimiseksi ilman, että satunnaismittojen äärettömän montaa atomia tarvitsee instantioida. Käytämme täysin satunnaisia mittoja (CRM) ja hierarkkisia CRM:iä Poisson-prosessien priorin määrittelyyn. Johdamme tuloksena olevan pisteprosessin marginaalijakauman, kun taustalla oleva CRM on marginalisoitu pois. Käyttämällä hyvin tunnettuja ominaisuuksia, jotka ovat ainutlaatuisia Poisson-prosesseille, pystyimme johtamaan tarkan lähestymistavan Poisson-prosessin instantioimiseksi hierarkkisen CRM:n priorin avulla. Lisäksi johdamme Gibbsin näytteenottostrategioita hierarkkisille CRM-malleille, jotka perustuvat kiinalaisen ravintolan franchising-näytteenottojärjestelmään. Esimerkkinä esittelemme yleistetyn gammaprosessin summan (SGGP) ja osoitamme sen sovelluksen aihepiirien mallintamisessa. Osoitamme, että aiheiden ja sanojen power-law-käyttäytyminen voidaan määrittää bayesiläisellä tavalla määrittelemällä SGGP:n parametreille ennakkoarvo.

**Tulos**

Hierarkkisten täysin satunnaisten mittausten romahtaneesta esittämisestä (Completely Random Measures)

**Esimerkki 2.1121**

väitteet trigrammit trigrammit riippuvuudet AMT-koulutetut kiinteät riippuvuudet 3,699% 4,697% 5,974% 18,18% 3,99% 2,797% 3,696% 2,597% 3,297% FEATURES (TFIDF:t) SVM-luokittelijan virhetaso DATASETTI

**Tulos**

Automaattisen patenttivaatimusten jäsentelyn parantaminen: Dataset, järjestelmä ja kokeet

**Esimerkki 2.1122**

Tässä artikkelissa tutkitaan symmetriaan perustuvaa hakuavaruuden pienentämistekniikkaa, joka voi nopeuttaa optimaalista polunhakua suuntaamattomilla tasakustannuksisilla ruudukkokartoilla jopa 38-kertaisesti. Tekniikkamme hajottaa ruudukkokartat joukoksi tyhjiä suorakulmioita poistamalla jokaisesta suorakulmiosta kaikki sisäpuoliset solmut ja mahdollisesti joitakin solmuja reunoilta. Tämän jälkeen lisäämme sarjan makrosärmiä valittujen jäljelle jääneiden kehän solmuparien välille, jotta voimme helpottaa todistettavasti optimaalista kulkua kunkin suorakulmion läpi. Kehitämme myös uudenlaisen online-karsintatekniikan, jolla nopeutamme hakua entisestään. Algoritmimme on nopea, muistitehokas ja säilyttää samat optimaalisuus- ja täydellisyystakeet kuin haku muokkaamattomalla ruutukartalla.

**Tulos**

Symmetriaan perustuva hakuavaruuden pienentäminen rasterikarttoja varten

**Esimerkki 2.1123**

Suosittelujärjestelmät hyödyntävät käyttäjien demografisia tietoja, kuten ikää, sukupuolta jne., suositusten personoimiseksi ja kohdennettujen mainosten sijoittamiseksi paremmin. Usein käyttäjät<lb>eivät vapaaehtoisesti anna näitä tietoja yksityisyyden suojaan liittyvien huolenaiheiden vuoksi tai<lb>ei oma-aloitteisesti täytä verkkoprofiilejaan.<lb>Kuvaamme uutta uhkaa, jossa suositusjärjestelmä oppii<lb>yksityisiä ominaisuuksia käyttäjiltä, jotka eivät vapaaehtoisesti paljasta niitä. Suunnittelemme sekä passiivisia että aktiivisia hyökkäyksiä, jotka<lb>pyytävät strategisesti valittujen kohteiden arvosanoja ja joita suositusjärjestelmä voisi siten<lb>käyttää tämän salaisen<lb>agendan toteuttamiseen. Menetelmämme perustuvat Bayesin<lb>matriisifaktoroinnin uuteen käyttöön aktiivisessa oppimisympäristössä. Arvioinnit<lb>t useilla tietokokonaisuuksilla osoittavat, että tällaiset hyökkäykset ovat<lb>toteutettavissa ja käyttävät huomattavasti vähemmän luokiteltuja kohteita kuin<lb>staattiset päättelymenetelmät. Tärkeää on, että ne onnistuvat uhraamatta käyttäjille annettujen suositusten laatua.

**Tulos**

Suositteleminen asialistan kanssa: Yksityisten attribuuttien aktiivinen oppiminen matriisitekijöiden avulla

**Esimerkki 2.1124**

Suurin osa verkkonäyttömainoksista tarjotaan reaaliaikaisen tarjouksen (RTB) avulla - jokainen näyttövaikutelma huutokaupataan reaaliaikaisesti, kun se on juuri syntynyt käyttäjän käynnin seurauksena. Jotta mainos voidaan sijoittaa automaattisesti ja optimaalisesti, mainostajien on tärkeää kehittää oppiva algoritmi, jolla mainosvaikutelma voidaan tarjota fiksusti reaaliaikaisesti. Useimmissa aiemmissa teoksissa tarjouspäätöstä tarkastellaan staattisena optimointiongelmana, jossa joko käsitellään kunkin vaikutelman arvoa itsenäisesti tai asetetaan tarjoushinta kullekin mainosvolyymin segmentille. Tietyn mainoskampanjan tarjoaminen tapahtuisi kuitenkin toistuvasti sen elinkaaren aikana ennen kuin budjetti loppuu. Näin ollen kukin tarjous on strategisesti riippuvainen rajoitetusta budjetista ja kampanjan kokonaistehokkuudesta (esim. tuotetuista klikkauksista saadut palkkiot), joka havaitaan vasta kampanjan päätyttyä. Näin ollen on erittäin kiinnostavaa laatia optimaalinen tarjousstrategia peräkkäin, jotta kampanjabudjetti voidaan jakaa dynaamisesti kaikille käytettävissä oleville näyttökerroille sekä välittömien että tulevien palkkioiden perusteella. Tässä asiakirjassa muotoilemme tarjouksen päätöksentekoprosessin vahvistusoppimisongelmaksi, jossa tilaavaruutta edustavat huutokauppatiedot ja kampanjan reaaliaikaiset parametrit, kun taas toiminta on asetettava tarjoushinta. Mallintamalla tilasiirtymä huutokauppakilpailun avulla luomme Markov-päätösprosessin kehyksen optimaalisen tarjouspolitiikan oppimiseksi mainonnan suorituskyvyn optimoimiseksi dynaamisessa reaaliaikaisessa tarjousympäristössä. Lisäksi suuren reaalimaailman huutokauppavolyymin ja kampanjabudjetin aiheuttama skaalautuvuusongelma on hyvin hoidettu tila-arvon approksimoinnilla neuroverkkojen avulla. Empiirinen tutkimus kahdella laajalla reaalimaailman tietokokonaisuudella ja A/B-testi kaupallisella alustalla ovat osoittaneet, että se on ylivoimaisen tehokas ja tehokas verrattuna uusimpiin menetelmiin.

**Tulos**

Reaaliaikainen tarjousten tekeminen vahvistusoppimisen avulla display-mainonnassa

**Esimerkki 2.1125**

Tärkeä tapa muodostaa suuria harjoitusjoukkoja on kerätä meluisia merkintöjä joukoista, jotka eivät ole asiantuntijoita. Ehdotamme menetelmää, jolla voidaan koota yhteen työntekijä- tai kommentoijien joukosta kerättyjä häiriömerkintöjä. Merkintöjen kerääminen on tärkeää esimerkiksi verkkohakujen laadun arvioinnissa ja tuotteiden luokittelussa. Menetelmämme olettaa, että merkinnät tuotetaan todennäköisyysjakaumalla kohteiden ja merkintöjen välillä. Muotoilemme menetelmän rinnastamalla Gaussin sekoitusmallit (Gaussian Mixture Models, GMM) ja rajoitetut Boltzmannin koneet (Restricted Boltzmann Machines, RBM) toisiinsa ja osoitamme, että äänten yhdistämisongelmaa voidaan pitää klusterointiongelmana. Käytämme K-RBM-malleja klusterointiin. Lopuksi esitämme joitakin empiirisiä arviointeja todellisilla tietokokonaisuuksilla.

**Tulos**

Äänten yhdistäminen klusterointiongelmana

**Esimerkki 2.1126**

Tässä artikkelissa esitellään tehokas geneettinen algoritmi (GA) matkustavan myyntimiehen ongelman (TSP) ratkaisemiseksi. Tehokkaan GA:n rakentamiseksi käytän reunojen vaihtoa (ES) ja paikallista hakumenettelyä määrittääkseni hyviä yhdistelmiä vanhempien ratkaisujen rakennuspalikoista, jotta voidaan tuottaa vielä parempia jälkeläisratkaisuja. Kokeelliset tulokset hyvin tutkituilla TSP-vertailukohteilla osoittavat, että ehdotettu GA on kilpailukykyinen löytämään erittäin laadukkaita ratkaisuja jopa 16 862 kaupunkia sisältäviin tapauksiin.

**Tulos**

Tehokas geneettinen algoritmi matkustavan myyntimiehen ongelmaan

**Esimerkki 2.1127**

Koneoppimisen tekniikoita käytetään laajalti turvallisuuteen liittyvissä sovelluksissa, kuten roskapostin ja haittaohjelmien havaitsemisessa. Tällaisissa ympäristöissä niiden on kuitenkin osoitettu olevan alttiita vastahyökkäyksille, kuten tietojen tarkoitukselliselle manipuloinnille testaushetkellä havaitsemisen välttämiseksi. Tässä työssä keskitymme lineaaristen luokittelijoiden haavoittuvuuteen kiertohyökkäyksille. Tätä voidaan pitää merkityksellisenä ongelmana, koska lineaarisia luokittelijoita on käytetty yhä enemmän sulautetuissa järjestelmissä ja mobiililaitteissa niiden vähäisen prosessointiajan ja muistivaatimusten vuoksi. Hyödynnämme viimeaikaisia havaintoja robustista optimoinnista tutkiessamme regularisoinnin ja lineaaristen luokittelijoiden turvallisuuden välistä yhteyttä hyökkäystyypistä riippuen. Analysoimme myös suotavaa käsittelykustannusten vähentämiseksi toivottavaa ominaisuuksien painojen harvuutta ja lineaaristen luokittelijoiden turvallisuutta. Lisäksi ehdotamme uutta kahdeksankulmaista regularisaattoria, jonka avulla voimme saavuttaa sopivan kompromissin niiden välillä. Lopuksi osoitamme empiirisesti, miten tämä regularisaattori voi parantaa luokittelijan turvallisuutta ja harvennusta todellisissa sovelluksissa, kuten roskapostin ja haittaohjelmien havaitsemisessa.

**Tulos**

Lineaaristen luokittelijoiden turvallisuudesta ja harvinaisuudesta vastakkaisasetelmissa

**Esimerkki 2.1128**

Yksi nykyaikaisten puhutun dialogin järjestelmien keskeisistä komponenteista on uskomustenseurantajärjestelmä, joka arvioi käyttäjän tavoitteen dialogin jokaisessa vaiheessa. Useimmilla nykyisillä lähestymistavoilla on kuitenkin vaikeuksia skaalautua suuremmille ja monimutkaisemmille dialogialueille. Tämä johtuu siitä, että ne ovat riippuvaisia joko a) puhutun kielen ymmärtämismalleista, jotka vaativat suuria määriä kommentoitua harjoitusdataa, tai b) käsin tehdyistä leksikoista, joilla voidaan tallentaa osa käyttäjien kielen kielellisestä vaihtelusta. Ehdotamme uudenlaista neuraalista uskomusten seurantaa (Neural Belief Tracking, NBT), joka ratkaisee nämä ongelmat hyödyntämällä viimeaikaisia edistysaskeleita representaatio-oppimisessa. NBT-mallit päättelevät valmiiksi koulutettujen sanavektorien yli ja oppivat kokoamaan ne käyttäjien lausumien ja dialogin kontekstin hajautetuiksi representaatioiksi. Arviointimme kahdella tietokokonaisuudella osoittaa, että tämä lähestymistapa ylittää aiemmat rajoitukset, sillä se vastaa käsin laadittuihin semanttisiin leksikoihin tukeutuvien huippumallien suorituskykyä ja on parempi kuin ne, kun tällaisia leksikoita ei ole saatavilla.

**Tulos**

Neuraalinen uskomusten jäljittäjä: Dialogin tilan seuranta

**Esimerkki 2.1129**

Bayesin uskomusverkot (Bayesian Belief Networks, BBN) ovat tehokas formalismi epävarmuustilanteissa tapahtuvaan päättelyyn, mutta niillä on joitakin vakavia rajoituksia: ne vaativat suuren määrän tietoa, ennen kuin päättelyprosessi voi alkaa, niiden kyky käsitellä ristiriitoja on rajallinen, ja niiden kyky antaa selityksiä johtopäätöksilleen on edelleen kiistanalainen. On olemassa päättelyjärjestelmien luokka, joita kutsutaan 11-todellisuuden ylläpitojärjestelmiksi (11-uth Maintenance Systems, TMS) ja jotka pystyvät käsittelemään osittain spesifioitua tietoa, antamaan perusteltuja selityksiä päätelmilleen sekä havaitsemaan ja käsittelemään ristiriitoja. TMS-järjestelmiä, joihin sisältyy epävarmuuden mittaaminen, kutsutaan uskomusten ylläpitojärjestelmiksi (Belief Maintenance Systems, BMss). Tässä artikkelissa kuvataan, miten todennäköisyyslogiikkaan perustuvaa BMS-järjestelmää voidaan soveltaa BBN-verkkoihin, jolloin saadaan käyttöön uusi BBN-verkkojen luokka, jota kutsutaan Ignorant Be lief Networks -verkkoiksi ja joka pystyy käsittelemään inkrementaalisesti osittain spesifioituja ehdollisia riippuvuuksia, antamaan selityksiä sekä havaitsemaan ja käsittelemään ristiriitoja.

**Tulos**

Uskomusten ylläpito Bayes-verkoissa

**Esimerkki 2.1130**

Analysoimme luokan estimaattoreita, jotka perustuvat koveraan relaksaatioon korkea-ulotteisten matriisien purkamisongelmien ratkaisemiseksi. Havainnot ovat lineaarisen muunnoksen X kohinallisia realisaatioita, jotka ovat (likimain) matalarivisen matriisin Θ⋆ ja toisen matriisin Γ⋆ summan lineaarinen muunnos, joka on varustettu täydentävällä matala-ulotteisella rakenteella; tähän asetelmaan sisältyy monia kiinnostavia tilastollisia malleja, mukaan lukien faktorianalyysin muodot, jaetun rakenteen omaava monitehtäväinen regressio ja robusti kovarianssin estimointi. Johdamme yleisen teoreeman, joka antaa ylärajat Frobeniusnormin virheelle parin (Θ⋆,Γ⋆) estimaatille, joka saadaan ratkaisemalla kovera optimointiongelma, jossa yhdistetään ydinnormi ja yleinen hajoava regularisaattori. Tuloksemme perustuvat "piikkisyysehdon" asettamiseen, joka liittyy singulaarivektorien inkoherenssiin mutta on sitä lievempi. Erikoistamme yleisen tuloksemme kahteen tapaukseen, joita on tutkittu aiemmissa töissä: matala arvoaste ja sisäkkäinen harva matriisi sekä matala arvoaste ja sarakkeellinen harva matriisi. Molemmille malleille teoriamme tuottaa ei-oireelliset Frobeniuksen virherajat sekä deterministisille että stokastisille kohinamatriiseille, ja se soveltuu matriiseille Θ⋆, jotka voivat olla täsmälleen tai suunnilleen matalarivisiä, ja matriiseille Γ⋆, jotka voivat olla täsmälleen tai suunnilleen harvoja. Lisäksi stokastisten kohinamatriisien ja identiteettihavaintooperaattorin tapauksessa asetamme minimax-virheelle sopivat alarajat ja osoitamme, että tuloksiamme ei voi parantaa kuin vakiokertoimilla. Numeeriset simulaatiot vahvistavat teoreettisten ennusteidemme terävyyden.

**Tulos**

Kohinaisten matriisien hajottaminen koveran relaksaation avulla: Optimaaliset hinnat suurissa ulottuvuuksissa

**Esimerkki 2.1131**

Tutkimme memristoriverkon käyttöä musiikin tuottamiseen hyödyntämällä memristorin muistia Markovin hypoteesia pidemmälle. Luodaan ja täytetään memristoriyhtälöiden avulla siemensiirtymämatriiseja, joiden osoitetaan tuottavan musiikillisia melodioita ja muuttuvan tyyliltään ajan mittaan siirtymämatriisiin annettavan palautteen seurauksena. Yksinkertaisten memristoriverkkojen piikitysominaisuuksia demonstroidaan ja niistä keskustellaan musiikin tekemisen sovellusten yhteydessä. Keskustellaan rajoituksista, joita liittyy memristoriverkkojen koostamisen simulointiin von Neumann -laitteistossa, ja esitellään fysikaalisiin memristorin ominaisuuksiin perustuva laitteistoratkaisu.

**Tulos**

Markovin ketjujen ohi, kohti mukautuvaa Memristor-verkkopohjaista musiikin tuottamista

**Esimerkki 2.1132**

Ehdotamme uutta oppimismenetelmää heterogeenisen toimialueen mukauttamiseen (HDA), jossa lähde- ja kohdealueen tiedot esitetään heterogeenisilla ominaisuuksilla, joilla on eri ulottuvuudet. Käyttämällä kahta eri projektiomatriisia muunnamme ensin kahdesta alueesta peräisin olevat tiedot yhteiseen aliavaruuteen, jotta voimme mitata kahden alueen tietojen samankaltaisuutta. Sen jälkeen ehdotamme kahta uutta ominaisuuksien kartoitusfunktiota, joilla muunnettuun dataan lisätään niiden alkuperäiset ominaisuudet ja nollat. Olemassa olevat oppimismenetelmät (esim. SVM ja SVR) voidaan helposti yhdistää uusiin ehdotettuihin laajennettuihin ominaisuuksien esityksiin, jotta molempien alojen tietoja voidaan hyödyntää tehokkaasti HDA:ssa. Käyttämällä SVM:n saranahäviöfunktiota esimerkkinä esittelemme yksityiskohtaisen tavoitefunktion menetelmässämme nimeltä Heterogeneous Feature Augmentation (HFA) lineaarista tapausta varten ja kuvaamme myös sen kernelöinnin, jotta voimme tehokkaasti käsitellä hyvin suurten ulottuvuuksien tietoja. Lisäksi kehitämme myös vuorottelevan optimointialgoritmin, jolla voidaan ratkaista tehokkaasti HFA-menetelmämme ei-triviaali optimointiongelma. Kattavat kokeet kahdella vertailutietoaineistolla osoittavat selvästi, että HFA päihittää nykyiset HDA-menetelmät.

**Tulos**

Oppiminen laajennetuilla ominaisuuksilla heterogeenisen toimialueen mukauttamista varten

**Esimerkki 2.1133**

Tässä artikkelissa esitellään uusi tekniikka nimeltä "Structural Crossing-Over", jolla syntetisoidaan päteviä tietoja koneoppimiseen perustuvan käsialan tunnistuksen harjoittelua varten. Ehdotettu tekniikka voi tarjota suuremman valikoiman harjoitusdatan malleja kuin nykyiset lähestymistavat, kuten elastinen vääristymä ja tangenttipohjainen affiinimuunnos. Valitaan pari harjoitusmerkkiä, sitten analysoidaan niiden samankaltaiset ja erilaiset rakenteet ja lopuksi ristiintaulukointi uusien merkkien luomiseksi. Kokeissa verrataan tangenttipohjaisen affiinimuunnoksen ja ehdotetun lähestymistavan suorituskykyä tuotettujen merkkien valikoiman ja tunnistusvirheiden prosenttiosuuden suhteen. Kokeissa käytetään vakiomuotoista MNIST-korpusta, joka sisältää 60 000 harjoitusmerkkiä ja 10 000 testimerkkiä. Ehdotettu tekniikka käyttää 1 000 merkkiä 60 000 merkin syntetisoimiseksi ja käyttää sitten näitä tietoja kouluttaakseen ja testatakseen vertailukohdan käsialan tunnistusjärjestelmää, joka hyödyntää Histogram of Gradient -menetelmää: HOG:tä ominaisuuksina ja tukivektorikonetta: SVM:ää tunnistimena. Kokeellinen tulos tuottaa 8,06 prosenttia virheistä. Se on huomattavasti parempi kuin tangenttipohjainen affiinimuunnos ja alkuperäinen MNIST-koulutusdata, jotka ovat 11,74 % ja 16,55 %.

**Tulos**

AUTOMAATTINEN HARJOITUSAINEISTON SYNTEESI KÄSIALAN TUNNISTUSTA VARTEN KÄYTTÄEN RAKENTEELLISTA RISTIINKYTKENTÄTEKNIIKKAA (CROSSING OVER)

**Esimerkki 2.1134**

Seuraamusten ja ristiriitojen ymmärtäminen on olennaisen tärkeää luonnollisen kielen ymmärtämiselle, ja seurausten ja ristiriitojen päätteleminen on arvokas testialusta semanttisten representaatioiden kehittämiselle. Tämän alan koneoppimistutkimusta on kuitenkin rajoittanut dramaattisesti laajamittaisten resurssien puute. Tämän ongelman ratkaisemiseksi esittelemme Stanford Natural Language Inference -korpuksen, joka on uusi, vapaasti saatavilla oleva kokoelma merkittyjä lausepareja, jotka ihmiset ovat kirjoittaneet kuvatekstien kirjoittamiseen perustuvaa uutta perusteltua tehtävää suorittaessaan. Korpus on 570 000 paria, mikä on kaksi kertaluokkaa suurempi kuin kaikki muut sen tyyppiset resurssit. Tämän mittakaavan kasvun ansiosta leksikaalisoidut luokittelijat voivat päihittää joitakin nykyisiä kehittyneitä päättelymalleja ja neuroverkkopohjainen malli voi ensimmäistä kertaa suoriutua kilpailukykyisesti luonnollisen kielen päättelyn vertailuanalyyseissä.

**Tulos**

Suuri annotoitu korpus luonnollisen kielen päättelyn oppimista varten

**Esimerkki 2.1135**

Esittelemme yleiskäyttöisen, syväoppimiseen perustuvan järjestelmän, joka purkaa kuvatekstien tuottamisen ja optisen merkintätunnistuksen (OCR) viimeaikaisiin edistysaskeliin perustuen kuvan esitystapamerkinnöiksi. Vaikka tämä tehtävä on hyvin tutkittu OCR-ongelma, menetelmämme lähestymistapa on luonteeltaan erilainen ja tietoon perustuva. Mallimme ei vaadi mitään tietoa taustalla olevasta merkintäkielestä, ja se yksinkertaisesti koulutetaan alusta loppuun todellisella esimerkkidatalla. Mallissa käytetään konvoluutioverkkoa tekstin ja ulkoasun tunnistamiseen yhdessä huomiopohjaisen neuraalisen konekäännösjärjestelmän kanssa. Mallin kouluttamiseksi ja arvioimiseksi otamme käyttöön uuden tietokokonaisuuden, joka koostuu reaalimaailman renderöidyistä matemaattisista lausekkeista ja LaTeX-merkinnöistä sekä synteettisestä tietokokonaisuudesta, joka koostuu verkkosivuista ja HTML-katkelmista. Kokeelliset tulokset osoittavat, että järjestelmä on yllättävän tehokas tuottamaan tarkkaa merkintää molemmissa aineistoissa. Kun tavallinen LaTeX OCR-järjestelmä saavuttaa noin 25 prosentin tarkkuuden, meidän mallimme jäljentää täsmällisesti renderöidyn kuvan 75 prosentissa esimerkeistä.

**Tulos**

Se mitä saat, on se mitä näet: Visual Markup Decompiler

**Esimerkki 2.1136**

Tässä artikkelissa laajennamme neuraalisen Turingin koneen (NTM) dynaamiseksi neuraaliseksi Turingin koneeksi (D-NTM) ottamalla käyttöön harjoiteltavan muistin osoitusjärjestelmän. Tämä järjestelmä ylläpitää kullekin muistisolulle kahta erillistä vektoria, sisältö- ja osoitevektoreita. Tämän ansiosta D-NTM voi oppia monenlaisia sijaintiin perustuvia osoitusstrategioita, mukaan lukien lineaariset ja epälineaariset strategiat. Toteutamme D-NTM:n sekä pehmeillä, differentioituvilla että kovilla, ei-differentioituvilla luku- ja kirjoitusmekanismeilla. Tutkimme mekanismeja ja vaikutuksia muistin lukemisen ja kirjoittamisen oppimiseen kokeilemalla Facebookin bAbI-tehtäviä käyttämällä sekä feedforward- että GRU-ohjainta. D-NTM:ää arvioidaan joukossa Facebookin bAbI-tehtäviä, ja sen osoitetaan päihittävän NTM- ja LSTM-perusasetukset.

**Tulos**

Dynaaminen neuraalinen Turingin kone, jossa on pehmeä ja kova osoitejärjestelmä

**Esimerkki 2.1137**

Tekoälyyhteisö on kiinnostunut suuresti siitä, että tietopohjasta erotetaan vain osa, joka on merkityksellinen aakkosten osajoukolle, joka tunnustetaan unohtamiseksi. Tavallisesta propositionaalisesta logiikasta puuttuu edelleen yleinen algoritmi unohtamiselle ja sen laskentaan suuntautuva tutkiminen erilaisissa fragmenteissa, joiden tyydyttävyys on saavutettavissa. Tämän artikkelin tavoitteena on täyttää tämä aukko. Tutkittuamme eräitä propositiologiikan unohtamisen perusominaisuuksia esitämme ratkaisupohjaisen unohtamisalgoritmin CNF-fragmenttia varten sekä eräitä kompleksisuustuloksia unohtamisesta propositiologiikan Horn-, uudelleennimettävissä Horn-, q-Horn-, Krom-, DNF- ja CNF-fragmenteissa.

**Tulos**

Unohtamisesta jäljitettävissä olevissa propositionaalisissa fragmenteissa

**Esimerkki 2.1138**

Perinteiset stokastisen optimoinnin algoritmit edellyttävät ratkaisun projisoimista jokaisella iteraatiokerralla tietylle alueelle sen toteutettavuuden varmistamiseksi. Kun vastassa on monimutkaisia alueita, kuten positiiviset puolimääritellyt kartiot, projisointioperaatio voi olla kallis, mikä johtaa korkeisiin laskentakustannuksiin iteraatiota kohti. Tässä artikkelissa esitellään uusi algoritmi, jolla pyritään vähentämään projisointien määrää stokastisessa optimoinnissa. Ehdotetussa algoritmissa yhdistyvät useiden stokastisen optimoinnin viimeaikaisten kehityskulkujen vahvuudet, mukaan lukien mini-batch-, extra-gradientti- ja epochegradienttilaskeutuminen, jotta voidaan tehokkaasti tutkia tasaisuutta ja voimakasta koveruutta. Osoitamme sekä odotusarvoisesti että suurella todennäköisyydellä, että kun kohdefunktio on sekä sileä että vahvasti kupera, ehdotettu algoritmi saavuttaa optimaalisen O(1/T ):n konvergenssinopeuden vain O(log T ):n projektioilla. Empiirinen tutkimuksemme vahvistaa teoreettisen tuloksen.

**Tulos**

O(logT ) projisoinnit sileiden ja voimakkaasti kuperaisten funktioiden stokastiseen optimointiin

**Esimerkki 2.1139**

Matriisin sijan minimointi affiinisten rajoitusten alaisena on perustavanlaatuinen ongelma, jolla on monia tärkeitä sovelluksia koneoppimisessa ja tilastotieteessä. Tässä artikkelissa ehdotamme yksinkertaista ja nopeaa algoritmia SVP (Singular Value Projection) rangin minimointiin affiinisilla rajoitteilla (ARMP) ja osoitamme, että SVP palauttaa pienimmän rangin ratkaisun affiinisille rajoitteille, jotka täyttävät rajoitetun isometrian ominaisuuden. Osoitamme menetelmämme kestävyyden kohinaa vastaan, ja geometrinen konvergenssinopeus on vahva myös kohinaisilla mittauksilla. Tuloksemme parantavat Rechtin, Fazelin ja Parillon [RFP07] äskettäistä läpimurtoa kolmella merkittävällä tavalla: 1) menetelmämme (SVP) on huomattavasti yksinkertaisempi analysoida ja helpompi toteuttaa, 2) annamme geometriset konvergenssitakuut SVP:lle ja, kuten empiirisesti osoitetaan, SVP on huomattavasti nopeampi todellisissa ja synteettisissä ongelmissa, 3) annamme optimaalisuus- ja geometriset konvergenssitakuut jopa ARMP:n kohinaiselle versiolle. Lisäksi käsittelemme käytännössä tärkeää matalarivisen matriisin täydentämisen ongelmaa, jota voidaan pitää ARMP:n erikoistapauksena. Matriisien täydentämisongelman määrittelevät affiiniset rajoitteet eivät kuitenkaan yleensä noudata rajoitetun isometrian ominaisuutta. Osoitamme empiirisesti, että algoritmimme palauttaa matalariviset epäkoherentit matriisit lähes optimaalisen määrän tasaisesti otettuja merkintöjä. Edistymme osittain tarkan palautuksen todistamisessa ja annamme jonkin verran intuitiota SVP:n suorituskyvylle sovellettuna matriisien täydentämiseen osoittamalla rajoitetumman isometriaominaisuuden. Algoritmimme on ARMP:n ja matriisien täydennysongelman osalta kertaluokkaa parempi kuin olemassa olevat menetelmät, kuten [RFP07, CR08, CT09, CCS08, KOM09], ja se on myös huomattavasti kestävämpi kohinan suhteen.

**Tulos**

Taattu sijan minimointi singulaariarvoprojektiolla (Singular Value Projection)

**Esimerkki 2.1140**

Automaattisesti tuotettu poliittinen tapahtumatieto on tärkeä osa yhteiskuntatieteiden datan ekosysteemiä. Lähestymistavat näiden tietojen tuottamiseksi ovat kuitenkin pysyneet pitkälti samoina kahden vuosikymmenen ajan. Tänä aikana laskennallisen kielitieteen ala on kehittynyt valtavasti. Tässä artikkelissa esitetään yleiskatsaus poliittiseen tapahtumadataan, mukaan lukien menetelmät ja ontologiat, sekä joukko kokeita, joilla selvitetään syvien neuroverkkojen soveltuvuutta poliittisten tapahtumien uutistekstistä poimimiseen.

**Tulos**

Poliittisesti merkityksellisten tapahtumatietojen tuottaminen

**Esimerkki 2.1141**

Viime vuosina syväoppimisesta on tullut ensisijainen ratkaisu monenlaisiin sovelluksiin, ja se on usein parempi kuin uusin tekniikka. Sekä teoreetikoiden että käytännön toimijoiden on kuitenkin tärkeää ymmärtää syvällisemmin yleisiin lähestymistapoihin ja algoritmeihin liittyviä vaikeuksia ja rajoituksia. Kuvaamme neljä ongelmaperhettä, joissa jotkin yleisesti käytetyistä nykyisistä algoritmeista epäonnistuvat tai kärsivät merkittävistä vaikeuksista. Havainnollistamme epäonnistumisia käytännön kokeilujen avulla ja tarjoamme teoreettisia näkemyksiä, joissa selitetään, mistä ne johtuvat ja miten ne voitaisiin korjata.

**Tulos**

Syväoppimisen epäonnistumiset

**Esimerkki 2.1142**

Jotta vuorovaikutteiset tarinajärjestelmät voivat kertoa tarinoita eri äänillä eri yleisöille, ne edellyttävät: (1) semanttisen esityksen tarinan rakenteesta ja (2) kyvyn luoda automaattisesti tarina ja dialogi tästä semanttisesta esityksestä jonkinlaisen luonnollisen kielen tuottamisen (nlg) avulla. On kuitenkin tutkittu vain vähän menetelmiä, joilla tarinan rakenteet voidaan yhdistää kohtausten ja tarinan tapahtumien kerronnallisiin kuvauksiin. Tässä artikkelissa esittelemme automaattisen menetelmän, jolla Scheherazaden tarinan aikomusgrafi, semanttinen esitys, muunnetaan personage nlg -moottorin tarvitsemaksi syötteeksi. Käyttämällä 36:aa Äsopin tarinaa, jotka on jaettu DramaBankiin, joka on kokoelma tarinakoodauksia, koulutamme käännössäännöt yhdelle tarinalle ja sitten testaamme näitä sääntöjä tuottamalla tekstiä lopuille 35:lle tarinalle. Tuloksia mitataan merkkijonojen samankaltaisuusmittareilla Levenshteinin etäisyys ja BLEU-pisteet. Tulokset osoittavat, että pystymme tuottamaan 35 tarinasta sisällöltään oikeat: testisarjan tarinat ovat keskimäärin lähellä Scheherazade-realisaattorin tulosta, joka on räätälöity tähän semanttiseen esitykseen. Esitämme joitakin esimerkkejä personagen tuottamista tarinavariaatioista. Tulevassa työssä kokeilemme eri äänillä tuotettujen samojen tarinoiden laadun mittaamista ja tekniikoita, joilla tarinankerronta voidaan tehdä vuorovaikutteiseksi.

**Tulos**

Erilaisten tarinakertomusten tuottaminen kertomuksen semanttisista esityksistä

**Esimerkki 2.1143**

Tässä artikkelissa esitetään teoreettinen selitys ei-negatiivisen matriisifaktoroinnin (NMF) klusterointinäkökulmasta. Todistamme, että vaikka NMF ei asettaisi ortogonaalisuutta eikä harvinaisuusrajoitusta perusmatriisille ja/tai kertoimimatriisille, NMF voi silti antaa klusterointituloksia, mikä antaa teoreettista tukea monille teoksille, esim. Xu et al. [1] ja Kim et al. [2], jotka osoittavat tavallisen NMF:n paremmuuden klusterointimenetelmänä. Avainsanat-rajoitteinen optimointi, klusterointimenetelmä, ei-konveksaalinen optimointi, ei-negatiivinen matriisifaktorointi, klusterointimenetelmä

**Tulos**

Nonnegatiivisen matriisifaktoroinnin klusterointinäkökulmasta

**Esimerkki 2.1144**

Yksilöllisten tunnisteiden antaminen n muuttujalle on tunnettu uudelleennimeämisongelma, johon on ehdotettu useita algoritmeja hajautettuja algoritmeja käsittelevässä kirjallisuudessa. Tietojemme mukaan kaikissa näissä algoritmeissa keskitytään kuitenkin vikasietoisuuteen ja jätetään huomiotta yksityisyyden suoja. Päinvastoin, tässä asiakirjassa emme ota huomioon agenttien vikoja, vaan tarvitsemme algoritmin, joka suojaa agenttien ja topologian yksityisyyden. Tätä tarkoitusta varten ehdotamme algoritmia 1, joka on muunnos verkkoliitteessä 2 esitetystä pseudopuun muodostamisalgoritmista ja parannettu versio Léautén ja Faltingsin (2009) ehdottamasta algoritmista. Kullekin muuttujalle x annetaan yksilöllinen numero idx, joka vastaa järjestystä, jossa se käydään ensimmäisenä rajoitteiden graafin hajautetun läpikäynnin aikana (tai tarkemmin sanottuna sen ylärajaa). Tämä tehdään liittämällä jokaiseen CHILD-viestiin toistaiseksi käytyjen muuttujien numero id (rivit 8, 29 ja 31). Kukin muuttuja lisää id:hen satunnaisluvun, jotta sen naapureiden lukumäärästä ei vuotaisi mitään hyödyllistä ylärajaa (rivit 5 ja 15). Tämän algoritmin lopussa juurimuuttuja löytää muuttujien kokonaismäärää koskevan ylärajan n+ ja paljastaa sen kaikille (rivit 35 ja 22-24).

**Tulos**

Yksityisyyden suojaaminen hajautetun laskennan avulla moniagenttipäätöksiä tehtäessä verkossa Liite 3: Yksilöllisen tunnisteen luomisalgoritmi

**Esimerkki 2.1145**

Keinotekoinen objektin havaitseminen perustuu yleensä ennalta määriteltyihin malleihin ja ominaisuuksien poiminta-algoritmeihin. Tutkimme, miten objektin käsite voidaan perustaa naiivin agentin sensomotoriseen kokemukseen. Ilman mitään tietoa itsestään tai maailmasta, johon se on uppoutunut, agentti tutkii sensomotorista avaruuttaan ja tunnistaa objektit kontekstista riippumattomina sensomotoristen siirtymien johdonmukaisina verkostoina. Tällaisten verkostojen syntymisen selittämiseksi kehitysnäkökulmasta oletetaan olevan perustavanlaatuinen ennustamisen pyrkimys. Lähestymistavan havainnollistamiseksi ehdotetaan ja testataan algoritmia.

**Tulos**

Esineiden havaitsemisen perustaminen naiivin agentin sensomotoriseen kokemukseen.

**Esimerkki 2.1146**

Tässä artikkelissa tutkitaan painotettujen automaattien oppimisen ongelmaa rajallisesta merkitystä harjoitusnäytteestä. Tarkastelemme useita yleisiä painotettujen automaattien perheitä, jotka on määritelty kolmen eri mittarin avulla: automaatin painojen normi, automaatin laskeman funktion normi tai vastaavan Hankelin matriisin normi. Esitämme uusia datasta riippuvia yleistystakeita painotettujen automaattien oppimiselle ilmaistuna näiden perheiden Rademacher-kompleksisuutena. Esitämme lisäksi näiden Rademacher-kompleksisuuksien ylärajat, jotka paljastavat keskeisiä uusia datasta riippuvia termejä, jotka liittyvät painotettujen automaattien oppimisen kompleksisuuteen.

**Tulos**

Painotettujen automaattien yleistysrajat

**Esimerkki 2.1147**

Ehdotamme uutta menetelmää, jolla voidaan soveltaa ennakkoarvioita ei-negatiivisen matriisifaktoroinnin (NMF) ratkaisuun. Ehdotettua algoritmia voidaan käyttää äänenvaimennukseen tai yhden kanavan lähteiden erotteluun (SCSS). NMF-ratkaisu ohjataan noudattamaan MMSE-estimaatteja (Minimum Mean Square Error) Gaussin seosmallien (GMM) perusteella lähdesignaalin osalta. SCSS-sovelluksissa havaitun sekasignaalin spektrit puretaan NMF:n avulla kunkin lähteen koulutettujen perusvektoreiden painotettuna lineaarisena yhdistelmänä. Tässä työssä NMF-dekomposition painomatriiseja käsitellään vääristyneenä kuvana vääristymäoperaattorilla, joka opitaan suoraan havaituista signaaleista. Tämän jälkeen löydetään painomatriisin MMSE-estimaatti GMM-prepriorilla ja log-normijakaumalla vääristymälle NMF-dekomposition tulosten parantamiseksi. MMSE-estimaatti sisällytetään optimointitavoitteeseen, jolloin muodostuu uusi regularisoitu NMF-kustannusfunktio. Uusien tavoitteiden vastaavat päivityssäännöt johdetaan tässä asiakirjassa. Kokeelliset tulokset osoittavat, että ehdotettu regularisoitu NMF alPreprint submitted to Elsevier March 1, 2013 ar X iv :1 30 2. 72 83 v1 [ cs .L G ] 2 8 Fe b 20 13 goritmi parantaa lähteiden erottelutehoa verrattuna NMF:n käyttämiseen ilman ennakkomallia tai muiden ennakkomallien kanssa.

**Tulos**

Lähteiden erottelu käyttämällä säännöllistettyä NMF:ää ja MMSE-estimaatteja GMM-prioreilla, joissa on online-oppiminen epävarmuustekijöiden varalta.

**Esimerkki 2.1148**

Kuvaamme strategian, jolla hankitaan harjoitustietoja, joita tarvitaan sosiaaliseen mediaan perustuvan varhaishavainnointijärjestelmän rakentamiseksi henkilöille, joilla on riski sairastua (ehkäistävissä olevaan) tyypin 2 diabetes mellitukseen (T2DM). Strategiassa käytetään pelimäistä tietovisaa, jonka tiedot ja kysymykset hankitaan puoliautomaattisesti Twitteristä. Kysymykset on suunniteltu innostamaan osallistujia sitoutumaan ja keräämään asiaankuuluvia tietoja, joiden avulla voidaan kouluttaa yksilöihin sovellettavaa kansanterveysmallia. Aiemmat järjestelmät, jotka on suunniteltu käyttämään sosiaalista mediaa, kuten Twitteriä, liikalihavuuden (T2DM:n riskitekijä) ennustamiseen, toimivat kokonaisissa yhteisöissä, kuten osavaltioissa, maakunnissa tai kaupungeissa, jotka perustuvat valtion virastojen keräämiin tilastoihin. Koska yksilöiden välillä on huomattavaa vaihtelua näiden ryhmien sisällä, yksilötason koulutustiedot olisivat tehokkaampia, mutta näitä tietoja on vaikea hankkia. Tässä ehdotetulla lähestymistavalla pyritään ratkaisemaan tämä ongelma. Strategiassamme on kaksi vaihetta. Ensin koulutimme satunnaismetsäluokittimen (julkisista) Twitter-statuksista ja osavaltiotason tilastoista kerätyillä tiedoilla, joiden tarkkuus oli huippuluokkaa. Sen jälkeen muunnimme tämän luokittimen 20 kysymystä sisältäväksi tietokilpailuksi ja asetimme sen saataville verkossa. Näin saimme aikaan suuren sitoutumisen tietovisaan osallistuneiden henkilöiden kanssa ja samalla muodostimme vapaaehtoisesti toimitetuista yksilötason tiedoista harjoitusjoukon tulevaa luokittelua varten.

**Tulos**

Kohti sosiaalisen median käyttöä ehkäistävissä olevien kroonisten sairauksien riskiryhmiin kuuluvien yksilöiden tunnistamisessa

**Esimerkki 2.1149**

RoboCup 2D Simulation League sisältää useita haastavia ominaisuuksia, jotka asettavat tekoälyn (AI) mittapuuksi. Tässä artikkelissa kuvaamme joitakin ideoita ja työkaluja, jotka liittyvät Gliders2012-joukkueemme kehittämiseen. Kuvauksessamme keskitymme arviointitoimintoon, joka on yksi keskeisistä toimintojen valintamekanismeista. Esittelemme myös uuden kehyksen lokitiedostojen katseluun verkkoselaimella, jonka julkaisemme RoboCup-yhteisön käyttöön ja jatkokehitykseen. Lopuksi teemme myös yhteenvedon RoboCup 2012:n aikana pelaamiemme ryhmä- ja loppuotteluiden tuloksista, joissa Gliders2012 sijoittui neljänneksi 19 joukkueen joukosta.

**Tulos**

Purjelentokoneet2012: Kehitys ja kilpailutulokset

**Esimerkki 2.1150**

Esittelemme ensimmäisen yleiskäyttöisen viipalevalintaisen päättelymoottorin todennäköisyyspohjaisille ohjelmille. Tämä moottori julkaistaan osana StocPy:tä, uutta Turing-Complete-todennäköisyysohjelmointikieltä, joka on saatavilla Python-kirjastona. Esitämme siivunäytteenoton transdimensionaalisen yleistyksen, joka on välttämätön, jotta päättelymoottori voi työskennellä jälkien kanssa, joissa on eri määrä satunnaismuuttujia. Osoitamme, että StocPy on joustavuudeltaan ja käytettävyydeltään parempi kuin muut PPL-kielet ja että slice sampling voi päihittää aiemmin esitetyt päättelymenetelmät. Kokeemme sisältävät logistisen regression, HMM:n ja Bayesin neuroverkon.

**Tulos**

Viipaleen näytteenotto todennäköisyysohjelmoinnissa (Slice Sampling for Probabilistic Programming)

**Esimerkki 2.1151**

Tässä artikkelissa käsitellään asiakirjojen ajallista erottelua, kuten sen määrittämistä, milloin asiakirja on laadittu tai milloin se on kirjoitettu, pelkästään sen tekstin perusteella. Sovellamme tiedonhaun tekniikoita, jotka ennustavat päivämääriä kielimallien avulla diskreetillä aikajanalla. Toisin kuin useimmissa aiemmissa töissä, me luotamme ainoastaan tekstin sisältämiin ajallisiin vihjeisiin. Tarkastelemme sekä asiakirjan todennäköisyyteen että eroavaisuuteen perustuvia tekniikoita ja useita tasoitusmenetelmiä molempia varten. Paras mallimme ennustaa yksilöiden elämän puolivälin mediaanilla 22 ja keskivirheellä 36 vuotta Wikipedian elämäkerroissa vuodesta 3800 eaa. nykypäivään. Osoitamme myös, että tämä lähestymistapa toimii hyvin, kun sitä harjoitellaan tällaisilla elämäkerroilla ja ennustetaan päivämääriä sekä tiettyjä vuosia käsitteleville Wikipedian muille kuin elämäkerrallisille sivuille (500 eaa. - 2010 jKr.) että novellien julkaisupäiville (1798 - 2008). Yhteenlaskettuna työmme osoittaa, että myös ilman ajallisia louhintaresursseja on mahdollista saavuttaa huomattava ajallinen paikallisuus erilaisissa teksteissä.

**Tulos**

Treffitekstit ilman eksplisiittisiä ajallisia vihjeitä

**Esimerkki 2.1152**

Vahvistusoppimisessa tavoitteet määritellään usein määrittelemällä palkkiot halutuissa tiloissa. Tämän lähestymistavan ongelmana on, että palkkiot on yleensä määriteltävä uudelleen aina, kun tavoite muuttuu, mikä edellyttää usein jonkinlaista ymmärrystä agentin ympäristön ratkaisusta. Kun ihmiset oppivat suorittamaan tehtäviä, käytämme säännöllisesti vaihtoehtoisia lähteitä, jotka ohjaavat ongelman ymmärtämistä. Tällaiset tehtäväesitykset mahdollistavat tavoitteiden määrittelyn niiden omilla ehdoilla, jolloin saadaan määrittelyjä, joita voidaan tulkita asianmukaisesti eri ympäristöissä. Tämä motivoi omaa työtämme, jossa representoimme tavoitteita ympäristöissä, jotka poikkeavat agentin ympäristöstä. Otamme käyttöön CDPR (Cross-Domain Perceptual Reward) -funktiot, opitut palkkiot, jotka edustavat agentin tilan ja toimialueiden välisen tavoitekuvan visuaalista samankaltaisuutta. Raportoimme tuloksia CDPR-funktioiden oppimisesta syvällä neuroverkolla ja niiden käyttämisestä kahden tehtävän ratkaisemiseen syvällä vahvistusoppimisella.

**Tulos**

Monialaiset havaintopalkitsemistoiminnot (Cross-Domain Perceptual Reward Functions)

**Esimerkki 2.1153**

Esittelemme uuden paikkatietorakenteen korkea-ulotteisille tiedoille, jonka nimi on likimääräinen pääsuuntapuu (APD-puu) ja joka mukautuu tietojen luontaiseen ulottuvuuteen. Algoritmimme takaa vektorikvantisointitarkkuuden, joka vastaa laskennallisesti kalliiden PCA-puiden tarkkuutta, ja sen aikakompleksisuus on samankaltainen kuin epätarkempien RP-puiden. APD-puut käyttävät pientä määrää powermethod-iteraatioita löytääkseen jakotasot datan rekursiivista jakamista varten. Näin ollen ne tarjoavat luonnollisen kompromissin RP- ja PCA-puiden saavuttaman suoritusajan ja tarkkuuden välillä. Teoreettiset tuloksemme osoittavat a) vahvat suorituskykytakeet riippumatta powermethodin konvergenssinopeudesta ja b) että O(log d)-iteraatiot riittävät PCA-puiden takuuseen, kun sisäinen ulottuvuus on d. Osoitamme tämän kompromissin ja tietorakenteemme tehokkuuden sekä CPU:lla että GPU:lla.

**Tulos**

Likimääräinen pääsuunta Puut

**Esimerkki 2.1154**

Koneiden saavuttama edistys kielen ja kuvien ymmärtämisessä on herättänyt tutkimusyhteisön kiinnostuksen avoimempiin, kokonaisvaltaisempiin tehtäviin ja antanut uutta puhtia vanhaan tekoälyn unelmaan älykkäiden koneiden rakentamisesta. Keskustelemme muutamista merkittävistä haasteista, jotka luonnehtivat tällaisia kokonaisvaltaisia tehtäviä, ja esitämme, että "kuvakysymyksiin vastaaminen" on erityisen houkutteleva esimerkki tällaisesta kokonaisvaltaisesta tehtävästä. Huomautamme erityisesti, että kyseessä on Turingin testin versio, joka on todennäköisesti kestävämpi ylitulkinnoille, ja asetamme sen vastakkain sellaisten tehtävien kanssa kuin maadoitus ja kuvausten tuottaminen. Lopuksi keskustelemme välineistä, joilla voidaan mitata edistystä tällä alalla.

**Tulos**

Vaikea huijata: Turingin testi, joka perustuu kuvia koskeviin kysymyksiin vastaamiseen.

**Esimerkki 2.1155**

Käyttäjätietojen kerääminen ja analysointi parantaa sovellus- ja verkkoekosysteemejä, mutta siihen liittyy myös yksityisyyden suojaan liittyviä riskejä. Tässä artikkelissa tarkastellaan diskreetin jakauman estimointia paikallisen yksityisyyden suojan puitteissa, jolloin palveluntarjoajat voivat oppia kiinnostavan kategorisen tilaston jakauman keräämättä sen taustalla olevia tietoja. Esittelemme uusia mekanismeja, kuten hashed k-ary Randomized Response (k-RR), jotka empiirisesti täyttävät tai ylittävät nykyisten mekanismien hyötysuhteen kaikilla yksityisyydensuojan tasoilla. Uudet teoreettiset tulokset osoittavat k-RR:n ja nykyisen RAPPOR-mekanismin järjestysoptimaalisuuden eri yksityisyysasteilla.

**Tulos**

Diskreetin jakauman estimointi paikallisen yksityisyyden suojan puitteissa

**Esimerkki 2.1156**

Yksi tärkeä haaste, jonka avulla agenttijoukko voi tehostaa yhteistyötään, on se, että agentit ylläpitävät asianmukaisia malleja toisistaan. Tärkeä näkökohta näissä toisia agentteja koskevissa malleissa on se, että ne ovat usein osittaisia ja epätäydellisiä. Toistaiseksi agenttimalleista on olemassa kaksi yleistä esitystapaa: MDP-pohjaiset ja toimintapohjaiset, jotka molemmat perustuvat toiminnan mallintamiseen. Monissa sovelluksissa agenttimalleja ei välttämättä ole annettu, joten ne on opittava. Vaikka voi tuntua kätevältä käyttää oppimiseen joko MDP-pohjaisia tai toimintapohjaisia malleja, tässä artikkelissa esitellään uusi kyvykkyysmalleihin perustuva esitys, jolla on useita ainutlaatuisia etuja. Ensinnäkin osoitamme, että kyvykkyysmallien oppiminen voidaan suorittaa tehokkaasti verkossa Bayesin oppimisen avulla, ja oppimisprosessi on kestävä suunnitelman suoritusjäljen suurelle epätäydellisyydelle (esim. kun on vain alku- ja lopputilat). Vaikka suunnitelman suoritusjäljen suuri epätäydellisyys aiheuttaa haasteita MDP-pohjaisten ja toimintapohjaisten mallien oppimiselle, kyvykkyysmallit voivat silti oppia abstrahoimaan hyödyllistä tietoa näistä jäljistä. Tämän vuoksi kyvykkyysmallit ovat hyödyllisiä sovelluksissa, joissa tällainen epätäydellisyys on yleistä, esimerkiksi robotin oppiessa ihmismallia havainnoista ja vuorovaikutuksesta. Lisäksi kun niitä käytetään usean agentin suunnittelussa (jossa jokainen agentti mallinnetaan erikseen), kyvykkyysmallit mahdollistavat joustavan toimintojen abstrahoinnin. Rajoituksena on kuitenkin se, että syntetisoitu suunnitelma on epätäydellinen ja abstrakti.

**Tulos**

Agenttien kyvykkyysmallien oppiminen ja sovellukset moniagenttisuunnittelussa

**Esimerkki 2.1157**

Esittelemme vaihtoehtoisen menetelmän algoritmien analysoimiseksi, joka perustuu odotetun diskontatun palkkion käsitteeseen. Tämä menetelmä käsittelee luonnollisesti algoritmeja, jotka eivät aina pääty, joten sitä voidaan (teoreettisesti) käyttää ratkaisemattomien ongelmien osittaisalgoritmeihin, kuten esimerkiksi keinotekoisessa yleisessä älykkyydessä (AGI) ja automaattisessa teoreemantarkastuksessa. Mainitsemme lähestymistavan itseparantuvaan AGI:hen, jonka tämä menetelmä mahdollistaa.

**Tulos**

Algoritmien ja osittaisalgoritmien analyysi

**Esimerkki 2.1158**

Tässä artikkelissa ehdotamme menetelmää, jossa käytetään puolivalvottuja konvoluutio-neuraaliverkkoja (CNN) valitsemaan harjoitteluaineistoa tilastollista konekääntämistä varten. Tämä lähestymistapa on erityisen tehokas silloin, kun käytettävissä on vain pieniä määriä verkkotunnuksen sisäistä dataa. Sisäisen datan ja satunnaisesti poimitun yleisen datan avulla koulutetaan datanvalintamalli puolivalvotulla CNN:llä, minkä jälkeen tämä malli laskee aluerelevanssipisteet kaikille yleisen datan lauseille. Järjestelmän kouluttamista varten valitaan lauseparit, joilla on parhaat pisteet. Teemme kokeita neljällä kielisuunnalla ja kolmella testialueella. Verrattuna vahvoihin perusjärjestelmiin, jotka on koulutettu suurella tietomäärällä, tämä menetelmä voi parantaa suorituskykyä jopa 3,1 BLEU:lla. Sen suorituskyky on huomattavasti parempi kuin kolmen uusimman kielimalliin perustuvan datanvalintamenetelmän. Osoitamme myös, että valintamallin kouluttamiseen käytettävä aineisto voi olla vain 100 lausetta, mikä mahdollistaa hienojakoisen, aihepiiristä riippuvaisen käännösten mukauttamisen.

**Tulos**

Puolivalvotut konvoluutioverkot käännöksen mukauttamiseen pienellä määrällä sisäistä dataa

**Esimerkki 2.1159**

Useimpien strukturoitujen lähtöennustajien on ratkaistava esikuvaongelma päättelyn aikana. Merkkijonoytimissä tämä ongelma vastaa tiettyyn syötteeseen liittyvän merkkijonon löytämistä. Algoritmilla, joka pystyy ratkaisemaan tämän ongelman tai löytämään sille hyvän approksimaation, olisi monia sovelluksia laskennallisessa biologiassa ja muilla aloilla. Tässä työssä käytetään äskettäin saatua tulosta merkkijonoytimiin perustuvien lineaaristen ennustimien kombinatorisesta optimoinnista, jotta esikuvaukselle voidaan kehittää matalan kompleksisuuden yläraja, joka on voimassa monille merkkijonoytimille. Tätä ylärajaa käytetään menestyksekkäästi haara- ja rajahakualgoritmissa. Sovelluksia ja tuloksia lääkkeeksi soveltuvien peptidien löytämisessä esitellään ja niistä keskustellaan.

**Tulos**

String Kernel Pre-Image -ongelmasta ja sen sovelluksista lääkkeiden löytämisessä

**Esimerkki 2.1160**

Tarkastelemme aktiivisen verkko-oppimisen ongelmaa tietojen keräämiseksi regressiomallinnusta varten. Tarkemmin sanottuna tarkastelemme päätöksentekijää, jolla on rajallinen kokeilubudjetti ja jonka on tehokkaasti opittava taustalla oleva lineaarinen populaatiomalli. Tärkein panoksemme on uusi kynnysarvoihin perustuva algoritmi informatiivisimpien havaintojen valitsemiseksi; kuvaamme sen suorituskykyä ja perustavanlaatuisia alarajoja. Laajennamme algoritmin ja sen takuut koskemaan harvaa lineaarista regressiota korkea-ulotteisissa ympäristöissä. Simulaatiot osoittavat, että algoritmi on huomattavan vankka: se tarjoaa merkittäviä etuja passiiviseen satunnaisotantaan verrattuna todellisissa tietokokonaisuuksissa, joissa on paljon epälineaarisuutta ja suurta ulottuvuutta - se vähentää merkittävästi sekä neliövirheen keskiarvoa että varianssia.

**Tulos**

Aktiivinen lineaarinen regressio verkossa kynnysarvon avulla

**Esimerkki 2.1161**

Amerikkalaiset viettävät noin kolmanneksen ajastaan verkossa, ja monet heistä osallistuvat yhteiskunnallisista ja poliittisista kysymyksistä käytäviin verkkokeskusteluihin. Oletamme, että sosiaalisessa mediassa esitetyt argumentit tällaisista kysymyksistä voivat olla kiinnostavampia ja vakuuttavampia kuin perinteiset tiivistelmät ja että tietyntyyppiset ihmiset voivat olla enemmän tai vähemmän vakuuttuneita tietyistä argumenttityyleistä, esimerkiksi tunneperäiset argumentit voivat olla joidenkin persoonallisuuksien mieleen, kun taas asialliset argumentit ovat toisten mieleen. Raportoimme kokeista, joissa testataan laajassa mittakaavassa, miten yleisömuuttujat ja argumenttityyli vaikuttavat vuorovaikutuksessa argumentin vakuuttavuuteen, mikä on luonnollisen kielen prosessoinnissa vähän tutkittu aihe. Osoitamme, että persoonallisuustekijät vaikuttavat uskomusten muuttumiseen, ja tunnolliset, avoimet ja miellyttävät ihmiset vakuuttuvat enemmän tunneperäisillä argumenteilla.

**Tulos**

Argumentin vahvuus on katsojan silmässä: Yleisövaikutukset vakuuttamisessa

**Esimerkki 2.1162**

Tässä asiakirjassa esitellään itseorganisoituva liikennevalojärjestelmä kaupunkien tieverkkoa varten. Järjestelmän keskeisiä elementtejä ovat agentit, jotka ohjaavat liikennevaloja risteyksissä. Kukin agentti käyttää intervallimikroskooppista liikennemallia ennustaakseen mahdollisten ohjaustoimiensa vaikutukset lyhyellä aikajänteellä. Toteutettava ohjaustoimi valitaan ennustettujen viivytysvälien perusteella. Koska ennustustulokset esitetään intervalleina, agentit voivat tunnistaa ja keskeyttää ne ohjaustoimet, joiden myönteinen vaikutus liikenteenohjauksen suorituskykyyn on epävarma. Ehdotetun liikenteenohjausjärjestelmän arviointi suoritettiin simulointiympäristössä. Simulointikokeet ovat osoittaneet, että ehdotettu lähestymistapa parantaa suorituskykyä erityisesti epätasaisilla liikennevirroilla.

**Tulos**

Ennakoivaan intervallimikroskooppiseen malliin perustuva itseorganisoituva järjestelmä kaupunkiliikenteen valvontaa varten

**Esimerkki 2.1163**

Imitaatio-oppimista on perinteisesti sovellettu yksittäisen tehtävän oppimiseen sen demonstraatioiden perusteella. Vaatimus strukturoiduista ja erillisistä demonstraatioista rajoittaa imitaatio-oppimismenetelmien skaalautuvuutta, sillä niitä on vaikea soveltaa todellisissa tilanteissa, joissa robottien on kyettävä suorittamaan useita tehtäviä. Tässä artikkelissa ehdotamme multimodaalista imitaatio-oppimiskehystä, joka pystyy segmentoimaan ja imitoimaan taitoja merkitsemättömistä ja jäsentymättömistä demonstraatioista oppimalla taitojen segmentointia ja imitaatio-oppimista yhdessä. Laajat simulointitulokset osoittavat, että menetelmämme voi tehokkaasti erottaa demonstraatiot yksittäisiin taitoihin ja oppia jäljittelemään niitä käyttämällä yhtä multimodaalista käytäntöä. Video kokeiluistamme on saatavilla osoitteessa http://sites.google.com/view/nips17intentiongan.

**Tulos**

Monimodaalinen imitaatio-oppiminen strukturoimattomista esityksistä käyttäen generatiivisia adversaaliverkkoja.

**Esimerkki 2.1164**

Conditional Simple Temporal Network (CSTN) on rajoitteisiin perustuva graafi-formalismi ehdollista ajallista suunnittelua varten. Se tarjoaa joustavamman formalismin kuin Tsamardinosin, Vidalin ja Pollackin vastaava CSTP-malli, josta se on johdettu lähinnä äänellisenä formalisointina. CSTN:lle ja CSTP:lle syntyy kolme johdonmukaisuuden käsitettä: heikko, vahva ja dynaaminen. Dynaaminen johdonmukaisuus on mielenkiintoisin käsite, mutta se on myös haastavin, ja sen arveltiin olevan vaikeasti arvioitavissa. Tsamardinos, Vidal ja Pollack antoivat kaksi kertaa eksponentiaalisen ajan algoritmin, jolla voidaan päättää, onko CSTN dynaamisesti johdonmukainen, ja tuottaa positiivisessa tapauksessa eksponentiaalisen kokoinen dynaaminen suoritusstrategia. Tässä työssä esitämme todisteen siitä, että sen päättäminen, onko CSTN dynaamisesti johdonmukainen, on coNP-vaikeaa, ja tarjoamme ensimmäisen yksikäsitteisen eksponentiaaliajan algoritmin tälle ongelmalle, joka tuottaa myös dynaamisen suoritusstrategian aina kun syötetty CSTN on dynaamisesti johdonmukainen. Algoritmi perustuu uuteen yhteyteen Mean Payoff Games -peleihin, jotka ovat äärellisillä graafeilla pelattavien kahden pelaajan äärettömien pelien perhe, joka tunnetaan hyvin sovelluksistaan mallintarkistuksessa ja muodollisessa verifioinnissa. Tällaisen yhteyden esittelyssä käytetään Hyper Temporal Network -mallia, joka on helposti lähestyttävä yleistys Simple Temporal Networks -mallista, jonka johdonmukaisuuden tarkistaminen vastaa Mean Payoff Games -pelien määrittämistä. Algoritmin analysoimiseksi otamme käyttöön tarkennetun dynaamisen johdonmukaisuuden käsitteen, nimeltään -dynaaminen johdonmukaisuus, ja esitämme terävän analyysin reaktioajan ε̂ kriittisestä arvosta, jossa CSTN siirtyy dynaamisesti johdonmukaisesta olemattomaksi. Tässä ε̂:n analyysissä esitettyä todistustekniikkaa voidaan soveltaa yleisemmin käsiteltäessä lineaarisia erotusrajoituksia, jotka sisältävät tiukkoja epätasa-arvoja.

**Tulos**

Ehdollisten yksinkertaisten ajallisten verkostojen dynaaminen johdonmukaisuus keskimääräisen voiton pelien avulla: yksikäsitteisen eksponentiaalisen ajan DC-tarkistus.

**Esimerkki 2.1165**

Koneoppimistekniikoita on viime aikoina käytetty näyttävin tuloksin artefaktien, kuten musiikin tai tekstin, tuottamiseen. Näillä tekniikoilla ei kuitenkaan vielä pystytä vangitsemaan ja tuottamaan vakuuttavasti jäsenneltyjä artefakteja. Tässä artikkelissa esitellään lähestymistapa strukturoitujen musiikkisekvenssien tuottamiseen. Esittelemme mekanismin, jonka avulla voidaan tehokkaasti ottaa näytteitä musiikkisekvenssien variaatioista. Kun syötesekvenssi ja tilastollinen malli on annettu, mekanismi ottaa näytteitä joukosta sekvenssejä, joiden etäisyys syötesekvenssiin on suunnilleen tietyissä rajoissa. Mekanismi on toteutettu uskomuspropagoinnin laajennuksena, ja siinä käytetään paikallisia kenttiä tuottamisen suuntaamiseen. Osoitamme kokeellisesti, että näytteeksi otetut sekvenssit ovat todellakin läheisessä korrelaatiossa Mongeaun ja Sankoffin määrittelemän vakiomuotoisen musiikillisen samankaltaisuusmitan kanssa. Sitten näytämme, miten tätä mekanismia voidaan käyttää sellaisten sävellysstrategioiden toteuttamiseen, jotka pakottavat mielivaltaisen rakenteen musiikilliseen lead sheet -generaatio-ongelmaan.

**Tulos**

Lyijylevyjen näytteenottovariaatiot

**Esimerkki 2.1166**

Tietojen representaatioiden oppiminen ja erityisesti ominaisuuksien oppiminen myöhempää ennustustehtävää varten on ollut hedelmällinen tutkimusalue, josta on viime vuosina saatu vaikuttavia empiirisiä tuloksia. Suhteellisen vähän tiedetään kuitenkin siitä, mikä tekee representaatiosta "hyvän". Ehdotamme ajatusta edustuksen oppimisen aiheuttamasta riskivajeesta tietyssä ennustuskontekstissa, joka mittaa eroa riskissä, jonka joku oppija saa käyttämällä opittuja ominaisuuksia verrattuna alkuperäisiin syötteisiin. Kuvaamme joukon riittäviä ehtoja, joiden täyttyessä valvomaton edustuksen oppiminen tuottaa riskikuilun avulla mitattua hyötyä. Nämä ehdot hajottavat ongelman, milloin edustusoppiminen toimii, sen osiin, jotka voidaan arvioida erikseen käyttämällä merkitsemätöntä näytettä, sopivia alaan liittyviä oletuksia yhteisestä jakaumasta sekä ominaisuuksien oppijan ja sitä seuraavan valvotun oppijan analyysia. Esitämme kaksi esimerkkiä tällaisista ehdoista merkitsemättömän jakauman erityisominaisuuksien yhteydessä, nimittäin silloin, kun data on lähellä matalaulotteista moninaisuutta ja kun se muodostaa klustereita. Vertaamme lähestymistapaamme hiljattain ehdotettuun puolivalvotun oppimisen analyysiin.

**Tulos**

Modulaarinen teoria ominaisuuksien oppimisesta

**Esimerkki 2.1167**

Monissa todellisissa sovelluksissa, joissa käytetään ja analysoidaan verkkotietoja, verkostograafin linkit voivat olla virheellisiä tai ne on johdettu todennäköisyystekniikoiden avulla. Tällaisissa tapauksissa solmujen luokitusongelma voi olla haastava, koska linkkien epäluotettavuus voi vaikuttaa luokitusprosessin lopputuloksiin. Jos linkkien luotettavuutta koskevaa tietoa ei käytetä eksplisiittisesti, luokittelutarkkuus taustalla olevassa verkossa voi heikentyä. Tässä artikkelissa keskitymme tilanteisiin, jotka edellyttävät graafin rakenteessa esiintyvän epävarmuuden analysointia. Tutkimme uudenlaista ongelmaa, joka koskee solmujen luokittelua epävarmoissa graafeissa, käsittelemällä epävarmuutta ensimmäisen luokan kansalaisena. Ehdotamme kahta Bayes-malliin ja automaattiseen parametrien valintaan perustuvaa tekniikkaa ja osoitamme, että epävarmuuden sisällyttäminen luokitteluprosessiin ensimmäisen luokan kansalaisena on hyödyllistä. Arvioimme ehdotettua lähestymistapaa kokeellisesti käyttämällä erilaisia todellisia tietokokonaisuuksia ja tutkimme algoritmien käyttäytymistä eri olosuhteissa. Tulokset osoittavat lähestymistapamme tehokkuuden ja toimivuuden.

**Tulos**

Solmujen luokittelu epävarmoissa graafeissa

**Esimerkki 2.1168**

Tässä artikkelissa jatketaan Poincaren ja Russelin noidankehäperiaatteen (VCP) tutkimista joukkoihin perustuvien loogisten ohjelmointikielten suunnittelun yhteydessä. Laajennamme aiemmin esiteltyä kieltä Alog aggregaateilla sallimalla äärettömät joukot ja useita muita joukkoihin liittyviä konstruktioita, jotka ovat hyödyllisiä tiedon esittämisessä ja opetuksessa. Lisäksi ehdotamme alkuperäisen VCP:n vaihtoehtoista formalisointia ja sisällytämme sen uuden kielen, Slogin, semantiikkaan, joka sallii joukkojen vapaamman rakentamisen ja niiden käytön ohjelmointisäännöissä. Osoitamme, että ohjelmille, joissa ei ole disjunktiota eikä äärettömiä joukkoja, aggregaattien muodollinen semantiikka Slogissa on yhteneväinen useiden muiden tunnettujen kielten semantiikan kanssa. Niiden intuitiivinen ja muodollinen semantiikka perustuvat kuitenkin aivan erilaisiin ajatuksiin ja näyttävät olevan monimutkaisempia kuin Slogin semantiikka.

**Tulos**

Noidankehäperiaate ja joukkojen muodostaminen ASP-pohjaisissa kielissä

**Esimerkki 2.1169**

Tietohierarkiat ovat organisaatiorakenteita, joita käytetään usein suurten ja monimutkaisten tietojen järjestämiseen ja esittämiseen sekä tehokkaan navigointimekanismin tarjoamiseen ihmisille. Onneksi on olemassa monia tilastollisia ja laskennallisia malleja, jotka luovat hierarkioita automaattisesti; nykyisissä lähestymistavoissa ei kuitenkaan oteta huomioon tietoverkkojen yhteyksiä, jotka ovat yhä yleisempiä todellisissa tilanteissa. Nykyiset lähestymistavat pyrkivät myös esittämään aiheet abstrakteina, todennäköisesti sanoihin jne. kohdistuvina jakaumina sen sijaan, että ne esitettäisiin alkuperäisen verkon konkreettisina solmuina. Lisäksi monissa aiemmissa töissä käytetyt tilastolliset tekniikat eivät vielä pysty käsittelemään verkkomittakaavan tietoja. Tässä asiakirjassa esitellään hierarkkinen dokumenttien aihepiirimalli (HDTM), jossa käytetään hajautettua vertex-ohjelmointiprosessia ei-parametrisen Bayesin generatiivisen mallin laskemiseen. Kokeet kolmella keskikokoisella aineistolla ja koko Wikipedian aineistolla osoittavat, että HDTM pystyy päättelemään tarkkoja hierarkioita jopa suurissa tietoverkoissa.

**Tulos**

Skaalautuvat mallit tietoverkkojen hierarkioiden laskemiseen

**Esimerkki 2.1170**

Tutkitaan maamerkkeihin perustuvaa heuristiikkaa, jolla voidaan vähentää todennäköisyysperusteisen tiekartan (PRM) liikesuunnittelumenetelmän kyselyvaiheen suoritusaikaa. Heuristiikka luodaan tallentamalla PRM-graafin pienestä määrästä kärkipisteitä minimitilannepuita ja käyttämällä näitä puita arvioimaan lyhyimmän polun kustannuksia minkä tahansa graafin kahden kärkipisteen välillä. Graafin esikäsittelyn välivaihe lisää klassisen liikesuunnittelutekniikan aika- ja muistivaatimuksia, mutta nopeuttaa yksittäisiä kyselyjä, mikä tekee menetelmästä edullisen monikyselysovelluksissa. Tässä artikkelissa tutkitaan näitä kompromisseja satunnaistetuissa ympäristöissä rakennetuilla PRM-grafeilla sekä käytännön manipulaattorisimulaatiolla. Tuloksena on, että menetelmä on parempi kuin Dijkstran algoritmi tai A∗-algoritmi perinteisillä heuristiikoilla monikyselysovelluksissa.

**Tulos**

Maamerkkiohjatut todennäköisyyskarttakyselyt (Landmark Guided Probabilistic Roadmap Queries)

**Esimerkki 2.1171**

Ehdotamme DENN-verkkoa (Diverse Embedding Neural Network), joka on uudenlainen arkkitehtuuri kielimalleille (LM). DENNLM projisoi syötetyn sanahistoriavektorin useisiin erilaisiin matalaulotteisiin aliavaruuksiin sen sijaan, että se projisoi yhden korkeampiulotteisen aliavaruuden, kuten tavanomaisissa feed-forward-neuraaliverkkomalleissa. Kannustamme näitä aliavaruuksia olemaan erilaisia verkon harjoittelun aikana täydennetyn tappiofunktion avulla. Penn Treebank -datasarjalla tekemämme kielimallinnuskokeet osoittavat, että DENNLM:n käyttö on suorituskykyistä.

**Tulos**

MONIPUOLISET UPOTETUT NEUROVERKON KIELIMALLIT

**Esimerkki 2.1172**

Ei-negatiivinen matriisifaktorointi (NMF) on luonnollinen sekoittumisen malli, ja sitä käytetään laajalti tieteessä ja tekniikassa. NMF:n ratkaisemiseksi on kehitetty lukuisia algoritmeja, mutta ongelman ei-konveksaalisen luonteen vuoksi näiden menetelmien toimivuudesta ei ole juurikaan takeita. Viime aikoina tutkimukset ovat keskittyneet hyvin rajoitettuun NMF-luokkaan, jota kutsutaan erotettavaksi NMF:ksi ja johon on kehitetty todistettavasti oikeita algoritmeja. Tässä artikkelissa ehdotamme osajoukko-erotettavan NMF:n käsitettä, joka olennaisesti yleistää erotettavuuden ominaisuutta. Osoitamme, että osajoukko-ositettavuus on luonnollinen välttämätön ehto sille, että faktorointi on yksikäsitteinen tai että sillä on minimitilavuus. Kehitimme Face-Intersect-algoritmin, joka todisteellisesti ja tehokkaasti ratkaisee osajoukko-erotettavissa olevan NMF:n luonnollisissa olosuhteissa, ja todistamme, että algoritmimme on kestävä pienelle kohinalle. Tutkimme Face-Intersectin suorituskykyä simulaatioissa ja keskustelemme asetuksista, joissa se empiirisesti päihitti nykyiset menetelmät. Työmme on askel kohti todistettavasti oikeita algoritmeja, jotka ratkaisevat suuria NMF-ongelmaluokkia.

**Tulos**

Risteävät kasvot: Ei-negatiivinen matriisifaktorointi uusilla takeilla

**Esimerkki 2.1173**

Esittelemme uudenlaisen lähestymistavan rajoituksiin perustuvaan kausaalisten suhteiden löytämiseen, joka toteutetaan suoraviivaisena loogisena päättelynä, jota sovelletaan luetteloon yksinkertaisista, loogisista lausumista kausaalisuhteista, jotka on johdettu suoraan havaituista (epä)riippuvuuksista. Se on sekä järkevä että täydellinen siinä mielessä, että kaikki vastaavan osittaisen esi-isäkaavion (PAG) muuttumattomat piirteet tunnistetaan, vaikka mukana olisi latentteja muuttujia ja valintaharhaa. Lähestymistapa osoittaa, että jokainen tunnistettava kausaalisuhde vastaa yhtä kahdesta perusmuodosta. Vielä tärkeämpää on se, että koska menetelmän perusrakenteet eivät perustu vastaavan PAG:n yksityiskohtaiseen (graafiseen) rakenteeseen, se avaa monia uusia mahdollisuuksia, kuten vankemman päättelyn, yksityiskohtaisemman tilivelvollisuuden ja soveltamisen suuriin malleihin.

**Tulos**

Looginen luonnehdinta rajoituksiin perustuvasta kausaalisen löydön löytämisestä.

**Esimerkki 2.1174**

\*Rajdeep Borgohain Department of Computer Science and Engineering, Dibrugarh University Institute of Engineering and Technology, Dibrugarh, Assam Sähköposti: rajdeepgohain@gmail.com Sugata Sanyal School of Technology and Computer Science, Tata Ins titute of Fundamental Research, Mumbai, Intia Sähköposti: sanyals@gmail.com \*Corresponding Author -------------------------------------------------------------------ABSTRACT-----------------------------------------------------------------Tekoälyä käytetään paitsi keskeisillä tietotekniikka-aloilla myös monitieteisillä aloilla, mukaan lukien lääketieteellinen diagnostiikka. Tässä artikkelissa esittelemme sääntöpohjaisen asiantuntijajärjestelmän, jota käytetään aivohalvauksen diagnosointiin. Asiantuntijajärjestelmä ottaa käyttäjän syötteen ja diagnosoi potilaan oireiden perusteella, kärsiikö potilas aivohalvauksesta. Asiantuntijajärjestelmä myös luokittelee aivohalvauksen esitettyjen oireiden perusteella lieväksi, keskivaikeaksi tai vaikeaksi.

**Tulos**

Sääntöpohjainen asiantuntijajärjestelmä aivohalvausdiagnoosia varten

**Esimerkki 2.1175**

Vaikka päästä päähän neuraalinen konekääntäminen (NMT) on viime aikoina edistynyt huomattavasti, NMT-järjestelmät perustuvat parametrien arvioinnissa vain rinnakkaisiin korpuksiin. Koska rinnakkaiskorporaatioiden määrä, laatu ja kattavuus ovat yleensä rajalliset, erityisesti vähäisten resurssien kielissä, on houkuttelevaa hyödyntää yksikielisiä korporaatioita NMT:n parantamiseksi. Ehdotamme puolivalvottua lähestymistapaa NMT-mallien kouluttamiseen leimattujen (rinnakkaiset korporaatiot) ja leimaamattomien (yksikieliset korporaatiot) tietojen yhdistelmällä. Keskeinen ajatus on rekonstruoida yksikieliset korporaatiot käyttämällä autokooderia, jossa lähde-kohde- ja kohde-lähde-käännösmallit toimivat vastaavasti kooderina ja dekooderina. Lähestymistapamme ei voi hyödyntää ainoastaan kohdekielen vaan myös lähdekielen yksikielisiä korpuksia. Kokeet kiinalais-englantilaisella tietokokonaisuudella osoittavat, että lähestymistapamme parantaa merkittävästi nykyaikaisia SMT- ja NMT-järjestelmiä.

**Tulos**

Puolivalvottu oppiminen neuraalista konekääntämistä varten

**Esimerkki 2.1176**

Bayesin verkko (BN) on muutakin kuin tiivis tapa koodata todennäköisyysjakauma; se vastaa myös funktiota, jota käytetään kyselyihin vastaamiseen. BN:ää voidaan siis arvioida sen palauttamien vastausten tarkkuuden perusteella. Monet BN-verkkojen oppimiseen tarkoitetut algoritmit pyrkivät kuitenkin optimoimaan toista kriteeriä (tavallisesti todennäköisyyttä, mahdollisesti täydennettynä regularisoivalla termillä), joka on riippumaton esitettyjen kyselyjen jakaumasta. Tässä artikkelissa otetaan "suorituskykyperusteet" vakavasti ja tarkastellaan haastetta löytää BN, jonka suorituskyky "tarkkuus kyselyjen jakauman suhteen" on optimaalinen. Osoitamme, että monet tämän oppimistehtävän osa-alueet ovat vaikeampia kuin vastaavat osatehtävät standardimallissa.

**Tulos**

Hyvin toimivien Bayes-verkkojen oppiminen

**Esimerkki 2.1177**

Viime aikoina on kehitetty runsaasti algoritmeja, joilla on todistettavat takuut aihepiirien mallintamiseen. Tavanomaisissa aihepiirimalleissa aihepiiri (kuten urheilu, liike-elämä tai politiikka) nähdään sanojen todennäköisyysjakaumana ai, ja dokumentti luodaan valitsemalla ensin aihepiirien seos w ja luomalla sitten sanat i.i.d. siihen liittyvästä seoksesta Aw. Kun otetaan huomioon suuri kokoelma tällaisia asiakirjoja, tavoitteena on palauttaa aihevektorit ja sitten luokitella uudet asiakirjat oikein niiden aiheseoksen mukaan. Tässä työssä tarkastelemme tämän kehyksen laajaa yleistystä, jossa sanojen ei enää oleteta olevan i.i.d. ja sen sijaan aihe on monimutkainen jakauma kappaleiden sarjoihin. Koska tällaista jakaumaa ei voida edes toivoa yleisesti edustettavan (vaikka kappaleet annettaisiinkin jonkin luonnollisen piirre-esityksen avulla), pyrimme sen sijaan oppimaan suoraan asiakirjaluokittelijan. Toisin sanoen pyrimme oppimaan ennustajan, joka ennustaa uuden asiakirjan perusteella tarkasti sen aihepiirin sekoituksen ilman, että jakaumia opetetaan eksplisiittisesti. Esittelemme useita luonnollisia olosuhteita, joissa tämä voidaan tehdä tehokkaasti, ja keskustelemme esimerkiksi kohinansietokyvystä ja näytteiden monimutkaisuudesta tässä mallissa. Yleisemmin malliamme voidaan pitää koneoppimisen moninäkökulma- tai yhteiskoulutusasetelman yleistyksenä. ∗ Osittain tuettu National Science Foundationin apurahoilla CCF-1525971 ja CCF-1535967. †Tuettu osittain National Science Foundationin apurahalla CCF-1525971 sekä Microsoft Research Graduate Fellowship -apurahalla ja IBM:n Ph.D Fellowship -apurahalla. ar X iv :1 61 1. 01 25 9v 1 [ cs .L G ] 4 N ov 2 01 6

**Tulos**

Yleistetty aihepiirien mallintaminen

**Esimerkki 2.1178**

Käsittelemme ongelmaa, joka liittyy taloudellisten tapahtumien jäsenneltyjen esitysten poimimiseen laajasta uutisartikkelien joukosta, käyttämällä luonnollisen kielen käsittely- ja koneoppimistekniikoiden yhdistelmää. Kehitettyjen tekniikoiden avulla voidaan luoda puoliautomaattisesti taloudellinen tietopohja, jota puolestaan voidaan käyttää tukemaan erilaisia tiedonlouhinta- ja tutkimustehtäviä. Keskeinen haaste tällä alalla on se, että samasta tapahtumasta raportoidaan usein useita kertoja, ja yksityiskohdat ovat vaihtelevan tarkkoja. Ratkaisemme tämän haasteen keräämällä ensin kaikki tiettyyn tapahtumaan liittyvät tiedot koko korpuksesta, tarkastelemalla sitten kaikkia mahdollisia tapahtuman esitystapoja ja asettamalla nämä esitykset valvotun oppimismenetelmän avulla paremmuusjärjestykseen niihin liittyvien luotettavuuspisteiden perusteella. Lähestymistapamme tärkein innovatiivinen elementti on se, että siinä poimitaan ja tallennetaan kaikki tapahtuman attribuutit yhdessä ainoana esityksenä (viisinkertaisena). Käyttämällä tarkoitusta varten luotua testijoukkoa osoitamme, että valvotun oppimisen lähestymistapamme voi parantaa F1-pistemäärää 25 prosenttia verrattuna perusmenetelmiin, jotka ottavat huomioon tapahtuman varhaisimman, viimeisimmän tai yleisimmän raportoinnin.

**Tulos**

Kohti rahaliikenteen tietopohjan rakentamista uutiskokoelmasta

**Esimerkki 2.1179**

Monet luonnollisen kielen ymmärtämistehtävät (NLU), kuten matala jäsennys (eli tekstin pilkkominen) ja semanttisten aukkojen täyttäminen, edellyttävät edustavien merkintöjen antamista lauseen merkityksellisille palasille. Useimmissa nykyisissä syviin neuroverkkoihin (deep neural network, DNN) perustuvissa menetelmissä näitä tehtäviä tarkastellaan sekvenssien merkintäongelmana, jossa sana, ei niinkään katkelma, on merkinnän perusyksikkö. Kappaleet päätellään sitten vakiomuotoisten IOB (Inside-OutsideBeginning) -merkintöjen avulla. Tässä asiakirjassa ehdotamme vaihtoehtoista lähestymistapaa tutkimalla DNN:n käyttöä sekvenssien kappalemäärittelyssä ja ehdotamme kolmea neuromallia, jotta jokaista kappaletta voidaan käsitellä täydellisenä yksikkönä merkitsemistä varten. Kokeelliset tulokset osoittavat, että ehdotetut neuraaliset sekvenssin pilkkomismallit voivat saavuttaa huippuluokan suorituskyvyn sekä tekstin pilkkomis- että aukkotäytteistämistehtävissä.

**Tulos**

Neuraaliset mallit sekvenssien pilkkomiseen

**Esimerkki 2.1180**

Suurin osa tilastollisen relaatio-oppimisen (SRL) alalla tehdystä työstä keskittyy diskreetteihin tietoihin, vaikka on ehdotettu muutamia lähestymistapoja hybrideille SRL-malleille, joissa yhdistetään numeerisia ja diskreettejä muuttujia. Tässä asiakirjassa erotetaan numeeriset satunnaismuuttujat, joille malli määrittelee todennäköisyysjakauman, numeerisista syötemuuttujista, joita käytetään vain diskreettien vastemuuttujien jakauman ehdollistamiseen. Osoitamme, miten numeerisia syöttösuhteita voidaan hyvin helposti käyttää Relational Bayesian Network -kehyksessä ja että nykyiset päättely- ja oppimismenetelmät tarvitsevat vain pieniä muutoksia, jotta niitä voidaan soveltaa tässä yleistetyssä ympäristössä. Tuloksena syntyvä kehys tarjoaa luonnollisia relaatioperusteisia laajennuksia klassisiin todennäköisyysmalleihin kategorista dataa varten. Osoitamme RBN-mallien hyödyllisyyden numeeristen syöttörelaatioiden kanssa useiden esimerkkien avulla. Erityisesti käytämme laajennettua RBN-kehystä määritellessämme todennäköisyysmalleja monisuhteisille (sosiaalisille) verkoille, joissa kahden solmun välisen yhteyden todennäköisyys riippuu solmuihin liittyvistä numeerisista latenttisista ominaisuusvektoreista. Yleistä oppimismenettelyä voidaan käyttää mallin parametrien ja latenttien ominaisuuksien arvojen suurimman todennäköisyyden sovittamiseen erilaisille malleille, jotka voidaan ilmaista korkean tason RBN-edustuksella. Ehdotamme erityisesti mallia, jonka avulla voimme tulkita opitut latenttien ominaisuuksien arvot yhteisön keskeisyysasteiksi, joiden avulla voimme tunnistaa solmut, jotka ovat keskeisiä yhdelle yhteisölle, jotka ovat solmukohtia yhteisöjen välillä tai jotka ovat eristyneitä solmuja. Monisuhteisessa ympäristössä malli tarjoaa myös kuvauksen siitä, miten eri suhteet liittyvät kuhunkin yhteisöön.

**Tulos**

Numeeriset syöttösuhteet suhteellista oppimista varten ja sovellukset yhteisön rakenneanalyysiin

**Esimerkki 2.1181**

Viime vuosina on edistytty merkittävästi rekurrenssia neuroverkkoja (RNN) on onnistuttu kouluttamaan menestyksekkäästi sekvenssioppimisongelmiin, joihin liittyy pitkän aikavälin ajallisia riippuvuuksia. Edistystä on saavutettu kolmella rintamalla: a) algoritmien parannukset, joihin liittyy kehittyneitä optimointitekniikoita, b) verkon suunnittelu, johon liittyy monimutkaisia piilokerroksen solmuja ja erikoistuneita rekurrenssikerroksen yhteyksiä, ja c) painojen alustamismenetelmät. Tässä artikkelissa keskitymme hiljattain ehdotettuun painojen alustamiseen identiteettimatriisilla RNN:n rekursiivisille painoille. Tätä alustusta on ehdotettu erityisesti piilosolmuja varten, joissa on ReLU (ReLU) -epälineaarisuus (Rectified Linear Unit). Tarjoamme yksinkertaisen dynaamisten järjestelmien näkökulman painojen alustamisprosessiin, jonka avulla voimme ehdottaa muunnettua painojen alustamisstrategiaa. Osoitamme, että tämä alustustekniikka johtaa ReLU:ista koostuvien RNN:ien onnistuneeseen koulutukseen. Osoitamme, että ehdotuksemme tuottaa vertailukelpoisen tai paremman ratkaisun kolmeen leikkiongelmaan, joihin liittyy pitkän aikavälin ajallinen rakenne: yhteenlaskuongelmaan, kertolaskuongelmaan ja MNIST-luokitusongelmaan, jossa käytetään pikseleiden sekvenssiä. Lisäksi esitämme tuloksia vertailuongelmasta, joka koskee toiminnan tunnistamista.

**Tulos**

TOISTUVAN NEUROVERKON SUORITUSKYVYN PARANTAMINEN RELU EPÄLINEAARISUUDEN KANSSA

**Esimerkki 2.1182**

Tässä artikkelissa tutkitaan neuraalista merkkipohjaista morfologista merkintää kielissä, joissa on monimutkainen morfologia ja suuret merkkijoukot. Tutkimme järjestelmällisesti erilaisia neuroarkkitehtuureja (DNN, CNN, CNNHighway, LSTM, BLSTM) saadaksemme merkkipohjaisia sanavektoreita yhdistettynä kaksisuuntaisiin LSTM:iin sanojen välisen kontekstin mallintamiseksi päästä päähän -asetelmassa. Tutkimme sanapohjaisten vektoreiden lisäkäyttöä, jotka on koulutettu suurilla määrillä merkitsemätöntä dataa. Morfologista merkintää koskevat kokeilumme osoittavat, että "yksinkertaisissa" mallikokoonpanoissa verkkoarkkitehtuurin valinta (CNN vs. CNNHighway vs. LSTM vs. BLSTM) tai täydennys valmiiksi koulutetuilla sanasulkeumilla voi olla tärkeää ja vaikuttaa selvästi tarkkuuteen. Mallin kapasiteetin kasvattaminen esimerkiksi lisäämällä syvyyttä ja optimoimalla neuroverkkoja huolellisesti voi johtaa huomattaviin parannuksiin, ja erot tarkkuudessa (mutta ei harjoittelussa kuluvassa ajassa) muuttuvat paljon pienemmiksi tai jopa mitättömiksi. Kaiken kaikkiaan parhaat morfologiset taggerimme saksan ja tšekin kielille ylittävät kirjallisuudessa raportoidut parhaat tulokset selvästi.

**Tulos**

Neurologinen morfologinen merkintä merkkien perusteella morfologisesti rikkaissa kielissä

**Esimerkki 2.1183**

Uudelleenjärjestäminen on suuri haaste konekääntämisessä kahden sellaisen kielen välillä, joiden sanajärjestyksessä on merkittäviä eroja. Tässä artikkelissa esitellään uusi uudelleenjärjestämismenetelmä, jossa käytetään riippuvuussanapareihin perustuvia harvalukuisia piirteitä. Kukin näiden ominaisuuksien instanssi kuvaa, noudattavatko kaksi sanaa, jotka liittyvät toisiinsa riippuvuuslinkin avulla lähdelauseen riippuvuussuhteiden jäsennyspuussa, samaa järjestystä vai vaihdetaanko ne käännöstulosteessa. Kokeet kiinan kielen kääntämisessä englanniksi osoittavat, että lähestymistapamme avulla saavutetaan tilastollisesti merkittävä 1,21 BLEU-pisteen parannus verrattuna uusimpaan tilastolliseen tekstinmuunnosjärjestelmään, joka sisältää aiempia uudelleenjärjestämismenetelmiä.

**Tulos**

Vaihtaa vai olla vaihtamatta? Riippuvuussanaparien hyödyntäminen uudelleenjärjestämisessä tilastollisessa konekäännöksessä

**Esimerkki 2.1184**

Tutkimme hajautettua laskentaa, jossa on useita palvelimia, joilla kullakin on pistejoukko, ja jotka haluavat laskea funktioita pistejoukkojensa liitosta. Keskeinen tehtävä tässä ympäristössä on pääkomponenttianalyysi (PCA), jossa palvelimet haluavat laskea matalaulotteisen aliavaruuden, joka kattaa mahdollisimman suuren osan pistejoukkojensa liiton varianssista. Kun PCA:n approksimaatiomenetelmä on olemassa, sitä voidaan käyttää k-means-klusteroinnin ja matalan arvon approksimaation kaltaisten ongelmien approksimointiin. Approksimatiivisen hajautetun PCA-algoritmin olennaisia ominaisuuksia ovat sen viestintäkustannukset ja laskennallinen tehokkuus, kun haluttu tarkkuus halutaan saavuttaa myöhemmissä sovelluksissa. Esitämme uusia algoritmeja ja analyysejä hajautetulle PCA:lle, jotka johtavat parempiin viestintä- ja laskentakustannuksiin k-means-klusterointia ja siihen liittyviä ongelmia varten. Empiirinen tutkimuksemme reaalimaailman tiedoilla osoittaa, että nopeutuminen on suuruusluokkaa, ja viestintä säilyy, mutta ratkaisun laatu heikkenee vain vähän. Jotkin näistä kehittämistämme tekniikoista, kuten yleinen muunnos vakion onnistumistodennäköisyyden alitilan sulautuksesta korkean onnistumistodennäköisyyden alitilan sulautukseen, jonka ulottuvuus ja harvuus ovat riippumattomia onnistumistodennäköisyydestä, voivat olla itsenäisesti kiinnostavia.

**Tulos**

Parannettu hajautettu pääkomponenttianalyysi

**Esimerkki 2.1185**

Syvät neuroverkot (Deep Neural Networks, DNN) ovat uusinta tekniikkaa monissa teknisissä ongelmissa, kuten tietokonenäkemisessä ja kuulonvarmistuksessa. Avaintekijä DNN:n menestyksessä on skaalautuvuus - suuremmat verkot toimivat paremmin. Syytä tähän skaalautuvuuteen ei kuitenkaan vielä ymmärretä hyvin. Tulkitsemme DNN:n diskreetiksi järjestelmäksi, joka koostuu lineaarisista suodattimista, joita seuraavat epälineaariset aktivoinnit, ja johon sovelletaan näytteenottoteorian lakeja. Tässä yhteydessä osoitamme, että ylinäytteistetyt verkot ovat valikoivampia, oppivat nopeammin ja oppivat kestävämmin. Tuloksemme voivat lopulta yleistyä ihmisaivoihin.

**Tulos**

Ylinäytteenotto syvässä neuroverkossa\_arxiv

**Esimerkki 2.1186**

Tarkastelemme suunnittelun oppimisen ongelmaa, jossa suunnittelun aikana hankittua tietoa käytetään uudelleen nopeamman suunnittelun mahdollistamiseksi uusissa ongelmatapauksissa. Muun muassa robottitehtävissä suunnitelman toteutus voidaan tallentaa visuaalisten kuvien sarjana. Ehdotamme, että tällaisilla aloilla käytetään syviä neuroverkkoja suunnittelun oppimiseen, joka perustuu reaktiivisen politiikan oppimiseen, joka jäljittelee suunnittelijan tuottamia suoritusjälkiä. Tutkimme syvien neuroverkkojen arkkitehtuurisia ominaisuuksia, jotka soveltuvat pitkän horisontin suunnittelukäyttäytymisen oppimiseen, ja tutkimme, miten politiikan lisäksi voidaan oppia heuristinen funktio, jota voidaan käyttää klassisten suunnittelijoiden tai hakualgoritmien, kuten A∗:n, kanssa. Haastavalla Sokoban-domainilla saamamme tulokset osoittavat, että sopivalla verkkosuunnittelulla voidaan jäljittelemällä oppia monimutkaisia päätöksentekopolitiikkoja ja tehokkaita heuristisia funktioita. Videot ovat saatavilla osoitteessa https://sites.google.com/site/learn2plannips/.

**Tulos**

Yleistettyjen reaktiivisten toimintatapojen oppiminen syvien neuroverkkojen avulla

**Esimerkki 2.1187**

Osoitamme, että yhteistoiminnallista suodatusta voidaan tarkastella sekvenssin ennustamisongelmana ja että tämän tulkinnan perusteella toistuvat neuroverkot tarjoavat erittäin kilpailukykyisen lähestymistavan. Tutkimme erityisesti, miten pitkää lyhytkestoista muistia (LSTM) voidaan soveltaa yhteissuodatukseen ja miten se vertautuu elokuvien suosittelussa tavanomaisiin lähimpiin naapureihin ja matriisifaktorointimenetelmiin. Osoitamme, että LSTM on kilpailukykyinen kaikilta osin, ja se päihittää suurelta osin muut menetelmät kohteiden kattavuuden ja lyhyen aikavälin ennusteiden osalta.

**Tulos**

Yhteistoiminnallinen suodatus toistuvien neuroverkkojen avulla

**Esimerkki 2.1188**

Viime aikoina monet varianssia pienentävät stokastiset vuorottelevan suunnan kertolaskumenetelmät (ADMM) (esim. SAGADMM, SDCA-ADMM ja SVRG-ADMM) ovat saavuttaneet jännittävää edistystä, kuten lineaarisen konvergenssinopeuden vahvasti koverille ongelmille. Paras tunnettu konvergenssinopeus yleisille koverille ongelmille on kuitenkin O(1/T ), toisin kuin nopeutettujen eräalgoritmien O(1/T ), jossa T on iteraatioiden lukumäärä. Näin ollen nykyisten stokastisten ADMM-algoritmien ja eräalgoritmien lähentymisnopeuksissa on edelleen eroa. Tämän kuilun kuromiseksi umpeen otamme eräoptimointiin tarkoitetun momenttikiihdytystempun käyttöön stokastiseen varianssia pienentävään gradienttiin perustuvassa ADMM:ssä (SVRG-ADMM), mikä johtaa kiihdytettyyn (ASVRG-ADMM) menetelmään. Sitten suunnittelemme kaksi erilaista momenttitermien päivityssääntöä vahvasti koveria ja yleisesti koveria tapauksia varten. Todistamme, että ASVRG-ADMM konvergoi lineaarisesti voimakkaasti koverille ongelmille. Sen lisäksi, että ASVRG-ADMM-menetelmällä on alhainen iteraatiokohtainen monimutkaisuus kuin nykyisillä stokastisilla ADMM-menetelmillä, se parantaa konvergenssinopeutta yleisissä koverissa ongelmissa O(1/T ):stä O(1/T ):een. Kokeelliset tulokset osoittavat ASVRG-ADMM:n tehokkuuden. Johdanto Tässä artikkelissa tarkastelemme luokkaa yhdistettyjä koveria koveria optimointiongelmia min x∈Rd1 f(x) + h(Ax), (1) jossa A∈Rd2×d1 on annettu matriisi, f(x) := 1 n ∑n i=1fi(x), jokainen fi(x) on kovera funktio ja h(Ax) on kovera mutta mahdollisesti epäsileä. Mitä tulee h(-):hen, olemme kiinnostuneita harvennusta aiheuttavasta regularisaattorista, esim. `1-normista, ryhmä-Lassosta ja ydinnormista. Kun A on identtisyysmatriisi eli A = Id1, edellä esitetty muotoilu (1) esiintyy monin paikoin koneoppimisessa, tilastotieteessä ja operaatiotutkimuksessa (Bubeck 2015), kuten logistisessa regressiossa, Lassossa ja tukivektorikoneessa (SVM). Keskitymme pääasiassa suuren otoksen järjestelmään. Tässä regiimissä jopa ensimmäisen kertaluvun erämenetelmät, esimerkiksi FISTA (Beck ja Teboulle 2009), muuttuvat laskennallisesti raskaiksi niiden O(nd1):n iteraatiokohtaisen monimutkaisuuden vuoksi. Tämän seurauksena stokastinen gradienttilaskeutuminen (SGD), jonka iteraatiokohtainen monimutkaisuus on O(d1), on Copyright c © 2017, Association for the Advancement of Artificial Intelligence (www.aaai.org). Kaikki oikeudet pidätetään. on viime vuosina tapahtunut valtavaa edistystä. Erityisesti on ehdotettu useita stokastisia varianssia vähentäviä gradienttimenetelmiä, kuten SAG (Roux, Schmidt ja Bach 2012), SDCA (Shalev-Shwartz ja Zhang 2013) ja SVRG (Johnson ja Zhang 2013), joilla on onnistuttu ratkaisemaan gradienttiestimaatin suuren varianssin ongelma tavallisessa SGD:ssä, mikä johtaa lineaariseen konvergenssivauhtiin (vahvasti konveksissa ongelmissa) SGD:n sublineaaristen nopeuksien sijaan. Viime aikoina Nesterovin kiihdytystekniikka (Nesterov 2004) otettiin käyttöön (Allen-Zhu 2016; Hien et al. 2016) stokastisten varianssivähennettyjen algoritmien nopeuttamiseksi entisestään, mikä johtaa parhaisiin tunnettuihin konvergenssilukuihin sekä vahvasti koverille että yleisille koverille ongelmille. Tämä motivoi meitä integroimaan momenttikiihdytystempun jäljempänä esitettävään stokastiseen vaihtoehtosuuntaiseen kertojamenetelmään (ADMM). Kun A on yleisempi matriisi, eli A 6=Id1 , muotoilusta (1) tulee monia monimutkaisempia ongelmia, jotka johtuvat koneoppimisesta, esim. graafiohjatusta sytytetystä Lassosta (Kim, Sohn ja Xing 2009) ja yleistetystä Lassosta (Tibshirani ja Taylor 2011). Tämän luokan yhdistettyjen optimointiongelmien ratkaisemiseksi, joissa on apumuuttuja y = Ax ja jotka ovat yleisen ADMM-muodon erikoistapaus, min x∈Rd1,y∈Rd2 f(x) + h(y), s.t. Ax+By = c, (2) ADMM on tehokas optimointityökalu (Boyd et al. 2011), ja se on osoittanut houkuttelevaa suorituskykyä monissa reaalimaailman ongelmissa, kuten big datan luokituksessa (Nie et al. 2014). Ratkaistakseen ongelman, joka liittyy eräkohtaisen (deterministisen) ADMM:n (suosittuna ensimmäisen kertaluvun optimointimenetelmänä) suureen periteraatiokompleksisuuteen, Wang ja Banerjee (2012), Suzuki (2013) ja Ouyang et al. (2013) ehdottivat joitakin online- tai stokastisia ADMM-algoritmeja. Kaikilla näillä muunnelmilla saavutetaan kuitenkin vain O(1/ √ T ) konvergenssinopeus yleisille koverille ongelmille ja O(log T/T ) vahvasti koverille ongelmille verrattuna nopeutettujen eräalgoritmien (Nesterov 1983), esimerkiksi FISTA:n, O(1/T ) ja lineaariseen konvergenssinopeuteen, jossa T on iteraatioiden lukumäärä. Tähän mennessä on ehdotettu useita nopeutettuja ja nopeammin konvergoivia versioita stokastisesta ADMM:stä, jotka kaikki perustuvat varianssin vähennystekniikoihin, esimerkiksi SAG-ADMM (Zhong ja Kwok 2014b), SDCA-ADMM (Suzuki 2014) ja SVRG-ADMM (Zheng ja Kwok 2016). Vahvasti koverien ongelmien osalta ar X iv :1 70 7. 03 19 0v 1 [ cs .L G ] 1 1 Ju l 2 01 7 Taulukko 1: Joidenkin stokastisten ADMM-algoritmien konvergenssinopeuksien ja muistivaatimusten vertailu. Yleinen konveksinen Vahvasti konveksinen Tilantarve SAG-ADMM O(1/T ) tuntematon O(d1d2+nd1) SDCA-ADMM tuntematon lineaarinen nopeus O(d1d2+n) SCAS-ADMM O(1/T ) O(1/T ) O(d1d2) SVRG-ADMM O(1/T ) lineaarinen nopeus O(d1d2) ASVRG-ADMM O(1/T ) lineaarinen nopeus O(d1d2) Suzuki (2014) ja Zheng ja Kwok (2016) osoittivat, että lineaarinen konvergenssi voidaan saada ADMM:n erityismuodolle (i.eli B = -Id2 ja c = 0) ja yleisen ADMM-muodon osalta. SAG-ADMM:ssä ja SVRG-ADMM:ssä voidaan taata O(1/T ) konvergenssinopeus yleisille koverille ongelmille, mikä merkitsee, että stokastisten ADMM-algoritmien ja kiihdytettyjen eräalgoritmien välille jää edelleen ero konvergenssinopeuksissa. Tämän kuilun kuromiseksi umpeen integroimme (Tseng 2010) deterministisen optimoinnin momenttikiihdytystempun (Tseng 2010) stokastiseen varianssin vähennysgradienttiin (SVRG) perustuvaan stokastiseen ADMM:ään (SVRG-ADMM). Ehdotetulla menetelmällä on luonnollisesti alhainen iteraatiokohtainen aikakompleksisuus kuin olemassa olevilla stokastisilla ADMM-algoritmeilla, eikä se vaadi kaikkien gradienttien (tai duaalimuuttujien) tallentamista, kuten SCAS-ADMM:ssä (Zhao, Li ja Zhou 2015) ja SVRGADMM:ssä (Zheng ja Kwok 2016), kuten taulukosta 1 käy ilmi. Seuraavassa esitetään yhteenveto tärkeimmistä panoksistamme. - Ehdotamme kiihdytettyä varianssia vähentävää stokastista ADMM-menetelmää (ASVRG-ADMM), joka integroi sekä (Tseng 2010) eräoptimoinnissa käytettävän momentumkiihdytystempun että SVRG:n (Johnson ja Zhang 2013) varianssin vähennystekniikan. - Osoitamme, että ASVRG-ADMM saavuttaa lineaarisen konvergenssinopeuden vahvasti kuperaongelmille, mikä on yhdenmukainen SDCA-ADMM:n (Suzuki 2014) ja SVRG-ADMM:n (Zheng ja Kwok 2016) parhaan tunnetun tuloksen kanssa. - Todistamme myös, että ASVRG-ADMM:n konvergenssinopeus on O(1/T ) ei-vahvasti koverille ongelmille, mikä on kertoimen T nopeampi kuin SAG-ADMM:n ja SVRG-ADMM:n, joiden konvergenssinopeudet ovat O(1/T ). - Kokeelliset tuloksemme vahvistavat lisäksi, että ASVRG-ADMM-menetelmämme suorituskyky on paljon parempi kuin nykyaikaisilla stokastisilla ADMM-menetelmillä. Aiheeseen liittyvä työ Kun y = Ax ∈R2, ongelmasta (1) tulee min x∈Rd1,y∈Rd2 f(x) + h(y), s.t. Ax- y = 0. (3) Vaikka (3) on vain yleisen ADMM-muodon (2) erikoistapaus, kun B = -Id2 ja c = 0, stokastiset (tai online) ADMM-algoritmit ja teoreettiset tulokset (Wang ja Banerjee 2012; Ouyang et al. 2013; Zhong ja Kwok 2014b; Zheng ja Kwok 2016) ja tämä työ ovat kaikki yleisempää ongelmaa (2) varten. (2) minimoimiseksi yhdessä kaksoismuuttujan λ kanssa erä-ADMM:n päivitysvaiheet ovat yk = argminy h(y) + β 2 ‖Axk-1+By-c+λk-1‖ , (4) xk = argminx f(x) + β 2 ‖Ax+Byk-c+λk-1‖ , (5) λk = λk-1 +Axk +Byk - c, (6) missä β>0 on rangaistusparametri. Jotta erä-ADMM voidaan laajentaa online- ja stokastisiin asetuksiin, yk:n ja λk:n päivitysvaiheet pysyvät muuttumattomina. Teoksessa (Wang ja Banerjee 2012; Ouyang et al. 2013) xk:n päivitysaskel approksimoidaan seuraavasti: xk = argmin x x x∇fik(xk-1) + 1 2ηk ‖x- xk-1‖G + β 2 ‖Ax+Byk-c+λk-1‖, (7) jossa arvotaan ik tasaisesti satunnaisesti valikosta [n] := {1, . . . . . , n}, ηk ∝ 1/ √ k on askelkoko, ja ‖z‖G = zGz, kun on annettu positiivinen puolimäärätön matriisi G, esimerkiksi G = Id1 (Ouyang et al. 2013). SGD:n tapaan stokastiset ADMM-muunnokset käyttävät jokaisella iteraatiokerralla gradientin puolueetonta estimaattia. Kaikkien näiden algoritmien konvergenssinopeus on kuitenkin paljon hitaampi kuin niiden eräkohtaisen vastineen, kuten edellä mainittiin. Tämä este johtuu pääasiassa gradienttien stokastisuuden aiheuttamasta varianssista. Lisäksi ne käyttävät konvergenssin takaamiseksi askelkoon ηk pienenevää sarjaa, mikä puolestaan vaikuttaa nopeuksiin. Viime aikoina on ehdotettu useita varianssia vähentäviä stokastisia ADMM-menetelmiä (esim. SAG-ADMM, SDCA-ADMM ja SVRG-ADMM), jotka ovat saavuttaneet jännittävää edistystä, kuten lineaarisen konvergenssinopeuden. SVRG-ADMM (Zheng ja Kwok 2016) on erityisen houkutteleva tässä yhteydessä, koska sen tallennustarve on pieni verrattuna algoritmeihin (Zhong ja Kwok 2014b; Suzuki 2014). SVRG-ADMM:n jokaisessa epookissa lasketaan ensin koko gradientti p̃ =∇f(x̃), jossa x̃ on edellisen epookin keskimääräinen piste. Tämän jälkeen ∇fik(xk-1) ja ηk kohdassa (7) korvataan ∇̃fIk(xk-1) = 1 |Ik| ∑ ik∈Ik (∇fik(xk-1)-∇fik(x̃))) + p̃ (8) ja vastaavasti vakioaskeleen η, jossa Ik ⊂ [n] on b-kokoinen minierä (mikä on hyödyllinen tekniikka varianssin pienentämiseksi). Itse asiassa ∇̃fIk(xk-1) on gradientin ∇f(xk-1) puolueeton estimaattori, eli E[∇̃fIk(xk-1)]=∇f(xk-1). Nopeutettu varianssia pienentävä stokastinen ADMM Tässä luvussa suunnitellaan nopeutettu varianssia pienentävä stokastinen ADMM-menetelmä sekä vahvasti koverille että yleisille koverille ongelmille. Aluksi tehdään seuraavat oletukset: Jokainen konveksinen fi(-) on Li-sileä, eli on olemassa vakio Li>0 siten, että ‖∇fi(x)-∇fi(y)‖≤Li‖x-y‖, ∀x, y ∈ R, ja L , maxi Li; f(-) on μ-vahvasti konveksinen, i.eli on olemassa μ > 0 siten, että f(x) ≥ f(y) +∇f(y)(x- y)+ μ2 ‖x-y‖ 2 kaikille x, y ∈R; Matriisilla A on täysi riviarvo. Kaksi ensimmäistä oletusta ovat yleisiä ensimmäisen kertaluvun optimointimenetelmien analyysissä, kun taas viimeisellä on Algoritmi 1 ASVRG-ADMM vahvasti konvekseille tapauksille Syöttö: m, η, β > 0, 1 ≤ b ≤ n. Alusta: x̃= z̃, ỹ, θ, λ̃=- 1 β (A T )∇f(x̃0). 1: s = 1, 2, . . . . , T do 2: x0 = z s 0 = x̃ s-1, y 0 = ỹ s-1, λ0 = λ̃ s-1; 3: p̃ = ∇f(x̃s-1); 4: for k = 1, 2, . . . . ,m do 5: Valitse Ik⊆ [n], jonka koko on b, tasaisesti satunnaisesti; 6: y k=argminy h(y)+ β 2 ‖Az s k-1+By- c+λk-1‖; 7: z k=z s k-1- η(∇̃fIk(x s k-1)+βA (Az k-1+By s k-c+λ s k-1)) γθ ; 8: xk=(1- θ)x̃s-1 + θz k; 9: λk=λ s k-1 +Az s k +By s k - c; 10: end for 11: x̃= 1 m ∑m k=1x s k, ỹ s=(1-θ)ỹs-1+ θ m ∑m k=1y s k, 12: λ̃=- 1 β (A T )†∇f(x̃s); 13: end for Tulos: x̃ , ỹ . on käytetty erä-ADMM:n (?; Nishihara et al. 2015; Deng ja Yin 2016) ja stokastisen ADMM:n (Zheng ja Kwok 2016) konvergenssianalyysissä. Vahvasti konveksinen tapaus Tässä osassa tarkastellaan (2) tapausta, jossa jokainen fi(-) on konveksinen,L-sileä ja f(-) on μ-vahvasti konveksinen. Muistutetaan, että tähän ongelmaluokkaan kuuluvat graafiohjattu logistinen regressio ja SVM huomattavina esimerkkeinä. Tämän ongelmaluokan ratkaisemiseksi tehokkaasti sisällytämme stokastiseen ADMM:ään sekä vauhtikiihdytys- että varianssin pienentämistekniikat. Algoritmimme on jaettu T epookkiin, ja kukin epookki koostuu m stokastisesta päivityksestä, jossa m valitaan yleensä O(n):ksi kuten (Johnson ja Zhang 2013). Olkoon z tärkeä apumuuttuja, jonka päivityssääntö on seuraava. Samoin kuin (Zhong ja Kwok 2014b; Zheng ja Kwok 2016), käytämme myös epätäsmällistä Uzawa-menetelmää (Zhang, Burger ja Osher 2011) approksimoidaksemme osaongelmaa (7), jolla voidaan välttää matriisin ( 1 η Id1+βA A) käänteislaskenta. Lisäksi momenttipaino 0≤ θs ≤ 1 (θs:n päivityssääntö esitetään jäljempänä) lisätään proksimaalitermiin 1 2η‖x-xk-1‖ 2 G samalla tavalla kuin (7), jolloin osaongelma z:n suhteen muotoillaan seuraavasti: min z (z -z k-1) ∇̃fIk(xk-1)+ θs-1 2η ‖z -z k-1‖G + β 2 ‖Az +By k - c+ λk-1‖, (9) missä ∇̃fIk(xk-1) on määritelty kohdassa (8), η < 1 2L , ja G = γId1- ηβ θs-1 AA, jossa γ ≥ γmin ≡ ηβ‖A A‖2 θs-1 +1 sen varmistamiseksi, että G I samanlainen kuin (Zheng ja Kwok 2016), jossa ‖-‖2 on spektrinormi, i.eli matriisin suurin singulaariarvo. Lisäksi x:n päivityssääntö on xk= x̃ +θs-1(z s k- x̃)=(1-θs-1)x̃+θs-1z k, (10) jossa θs-1(z k - x̃s-1) on keskeinen momenttitermi (samanlainen kuin kiihdytetyissä erämenetelmissä (Nesterov 2004)), jonka avulla algoritmimme kiihtyy käyttämällä edellisen epookin iteraattia, i.x̃s-1. Samoin kuin xk, ỹ s = (1- θs-1)ỹ s-1+ θs-1 m ∑m k=1y s k. Lisäksi θs voidaan asettaa vakioksi θ kaikissa algoritmimme epookeissa, jonka on täytettävä seuraavat vaatimukset: 0 ≤ θ ≤ 1- δ(b)/(α-1), missä α = 1 Lη > 1+ δ(b), ja δ(b) määritellään jäljempänä. θ:n optimaalinen arvo on esitetty seuraavassa lauseessa 1. Yksityiskohtainen menettely on esitetty algoritmissa 1, jossa käytämme λ̃:n osalta samaa alustusmenetelmää kuin (Zheng ja Kwok 2016), ja (-)† on pseudoinversio. Huomaa, että kun θ=1, ASVRG-ADMM degeneroituu SVRG-ADMM:ään (Zheng ja Kwok 2016). Ei-vahvasti konveksinen tapaus Tässä osassa tarkastelemme yleisiä konveksisia ongelmia, jotka ovat muotoa (2), kun kukin fi(-) on konveksinen, L-sileä ja h(-) ei välttämättä ole vahvasti konveksinen (mutta mahdollisesti ei-sileä). Vahvasti koverasta tapauksesta poiketen momenttipainon θs edellytetään täyttävän seuraavat epäyhtälöt: 1- θs θ2 s ≤ 1 θ2 s-1 ja 0 ≤ θs ≤ 1- δ(b) α- 1 , (11) missä δ(b) := n-b b(n-1) on pienen erän koon b suhteen pienenevä funktio. Ehto (20) sallii momenttipainon pienentyä, mutta ei liian nopeasti, mikä on samankaltainen kuin klassisen SGD:n ja stokastisen ADMM:n (?) askelmäärää ηk koskeva vaatimus. Toisin kuin eräkiihtyvyysmenetelmissä, painon on täytettävä molemmat ehdon (20) epäyhtälöt. Eräoptimoinnin momentumkiihdytystekniikoiden (Tseng 2010; Nesterov 2004) innoittamana annamme painon θs päivityssäännön minierän tapaukselle: θs = √ θ4 s-1+ 4θ 2 s-1 - θ s-1 2 ja θ0 = 1- δ(b) α- 1 . (12) Erikoistapauksessa b = 1 on δ(1) = 1 ja θ0 = 1- 1 α-1 , kun taas b = n (eli eräversio), δ(n)=0 ja θ0 = 1. Koska {θs} on vähenevä, θs ≤ 1- δ(b) α-1 täyttyy. Yksityiskohtainen menettely on esitetty algoritmissa 2, jossa on monia pieniä eroja algoritmiin 1 verrattuna kunkin epookin alustuksessa ja tulostuksessa. Lisäksi niiden keskeinen ero on momenttipainon θs päivityssääntö. Algoritmissa 1 θs voidaan asettaa vakioksi, kun taas algoritmissa 2 sitä säädetään adaptiivisesti (12) mukaisesti. Konvergenssianalyysi Tässä luvussa esitetään ASVRG-ADMM-algoritmien (eli algoritmien 1 ja 2) konvergenssianalyysi vahvasti koverille ja yleisesti koverille ongelmille. (Zheng ja Kwok 2016) mukaisesti esittelemme ensin seuraavan funktion P (x, y) := f(x)-f(x∗)-∇f(x∗)T(x- x∗)+h(y)-h(y∗)-h′(y∗)T(y- y∗) konvergenssikriteeriksi, missä h′(y) tarkoittaa h(-):n (ali)gradienttia pisteessä y. Todellakin, P (x, y)≥ 0 kaikille x, y ∈ R. Seuraavassa esitämme analyysimme keskeiset välitulokset. Algoritmi 2 ASVRG-ADMM yleisessä konveksisessa tapauksessa Syöttö: m, η, β > 0, 1 ≤ b ≤ n. Alusta: x̃ = z̃, ỹ, λ̃, θ0 = 1- Lηδ(b) 1-Lη . 1: s = 1, 2, . . . . , T do 2: x0=(1-θs-1)x̃+θs-1z̃, y 0= ỹs-1, λ0= λ̃s-1; 3: p̃ = ∇f(x̃s-1), z 0= z̃s-1; 4: for k = 1, 2, ... . ,m do 5: Valitse Ik⊆ [n], jonka koko on b, tasaisesti satunnaisesti; 6: y k=argminy h(y)+ β 2 ‖Az s k-1+By- c+λk-1‖; 7: z k=z s k-1- η(∇̃fIk(x s k-1)+βA (Az k-1+By s k-c+λ s k-1)) γθs-1 ; 8: xk=(1- θs-1)x̃ + θs-1z k; 9: λk=λ s k-1 +Az s k +By s k - c; 10: end for 11: x̃= 1 m ∑m k=1x s k, ỹ =(1-θs-1)ỹ+ θs-1 m ∑m k=1y s k, 12: λ̃=λm, z̃ =z m, θs= √ θ4 s-1+4θ 2 s-1-θ 2 s-1 2 ; 13: end for Tulos: x̃ , ỹ , . Lemma 1. E[‖∇̃fIk(xk-1)-∇f(xk-1)‖] ≤2Lδ(b) [ f(x̃)-f(xk-1)+(xk-1- x̃)∇f(xk-1) ] , missä δ(b)= n-b b(n-1)≤1 ja 1 ≤ b ≤ n. Lemma 2. Käytetään samaa merkintätapaa kuin Lemmassa 1. Olkoon (x∗, y∗, λ∗) ongelman (2) optimaalinen ratkaisu ja {(z k, xk, y k, λk, x̃, ỹ)} algoritmin 1 tai 2 tuottama sarja, jossa θs ≤ 1- δ(b) α-1, missä α= 1 Lη . Tällöin pätee seuraava kaikille k:lle, E [ P [ P (x̃, ỹ)- θs-1 m m ∑ k=1 ( (x∗-zs k)Aφk+(y-y k)Bφk )] ≤E [ P [ P (x̃s-1, ỹs-1) 1/(1-θs-1) + θ s-1 ( ‖x∗- z 0‖G-‖x- z m‖G )

**Tulos**

Nopeutettu varianssia vähentävä stokastinen ADMM (Accelerated Variance Reduced Stochastic ADMM)

**Esimerkki 2.1189**

Korkean tason reittisuunnittelussa ympäristöt mallinnetaan yleensä etäisyysgrafiikoiksi, ja reittisuunnitteluongelmat pelkistetään lyhimmän polun etsimiseen etäisyysgrafiikoissa. Yksi tämän mallinnuksen suurimmista haitoista on kyvyttömyys mallintaa epävarmuustekijöitä, joita käytännössä esiintyy kymmenen. Tässä artikkelissa ehdotetaan uutta työkalua, U-graafia, ympäristön mallintamiseen. U-graafi on etäisyysgraafien muunnos, joka pystyy käsittelemään eräänlaista epävarmuutta. Mallintamalla epävarmaa ympäristöä U-graafina ja navigointiongelmaa Markovin päätösprosessina voimme määritellä tarkasti uuden optimaalisuuskriteerin navigointisuunnitelmille ja, mikä tärkeämpää, voimme kehittää yleisen algoritmin navigointitehtävien optimaalisten suunnitelmien laskemiseen.

**Tulos**

Korkean tason reittisuunnittelu epävarmuuden kanssa

**Esimerkki 2.1190**

Monissa sovelluksissa tiedot esitetään luonnollisesti joidenkin peruselementtien tai symbolien järjestyksinä. Tällaisten tietojen päättely edellyttää samankaltaisuuden käsitettä, joka pystyy käsittelemään eripituisia sekvenssejä. Tässä artikkelissa kuvaamme Mercerin ydinfunktioiden perheen tällaiselle peräkkäisesti jäsennellylle tiedolle. Perheelle on ominaista symboli- ja rakennetason samankaltaisuuksien suhteen hajotettava rakenne, joka edustaa erityistä ytimien yhdistelmää, joka mahdollistaa tehokkaan laskennan. Esitämme kokeellisen arvioinnin sekvenssiluokittelutehtävissä, joissa verrataan ydinperheeseemme kuuluvia ytimiä uusimpaan sekvenssiytymisytimeen nimeltä Global Alignment kernel, jonka on osoitettu olevan parempi kuin Dynamic Time Warping.

**Tulos**

Sarjojen hajoavien ytimien perheestä Sarjojen hajoavien ytimien perheestä

**Esimerkki 2.1191**

Tässä artikkelissa esitellään reaktiiviset monikontekstijärjestelmät (rMCS), jotka ovat kehys reaktiiviselle päättelylle heterogeenisten tietolähteiden läsnä ollessa. Näytämme erityisesti, miten tietovirtoja voidaan integroida monikontekstijärjestelmiin ja miten järjestelmien dynamiikkaa voidaan mallintaa kahdenlaisten siltasääntöjen perusteella. Havainnollistamme, miten useita tyypillisiä tietovirtojen päättelyyn liittyviä ongelmia voidaan käsitellä kehyksemme avulla. Useisiin tietolähteisiin perustuvassa päättelyssä, jotka on integroitava, on ongelmana mahdolliset epäjohdonmukaisuudet. Keskustelemme erilaisista menetelmistä epäjohdonmukaisuuksien käsittelemiseksi, ja keskitymme erityisesti tasapainojen olemattomuuteen. Näytämme erityisesti, miten hallituille MCS-järjestelmille kehitetyt menetelmät voidaan yleistää rMCS-järjestelmiin. Tutkimme myös epämääräisyyttä rMCS-järjestelmissä. Yksi tapa välttää epädeterminismi on soveltaa vaihtoehtoista, skeptistä semantiikkaa. Näytämme, miten tällainen semantiikka, jota kutsutaan hyvin perustelluksi semantiikaksi, voidaan määritellä rMCS:ille ja mitä vaikutuksia on sillä, että tätä semantiikkaa käytetään alkuperäisen semantiikan sijasta. Tutkimme erilaisten rMCS:iin liittyvien päättelyongelmien monimutkaisuutta. Lopuksi keskustelemme aiheeseen liittyvistä töistä ja keskitymme erityisesti kahteen tärkeimpään virran päättelyyn liittyvään lähestymistapaan, nimittäin LARSiin ja STARQLiin.

**Tulos**

Reaktiiviset monikontekstiset järjestelmät: Heterogeeninen päättely dynaamisissa ympäristöissä.

**Esimerkki 2.1192**

Kahden aikaskaalan stokastisia approksimaatioalgoritmeja (SA) käytetään laajalti vahvistusoppimisessa (RL). Niiden iteraateissa on kaksi osaa, jotka päivitetään eri askeltiheyksillä. Tässä työssä tarjoamme reseptin kahden aikamittakaavan SA:n analysoimiseksi. Sen avulla kehitämme niille ensimmäisen konvergenssinopeuden tuloksen. Tästä tuloksesta poimimme keskeisiä oivalluksia askelkoon valinnasta. Sovelluksena saamme konvergenssiluvut kahden aikamittakaavan RL-algoritmeille, kuten GTD(0), GTD2 ja TDC.

**Tulos**

Kahden aikamittakaavan stokastisen approksimaation konvergenssiasteet ja sovelluksia vahvistusoppimiseen

**Esimerkki 2.1193**

Yksi lupaavimmista lähestymistavoista monimutkaisten teknisten järjestelmien analysointiin on ensemble-luokitusmenetelmät. Ensemble-menetelmien avulla voidaan rakentaa luotettavia päätöksentekosääntöjä ominaisuusavaruuden luokittelua varten, kun järjestelmässä on monia mahdollisia tiloja. Tässä asiakirjassa käytetään päätöspuihin perustuvia uusia tekniikoita sähkövoimajärjestelmien järjestelmän luotettavuuden arviointiin. Ehdotimme hybridilähestymistapaa, joka perustuu satunnaismetsämalleihin ja boosting-malleihin. Tällaisia tekniikoita voidaan soveltaa ennustamaan lisääntyvän uusiutuvan energian, kantaverkkolaitteiden ja älykkäiden kotitalouslaitteiden, varaajien ja ilmastointilaitteiden sekä sähköajoneuvojen älykkäiden kuormien vaihtamisen vuorovaikutusta verkon kanssa tehostettua päätöksentekoa varten. Ensemble-luokittelumenetelmiä testattiin muutetussa 118-väyläisessä IEEE-sähköjärjestelmässä, mikä osoittaa, että ehdotettua tekniikkaa voidaan käyttää sen tutkimiseen, onko sähköjärjestelmä turvattu vakaan tilan toimintaolosuhteissa.

**Tulos**

Ensemble-luokitusmenetelmät sähköjärjestelmien turvallisuuden arvioinnissa

**Esimerkki 2.1194**

Saavuttaakseen huippuluokan tuloksia visuaalisen näkemisen haasteissa konvolutiiviset neuroverkot oppivat paikallaan pysyviä suodattimia, jotka hyödyntävät taustalla olevaa kuvan rakennetta. Tarkoituksenamme on ehdottaa tehokasta kerrosmuotoilua, joka laajentaa tämän ominaisuuden mihin tahansa graafin kuvaamaan alueeseen. Käytämme nimittäin sen vierekkäisyysmatriisin tukea suunnitellaksemme opittavia painonjakosuodattimia, jotka pystyvät hyödyntämään signaalien taustalla olevaa rakennetta. Ehdotettu muotoilu mahdollistaa suodattimen painojen oppimisen sekä järjestelmän, joka ohjaa, miten ne jaetaan graafissa. Suoritamme validointikokeita kuvatietokannoilla ja osoitamme, että nämä suodattimet ovat suorituskyvyltään verrattavissa konvoluutiosuodattimiin.

**Tulos**

Paikallisten reseptivien kenttien oppiminen ja niiden painonjakojärjestelmä graafeissa

**Esimerkki 2.1195**

Tiivistelmä Lähimmän naapurin sääntö on yksi yleisimmin käytetyistä luokittelumalleista, ja prototyyppitapausten kompaktin joukon valitseminen on yksi sen sovellusten ensisijaisista haasteista. Monet nykyiset lähestymistavat prototyyppien valintaan hyödyntävät instanssipohjaisia analyysejä ja paikallisesti määriteltyjä luokkajakauman kriteerejä, jotka ovat hankalia numeerisille optimointitekniikoille. Tässä artikkelissa tutkimme parametrista kehystä, jossa on mukautettu lähimmän naapurin sääntö, jossa naapuriprototyyppien valintaa muutetaan niiden parametrien perusteella. Kehyksen avulla voimme muotoilla minimointiongelman, joka koskee mukautetun lähimmän naapurin säännön rikkomista harjoitusjoukon yli numeeristen parametrien suhteen. Osoitamme, että ongelma redusoituu suuren marginaalin periaatteelliseen oppimiseen, ja osoitamme sen edut empiirisillä vertailuilla viimeaikaisiin uusimpiin uusimpiin menetelmiin käyttäen julkista vertailuaineistoa.

**Tulos**

Prototyyppijoukon diskriminoiva oppiminen lähimmän naapurin luokittelua varten

**Esimerkki 2.1196**

Semanttisia jäsennysmenetelmiä käytetään tekstin semanttisen merkityksen tallentamiseen ja esittämiseen. Tekstin kaikki käsitteet kattava merkityksen esitys ei aina ole käytettävissä tai se ei välttämättä ole riittävän täydellinen. Ontologiat tarjoavat jäsennellyn ja päättelyyn kykenevän tavan mallintaa tekstikokoelman sisältöä. Tässä työssä esitellään uusi lähestymistapa ontologian ja semanttisen jäsentimen yhteiseen oppimiseen tekstistä. Menetelmä perustuu kontekstittoman kieliopin puoliautomaattiseen induktioon semanttisesti annotoidusta tekstistä. Kielioppi jäsentää tekstin semanttisiksi puiksi. Sekä kielioppia että semanttisia puita käytetään ontologian oppimiseen useilla tasoilla - luokat, instanssit, taksonomiset ja ei-taksonomiset suhteet. Lähestymistapaa arvioitiin Wikipedian henkilöitä kuvaavien sivujen ensimmäisillä lauseilla.

**Tulos**

Ontologian ja semanttisen jäsentimen yhteinen oppiminen tekstistä

**Esimerkki 2.1197**

Seuraavassa biologisesti inspiroituneessa järjestelmässä todettiin, että ihmisen toiminnan tunnistaminen nisäkkäiden aivoissa johtaa mallissa kahteen eri reittiin, jotka ovat erikoistuneet liikkeen (optisen virtauksen) ja muototiedon analysointiin. Periaatteessa olemme määritelleet uudenlaiset ja vankat muoto-ominaisuudet, joissa käytetään aktiivista perusmallia muodon poimijana biologisesti inspiroituneen mallin muotopolulla. Epätasapainoinen synerginen neuroverkko luokittelee ihmisen esineiden muodot ja rakenteet sekä virittää sen huomioparametria kvanttihiukkasparven optimoinnilla (QPSO) käynnistämällä Centroidal Voronoi Tessellations. Näitä työkaluja hyödynnetään ja perustellaan vahvoina työkaluina biologisen järjestelmän mallin seuraamiseksi muotopolulla. Lopullinen päätös on kuitenkin tehty yhdistämällä molempien polkujen lopulliset tulokset sumean päättelyn avulla, mikä lisää ehdotetun mallin uutuutta. Näiden kahden aivoreitin yhdistäminen tehdään ottamalla huomioon kukin ominaisuusjoukko Gaussin jäsenyysfunktioissa sumean tuotteen päättelymenetelmän avulla. Aivopolulle on ehdotettu kahta konfiguraatiota: moniprototyyppisten ihmisen toimintamallien soveltaminen käyttämällä kahden ajan synergistä neuroverkkoa yhtenäisen mallin saamiseksi kunkin toiminnon osalta, ja toinen skenaario, jossa käytetään ihmisen toiminnan abstrahointia neljään avainkehykseen. Kokeelliset tulokset osoittivat lupaavaa tarkkuutta eri tietokokonaisuuksissa (KTH ja Weizmann).

**Tulos**

Biologisesti inspiroituneen ihmisen toiminnan tunnistaminen käyttäen hybridi Max-Product Neuro-Fuzzy-luokittelijaa ja Quantum-Behaved PSO:ta.

**Esimerkki 2.1198**

Esittelemme mallin, joka oppii aktiivisia oppimisalgoritmeja metallioppimisen avulla. Mallimme oppii yhdessä tietylle joukolle toisiinsa liittyviä tehtäviä: datan esityksen, kohteiden valintaheuristiikan ja menetelmän ennustinfunktioiden rakentamiseksi merkityistä harjoitusjoukoista. Mallimme käyttää kohteiden valintaheuristiikkaa kerätäkseen merkittyjä harjoitusjoukkoja, joista voidaan rakentaa ennustinfunktioita. Omniglot- ja MovieLens-tietokantojen avulla testaamme malliamme synteettisissä ja käytännön tilanteissa.

**Tulos**

Aktiivisen oppimisen oppimisalgoritmit

**Esimerkki 2.1199**

Satunnaistetut matriisipakkaustekniikat, kuten Johnson-Lindenstrauss-muunnos, ovat osoittautuneet tehokkaaksi ja käytännölliseksi tavaksi ratkaista suuria ongelmia tehokkaasti. Kun keskitytään laskennalliseen tehokkuuteen, on kuitenkin tehtävä kompromissi ratkaisujen laadusta ja tarkkuudesta. Tässä artikkelissa tutkimme pakattuja pienimmän neliösumman ongelmia ja ehdotamme uusia malleja ja algoritmeja, jotka käsittelevät pakkauksen aiheuttamia virheitä ja kohinaa. Samalla kun mallimme säilyttävät laskentatehokkuuden, ne tarjoavat vankkoja ratkaisuja, jotka ovat tarkempia - suhteessa pakkaamattomien pienimmän neliösumman ongelmien ratkaisuihin - kuin klassisten pakattujen vaihtoehtojen ratkaisut. Otamme käyttöön robustin optimoinnin työkaluja yhdessä osittaisen pakkauksen kanssa, jotta voimme parantaa pakattujen pienimmän neliösumman ratkaisujen virhe-aika-sopimuksia. Kehitämme tehokkaan ratkaisualgoritmin robustille osittain pakatulle (RPC) mallillemme, joka perustuu pelkistämiseen yksiulotteiseen hakuun. Johdamme myös ensimmäiset approksimaatiovirherajat osittain pakatuille pienimmän neliösumman ratkaisuille. Empiiriset tulokset, joissa verrataan lukuisia vaihtoehtoja, viittaavat siihen, että robustit ja osittain pakatut ratkaisut ovat tehokkaasti suojattuja aggressiivisia satunnaistettuja muunnoksia vastaan.

**Tulos**

Vankka osittain pakattu Least-Squares-menetelmä (osittain pakattu)

**Esimerkki 2.1200**

Ehdotamme satunnaismetsän (RF) karsimista resurssirajoitteista ennustamista varten. Rakennamme ensin RF:n ja karsimme sen sitten optimoidaksemme odotetut ominaisuuksien kustannukset ja tarkkuuden. Esitämme metsän karsimisen uudenlaisena 0-1 kokonaislukuohjelmana, jossa on lineaarisia rajoitteita, jotka kannustavat ominaisuuksien uudelleenkäyttöä. Todistamme, että rajoitusjoukko on täysin unimodulaarinen ja että vastaava LP-relaksaatio ratkaisee alkuperäisen kokonaislukuohjelman. Sen jälkeen hyödynnämme yhteyksiä kombinatoriseen optimointiin ja kehitämme tehokkaan primaari-duaalialgoritmin, joka skaalautuu suurille tietokokonaisuuksille. Toisin kuin meidän alhaalta ylöspäin suuntautuva lähestymistapamme, joka hyötyy hyvästä RF-alkuistutuksesta, tavanomaiset menetelmät ovat ylhäältä alaspäin suuntautuvia ja hankkivat piirteitä niiden hyötyarvon perusteella, ja se on yleensä vaikeasti ratkaistavissa ja vaatii heuristiikkoja. Empiirisesti karsinta-algoritmimme on parempi kuin nykyiset resurssirajoitteiset algoritmit.

**Tulos**

Satunnaismetsiköiden karsiminen ennusteiden laatimiseksi pienellä budjetilla

**Esimerkki 2.1201**

Mielipidetekstien luokittelusta positiivisiin ja negatiivisiin teksteihin on tulossa suuri kiinnostuksen kohde sentimenttianalyysissä. Monien leimattujen mielipiteiden olemassaolo motivoi tilastollisten ja koneoppimismenetelmien käyttöä. Ensimmäisen asteen tilastot ovat osoittautuneet hyvin rajallisiksi tällä alalla. Opinum-lähestymistapa perustuu sanojen järjestykseen käyttämättä syntaktista ja semanttista tietoa. Se koostuu yhden todennäköisyysmallin rakentamisesta myönteisille ja toisen kielteisille mielipiteille. Sitten testimielipiteitä verrataan molempiin malleihin ja lasketaan päätös ja luottamusaste. Harjoitusjoukon monimutkaisuuden vähentämiseksi tekstit lemmatoidaan ensin ja suurin osa nimityksistä korvataan jokerimerkillä. Opinum tuottaa yli 81 prosentin tarkkuuden espanjalaisten mielipiteiden osalta rahoitustuotteiden alalla. Tässä työssä käsitellään tärkeimpiä tekijöitä, jotka vaikuttavat luokittelutulokseen.

**Tulos**

Tilastollinen tunneanalyysi Opinumissa

**Esimerkki 2.1202**

Ennustemarkkinat ovat hyödyllinen keino koota yhteen tietoa tulevasta tapahtumasta. Toimiakseen markkinat tarvitsevat luotettavan tahon, joka lopulta varmistaa todellisen lopputuloksen. Hajautettujen ennustemarkkinoiden hiljattaisen käyttöönoton innoittamana esittelemme mekanismin, jonka avulla lopputulos voidaan määrittää sellaisten välimiesten äänillä, joilla voi itselläänkin olla osuuksia markkinoilla. Mahdollisesta eturistiriidasta huolimatta johdamme ehdot, joilla voimme kannustaa välimiehiä äänestämään totuudenmukaisesti käyttämällä markkinamaksuista kerättyjä varoja vertaisennustusmekanismin toteuttamiseen. Lopuksi tutkimme, mitä parametriarvoja voitaisiin käyttää mekanismimme todellisessa toteutuksessa.

**Tulos**

Joukkoistettu tulosten määrittäminen ennustemarkkinoilla

**Esimerkki 2.1203**

Tässä artikkelissa käsitellään sitä, miten SP-älykkyysteoriaa ja sen toteuttamista SP-koneessa voidaan eduksi soveltaa suurten tietojen hallintaan ja analysointiin. Artikkelissa esitelty ja muualla kokonaisuudessaan kuvattu SP-järjestelmä voi auttaa ratkaisemaan big datan monimuotoisuuden ongelman: sillä on potentiaalia universaalina kehyksenä erilaisten tietämystyyppien (UFK) esittämiseen ja käsittelyyn, mikä auttaa vähentämään tietämyksen formalismien ja formaattien moninaisuutta ja erilaisia tapoja, joilla niitä käsitellään. Sen vahvuuksia ovat muun muassa valvomaton oppiminen tai rakenteen löytäminen datasta, hahmontunnistus, luonnollisen kielen jäsentäminen ja tuottaminen sekä monenlainen päättely. Se soveltuu hyvin suoratoistodatan analysointiin, mikä auttaa ratkaisemaan suurten datamäärien nopeusongelman. Keskeistä järjestelmän toiminnassa on tiedon häviötön pakkaaminen: se pienentää isoja tietoja ja vähentää tallennus- ja hallintaongelmia. Tiedonsiirrossa voidaan saavuttaa huomattavia säästöjä, laskennan energiankäyttöä voidaan vähentää huomattavasti, käsittelyä voidaan nopeuttaa ja tietokoneita voidaan pienentää ja keventää. Järjestelmän avulla voidaan ratkaista suurten datamäärien todenperäisyysongelma, ja sen avulla voidaan auttaa hallitsemaan tietojen virheitä ja epävarmuustekijöitä. Se soveltuu hyvin tietorakenteiden ja päätelmäprosessien visualisointiin. Korkean rinnakkaistason avoimen lähdekoodin versio SP-koneesta tarjoaisi tutkijoille kaikkialla maailmassa mahdollisuuden tutkia, mitä järjestelmällä voidaan tehdä, ja luoda siitä uusia versioita.

**Tulos**

Big Data ja SP-teoria älykkyydestä

**Esimerkki 2.1204**

Retorinen rakennepuu (RS-puu) on esitys diskurssin alkeisyksiköiden (EDU) välisistä diskurssisuhteista. RS-puu on erittäin hyödyllinen monissa tekstinkäsittelytehtävissä, joissa käytetään EDU:iden välisiä suhteita, kuten tekstin ymmärtämisessä, tiivistämisessä ja kysymyksiin vastaamisessa. Thaimaalainen kieli ja sen ainutlaatuiset kielelliset piirteet edellyttävät ainutlaatuista RS-puun rakentamistekniikkaa. Tässä artikkelissa ehdotetaan thaimaalaisen RS-puun rakentamiseen lähestymistapaa, joka koostuu kolmesta päävaiheesta: EDU:n segmentointi, thaimaalaisen RS-puun rakentaminen ja diskurssisuhteiden (DR) tunnistaminen. Kieliopillisista säännöistä johdettuja kahta piilomarkovin mallia käytetään EDU:iden segmentointiin, thaimaalaisten semanttisten sääntöjen perusteella johdettua klusterointitekniikkaa ja sen kaltaisuusmittaa käytetään thaimaalaisen RS-puun rakentamiseen, ja EDU:iden välisten DR:iden määrittämiseen käytetään päätöspuuta, jonka piirteet on poimittu säännöistä. Ehdotettua tekniikkaa arvioidaan käyttämällä kolmea thaimaalaista korporaatiota. Tulokset osoittavat, että thaimaalaisen RS-puun rakentamisen tehokkuus on 94,90 prosenttia ja DR:n tunnistamisen tehokkuus 82,81 prosenttia. AvainsanatThain kieli, elementtidiskurssiyksikkö, retorinen rakennepuu, diskurssisuhde.

**Tulos**

THAIMAALAINEN RETORINEN RAKENNEANALYYSI

**Esimerkki 2.1205**

L'articolo presenta un'introduzione all'Intelligenza Artificiale (IA) in forma divulgativa e informale ma precisa. L'articolo affronta prevalentemente gli aspetti informatici della disciplina, presentando varie tecniche usate nei sistemi di IA and divendole in simboliche e subsimboliche. L'articolo si conclude presentando il dibattito in corso sull'IA e in particolare sui vantaggi e i pericoli che sono stati individuati, terminalando con l'opinion dell'autore al riguardo.

**Tulos**

Introduzione all'Intelligenza Artificiale∗∗.

**Esimerkki 2.1206**

Suunnitelmien tunnistamiseen epävarmuuden vallitessa käytettävät tekniikat edellyttävät stokastista mallia suunnitelman laatimisprosessista. Otamme käyttöön todennäköisyyteen perustuvia tilariippuvaisia kielioppeja (PSDG) edustamaan agentin suunnitelmanmuodostusprosessia. PSDG-kielimalli laajentaa probabilistisia kontekstivapaita kielioppeja (PCFG) sallimalla tuotantotodennäköisyyksien riippuvuuden suunnitteluagentin sisäisen ja ulkoisen tilan eksplisiittisestä mallista. Kun suunnitelman tuottamisprosessin PSDG-kuvaus on annettu, voimme käyttää päättelyalgoritmeja, jotka hyödyntävät PSDG-kielen erityisiä riippumattomuuden oikeita sidoksia vastataksemme tehokkaasti suunnitelman tunnistuskyselyihin. PSDG-kielen mallin ja päättelyalgoritmien yhdistelmä laajentaa niiden suunnitelmien tunnistamisalueiden valikoimaa, joilla käytännön todennäköisyyspohjainen päättely on mahdollista, kuten liikennevalvonnan ja ilmataistelun sovellukset osoittavat.

**Tulos**

Todennäköisyysriippuvaiset kieliopit suunnitelmien tunnistamista varten

**Esimerkki 2.1207**

Esittelemme uudenlaisen neuroverkkomallin, joka oppii POS-tunnistusta ja graafipohjaista riippuvuuksien jäsentelyä yhdessä. Mallimme käyttää kaksisuuntaisia LSTM-malleja oppiakseen piirre-edustuksia, jotka ovat yhteisiä sekä POS-tunnistamis- että riippuvuuksien jäsentämistehtäviä varten, ja siten käsittelemme piirre-insinööriongelmaa. Laajat kokeilumme Universal Dependencies -hankkeen 19 kielellä osoittavat, että mallimme päihittää uusimman neuroverkkopohjaisen pino-propagointimallin POS-tunnistamisen ja siirtymäpohjaisen riippuvuuksien jäsennyksen yhteistoiminnalle, mikä on uutta tekniikan tasoa. Koodimme on avointa lähdekoodia ja saatavilla osoitteessa: https://github.com/ datquocnguyen/jPTDP.

**Tulos**

Uusi neuroverkkomalli POS-taggauksen ja graafipohjaisen riippuvuuksien jäsennyksen yhteiseen käsittelyyn.

**Esimerkki 2.1208**

Luonnollisten kuvien luokilla on taipumus noudattaa pitkän hännän jakaumia. Tämä on ongelmallista, kun harvinaisia luokkia varten ei ole riittävästi harjoitusesimerkkejä. Tämä vaikutus korostuu yhdistelmäluokissa, joissa yhdistyvät useat käsitteet, kuten toimintojen tunnistamiseen liittyvissä tietokokonaisuuksissa. Tässä artikkelissa ehdotamme tämän ongelman ratkaisemista oppimalla hyödyntämään yleisiä visuaalisia käsitteitä, joita on helposti saatavilla. Tunnistamme kuvissa esiintyvät näkyvät käsitteet ja käytämme niitä kohdetarrojen päättelyyn sen sijaan, että käyttäisimme suoraan visuaalisia piirteitä, ja yhdistelemme näin visuaalisen näkemisen ja luonnollisen kielen käsittelyn työkaluja. Validoimme menetelmämme hiljattain käyttöönotetulla HICO-tietokannalla, jonka mAP on 31,54 %, ja Stanford40 Actions -tietokannalla, jossa ehdotettu menetelmä on nykyistä tekniikan tasoa parempi ja jossa yhdistettynä suoriin visuaalisiin piirteisiin saavutetaan tarkkuus 83,12 %. Lisäksi menetelmä tarjoaa kullekin luokalle semanttisesti merkityksellisen luettelon avainsanoista ja merkityksellisistä kuva-alueista, jotka liittyvät luokkaan kuuluviin käsitteisiin.

**Tulos**

Toimintojen luokittelu käsitteiden ja attribuuttien avulla

**Esimerkki 2.1209**

Moniin luokitusongelmiin liittyy toisiinsa linkitettyjä tietoinstansseja, kuten hyperlinkkien yhdistämiä verkkosivuja. Kollektiivisen luokittelun (CC) tekniikat parantavat usein tarkkuutta tällaisissa datagraafeissa, mutta vaativat yleensä täysin merkityn harjoitusgraafin. Sitä vastoin tutkimme, miten CC-mallien puolivalvottua oppimista voidaan parantaa, kun käytössä on vain harvoin merkitty graafi, mikä on yleinen tilanne. Aluksi kuvaamme, miten voidaan käyttää uusia luokittelijoiden yhdistelmiä, jotta voidaan hyödyntää relaatiopiirteiden ja ei-relaatiopiirteiden erilaisia ominaisuuksia. Laajennamme myös etikettisäännöstelyn ideat tällaisiin hybridiluokittelijoihin, jolloin ne voivat hyödyntää merkitsemätöntä dataa oppimisprosessin vääristämiseksi. Huomaamme, että nämä tekniikat, jotka ovat tehokkaita ja helppoja toteuttaa, parantavat merkittävästi tarkkuutta kolmessa todellisessa tietokokonaisuudessa. Lisäksi tuloksemme selittävät aiempien aiheeseen liittyvien tutkimusten ristiriitaisia havaintoja.

**Tulos**

Puolivalvottu kollektiivinen luokittelu hybridi-merkintäsäännöstelyn avulla

**Esimerkki 2.1210**

Esittelemme konvergoituneen algoritmin Tikhonovin regularisoidulle ei-negatiivisen matriisin faktoroinnille (NMF). Valitsimme erityisesti tämän regularisoinnin, koska tiedetään, että Tikhonovin regularisoitu pienin neliö (LS) on perinteistä LS:ää suositeltavampi muoto lineaaristen käänteisongelmien ratkaisemisessa. Koska NMF-ongelma voidaan purkaa LS-osaongelmiksi, voidaan odottaa, että Tikhonovin regularisoitu NMF on sopivampi lähestymistapa NMF-ongelmien ratkaisemiseen. Algoritmi on johdettu käyttämällä additiivisia päivityssääntöjä, joilla on osoitettu olevan konvergenssitakuu. Varustamme algoritmin mekanismilla, jolla regularisointiparametrit määritetään automaattisesti L-käyrän perusteella, mikä on tunnettu käsite käänteisongelmien yhteisössä, mutta melko tuntematon NMF-tutkimuksessa. Tämän algoritmin käyttöönotto ratkaisee siten kaksi Tikhonov-regularisoitujen NMF-algoritmien tutkimukseen liittyvää ongelmaa, eli konvergenssitakuun ja regularisointiparametrien määrittämisen.

**Tulos**

Konvergoitunut algoritmi Tikhonov-säännösteltyyn ei-negatiiviseen matriisifaktorointiin automaattisella säännöstelyparametrien määrittämisellä

**Esimerkki 2.1211**

Esittelemme T-LESSin, uuden julkisen tietokokonaisuuden tekstuurittomien jäykkien objektien 6D-asennon, eli translaation ja rotaation, arvioimiseksi. Tietokanta sisältää kolmekymmentä teollisuuden kannalta merkityksellistä objektia, joilla ei ole merkittävää tekstuuria eikä erottelevia väri- tai heijastusominaisuuksia. Objekteilla on symmetriaa ja keskinäisiä samankaltaisuuksia muodon ja/tai koon suhteen. Muihin tietokokonaisuuksiin verrattuna ainutlaatuinen ominaisuus on se, että osa objekteista on toisten osia. Tietokanta sisältää harjoitus- ja testikuvia, jotka on otettu kolmella synkronoidulla anturilla, erityisesti strukturoidun valon ja valoaikaisen RGB-D-anturin sekä korkean resoluution RGB-kameran avulla. Kustakin sensorista on noin 39K harjoitus- ja 10K testikuvaa. Lisäksi kustakin kohteesta on kaksi erilaista 3D-mallia, käsin luotu CAD-malli ja puoliautomaattisesti rekonstruoitu malli. Harjoituskuvat kuvaavat yksittäisiä esineitä mustaa taustaa vasten. Testikuvat ovat peräisin kahdestakymmenestä testikohtauksesta, joiden monimutkaisuus vaihtelee yksinkertaisista kohtauksista, joissa on useita yksittäisiä esineitä, erittäin haastaviin kohtauksiin, joissa on useita eri esineitä ja joissa on paljon epäselvyyttä ja peittävyyttä. Kuvat otettiin systemaattisesti näytteistetyltä näkymäpallolta kohteen/näyttämön ympäriltä, ja niihin on merkitty kaikkien mallinnettujen kohteiden tarkat 6D-poseeraukset. Alustavat arviointitulokset osoittavat, että 6D-objektien asentojen estimoinnissa on paljon parantamisen varaa, erityisesti vaikeissa tapauksissa, joissa on paljon peittävyyttä. T-LESS-tietokanta on saatavilla verkossa osoitteessa cmp.felk.cvut.cz/t-less.

**Tulos**

T-LESS: RGB-D-tietokanta tekstuurittomien kohteiden 6D-poseerauksen arvioimiseksi.

**Esimerkki 2.1212**

Kyky ymmärtää automaattisesti ja nopeasti käyttäjäkonteksti istunnon aikana on yksi tärkeimmistä ongelmista suosittelujärjestelmissä. Ensimmäisenä askeleena kohti tätä<lb>tavoitetta ehdotamme mallia, joka havainnoi reaaliaikaisesti kunkin kohteen tuoman<lb>moninaisuuden suhteessa lyhyeen sarjaan<lb>konsultaatioita, jotka vastaavat käyttäjän viimeaikaista historiaa. Mallimme<lb>monimutkaisuus on vakioaikaa, ja se on yleinen, koska<lb>se soveltuu mihin tahansa verkkopalvelun kohdetyyppiin (esim. <lb>profiilit, tuotteet, musiikkikappaleet) ja mihin tahansa sovellusalueeseen (sähköinen<lb>kauppa, sosiaalinen verkosto, musiikin suoratoisto), kunhan meillä on<lb>osittaiset kohdekuvaukset. Havainnoimalla monimuotoisuuden tasoa<lb>aikojen kuluessa voimme havaita implisiittiset muutokset. Pitkällä aikavälillä<lb>suunnittelemme kontekstin karakterisointia eli yhteisten piirteiden<lb>löytämistä<lb>yhteenkuuluvasta kohteiden osajaksosta kahden mallimme määrittämän kontekstin muutoksen<lb> välillä. Näin voimme<lb>tehdä kontekstin huomioon ottavia ja yksityisyyden suojaa kunnioittavia suosituksia,<lb>selittää niitä käyttäjille. Koska kyseessä on meneillään oleva tutkimus,<lb>ensimmäinen vaihe on tutkia mallimme kestävyyttä<lb>yhteyden muutosten havaitsemisessa. Tätä varten käytämme<lb>musiikkikorpusta, jossa on 100 käyttäjää ja yli 210 000 konsultaatiota<lb>(soitettujen kappaleiden määrä globaalissa historiassa). Validoimme<lb>havaintojemme relevanssia löytämällä yhteyksiä<lb>kontekstin muutosten ja tapahtumien, kuten istunnon päättymisen, välillä. Nämä tapahtumat ovat tietysti osajoukko<lb> mahdollisista kontekstin muutoksista, koska istunnon sisällä voi olla useita konteksteja.<lb>Muutimme korpuksemme laatua useilla tavoilla testataksemme mallimme suorituskykyä<lb>harvinaisuuden ja erityyppisten kohteiden<lb>suhteissa. Tulokset osoittavat, että<lb>mallimme on kestävä ja lupaava lähestymistapa. Avainsanat-käyttäjämallinnus; monimuotoisuus; konteksti; reaaliaikainen<lb>navigointipolun analyysi; suosittelujärjestelmät.

**Tulos**

Kohti kestävää monimuotoisuuteen perustuvaa mallia kontekstin muutosten havaitsemiseksi.

**Esimerkki 2.1213**

Tässä artikkelissa esitellään uusi viestintätehokas rinnakkainen uskomuspropagointialgoritmi (CE-PBP) latentin Dirichlet-allokaation (LDA) kouluttamiseen. Synkronisen uskomusten etenemisalgoritmin (BP) pohjalta kehitämme ensin rinnakkaisen uskomusten etenemisalgoritmin (PBP) rinnakkaisarkkitehtuurille. Koska laaja tiedonsiirtoviive aiheuttaa usein alhaisen tehokkuuden rinnakkaisessa aihepiirien mallintamisessa, käytämme lisäksi Zipfin lakia vähentämään kokonaisviestintäkustannuksia PBP:ssä. Laajat kokeet eri aineistoilla osoittavat, että CE-PBP:llä saavutetaan korkeampi aiheiden mallinnustarkkuus ja vähennetään viestintäkustannuksia yli 80 prosenttia verrattuna uusimpaan rinnakkaiseen Gibbs sampling (PGS) -algoritmiin.

**Tulos**

Tietoliikennetehokas rinnakkainen uskomuksen eteneminen latenttia Dirichlet-allokaatiota varten

**Esimerkki 2.1214**

Monimuuttujaisten aikasarjojen ennustaminen on tärkeä koneoppimisen ongelma monilla aloilla, kuten aurinkovoimaloiden energiantuoton, sähkönkulutuksen ja liikenneruuhkatilanteen ennustamisessa. Näissä reaalimaailman sovelluksissa syntyviin aikatietoihin sisältyy usein sekoitus pitkän ja lyhyen aikavälin malleja, joihin perinteiset lähestymistavat, kuten autoregressiiviset mallit ja Gaussin prosessi, voivat epäonnistua. Tässä asiakirjassa ehdotimme uutta syväoppimiskehystä, nimittäin Longand Short-term Time-series network (LSTNet), vastaamaan tähän avoimeen haasteeseen. LSTNet käyttää konvoluutiohermoverkkoa (Convolution Neural Network, CNN) poimimaan muuttujien välisiä lyhyen aikavälin paikallisia riippuvuusmalleja ja rekursiivista hermoverkkoa (Recurrent Neural Network, RNN) löytämään pitkän aikavälin malleja ja suuntauksia. Arvioinnissamme reaalimaailman aineistolla, jossa oli monimutkaisia toistuvien kuvioiden sekoituksia, LSTNet saavutti merkittäviä parannuksia suorituskykyyn verrattuna useisiin uusimpiin perusmenetelmiin. Tietoaineisto ja kokeen koodi on ladattu Githubiin.

**Tulos**

Pitkän ja lyhyen aikavälin ajallisten kuvioiden mallintaminen syvien neuroverkkojen avulla

**Esimerkki 2.1215**

Tässä työssä esittelemme ehdollisen nopeutetun laiskan stokastisen gradientin laskeutumisalgoritmin, jolla on optimaalinen määrä kutsuja stokastiseen ensimmäisen järjestyksen oraakkeliin ja jonka konvergenssinopeus on O( 1 ε2 ) ja joka on parempi kuin Hazanin ja Kalen [2012] projektiovapaa, Online Frank-Wolfe -pohjainen stokastinen gradientin laskeutumisalgoritmi, jonka konvergenssinopeus on O( 1 ε4 ).

**Tulos**

Ehdollinen nopeutettu laiska stokastinen asteittainen laskeutuminen (Conditional Accelerated Lazy Stochastic Gradient Descent)

**Esimerkki 2.1216**

Reunamerkkien ennustamisen ongelmassa meille annetaan suunnattu graafi (joka edustaa sosiaalista verkostoa), ja tehtävämme on ennustaa reunojen binääriset merkinnät (eli sosiaalisten suhteiden positiivinen tai negatiivinen luonne). Monet menestyksekkäät heuristiikat tähän ongelmaan perustuvat trolli-luottamusominaisuuksiin, joilla arvioidaan jokaisessa solmussa lähtevien ja tulevien positiivisten/negatiivisten reunojen osuus. Osoitamme, että nämä heuristiikat voidaan ymmärtää ja analysoida tiukasti Bayesin optimaalisen luokittelijan approksimaattoreina yksinkertaiselle todennäköisyysmallille reunamerkinnöistä. Tämän jälkeen osoitamme, että tämän mallin suurimman todennäköisyyden estimaattori vastaa suunnilleen alkuperäisen sosiaalisen graafin muunnetun version Label Propagation -algoritmin ennusteita. Laajat kokeet useilla reaalimaailman tietokokonaisuuksilla osoittavat, että tämä algoritmi on kilpailukykyinen uusimpiin luokittelijoihin nähden sekä tarkkuuden että skaalautuvuuden osalta. Lopuksi osoitamme, että trolltrust-ominaisuuksien avulla voidaan myös johtaa online-oppimisalgoritmeja, joilla on teoreettiset takeet myös silloin, kun reunat on merkitty vastakkaisesti.

**Tulos**

Troll-Trust-malli reunamerkkien ennustamiseen sosiaalisissa verkostoissa

**Esimerkki 2.1217**

Tutkimme vastauksen valintaa usean käännöksen keskustelua varten hakupohjaisissa chat-roboteissa. Nykyiset teokset joko jättävät huomiotta lausumien väliset suhteet tai jättävät huomiotta tärkeän kontekstiin liittyvän tiedon, kun vastausta sovitetaan yhteen hyvin abstraktin kontekstivektorin kanssa. Ehdotamme uutta istuntopohjaista sovitusmallia molempien ongelmien ratkaisemiseksi. Malli sovittaa ensin vastauksen kuhunkin lausahdukseen useilla eri rakeisuusasteilla ja erottelee jokaisesta parista tärkeän sovitustiedon vektorina konvoluutio- ja yhdistämisoperaatioiden avulla. Tämän jälkeen vektorit kerätään kronologiseen järjestykseen rekursiivisen neuroverkon (RNN) avulla, joka mallintaa lausumien välisiä suhteita. Lopullinen vastaavuuspistemäärä lasketaan RNN:n piilotettujen tilojen avulla. Empiirinen tutkimus kahdella julkisella aineistolla osoittaa, että mallimme voi olla huomattavasti parempi kuin nykyaikaiset menetelmät vastauksen valinnassa monikielisessä keskustelussa.

**Tulos**

Sequential Match Network: Uusi arkkitehtuuri monikierroksisen vastauksen valintaan hakupohjaisissa chat-roboteissa.

**Esimerkki 2.1218**

Potilaiden aikasarjojen luokitteluun liittyy haasteita, jotka liittyvät suureen ulottuvuuteen ja puuttuvuuteen. Potilaiden samankaltaisuusteorian valossa tässä tutkimuksessa tutkitaan tehokkaita ajallisten ominaisuuksien suunnittelua ja vähentämistä, puuttuvien arvojen imputointia ja muutospisteiden havaitsemismenetelmiä, joilla voidaan tarjota samankaltaisuuteen perustuvia luokittelumalleja, joiden tarkkuus paranee toivotulla tavalla. Valitsemme kappaleittaisen aggregaatioapproksimaatiomenetelmän hienojakoisten ajallisten piirteiden poimimiseksi ja ehdotamme minimalistista menetelmää puuttuvien arvojen imputoimiseksi ajallisissa piirteissä. Dimensioiden vähentämiseksi otamme käyttöön gradienttihakumenetelmän ominaisuuksien painojen määrittämiseksi. Ehdotamme uusia potilaan tilan ja suunnanmuutoksen määritelmiä, jotka perustuvat lääketieteelliseen tietämykseen tai kliinisiin ohjeisiin potilaan eri tilatasojen arvoalueista, ja kehitämme menetelmän, jolla havaitaan positiivisia tai negatiivisia potilaan tilanmuutoksia osoittavat muutospisteet. Arvioimme ehdotettujen menetelmien tehokkuutta varhaisen teho-osastokuolleisuuden ennustamisen yhteydessä. Arviointitulokset osoittavat, että k-Nearest Neighbor -algoritmi, joka sisältää valitsemamme ja ehdottamamme menetelmät, on huomattavasti parempi kuin asiaankuuluvat vertailuarvot teho-osastokuolleisuuden varhaisessa ennustamisessa. Tämä tutkimus edistää aikasarjaluokitusta ja teho-osastokuolleisuuden varhaista ennustamista yksilöimällä ja parantamalla ajallisten ominaisuuksien suunnittelu- ja vähentämismenetelmiä samankaltaisuuteen perustuvaa aikasarjaluokitusta varten.

**Tulos**

Aikasarjadatan hyödyntäminen samankaltaisuuteen perustuvissa terveydenhuollon ennustemalleissa

**Esimerkki 2.1219**

Keskiarvokentän variatiivinen päättely on menetelmä Bayesin likimääräiseen posterioriseen päättelyyn. Se approksimoi täydellistä posteriorijakaumaa faktoroidulla jakaumien joukolla maksimoimalla marginaalisen todennäköisyyden alarajan. Tämä edellyttää kykyä integroida termien summa log yhteiseen todennäköisyyteen käyttäen tätä faktoroitua jakaumaa. Usein kaikki integraalit eivät ole suljetussa muodossa, mikä hoidetaan tyypillisesti käyttämällä alarajoja. Esitämme vaihtoehtoisen stokastiseen optimointiin perustuvan algoritmin, joka mahdollistaa suoraan variationaalisen alarajan optimoinnin. Tässä menetelmässä käytetään kontrollimuuttujia stokastisen hakugradientin varianssin pienentämiseen, jossa olemassa olevilla alarajoilla voi olla tärkeä rooli. Demonstraamme lähestymistapaa kahdella ei-konjugoituneella mallilla: logistisella regressiolla ja HDP:n approksimaatiolla.

**Tulos**

Variationaalinen Bayesin päättely stokastisen haun avulla

**Esimerkki 2.1220**

Syvät neuroverkot (Deep Neural Networks, DNN) ovat tehokkaita malleja, jotka ovat saavuttaneet erinomaisen suorituskyvyn vaikeissa oppimistehtävissä. Vaikka DNN:t toimivat hyvin aina, kun käytettävissä on suuria merkittyjä harjoitusjoukkoja, niitä ei voida käyttää sekvenssien kartoittamiseen sekvensseihin. Tässä artikkelissa esitämme yleisen, alusta loppuun ulottuvan lähestymistavan sekvenssien oppimiseen, jossa sekvenssin rakenteeseen kohdistuu vain vähän oletuksia. Menetelmämme käyttää monikerroksista LSTM-muistia (Long Short-Term Memory, pitkä lyhytkestoinen muisti) kartoittamaan syötesekvenssin vektoriksi, jolla on kiinteä ulottuvuus, ja sitten toista syvää LSTM-muistia purkamaan kohdesekvenssin vektorista. Tärkein tuloksemme on, että WMT-14-tietokannan englannista ranskaan kääntämistehtävässä LSTM:n tuottamat käännökset saavuttavat 34,8 BLEU-pisteen koko testijoukossa, jossa LSTM:n BLEU-pistemäärää rangaistiin sanaston ulkopuolisista sanoista. Lisäksi LSTM:llä ei ollut vaikeuksia pitkien lauseiden kanssa. Vertailun vuoksi mainittakoon, että vahva fraasipohjainen SMT-järjestelmä saa samassa aineistossa BLEU-pistemääräksi 33,3. Kun käytimme LSTM:ää edellä mainitun SMT-järjestelmän tuottamien 1000 hypoteesin uudelleenluokitteluun, sen BLEU-pistemäärä nousi 36,5:een, mikä ylittää aiemman tekniikan tason. LSTM oppi myös järkeviä lause- ja virkerepresentaatioita, jotka ovat herkkiä sanajärjestykselle ja ovat suhteellisen muuttumattomia aktiivin ja passiivin suhteen. Lopuksi havaitsimme, että sanojen järjestyksen kääntäminen päinvastaiseksi kaikissa lähdelauseissa (mutta ei kohdelauseissa) paransi LSTM:n suorituskykyä huomattavasti, koska se toi lähde- ja kohdelauseen välille monia lyhytaikaisia riippuvuuksia, jotka helpottivat optimointiongelmaa.

**Tulos**

Järjestyksestä järjestykseen oppiminen neuroverkkojen avulla

**Esimerkki 2.1221**

Suosittelujärjestelmien tutkimuksessa algoritmeja luonnehditaan usein joko yhteissuodatukseksi (Collaborative Filtering, CF) tai sisältöpohjaiseksi (Content Based, CB). CF-algoritmit koulutetaan käyttämällä tietokantaa, joka sisältää käyttäjien eksplisiittisiä tai implisiittisiä mieltymyksiä, kun taas CB-algoritmit perustuvat tyypillisesti tuoteprofiileihin. Nämä lähestymistavat hyödyntävät hyvin erilaisia tietolähteitä, joten myös tuloksena olevat suositellut kohteet ovat yleensä hyvin erilaisia. Tässä artikkelissa esitellään uusi malli, joka toimii siltana kohteiden sisällöstä niiden CF-edustuksiin. Esitämme usean syötteen syväregressiomallin, jolla ennustetaan kohteiden CF-latenttiset upotusvektorit niiden tekstimuotoisen kuvauksen ja metatietojen perusteella. Esittelemme ehdotetun mallin tehokkuuden ennustamalla elokuvien ja sovellusten CF-vektorit niiden tekstikuvausten perusteella. Lopuksi osoitamme, että mallia voidaan parantaa edelleen sisällyttämällä siihen metatietoja, kuten elokuvan julkaisuvuosi ja tunnisteet, jotka lisäävät tarkkuutta.

**Tulos**

Microsoft Word - cb2cf\_arxiv.docx

**Esimerkki 2.1222**

Äidin ja vauvan elektrokardiografiasignaalien (EKG tai EKG) erottaminen on haastava tehtävä, koska käytetään yhtä laitetta, joka vastaanottaa useiden sydämenlyöntien sekoitusta. Tässä asiakirjassa haluamme suunnitella suodattimen, jolla signaalit voidaan erottaa toisistaan.

**Tulos**

Elektrokardiografia Äidin ja vauvan erottaminen toisistaan

**Esimerkki 2.1223**

Tässä artikkelissa kuvataan uudenlainen tiedon esittämis- ja louhintajärjestelmä, jota kutsumme semanttiseksi tietograafiksi. Semanttisen tietämysgraafin ytimessä on käänteinen indeksi ja sitä täydentävä kääntämätön indeksi, jotka edustavat solmuja (termejä) ja reunoja (asiakirjoja, jotka ovat useiden termien/solmujen risteävien postituslistojen sisällä). Tämä tarjoaa indirektiokerroksen jokaisen solmuparin ja niitä vastaavan reunan välille, jolloin reunat voivat syntyä dynaamisesti taustalla olevien korpustilastojen perusteella. Tämän seurauksena minkä tahansa solmujen yhdistelmän reunat voivat muodostua mihin tahansa toiseen solmuun, ja ne voidaan pisteyttää, jotta voidaan paljastaa solmujen väliset piilevät suhteet. Tästä on monia etuja: tietämysgraafi voidaan rakentaa automaattisesti reaalimaailman tietopohjasta, uusia solmuja ja niiden yhdistettyjä reunoja voidaan välittömästi materialisoida mistä tahansa olemassa olevien solmujen mielivaltaisesta yhdistelmästä (joukko-operaatioiden avulla), ja kaikkien toimialan entiteettien välisten semanttisten suhteiden täysi malli voidaan esittää ja sitä voidaan dynaamisesti selata graafin erittäin kompaktin esityksen avulla. Tällaisella järjestelmällä on laajalti sovelluksia niinkin erilaisilla aloilla kuin tietämyksen mallintaminen ja päättely, luonnollisen kielen käsittely, poikkeamien havaitseminen, tietojen puhdistus, semanttinen haku, analytiikka, tietojen luokittelu, perimmäisten syiden analysointi ja suositusjärjestelmät. Tämän artikkelin tärkein panos on uudenlaisen järjestelmän, Semantic Knowledge Graphin, käyttöönotto, joka pystyy dynaamisesti löytämään ja pisteyttämään mielenkiintoisia suhteita minkä tahansa mielivaltaisen kokonaisuuksien (sanojen, lauseiden tai käsitteiden) yhdistelmän välillä materialisoimalla dynaamisesti solmuja ja särmiä kompaktista graafisesta esityksestä, joka on rakennettu automaattisesti tietämysaluetta edustavasta tietopaketista. Semanttisen tietograafin toteutuksen lähdekoodi julkaistaan tämän artikkelin yhteydessä, jotta työn jatkotutkimus ja laajentaminen olisi helpompaa.

**Tulos**

Semanttinen tietämysgraafi: Kompakti, automaattisesti luotu malli, jolla voidaan reaaliaikaisesti selata ja asettaa kaikki suhteet verkkotunnuksen sisällä paremmuusjärjestykseen.

**Esimerkki 2.1224**

Tässä artikkelissa tutkitaan uutta menetelmää, jolla parannetaan asiantuntijoiden sekoitusmallin (Mixture of Experts, ME) oppimisalgoritmia käyttämällä muunnetun käkikukkohakumenetelmän (Modified Cuckoo Search, MCS) ja konjugaattigradientin (Conjugate Gradient, CG) hybridiä toisen asteen optimointitekniikkana. CG-tekniikka yhdistetään Back-Propagation (BP) -algoritmiin, jolloin saadaan paljon tehokkaampi oppimisalgoritmi ME-rakenteelle. Lisäksi tehostetun mallin asiantuntijat ja porttiverkot korvataan CG-pohjaisilla monikerroksisilla perceptroneilla (MLP) nopeamman ja tarkemman oppimisen aikaansaamiseksi. CG riippuu huomattavasti keinotekoisen neuroverkon (ANN) yhteyksien alkupainoista, joten optimaalisten painojen valitsemiseksi käytetään metaheuristista algoritmia, niin sanottua modifioitua Cuckoo-hakua. Ehdotetun menetelmän suorituskykyä verrataan Gradient Decent Based ME:hen (GDME) ja Conjugate Gradient Based ME:hen (CGME) luokitus- ja regressio-ongelmissa. Kokeelliset tulokset osoittavat, että MSC:hen ja CG:hen perustuvan ME:n (MCS-CGME) hybridi on nopeampi konvergenssi ja parempi suorituskyky käytetyissä vertailutietoaineistoissa.

**Tulos**

MLP-asiantuntijoiden laajennettu sekoitus konjugaattisen gradienttimenetelmän ja modifioidun Cuckoo-hakumenetelmän hybridillä

**Esimerkki 2.1225**

Monissa robottisovelluksissa järjestelmän dynamiikan jotkin osatekijät voidaan mallintaa tarkasti, kun taas toisia on vaikea saada tai mallintaa. Esittelemme uudenlaisen vahvistusoppimismenetelmän (RL) jatkuville tila- ja toimintaavaruuksille, joka oppii osittaisella tietämyksellä järjestelmästä ja ilman aktiivista etsintää. Se ratkaisee lineaarisesti ratkaistavia Markovin päätösprosesseja (L-MDP), jotka soveltuvat hyvin jatkuviin tila- ja toiminta-avaruuksiin ja perustuvat toimijakriittiseen arkkitehtuuriin. Verrattuna aiempiin L-MDP:iden RL-menetelmiin ja polkuintegraalimenetelmiin, jotka perustuvat malliin, toimijakriittinen oppiminen ei tarvitse mallia hallitsemattomasta dynamiikasta eikä, mikä on tärkeää, siirtymäkohinatasoista; se edellyttää kuitenkin ongelman valvontadynamiikan tuntemista. Arvioimme menetelmäämme kahdella synteettisellä testiongelmalla ja yhdellä todellisella ongelmalla simuloinnissa ja käyttämällä todellista liikennedataa. Kokeemme osoittavat, että oppiminen ja toimintaperiaatteiden suorituskyky ovat parantuneet.

**Tulos**

Toimijakriittisyys lineaarisesti ratkaistavissa oleville jatkuville MDP:ille, joiden dynamiikka tunnetaan osittain.

**Esimerkki 2.1226**

Tässä artikkelissa ehdotetaan todennäköisyysmallin tarkistuksen käyttöä optimaalisten robottikäytäntöjen syntetisoimiseksi monitehtäväisissä autonomisissa järjestelmissä, jotka ovat alttiita ihmisen ja robotin väliselle vuorovaikutukselle. Kun otetaan huomioon vakuuttavat empiiriset todisteet siitä, että ihmisen käyttäytyminen voidaan liittää vahvistusmalleihin, otamme syötteenä hyvin tutkittua Q-taulukkomallia ihmisen käyttäytymisestä joustavissa skenaarioissa. Aluksi kuvaamme automaattisen menettelyn, jolla tislataan Markovin päätösprosessi (MDP) ihmiselle mielivaltaisessa mutta kiinteässä skenaariossa. Erikoisuutena on se, että - toisin kuin nykyisissä malleissa - ihmisen käyttäytymisen alispesifiointi on otettu huomioon. Ihmisen käyttäytymisen ennustamiseen käytetään todennäköisyystarkastelua. Lopuksi MDP-mallia laajennetaan robottimallilla. Optimaaliset robottikäytännöt syntetisoidaan analysoimalla tuloksena syntyvää kahden pelaajan stokastista peliä. Kokeelliset tulokset, jotka on saatu PRISM-ohjelmalla toteutetun prototyyppitoteutuksen avulla, osoittavat lupaavia tuloksia.

**Tulos**

Todennäköisyysmallin tarkistus monimutkaisia kognitiivisia tehtäviä varten

**Esimerkki 2.1227**

Multilabel-luokittelu on suhteellisen uusi koneoppimisen osa-alue. Toisin kuin klassisessa lähestymistavassa, jossa tapaukset merkitään vain yhdellä luokalla, monimerkkiluokittelussa tapauksen merkitsemiseen valitaan mielivaltainen määrä luokkia. Ongelman monimutkaisuuden vuoksi (ratkaisu on yksi eksponentiaalisen määrän vaihtoehtojen joukossa) käytetään usein hyvin yleistä ratkaisua (binäärimenetelmä), jossa jokaiselle luokalle opetetaan binääriluokittelija ja yhdistetään ne kaikki sen jälkeen. Tässä ratkaisussa käytetty oletus ei ole realistinen, ja tässä työssä annamme esimerkkejä, joissa kaikkia merkintöjä koskevia päätöksiä ei tehdä toisistaan riippumatta, ja näin ollen valvotun lähestymistavan pitäisi oppia luokkien väliset suhteet paremman luokituksen aikaansaamiseksi. Siksi esitämme tässä yleisen menetelmän, jolla voidaan parantaa riippumattomien todennäköisyysluokittelijoiden avulla saatuja tuloksia käyttämällä yhdistelmämenetelmää, jossa käytetään luokittelijan kanssa luokittelua, joka on koulutettu merkintöjen yhteisesiintyvyyksien perusteella. Esitämme kattavan kokeilun kolmella eri standardikorpuksella leimattuja asiakirjoja (Reuters-21578, Ohsumed-23 ja RCV1), joissa kaikissa on havaittavissa huomattavia parannuksia, kun menetelmäämme käytetään kolmessa probabilistisessa perusluokittimessa.

**Tulos**

Todennäköisyysperusteinen menetelmä monimerkkiluokittelua varten

**Esimerkki 2.1228**

Lähimmän naapurin menetelmät ovat suosittu ei-parametristen estimaattoreiden luokka, jolla on useita<lb>toivottuja ominaisuuksia, kuten mukautuvuus eri etäisyysasteikoille<lb>avaruuden eri alueilla. Lähimmän naapurin luokittelun konvergenssinopeuksia koskeva aiempi työ ei ole täysin<lb>tarkoittanut näitä hienovaraisia ominaisuuksia. Analysoimme näiden estimaattoreiden käyttäytymistä metrisissä<lb>avaruuksissa ja annamme äärellisestä otoksesta ja jakaumasta riippuvat konvergenssinopeudet minimi<lb>oletuksilla. Sivutuotteena pystymme osoittamaan<lb>lähimmän naapurin universaalin johdonmukaisuuden laajemmissa data-avaruuksissa kuin aiemmin tiedettiin. Havainnollistamme ylä- ja alarajojamme ottamalla käyttöön sileysluokkia, jotka on räätälöity<lb>lähimmän naapurin luokittelua varten.

**Tulos**

Lähimmän naapurin luokittelun lähentymisasteet

**Esimerkki 2.1229**

Analysoimme shakin tilaavaruuden rakennetta siirtymäpolkujen näytteenoton Monte Carlo -simuloinnin avulla. Perustuen tyypilliseen siirtojen määrään, joka tarvitaan tietyn shakkinappuloiden kokoonpanon siirtämiseen toiseen, päätämme, että tilaavaruus koostuu useista taskuista, joiden välillä siirtymät ovat harvinaisia. Taitavat pelaajat tutkivat vielä pienempää osajoukkoa asemia, jotka asuttavat joitakin näistä taskuista vain hyvin harvoin. Nämä tulokset viittaavat siihen, että tavanomaiset mittarit, joilla arvioidaan sekä tilaavaruuden kokoa että laillisten siirtojen puun kokoa, eivät ole yksiselitteisiä pelin monimutkaisuuden indikaattoreita, vaan että topologiset näkökohdat ovat yhtä tärkeitä. Shakki on kahden pelaajan lautapeli, jossa on pieni joukko sääntöjä, joiden mukaan nappuloita voidaan siirtää. Se kuuluu täydellisen informaation omaavien pelien luokkaan, joita ei ole vielä ratkaistu sen tilaavaruuden suuren koon vuoksi. Shakin analysointi tietokoneella alkoi Claude Shannonin vuonna 1950 ilmestyneestä uraauurtavasta artikkelista [1], ja noin vuodesta 2000 lähtien tietokoneohjelmat ovat säännöllisesti voittaneet huipputason ihmispelaajia [2]. Ne tekevät sen käyttämällä pelin tilojen hyvin räätälöityjä heuristisia arviointifunktioita, joiden avulla voidaan lyhentää mahdollisten siirtojen laajan pelipuun tutkimista. Tässä yhteydessä shakkia verrataan usein Go -peliin, jossa tietokoneet ovat vasta aivan hiljattain alkaneet saavuttaa ihmismestareiden suorituskykyä [3]. Ero johtuu yleensä pelien tila-avaruuksien erilaisesta koosta: Go:n pelipuun monimutkaisuus ylittää shakin noin 200 suuruusluokkaa. Vaikka koko on tärkeä tekijä pelin monimutkaisuuden määrittelyssä, myös tilaavaruuden topologia voi olla yhtä tärkeä tekijä. Intuitiivisesti eri shakkinappuloiden tekemät erilaiset siirrot aiheuttavat erittäin epätriviaalin (ja suunnatun) topologian. Ei ole lainkaan yksinkertaista todeta, onko tiettyyn tilaavaruuden pisteeseen mahdollista päästä toisesta pisteestä laillisten siirtojen sarjalla. Edessämme on siis mielenkiintoinen näytteenotto-ongelma: voidaanko kahden shakkikonfiguraation perusteella todeta, ovatko ne yhteydessä toisiinsa, eli onko olemassa laillisten siirtojen sarja, joka muuttaa ensimmäisen konfiguraation toiseksi? Lisäksi, mikä on tyypillinen etäisyys (kerroksina tai puolikkaina siirtoina) tällaisten konfiguraatioiden välillä? On selvää, että suora luetteleminen tai tavanomainen Monte Carlo -näytteenotto eivät ole mahdollisia: jokaisen kerroksen jälkeen pelipuu haarautuu arviolta 30-35 alipuuhun [1]. Tässä osoitamme, että shakin tilaavaruuden topologista rakennetta on mahdollista analysoida stokastisen prosessin harvinaisten tapahtumien näytteenotolla (SPRES) [4]. SPRES on siirtymäpolku-Monte Carlo -näytteenottojärjestelmä, joka toimii täysin epätasapaino-olosuhteissa, joissa dynamiikka ei ole stationaarista eikä palautuvaa. 1 Yhdistämällä SPRES:n optimoituun shakkiliikkeen generaattoriin [5] estimoimme polkujen pituuksien jakauman sekä satunnaisesti luotujen kokoonpanojen että ihmisten pelaamissa peleissä esiintyvien kokoonpanojen välillä. Analysoimalla näitä jakaumia satunnaisgrafioteorian kannalta oletamme, että shakin tilaavaruus koostuu useista erillisistä taskuista, joita yhdistävät suhteellisen harvat polut. Näissä taskuissa on vain hyvin harvoin sellaisia tiloja, joilla on merkitystä taitavan pelin kannalta. Aikaisemmat shakin tilastofysiikan analyysit ovat keskittyneet lähinnä siirtojen jakautumiseen ihmisten pelatessa tai tietokoneiden shakkimoottoreiden pelaamiin peleihin. Esimerkiksi avausjaksojen suosio noudattaa Zipfin lain [6] mukaista voimalaajajuusjakaumaa (tässä 1Shakkianalyysimme osoittaa myös SPRESin monipuolisuutta ja tehoa tekniikkana, jota voidaan soveltaa abstraktiin ei-fysikaaliseen dynamiikkaan. p-1 ar X iv :1 60 9. 04 64 8v 1 [ cs .A I] 1 4 Se p 20 16 A. Atashpendar, T. Schilling ja Th. Voigtmannin kontekstissa Go on melko samankaltainen [7]), joka on erittäin riippuvainen osallistuvien pelaajien taidoista [8, 9]. Optimaalista pelaamista (siinä mielessä, että nykyaikaiset tietokoneshakkimoottorit arvioivat siirrot suotuisiksi) on myös analysoitu vapaiden energiamaisemien kielellä [10]. Lähestymistapamme on täysin erilainen: tarkastelemme kaikkien laillisten siirtojen joukkoa riippumatta niiden moottoriarvioinnista, jotta voimme määrittää shakin tilaavaruuden kytkeytyneisyyden. Tämän avaruuden sisällä tutkimme sitten myös shakkimestareiden pelaamissa peleissä esiintyvien asemien osajoukon suhteellista kokoa ja rakennetta. Shakkipelin tilaa missä tahansa vaiheessa kuvaa täysin laudan kokoonpano (kaikkien shakkinappuloiden sijainnit), pieni joukko lisämuuttujia, joilla seurataan erikoisliikkeiden mahdollisuutta (linnoitus tai kaato), ja tieto siitä, minkä pelaajan vuoro on meneillään. Mahdollisten tilojen joukon muodostavat kaikki tilat, joissa on enintään 16 shakkinappulaa väriä kohti (nappuloiden vuoksi niitä voi olla vähemmän, ja nappuloiden ja sotilaiden määrä voi muuttua sotilaiden siirtojen vuoksi). Vain osajoukko kaikista mahdollisista tiloista on laillisia, sillä esimerkiksi kaksi kuningasta ei voi olla samaan aikaan shakissa. Seuraavassa kiinnostavat tilat, jotka ovat laillisia ja jotka ovat myös saavutettavissa annetusta lähtökonfiguraatiosta. Esimerkkinä saavuttamattomasta mutta laillisesta tilasta voidaan tarkastella tapausta, jossa piispan asema poikkeaa sen alkuasennosta, mutta sotilaiden asemat eivät poikkea. Tämä tila on saavuttamattomissa, koska sotilaat asetetaan alun perin oman värinsä muiden nappuloiden eteen, niiden siirrot ovat aina peruuttamattomia ja muut nappulat (ratsuja lukuun ottamatta) eivät voi hypätä sotilaiden yli. Näin ollen, vaikka tila on laillinen, siihen ei voi päästä laillisilla siirroilla. Otantaa varten tilaavaruuden rakenteesta luodaan saavutettavissa olevien tilojen sarjoja arpomalla satunnaisesti siirtoja tasaisesti kaikista laillisista siirroista (Monte Carlo, MC). Useimmat näistä tiloista aiheuttavat dramaattisia haittoja ainakin toiselle osapuolelle. Siksi optimaalisen strategian pelissä esiintyvien tilojen joukko on huomattavasti pienempi kuin otoksemme. Näiden tuntemattomien optimaalisten tilojen korvikkeena käytämme tietokannan (DB) tiloja, jotka on poimittu noin kahden miljoonan ihmisen pelaaman pelin tietokannasta [11]. Molemmissa tapauksissa (MC ja DB) valitsemme sitten satunnaisesti tilapareja ja määrittelemme niiden liitettävyyden pelipuun suhteen kaikilla laillisilla (MC) siirroilla, eli riippumatta siitä, sisältääkö yhdistävä polku pelattavuuden kannalta epäedullisia asemia. Lähtökonfiguraation läheisyydessä monet satunnaisesti arvotut asemaparit ovat välttämättä irrallisia, koska pelinappulat liikkuvat vain eteenpäin ja monien nappuloiden on vielä saatava liikkumisvapaus. Pelin toisessa päässä paritusasemat toimivat absorboivina tiloina. Lisäksi MC-dynamiikassa on joukko absorboivia tiloja, joissa pelilaudalle jää vain kuninkaita. Jotta voisimme ottaa otoksen tiloista, jotka heijastavat tilaavaruuden luontaista topologiaa, rajoitamme keskustelun siten tilapareihin, jotka on poimittu pelin syvyydeltä 5 ja 50 lyönnin väliseltä syvyydeltä. Tämä vastaa löyhästi shakinpelaajien käsitystä keskipelistä. Tämän ikkunan sisällä emme havainneet selvää korrelaatiota sen kerroksellisuuden, josta tilapari piirrettiin, ja niiden välisen etäisyyden välillä. Otimme otoksen tilojen välisistä poluista SPRES-ohjelmalla [4]. Tässä menetelmässä tila-avaruuden rajapinnat määritellään skalaarisen reaktiokoordinaatin vakioarvoilla, jotka kvantifioivat tilasta toiseen etenemistä. Tämän jälkeen dynaamisten polkujen adaptiivinen näytteenotto suoritetaan siten, että näiden rajapintojen välille saadaan vakiomäärä eteenpäin siirtymisiä. Kun näytteenotto on saatu päätökseen, havaintomuuttujien keskiarvo voidaan laskea näytteenottokäytävien kokonaisuudesta. Shakin tapauksessa meitä kiinnostaa erityisesti kahden konfiguraation välisen lyhimmän polun pituus (kerrosten lukumäärä). Vaikka optimaalisen reaktiokoordinaatin valinta on oma aiheensa [10], me hyödynnämme sitä, että SPRES poimii polkuja uskollisesti myös epäoptimaalisilla valinnoilla [4]. Reaktiokoordinaatiksi valitsimme euklidisen geometrisen etäisyysmitan tavoitekonfiguraatiosta. Kullekin kappaleelle lasketaan geometrinen etäisyys käyttäen metriikkaa, joka on mukautettu kyseisen kappaleen suorittamien siirtojen tyyppiin: Chebyshevin metriikka kuningattarille, kuninkaille ja piispoille, puolen Chebyshevin etäisyyden katto ratsuille, Manhattanin etäisyys torneille ja sijoitusero sotilaille. (Yksityiskohtaiset tiedot löytyvät viitteestä [5]). Parit hylätään irrallisina, jos ne ovat kauempana toisistaan kuin 120 pietiä; tämä approksimaatio on mukautettu todellisten shakkipelien tyypilliseen pituuteen. Triviaalisti irralliset parit hylätään alkutestillä, joka perustuu reaktiokoordinaatistoon, nappuloiden rakenteeseen ja nappuloiden lukumäärään. Polkujen pituuksien arvioimiseksi on poimittu 4000 (3000) MC:stä (DB:stä) luotua paria, jotka ovat läpäisseet tämän testin. Kuvassa 1 esitetään polkujen pituuksien histogrammi niiden satunnaisesti valittujen parien välillä, jotka ovat yhteydessä toisiinsa SPRES:n mukaan (vastaa 79 prosenttia kaikista satunnaisesti poimituista MC-pareista ja 85 prosenttia kaikista DB:stä poimituista pareista). MC:n avulla luotujen parien osalta polunpituusjakaumassa on kaksi erillistä osuutta, joista toisessa on huippu tasolla `1 ≈ 20 kerrosta ja toisessa pienempi tasolla `2 ≈ 45 kerrosta. Tietokannasta poimittujen parien välinen polunpituusjakauma painottuu pienempiin polunpituuksiin, ja siinä on vain yksi näkyvä huippu hieman alle `1:n polunpituuden kohdalla, `1 ≈ 18 kerrosta. Suuria etäisyyksiä kohti suuntautuva pyrstö näkyy edelleen jäänteenä MC-jakaumassa havaitusta toisesta piikistä. Huomaa, että SPRES:n löytämät polut DB-pareille kulkevat lähes varmasti muiden kuin DB-tilojen kautta (eli tilojen, joita ei yleensä esiinny ihmisten pelaamissa peleissä). Tyypillinen moottorin arviointifunktio (Stockfish [13]) osoittaa valtavia vaihteluita SPRES:n polkujen varrella, mikä osoittaa, että taitavat ihmispelaajat eivät todennäköisesti koskaan valitse näitä polkuja. Kuvassa 1 esitetyt tulokset paljastavat, että todelliset shakkipelit tapahtuvat aliavaruudessa, joka on paljon tiiviimmin sidottu kuin saavutettavien tilojen avaruus. Kaksoishuippuinen histogrammi viittaa "blob"-rakenteeseen (ks. luonnos).

**Tulos**

epl-luonnos Sekvenssisakki

**Esimerkki 2.1230**

Ehdotamme kuvien strukturoitua ennustusarkkitehtuuria, jonka keskiössä ovat syvät toistuvat neuroverkot. Ehdotettu verkko, nimeltään ReSeg, perustuu hiljattain esittelemäänsä ReNet-malliin objektien luokittelua varten. Muokkaamme ja laajennamme sitä suorittamaan objektien segmentointia ja toteamme, että poolingin välttäminen voi yksinkertaistaa huomattavasti kuvien pikselikohtaisia tehtäviä. ReSeg-kerros koostuu neljästä rekursiivisesta neuroverkosta, jotka pyyhkäisevät kuvaa vaaka- ja pystysuunnassa molempiin suuntiin, sekä viimeisestä kerroksesta, joka laajentaa ennusteen takaisin kuvan alkuperäiseen kokoon. ReSeg yhdistää useita ReSeg-kerroksia useisiin mahdollisiin syöttökerroksiin sekä loppukerrokseen, joka laajentaa ennusteen takaisin alkuperäiseen kuvakokoon, joten se soveltuu monenlaisiin strukturoituihin ennustustehtäviin. Arvioimme ReSegiä erityisessä objektin segmentointitehtävässä kolmella laajalti käytetyllä kuvan segmentointitietokannalla, jotka ovat Weizmann Horse, Fashionista ja Oxford Flower. Tulokset viittaavat siihen, että ReSeg voi haastaa esineiden segmentoinnin nykytilanteen, ja sillä voi olla lisäsovelluksia strukturoidussa ennustamisessa laajemminkin.

**Tulos**

RESEG: TOISTUVA NEUROVERKKO KOHTEIDEN SEGMENTOINTIIN

**Esimerkki 2.1231**

Keinotekoiset neuroverkot ovat tehokkaita kuvioiden luokittelijoita, mutta niiden tarkkuuden ovat kuitenkin ohittaneet sellaiset menetelmät kuin tukivektorikoneet ja satunnaismetsät, jotka ovat myös helpompia käyttää ja nopeampia kouluttaa. Keinotekoisten neuroverkkojen kouluttamiseen käytettävä backpropagation kärsii laumavaikutuksesta, joka johtaa pitkiin koulutusaikoihin ja rajoittaa luokittelutarkkuutta. Käytämme disjunktiivista normaalimuotoa ja approksimoimme boolen konjunktio-operaatioita tuotteilla uuden verkkoarkkitehtuurin rakentamiseksi. Ehdotettu malli voidaan kouluttaa minimoimalla virhefunktio, ja se mahdollistaa tehokkaan ja intuitiivisen alustuksen, joka ratkaisee backpropagationiin liittyvän laumavaikutusongelman. Tämä johtaa uusimman tason luokittelutarkkuuteen ja nopeisiin koulutusaikoihin. Lisäksi mallimme voidaan optimoida yhdessä konvoluutio-ominaisuuksien kanssa yhtenäisessä rakenteessa, mikä johtaa huippuluokan tuloksiin tietokonenäköongelmissa nopeilla konvergenssivauhdilla. Saatavilla on myös LDNN:n GPU-toteutus valinnaisilla konvoluutio-ominaisuuksilla.

**Tulos**

Disjunktiiviset normaaliverkot

**Esimerkki 2.1232**

Ehdotamme uutta koulutusmenetelmää feedforward-neuraaliverkolle, jonka aktivointifunktioilla on geometrinen supistumisominaisuus. Menetelmässä rakennetaan uusi funktio, joka on vähemmän epälineaarinen verrattuna klassiseen funktioon, poistamalla aktivointifunktioiden epälineaarisuus ulostulokerroksesta. Validoimme tämän uuden menetelmän useilla kokeilla, jotka osoittavat, että oppimisnopeus on parantunut ja luokitteluvirhe on parempi. MSC: 92B20, 68T05.

**Tulos**

Aktivointifunktioiden geometriseen supistumisominaisuuteen perustuva uusi koulutusmenetelmä feedforward-neuraaliverkoille

**Esimerkki 2.1233**

Ensimmäisen asteen faktakysymyksiin vastaamisessa oletetaan, että kysymykseen voidaan vastata tietopohjassa olevan yksittäisen faktan avulla. Vaikka tämä ei vaikuta haastavalta tehtävältä, monissa viimeaikaisissa yrityksissä, joissa on käytetty joko monimutkaista kielellistä päättelyä tai syviä neuroverkkoja, on saavutettu 35-65 prosentin tarkkuus vertailujoukoissa. Meidän lähestymistapamme muotoilee tehtävän kahdeksi koneoppimisongelmaksi: kysymyksen sisältämien entiteettien tunnistaminen ja kysymyksen luokittelu yhdeksi tietokannan relaatiotyypeistä. Tämän rakenneoletuksen perusteella yksinkertainen mutta tehokas lähestymistapamme kouluttaa kaksi rekursiivista neuroverkkoa, joiden avulla saavutetaan huomattavasti parempi tulos kuin nykyisillä tekniikoilla - suhteellinen parannus on 16 prosenttia WebQuestions-kysymyksissä ja yli 38 prosenttia SimpleQuestions-kysymyksissä.

**Tulos**

Yksinkertainen ja tehokas kysymyksiin vastaaminen toistuvien neuroverkkojen avulla

**Esimerkki 2.1234**

Luonnollisen kielen tuottamisella on ratkaiseva merkitys kaikissa puhutun dialogin järjestelmissä. Esittelemme uuden lähestymistavan luonnollisen kielen tuottamiseen käyttämällä rekursiivisia neuroverkkoja koodaajan ja dekoodaajan välissä. Aiemmista töistä poiketen mallimme käyttää sekä leksikalisoituja että deleksikalisoituja versioita paikka-arvopareista kunkin dialogin kohdalla. Näin mallimme voi oppia kaikesta saatavilla olevasta datasta sen sijaan, että se olisi rajoitettu oppimaan vain deleksikalisoiduista slot-value-pareista. Osoitamme, että tämä auttaa malliamme tuottamaan luonnollisempia lauseita, joissa on parempi kielioppi. Parannamme mallimme suorituskykyä entisestään alustamalla sen painot esivalmistellusta kielimallista. Parhaiten suoriutuneen mallimme inhimillinen arviointi osoittaa, että se tuottaa lauseita, joita käyttäjät pitävät luonnollisempina ja miellyttävämpinä.

**Tulos**

Luonnollisen kielen tuottaminen vuoropuhelussa käyttämällä leksikalisoitua ja deleksikalisoitua dataa.

**Esimerkki 2.1235**

Ehdotamme moniagenttialgoritmia, joka pystyy automaattisesti löytämään relevantteja säännönmukaisuuksia tietyssä tietokokonaisuudessa ja määrittämään samalla hyväksytyn parametrisen eroavaisuusmitan konfiguraatioiden joukon, joka tuottaa kompakteja ja erillisiä klustereita. Kukin agentti toimii itsenäisesti suorittamalla Markovin satunnaiskävelyä syötetiedoston sopivassa painotetussa graafiesityksessä. Tällainen painotettu graafiesitys saadaan aikaan agentin hyväksymän erilaisuusmitan erityisellä parametrikokoonpanolla, ja agentti etsii ja tekee päätöksiä itsenäisesti yhden klusterin osalta kerrallaan. Tulokset osoittavat, että algoritmi pystyy löytämään parametrikonfiguraatioita Vastaava kirjoittaja Sähköpostiosoitteet: filippo.binachi@ryerson.ca (Filippo Maria Bianchi), enrico.maiorino@uniroma1.it (Enrico Maiorino), llivi@scs.ryerson.ca (Lorenzo Livi), antonello.rizzi@uniroma1.it (Antonello Rizzi), asadeghi@ryerson.ca (Alireza Sadeghian) URL: https://sites.google.com/site/lorenzlivi/ (Lorenzo Livi), http://infocom.uniroma1.it/~rizzi/ (Antonello Rizzi), http://www.scs.ryerson.ca/~asadeghi/ (Alireza Sadeghian) Preprint submitted to Information Sciences September 18, 2014, jotka tuottavat johdonmukaisen ja tulkittavissa olevan kokoelman klustereita. Lisäksi osoitamme, että algoritmimme osoittaa vertailukelpoista suorituskykyä muiden vastaavien huipputason algoritmien kanssa, kun vastassa on tiettyjä klusterointiongelmia.

**Tulos**

Agenttipohjainen algoritmi, joka hyödyntää useita paikallisia eroavaisuuksia klustereiden louhinnassa ja tiedon löytämisessä.

**Esimerkki 2.1236**

Monet tutkijat ja kaupalliset kehittäjät ovat hyödyntäneet Microsoft Kinect -kameraa ja sen luustonseurantaominaisuuksia erilaisissa reaaliaikaisissa ihmisen liikkeiden analysointisovelluksissa. Tässä artikkelissa arvioidaan ihmisen kinemaattisen liiketiedon tarkkuutta Kinect-järjestelmän ensimmäisessä ja toisessa sukupolvessa ja verrataan tuloksia optiseen liikkeenkaappausjärjestelmään. Keräsimme liikedataa 12 harjoituksessa 10 eri koehenkilöltä ja kolmesta eri näkökulmasta. Raportoimme Kinectin luurankojen nivelten lokalisoinnin ja luun pituuden arvioinnin tarkkuudesta verrattuna liikkeenkaappaukseen. Analysoimme myös nivelten lokalisointipoikkeamien jakaumaa sovittamalla Gaussin ja tasaisen jakauman mallien seosta, jotta voimme määrittää Kinectin liikedatan poikkeamat. Analyysimme osoittaa, että Kinect 2:lla on kaiken kaikkiaan vankempi ja tarkempi ihmisen asennon seuranta kuin Kinect 1:llä.

**Tulos**

Microsoft Kinectin ensimmäisen ja toisen sukupolven asennonseurannan tarkkuuden arviointi

**Esimerkki 2.1237**

Multimediasisältö tai puhuttu sisältö tarjoaa houkuttelevampaa tietoa kuin pelkkä tekstisisältö, mutta ensin mainittua on vaikeampi näyttää näytöllä ja käyttäjän on vaikeampi valita se. Tämän vuoksi suurten kokoelmien käyttö on ihmisille paljon vaikeampaa ja aikaa vievämpää kuin jälkimmäisen. Siksi on erittäin houkuttelevaa kehittää koneita, jotka pystyvät automaattisesti ymmärtämään puhuttua sisältöä ja tiivistämään keskeiset tiedot ihmisten selattavaksi. Tässä pyrkimyksessä ehdotettiin hiljattain uutta puhutun sisällön koneellisen ymmärtämisen tehtävää. Alkuperäiseksi tavoitteeksi määriteltiin TOEFL-testin kuullun ymmärtämiskoe, joka on haastava akateeminen englannin kielen koe englanninoppijoille, joiden äidinkieli ei ole englanti. Tähän tehtävään ehdotettiin myös Attention-based Multi-hop Recurrent Neural Network (AMRNN) -arkkitehtuuria, jossa otettiin huomioon vain puheilmaisujen sisällä olevat peräkkäiset suhteet. Tässä artikkelissa ehdotamme uutta hierarkkista huomiomallia (Hierarchical Attention Model, HAM), joka rakentaa monihyppymekanismin puustruktuuristen eikä peräkkäisten lausumien esitysten päälle. Ymmärtämissuorituskyky parani ASR-virheiden suhteen.

**Tulos**

HIERARKKINEN HUOMIOMALLI PUHUTUN SISÄLLÖN PAREMPAA KONEELLISTA YMMÄRTÄMISTÄ VARTEN

**Esimerkki 2.1238**

Yksi suosittu menetelmä laajojen tietokokonaisuuksien käsittelyyn on otanta. Esimerkiksi käyttämällä empiirisiä tilastollisia vipuvaikutuspistemääriä tärkeysotantajakaumana algoritmisen vipuvaikutuksen menetelmä ottaa näytteitä ja skaalaa uudelleen datamatriisien rivit/sarakkeet datan koon pienentämiseksi ennen laskelmien suorittamista osaongelmalle. Menetelmä on onnistunut parantamaan matriisiongelmien algoritmien, kuten pienimmän neliösumman approksimaation, pienimmän absoluuttisen poikkeaman approksimaation ja matalien matriisien approksimaation, laskennallista tehokkuutta. Aikaisemmassa työssä on keskitytty algoritmisiin kysymyksiin, kuten pahimman tapauksen suoritusaikoihin ja numeerisiin kysymyksiin, jotka liittyvät korkealaatuisten toteutusten tarjoamiseen, mutta mikään niistä ei käsittele tämän menetelmän tilastollisia näkökohtia. Tässä artikkelissa tarjoamme yksinkertaisen mutta tehokkaan kehyksen algoritmisen vipuvaikuttamisen tilastollisten ominaisuuksien arvioimiseksi lineaarisen regressiomallin parametrien estimoinnin yhteydessä, kun ennustajia on kiinteä määrä. Erityisesti johdamme useille vipuvaikutuspohjaisen otannan versioille tuloksia harhasta ja varianssista sekä ehdollisesti että ehdoitta havaittujen tietojen suhteen. Osoitamme, että tilastollisen harhan ja varianssin näkökulmasta kumpikaan, vipuvaikutukseen perustuva otanta tai yhtenäinen otanta, ei hallitse toista. Tämä tulos on erityisen silmiinpistävä, kun otetaan huomioon tunnettu tulos, jonka mukaan algoritmisesta näkökulmasta pahimman tapauksen analyysin kannalta vipuvaikutukseen perustuva otanta tuottaa tasaisesti parempia pahimman tapauksen algoritmisia tuloksia kuin yhtenäinen otanta. Näiden teoreettisten tulosten perusteella ehdotamme ja analysoimme kahta uutta vipuvaikutusalgoritmia: toinen rakentaa pienemmän pienimmän neliösumman ongelman "kutistetuilla" vipuvaikutuspisteillä (SLEV) ja toinen ratkaisee pienemmän ja painottamattoman (tai puolueellisen) pienimmän neliösumman ongelman (LEVUNW). Nykyisten vipuvaikutukseen perustuvien menetelmien sekä näiden kahden uuden menetelmän yksityiskohtainen empiirinen arviointi suoritetaan sekä synteettisillä että todellisilla aineistoilla. Empiiriset tulokset osoittavat, että teoriamme on hyvä ennustaja nykyisten ja uusien vipupohjaisten algoritmien käytännön suorituskyvylle ja että uudet algoritmit parantavat suorituskykyä. Esimerkiksi ehdotettu SLEV johtaa tyypillisesti parempiin harhapainoihin ja variansseihin sekä ehdoitta että ehdollisesti (havainnoitujen tietojen perusteella) ja ehdotettu LEVUNW johtaa tyypillisesti parempiin ehdottomiin harhapainoihin ja variansseihin, kun laskennan vähennys on sama kuin alkuperäisessä algoritmisessa vipuvaikutusmenetelmässä. ∗ Tilastotieteen laitos, Illinoisin yliopisto Urbana-Champaignissa, Champaign, IL 61820. Sähköposti: pingma@illinois.edu. †Matematiikan laitos, Stanfordin yliopisto, Stanford, CA 94305. Sähköposti: mmahoney@cs.stanford.edu. ‡Tilastotieteen laitos, Kalifornian yliopisto Berkeleyssä, Berkeley, CA 94720. Sähköposti: binyu@stat.berkeley.edu. 1 ar X iv :1 30 6. 53 62 v1 [ st at .M E ] 2 3 Ju n 20 13

**Tulos**

Tilastollinen näkökulma algoritmiseen vipuvaikuttamiseen

**Esimerkki 2.1239**

Viimeaikainen kehitys on osoittanut, että MSBN-verkkoja (Multiply Sectioned Bayesian Networks) voidaan käyttää luonnollisten järjestelmien diagnosointiin sekä keinotekoisten järjestelmien mallipohjaiseen diagnosointiin. Niitä voidaan soveltaa sekä yhden agentin päättelyjärjestelmiin että usean agentin hajautettuihin päättelyjärjestelmiin. Uskomusten eteneminen aliverkkoparin välillä on keskeisessä asemassa MSBN:n globaalin johdonmukaisuuden ylläpitämisessä. Tässä artikkelissa tutkitaan alun perin MSBN:ien yhteydessä esiteltyä UpdateBelief-operaatiota aliverkkojen väliseen uskomusten levittämiseen. Analysoimme, miten operaatio saavuttaa sille tarkoitetun toiminnon, mikä antaa vihjeitä sen tehokkuuden parantamiseksi. Tämän jälkeen kehitetään uusia versioita UpdateBeliefistä, jotka vähentävät aliverkkojen välisen propagandan laskenta-aikaa. Yksi niistä on optimaalinen siinä mielessä, että monilinkkisen uskomuspropagandan koordinointiin tarvitaan mahdollisimman vähän laskentaa. Optimointiongelma ratkaistaan ratkaisemalla graafiteoreettinen ongelma: pienimmän painon avoin kierros puussa.

**Tulos**

Aliverkkojen välisen uskomuksen päivityksen optimointi moninkertaisesti jaksotetuissa Bayesin verkoissa

**Esimerkki 2.1240**

Tietojen laatu on olennaisen tärkeää, jotta voidaan varmistaa tietojen luotettavuus, jotta sidosryhmät voivat tehdä päätöksiä. Todellisissa sovelluksissa, kuten äärimmäisten ympäristöjen tieteellisessä tutkimuksessa, on epärealistista vaatia, että kerätyt raakatiedot ovat täydellisiä. Kun tiedonlouhijoina ei ole mahdollista tietää fyysisesti miksi ja miten tietojen puhdistamiseksi, ehdotamme, että etsimme signaalin luontaista rakennetta monimuuttujaisten tietojen yhteisten tekijöiden tunnistamiseksi. Käyttämällä uutta tietoon perustuvaa oppimismenetelmäämme - yhteisten tekijöiden tiedonpuhdistusmenetelmää - vastaamme monimuuttujaisten tietojen puhdistamiseen liittyvään monitieteelliseen haasteeseen, kun monimutkaiset ulkoiset vaikutukset näyttävät häiritsevän useita tietomittauksia. Nykyiset data-analyysit käsittelevät tyypillisesti yhtä signaalimittausta kerrallaan ottamatta huomioon kaikkien signaalien välisiä yhteyksiä. Analysoimme kaikki signaalimittaukset samanaikaisesti löytääksemme piilotetut yhteiset tekijät, jotka saavat kaikki mittaukset vaihtelemaan yhdessä, mutta eivät todellisten datamittausten seurauksena. Käytämme yhteisiä tekijöitä vähentämään datan vaihtelua muuttamatta datan peruskeskitasoa, jotta fyysinen merkitys ei muuttuisi. Olemme analysoineet uudelleen NASA:n Mars Phoenix -lennon datan, jota Kounavesin (Phoenixin märkäkemian kokeen johtava tutkija) [1, 2] johtoryhmässä käytettiin ehdotetulla menetelmällämme osoittaaksemme tuloksena olevat erot. Osoitamme, että tämä uusi yhteistekijämenetelmä auttaa onnistuneesti vähentämään systemaattista kohinaa ilman, että sen lähde ymmärretään lopullisesti, ja heikentämättä signaalin fysikaalista merkitystä.

**Tulos**

Yhteisten tekijöiden lähestymistapa monimuuttujaisen datan puhdistukseen ja sen soveltaminen Mars Phoenix -lennon dataan.

**Esimerkki 2.1241**

Esittelemme menetelmän, jolla voidaan oppia puunleveyden rajoittamia Bayes-verkkoja tuhansia muuttujia sisältävistä datajoukoista. Bayes-verkon puunleveyden rajoittaminen vähentää huomattavasti päätelmien monimutkaisuutta. Koska se on kuitenkin graafin globaali ominaisuus, se lisää huomattavasti oppimisprosessin vaikeutta. Ehdotamme tähän tehtävään uutta algoritmia, joka pystyy skaalautumaan suurille alueille ja suurille puunleveyksille. Uudenlainen lähestymistapamme on johdonmukaisesti parempi kuin nykyinen tekniikka jopa kymmenentuhannen muuttujan aineistoissa.

**Tulos**

Rajoitetun puunleveyden Bayes-verkkojen oppiminen tuhansilla muuttujilla

**Esimerkki 2.1242**

Tässä artikkelissa kuvataan uusi lähestymistapa termipainotusjärjestelmien (TWS) oppimiseen tekstiluokittelun yhteydessä. Tekstinlouhinnassa TWS määrittää tavan, jolla dokumentit esitetään vektoriavaruusmallissa, ennen kuin luokittelijaa sovelletaan. Vaikka tavanomaisilla TWS-järjestelmillä (esim. Boolen ja termien frekvenssijärjestelmillä) on saavutettu hyväksyttävää suorituskykyä, TWS-järjestelmien määrittely on perinteisesti ollut taidetta. Lisäksi on edelleen vaikeaa määrittää, mikä on paras TWS-järjestelmä tiettyyn ongelmaan, eikä ole vielä selvää, voidaanko tunnettuja TWS-järjestelmiä yhdistelemällä luoda nykyisiä järjestelmiä parempia järjestelmiä. Ehdotamme tässä artikkelissa geneettistä ohjelmaa, jonka tavoitteena on oppia tehokkaita TWS-järjestelmiä, jotka voivat parantaa nykyisten järjestelmien suorituskykyä tekstiluokittelussa. Geneettinen ohjelma oppii yhdistämään joukon perusyksiköitä, jotta saadaan aikaan erottelevia TWS-järjestelmiä. Raportoimme laajasta kokeellisesta tutkimuksesta, joka käsittää aineistoja temaattisesta ja ei-teemaattisesta tekstiluokittelusta sekä kuvien luokittelusta. Tutkimuksemme osoittaa ehdotetun menetelmän pätevyyden; itse asiassa osoitamme, että geneettisen ohjelman avulla opitut TWS:t päihittävät perinteiset järjestelmät ja muut Vastaava kirjoittaja. Sähköpostiosoitteet: hugojair@inaoep.mx (Hugo Jair Escalante), mauricio.garcia.cs@gmail.com (Mauricio A. Garćıa-Limón), a.morales@inaoep.mx (Alicia Morales-Reyes), mgraffg@gmail.com (Mario Graff), mmontesg@inaoep.mx (Manuel Montes-y-Gómez), emorales@inaoep.mx (Eduardo F. Morales) Preprint submitted to Elsevier 8. lokakuuta 2014 Viimeaikaisissa töissä ehdotetut TWS:t. Lisäksi osoitamme, että tietyllä alalla opittuja TWS-järjestelmiä voidaan käyttää tehokkaasti muihin tehtäviin.

**Tulos**

Termipainotuksen oppiminen geneettisen ohjelmoinnin avulla tekstin luokittelua varten

**Esimerkki 2.1243**

Tässä artikkelissa kuvataan (ko)induktiivisten konstruktioiden laskennan (Calculus of (Co)Inductive Constructions (CIC)) tarkennusalgoritmi, joka on toteutettu interaktiivisessa teoreemantodistajassa Matita. Jalostusalgoritmin tehtävänä on antaa merkitys termeille, tyypeille ja todistustermille, jotka käyttäjä on kirjoittanut suoraan tai jotka on tuotettu käyttämällä taktiikoita, päätösmenettelyjä tai yleistä automaatiota. Termit kirjoitetaan "ulkoisella syntaksilla", jonka on tarkoitus olla käyttäjäystävällinen ja joka sallii tietojen poisjättämisen, tyypittelemättömät sidokset ja käyttäjän määrittelemien alatyyppien vapaan käytön. Jalostaja muokkaa termejä saadakseen niihin liittyviä hyvin tyypitettyjä termejä ITP:n ytimen ymmärtämällä sisäisellä syntaksilla. Se toimii erityisesti tyypin päättelyalgoritmina, kun kaikki sidokset ovat tyypittömiä. Ehdotettu algoritmi on kaksisuuntainen: kun termi on ulkoisessa syntaksissa ja termille odotetaan tyyppiä, se levittää mahdollisimman paljon tyypitystietoa termin lehtiin. Perinteiset yksisuuntaiset algoritmit sen sijaan etenevät alhaalta ylöspäin päättelemällä alatermin tyypin ja vertaamalla (yhdistämällä) sitä sen kontekstissa odotettuun tyyppiin vasta lopussa. Ehdotamme CIC:lle joitakin uusia kaksisuuntaisia sääntöjä, jotka ovat erityisen tehokkaita. Kaksisuuntaisuuden eduista mainittakoon parempi virheilmoitusten raportointi ja parempi riippuvaisten tyyppien päättely. Lisäksi kaksisuuntaisuuden ansiosta alatyypin määrittelyn pakkokeinojärjestelmä on tehokkaampi, ja tyypin päättely tuottaa yksinkertaisempia yhdistämisongelmia, jotka voidaan todennäköisemmin ratkaista toteutetuilla korkeamman asteen yhdistämisalgoritmeilla, jotka ovat luonnostaan epätäydellisiä. Lopuksi otamme ulkoiseen syntaksiin käyttöön paikanvaraajien vektorin käsitteen, jonka avulla voidaan jättää kerralla pois mielivaltainen määrä argumentteja. Paikanhaltijoiden vektorit mahdollistavat implisiittisten argumenttien triviaalin toteuttamisen ja yksinkertaistavat huomattavasti alkeellisten ja yksinkertaisten taktiikoiden toteuttamista. 1998 ACM Subject Classification: D.3.1, F.3.0.

**Tulos**

KAKSISUUNTAINEN TARKENNUSALGORITMI (KO)INDUKTIIVISTEN KONSTRUKTIOIDEN LASKENTAA VARTEN.

**Esimerkki 2.1244**

Kaksoisvirtalähteisten ajoneuvojen, kuten hybridisähköajoneuvojen, pistokehybridisähköajoneuvojen ja polttokennosähköajoneuvojen, energiankulutuksen globaali optimointi edellyttää, että reitin kaikki ominaisuudet tunnetaan matkan alussa. Yksi tärkeimmistä ominaisuuksista on ajoneuvon nopeusprofiili koko reitillä. Profiili vaikuttaa suoraan tietyn ajoneuvon energiantarpeeseen. Kuljettajan valitsema ajoneuvon nopeus vaihtelee kuitenkin kuljettajakohtaisesti ja ajoittain, ja se voi olla hitaampi, yhtä suuri tai nopeampi kuin keskimääräinen liikennevirta. Jos kuljettajan erityinen nopeusprofiili voidaan ennustaa, energiankäyttö voidaan optimoida valitulla reitillä. Tämän artikkelin tarkoituksena on tutkia syväoppimistekniikoiden soveltamista tähän ongelmaan, jotta voidaan tunnistaa ajosyklin alussa kuljettajakohtainen ajoneuvon nopeusprofiili yksittäisen kuljettajan toistuvalle ajosyklille, jota voidaan käyttää optimointialgoritmissa matkan aikana käytetyn fossiilisen polttoaineen energiamäärän minimoimiseksi. Avainsanat: Syväoppiminen, pinotut autokooderit, neuroverkot, liikenteen ennustaminen.

**Tulos**

Ajoneuvon nopeuden ennustaminen syväoppimisen avulla

**Esimerkki 2.1245**

Aikaisemmassa työssä otimme käyttöön joustavan päättelyn ja päätöksentekoteoreettisen metapohdinnan, joilla pyrimme ratkaisemaan normatiivisen päättelyn vaikeasti ratkaistavissa olevan ongelman. Sen sijaan, että pyrkisimme laskemaan uskomuksia ja toimia epävarmoja tapahtumia koskevista erotteluista koostuvien päätöksentekomallien avulla, tarkastelemme tässä menetelmiä, joilla voidaan päätellä uskomuksia matemaattisesta totuudesta ennen kuin automaattinen teoreemantarkistaja saa todistuksen valmiiksi. Käytämme Bayesin analyysia totuuden uskottavuuden päivittämiseksi teoreemantarkastuksen edistymisen perusteella ja osoitamme, miten päätösteoreettisia menetelmiä voidaan käyttää sen määrittämiseen, onko harkinnan jatkaminen arvokkaampaa kuin välitön toiminta aikakriittisissä tilanteissa.

**Tulos**

Tutkimukset teoreemojen todistamisesta rajallisten resurssien vallitessa

**Esimerkki 2.1246**

Ihmisten välisten suhteiden kuvaaminen on olennaisen tärkeää kertomusten ymmärtämisen kannalta. Tässä työssä käsittelemme ongelmaa, joka liittyy henkilöiden välisten suhteiden polariteetin päättelyyn kertomusten tiivistelmissä. Muotoilemme ongelman kunkin kertomuksen yhteiseksi rakenteelliseksi ennusteeksi ja esittelemme mallin, jossa yhdistetään kielellisistä ja semanttisista piirteistä saatava todistusaineisto sekä tekstin sosiaalisen yhteisön rakenteeseen perustuvat piirteet. Esitämme myös klusterointiin perustuvan lähestymistavan, jolla voidaan hyödyntää kertomustyyppien säännönmukaisuuksia, esimerkiksi oppia mieltymys rakkauskolmioihin romanttisissa tarinoissa. Wikipediasta poimittujen elokuvakertomusten tietokannassa strukturoidut mallit vähentävät virheitä yli 30 prosenttia verrattuna kilpailevaan perusmalliin, jossa hahmopareja tarkastellaan erikseen.

**Tulos**

Ihmisten välisten suhteiden päätteleminen kerronnallisessa tiivistelmässä (Inferring Interpersonal Relations in Narrative Summarie)

**Esimerkki 2.1247**

Kiinteistöarviointi, joka on kiinteistöjen hinnan arviointiprosessi, on ratkaisevan tärkeä sekä ostajille että myyjille neuvottelujen ja kaupan perustana. Perinteisesti kiinteistöjen hinnan arvioinnissa on laajalti käytetty toistuvan myynnin mallia. Se edellyttää kuitenkin monimutkaisen taloudelliseen tilanteeseen liittyvän indeksin suunnittelua ja laskentaa, jonka tarkka arviointi on haastavaa. Nykyään kiinteistönvälittäjät tarjoavat asiakkailleen helposti yksityiskohtaisia verkkotietoja kiinteistöistä. Olemme kiinnostuneita arvioimaan kiinteistöjen hintaa näiden suurten, helposti saatavilla olevien tietomäärien perusteella. Analysoimme erityisesti verkossa olevien talokuvien ennustusvoimaa, sillä ne ovat yksi tärkeimmistä tekijöistä, joiden perusteella internetin käyttäjät tekevät mahdollisen käyntipäätöksen. Vankkojen tietokonenäköalgoritmien kehittäminen tekee visuaalisen sisällön analysoinnin mahdolliseksi. Tässä työssä käytämme rekursiivista neuroverkkoa (Recurrent Neural Network, RNN) ennustamaan kiinteistöjen hintoja käyttämällä uusimpia visuaalisia ominaisuuksia. Kokeelliset tulokset osoittavat, että mallimme päihittää useat muut huipputason perusalgoritmit sekä absoluuttisen keskivirheen (MAE) että absoluuttisen prosentuaalisen keskivirheen (MAPE) osalta.

**Tulos**

Kiinteistöjen kuvaperusteinen arviointi

**Esimerkki 2.1248**

Mittasimme entropiaa ja symbolista monimuotoisuutta englannin- ja espanjankielisistä teksteistä, jotka sisälsivät kirjallisuuden Nobel-palkittuja ja muita kuuluisia kirjailijoita. Entropiaa, symbolien monimuotoisuutta ja symbolien taajuusprofiileja verrattiin näiden neljän ryhmän osalta. Rakensimme myös asteikon, joka on herkkä kirjoituksen laadulle, ja arvioimme sen suhdetta englannin kielen Fleschin ́s readability indexiin ja espanjan kielen Szigrisztin ́s perspicuity indexiin. Tulokset viittaavat siihen, että entropian ja sanojen monimuotoisuuden ja kirjoituksen laadun välillä on yhteys. Myös tekstilaji vaikuttaa tekstin entropiaan ja monimuotoisuuteen. Tulokset viittaavat tekstien automaattisen laadunarvioinnin uskottavuuteen.

**Tulos**

Kirjallisuuden laadun kvantifiointi monimutkaisuuskriteerien avulla

**Esimerkki 2.1249**

Toistuvien neuroverkkojen kouluttaminen pitkän aikavälin riippuvuuksien mallintamiseen on vaikeaa. Siksi ehdotamme, että ulkoista kielitietoa käytetään eksplisiittisenä signaalina, joka ilmoittaa mallille, mitä muistoja sen pitäisi käyttää. Ulkoista tietoa käytetään erityisesti lisäämään sekvenssiä mielivaltaisen kaukana toisistaan olevien elementtien välisillä tyypitetyillä reunoilla, ja tuloksena syntyvä graafi puretaan suunnattuihin asyklisiin aligraafeihin. Esittelemme mallin, joka koodaa tällaiset graafit eksplisiittisenä muistina rekursiivisissa neuroverkoissa, ja käytämme sitä tekstin ydinviittaussuhteiden mallintamiseen. Sovellamme malliamme useisiin tekstinymmärtämistehtäviin ja saavutamme uusia huipputuloksia kaikilla tarkastelluilla vertailuarvoilla, mukaan lukien CNN, bAbi ja LAMBADA. Mallimme ratkaisee bAbi QA -tehtävissä 15 tehtävää 20:stä vain 1000 harjoitusesimerkillä tehtävää kohden. Opittujen representaatioiden analyysi osoittaa lisäksi mallimme kyvyn koodata hienojakoista entiteettitietoa koko asiakirjasta.

**Tulos**

Kielitieto rekursiivisten neuroverkkojen muistina

**Esimerkki 2.1250**

Variationaaliset autokooderit (VAE) ovat suunnattuja generatiivisia malleja, jotka oppivat faktoriaalisia latentteja muuttujia. Kuten Burda et al. (2015) huomasivat, näissä malleissa esiintyy faktorien ylikarsimisen ongelma, jossa huomattava määrä stokastisia faktoreita ei opi mitään ja jäävät toimimattomiksi. Tämä voi rajoittaa niiden mallinnustehoa ja kykyä oppia monipuolisia ja merkityksellisiä latentteja representaatioita. Tässä artikkelissa arvioimme useita menetelmiä tämän ongelman ratkaisemiseksi ja ehdotamme tehokkaampaa mallipohjaista lähestymistapaa, jota kutsutaan epitomic variational autoencoderiksi (eVAE). Tämän mallin niin sanotut epitomit ovat toisiaan poissulkevien latenttien tekijöiden ryhmiä, jotka kilpailevat tietojen selittämisestä. Tämä lähestymistapa auttaa estämään inaktiivisten yksiköiden syntymisen, koska kukin ryhmä joutuu selittämään dataa. Vertailemme lähestymistapoja kvalitatiivisilla ja kvantitatiivisilla tuloksilla MNIST- ja TFD-tietokannoissa. Tuloksemme osoittavat, että eVAE hyödyntää mallin kapasiteettia tehokkaasti ja yleistyy paremmin kuin VAE.

**Tulos**

Ylikarsinnan torjuminen varioivissa autokoodereissa

**Esimerkki 2.1251**

Asiantuntijajärjestelmät osoittautuvat sopiviksi ihmisasiantuntijoiden korvaajiksi silloin, kun ihmisasiantuntijat eivät ole käytettävissä eri syistä. Erilaisia asiantuntijajärjestelmiä on kehitetty monenlaisia sovelluksia varten. Kalastus- ja vesiviljelyalalla on kehitetty joitakin asiantuntijajärjestelmiä, mutta järjestelmää, joka auttaisi käyttäjää valitsemaan uutta lisäystä akvaarioaltaaseen, ei ole koskaan suunniteltu. Tässä asiakirjassa ehdotetaan asiantuntijajärjestelmää, joka ehdottaa uutta lisäystä akvaarioaltaaseen akvaarion nykyisen ympäristön tilan ja akvaariossa tällä hetkellä olevien kalojen perusteella. Järjestelmä ehdottaa akvaarion olosuhteisiin parhaiten sopivaa ja muiden kalojen kanssa parhaiten yhteensopivaa kalaa.

**Tulos**

Asiantuntijajärjestelmä, joka suosittelee akvaarioon sopivia koristekaloja akvaarion tilan perusteella

**Esimerkki 2.1252**

Tosielämän tiedot sisältävät usein tietoa eri kanavista. Esimerkiksi tietokonenäössä voimme kuvata kuvaa käyttämällä erilaisia kuvan ominaisuuksia, kuten pikselin intensiteettiä, väriä, HOG-, GIST- ja SIFT-ominaisuuksia jne. Näitä samojen kohteiden eri näkökohtia kutsutaan usein moninäkökulmaiseksi (tai multimodaaliseksi) dataksi. Lowrank-regressiomalli on osoittautunut tehokkaaksi oppimismekanismiksi tutkimalla tosielämän datan low-rank-rakennetta. Aiemmat matalarivin regressiomallit toimivat kuitenkin vain yhden näkymän datalla. Tässä artikkelissa ehdotetaan moninäkymäistä matalarivin regressiomallia asettamalla matalarivin rajoitteita moninäkymäiseen regressiomalliin. Tärkeintä on, että tarjoamme suljetun muodon ratkaisun moninäkymäiselle matalan rankin regressiomallille. Laajat kokeet neljällä eri näkymiä sisältävällä tietokokonaisuudella osoittavat, että usean näkymän matalarivinen regressiomalli päihittää yhden näkymän regressiomallin ja paljastaa, että usean näkymän matalarivinen rakenne on erittäin hyödyllinen. Johdanto Monissa tehtävissä yhtä kohdetta voidaan kuvata käyttämällä eri kanavien (tai näkymien) tietoja. Esimerkiksi kolmiulotteista kohdetta voidaan kuvata käyttämällä eri kuvakulmista otettuja kuvia, verkkosivustoa voidaan kuvata sen sisältämien sanojen ja hyperlinkkien avulla, kuvaa voidaan kuvata käyttämällä erilaisia ominaisuuksia, kuten SIFT-ominaisuutta ja HOG-ominaisuutta, ja jokapäiväisessä elämässä henkilöä voidaan luonnehtia iän, pituuden, painon ja niin edelleen avulla. Kaikki nämä tiedot ovat peräisin eri näkökulmista ja kanavista. Moninäkökulmaongelmissa pyritään parantamaan olemassa olevaa yhden näkymän mallia oppimalla malli, jossa hyödynnetään useista kanavista kerättyjä tietoja (Rüping ja Scheffer 2005) (de Sa 2005) (Zhou ja Burges 2007). Low-rank-regressiomalli on osoittautunut tehokkaaksi oppimismekanismiksi tutkimalla tosielämän datan low-rank-rakennetta (Xiang et al. 2012) (Evgeniou ja Pontil 2007) (Cai et al. 2013). Nykyiset regressiomallit toimivat vain yhden näkymän datalla. Tarkemmin sanottuna lineaarisella regressiolla etsitään lineaarinen malli yhden näkymän ominaisuustietojen suhteen, jotta se sopisi kohdeluokan tietoihin (Seber ja Lee 2012). Olkoon matriisi B ∈ <p×c lineaarisen mallin parametri. Lineaarinen regressio ratkaisee ongelman minB ||Y - XB||F , jossa Copyright c © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). Kaikki oikeudet pidätetään. X = [x1,x2, ...,xn] ∈ <p×n on yhden näkymän ominaisuustietomatriisi ja Y ∈ <n×c on kohdeluokan indikaattorimatriisi. Ridge-regressiolla voidaan saavuttaa parempia tuloksia lisäämällä Frobeniuksen normiin perustuva regularisointi lineaarisen regression tappiotavoitteeseen (Hoerl ja Kennard 1970) (Marquaridt 1970). Ridge-regressio ratkaisee ongelman minB ||Y -XB||F + λ||B||F , jossa λ on regularisointipainoparametri. Cai (Cai et al. 2013) osoitti, että kunB on matalarivinen, regressio vastaa lineaariseen diskriminaatioanalyysiin perustuvaa regressiota. Kaikki nämä työt toimivat kuitenkin vain yhden näkymän ongelmille. Tässä artikkelissa ehdotamme moninäkymäistä matalan rankin regressiomallia asettamalla regressiomalliin matalan rankin rajoituksia. Tämä malli voidaan ratkaista suoraan suljetussa muodossa. Lineaarisessa regressiossa matalarivinen parametrimatriisi B on riippuvainen näkymästä ν. Teoreettisen analyysin avulla osoitamme, että moninäkymäinen matalarivinen regressiomalli vastaa regressiota kunkin näkymän aliavaruudessa. Toisin sanoen, olkoon B = AνB, ja se vastaa jaetun regressioparametrimatriisin B löytämistä aliavaruuden muunnoksen Aν alla näkymän ν suhteen. Laajat kokeet, jotka on tehty neljällä eri näkymiä sisältävällä tietokokonaisuudella, osoittavat, että ehdotettu malli päihittää yhden näkymän regressiomallin, ja paljastaa, että matalarivirtainen rakenne voi parantaa täysirivisen mallin luokittelutulosta. Merkinnät. Tässä artikkelissa matriisit kirjoitetaan isoilla kirjaimilla, kuten X, Y . Vektorit kirjoitetaan lihavoidulla pienellä kirjaimella, kuten x, y. Tr(X) tarkoittaa matriisin X jäljitysoperaatiota. Oletetaan, että on olemassa v näkymää ja c luokkaa, pν on näkymän ν ulottuvuus, nj on j:nnen luokan otoskoko ja n on koko otoskoko. Olkoon Xν = [x1 , ...,x ν n] ∈ <pν×n näkymän ν tietomatriisi, ν = 1, 2, ..., v, ja Y = [y1, ...,yc] ∈ <n×c on normalisoitu luokkaindikaattorimatriisi, eli Yij = 1/ √ nj, jos i:nnen datapiste kuuluu j:nteen luokkaan, ja Yij = 0 muuten. Yritetään minimoida matalan luokan regressiomallin residuaali kussakin luokassa ja kussakin näkymässä. Loss function of multiProceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence 1973 ar X iv :1 61 0. 04 66 8v 1 [ cs .L G ] 1 4 O ct 2 01 6 view low rank ridge regression voidaan ehdottaa yhtälön (1) mukaisesti:

**Tulos**

Suljetussa muodossa oleva ratkaisu moninäkymäiseen matalarankaregressioon (Multi-View Low-Rank Regression)

**Esimerkki 2.1253**

Tässä artikkelissa ehdotamme monitehtäväoppimista varten monikonveksaalista kehystä, joka parantaa ennusteita oppimalla sekä tehtävien että ominaisuuksien välisiä suhteita. Kehyksemme on yleistetty monitehtäväoppimisen menetelmistä, jotka oppivat joko tehtävien tai ominaisuuksien suhteita, mutta eivät molempia. Aloitamme hierarkkisesta Bayes-mallista ja käytämme empiiristä Bayes-menetelmää muuttaaksemme taustalla olevan päättelyongelman monikonveksaaliseksi optimointiongelmaksi. Ehdotamme koordinaattiviisasta minimointialgoritmia, jolla on suljettu ratkaisu jokaiselle lohkon osaongelmalle. Naiivisti nämä ratkaisut olisivat kalliita laskea, mutta käyttämällä kaksinkertaisesti stokastisten matriisien teoriaa pystymme pelkistämään taustalla olevan matriisioptimoinnin osaongelman minimipainoiseksi täydellisen kaksijakoisen graafin täydellisen yhteensovittamisen ongelmaksi ja ratkaisemaan sen analyyttisesti ja tehokkaasti. Painojen oppimisen osaongelman ratkaisemiseksi ehdotamme kolmea eri strategiaa, mukaan lukien gradienttimenetelmä, jolla on lineaarinen konvergenssitakuu, kun instansseja ei ole jaettu useiden tehtävien kesken, ja Sylvesterin yhtälöön perustuva numeerinen ratkaisu, kun instanssit on jaettu. Osoitamme menetelmämme tehokkuuden sekä synteettisillä että todellisilla tietokokonaisuuksilla. Kokeet osoittavat, että ehdotettu optimointimenetelmä on kertaluokkaa nopeampi kuin valmis projisoitu gradienttimenetelmä, ja mallimme pystyy hyödyntämään useiden tehtävien ja ominaisuuksien välisiä korrelaatiorakenteita.

**Tulos**

Tehokas monitehtäväinen ominaisuuksien ja suhteiden oppiminen

**Esimerkki 2.1254**

Tässä artikkelissa esitellään menetelmiä, joilla voidaan hyökätä ja puolustaa k-grammien tilastollisia analyysitekniikoita, joita käytetään esimerkiksi verkkoliikenteen analysoinnissa ja peitekanavien havaitsemisessa. Tärkein uusi tulos on osoituksemme siitä, miten käyttäytymisen tai prosessin k-järjestyksen tilastoja voidaan käyttää sellaisen stokastisen prosessin rakentamiseen, jolla on samat k-järjestyksen stationääriset tilastot mutta jolla on haluttaessa erilaiset, tarkoituksellisesti suunnitellut (k + 1)-järjestyksen tilastot. Tällainen malli toteuttaa prosessin tai käyttäytymisen "monimutkaistamisen", jota puolustaja voi käyttää seuratakseen, muokkaako hyökkääjä käyttäytymistä. Ottamalla tarkoituksellisesti käyttöön suunniteltuja (k + 1)-järjestyksen käyttäytymismalleja, puolustaja voi tarkistaa, esiintyykö kyseisiä käyttäytymismalleja datassa. Kehitämme myös rakenteita lähdekoodeille, jotka kunnioittavat prosessin k-järjestyksen tilastoja samalla kun ne koodaavat salattua tietoa. Yksi näiden tulosten perustavanlaatuinen seuraus on, että tietyntyyppiset käyttäytymisen analysointitekniikat johtavat kilpavarusteluun siinä mielessä, että etulyöntiasema menee sille osapuolelle, jolla on enemmän laskentaresursseja ongelman ratkaisemiseksi. Tässä asiakirjassa esitetyt näkökannat ovat kirjoittajien omia, eivätkä ne välttämättä edusta rahoittajavirastojen tai Yhdysvaltain hallituksen virallista kantaa. V. Crespi työskentelee tietojenkäsittelytieteen laitoksella, California State University at Los Angeles, Los Angeles CA, 90032 USA. sähköposti: vcrespi@calstatela.edu. Crespin työtä on osittain tuettu AFOSR-avustuksella FA9550-07-1-0421 ja NSF-avustuksella HRD-0932421. G. Cybenko on Thayer School of Engineering, Dartmouth College, Hanover NH 03755. Sähköposti: gvc@dartmouth.edu. Cybenkon työtä ovat osittain tukeneet Air Force Research Laboratoryn sopimukset FA8750-10-1-0045 ja FA8750-09-1-0174, AFOSR:n sopimus FA9550-07-1-0421, Yhdysvaltain sisäisen turvallisuuden ministeriön avustus 2006-CS-001-000001 ja DARPA:n sopimus HR001-06-1-0033. A. Giani työskentelee Kalifornian yliopiston EECS-osastolla Berkeleyssä, Berkeley CA 94720. Sähköpostiosoite: agiani@eecs.berkeley.edu. Gianin työtä on osittain tuettu Yhdysvaltain sisäisen turvallisuuden ministeriön myöntämällä avustuksella 2006-CS001-000001 ja DARPA-sopimuksella HR001-06-1-0033 hänen ollessaan tohtorikoulutettavana Dartmouthissa 28. huhtikuuta 2011 LUONNOS ar X iv :1 10 4. 50 71 v1 [ cs .L G ] 2 7 A pr 2 01 1 CRESPI, CYBENKO, GIANI 2

**Tulos**

Peitekanavien ja käyttäytymismallien hyökkäys ja puolustaminen

**Esimerkki 2.1255**

Ontologiat kärsivät yleensä semanttisesta heterogeenisuudesta, kun niitä käytetään samanaikaisesti tiedon jakamiseen, yhdistämiseen, integrointiin ja kyselyihin. Siksi käytettyjen ontologioiden samankaltaisuuden tunnistamisesta tulee pakollinen tehtävä kaikissa näissä prosesseissa semanttisen heterogeenisuuden ongelman käsittelemiseksi. Tässä artikkelissa ehdotetaan tehokasta tekniikkaa kahden ontologian välisen samankaltaisuuden mittaamiseen. Ehdotettu tekniikka tunnistaa kaikki samankaltaisten käsitteiden ehdokasparit jättämättä yhtään samankaltaista paria pois. Ehdotettua tekniikkaa voidaan käyttää erityyppisissä ontologioihin kohdistuvissa operaatioissa, kuten yhdistämisessä, kartoittamisessa ja yhdenmukaistamisessa. Analysoimalla sen tuloksia on saavutettu kohtuullinen parannus tulosten täydellisyyden, oikeellisuuden ja yleisen laadun suhteen.

**Tulos**

Tehokas tekniikka ontologioiden välisen samankaltaisuuden tunnistamiseen

**Esimerkki 2.1256**

Pistepilven perusteella tarkastelemme 3D-nivellettyjen objektien, kuten laatikoiden, kinemaattisten mallien päättelyä niiden manipulointia varten. Vaikka aiemmissa töissä on osoitettu, miten voidaan poimia tasomainen kinemaattinen malli (joka esitetään usein lineaarisena ketjuna), tällaiset tasomallit eivät sovellu 3D-objekteihin, jotka koostuvat segmenteistä, jotka ovat usein yhteydessä toisiin segmentteihin syklisissä kokoonpanoissa. Esitämme lähestymistavan sellaisen mallin rakentamiseksi, joka kuvaa syötetyn pistepilven piirteiden ja objektin segmentin välistä suhdetta sekä objektin naapurisegmenttien välistä suhdetta. Käytämme ehdollista satunnaiskenttää, jonka avulla voimme mallintaa kohteen eri segmenttien väliset riippuvuudet. Testaamme lähestymistapaamme kinemaattisen rakenteen päättelemiseksi osittaisista ja kohinaisista pistepilvitiedoista monenlaisille laatikoille, kuten kakkulaatikoille, pizzalaatikoille ja erikokoisille pahvilaatikoille. Saavutetun rakenteen avulla robottimme pystyy sulkemaan nämä laatikot onnistuneesti manipuloimalla läppiä.

**Tulos**

3D-nivellettyjen mallien päätteleminen laatikkopakkausrobotille

**Esimerkki 2.1257**

Verkkopäätösongelmassa tehdään päätöksiä usein asiantuntijoiden kanssa, mutta ilman tietoa tulevaisuudesta. Jokaisen vaiheen jälkeen maksetaan kustannus, joka perustuu päätökseen ja havaittuun korkoon. Yksi järkevä tavoite olisi suoriutua yhtä hyvin kuin poolin paras asiantuntija. Nykyaikainen ja tunnettu tapa saavuttaa tämä tavoite on eksponentiaalisen painotuksen algoritmi. Hiljattain on kuitenkin kehitetty toinen algoritmi nimeltä follow the perturbed leader, jolla saavutetaan suunnilleen sama suorituskyky. Työssämme osoitamme ensin näiden kahden algoritmin yhteiset ominaisuudet, jotka selittävät suorituskyvyn samankaltaisuuden. Seuraavaksi osoitamme, että tietyllä häiriöllä nämä kaksi algoritmia ovat identtisiä. Lopuksi osoitamme joidenkin esimerkkien avulla, että follow-the-leader -tyyliset algoritmit ulottuvat luonnollisesti suureen luokkaan strukturoituja online-ongelmia, joihin eksponentiaaliset algoritmit ovat tehottomia. 1 Online-ongelma-asetelma Online-päätösongelmassa tehdään päätöksiä usein asiantuntijoiksi kutsutun päätöksentekosarjan kanssa, mutta ilman tietoa tulevaisuudesta. Jokaisen vaiheen jälkeen maksetaan kustannus, joka perustuu päätökseen ja havaittuun tilaan. Koska ei ole ennakkotietoa pooliin kuuluvien asiantuntijoiden tarkkuudesta, yksi järkevä tavoite tässä yleisessä ongelmassa olisi toimia yhtä hyvin kuin pooliin kuuluva paras asiantuntija useiden vaiheiden jälkeen. Tarkemmin sanottuna tarkastelemme seuraavaa matemaattista ongelmaa: - Annetaan joukko S asiantuntijoita. - Algoritmi on vuorovaikutuksessa vastustajan kanssa sarjassa T askelta. - Kussakin vaiheessa j algoritmi valitsee asiantuntijan xj ∈ S , ja vastustaja valitsee kustannusfunktion cj: S → R. Vastustaja voi olla mukautuva, jolloin cj voi riippua {xi : i < j}. - Algoritmi aiheuttaa kustannuksia , ja saa palautteena cj(xj):n arvon. - Minimoidaan algoritmin katumus, joka määritellään algoritmin valintasarjan ja S:n parhaan kiinteän asiantuntijan valintasarjan odotettujen kustannusten erotuksena:

**Tulos**

Siirtyminen Perturbed-järjestelmästä eksponentiaaliseen painotuskeskiarvoon.

**Esimerkki 2.1258**

Kuvaamme avoimen lähdekoodin työkalupaketin neuraalista konekääntämistä (NMT) varten. Työkalupakissa asetetaan etusijalle tehokkuus, modulaarisuus ja laajennettavuus tavoitteena tukea NMT-tutkimusta malliarkkitehtuurien, ominaisuuksien esittämisen ja lähdemodaliteettien osalta säilyttäen samalla kilpailukykyinen suorituskyky ja kohtuulliset koulutusvaatimukset. Työkalupaketti koostuu mallinnus- ja käännöstuesta sekä yksityiskohtaisesta pedagogisesta dokumentaatiosta taustalla olevista tekniikoista.

**Tulos**

OpenNMT: Avoimen lähdekoodin työkalupakki neuraalista konekääntämistä varten

**Esimerkki 2.1259**

Viime vuosikymmenen aikana ihmiset ovat kokeneet räjähdysmäisen kasvun verkkoresurssien, erityisesti sosiaalisen median ja mikrobloggaussivustojen, kuten Facebookin, Twitterin ja YouTuben, sekä mobiilisovellusten, kuten WhatsAppin ja Linen, käytössä. Monet yritykset ovat tunnistaneet nämä resurssit markkinointitiedon rikkaaksi kaivokseksi. Tämä tieto tarjoaa arvokasta palautetta, jonka avulla ne voivat kehittää tuotteensa seuraavaa sukupolvea. Tässä asiakirjassa tuotteen sentimenttianalyysi suoritetaan poimimalla kyseistä tuotetta koskevat twiitit ja luokittelemalla twiitit, jotka osoittavat sen positiivisen ja negatiivisen sentimentin. Kirjoittajat ehdottavat hybridi-lähestymistapaa, jossa yhdistetään valvomaton oppiminen K-means-klusteroinnin muodossa twiittien klusteroimiseksi ja sen jälkeen valvotut oppimismenetelmät, kuten päätöspuut ja tukivektorikoneet, luokittelua varten.

**Tulos**

Parannettu Twitter-tunnelmien ennustaminen "Cluster-then-Predict Model" -mallin avulla

**Esimerkki 2.1260**

MOBA (Multiplayer Online Battle Arena) on nykyään yksi pelatuimmista peligenreistä. Tämän lajityypin lisääntyessä on tarpeen kehittää tehokkaita älykkäitä agentteja pelaamaan ihmispelaajien rinnalla tai heitä vastaan. Tässä artikkelissa käsittelemme MOBA-pelien agenttien kehittämisen ongelmaa. Toteutamme kaksikerroksisen arkkitehtuurin agentin, joka huolehtii sekä navigoinnista että pelimekaniikasta. Tämä arkkitehtuuri perustuu vaikutuskarttojen käyttöön, joka on laajalti käytetty lähestymistapa taktiseen analyysiin. Useita kokeita tehtiin käyttäen League of Legends -peliä testialustana, ja ne osoittavat lupaavia tuloksia tässä erittäin dynaamisessa reaaliaikaympäristössä.

**Tulos**

Älykkäiden agenttien kehittäminen MOBA-pelejä varten

**Esimerkki 2.1261**

Esittelemme uudenlaisen ulottuvuuskohtaisen oppimisnopeusmenetelmän gradienttilaskeutumista varten nimellä ADADELTA. Menetelmä mukautuu dynaamisesti ajan mittaan käyttäen vain ensimmäisen kertaluvun tietoa, ja sen laskennallinen yleiskustannus on minimaalinen verrattuna vaniljaiseen stokastiseen gradienttilaskeutumiseen. Menetelmä ei vaadi oppimisnopeuden manuaalista virittämistä, ja se vaikuttaa kestävältä kohinaiselle gradienttitiedolle, erilaisille malliarkkitehtuurivalinnoille, erilaisille datamodaliteeteille ja hyperparametrien valinnalle. Näytämme lupaavia tuloksia muihin menetelmiin verrattuna MNIST-digitaaliluokitustehtävässä yhdellä koneella ja laajamittaisessa äänitietoaineistossa hajautetussa klusteriympäristössä.

**Tulos**

ADADELTA: MUKAUTUVA OPPIMISNOPEUSMENETELMÄ

**Esimerkki 2.1262**

Monikieliset puhutun dialogin järjestelmät ovat viime aikoina yleistyneet, mikä on edellyttänyt kielentunnistusjärjestelmää (LID). Useimmat nykyisistä LID-järjestelmistä perustuvat kielen erottelevan tiedon mallintamiseen matalan tason akustisista piirteistä. Puheessa esiintyvien vaihteluiden (puhujan ja tunteiden vaihtelut jne.) vuoksi matalan tason akustisia piirteitä käyttävät laajamittaiset LID-järjestelmät kärsivät suorituskyvyn heikkenemisestä. Tässä lähestymistavassa olemme pyrkineet mallintamaan korkeamman tason kielen erottelevaa fonotaktista tietoa LID-järjestelmän kehittämiseksi. Tässä työssä syötetty puhesignaali merkitään puhelinsekvensseiksi kielestä riippumattoman puhelintunnistimen avulla. Saatujen puhelinsekvenssien kielidiskriminoivaa fonotaktista tietoa mallinnetaan tilastollisilla ja rekursiivisiin neuroverkkoihin perustuvilla kielimallinnusmenetelmillä. Koska tämä lähestymistapa perustuu korkeamman tason fonotaktiseen informaatioon, se on kestävämpi puheen vaihteluille. Ehdotettu lähestymistapa on laskennallisesti kevyt, hyvin skaalautuva ja sitä voidaan käyttää nykyisten LID-järjestelmien rinnalla.

**Tulos**

KIELIMALLIIN PERUSTUVA LÄHESTYMISTAPA LAAJAMITTAISIIN JA KEVYISIIN KIELEN TUNNISTUSJÄRJESTELMIIN

**Esimerkki 2.1263**

Tärkeä Wikipedian muokkauskäytäntö on, että Wikipedian sivuille lisätyille lausunnoille annetaan viittaukset, kun lausunnot voivat olla mielivaltaisia tekstikappaleita, jotka voivat olla lauseesta kappaleeseen. Monissa tapauksissa viittaukset ovat joko vanhentuneita tai puuttuvat kokonaan. Tässä työssä käsittelemme ongelmaa, joka koskee uutisten viittausten löytämistä ja päivittämistä kokonaisuussivujen lausumille. Ehdotamme tähän ongelmaan kaksivaiheista valvottua lähestymistapaa. Ensimmäisessä vaiheessa rakennamme luokittelijan, jolla selvitetään, tarvitaanko lausumiin uutisviittaus vai muunlaisia viittauksia (verkko-, kirja-, aikakauslehti- jne.). Toisessa vaiheessa kehitämme Wikipedian lausuntoja varten uutisviittausalgoritmin, joka suosittelee sopivia viittauksia tietystä uutiskokoelmasta. IR-tekniikoiden lisäksi, jotka käyttävät lausuntoa uutiskokoelman kyselyyn, formalisoimme myös kolme sopivaa viittauksen ominaisuutta: (i) viittauksen tulisi sisältää Wikipedia-lausunto, (ii) lausunnon tulisi olla keskeinen viittauksen kannalta ja (iii) viittauksen tulisi olla peräisin arvovaltaisesta lähteestä. Suoritamme molempien vaiheiden laajan arvioinnin käyttäen 20 miljoonaa artikkelia todellisesta uutiskokoelmasta. Tuloksemme ovat varsin lupaavia ja osoittavat, että voimme suorittaa tämän tehtävän suurella tarkkuudella ja mittakaavassa.

**Tulos**

Wikipedian uutissitaattien löytäminen

**Esimerkki 2.1264**

Syviin neuroverkkoihin perustuvien menetelmien viimeaikainen mukauttaminen vahvistusoppimiseen ja suunnitteluun on tuottanut huomattavaa edistystä yksittäisissä tehtävissä. Edistyminen tehtävästä toiseen tapahtuvassa siirrossa on kuitenkin edelleen vähäistä. Tehokkaan ja vankan yleistyksen saavuttamiseksi esittelemme Schema Network -verkon, joka on objektikeskeinen generatiivinen fysiikkasimulaattori, joka pystyy erottelemaan tapahtumien useita syitä ja päättelemään syiden kautta taaksepäin tavoitteiden saavuttamiseksi. Schema Networkin rikkaasti jäsennelty arkkitehtuuri voi oppia ympäristön dynamiikan suoraan datasta. Vertailemme Schema-verkkoja asynkroniseen Advantage Actor-Criticiin ja progressiivisiin verkkoihin Breakout-muunnelmissa, ja raportoimme tulokset koulutuksen tehokkuudesta ja nollapisteen yleistyksestä, jotka osoittavat johdonmukaisesti nopeampaa, vankempaa oppimista ja parempaa siirtoa. Väitämme, että yleistäminen rajallisista tiedoista ja kausaalisuhteiden oppiminen ovat olennaisia kykyjä matkalla kohti yleisesti älykkäitä järjestelmiä.

**Tulos**

Skeemaverkot: Intuitiivisen fysiikan generatiivisen kausaalimallin avulla tapahtuva nollapistesiirto

**Esimerkki 2.1265**

Aiemmin tunnetun rakenteellisen tiedon hyödyntäminen on monien kuvien rekonstruktiomenetelmien ydin, jotka voidaan ilmaista käänteisongelmina. Synteesimalli, jossa oletetaan, että kuvat voidaan hajottaa jonkin sanakirjan hyvin harvojen atomien lineaariseksi yhdistelmäksi, on nykyään vakiintunut työkalu kuvan rekonstruktio-algoritmien suunnittelussa. Mielenkiintoinen vaihtoehto on analyysimalli, jossa signaali kerrotaan analyysioperaattorilla ja tuloksen oletetaan olevan harva. Tämä lähestymistapa on vasta viime aikoina herättänyt kasvavaa kiinnostusta. Analyysimalliin perustuvien rekonstruktiomenetelmien laatu riippuu suuresti sopivan operaattorin oikeasta valinnasta. Tässä työssä esitellään algoritmi analyysioperaattorin oppimiseksi harjoituskuvista. Menetelmämme perustuu `p-normin minimointiin joukolle full rank -matriiseja, joiden sarakkeet on normalisoitu. Esittelemme huolellisesti käytetyn konjugoituneen gradientin menetelmän moninaisuuksilla ja selitämme rajoitusten taustalla olevan geometrian. Lisäksi vertaamme lähestymistapaamme uusimpiin menetelmiin, joita käytetään kuvien kohinanpoistossa, maalauksessa ja yksittäisen kuvan superresoluutiossa. Numeeriset tuloksemme osoittavat, että yleinen lähestymistapamme on kilpailukykyinen kaikissa esitetyissä sovelluksissa verrattuna erikoistuneisiin uusimpiin tekniikoihin.

**Tulos**

Analyysioperaattorin oppiminen ja sen soveltaminen kuvan rekonstruktioon

**Esimerkki 2.1266**

Esittelemme kaksi algoritmia tarkkaa ja ap-läheistä päättelyä varten kausaaliverkoissa. Ensimmäinen algoritmi, dynaaminen ehdollistaminen, on leikkausjoukkoehdollistamisen parannus, jonka monimutkaisuus joissakin verkoissa, joissa leikkausjoukkoehdollistaminen on eksponentiaalinen. Toinen algoritmi, B-conditioning, on approksimatiivisen päättelyn algoritmi, jonka avulla voidaan tehdä kompromissi approksimaatioiden laadusta ja laskenta-ajasta. Esitämme myös joitakin kokeellisia tuloksia, jotka havainnollistavat ehdotettujen algoritmien ominaisuuksia.

**Tulos**

Ehdolliset algoritmit täsmällistä ja likimääräistä päättelyä varten kausaaliverkostoissa.

**Esimerkki 2.1267**

Esittelemme mallin kuvien ja lauseiden kaksisuuntaiseen hakuun visuaalisen ja luonnollisen kielitiedon multimodaalisen upottamisen avulla. Toisin kuin aiemmat mallit, jotka kartoittavat kuvia tai lauseita suoraan yhteiseen upotusavaruuteen, meidän mallimme toimii hienojakoisemmalla tasolla ja upottaa kuvien fragmentteja (objekteja) ja lauseiden fragmentteja (tyypitettyjä riippuvuuspuusuhteita) yhteiseen avaruuteen. Aiemmissa töissä nähdyn luokittelutavoitteen lisäksi voimme näin lisätä uuden fragmenttien kohdistamistavoitteen, joka oppii yhdistämään nämä fragmentit suoraan eri modaliteettien välillä. Laaja kokeellinen arviointi osoittaa, että päättely sekä kuvien ja lauseiden yleisellä tasolla että niiden fragmenttien tarkemmalla tasolla parantaa merkittävästi suorituskykyä kuvien ja lauseiden hakutehtävissä. Lisäksi mallimme tuottaa tulkittavissa olevia ennusteita, koska modaliteettien välinen fragmenttien kohdistaminen on eksplisiittistä.

**Tulos**

Syvä fragmenttien sulauttaminen kaksisuuntaiseen kuvien lauseiden kartoitukseen

**Esimerkki 2.1268**

Esittelemme lähestymistavan, jonka avulla voidaan luoda uusia tietokonepelien tasoja, joissa yhdistyvät erilaiset pelikäsitteet ilman valvontaa. Ensisijainen panoksemme on analoginen päättelyprosessi, jonka avulla voidaan rakentaa pelivideoista opittujen tasosuunnittelumallien välisiä sekoituksia. Mallit edustavat pelin elementtien välisiä todennäköisyyssuhteita. Analoginen päättelyprosessi kartoittaa piirteitä kahden mallin välillä ja tuottaa sekoitettuja malleja, joiden avulla voidaan luoda uusia tasokokonaisuuksia. Proof-of-concept-menetelmänä koulutamme järjestelmäämme klassisella tasohyppelypelillä Super Mario Bros., koska sen tasosuunnittelua arvostetaan ja ymmärretään hyvin. Arvioimme, missä määrin mallit edustavat tyylillistä tasosuunnittelutietämystä, ja osoitamme järjestelmämme kyvyn selittää tasoja, jotka on sekoitettu ihmisten asiantuntijasuunnittelijoiden toimesta.

**Tulos**

Tietokonepelien tasojen yhdistämisen opettelu

**Esimerkki 2.1269**

Standardialgoritmit lyhimmän polun löytämiseksi graafista edellyttävät, että polun kustannukset ovat additiivisia reunakustannuksiin nähden, ja yleensä oletetaan, että kustannukset ovat deterministisiä. Tarkastelemme ongelmaa, jossa reunakustannukset ovat epävarmoja ja kustannusten välillä on mahdollisia todennäköisyyteen perustuvia riippuvuuksia. Vaikka nämä riippuvuudet rikkovat tavanomaisen dynaamisen ohjelmoinnin purkamisen, tunnistamme heikomman stokastisen johdonmukaisuusehdon, joka oikeuttaa yleistetyn dynaamisen ohjelmoinnin lähestymistavan, joka perustuu stokastiseen dominanssiin. Esittelemme tarkistetun polkusuunnittelualgoritmin ja todistamme, että se tuottaa optimaalisia polkuja ajasta riippuvien epävarmojen kustannusten vallitessa. Testaamme algoritmia soveltamalla sitä stokastisten linja-autoverkkojen malliin ja esitämme empiirisiä suorituskykytuloksia, joissa sitä verrataan joihinkin vaihtoehtoihin. Lopuksi tarkastelemme näiden käsitteiden soveltamista yleisempiin heuristisen etsinnän ongelmiin epävarmuuden vallitessa.

**Tulos**

Reitin suunnittelu ajasta riippuvan epävarmuuden vallitessa

**Esimerkki 2.1270**

Miten rakennamme automaattisesti tietämysgraafin (Knowledge Graph, KG) tietylle alalle, kun meillä on joukko tietyn alan asiakirjoja (esim. lääketieteelliset tutkimuslehdet)? Relaatioiden ja niiden skeemojen automaattinen tunnistaminen, eli relaatioiden argumenttien (esim. undergo(Patient, Surgery)) tyyppisignatuuri, on tärkeä ensimmäinen askel kohti tätä tavoitetta. Kutsumme tätä ongelmaa nimellä Relation Schema Induction (RSI). Tässä artikkelissa ehdotamme Schema Induction using Coupled Tensor Factorization (SICTF) -menetelmää, joka on uusi tensorifaktorointimenetelmä relaatioskeemojen induktiota varten. SICTF faktorisoi Open Information Extraction (OpenIE) -kolmioita, jotka on poimittu verkkotunnuksen korpuksesta, yhdessä ylimääräisen sivutiedon kanssa periaatteellisella tavalla relaatioskeemojen indusoimiseksi. Tietojemme mukaan tämä on ensimmäinen tensorifaktoroinnin sovellus RSI-ongelmaan. Laajoissa kokeissa useilla reaalimaailman tietokokonaisuuksilla havaitsimme, että SICTF on paitsi tarkempi kuin uusimmat perusratkaisut, myös huomattavasti nopeampi (noin 14 kertaa nopeampi).

**Tulos**

Relaatioskeeman induktio käyttämällä tensorifaktorointia sivutiedon kanssa

**Esimerkki 2.1271**

Tutkijat ovat kiinnittäneet huomiota asennon havaitsemiseen eli puhujan mielipiteen tunnistamiseen tiettyä kohdetta kohtaan. Tässä artikkelissa kuvataan uusi lähestymistapa kannan havaitsemiseen Twitterissä. Määrittelemme joukon piirteitä, joiden avulla voidaan ottaa huomioon kiinnostuksen kohteena olevaa kohdetta ympäröivä konteksti, ja lopullisena tavoitteena on kouluttaa malli, jonka avulla voidaan ennustaa asennoituminen mainittuja kohteita kohtaan. Olemme erityisesti kiinnostuneita tutkimaan poliittisia keskusteluja sosiaalisessa mediassa. Tästä syystä arvioimme lähestymistapaamme keskittyen SemEval-2016 Task 6 on Detecting stance in tweets -tehtävän kahteen kohteeseen, jotka liittyvät Yhdysvaltain vuoden 2016 presidentinvaalien poliittiseen kampanjaan: Hillary Clinton vs. Donald Trump. Vertailun vuoksi tekniikan tasoon arvioimme malliamme SemEval-2016 Task 6 -tehtäväkilpailussa julkaistua tietokokonaisuutta vastaan. Tuloksemme ylittävät osallistuvien tiimien saamat parhaat tulokset ja osoittavat, että tiedot poliitikkojen vihollisista ja ystävistä auttavat havaitsemaan asennoitumisen heitä kohtaan.

**Tulos**

Clintonin ja Trumpin ystävät ja viholliset: Context for Detecting Stance in Political Tweets: Using Context for Detecting Stance in Political Tweets".

**Esimerkki 2.1272**

Valtavat määrät verkkotekstiä, kuten luokitellut ilmoitukset, huutokauppaluettelot, foorumikirjoitukset jne. ovat jäsentymätöntä ja epäkieliopillista tekstiä. Kutsumme tällaista tekstiä "viesteiksi". Epäjohdonmukaisesta rakenteesta ja kieliopin puutteesta huolimatta viestit ovat täynnä hyödyllistä tietoa. Tässä artikkelissa esitellään työtä, jossa puoliautomaattisesti rakennetaan relaatiotietotaulukoita, joita kutsutaan "viitejoukoiksi", analysoimalla tällaisia viestejä suoraan. Viitejoukkoja voidaan soveltaa useisiin tehtäviin, kuten ontologian ylläpitoon ja tiedon louhintaan. Viitejoukkojen rakentamismenetelmämme lähtee liikkeelle vain pienestä määrästä taustatietoa ja muodostaa viestien sisältämiä entiteettejä kuvaavia tupleja viitejoukon muodostamiseksi. Kuvaamme myös tämän lähestymistavan laajennuksen erityistapausta varten, jossa edes tätä pientä määrää taustatietoa on mahdotonta löytää ja käyttää. Arvioidaksemme koneellisesti muodostettujen viitejoukkojen hyödyllisyyttä vertaamme niitä manuaalisesti muodostettuihin viitejoukkoihin viitejoukkoihin perustuvan tiedonlouhinnan yhteydessä. Tuloksemme osoittavat, että menetelmällämme rakennetut viitejoukot ovat parempia kuin manuaalisesti rakennetut viitejoukot. Vertaamme myös viitejoukkopohjaista tiedonlouhintamenetelmää, jossa käytetään koneellisesti rakennettuja viitejoukkoja, valvottuihin tiedonlouhintamenetelmiin, joissa käytetään yleisiä piirteitä. Nämä tulokset osoittavat, että koneellisesti muodostettujen viitejoukkojen käyttö on parempi kuin valvotut menetelmät, vaikka valvotut menetelmät vaativat harjoitusdataa.

**Tulos**

Viitejoukkojen rakentaminen jäsentymättömästä ja epäkieliopillisesta tekstistä.

**Esimerkki 2.1273**

Hiljattain esitelty sarja DLLite-nimellä kulkevia kuvauslogiikoita on herättänyt huomiota kuvauslogiikka- ja semanttisen webin yhteisöissä johtuen yhtäältä päättelyn alhaisesta laskennallisesta monimutkaisuudesta ja toisaalta kyvystä esittää käsitteellisiä mallinnusformalismeja. Tämän artikkelin päätavoitteena on tutkia perusteellisesti ja systemaattisesti päättelyä alkuperäisen DL-Lite-logiikan laajennuksissa viidellä akselilla: (i) lisäämällä Boolen konnektiivit ja (ii) lukumäärärajoitukset käsitekonstruktioihin, (iii) sallimalla roolihierarkiat, (iv) sallimalla roolien disjointness-, symmetria-, epäsymmetria-, refleksiivisyys-, irrefleksiivisyys- ja transitiivisuusrajoitukset sekä (v) hyväksymällä tai poistamalla uniikkinimiehdotus. Analysoimme tuloksena syntyvien logiikoiden tyydyttävyyden yhdistettyä monimutkaisuutta sekä instanssien tarkistamisen ja positiivisiin eksistentiaalisiin kyselyihin vastaamisen datakompleksisuutta. Lähestymistapamme perustuu DL-Lite-logiikoiden upottamiseen sopiviin yhden muuttujan ensimmäisen järjestyksen logiikan fragmentteihin, mikä antaa hyödyllistä tietoa niiden ominaisuuksista ja erityisesti laskennallisesta käyttäytymisestä.

**Tulos**

DL-Lite-perhe ja suhteet

**Esimerkki 2.1274**

Konekääntämisessä on yleinen ilmiö, että konekieliset sanakirjat ja vakiomuotoiset jäsennyssäännöt eivät riitä varmistamaan tarkkuutta englanninkielisten lauseiden jäsennyksessä ja kääntämisessä korean kielelle, mikä näkyy harhaanjohtavina käännöstuloksina, jotka johtuvat rakenteellisista ja semanttisista epäselvyyksistä. Tämän artikkelin tavoitteena on ehdottaa ratkaisua rakenteellisiin ja semanttisiin epäselvyyksiin, jotka johtuvat englannin kielessä yleisesti käytettyjen fraasien idiomaattisuudesta ja epäkieliopillisuudesta, soveltamalla kaksikielistä fraasitietokantaa englannin ja korean konekääntämisessä (EKMT). Tässä asiakirjassa selvitetään ensinnäkin, mikä on EKMT:n lauseyksikkö englanninkielisen lauseen määritelmän perusteella, toiseksi selvitetään, millainen kieliyksikkö voi olla EKMT:n lausetietokannan kohde, kolmanneksi ehdotetaan tapaa rakentaa lausetietokanta esittelemällä esimerkkien avulla lausetietokannan muoto ja lopuksi käsitellään lyhyesti menetelmää, jolla tätä kaksikielistä lausetietokantaa voidaan soveltaa EKMT:ssä rakenteelliseen ja semanttiseen disambiguointiin.

**Tulos**

Lause-tietokanta Lähestymistapa rakenteelliseen ja semanttiseen disambiguointiin englannin ja korean kielen konekäännöksessä.

**Esimerkki 2.1275**

Tieto suurimmasta nopeudesta, jolla dataa voidaan lähettää päästä päähän -reitillä siten, että poistumisnopeus on suurella todennäköisyydellä yhtä suuri kuin saapumisnopeus, voi olla hyvin käytännöllinen, kun valitaan siirtonopeuksia videon suoratoistossa tai vertaisverkkosovellusten vertaisvertaisverkkosovelluksissa. Otamme käyttöön probabilistisen käytettävissä olevan kaistanleveyden, joka määritellään polun liikenteen sisäänmeno- ja ulostulonopeuksien perusteella eikä polun muodostavien linkkien kapasiteetin ja käyttöasteen perusteella, kuten tavallinen käytettävissä olevan kaistanleveyden metriikka. Tässä artikkelissa kuvaamme todennäköisyysgrafiseen malliin ja Bayesin aktiiviseen oppimiseen perustuvan hajautetun algoritmin, jolla voidaan samanaikaisesti arvioida useiden verkon läpi kulkevien polkujen todennäköistä käytettävissä olevaa kaistanleveyttä. Menetelmämme hyödyntää sitä, että jokainen pakettijuna antaa tietoa paitsi kulkemistaan poluista myös kaikista poluista, joilla on yhteinen linkki seurattavan polun kanssa. Simulaatiot ja PlanetLab-kokeet osoittavat, että tämä menettely voi vähentää huomattavasti tarkkojen arvioiden tuottamiseen tarvittavien koettimien määrää.

**Tulos**

Käytettävissä olevan kaistanleveyden todennäköisyysarviointi usean polun kautta Bayesin aktiivisen oppimisen avulla.

**Esimerkki 2.1276**

<lb>Tutkimme parhaan käsivarren tunnistamisongelmaa (Best-1-Arm), joka on määritelty seuraavasti.<lb>Meille annetaan n stokastista bandit-kättä. Seitsemännellä käsivarrella on palkitsemisjakauma<lb>Di, jonka<lb>keskiarvo μi on tuntematon. Jokaisesta i:nnen käsivarren pelistä saamme palkkion, joka on i.i.d. otos<lb>Di:stä. Haluamme tunnistaa käden, jolla on suurin keskiarvo, todennäköisyydellä vähintään 1 - δ,<lb> käyttäen mahdollisimman vähän näytteitä. Esitämme ei-triviaalin algoritmin Best-1-Armille, joka<lb>parantaa useita aiempia ylärajoja samalle ongelmalle. Tutkimme myös tärkeää<lb>erityistapausta, jossa on vain kaksi vartta, jota kutsumme Sign-ξ-ongelmaksi. Esitämme<lb>uuden alemman rajan Sign-ξ:lle, yksinkertaistamme ja laajennamme merkittävästi Farrellin<lb> vuonna 1964 tekemää klassista tulosta ja esitämme täysin uuden todistuksen. Sign-ξ:n uuden alarajan avulla saamme<lb>ensimmäisen alarajan Best-1-Armille, joka ylittää klassisen Mannor-Tsitsiklisin alarajan,<lb> mielenkiintoisella reduktiolla Sign-ξ:stä Best-1-Armiin. Esitämme mielenkiintoisen olettamuksen<lb>, joka koskee Best-1-Armin optimaalista näytteenottokompleksisuutta instanssiviisaan<lb>optimaalisuuden näkökulmasta.

**Tulos**

Optimaalinen näytteen monimutkaisuus parhaan käsivarren tunnistamista varten

**Esimerkki 2.1277**

Automaattista tekstin tiivistämistä pidetään yleisesti erittäin vaikeana ongelmana, mikä johtuu osittain siitä, että suuria tekstin tiivistämistietoaineistoja ei ole saatavilla. Koska kokotekstin laajamittaisten tiivistelmien laatiminen on suuri haaste, tässä asiakirjassa esitellään suuri kiinalaisen lyhyen tekstin tiivistämistietokokonaisuus, joka on luotu kiinalaisesta mikroblogisivustosta Sina Weibo, joka on julkaistu yleisölle1. Tämä korpus koostuu yli 2 miljoonasta todellisesta kiinalaisesta lyhyestä tekstistä, joissa on kunkin tekstin kirjoittajan antamat lyhyet tiivistelmät. Merkitsimme myös manuaalisesti 10 666 lyhyen tiivistelmän relevanssin vastaaviin lyhyisiin teksteihin. Korpuksen perusteella otamme käyttöön rekursiivisen neuroverkon tiivistelmän tuottamiseen ja saamme lupaavia tuloksia, mikä osoittaa ehdotetun korpuksen hyödyllisyyden lyhyiden tekstien tiivistämisen tutkimuksessa ja tarjoaa myös perustan aiheen jatkotutkimukselle.

**Tulos**

LCSTS: Laaja kiinalainen lyhyen tekstin tiivistämistietokanta: Large Scale Chinese Short Text Summarization Dataset

**Esimerkki 2.1278**

Hiljattain esitelty Deep Q-Networks (DQN) -algoritmi on saanut huomiota yhtenä ensimmäisistä onnistuneista syvien neuroverkkojen ja vahvistusoppimisen yhdistelmistä. Sen lupaukset osoitettiin Arcade Learning Environment (ALE) -ohjelmassa, joka on kymmenistä Atari 2600 -peleistä koostuva haastava kehys, jota käytetään tekoälyn yleisen osaamisen arviointiin. Se saavutti huomattavasti parempia tuloksia kuin aiemmat lähestymistavat, mikä osoittaa, että sen kyky oppia hyviä representaatioita on varsin vankka ja yleinen. Tässä artikkelissa pyritään ymmärtämään periaatteita, jotka ovat DQN:n vaikuttavan suorituskyvyn taustalla, ja selventämään paremmin sen menestystä. Arvioimme systemaattisesti DQN:n verkon koodaamien keskeisten representaatioharhojen merkitystä ehdottamalla yksinkertaisia lineaarisia representaatioita, jotka hyödyntävät näitä käsitteitä. Kun nämä ominaisuudet otetaan huomioon, saadaan laskennallisesti käytännöllinen ominaisuusjoukko, jolla saavutetaan DQN:n kanssa kilpailukykyinen suorituskyky ALE:ssa. Sen lisäksi, että tarjoamme tietoa DQN:n vahvuuksista ja heikkouksista, tarjoamme ALE:lle yleisen esityksen, joka vähentää merkittävästi taakkaa, joka aiheutuu esityksen oppimisesta jokaista peliä varten. Lisäksi tarjoamme myös yksinkertaisen, toistettavissa olevan vertailukohteen, jota voidaan verrata tulevaan työhön ALE:n parissa.

**Tulos**

Atari-pelien uusin ohjaus käyttäen matalaa vahvistusoppimista (Shallow Reinforcement Learning)

**Esimerkki 2.1279**

Verkko-opetuksessa pyritään lähes yhtä hyvään suoritukseen kuin parhaassa hypoteesissa jälkikäteen. Joillekin hypoteesiluokille parhaan hypoteesin löytäminen offline-oppimisessa on kuitenkin haastavaa. Tällaisissa offline-tapauksissa käytetään usein paikallisia hakutekniikoita ja taataan vain paikallinen optimi. Tällaisia hypoteesiluokkia koskevaa online-päätöksentekoa varten otamme käyttöön paikallisen katumuksen, katumuksen yleistyksen, jonka tavoitteena on suoriutua lähes yhtä hyvin kuin vain lähellä olevat hypoteesit. Sitten esittelemme yleisen algoritmin, jolla minimoidaan paikallinen katumus mielivaltaisilla paikallisuusgrafiikoilla. Näytämme myös, miten graafirakennetta voidaan hyödyntää oppimisen nopeuttamiseksi huomattavasti. Näitä algoritmeja demonstroidaan sitten erilaisilla online-ongelmilla: online-disjunktio-oppiminen, online-Max-SAT ja online-päätöspuuoppiminen.

**Tulos**

Paikallisesta katumuksesta

**Esimerkki 2.1280**

Schatten-p-kvasi-normia (0<p<1) käytetään yleensä korvaamaan vakioydinnormi, jotta rank-funktiota voidaan approksimoida tarkemmin. Nykyiset Schattenp-kvasinormin minimointialgoritmit sisältävät kuitenkin singulaariarvojen hajotuksen (SVD) tai ominaisarvojen hajotuksen (EVD) jokaisessa iteraatiossa, joten niistä voi tulla hyvin hitaita ja epäkäytännöllisiä suurissa ongelmissa. Tässä artikkelissa määritellään ensin kaksi helposti lähestyttävää Schatten-kvasinormia, Frobenius/nukleaarinen hybridi ja bi-ukleaarinen kvasinormi, ja sitten osoitetaan, että ne ovat pohjimmiltaan Schatten-2/3- ja 1/2-kvasinormit, jotka johtavat erittäin tehokkaiden algoritmien suunnitteluun, joissa tarvitsee päivittää vain kaksi paljon pienempää tekijämatriisia. Suunnittelemme myös kaksi tehokasta proksimaalista vuorottelevaa linearisoitua minimointialgoritmia edustavien matriisien täydennysongelmien ratkaisemiseksi. Lopuksi annamme globaalin konvergenssin ja suorituskykytakeet algoritmeillemme, joilla on paremmat konvergenssiominaisuudet kuin nykyisillä algoritmeilla. Synteettisellä ja todellisella datalla tehdyt kokeelliset tulokset osoittavat, että algoritmimme ovat tarkempia kuin nykyiset menetelmät ja kertaluokkaa nopeampia. Johdanto Viime vuosina matriisien sijan minimointiongelma on esiintynyt monissa eri sovelluksissa, kuten matriisien täydentämisessä, robustissa pääkomponenttianalyysissä, matalan sijan esityksessä, monimuuttujaregressiossa ja monitehtäväoppimisessa. Tällaisten ongelmien ratkaisemiseksi Fazel, Hindi ja Boyd; Candès ja Tao; Recht, Fazel ja Parrilo (2001; 2010; 2010) ovat ehdottaneet, että rank-funktiota relaksoidaan sen koveralla kuorella eli ydinnormilla. Ydinnormi vastaa itse asiassa matriisin singulaariarvojen l1-normia, ja siten se edistää matalarivistä ratkaisua. Kuitenkin on osoitettu (Fan ja Li 2001), että l1-normin regularisointi rankaisee liikaa vektoreiden suuria merkintöjä ja johtaa vinoutuneeseen ratkaisuun. Toteuttamalla niiden välisen läheisen suhteen ydinnormin rangaistus myös ylikorostaa suuria singulaariarvoja, eli se voi saada ratkaisun poikkeamaan alkuperäisestä ratkaisusta, kuten l1-normi tekee (Nie, Huang ja Ding 2012; Lu et al. 2015). Ydinnormiin verrattuna Schatten-p-kvasi-normi, kun 0 < p < 1, on lähempänä Copyright c © 2016, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved. approksimaatio rank-funktiolle. Näin ollen Schatten-p-kvasi-normin minimointi on herättänyt paljon huomiota kuvien talteenotossa (Lu ja Zhang 2014; Lu et al. 2014), yhteistoiminnallisessa suodatuksessa (Nie et al. 2012; Lu et al. 2015; Mohan ja Fazel 2012) ja MRI-analyysissä (Majumdar ja Ward 2011). Lisäksi monia ei-konveksaalisia l0-normin korvikefunktioita, jotka on lueteltu (Lu et al. 2014; Lu et al. 2015), on laajennettu approksimoimaan rank-funktiota, kuten SCAD (Fan ja Li 2001) ja MCP (Zhang 2010). Kaikki edellä mainitut ei-konveksaaliset sijaisfunktiot matalan rankin minimointia varten johtavat joihinkin ei-konveksaalisiin, ei-sileisiin, jopa ei-Lipschitzin optimointiongelmiin. Siksi on ratkaisevan tärkeää kehittää nopeita ja skaalautuvia algoritmeja, jotka ovat erikoistuneet ratkaisemaan joitakin vaihtoehtoisia muotoiluja. Tähän mennessä Lai, Xu ja Yin (2013) ehdottivat iteratiivista uudelleen painotettua vuokrannormineliöalgoritmia (IRucLq) Schatten-p:n kvasinormin minimointiongelman approksimoimiseksi ja todistivat, että heidän algoritminsa tuottaman minkä tahansa konvergentin osajonon rajakohta on kriittinen piste. Lisäksi Lu et al. (2014) ehdottivat iteratiivisesti uudelleenpainotettua ydinnormin (IRNN) algoritmia monien ei-konveksaalisten sijaisnormin minimointiongelmien ratkaisemiseksi. Matriisien täydennysongelmissa Schatten-p-kvasi-normin on osoitettu olevan empiirisesti ydinnormia parempi (Marjanovic ja Solo 2012). Lisäksi Zhang, Huang ja Zhang (2013) osoittivat teoreettisesti, että Schatten-p-kvasi-normin minimointi pienellä p:llä vaatii huomattavasti vähemmän mittauksia kuin kovera ydinnormiminimointi. Kaikki olemassa olevat algoritmit on kuitenkin ratkaistava iteratiivisesti ja niihin liittyy SVD tai EVD jokaisessa iteraatiossa, mikä aiheuttaa suuria laskentakustannuksia ja on liian kallista suurten ongelmien ratkaisemiseen (Cai ja Osher 2013; Liu et al. 2014). Sitä vastoin ydinnormin vaihtoehtoisena ei-konveksaalisena muotoiluna bilineaarista spektraalista regularisointia, kuten (Srebro, Rennie ja Jaakkola 2004; Recht, Fazel ja Parrilo 2010), on sovellettu menestyksekkäästi monissa laajamittaisissa sovelluksissa, esimerkiksi yhteistoiminnallisessa suodatuksessa (Mitra, Sheorey ja Chellappa 2010). Koska Schatten-p-kvasinormi vastaa lp-kvasinormia matriisin singulaariarvoille, on luontevaa esittää seuraava kysymys: voimmeko suunnitella ekvivalentteja matriisifaktorisointimuotoja Schatten-kvasinormin tapauksille, esim. p = 2/3 tai 1/2? Vastataksemme edellä esitettyyn kysymykseen määrittelemme tässä artikkelissa<lb>ensimmäiseksi kaksi käsiteltävää Schatten-kvasi-mormia, eli<lb>Frobenius/nukleaarisen hybridin ja bi-ydin-kvasi-muodon. Todistamme sitten, että ne ovat pohjimmiltaan Schatten-2/3- ja 1/2<lb>kvasinormit, joiden minimoinnin ratkaisemiseksi<lb>meidän tarvitsee suorittaa SVD:t vain kahdelle paljon pienemmälle fac-<lb>tor-matriisille toisin kuin nykyisissä<lb>algoritmeissa, esimerkiksi IRNN:ssä, käytetyissä suuremmissa matriiseissa. Siksi menetelmämme on erityisen<lb>hyödyllinen monissa "big data" -sovelluksissa, joissa täytyy<lb>käsitellä suuria, korkea-ulotteisia tietoja, joissa on puuttuvia arvoja.<lb>Tietojemme mukaan tämä on ensimmäinen artikkeli, jossa skaalataan<lb>Schattenin kvasinorminratkaisijoita Netflix-tietokantaan. Lisäksi annamme algoritmeillemme globaalin konvergenssin ja palautumisen per-<lb>suorituskyvyn takuut. Toisin sanoen, tämä<lb>on paras taattu konvergenssi algoritmeille, jotka ratkaisevat<lb>sellaisia haastavia ongelmia. Merkinnät ja tausta<lb>Matriisin X ∈ Rm×n Schatten-p normi (0 < p < ∞)<lb>.(m ≥ n) määritellään seuraavasti: ‖X‖Sp ,<lb>(<lb>n<lb>∑ i=1<lb>σ<lb>i (X)<lb>)1/p

**Tulos**

Skaalautuvat algoritmit hallittavissa olevaan Schatten-kvasinormin minimointiin

**Esimerkki 2.1281**

Esittelemme peräkkäisen mallin, jolla voidaan luokitella ajallisia suhteita lauseen sisäisten tapahtumien välillä. Keskeinen havainto on, että tapahtumien välisen monisanaisen kontekstin yleinen syntaktinen rakenne ja koostumukselliset merkitykset ovat tärkeitä hienojakoisten ajallisten suhteiden erottamisessa. Tarkemmin sanottuna lähestymistapamme poimii ensin kontekstisanojen sarjan, joka osoittaa kahden tapahtuman välisen ajallisen suhteen, joka vastaa hyvin kahden tapahtumailmoituksen välistä riippuvuuspolkua. Tämän jälkeen kontekstisanasekvenssi sekä sanasekvenssiä vastaava sananosien merkintäsekvenssi ja riippuvuussuhdesekvenssi, jotka luodaan sanasekvenssiä vastaaviksi, annetaan syötteenä kaksisuuntaisille rekursiivisille neuroverkkomalleille (LSTM). Neuroverkot oppivat koostumuksellisia syntaktisia ja semanttisia representaatioita kahta tapahtumaa ympäröivistä konteksteista ja ennustavat niiden välisen ajallisen suhteen. Ehdotetun lähestymistavan arviointi TimeBank-korpuksella osoittaa, että sekventiaalinen mallintaminen pystyy tunnistamaan tarkasti tapahtumien väliset ajalliset suhteet, mikä päihittää aiempia ominaisuuksiin perustuvia malleja jäljittelevän neuroverkkomallin, joka käyttää syötteenä erilaisia erillisiä piirteitä.

**Tulos**

Sekventiaalinen malli lauseen sisäisten tapahtumien välisten ajallisten suhteiden luokittelemiseksi.

**Esimerkki 2.1282**

Puheentunnistus etäältä on haasteellista erityisesti siksi, että puhujan ja mikrofonin välisten suurten etäisyyksien aiheuttama jälkikaiunta turmelee puhesignaalit. Jotta selviydyttäisiin monenlaisesta kaikusta reaalimaailman tilanteissa, esitämme uusia lähestymistapoja akustiseen mallintamiseen, mukaan lukien syvien neuroverkkojen (DNN) kokonaisuus ja yhdessä koulutettujen DNN:ien kokonaisuus. Ensin perustetaan useita DNN:iä, joista kukin vastaa eri jälkikaiunta-aikaa 60 (RT60) asetusvaiheessa. Lisäksi jokainen DNN-akustisten mallien kokonaisuuden malli koulutetaan edelleen yhdessä, mukaan lukien sekä ominaisuuksien kartoitus että akustinen mallinnus, jossa ominaisuuksien kartoitus on suunniteltu kaikuisuuden poistoa varten etupäässä. Testausvaiheessa kaksi todennäköisintä DNN-mallia valitaan DNN-kokonaisuudesta käyttämällä MAP-todennäköisyyksiä (maximum a posteriori), jotka lasketaan online-muodossa käyttämällä ML-pohjaista sokeaa RT60-estimointia, ja sitten kahden DNN-mallin jälkitodennäköisyystulokset yhdistetään ML-pohjaisten painojen avulla yksinkertaisena keskiarvona. Laajat kokeet osoittavat, että ehdotettu lähestymistapa parantaa puheentunnistustarkkuutta huomattavasti perinteisiin DNN-perusjärjestelmiin verrattuna erilaisissa kaiuntaolosuhteissa.

**Tulos**

Yhdessä koulutettujen syvien neuroverkkopohjaisten akustisten mallien muodostama kokonaisuus kaikupuheen tunnistamiseen.

**Esimerkki 2.1283**

Suljettujen usein toistuvien mallien löytäminen on yksi tiedonlouhinnan perusongelmista. Viimeaikaiset rajoitusohjelmointi (CP) -lähestymistavat deklaratiiviseen joukkojen louhintaan ovat osoittaneet hyödyllisyytensä ja joustavuutensa. Nykyisissä CP-lähestymistavoissa käytettyjen rajoitusten laaja käyttö aiheuttaa kuitenkin monia vaikeuksia, kun halutaan selviytyä suuriulotteisista tietokokonaisuuksista. Tässä artikkelissa ehdotetaan CLOSEDPATTERN-rajoitteita, jotka eivät vaadi mitään uudelleen määriteltyjä rajoitteita eivätkä ylimääräisiä muuttujia koodatakseen tehokkaasti CFPM-rajoitteita (Closed Frequent Pattern Mining). CLOSEDPATTERN kattaa CFPM-ongelman erityisen semantiikan, jotta voidaan varmistaa polynominen karsinta-algoritmi, joka takaa toimialueen johdonmukaisuuden. Rajoitteemme laskennallisia ominaisuuksia analysoidaan ja niiden käytännön tehokkuutta arvioidaan kokeellisesti.

**Tulos**

Maailmanlaajuinen rajoitus suljettujen kohdejoukkojen louhintaa varten

**Esimerkki 2.1284**

Tässä artikkelissa ehdotamme kontekstitietoista avainsanojen havaitsemismallia, jossa käytetään merkkitason toistuvaa neuroverkkoa (RNN) puhuttujen termien havaitsemiseen jatkuvassa puheessa. RNN on koulutettu alusta loppuun konnektionistisella temporaalisella luokittelulla (CTC) merkkien ja sanarajojen merkintöjen todennäköisyyksien tuottamiseksi. Koulutuksessa ja testauksessa ei tarvita foneettista transkriptiota, senone-mallinnusta tai järjestelmän sanakirjaa. Avainsanoja voidaan myös helposti lisätä ja muuttaa muokkaamalla tekstipohjaista avainsanaluetteloa ilman RNN:n uudelleenkoulutusta. Lisäksi yksisuuntainen RNN käsittelee äärettömän pitkiä äänivirtoja ilman esisegmentointia, ja avainsanat havaitaan pienellä latenssilla ennen kuin lausuma on valmis. Kokeelliset tulokset osoittavat, että ehdotettu avainsanojen havaitsija on huomattavasti parempi kuin syvään neuroverkkoon (DNN) ja piilotettuun Markov-malliin (HMM) perustuva avainsanojen täyttäjämalli jopa vähemmillä laskutoimituksilla.

**Tulos**

Avainsanojen online-etsintä merkkitason rekursiivisen neuroverkon avulla

**Esimerkki 2.1285**

Monissa sovelluksissa, jotka ulottuvat anturiverkoista sosiaalisiin verkostoihin, liikennejärjestelmiin, geenisäätelyverkkoihin tai big dataan, kiinnostavat signaalit määritellään graafin kärkipisteissä. Tämän artikkelin tavoitteena on ehdottaa pienimmän neliösumman (LMS) strategiaa graafien yli määriteltyjen signaalien adaptiivista estimointia varten. Jos oletetaan, että graafisignaali on kaistarajoitteinen tunnetulla kaistanleveydellä, menetelmä mahdollistaa rekonstruktion, jonka suorituskyky on taattu keskineliövirheen osalta, ja seurannan rajoitetusta määrästä havaintoja osajoukon kärkipisteistä. Yksityiskohtainen keskiarvon neliöanalyysi osoittaa ehdotetun menetelmän suorituskyvyn ja johtaa useisiin oivalluksiin hyödyllisten näytteenottostrategioiden suunnittelemiseksi graafisignaaleille. Numeeriset tulokset vahvistavat teoreettiset havaintomme ja havainnollistavat ehdotetun menetelmän suorituskykyä. Lisäksi ehdotamme menetelmää, jossa kaistanleveys ei ole etukäteen tiedossa ja jossa tehdään harvennettu online-estimointi signaalin tuesta (graafin) taajuusalueella, mikä mahdollistaa graafin näytteenottostrategian online-sopeuttamisen. Lopuksi sovellamme ehdotettua menetelmää kognitiivisen verkkoympäristön tietyn toiminta-alueen tehon tilatiheyskartografian rakentamiseen.

**Tulos**

Graafisignaalien pienimmän keskiarvon neliöiden estimoinnit (Least Mean Squares Estimation of Graph Signals)

**Esimerkki 2.1286**

Tarkastelemme Bayesin likimääräisen päättelyn ongelmaa log-supermodulaarisissa malleissa. Nämä mallit kattavat tavalliset pareittaiset MRF-mallit, joissa on binäärimuuttujia, mutta ne mahdollistavat korkeamman asteen vuorovaikutuksen, joka on hankalaa nykyisille likimääräisille päättelytekniikoille, kuten uskomusten etenemiselle, keskiarvokentälle ja muunnoksille. Osoitamme, että hiljattain ehdotettu muunnoslähestymistapa log-supermodulaaristen mallien päättelyyn -L-FIELD- pelkistyy laajasti tutkittuun miniminormiongelmaan submodulaarista minimointia varten. Tämä oivallus mahdollistaa tehokkaiden olemassa olevien työkalujen hyödyntämisen ja siten variaatioproblematiikan ratkaisemisen suuruusluokkia tehokkaammin kuin aiemmin on ollut mahdollista. Tämän jälkeen esitämme toisen luonnollisen tulkinnan L-FIELDille ja osoitamme, että se minimoi täsmälleen tietynlaisen Rényin divergenssimäärän. Tämä oivallus valaisee L-FIELDin tuottamien muunnosapproksimaatioiden luonnetta. Lisäksi osoitamme, miten rinnakkainen päättely voidaan suorittaa viestien välittämisenä sopivassa faktorigraafissa lineaarisella konvergenssivauhdilla ilman, että kaikkien faktorin konfiguraatioiden yli tarvitsee laskea yhteen. Lopuksi sovellamme lähestymistapaamme haastavaan kuvan segmentointitehtävään. Kokeemme vahvistavat lähestymistapamme skaalautuvuuden, marginaalien korkean laadun ja korkeamman kertaluvun potentiaalien sisällyttämisestä saatavan hyödyn.

**Tulos**

Skaalautuva variatiivinen päättely log-supermodulaarisissa malleissa

**Esimerkki 2.1287**

Kehitämme uusia ensimmäisen ja toisen asteen piirteitä riippuvuuksien jäsennykseen Googlen Syntactic Ngrams -korpuksen pohjalta, joka on kokoelma skannattujen kirjojen jäsenneltyjen lauseiden alipuiden lukumääriä. Laajennamme myös aiempaa työtä n-grammiominaisuuksien kehittämiseksi Web1T-korpuksesta Google Books -korpukseen ja ensimmäisen asteen ominaisuuksista toisen asteen ominaisuuksiin vertailemalla ja analysoimalla suorituskykyä uutislehti- ja verkkopankkien avulla. Sekä pinta- että syntaktiset n-grammit tuottavat huomattavia ja toisiaan täydentäviä parannuksia jäsennystarkkuuteen eri aloilla. Our best system combines the two feature sets, achieving up to 0.8% absolute UAS improvements on newswire and 1.4% on web text.

**Tulos**

Verkkomittakaavan pinta- ja syntaktiset n-grammiominaisuudet riippuvuuksien jäsennystä varten

**Esimerkki 2.1288**

Järjestäytyneen tietoverkkorikollisuuden tapaukset lisääntyvät, koska rikolliset saavat suuria taloudellisia palkkioita, kun taas rikosten tekeminen maksaa vähän. Kun digitaalinen toimintaympäristö laajenee yhä useampien internet-yhteensopivien laitteiden ja tekniikoiden, kuten sosiaalisen median, myötä, yhä useammat tietoverkkorikolliset, joiden äidinkieli ei ole englanti, tunkeutuvat verkkoavaruuteen saadakseen rahaa nopeista hyväksikäytöistä. Tässä artikkelissa arvioidaan kolmen koneoppimisen luokittelijan suorituskykyä 419-huijausten havaitsemisessa kaksikielisessä nigerialaisessa verkkorikollisyhteisössä. Käytämme kolmea tekstinkäsittelyssä suosittua luokittelijaa, jotka ovat seuraavat: Naïve Bayes, k-nearest neighbors (IBK) ja Support Vector Machines (SVM). Alustavat tulokset reaalimaailman tietokokonaisuudesta osoittavat, että SVM päihittää Naïve Bayesin ja IBK:n huomattavasti 95 prosentin luotettavuustasolla. Avainsanat: koneoppiminen; kaksikieliset verkkorikolliset; 419-huijaukset;

**Tulos**

Luokittelijoiden arviointi 419-huijausten havaitsemisessa kaksikielisissä verkkorikollisyhteisöissä.

**Esimerkki 2.1289**

Tässä artikkelissa esitellään vahvistusoppimisen automaattinen taitojen hankintakehys, johon kuuluu tietyn tehtävän hierarkkisen kuvauksen tunnistaminen abstraktien tilojen ja abstraktien tilojen välisten laajennettujen toimintojen avulla. Tällaisten tehtävässä esiintyvien rakenteiden tunnistaminen tarjoaa keinoja yksinkertaistaa ja nopeuttaa vahvistusoppimisen oppimisalgoritmeja. Nämä rakenteet auttavat myös yleistämään tällaisia algoritmeja useisiin tehtäviin ilman, että toimintatapoja tarvitsee opetella uudelleen tyhjästä. Käytämme dynaamisten järjestelmien ideoita löytääksemme metastabiileja alueita tilaavaruudesta ja liittääksemme ne abstrakteihin tiloihin. Spectral clustering algorithm PCCA+ -algoritmia käytetään tunnistamaan sopivia abstraktioita, jotka on kohdistettu taustalla olevaan rakenteeseen. Taidot määritellään tällaisten abstraktien tilojen välisten siirtymien avulla. PCCA+:n liitettävyystietoja käytetään näiden taitojen tai vaihtoehtojen luomiseen. Taidot ovat oppimistehtävästä riippumattomia, ja niitä voidaan käyttää tehokkaasti uudelleen useissa eri tehtävissä, jotka on määritelty yhteisen tila-avaruuden kautta. Lähestymistavan toinen merkittävä etu on, että se ei tarvitse MDP:n ennakkomallia ja toimii hyvin myös silloin, kun MDP:t on muodostettu näytteistetyistä liikeradoista. Lopuksi esittelemme yrityksemme laajentaa automaattista taitojen hankkimista koskevaa kehystä monimutkaisiin tehtäviin, kuten videopelien pelaamisen oppimiseen, jossa käytämme syväoppimistekniikoita representaatio-oppimiseen spatio-temporaalisen abstraktiokehyksemme avuksi. Proceedings of the 33 rd International Conference on Machine Learning, New York, NY, USA, 2016. JMLR: W&CP nide 48. Tekijän tai kirjoittajien tekijänoikeudet 2016. 1. Motivaatio ja johdanto Hierarkkisen vahvistusoppimisen ydinajatus on hajottaa vahvistusoppimisongelma osatehtäviin abstraktiohierarkian avulla. Tyypillisesti täydellisessä vahvistusoppimisongelmassa agentin oletetaan olevan Markovin päätösprosessin yhdessä tilassa jokaisella aika-askeleella. Tämän jälkeen agentti suorittaa yhden monista mahdollisista alkeistoimista. Agentin tilasta hetkellä t ja sen toimesta, jonka se tekee kyseisestä tilasta, määritetään agentin tila hetkellä t + 1. Tämän perusteella määritetään agentin tila hetkellä t+1. Suurissa ongelmissa tämä voi kuitenkin johtaa liian suureen rakeisuuteen: kun agentti joutuu päättämään jokaisesta primitiivisestä toimesta jokaisessa rakeisessa tilassa, se voi usein menettää kokonaiskuvan näkyvistä. Jos toimintasarja voidaan kuitenkin abstrahoida abstraktiksi toiminnaksi, agentti voi vain muistaa toimintasarjan, joka oli hyödyllinen sen saamiseksi ajallisesti kaukaiseen hyödylliseen tilaan alkutilasta. Tätä kutsutaan tyypillisesti vaihtoehdoksi tai taidoksi vahvistusoppimisen kirjallisuudessa. Hyvä analogia on ihminen, joka suunnittelee liikkumistaan kulkiakseen nykyisestä paikasta A määränpäähän B. Suunnitellessamme A:sta B:hen tunnistamme välitavoitteet Ci, jotka johtavat meidät A:sta B:hen, sen sijaan, että murehtisimme välittömän liikkumisen täsmällisiä mekanismeja A:ssa, jotka abstrahoidaan. Vaihtoehdot ovat kätevä tapa virallistaa tämä abstraktio. Vahvistusoppimisen yleisen filosofian mukaisesti haluamme rakentaa agentteja, jotka voivat automaattisesti löytää vaihtoehtoja ilman ennakkotietoa, pelkästään tutkimalla ympäristöä. Näin ollen lähestymistapamme kuuluu laajaan taitojen automaattisen löytämisen kategoriaan. Tehtävien rakenteen hyödyntämiseksi hierarkkinen hajottaminen ottaa käyttöön malleja, jotka on määritelty erillisillä toimintaperiaatteilla (joita kutsutaan myös ajallisesti laajennetuiksi toiminnoiksi, vaihtoehdoiksi tai taidoiksi), joiden suorittaminen voi kestää useita aika-askeleita. Taidot voivat hyödyntää rakenteen esittämistä esittämällä aliohjelmia, jotka suoritetaan useita kertoja tehtävän suorittamisen aikana. Tällaisia yhdessä tehtävässä opittuja taitoja voidaan käyttää uudelleen ar X iv :1 60 5. 05 35 9v 1 [ cs .L G ] 1 7 M ay 2 01 6 Submission and Formatting Instructions for ICML 2016 toisessa tehtävässä, kunhan se edellyttää saman aliohjelman suorittamista. Vaihtoehdot myös tehostavat etsintää tarjoamalla päätöksentekijälle korkean tason käyttäytymisen, jolla hän voi katsoa eteenpäin vastaavan aliohjelman suorittamista. Taitojen tai vaihtoehtojen automaattinen löytäminen on ollut aktiivinen tutkimusalue, ja siihen on ehdotettu useita lähestymistapoja. Nykyiset menetelmät voidaan karkeasti luokitella näytteenottoreittiin perustuviin ja osioihin perustuviin menetelmiin. Joitakin niistä ovat mm: - Tila-avaruuden pullonkaulojen tunnistaminen, jossa tila-aukio on jaettu joukkoihin ja kahden harvinaisen tilajoukon väliset siirtymät voidaan nähdä pullonkaulajoukkojen käyttöönottona tällaisten harvinaisten siirtymien kohdissa. Tällaisten tilojen saavuttamiseen tähtäävät toimintatavat tallennetaan välimuistiin vaihtoehtoina (McGovern & Barto, 2001). - Faktoroidussa tilakuvauksessa olevan rakenteen käyttäminen sellaisten toimintasekvenssien tunnistamiseen, jotka aiheuttavat muutoin harvinaisia muutoksia tilamuuttujissa: nämä sekvenssit tallennetaan välimuistiin vaihtoehtoina (Hengst, 2004). - Saadaan graafinen esitys agentin vuorovaikutuksesta ympäristönsä kanssa ja käytetään betweenness centrality -mittoja osatehtävien tunnistamiseen (Simsek & Barto, 2008). - Käyttämällä klusterointimenetelmiä (spektraalisia tai muita) Markovin päätösprosessin (MDP) eri vahvasti toisiinsa liittyvien komponenttien erottamiseksi toisistaan ja eri klustereita yhdistävien pääsytilojen tunnistamiseksi (Menache et al., 2002). Vaikka näillä menetelmillä on ollut vaihtelevaa menestystä, niissä on tiettyjä puutteita. Pullonkauloihin perustuvilla lähestymistavoilla ei ole luonnollista tapaa tunnistaa sitä tilaavaruuden osaa, jossa vaihtoehdot ovat sovellettavissa ilman ulkoista tietoa ongelma-alueesta. Spektraaliset menetelmät tarvitsevat jonkinlaista regularisointia estääkseen epätasa-arvoiset jaottelut, jotka saattavat johtaa tilaavaruuden mielivaltaiseen jakamiseen. Esittelemme kehyksen, joka havaitsee trajektorien perusteella estimoidusta MDP-mallista hyvin yhteenliitetyt tai metavakaat tilaavaruuden alueet. Käytämme PCCA+ -algoritmia, joka on konformidynamiikasta peräisin oleva spektrinen klusterointialgoritmi (Weber et al., 2004), joka paitsi jakaa MDP:n osiin myös palauttaa alueiden välisen kytkentätiedon. Sitten ehdotamme erittäin tehokasta tapaa koostaa vaihtoehtoja samaa kehystä käyttäen, jotta pääsemme metastabiililta alueelta toiselle, jolloin saamme politiikan ilmaiseksi. Kun meillä on nämä vaihtoehdot, voimme käyttää tavanomaisia vahvistusoppimisalgoritmeja oppiaksemme politiikan osatehtävien yli tietyn tehtävän ratkaisemiseksi. Näytämme erityisesti tuloksia käyttämällä SMDP Qlearning -oppimista 2-huoneen alueella. Kun yritämme laajentaa sitä korkeampiulotteisempiin tila-avaruustehtäviin, kuten Atari 2600 -videopeleihin, liitämme opitut vaihtoehdot alkeistoimintojen joukkoon käyttämällä Intra-Option Value -oppimista ja opettelemme siten toimintaperiaatteen, jolla ratkaistaan annettu tehtävä. Lähestymistavan suurimpana etuna on, että saamme vaihtoehtojen toimintaperiaatteet ilmaiseksi samalla kun teemme osiointia hyödyntämällä PCCA+:n palauttamia jäsenyysfunktioita. Lähestymistapamme pystyy oppimaan kohtuullisen hyviä taitoja jopa rajoitetulla otannalla, mikä tekee siitä hyödyllisen tilanteissa, joissa tutkimista rajoittavat ympäristökustannukset. Se tarjoaa myös tavan tarkentaa abstraktioita online-muodossa ilman, että koko MDP:tä rekonstruoidaan eksplisiittisesti. Vielä tärkeämpää on, että laajennamme sen koskemaan tapauksia, joissa tilaavaruus on niin suuri, että tarkka mallintaminen ei ole mahdollista. Tässä tapauksessa otamme mallioppimismallin oppimiseen (Oh et al., 2015) inspiraation viimeaikaisesta työstä, joka koskee etukäteisennusteita, ja käytämme syviä konvoluutio-neuraaliverkkoja (Deep Convolutional Neural Networks, CNN) ja pitkää lyhytaikaista muistia (Long Short Term Memory, LSTM) oppiaksemme tila-avaruuden spatio-temporaalisia representaatioita. Opittua representaatiota käyttäen suoritamme tilojen aggregoinnin klusterointitekniikoiden avulla ja estimoimme siirtymämallimme abstraktille avaruudelle näiden aggregoitujen tilojen perusteella. Seuraavassa luetellaan tämän lähestymistavan edut: - Taidot hankitaan verkossa, näytteeksi otetuista liikeradoista sen sijaan, että vaadittaisiin ennakkomalli MDP:stä. - Pullonkaulatilojen etsimisen sijasta etsitään hyvin toisiinsa yhteydessä olevia alueita, ja näin ollen löydetyt vaihtoehdot vastaavat paremmin tilaavaruuden rakennetta. - Lähestymistapa palauttaa metastabiilien alueiden välistä liitettävyystietoa, jota voidaan käyttää tilaavaruuden abstraktin graafin rakentamiseen, jolloin alueellinen ja ajallinen tieto yhdistyvät mielekkäästi. - Klusterointialgoritmi antaa jokaiselle tilalle sumean jäsenyyden kuulumisesta tiettyyn metastabiiliin alueeseen, mikä tarjoaa tehokkaan tavan koota vaihtoehtoja luonnollisesti. Järjestämme loput artikkelista seuraavasti: Ensin selvitämme ehdottamamme vaihtoehtojen luomisjärjestelmän ja syvennymme spektrisen klusterointialgoritmin PCCA+ tärkeisiin näkökohtiin sekä siihen, miten PCCA+:aa voidaan käyttää vaihtoehtojen luomisessa. Esitämme tuloksia 2-huoneen alueelta tämän kehyksen osalta. Artikkelin seuraavassa osassa keskitytään yritykseemme laajentaa tätä kehystä monimutkaisempiin tehtäviin, kuten Seaquestin kaltaisten videopelien pelaamiseen Atari 2600 -alueella vaihtoehtojen avulla. Selitämme syvien verkkojen motivaation ja käytön, tila-aggregaatioon käytetyt klusterointialgoritmit, minkä jälkeen esitämme käytetyn mallin ja ICML 2016:n alustavien tulosten toimittamis- ja muotoiluohjeet. Päätämme sitten keskusteluun tämän lähestymistavan haasteista ja myös vertailuun muiden viimeaikaisten hierarkkisen vahvistusoppimisen yritysten kanssa korkeamman ulottuvuuden tehtävissä. 2. Vaihtoehtojen tuottamisen kehys Jaamme tämän osan kahteen osaan: Ensin selitämme spektraalisen klusterointialgoritmin PCCA+ ja motivoimme sen käyttöä spatiaaliseen abstraktioon. Toisessa osassa käsitellään vaihtoehtojen tuottamista PCCA+:n avulla. 2.1. Spatiaalinen abstraktio PCCA+:n avulla Kun MDP:tä edustavan graafin algebrallinen esitys on annettu, haluamme löytää sopivia abstraktioita, jotka ovat linjassa taustalla olevan rakenteen kanssa. Käytämme tähän spektraalista klusterointialgoritmia. Keskeistä spektraalisen klusteroinnin ideassa on graafin Laplacian, joka saadaan samankaltaisuusgraafista. Graafien topologisten ominaisuuksien ja graafin Laplacian-matriisien välillä on monia kiinteitä yhteyksiä, joita spektraaliset klusterointimenetelmät hyödyntävät datan jakamiseksi klustereihin. Vaikka Laplacianin spektri säilyttääkin graafin rakenteelliset ominaisuudet, datan klusterointi Laplacianin ominaistilassa ei kuitenkaan takaa tätä. Esimerkiksi k-means-klusterointi (Ng et al., 2001) Laplacianin ominaistilassa toimii vain, jos klusterit sijaitsevat taustalla olevan ominaistilan epäyhtenäisissä koverissa joukoissa. Aineiston jakaminen klustereihin projisoimalla se suurimpiin k-eigenvektoreihin (Meila & Shi, 2001) ei säilytä aineiston topologisia ominaisuuksia Laplacianin ominaistilassa. Ehdotettu kehys vaatii spatiaalisen abstrahoinnin tehtävään klusterointimenetelmää, jossa hyödynnetään rakenteellisia ominaisuuksia sekä kohteiden konfiguraatioavaruudessa että spektraalisessa aliavaruudessa, aivan toisin kuin aiemmissa menetelmissä. Siksi otamme inspiraatiota konformidynamiikan kirjallisuudesta, jossa (Weber et al., 2004) tehdään samanlainen analyysi dynaamisen järjestelmän konformitilojen havaitsemiseksi. He ehdottavat spektraalista klusterointialgoritmia PCCA+, joka perustuu Perronin klusterianalyysin periaatteisiin systeemin siirtymärakenteesta. Laajennamme heidän analyysiaan havaitsemaan spatiaalisia abstraktioita autonomisesti ohjatuissa dynaamisissa järjestelmissä. Tässä lähestymistavassa rakennetaan Laplacian L:n spektri (johdettu vierekkäisyysmatriisista S) ja etsitään spektrin paras muunnos siten, että muunnettu perusta kohdistuu datapisteiden klustereihin omavaruudessa. (Weber et al., 2004) kuvattua projektiomenetelmää käytetään löytämään kunkin tilan kuuluminen muunnetun perustan päällä sijaitsevien erityisten pisteiden joukkoon, jotka tunnistetaan R-alavaruuden simpleksin kärkipisteiksi (Spectral Gap -menetelmää käytetään klusterien määrän k arvioimiseen). Ensimmäisen kertaluvun häiriön osalta simpleksi on vain lineaarinen muunnos origon ympärillä, ja simpleksin kärkipisteiden löytämiseksi on löydettävä k pistettä, jotka muodostavat kuperan rungon siten, että kaikkien pisteiden poikkeama tästä rungosta on minimoitu. Tämä saavutetaan etsimällä datapiste, joka on kauimpana origosta, ja tunnistamalla iteratiivisesti datapisteet, jotka ovat kauimpana nykyiseen kärkijoukkoon sopivasta hypertasosta. Kuva 1. Simplex Ensimmäisen ja ylemmän asteen häiriöalgoritmi 1 PCCA+ 1: Rakennetaan Laplacian L 2: Lasketaan n (kärkipisteiden lukumäärä) L:n ominaisarvoa alenevassa järjestyksessä 3: Valitaan ensimmäiset k ominaisarvoa, joiden ek-ek+1 1-ek+1 > tc (spektrinen aukkokynnys). 4: Lasketaan vastaavien ominaisarvojen (e1, e2, - - - , ek) ominaisvektorit ja pinotaan ne sarakevektoreiksi ominaisvektorimatriisiin Y . 5: Merkitään Y:n rivejä seuraavasti: Y(1),Y(2), - - - - ,Y(N) ∈ R 6: Määritellään π(1) indeksiksi, jolle ||Y (π(1))||2 on maksimi. Määritellään γ1 = span{Y (π(1))} 7: For i = 2, - - - - , k: Määritellään πi indeksiksi, jonka etäisyys hypertasoon γi-1, eli ||Y (πi) - γi-1||2 on suurin. Määritellään γi = span{Y (π1), - - - , Y (πi)}. ||Y (πi) - γi-1||2 = ||Y (πi)- γ i-1((γi-1γ i-1)γi-1)Y (πi) )|| PCCA+-algoritmi palauttaa jäsenyysfunktion χ, joka määrittelee kunkin tilan s jäsenyysasteen abstraktiin tilaan Sj . Kahden abstraktin tilan (Si, Sj) välinen liitettävyystieto annetaan χLχ:n (i, j) merkinnällä, kun taas diagonaaliset merkinnät antavat suhteellisen liitettävyystiedon klusterin sisällä. Liitettävyystietoa hyödynnetään abstraktien tilojen välisten päätöskäytäntöjen oppimiseen, jota kuvataan seuraavassa jaksossa. Ominaisarvojakauman jyrkkien piikkien esiintyminen (osoittaa hyvää klusterointia) palauttaa tietoa tilojen klusteroinnin hyvyydestä. ICML 2016 -julkaisun toimittamista ja muotoilua koskevat ohjeet 2.2. ICML 2016. Vaihtoehtojen tuottaminen PCCA+:sta

**Tulos**

Hierarkkinen vahvistusoppiminen käyttäen spatio-ajallisia abstraktioita ja syviä neuroverkkoja.

**Esimerkki 2.1290**

Viimeaikaiset tutkimukset ovat osoittaneet, että GPS-(paikannus)tiedoista voidaan oppia yllättävän monipuolisia malleja ihmisen toiminnasta. Suurin osa tähänastisista ponnisteluista on kuitenkin keskittynyt yksittäisten yksilöiden tai ihmisryhmien tilastollisten ominaisuuksien mallintamiseen. Lisäksi aiemmassa työssä on keskitytty ainoastaan mallintamaan kiinnostavien toimintojen todellisia onnistuneita suorituksia (eikä epäonnistuneita tai yritettyjä suorituksia). Me sitä vastoin otamme tehtäväksi ymmärtää ihmisten vuorovaikutusta, vuorovaikutusyrityksiä ja aikomuksia meluisista anturitiedoista täysin relationaalisessa moniagenttiympäristössä. Käytämme reaalimaailman "kaappaa lippu" -peliä havainnollistamaan lähestymistapaamme hyvin määritellyllä alalla, johon sisältyy monia erilaisia yhteistoiminnallisia ja kilpailullisia yhteisiä toimintoja. Mallinnamme alaa käyttämällä Markovin logiikkaa, tilastollis-relationaalista kieltä, ja opimme teorian, joka yhdessä denoisesaa datan ja päättelee korkean tason toimintojen esiintymiset, kuten pelaajan vangitsemisen vihollisen. Yhtenäisessä mallissamme yhdistetään todennäköisyysperusteisesti ja loogisesti järkevällä tavalla pelialueen geometrian, pelaajien liikemallin sekä pelin sääntöjen ja dynamiikan asettamat rajoitukset. Osoitamme, että vaikka moniagenttitoimintaa voi olla mahdotonta havaita suoraan anturien kohinan tai toimintahäiriön vuoksi, toiminnan tapahtuminen voidaan silti päätellä ottamalla huomioon sekä sen vaikutus siihen osallistuvien henkilöiden tulevaan käyttäytymiseen että tapahtumat, jotka ovat voineet edeltää sitä. Lisäksi osoitamme, että kun annamme mallin onnistuneesti suoritetuista moniagenttitoiminnoista sekä joukon esimerkkejä epäonnistuneista yrityksistä samoissa toiminnoissa, järjestelmämme oppii automaattisesti laajennetun mallin, joka pystyy tunnistamaan onnistumisen ja epäonnistumisen sekä ihmisten toiminnan tavoitteet suurella tarkkuudella. Vertaamme lähestymistapaamme muihin vaihtoehtoihin ja osoitamme, että yhtenäistetty mallimme, jossa otetaan huomioon paitsi yksittäisten pelaajien väliset suhteet myös toimintojen väliset suhteet koko pelin keston ajan, on huomattavasti tarkempi, vaikka se onkin laskennallisesti kalliimpi. Lopuksi osoitamme, että epäonnistuneiden yritysten eksplisiittinen mallintaminen parantaa suorituskykyä muissa tärkeissä tunnistustehtävissä.

**Tulos**

Sijaintiin perustuva päättely monimutkaisesta moniagenttikäyttäytymisestä

**Esimerkki 2.1291**

Kuvaamme tekniikan painotettujen puuautomaattien (WTA) minimoimiseksi, joka on tehokas formalismi, joka kattaa todennäköisyyspohjaiset kontekstittomat kieliopit (PCFG) ja latenttimuuttujan PCFG:t. Menetelmämme perustuu WTA:n määrittelemän Hankel-matriisin singulaariarvopurkuun. Tärkein teoreettinen tuloksemme on tehokas algoritmi äärettömän Hankel-matriisin SVD:n laskemiseksi, joka on esitetty implisiittisesti WTA:na. Analysoimme minimoinnin aiheuttamaa approksimaatiovirhettä ja arvioimme menetelmäämme reaalimaailman datalla, joka on peräisin uutislehtipankista. Osoitamme, että malli saavuttaa pienemmän perpleksiteetin kuin aiemmat PCFG:n minimointimenetelmät, ja se on myös paljon vakaampi, koska paikallisia optimeja ei esiinny.

**Tulos**

Painotettujen puuautomaattien approksimointi singulaariarvojen typistämisen avulla

**Esimerkki 2.1292**

Tutkimme parhaan käsivarren tunnistamisongelmaa lineaarisessa banditissa, jossa käsivarsien palkkiot riippuvat lineaarisesti tuntemattomasta parametrista θ ja tavoitteena on palauttaa käsivarsi, jolla on suurin palkkio. Luonnehdimme ongelman monimutkaisuutta ja esittelemme näytteenjakostrategioita, jotka vetävät käsivarret tunnistamaan parhaan käsivarren kiinteällä varmuudella minimoiden samalla näytteenottobudjetin. Erityisesti osoitamme, että on tärkeää hyödyntää globaalia lineaarista rakennetta, jotta voidaan parantaa lähes optimaalisten käsien palkkion estimaattia. Analysoimme ehdotettuja strategioita ja vertailemme niiden empiiristä suorituskykyä. Analyysimme sivutuotteena osoitamme yhteyden G-optimaalisuuskriteeriin, jota käytetään optimaalisessa koesuunnittelussa.

**Tulos**

Parhaan käsivarren tunnistaminen lineaarisissa rosvojoukkueissa

**Esimerkki 2.1293**

Useimmissa todistettavasti tehokkaissa vahvistusoppimisalgoritmeissa käytetään optimismia huonosti ymmärrettyjen tilojen ja toimintojen suhteen, mikä kannustaa tutkimiseen. Tutkimme vaihtoehtoista lähestymistapaa tehokkaaseen etsintään: posteriorinen näytteenotto vahvistusoppimisessa (posterior sampling for reinforcement learning, PSRL). Tämä algoritmi etenee toistuvissa jaksoissa, joiden kesto tunnetaan. Kunkin jakson alussa PSRL päivittää Markovin päätösprosessien ennakkojakauman ja ottaa yhden näytteen tästä jälkijakaumasta. Sen jälkeen PSRL noudattaa jakson aikana politiikkaa, joka on optimaalinen tälle otokselle. Algoritmi on käsitteellisesti yksinkertainen ja laskennallisesti tehokas, ja sen avulla agentti voi koodata ennakkotietoa luonnollisella tavalla. Määritämme Õ(τS √ AT ) -rajan odotetulle katumukselle, jossa T on aika, τ on jakson pituus ja S ja A ovat tila- ja toiminta-avaruuksien kardinaalisuudet. Tämä raja on yksi ensimmäisistä algoritmeille, jotka eivät perustu optimismiin, ja se on lähellä kaikkien vahvistusoppimisalgoritmien nykytasoa. Osoitamme simuloimalla, että PSRL päihittää merkittävästi olemassa olevat algoritmit, joilla on samanlaiset katumuksen rajat.

**Tulos**

(Lisää) Tehokas vahvistusoppiminen posteriorisen näytteenoton avulla

**Esimerkki 2.1294**

Tässä artikkelissa esitellään tietoteoreettinen lähestymistapa älykkyyden käsitteeseen laskennallisessa mielessä. Esitämme todennäköisyyspohjaisen kehyksen, jonka avulla osoitetaan, että laskennallinen älykkyys on entropiaa minimoiva prosessi paikallisella tasolla. Tätä uutta järjestelmää käyttäen kehitämme yksinkertaisen dataan perustuvan klusterointiesimerkin ja keskustelemme sen sovelluksista.

**Tulos**

Älykkyyden laskennallinen teoria: Informaatioentropia

**Esimerkki 2.1295**

Tässä asiakirjassa esitellään hybridijärjestelmään perustuva malli, jolla simuloidaan numeerisesti lentokoneen nousuvaihetta. Tätä mallia käytetään sitten lentoradan ennustustyökalussa. Lopuksi käytetään Covariance Matrix Adaptation Evolution Strategy (CMA-ES) -optimointialgoritmia viiden valitun parametrin virittämiseen ja siten mallin tarkkuuden parantamiseen. Kun malli liitetään lentoradan ennustustyökaluun, sitä voidaan käyttää ennustevirheen suuruusluokan määrittämiseen ajan kuluessa ja siten lentoradan ennusteen pätevyysalueen määrittämiseen. Ehdotetun mallin ensimmäinen validointikokeilu perustuu lentoonlähdön yhteydessä tehtävän kertaluonteisen lentoradan ennusteen virheisiin ajan suhteen teoreettisen BADA-mallin oletusarvoihin. Tämä kokeilu, jossa oletetaan täydelliset tiedot, osoittaa myös mallin rajat. Toisessa koeosassa esitellään on-line-ratojen ennustaminen, jossa ennustetta päivitetään jatkuvasti ilma-aluksen senhetkisen sijainnin perusteella. Tämä lähestymistapa herättää useita kysymyksiä, joita varten ehdotetaan perusmallin parannuksia, ja tuloksena saatu lentoradan ennustustyökalu osoittaa tilastollisesti merkittävästi tarkempia tuloksia kuin oletusmallin tulokset.

**Tulos**

Verkko-oppiminen maanpinnan lentoratojen ennustamisessa

**Esimerkki 2.1296**

Jos jakaumasta otetaan näytteitä, kuinka monta uutta elementtiä on odotettavissa, jos jatkamme näytteenottoa tästä jakaumasta? Tämä on tärkeä ja aktiivisesti tutkittu ongelma, jolla on monia sovelluksia, jotka ulottuvat näkymättömien lajien arvioinnista genomiikkaan. Yleistämme tämän ekstrapolaation ja siihen liittyvät näkymättömät estimointiongelmat usean populaation tilanteeseen, jossa populaatiolla j on tuntematon jakaumaDj, josta havaitsemme nj näytettä. Johdamme optimaalisen estimaattorin elementtien kokonaismäärälle, jonka odotamme löytyvän uusista näytteistä eri populaatioissa. Yllättäen todistamme, että estimaattorimme tarkkuus on riippumaton populaatioiden lukumäärästä. Kehitämme myös tehokkaan optimointialgoritmin, jolla voidaan ratkaista yleisempi ongelma, joka koskee usean populaation frekvenssijakaumien estimointia. Validoimme menetelmämme ja teoriamme laajoilla kokeilla. Lopuksi osoitamme todellisella ihmisgenomitietokannalla, joka koostuu useista eri syntyperistä, miten näkymättömän estimoinnin lähestymistapamme voi mahdollistaa kohorttisuunnittelun, jolla voidaan löytää kiinnostavia mutaatioita tehokkaammin.

**Tulos**

Näkymättömän arvioiminen useista populaatioista

**Esimerkki 2.1297**

Tieto siitä, missä ihmiset asuvat, on olennainen osa monia päätöksentekoprosesseja, kuten kaupunkikehitystä, tartuntatautien leviämisen estämistä, evakuointisuunnittelua, riskinhallintaa, luonnonsuojelusuunnittelua ja paljon muuta. Vaikka alhaalta ylöspäin suuntautuvilla, kyselytutkimuksiin perustuvilla väestölaskennoilla voidaan saada kattava kuva maan väestömaisemasta, niiden toteuttaminen on kallista, niitä tehdään harvoin ja niistä saadaan vain laajojen alueiden väestömäärät. Väestön erittelytekniikat ja väestöennustemenetelmät korjaavat näitä puutteita, mutta niillä on myös omat puutteensa. Vastataksemme yhdessä kysymyksiin "missä ihmiset asuvat" ja "kuinka monta ihmistä siellä asuu", ehdotamme syväoppimismallia, jolla luodaan korkean resoluution väestöarvioita satelliittikuvista. Koulutamme konvoluutio-neuraaliverkkoja ennustamaan Yhdysvaltojen väestöä 0,01◦ × 0,01◦ resoluutioruudukon avulla 1 vuoden Landsat-kuvista. Validoimme nämä mallit kahdella tavalla: kvantitatiivisesti vertaamalla mallimme piirikuntatasolla aggregoituja ruudukkosoluestimaatteja useisiin Yhdysvaltain väestönlaskennan piirikuntatason väestöennusteisiin ja kvalitatiivisesti tulkitsemalla mallin ennusteita suoraan satelliittikuvien syötteiden perusteella. Huomasimme, että mallimme estimaattien aggregointi antaa vertailukelpoisia tuloksia väestönlaskennan piirikuntatason väestöennusteiden kanssa ja että mallimme ennusteita voidaan tulkita suoraan, mikä antaa sille etuja perinteisiin väestönerittelymenetelmiin verrattuna. Yleisesti ottaen mallimme on esimerkki siitä, miten koneoppimistekniikat voivat olla tehokas väline tiedon poimimiseksi luonnostaan jäsentymättömästä kaukokartoitusaineistosta, jotta voidaan tarjota tehokkaita ratkaisuja sosiaalisiin ongelmiin.

**Tulos**

Syväoppimismenetelmä väestömäärän arvioimiseksi satelliittikuvista

**Esimerkki 2.1298**

<lb>Harkitsemme stokastisia bandit-ongelmia, joissa on jatkuva joukko käsivarsia ja joissa odotettu re-<lb>tulos on jatkuva ja unimodaalinen funktio käsivarsista. Odotetun palkkiofunktion tasaisuudesta ja rakenteesta ei tehdä muita oletuksia. Näille ongelmille ehdotamme<lb>stokastista pentakotomia-algoritmia (Stochastic Pentachotomy, SP) ja johdamme äärellisen ajan ylärajat sen katumukselle ja<lb>optimointivirheelle. Erityisesti osoitamme, että mille tahansa odotetun palkkion funktiolle μ, joka käyttäytyy<lb>μ(x) = μ(x)- C|x- x| paikallisesti sen maksimoitsijan x ympärillä jollakin ξ, C > 0, SP-algoritmi on<lb>järjestysoptimaalinen. Nimittäin sen katumus ja optimointivirhe skaalautuvat O(<lb>√<lb>T log(T )) ja O(<lb>√<lb>log(T )/T ),<lb>vastavasti, kun aikahorisontti T kasvaa suureksi. Nämä skaalaukset saavutetaan ilman tietoa<lb> ξ:stä ja C:stä. Algoritmimme perustuu asymptoottisesti optimaalisiin peräkkäisiin tilastollisiin testeihin, joita käytetään menestyksekkääseen<lb>resessiiviseen rajaamiseen intervalliin, joka sisältää suurella todennäköisyydellä parhaan varren. Tietojemme mukaan SP<lb>algoritmi on ensimmäinen peräkkäisen käsivarren valintasääntö, jolla saavutetaan katumus- ja optimointivirhe<lb>skaala O(<lb>√<lb>).T ) ja O(1/<lb>√<lb>T ), vastaavasti logaritmiseen tekijään asti ei-sileille odotetuille<lb>palkintofunktioille sekä sileille funktioille, joiden sileys on tuntematon.

**Tulos**

Unimodaaliset rosvot ilman tasaisuutta

**Esimerkki 2.1299**

Tässä artikkelissa analysoimme spektrin miehitystä käyttämällä erilaisia koneoppimistekniikoita. Sekä valvottuja tekniikoita (naiivi Bayesin luokittelija (NBC), päätöspuut (DT), tukivektorikone (SVM), lineaarinen regressio (LR)) että valvomatonta algoritmia (piilomarkovin malli (HMM)) tutkitaan parhaan tekniikan löytämiseksi, jolla on korkein luokittelutarkkuus (CA). Valvottuja ja valvomattomia algoritmeja verrataan yksityiskohtaisesti laskenta-ajan ja luokittelutarkkuuden suhteen. Luokiteltua miehitystilaa hyödynnetään myös arvioitaessa toissijaisen käyttäjän käyttökatkon todennäköisyyttä tulevien aikavälien osalta, ja järjestelmän suunnittelijat voivat käyttää sitä taajuuksien jakamisen ja taajuuksien yhteiskäytäntöjen määrittelyyn. Numeeriset tulokset osoittavat, että SVM on paras algoritmi kaikista valvotuista ja valvomattomista luokittelijoista. Tämän perusteella ehdotimme uutta SVM-algoritmia yhdistämällä sen Fire Fly -algoritmiin (FFA), joka osoittautui kaikkia muita algoritmeja paremmaksi. Indeksi Termit Fire fly -algoritmi, piilotettu Markovin malli, spektrin miehitys ja tukivektorikone. 25. maaliskuuta 2015 LUONNOS

**Tulos**

Taajuuksien käyttöasteen analysointi koneoppimisalgoritmien avulla

**Esimerkki 2.1300**

Kuvaamme TweeTIME-ohjelman, joka on ajallinen merkintäohjelma Twitterin aikailmaisujen tunnistamiseen ja normalisointiin. Useimmat aiemmat sosiaalisen median analysointiin liittyvät työt ovat perustuneet ajallisiin resolveriin, jotka on suunniteltu hyvin muokattua tekstiä varten ja jotka kärsivät siksi alojen yhteensopimattomuudesta johtuvasta heikentyneestä suorituskyvystä. Esittelemme minimaalisesti valvotun menetelmän, joka oppii suurista määristä merkitsemätöntä dataa eikä vaadi käsin kehitettyjä sääntöjä tai käsin annotoituja harjoitusjoukkoja. TweeTIME saavuttaa 0,68 F1-pistettä päivämäärämerkintöjen ratkaisutehtävässä, mikä on parempi tulos kuin monilla uusimmilla järjestelmillä.1

**Tulos**

Minimaalisesti valvottu menetelmä Twitterin aikailmaisujen tunnistamiseen ja normalisointiin

**Esimerkki 2.1301**

Esittelemme uuden teoreettisen kehyksen keinotekoisten neuroverkkojen analysointiin ja oppimiseen. Lähestymistapamme oppii samanaikaisesti ja adaptiivisesti sekä verkon rakenteen että sen painot. Menetelmä perustuu vahvoihin datasta riippuvaisiin teoreettisiin oppimistakuisiin ja niihin liittyy vahvoja datasta riippuvia teoreettisia oppimistakuita, joten lopullinen verkkorakenne mukautuu todistettavasti minkä tahansa ongelman monimutkaisuuteen.

**Tulos**

AdaNet: Adaptive Structural Learning of Artificial Neural Networks (Keinotekoisten neuroverkkojen mukautuva rakenteellinen oppiminen)

**Esimerkki 2.1302**

Tässä asiakirjassa esittelemme tutkimustuloksemme, jotka liittyvät tiedon ja kokemusten keräämiseen seniorituristien yhteisöstä. Käyttämällä laadullisen analyysin välineitä ja kirjallisuuskatsausta onnistuimme todentamaan joukon hypoteeseja, jotka liittyvät seniorituristien tuottamaan sisältöön heidän osallistuessaan verkkoyhteisöihin. Tuotimme myös koodikirjan, joka edustaa erilaisia teemoja, joita tällaisissa yhteisöissä voi esiintyä. Tämä koodikirja, joka on johdettu omasta laadullisesta tutkimuksestamme ja kirjallisuuskatsauksesta, toimii pohjana automaattisten tiedonlouhintatyökalujen jatkokehitykselle. Onnistuimme myös havaitsemaan, että iäkkäät aikuiset mainitsevat matkailufoorumeilla useammin kuin muut postaajat, mainitsevat ikänsä keskusteluissa, kertovat useammin kokemuksistaan ja motivaatiostaan matkustaa, mutta he eivät eroa toisistaan kuvaillessaan matkalla kohtaamiaan esteitä.

**Tulos**

Kultaiset vuodet, kultaiset rannat: A Study of Elders in Online Travel Communities (Tutkimus ikääntyneistä online-matkayhteisöissä)

**Esimerkki 2.1303**

Tässä asiakirjassa keskitytään tyylinsiirtoon ei-paralleelisen tekstin perusteella. Tämä on yksi osa laajempaa ongelmaperhettä, johon kuuluvat konekääntäminen, tulkitseminen ja tunteiden muokkaaminen. Keskeinen tekninen haaste on sisällön erottaminen halutuista tekstin ominaisuuksista, kuten tunnetilasta. Hyödynnämme latenttien representaatioiden hienostunutta kohdistamista yksikielisissä tekstikorpuksissa, joilla on erilaisia ominaisuuksia. Muokkaamme tarkoituksellisesti koodattuja esimerkkejä niiden ominaisuuksien mukaan, jolloin jäljitettyjen esimerkkien on vastattava käytettävissä olevia esimerkkejä, joilla on muutetut ominaisuudet populaationa. Osoitamme tämän ristiinsovittamismenetelmän tehokkuuden kolmessa tehtävässä: tunteiden muokkaaminen, sanojen korvaavien salakirjoitusten purkaminen ja sanajärjestyksen palauttaminen.

**Tulos**

Tyylinsiirto ei-rinnakkaisesta tekstistä ristiinkohdistamalla

**Esimerkki 2.1304**

Minimaalisen vertexin peiton (MinVC) ongelma on tunnettu NP-vaikea ongelma. Viime aikoina on oltu hyvin kiinnostuneita tämän ongelman ratkaisemisesta reaalimaailman massiivisissa graafeissa. Paikallinen haku on lupaava lähestymistapa optimaalisten tai lähes optimaalisten ratkaisujen löytämiseksi tällaisissa graafeissa. Tässä artikkelissa ehdotamme paikallishakualgoritmia, joka hyödyntää pelkistyssääntöjä ja tietorakenteita MinVC-ongelman ratkaisemiseksi tällaisissa graafeissa. Kokeelliset tulokset laajalla joukolla todellisia sanamassiivisia graafeja osoittavat, että algoritmimme löytää parempia peitteitä kuin nykyiset MinVC:n paikallishakualgoritmit. Esitämme myös mielenkiintoisia tuloksia joidenkin tunnettujen heuristiikkojen monimutkaisuudesta.

**Tulos**

Pelkistyssääntöjen ja tietorakenteiden hyödyntäminen: Paikallinen haku minimipisteiden peittämiseksi massiivisissa graafeissa

**Esimerkki 2.1305**

Suurena haasteena parafraasitutkimuksessa on rinnakkaisten korpusten puute. Tässä artikkelissa esitellään uusi menetelmä, jolla kerätään laajamittaisia lauseparafraaseja Twitteristä linkittämällä twiittejä jaettujen URL-osoitteiden kautta. Menetelmämme tärkein etu on sen yksinkertaisuus, sillä siinä päästään eroon luokittelijasta tai ihmisestä silmukassa, jota tarvitaan datan valintaan ennen annotointia ja sen jälkeen parafraasien tunnistusalgoritmien soveltamista aiemmissa töissä. Esittelemme tähän mennessä suurimman ihmisen merkitsemän 51 524 lauseparin korpuksen ja ensimmäisen monialaisen vertailuanalyysin automaattiselle parafraasien tunnistamiselle. Lisäksi osoitamme, että yli 30 000 uutta lauseparafraasia voidaan helposti ja jatkuvasti tunnistaa kuukausittain ∼70 prosentin tarkkuudella, ja osoitamme niiden hyödyllisyyden myöhemmissä NLP-tehtävissä fraasiparafraasien louhinnan avulla. Koodimme ja datamme ovat vapaasti saatavilla.1

**Tulos**

Jatkuvasti kasvava lauseenvastikkeiden tietokanta

**Esimerkki 2.1306**

Implisiittisten diskurssisuhteiden tunnistaminen on keskeinen osa automaattista diskurssitason analyysia ja luonnollisen kielen ymmärtämistä. Aiemmissa tutkimuksissa hyödynnetään erottelevia malleja, jotka perustuvat joko tehokkaisiin manuaalisiin piirteisiin tai syviin diskurssin representaatioihin. Tässä artikkelissa tutkitaan sen sijaan generatiivisia malleja ja ehdotetaan varioivaa neuraalista diskurssisuhdetunnistinta. Kutsumme tätä mallia nimellä VIRILE. VIRILE luo suunnatun todennäköisyysmallin, jossa on latentti jatkuva muuttuja, joka generoi sekä diskurssin että diskurssin kahden argumentin välisen suhteen. Tehokasta päättelyä ja oppimista varten otamme käyttöön neuraalisen diskurssisuhdemallin latentin muuttujan posteriorin approksimoimiseksi ja käytämme tätä approksimoitua posterioria optimoimaan uudelleenparametrisoitua variationaalista alarajaa. Näin VIRILE voidaan kouluttaa tavanomaisilla stokastisilla gradienttimenetelmillä. Vertailuaineistolla tehdyt kokeet osoittavat, että VIRILE pystyy saavuttamaan kilpailukykyisiä tuloksia verrattuna uusimpiin perusmenetelmiin.

**Tulos**

Variationaalinen neuraalinen diskurssisuhteiden tunnistin

**Esimerkki 2.1307**

Toiminta-alueilla, joilla toimijoilla voi olla virheellisiä uskomuksia, toimien vaikutuksia koskeva päättely edellyttää uskomusten muuttumista koskevaa päättelyä. Tässä artikkelissa käytämme siirtymäjärjestelmämenetelmää päättelyyn agentin uskomusten kehityksestä, kun toimia suoritetaan. Jotkin toimet saavat agentin tarkistamaan uskomuksiaan, kun taas toiset saavat agentin päivittämään uskomuksiaan, mutta tarkistuksen ja päivityksen välinen vuorovaikutus voi olla epäelementaarinen. Esitämme joukon rationaalisuusominaisuuksia, jotka kuvaavat tarkistuksen ja päivityksen vuorovaikutusta, ja esittelemme uuden luokan uskomusten muutosoperaattoreita, joiden avulla voidaan järkeillä tarkistusten ja päivitysten vuorottelevia sarjoja. Uskomusten muutosoperaattoreitamme voidaan luonnehtia luonnollisen siirto-operaation avulla, joka koskee tulkintojen kokonaisesijärjestyksiä. Vertaamme lähestymistapaamme vastaaviin töihin, jotka käsittelevät toiminnan aiheuttamaa iteroituvaa uskomusmuutosta, ja lopuksi esitämme joitakin tulevaisuuden tutkimussuuntia.

**Tulos**

Toimien ja havaintojen aiheuttama kerrannaismuutos uskomuksessa

**Esimerkki 2.1308**

Tässä artikkelissa esitellään malli sähköpostin tekijyyden tunnistamiseen (EAI) käyttämällä klusteripohjaista luokittelutekniikkaa (CCM). Perinteisesti tyylillisiä piirteitä on käytetty menestyksekkäästi erilaisissa tekijyyden analysointitehtävissä; laajennamme perinteistä piirrejoukkoa sisällyttämällä siihen joitakin mielenkiintoisempia ja tehokkaampia piirteitä sähköpostin tekijyyden tunnistamiseen (esim. sähköpostin viimeisin käytetty välimerkki, kirjoittajan taipumus käyttää isoja alkukirjaimia sähköpostin alussa tai välimerkit tervehdyksen tai jäähyväisten jälkeen). Lisäksi otimme mukaan Info Gain -ominaisuuden valintaan perustuvia sisältöominaisuuksia. On havaittu, että tällaisten piirteiden käyttö tekijyyden tunnistamisprosessissa vaikuttaa myönteisesti tekijyyden tunnistamistehtävän tarkkuuteen. Suoritimme kokeita perustellaksemme väitteemme ja vertasimme tuloksia muihin perusmalleihin. Kokeelliset tulokset osoittavat, että ehdotettu CCM-pohjainen sähköpostin tekijyyden tunnistamismalli yhdessä ehdotetun ominaisuusjoukon kanssa päihittää nykyaikaiset SVM-pohjaiset mallit (support vector machine) sekä Iqbalin ja muiden ehdottamat mallit [1, 2]. Ehdotettu malli saavuttaa Enron-tietokannassa 94 %:n tarkkuuden 10 kirjoittajan osalta, 89 %:n 25 kirjoittajan osalta ja 81 %:n 50 kirjoittajan osalta, kun taas 89,5 %:n tarkkuus on saavutettu kirjoittajien rakentamassa todellisen sähköpostin tietokannassa. Enron-tietokannan tulokset on saavutettu melko suurella kirjoittajamäärällä verrattuna Iqbalin ja muiden ehdottamiin malleihin [1, 2].

**Tulos**

CEAI: CCM-pohjainen sähköpostin tekijyyden tunnistamismalli

**Esimerkki 2.1309**

Tässä asiakirjassa esitellään erilaisten koneoppimisalgoritmien sovelluksia vesiviljelyssä. Koneoppimisalgoritmit oppivat malleja historiatiedoista. Vesiviljelyssä historiatiedot saadaan viljelykäytännöistä, tuotoista ja ympäristötietolähteistä. Näiden eri muuttujien välisiä yhteyksiä voidaan saada soveltamalla koneoppimisalgoritmeja historiatietoihin. Tässä artikkelissa esitellään erilaisten koneoppimisalgoritmien sovelluksia vesiviljelysovelluksissa.

**Tulos**

Koneoppimistekniikoiden soveltaminen vesiviljelyyn

**Esimerkki 2.1310**

Esittelemme decentralized rollout sampling policy iteration (DecRSPI) - uuden algoritmin DEC-POMDP:nä formalisoituja moniagenttipäätösongelmia varten. DecRSPI on suunniteltu parantamaan skaalautuvuutta ja ratkaisemaan ongelmia, joista puuttuu eksplisiittinen malli. Algoritmi käyttää MonteCarlo-menetelmiä tuottaakseen otoksen saavutettavissa olevista uskomustiloista. Sen jälkeen se laskee kullekin uskomustilalle yhteisen politiikan, joka perustuu vieritysarvioihin. Uuden politiikkarepresentaation avulla voimme esittää ratkaisut kompaktisti. Algoritmin keskeisiä etuja ovat sen lineaarinen aikakompleksisuus agenttien lukumäärän suhteen, sen rajattu muistin käyttö ja hyvä ratkaisujen laatu. Algoritmi pystyy ratkaisemaan suurempia ongelmia, jotka ovat vaikeasti ratkaistavissa nykyisille suunnittelualgoritmeille. Kokeelliset tulokset vahvistavat lähestymistavan tehokkuuden ja skaalautuvuuden.

**Tulos**

Rollout Sampling Policy Iteraatio hajautetuille POMDP:ille.

**Esimerkki 2.1311**

Syvät neuroverkot voivat olla törkeän tuhlailevia. Videota käsiteltäessä konvoluutioverkko käyttää kiinteän määrän laskentatehoa jokaista ruutua kohti ottamatta huomioon naapurikuvien samankaltaisuutta. Tämän seurauksena se päätyy tekemään toistuvasti hyvin samanlaisia laskutoimituksia. Tällaisen tuhlauksen lopettamiseksi otamme käyttöön SigmaDelta-verkot. Jokaisen uuden syötteen yhteydessä tämän verkon jokainen kerros lähettää seuraavalle kerrokselle diskretoidun muodon aktivointimuutoksestaan. Näin verkon suorittamien laskutoimitusten määrä skaalautuu pikemminkin syötteen ja kerrosten aktivaatioiden muutosten määrän kuin verkon koon mukaan. Esittelemme optimointimenetelmän, jolla mikä tahansa esivalmennettu syvä verkko voidaan muuntaa optimaalisen tehokkaaksi Sigma-Delta-verkoksi, ja osoitamme, että algoritmimme voi sopivalla laitteistolla ajettuna vähentää videodatan käsittelyn laskentakustannuksia vähintään kertaluokkaa.

**Tulos**

SIGMA-DELTA-KVANTISOIDUT VERKOT

**Esimerkki 2.1312**

Vaikka viime vuosikymmenen aikana on tehty melko paljon töitä tunneanalyysin (SA) ja mielipiteiden louhinnan (OM) järjestelmissä ja näiden järjestelmien suorituskyvyn osalta, mutta se ei edelleenkään ole toivottua suorituskykyä, erityisesti morfologisesti rikkaiden kielten (MRL), kuten arabian, osalta, koska kielten luonne on monimutkainen ja haasteellinen. Yksi näistä haasteista on idiomien tai sananlaskujen havaitseminen kirjoittajan tekstissä tai kommentissa. Idiomi tai sananlasku on puheen muoto tai ilmaisu, joka on omaleimainen. Kieliopillisesti sitä ei voida ymmärtää sen osien yksittäisten merkitysten perusteella, ja se voi tuottaa erilaisia tunteita, kun sitä käsitellään erillisinä sanoina. Näin ollen tässä asiakirjassa esitellään AIPSeLEX, joka on uusi idiomien ja sananlaskujen tunnesanasto nykyaikaiselle arabian standardille (MSA) ja puhekielelle, jotta voitaisiin helpottaa sanojen havaitsemista ja luokittelua automaattisia sanojen ja sanojen automaatiojärjestelmiä varten. AIPSeLEX on kerätty manuaalisesti ja annotoitu lausetasolla semanttisella suuntautumisella (positiivinen tai negatiivinen). Leksikon manuaaliseen rakentamiseen ja annotointiin liittyvät ponnistelut raportoidaan. Lisäksi rakennetaan luokittelija, joka poimii tekstistä idiomit, sananlaskut ja fraasit käyttäen n-grammi- ja samankaltaisuusmittausmenetelmiä. Lopuksi tehtiin useita kokeita erilaisilla aineistoilla, mukaan lukien arabialaiset twiitit ja arabialaiset mikroblogit (hotellivaraukset, tuotearvostelut ja televisio-ohjelmien kommentit) julkisesti saatavilla olevilta arabiankielisiltä verkkoarvostelusivustoilta (sosiaalinen media, blogit, foorumit, sähköisen kaupankäynnin sivustot), jotta AIPSeLEXin kattavuutta ja tarkkuutta voitiin arvioida. Yleiset termit Sentimenttianalyysi, moderni standardiarabia, puhekieli, luonnollisen kielen käsittely.

**Tulos**

Idiomit-sanonnat - sananlaskujen sanasto nykyaikaista arabiaa ja puhekielistä tunneanalyysia varten

**Esimerkki 2.1313**

Yksi tapa lähestyä päästä päähän autonomista ajamista on oppia toimintatapafunktio, joka kartoittaa aistitiedon, kuten etukameran kuvakehyksen, ajotapahtumaksi jäljittelemällä asiantuntijakuljettajaa tai referenssikäytäntöä. Tämä voidaan tehdä valvotulla oppimisella, jossa toimintatapafunktio viritetään siten, että ennustetun ja todellisen toiminnan välinen ero minimoidaan. Tällä tavoin koulutetun toimintatapafunktion tiedetään kuitenkin kärsivän odottamattomasta käyttäytymisestä, joka johtuu epäsuhtaisuudesta referenssipolitiikan ja koulutettujen toimintatapafunktioiden saavutettavissa olevien tilojen välillä. Kehittyneemmissä jäljittelyoppimisalgoritmeissa, kuten DAggerissa, tämä ongelma ratkaistaan keräämällä iteratiivisesti harjoitusesimerkkejä sekä referenssi- että koulutetuista politiikoista. Nämä algoritmit vaativat usein suuren määrän kyselyitä referenssipolitiikalle, mikä ei ole toivottavaa, koska referenssipolitiikka on usein kallista. Tässä artikkelissa ehdotamme DAggerin laajennusta, SafeDAggeria, joka on kyselytehokas ja soveltuu paremmin päästä päähän tapahtuvaan autonomiseen ajamiseen. Arvioimme ehdotettua SafeDAggeria autokilpailusimulaattorissa ja osoitamme, että se todellakin vaatii vähemmän kyselyjä vertailupolitiikalle. Havaitsemme, että konvergenssi nopeutuu merkittävästi, minkä oletamme johtuvan automaattisen opetussuunnitelman oppimisen vaikutuksesta.

**Tulos**

Query-Efficient Imitation Learning for End-to-End Autonomous Driving -oppiminen päästä päähän -autonomiseen ajamiseen

**Esimerkki 2.1314**

NLP-tehtävät eroavat toisistaan niiden vaatiman semanttisen tiedon suhteen, eikä tällä hetkellä mikään yksittäinen semanttinen esitys täytä kaikkia vaatimuksia. Logiikkaan perustuvat esitykset kuvaavat lauserakennetta, mutta ne eivät tavoita merkityksen asteittaista ulottuvuutta. Jakaumamallit antavat sanoille ja lauseille luokiteltuja samankaltaisuusluokituksia, mutta ne eivät kuvaa riittävästi lauseen yleistä rakennetta. On siis väitetty, että nämä kaksi mallia täydentävät toisiaan. Tässä artikkelissa omaksumme hybridilähestymistavan, jossa yhdistetään logiikkapohjainen ja distributiivinen semantiikka todennäköisyyteen perustuvan logiikan avulla Markovin loogisissa verkoissa (MLN). Keskitymme tekstuaaliseen päättelyyn (RTE), joka on tehtävä, jossa voidaan hyödyntää molempien representaatioiden vahvuuksia. Järjestelmämme koostuu kolmesta osasta: 1) jäsennys ja tehtävän esittäminen, jossa syötetyt RTE-ongelmat esitetään todennäköisyyslogiikalla. Tämä eroaa huomattavasti niiden esittämisestä tavallisella ensimmäisen järjestyksen logiikalla. 2) tietopohjan rakentaminen painotettujen päättelysääntöjen muodossa eri lähteistä, kuten WordNetistä, parafraasikokoelmista sekä lennossa generoiduista leksikaalisista ja fraasijakautumissäännöistä. Käytämme Robinson-resoluution muunnelmaa tarvittavien päättelysääntöjen määrittämiseen. Lisää lähteitä voidaan helposti lisätä kartoittamalla ne loogisiin sääntöihin; järjestelmämme oppii resurssikohtaisen painon, joka tasoittaa resurssien välisiä skaalauseroja. 3) päättely, jossa näytämme, miten päättelyongelmat ratkaistaan tehokkaasti. Tässä asiakirjassa keskitymme SICK-tietokantaan, ja saavutamme huipputason tuloksen. Järjestelmämme käsittelee logiikassa yleistä lauserakennetta ja negaation kaltaisia ilmiöitä ja käyttää sitten Robinson-ratkaisuvaihtoehtoamme sanoja ja lyhyitä lauseita koskevien distribuutiojärjestelmien kyselyyn. Näin ollen käytämme järjestelmäämme arvioimaan distributiivisia leksikaalisia päättelymenetelmiä. Julkaisemme myös SICK-tietokannasta kysyttyjen sääntöjen joukon, joka voi olla hyvä resurssi niiden arvioimiseksi.

**Tulos**

Merkityksen esittäminen loogisen muodon ja vektoreiden yhdistelmällä

**Esimerkki 2.1315**

Sumeiden järjestelmien käyttäytymisen vertailemiseksi on pitkään käytetty sumean kielen ekvivalenssia, mutta tällä tasolla vertailu on liian karkeaa. Äskettäin sumeille äärellisille automaateille on otettu käyttöön hienovaraisempi käyttäytymismitta, bisimulaatio. Saadut tulokset soveltuvat kuitenkin vain äärellisten tilojen järjestelmiin. Tässä artikkelissa tarkastelemme bisimulaatiota yleisille sumeille järjestelmille, jotka voivat olla äärettömän tilaisia tai äärettömän tapahtumarikkaita, mallintamalla ne sumeina siirtymäjärjestelminä. Jotta bisimulaatio olisi helpompi ymmärtää ja tarkistaa, kuvaamme sitä kolmella tavalla: luettelemalla kokonaisia siirtymiä, vertaamalla yksittäisiä siirtymiä ja käyttämällä monotonista funktiota. Lisäksi käsittelemme sumeiden siirtymäjärjestelmien kompositio-operaatioita, osajärjestelmiä, kveenejä ja homomorfismeja ja keskustelemme niiden bisimulaatioon liittyvistä ominaisuuksista. Tässä esitetyt tulokset ovat hyödyllisiä yleisten sumeiden järjestelmien käyttäytymisen vertailussa. Erityisesti tämä mahdollistaa sen, että ääretön sumea järjestelmä voidaan suhteuttaa äärelliseen, helpommin analysoitavaan, samalla tavalla käyttäytyvään järjestelmään.

**Tulos**

Sumeiden siirtymäjärjestelmien bisimulaatiot

**Esimerkki 2.1316**

Tässä artikkelissa esitellään DifferenceBased Causality Learner (DBCL), algoritmi diskreettiaikaisten dynaamisten mallien luokan oppimiseen, joka edustaa kaikkia ajallisia kausaalisuhteita järjestelmässä tapahtuvia muutoksia ohjaavien erotusyhtälöiden avulla. Motivoimme tätä esitystä todellisten mekaanisten järjestelmien avulla ja todistamme, että DBCL on oikea oppimaan rakennetta aikasarjadatasta, jota vaikeuttaa havaittavien latenttien derivaattojen olemassaolo. Todistamme myös, että kausaalisten tekijöiden havaitsemista koskevien yleisten oletusten mukaisesti DBCL tunnistaa takaisinkytkentäsilmukoiden olemassaolon tai puuttumisen, mikä tekee mallista hyödyllisemmän ennustettaessa muuttujien manipuloinnin vaikutuksia, kun järjestelmä on tasapainossa. Argumentoimme analyyttisesti ja osoitamme empiirisesti DBCL:n edut verrattuna vektoriautoregressio- (VAR) ja Granger-kausaliteettimalleihin sekä Bayesin ja rajoitteisiin perustuvien rakenteiden löytämisalgoritmien muunnettuihin muotoihin. Lopuksi osoitamme, että algoritmimme voi löytää ihmisaivojen alfarytmien kausaalisia suuntia EEG-datasta.

**Tulos**

Miksi asiat muuttuvat: Eroihin perustuva syy-seuraus-oppija.

**Esimerkki 2.1317**

Tämän artikkelin tavoitteena on esitellä uusi algoritmi automaattiseen puheen ja tekstin väliseen tiivistämiseen, joka perustuu todennäköisyyksien ja graafien tilastollisiin eroavuuksiin. Syötteenä on kohinaa sisältävästä puhekeskustelusta peräisin oleva teksti ja tuloksena on tiivis tekstin tiivistelmä. Tuloksemme pilottitehtävässä CCCS Multiling 2015 French corpus ovat hyvin rohkaisevia.

**Tulos**

LIA-RAG: graafien ja todennäköisyyksien eroavaisuuksiin perustuva järjestelmä, jota sovelletaan puhe-tekstiyhteenvetoon.

**Esimerkki 2.1318**

Tekstien tuottaminen strukturoidusta datasta (esim. taulukosta) on tärkeää erilaisissa luonnollisen kielen käsittelytehtävissä, kuten kysymyksiin vastaamisessa ja dialogijärjestelmissä. Viimeaikaisissa tutkimuksissa tutkijat ovat käyttäneet neuraalisia kielimalleja ja koodaaja-dekooderikehyksiä taulukosta tekstiksi tuottamiseen. Näissä neuroverkkopohjaisissa lähestymistavoissa ei kuitenkaan mallinneta sisällön järjestystä tekstin tuottamisen aikana. Kun ihminen kirjoittaa yhteenvetoa tietyn taulukon perusteella, hän todennäköisesti ottaa huomioon sisällön järjestyksen ennen sanamuotoa. Esimerkiksi elämäkerrassa henkilön kansallisuus mainitaan tyypillisesti ennen ammattia. Tässä artikkelissa ehdotamme järjestyksen suunnittelua hyödyntävää tekstinmuodostusmallia, jolla eri kenttien väliset suhteet voidaan vangita ja käyttää näitä suhteita, jotta tuotetusta tekstistä saadaan sujuvampaa ja sulavampaa. Teimme kokeita WIKIBIO-tietokannalla ja saavutimme huomattavasti paremman suorituskyvyn kuin aiemmat menetelmät BLEU-, ROUGE- ja NIST-pisteiden osalta.

**Tulos**

Järjestyssuunnitteleva neurotekstin tuottaminen strukturoidusta datasta

**Esimerkki 2.1319**

Ensimmäinen vaihe kysymyksen käsittelyssä kysymysten vastausjärjestelmissä on kysymyksen yksityiskohtainen analyysi, jonka tarkoituksena on selvittää, mitä kysymyksessä kysytään ja miten siihen voidaan vastata täydellisesti. Kysymysten analysoinnissa käytetään useita tekniikoita luonnollisella kielellä esitettyjen kysymysten analysointiin: Stanford POS Tagger & parser arabian kielelle, nimettyjen entiteettien tunnistin, tokenizer, stop-sanojen poisto, kysymysten laajentaminen, kysymysten luokittelu ja kysymysten painopisteiden erottaminen. Käytämme lukuisia havaintosääntöjä ja koulutettua luokittelijaa, joka käyttää analyysin piirteitä havaitakseen kysymyksen tärkeät elementit, mukaan lukien: 1) kysymyksen osa, joka viittaa vastaukseen (fokus); 2) kysymyksen eri termit, jotka tunnistavat, minkä tyyppistä kokonaisuutta kysytään (leksikaaliset vastaustyypit); 3) kysymyksen laajentaminen ; 4) kysymyksen luokitteluprosessi yhdeksi tai useammaksi useista ja erilaisista tyypeistä; ja Kuvaamme, miten nämä elementit tunnistetaan, ja arvioimme täsmällisen tunnistuksen vaikutusta kysymysten vastausjärjestelmäämme käyttämällä Mean Reciprocal Rank(MRR)-tarkkuusmittaria.

**Tulos**

KYSYMYSANALYYSI ARABIAN KIELEN KYSYMYKSIIN VASTAAVIA JÄRJESTELMIÄ VARTEN

**Esimerkki 2.1320**

Tässä artikkelissa perustelemme teoreettisesti Higgs Boson Machine Learning Challenge -kilpailun osallistujien keskuudessa suositun lähestymistavan, jolla optimoidaan likimääräinen mediaanimerkitsevyys (AMS). Lähestymistapa perustuu seuraavaan kaksivaiheiseen menettelyyn. Ensin opitaan reaaliarvoinen funktio f minimoimalla binääriluokittelun korvaava häviö, kuten logistinen häviö, harjoitusnäytteellä. Sitten, kun f on annettu, kynnysarvo θ̂ viritetään erillisellä validointinäytteellä AMS:n suoralla optimoinnilla. Osoitamme, että tuloksena saadun luokittelijan (joka on saatu kynnyksen määrittämisestä f:lle θ̂:llä) katumus mitattuna AMS:n neliön suhteen on ylempänä kuin f:n katumus mitattuna logistisen tappion suhteen. Näin ollen todistamme, että logistisen surrogaatin minimointi on johdonmukainen tapa optimoida AMS.

**Tulos**

AMS:n johdonmukainen optimointi logistisen tappion minimoinnin avulla

**Esimerkki 2.1321**

Nykyinen suuntaus kohteiden havaitsemisessa ja paikantamisessa on ennusteiden oppiminen suurikapasiteettisilla syvillä neuroverkoilla, jotka on koulutettu hyvin suurella määrällä kommentoitua dataa ja jotka käyttävät paljon prosessointitehoa. Tässä työssä ehdotamme uutta neuromallia, joka ennustaa suoraan bounding box -koordinaatteja. Osuutemme erityispiirre on ennusteiden paikallinen laskenta uudenlaisella paikallisten parametrien jakamisella, joka pitää koulutettavien parametrien kokonaismäärän alhaisena. Mallin keskeiset osat ovat spatiaaliset 2D-LSTM-rekurrenssikerrokset, jotka välittävät kuvan alueiden välistä kontekstuaalista tietoa. Osoitamme, että tämä malli on nykyistä tekniikan tasoa tehokkaampi sovelluksissa, joissa harjoitusdataa ei ole yhtä runsaasti kuin klassisessa kokoonpanossa luonnollisten kuvien ja Imagenet/Pascal VOC -tehtävissä. Tavoitteenamme on erityisesti tekstin havaitseminen asiakirjakuvissa, mutta menetelmämme ei rajoitu tähän asetelmaan. Ehdotettu malli helpottaa myös monien kohteiden havaitsemista yhdessä kuvassa ja pystyy käsittelemään erikokoisia syötteitä ilman koon muuttamista.

**Tulos**

Oppiminen havaitsemaan ja paikallistamaan monia kohteita harvoista esimerkeistä

**Esimerkki 2.1322**

Huomiomallista on tullut neuraalisen konekääntämisen (NMT) vakiokomponentti, ja se ohjaa käännösprosessia keskittymällä valikoivasti lähdelauseen osiin ennustettaessa kutakin kohdesanaa. Huomaamme kuitenkin, että kohdesanan muodostaminen ei riipu ainoastaan lähdelauseesta, vaan se riippuu myös suuresti aiemmin muodostetuista kohdesanoista, erityisesti kaukana sijaitsevista sanoista, joita on vaikea mallintaa rekurrenssia neuroverkkoja käyttämällä. Tämän ongelman ratkaisemiseksi ehdotamme tässä artikkelissa NMT:ssä uudenlaista look-ahead-huomiointimekanismia, jonka tarkoituksena on ottaa suoraan huomioon kohdesanojen välinen riippuvuussuhde. Lisäksi suunnittelemme kolme mallia, joilla integroimme look-ahead-huomion perinteiseen huomiomalliin. Kokeet NIST Chinese-to-English- ja WMT English-to-German -käännöstehtävissä osoittavat, että ehdotettu look-ahead-huomiointimekanismimme parantaa huomattavasti tilannetta verrattuna uusimpiin perusmalleihin.

**Tulos**

Look-ahead Attention for Generation in Neural Machine Translation (Ennakoiva tarkkaavaisuus neuraalisessa konekääntämisessä)

**Esimerkki 2.1323**

Arvioimme merkkitason käännösmenetelmää neuraalista semanttista jäsennystä varten laajalla lauseiden korpuksella, joka on merkitty abstrakteilla merkitysrepresentaatioilla (AMR). Käyttämällä seq2seq-mallia ja AMR-merkitysten triviaalia esi- ja jälkikäsittelyä saamme perustarkkuudeksi 53,1 (F-tulos AMR-kolmioilla). Tutkimme neljää eri lähestymistapaa tämän perustuloksen parantamiseksi: (i) AMR-haarojen järjestäminen uudelleen vastaamaan syötteen lauseen sanajärjestystä parantaa suorituskykyä 58,3:een; (ii) myös (automaattisesti tuotettujen) puheosamerkkien lisääminen syötteeseen parantaa suorituskykyä (57,2); (iii) samoin supermerkkien lisääminen (usein toistuvien merkkisarjojen yhdistäminen yhdeksi merkiksi), jolloin tulos on 57,4; (iv) suurimman parannuksen saa lisäämällä hopeastandardin mukaista harjoitusdataa, joka on saatu valmiilla jäsentimellä, jolloin F-tulokseksi saadaan 64,0. Kaikkien neljän tekniikan yhdistäminen johtaa F-tulokseen 69,0, joka on AMR-jäsennyksen huippuluokkaa. Tämä on huomattavaa, koska lähestymistapa on suhteellisen yksinkertainen: ainoa käyttämämme eksplisiittinen kielitieto ovat puhekieliset tunnisteet.

**Tulos**

Neuraalinen semanttinen jäsennys merkkipohjaisen käännöksen avulla: Kokeet abstrakteilla merkitysrepresentaatioilla

**Esimerkki 2.1324**

Dendriittisolualgoritmin (DCA) teoreettiset analyysit ovat tuottaneet useita kritiikkejä sen rakenteesta ja toiminnasta. Tämän seurauksena kirjallisuudessa on ehdotettu useita muutoksia ja korjauksia näiden havaintojen korjaamiseksi. Tässä työssä tutkitaan, mitä vaikutuksia on sillä, että DCA:n luokitteluvaihe (jonka tiedetään olevan virheellinen) korvataan perinteisellä koneoppimistekniikalla. Lisäksi tässä työssä kyseenalaistetaan niiden DCA:n ainutlaatuisten ominaisuuksien ansiot, joita ei ole vielä analysoitu perusteellisesti. Jos yhdestäkään näistä ominaisuuksista ei voida todeta olevan hyötyä perinteisiin lähestymistapoihin verrattuna, DCA:n "korjaaminen" on todennäköisesti vähemmän tehokasta kuin uuden algoritmin luominen. Tässä työssä tarkastellaan DCA:n dynaamista suodatusominaisuutta ja kyseenalaistetaan tämän ainutlaatuisen ominaisuuden hyödyllisyys poikkeamien havaitsemisongelmassa. Havaitaan, että vaikka tämä ominaisuus on edullinen meluisessa, ajallisesti järjestetyssä luokittelussa, se ei ole yhtä hyödyllinen kuin perinteinen staattinen suodatin synteettisen tietokokonaisuuden käsittelyssä. Johtopäätöksenä todetaan, että DCA:n ainutlaatuisia ominaisuuksia on vielä tutkimatta. Ehdotetaan alueita, joista voi olla hyötyä keinotekoisten immuunijärjestelmien yhteisölle.

**Tulos**

Hiljaisuus luokassa : luokittelu, melu ja dendriittisolualgoritmi

**Esimerkki 2.1325**

Esittelemme algoritmikokonaisuuden ulottuvuudesta riippumatonta samankaltaisuuslaskentaa (Dimension Independent Similarity Computation, DISCO) varten, jolla lasketaan kaikki pareittaiset samankaltaisuudet erittäin suurten ulottuvuuksien harvalukuisten vektoreiden välillä. Kaikki tuloksemme ovat todistettavasti riippumattomia ulottuvuudesta, mikä tarkoittaa, että lukuun ottamatta datan triviaalin lukemisen alkukustannuksia, kaikki myöhemmät operaatiot ovat riippumattomia ulottuvuudesta, joten ulottuvuus voi olla hyvin suuri. Tutkimme Cosine-, Dice-, Overlap-, Conditional- ja Jaccard- samankaltaisuusmittoja. Jaccardin samankaltaisuutta varten käytämme parannettua versiota MinHashista. Tuloksemme on suunnattu MapReduce-kehykseen. Validoimme teoreemamme empiirisesti suuressa mittakaavassa käyttämällä sosiaalisen verkostoitumisen sivuston Twitterin tietoja.

**Tulos**

Ulottuvuudesta riippumaton samankaltaisuuden laskenta

**Esimerkki 2.1326**

Autokooderit ovat ehkä tunnetuimpia ei-todennäköisyyspohjaisia menetelmiä representaatioiden oppimisessa. Ne ovat käsitteellisesti yksinkertaisia ja helppoja kouluttaa. Viimeaikaiset teoreettiset työt ovat valottaneet niiden kykyä vangita moninaisia rakenteita ja luoneet yhteyksiä tiheysmallinnukseen. Tämä on motivoinut tutkijoita etsimään tapoja autokooderin pisteytykseen, mikä on edistänyt niiden käyttöä luokittelussa. Gated autokooderit (Gated auto-encoders, GAE) ovat mielenkiintoinen ja joustava laajennus autokoodereihin, jotka voivat oppia eri kuvien välisiä muunnoksia tai kuvien sisäisiä pikselikovariansseja. Niitä on kuitenkin tutkittu teoreettisesti ja empiirisesti paljon vähemmän. Tässä työssä sovellamme dynaamisten järjestelmien näkökulmaa GAE:iin, johdamme pisteytysfunktion ja luomme yhteyksiä rajoitettuihin Boltzmannin koneisiin. Lisäksi osoitamme niiden tehokkuuden syvän oppimisen vertailuarvojen avulla yhden ja usean merkin luokittelussa.

**Tulos**

Pisteytys ja luokittelu Gated-autokoodereilla

**Esimerkki 2.1327**

Ehdotamme menetelmää, jolla kaksiulotteiset sijainnit voidaan upottaa jatkuvaan vektoriavaruuteen käyttämällä neuroverkkopohjaista mallia, joka sisältää Gaussin jakaumien sekoituksia, ja esitämme kaksi mallivaihtoehtoa tekstipohjaista maantieteellistä paikannusta ja leksikaalista dialektologiaa varten. Twitter-tiedoilla arvioituna ehdotettu malli päihittää perinteisen regressiopohjaisen maantieteellisen sijainnin ja antaa paremman arvion epävarmuudesta. Osoitamme myös esityksen tehokkuuden sanojen ennustamisessa sijainnin perusteella leksikaalisessa dialektologiassa ja arvioimme sitä DARE-tietokannan avulla.

**Tulos**

Sijainnin jatkuva esittäminen geopaikannusta ja leksikaalista dialektologiaa varten sekoitustiheysverkkojen avulla.

**Esimerkki 2.1328**

Tutkimme demonstraatioista oppimista koskevaa nykyistä työtä pidemmälle esittämällä seuraavan kysymyksen: "Toisin sanoen, voiko robotti itsenäisesti oppia opetuskäytännön asiantuntijan demonstraatiosta ja käyttää sitä opastaakseen tai tehdäkseen yhteistyötä ihmisten kanssa monimutkaisten tehtävien suorittamisessa epävarmoissa ympäristöissä? Tässä artikkelissa pyrimme löytämään ratkaisun tähän ongelmaan hyödyntämällä ajatusta, että ihmiset usein implisiittisesti hajottavat korkeamman tason tehtävän useisiin alatavoitteisiin, joiden suorittaminen vie tehtävän lähemmäksi sen loppuunsaattamista. Ehdotamme Dirichlet-prosessiin perustuvaa ei-parametrista käänteistä vahvistusoppimista (Dirichlet process based non-parametric Inverse Reinforcement Learning, DPMIRL) palkkiopohjaista, valvomatonta tehtäväalueen klusterointia alatavoitteisiin. Tämän lähestymistavan osoitetaan vangitsevan piilevät osatavoitteet, joita ihmisopettaja olisi käyttänyt aloittelijan kouluttamiseen. Käsite "toimintoalku" esitellään keinona välittää opettamiskäytäntö ihmisille mahdollisimman yksinkertaisella tavalla ja laskennallisesti tehokkaana välineenä demonstraatiodatan segmentoimiseksi. Arvioimme lähestymistapaamme kokeilla hydraulisesti toimivalla kaivinkoneen pienoismallilla ja arvioimme ja vertailemme erilaisia robotin käyttämiä opetusstrategioita.

**Tulos**

Voivatko yhteisrobotit oppia opettamaan?

**Esimerkki 2.1329**

Verkkoteknologian kehittymisen ja kasvun myötä Internetissä on valtava määrä tietoa, jota internetin käyttäjät voivat käyttää, ja myös tietoa syntyy paljon. Internetistä on tullut verkko-oppimisen, ajatustenvaihdon ja mielipiteiden jakamisen foorumi. Sosiaaliset verkostoitumissivustot, kuten Twitter, Facebook ja Google+, kasvattavat nopeasti suosiotaan, sillä niiden avulla ihmiset voivat jakaa ja ilmaista näkemyksiään aiheista, keskustella eri yhteisöjen kanssa tai lähettää viestejä ympäri maailmaa. Twitter-tietojen sentimenttianalyysin alalla on tehty paljon työtä. Tässä tutkimuksessa keskitytään pääasiassa Twitter-tietojen sentimenttianalyysiin, jonka avulla voidaan analysoida tietoja twiiteissä, joissa mielipiteet ovat hyvin jäsentymättömiä, heterogeenisiä ja joko myönteisiä tai kielteisiä tai joissain tapauksissa neutraaleja. Tässä asiakirjassa esitetään katsaus ja vertaileva analyysi nykyisistä mielipiteiden louhintatekniikoista, kuten koneoppimisesta ja leksikonipohjaisista lähestymistavoista, sekä arviointimittareita. Käyttämällä erilaisia koneoppimisalgoritmeja, kuten Naive Bayes, Max Entropy ja Support Vector Machine, tutkimme Twitterin tietovirtoja. Olemme myös keskustelleet Twitterin tunneanalyysin yleisistä haasteista ja sovelluksista.

**Tulos**

Twitter-datan tunneanalyysi: katsaus tekniikoihin

**Esimerkki 2.1330**

Polkusuunnittelua pidetään tekoälyssä tyypillisesti graafihakuun perustuvana ongelmana, ja R\* on sen ratkaisemiseen räätälöity huippuluokan algoritmi. Algoritmi hajottaa tietyn polunhakutehtävän sarjaksi osatehtäviä, joista jokainen voidaan helposti (laskennallisesti) ratkaista tunnetuilla menetelmillä (kuten A\*). Purkamisen suorittamiseen käytetään parametrisoitua satunnaisvalintaa, minkä vuoksi R\*:n suorituskyky riippuu pitkälti sen syöttöparametrien valinnasta. Työssämme muotoilemme erilaisia oletuksia, jotka koskevat R\*-parametrien mahdollisia ylä- ja alarajoja, niiden keskinäistä riippuvuutta ja niiden vaikutusta R\*:n suorituskykyyn. Sitten arvioimme näitä oletuksia suorittamalla suuren määrän kokeita. Tuloksena muotoilemme joukon heuristisia sääntöjä, joita voidaan käyttää R\*-parametrien arvojen alustamiseen tavalla, joka johtaa algoritmin parhaaseen suorituskykyyn.

**Tulos**

Satunnaistetun heuristisen haun hienosäätö 2D-polun suunnittelua varten: R\*-algoritmin parhaiden syöttöparametrien löytäminen kokeilujen avulla.

**Esimerkki 2.1331**

Tässä asiakirjassa esitellään lähestymistapa, jolla formalisoidaan ja varmistetaan tietosuojan käyttöominaisuuksien luokka tietoon perustuvissa järjestelmissä. Aiemmista töistä poiketen keskitymme suojattujen tietotyyppien edustajien (eli vahvojen ennustajien) käyttörajoituksiin. Määritelmämme liittää välittäjän käytön ohjelmassa tapahtuviin välilaskutoimituksiin ja tunnistaa kaksi olennaista ominaisuutta, jotka luonnehtivat tätä käyttäytymistä: 1) sen tulos liittyy vahvasti kyseiseen suojattuun tietotyyppiin ja 2) se todennäköisesti vaikuttaa kausaalisesti ohjelman lopulliseen tulokseen. Tämän määritelmän tiettyä esimerkkiä varten esittelemme ohjelma-analyysitekniikan, joka havaitsee mallissa esiintyvät välityskäytön tapaukset ja tarjoaa todistajan, joka tunnistaa, missä vastaavan ohjelman osissa kyseinen käyttäytyminen esiintyy. Koska kaikki suojatun tietotyypin välityskäytön tapaukset eivät ole sopimattomia, käytämme normatiivista tuomio-orakkelia, joka tekee tämän sopimattomuuden määrityksen tietyn todistajan osalta. Korjausalgoritmimme käyttää epäsopivaa välityskäytön todistajaa muuttamaan mallin sellaiseksi, jossa ei todistettavasti esiinny välityskäyttöä, ja välttää samalla muutoksia, jotka vaikuttavat kohtuuttomasti luokittelutarkkuuteen. Sosiaalisten tietokokonaisuuksien avulla tekemämme arviointi osoittaa, että nämä algoritmit pystyvät havaitsemaan välityskäytön tapaukset, joita olisi vaikea löytää nykyisillä tekniikoilla, ja poistamaan ne säilyttäen samalla hyväksyttävän luokittelutuloksen.

**Tulos**

Tietosuojan käyttö tietoon perustuvissa järjestelmissä

**Esimerkki 2.1332**

Kielestä riippumatonta Stemmeriä on aina etsitty. Yhden N-grammin merkintätekniikka toimii hyvin, mutta se tuottaa usein kantoja, jotka alkavat pikemminkin välimerkeillä kuin alkumerkillä. Esittelemme uuden tekniikan, joka vie N-grammien käsitteen askeleen eteenpäin, ja vertaamme menetelmäämme alan vakiintuneeseen algoritmiin Porterin Stemmeriin. Porter's Stemmer on kieliriippuvainen, eikä ehdotetun menetelmän suorituskyky ole huonompi kuin sen.

**Tulos**

Kielestä riippumattoman kantojen poistoalgoritmin luominen, toteuttaminen ja arviointi

**Esimerkki 2.1333**

Suunnittelun laskennallisen monimutkaisuuden varhaiset luokittelut STRIPS- (Bylander) ja SAS-ohjelmissa (Bäckström ja Nebel) ovat vaikuttaneet monin tavoin suunnittelun seuraavaan tutkimukseen. Menemme takaisin ja analysoimme uudelleen niiden alaluokkia, mutta tällä kertaa käyttäen uudenaikaisempaa parametrisoidun kompleksisuusanalyysin työkalua. Näin saadaan uusia tuloksia, jotka yhdessä vanhojen tulosten kanssa antavat yksityiskohtaisemman kuvan kompleksisuusmaisemasta. Osoitamme erottelutuloksia, jotka eivät ole mahdollisia tavanomaisen kompleksisuusteorian avulla, mikä osaltaan selittää, miksi tietyt suunnittelutapaukset ovat käytännössä näyttäneet yksinkertaisemmilta kuin teoria on ennustanut. Näytämme erityisesti, että tietyt käytännön kannalta kiinnostavat rajoitukset ovat parametroidussa merkityksessä käsiteltävissä ja että yksinkertainen heuristiikka riittää saamaan tunnetun osittaisjärjestyksen suunnittelijan hyödyntämään tätä tosiasiaa.

**Tulos**

Suunnittelun monimutkaisuus uudelleen tarkasteltuna - parametrisoitu analyysi

**Esimerkki 2.1334**

Asiakasarvostelujen perusteella tapahtuva mielipiteiden louhinta on yleistynyt viime vuosina. Arvostelujen lauseet luokitellaan kuitenkin yleensä itsenäisesti, vaikka ne ovat osa arvostelun argumentatiivista rakennetta. Arvostelun lauseet rakentuvat ja kehittyvät toistensa pohjalta, joten kunkin lauseen luokittelussa olisi käytettävä tietoa arvostelun rakenteesta ja lauseyhteydestä. Osoitamme tämän hypoteesin näkökulmapohjaisen sentimenttianalyysin tehtävän osalta mallintamalla arvostelun lauseiden keskinäiset riippuvuudet hierarkkisen kaksisuuntaisen LSTM:n avulla. Osoitamme, että hierarkkinen malli on parempi kuin kaksi ei-hierarkkista perusmallia, että sillä saavutetaan huipputekniikan kanssa kilpailukykyisiä tuloksia ja että se on parempi kuin huipputekniikka viidessä monikielisessä ja monialaisessa tietokokonaisuudessa ilman käsin kehitettyjä piirteitä tai ulkoisia resursseja.

**Tulos**

Arvostelujen hierarkkinen malli näkökulmapohjaista tunneanalyysia varten

**Esimerkki 2.1335**

Tarkastelemme tehtävää, joka koskee asenteiden tunnistamista tekstistä tiettyä joukkoa entiteettejä kohtaan. Tavanomaisesti tämä tehtävä jaetaan kahteen erilliseen osatehtävään: kohteen tunnistaminen, jossa tunnistetaan, mainitaanko kukin entiteetti tekstissä joko eksplisiittisesti tai implisiittisesti, ja polariteettiluokittelu, jossa tunnistettua entiteettiä (kohdetta) koskeva tarkka tunnetila luokitellaan positiiviseksi, negatiiviseksi tai neutraaliksi. Sen sijaan osoitamme, että asenteiden tunnistaminen voidaan ratkaista kokonaisvaltaisella koneoppimisarkkitehtuurilla, jossa nämä kaksi osatehtävää on lomitettu toisiinsa syvämuistiverkon avulla. Tällä tavoin kohteen tunnistamisessa tuotetut signaalit antavat vihjeitä polariteetin luokittelua varten, ja päinvastoin ennustettu polariteetti antaa palautetta kohteiden tunnistamiseen. Lisäksi kohdejoukon käsittelyt vaikuttavat myös toisiinsa - opitut representaatiot voivat jakaa saman semantiikan joidenkin kohteiden kohdalla mutta vaihdella toisten kohdalla. Ehdotettu syvämuistiverkko AttNet päihittää menetelmät, jotka eivät ota huomioon osatehtävien välisiä tai kohteiden välisiä vuorovaikutuksia, mukaan lukien perinteiset koneoppimismenetelmät ja uusimmat syväoppimismallit.

**Tulos**

Syvämuistiverkot asennon tunnistamiseen

**Esimerkki 2.1336**

Taksonomisen ontologian kahden käsitteen välistä lyhintä polkua käytetään yleisesti kuvaamaan käsitteiden välistä semanttista etäisyyttä reunapohjaisissa semanttisissa samankaltaisuusmittauksissa. Aiemmin reuna-arvojen laskentaa on pidetty oletusmenetelmänä polun laskemisessa, sillä se on yksinkertainen, intuitiivinen ja vähäinen laskennallinen monimutkaisuus. Suuressa leksikaalisessa taksonomiassa, kuten WordNetissä, on kuitenkin epäsäännöllinen tiheys käsitteiden välisiä linkkejä, koska se on laaja-alainen. Reunojen laskentaan perustuva polkujen laskenta ei pysty ratkaisemaan tätä epäyhtenäisyysongelmaa. Tässä asiakirjassa esitämme, että polun laskenta voidaan erottaa reunoihin perustuvista samankaltaisuusmittauksista ja muodostaa erilaisia yleisiä laskentamalleja. Ratkaistaksemme käsitteiden tiheyden epäyhtenäisyysongelman suuressa taksonomisessa ontologiassa ehdotamme siksi uutta polunlaskentamallia, joka perustuu käsitteiden paikallisen aluetiheyden kompensointiin, joka on yhtä suuri kuin lyhimmän polun käsitteiden alasummien suorien hyponyymien määrä. Tässä polkumallissa käsitteiden paikallista aluetiheyttä pidetään reunapohjaisen polun laajennuksena ja muunnetaan paikallinen aluetiheys jaettuna niiden syvyydellä reunapohjaisen polun kompensaatioksi säädettävän parametrin avulla, minkä ajatuksen on osoitettu olevan sopusoinnussa tietoteorian kanssa. Tämä malli on yleinen polun laskentamalli, ja sitä voidaan soveltaa erilaisissa reunapohjaisissa samankaltaisuusalgoritmeissa. Kokeiden tulokset osoittavat, että ehdotettu polkumalli parantaa keskimääräistä korrelaatiota reunapohjaisten mittausten ja ihmisen arvioiden välillä Miller- ja Charles-vertailussa alle 0,8:sta yli 0,85:een, ja sillä on suuri hyöty tehokkuudessa verrattuna informaatiosisällön (IC) laskentaan dynaamisessa ontologiassa, ja siten se ratkaisee onnistuneesti taksonomisen ontologian epäyhtenäisyysongelman.

**Tulos**

Tiheyskompensaatioon perustuva polkujen laskentamalli semanttisen samankaltaisuuden mittaamiseen

**Esimerkki 2.1337**

Todennäköisyysmallien oppimista haittaa usein huomattavasti normalisointikertoimen ja sen johdannaisten yleinen vaikeasti ratkaistavissa oleva ongelma. Tässä ehdotamme uutta oppimistekniikkaa, jolla vältetään tarve laskea hankala normalisointikerroin tai ottaa näyte mallin tasapainojakaumasta. Tämä saavutetaan luomalla dynamiikka, joka muuttaisi havaitun datajakauman mallijakaumaksi, ja asettamalla sitten tavoitteeksi datajakaumasta poispäin suuntautuvan alkutodennäköisyysvirran minimointi. Pistemäärän sovittaminen, miniminopeusoppiminen ja tietyt kontrastiivisen divergenssin muodot ovat tämän oppimistekniikan erikoistapauksia. Osoitamme minimitodennäköisyysvirran oppimisen soveltamisen Ising-mallien, syvien uskomusverkkojen, monimuuttujaisten Gaussin jakaumien ja jatkuvan mallin, jossa on hyvin yleinen potenssisarjana määritelty energiafunktio, parametrien estimointiin. Ising-mallin tapauksessa minimitodennäköisyysvirtauksen oppiminen on oppimisaikana noin kaksi kertaluokkaa parempi kuin nykyiset tekniikat, kun palautettujen parametrien virheet ovat vertailukelpoisia. Toivomme, että tämä tekniikka lieventää nykyisiä rajoituksia, jotka koskevat käytännöllisesti katsoen käyttökelpoisten todennäköisyysmallien luokkia.

**Tulos**

Minimitodennäköisyys Virtauksen oppiminen

**Esimerkki 2.1338**

Tutkimme sellaisten todennäköisyyslogiikkaohjelmien merkitystä ja monimutkaisuutta, jotka koostuvat joukosta sääntöjä ja joukosta riippumattomia todennäköisyysfaktoja (eli Saton jakaumasemantiikkaan perustuvista ohjelmista). Keskitymme kahteen semantiikkaan, jotka perustuvat stabiileihin ja hyvin perusteltuihin malleihin. Osoitamme, että stabiileihin malleihin perustuva semantiikka (jota kutsutaan "uskottavuussemantiikaksi") tuottaa todennäköisyysmallien joukot, jotka hallitsevat äärettömän monotonisia Choquet-kapasiteetteja; kuvaamme useita hyödyllisiä seurauksia tästä tuloksesta. Tämän jälkeen tarkastelemme todennäköisyyslogiikkaohjelmilla tehtävän päättelyn monimutkaisuutta. Erottelemme toisistaan päättelyn monimutkaisuuden, kun probabilistinen ohjelma ja kysely on annettu (päättelyn monimutkaisuus), ja päättelyn monimutkaisuuden, kun probabilistinen ohjelma on kiinteä ja kysely on annettu (kyselyn monimutkaisuus, joka on sukua tietokantateoriassa käytetylle tiedon monimutkaisuudelle). Saamme tuloksia päättely- ja kyselykompleksisuudesta asyklisille, kerrostuneille ja syklisille propositionaalisille ja relaatiopohjaisille ohjelmille; kompleksisuus yltää laskentahierarkian eri tasoille ja jopa eksponentiaalisille tasoille.

**Tulos**

Todennäköisyyslogiikkaohjelmien semantiikasta ja kompleksisuudesta

**Esimerkki 2.1339**

MIML-oppiminen (Multi-instance multi-label) on monelta kannalta haastava ongelma. Tällaiset oppimismenetelmät voivat olla hyödyllisiä monissa lääketieteellisissä diagnoosisovelluksissa, kuten rintasyövän havaitsemisessa ja luokittelussa. Tässä tutkimuksessa käytetään digiPATH-tietokannan osajoukkoa (digitaalisia rintasyövän histopatologisia kuvia koko dia-aineistosta) kuuden uusimman MIML-menetelmän harjoitteluun ja arviointiin. Lopuksi vertaillaan näiden menetelmien suorituskykyä tehokkaiden arviointimittareiden avulla. Osoittautuu, että MIML-kNN saavuttaa parhaan tuloksen eli 65,3 prosentin keskimääräisen tarkkuuden, ja useimmat muutkin menetelmät saavuttavat hyväksyttäviä tuloksia.

**Tulos**

Arviointi yhteisestä monitilanneoppimisesta (Multi-Instance Multi-Label Learning) rintasyövän diagnosointia varten.

**Esimerkki 2.1340**

Esittelemme uuden lähestymistavan epäselvyyksien havaitsemiseen käyttämällä kaksisuuntaista pitkäaikaista ja lyhytaikaista muistia (BLSTM) hyödyntävää neuroverkkoa. Sanasekvenssin lisäksi malli ottaa syötteeksi kuvion vastaavuusominaisuuksia, jotka kehitettiin vähentämään herkkyyttä sanaston koolle harjoittelussa, mikä johtaa parempaan suorituskykyyn kuin pelkkä sanasekvenssi. BLSTM hyödyntää eksplisiittisiä korjaustiloja tavanomaisten korjaustilojen lisäksi. Lopullisessa tuotoksessa hyödynnetään kokonaislukuista lineaarista ohjelmointia, jotta voidaan ottaa huomioon hajanaisuuden rakenteen rajoitukset. Switchboard-korpuksella tehdyissä kokeissa malli saavutti parhaan mahdollisen suorituskyvyn sekä tavallisessa epäselvyyksien havaitsemistehtävässä että korjausten havaitsemistehtävässä. Analyysi osoittaa, että malli havaitsee muita kuin toistuvia epäsuoruuksia paremmin, sillä niitä on yleensä paljon vaikeampi havaita.

**Tulos**

Disfluting-tunnistus kaksisuuntaisen LSTM:n avulla

**Esimerkki 2.1341**

Tässä artikkelissa tutkitaan telugun käsikirjoituksen entropiaa. Koska tämä kirjoitus on tavukirjoitus eikä aakkoskirjoitus, entropian laskeminen on hieman monimutkaista.

**Tulos**

Telugun entropia

**Esimerkki 2.1342**

Tässä artikkelissa esittelemme RankPL:n, mallinnuskielen, jota voidaan pitää todennäköisyysohjelmointikielen laadullisena muunnoksena, jonka semantiikka perustuu Spohnin ranking-teoriaan. Laajasti ottaen RankPL:ää voidaan käyttää sellaisten prosessien esittämiseen ja päättelyyn, joihin liittyy epävarmuutta, joka voidaan ilmaista erottamalla "normaalit" ja "yllättävät" tapahtumat toisistaan. RankPL mahdollistaa (iteroidun) rankingin tarkistamisen vaihtoehtoisten ohjelmatilojen yli ja tukee erilaisia päättelytapoja, kuten abduktiota ja kausaalista päättelyä. Esittelemme kielen, sen merkityssemantiikan ja useita käytännön esimerkkejä. Käsittelemme myös RankPL:n toteutusta, joka on ladattavissa.

**Tulos**

RankPL: Kvalitatiivinen todennäköisyysohjelmointikieli (Qualitative Probabilistic Programming Language)

**Esimerkki 2.1343**

Luomme siirtymiin perustuvan riippuvuuksien jäsentäjän käyttämällä yleiskäyttöistä oppimis- ja hakujärjestelmää. Tuloksena on nopea ja tarkka jäsentäjä monille kielille. Verrattuna muihin siirtymiin perustuviin riippuvuuksien jäsennysmenetelmiin jäsennyksemme tarjoaa samanlaisen tilastollisen ja laskennallisen suorituskyvyn kuin tunnetuimmat lähestymistavat ja välttää samalla erilaisia haittoja, kuten satunnaistamista, ylimääräisiä ominaisuusvaatimuksia ja mukautettuja oppimisalgoritmeja. Osoitamme, että riippuvuuksien jäsentäjä on mahdollista toteuttaa avoimen lähdekoodin oppimis- ja hakukirjaston avulla noin 300 rivillä C++-koodia, kun taas nykyiset järjestelmät vaativat usein useita tuhansia rivejä.

**Tulos**

Riippuvuuksien etsimisen opettelu

**Esimerkki 2.1344**

Tietokannan (KB) täydentäminen lisää uusia tietoja KB:hen tekemällä päätelmiä olemassa olevista tiedoista, esimerkiksi päättelemällä suurella todennäköisyydellä kansalaisuus(X,Y) syntynytIn(X,Y) -tiedoista. Useimmat aiemmat menetelmät päättelevät yksinkertaisia yhden askeleen relaatiosynonyymejä tai käyttävät todisteena usean askeleen relaatiopolkua, jota käsitellään atomisena ominaisuutena, kuten bornIn(X,Z)→ containedIn(Z,Y). Tässä artikkelissa esitellään lähestymistapa, jossa monihyppyrelaatioiden yhteyksiä perustellaan ei-atomisesti, vaan polun implikaatiot muodostetaan käyttämällä toistuvaa neuroverkkoa (RNN), joka ottaa syötteenä polun binäärisen relaation vektorin upotukset. Näin voidaan yleistää polkuihin, joita ei ole nähty koulutushetkellä, mutta myös ennustaa yhdellä suuren kapasiteetin RNN:llä uusia relaatiotyyppejä, joita ei ole nähty, kun koostumismallia koulutettiin (zero-shot learning). Kokosimme uuden tietokokonaisuuden, joka sisältää yli 52 miljoonaa relaatiokolmiota, ja osoitimme, että menetelmämme parantaa perinteistä luokittelijaa 11 prosenttia ja ennalta koulutettuja upotuksia hyödyntävää menetelmää 7 prosenttia.

**Tulos**

Kompositionaaliset vektoriavaruusmallit tietopohjan täydentämistä varten

**Esimerkki 2.1345**

Sequence-to-sequence-neuraaliset käännösmallit oppivat lauseparien välisiä semanttisia ja syntaktisia suhteita optimoimalla kohteen todennäköisyyttä lähteen perusteella, eli p(y|x), mikä on tavoite, joka jättää huomiotta muut mahdollisesti hyödylliset tietolähteet. Esitämme neuraaliselle MT:lle vaihtoehtoisen tavoitefunktion, joka maksimoi lähde- ja kohdelauseiden keskinäisen informaation ja mallintaa lähteiden ja kohteiden kaksisuuntaista riippuvuutta. Toteutamme mallin yksinkertaisella uudelleenjärjestämismenetelmällä ja esittelemme myös dekoodausalgoritmin, joka lisää ensimmäisen läpikäynnin tuottaman N-parhaan luettelon monimuotoisuutta. Kun ehdotettuja malleja sovelletaan WMT German/English- ja French/English-tehtäviin, ne parantavat johdonmukaisesti suorituskykyä sekä tavallisiin LSTM- että tarkkaavaisuuspohjaisiin neuraalisiin MT-arkkitehtuureihin verrattuna.

**Tulos**

Keskinäinen informaatio ja monipuolinen dekoodaus parantavat neuraalista konekääntämistä.

**Esimerkki 2.1346**

Tässä artikkelissa esitellään Centre for Development of Advanced Computing Mumbain (CDACM) osallistuminen NLP Tools Contest on Part-Of-Speech (POS) Tagging For Code-mixed Indian Social Media Text (POSCMISMT) 2015 -kilpailuun (joka järjestetään yhdessä ICON 2015 -tapahtuman kanssa). Toimitimme tuloksia hindin (hi), bengalin (bn) ja telugun (te) kielille, jotka on sekoitettu englannin (en) kanssa. Tässä artikkelissa kuvailimme lähestymistapojamme POS-tunnistustekniikoihin, joita hyödynsimme tässä tehtävässä. Koneellista oppimista on käytetty sekakielisen tekstin POS-tunnistamiseen. POS-taggausta varten on kokeiltu sanojen hajautettuja esityksiä vektoriavaruudessa (word2vec) ominaisuuksien louhintaa varten ja log-lineaarisia malleja. Raportoimme työstämme kaikilla kolmella kielellä hi, bn ja te, jotka on sekoitettu en-kielen kanssa.

**Tulos**

Kokeet POS Tagging Code-sekoitetulla intialaisella sosiaalisen median tekstillä

**Esimerkki 2.1347**

Tilastotieteen ja koneoppimisen keskeisenä kysymyksenä on valita automaattisesti "oikea" mallin monimutkaisuus, esimerkiksi keskiarvoistettavien naapureiden määrä k lähimmän naapurin (kNN) regressiossa tai polynomiaste polynomeja sisältävässä regressiossa. Ehdotamme uutta periaatetta, Loss Rank Principle (LoRP), mallin valintaa varten regressiossa ja luokittelussa. Se perustuu häviöarvoon, joka laskee, kuinka moni muu (kuvitteellinen) aineisto sopisi paremmin. LoRP valitsee mallin, jonka tappioaste on minimaalinen. Toisin kuin useimmat rangaistujen maksimiluotettavuusmallien muunnokset (AIC, BIC, MDL), LoRP riippuu vain regressiofunktioista ja häviöfunktiosta. Se toimii ilman stokastista kohinamallia, ja sitä voidaan soveltaa suoraan mihin tahansa ei-parametriseen regressoriin, kuten kNN:ään.

**Tulos**

Mallien valinta häviöjärjestysperiaatteella

**Esimerkki 2.1348**

Suorassa kvantiiliregressiossa estimoidaan vastemuuttujan tiettyä kvantiilia syötemuuttujien funktiona. Esittelemme uuden kehyksen suoralle kvantiiliregressiolle, jossa opitaan Gaussin prosessimalli, joka minimoi odotetun kallistetun tappiofunktion. Oppimisessa tarvittava integrointi ei ole analyyttisesti käsiteltävissä, joten oppimisen nopeuttamiseksi käytämme odotusarvojen leviämisalgoritmia. Kuvaamme, miten tämä työ liittyy muihin kvantiiliregressiomenetelmiin, ja sovellamme menetelmää sekä synteettisiin että todellisiin aineistoihin. Menetelmän osoitetaan olevan kilpailukykyinen nykyisten menetelmien kanssa, mutta samalla se mahdollistaa Gaussin prosessin todennäköisyyskehyksen hyödyntämisen.

**Tulos**

Suora Gaussin prosessin kvantiiliregressio odotusarvojen etenemisen avulla

**Esimerkki 2.1349**

Tämä artikkeli on empiirinen tutkimus hajautetusta syväoppimisesta kysymykseen vastaamisen osatehtävää varten: vastauksen valinta. SGD-, MSGD-, DOWNPOUR- ja EASGD/EAMSGD-algoritmien vertailututkimukset on esitetty. Kokeelliset tulokset osoittavat, että viestinvälitysrajapintaan perustuva hajautettu kehys voi nopeuttaa konvergenssinopeutta epälineaarisessa mittakaavassa. Tämä artikkeli osoittaa hajautetun harjoittelun merkityksen: 120 työntekijällä voidaan saavuttaa 83-kertainen nopeus ja ajoaika lyhenee 107,9 tunnista 1,3 tuntiin, mikä hyödyttää tuottavuutta merkittävästi.

**Tulos**

Hajautettu syväoppiminen vastausten valintaa varten

**Esimerkki 2.1350**

Tässä artikkelissa esitellään käsitteellisiä suhteita tlrat syntlresizc utilitaristisia ja loogisia käsitteitä, jotka laajentavat [{escher. Määritellään ensin mahdollisten maailmojen tilassaL rajoituksista riippuvaisia mittareita, jotka kvantifioivat päättelyongelmien al tematiivisten ratkaisujen suhteellista laatua tai erilaisten toimintatapojen suhteellista toivottavuutta päätöksenteko- ja suunnitteluproblematiikoissa (coutrol, decision ja planning problerns). \Yc sho\\\' tbat. nämä llH� measures rnay interpreted totuusarvoina multivaluecl logiikka ami ehdottaa mekanismeja monimutkaisten rajoitusten esittämiseksi yksinkertaisempien rajoitusten yhdistelminä. Nämä laajennetut log iced-operaatiot mahdollistavat myös tavoitekohtaisten laatulukujen yhdistämisen ja Rggregoinnin. sureita globaaleiksi hyötyjen mittareiksi . Tunnistamme myös relaatioita, jotka edustavat differentiaalisia prd!'rcnces hdll "ecn ill t ernativ<' ratkaisuja ja rel<1tc thcrrr t.IJ<- pr< ' viously defined desir ability tne<IS\nes.

**Tulos**

TOTUUS HYÖDYLLISYYTENÄ: KÄSITTEELLINEN SYNTEESI

**Esimerkki 2.1351**

Huolimatta konfiguroitavien ohjelmistojärjestelmien valtavasta levinneisyydestä ja taloudellisesta merkityksestä, näiden järjestelmien koko potentiaalin hyödyntäminen suorituskykyoptimaalisten kokoonpanojen löytämiseksi ei ole saanut tyydyttävää tukea. Aikaisemmat työt ohjelmistokokoonpanojen suorituskyvyn ennustamiseksi kärsivät siitä, että joko a) vaadittiin aivan liian monta esimerkkikokoonpanoa tai b) ennusteissa oli suuria hajontoja. Molemmat ongelmat voidaan välttää käyttämällä WHAT-spektrioppijaa. WHATin innovaatio on konfiguroitavan ohjelmistojärjestelmän konfiguraatioiden välisen etäisyysmatriisin spektrin (ominaisarvojen) käyttäminen dimensioiden vähentämiseen. Tässä pienennetyssä konfiguraatioavaruudessa voidaan tutkia monia läheisesti toisiinsa liittyviä konfiguraatioita suorittamalla vain muutama esimerkkikonfiguraatio. Tässä tutkittujen järjestelmien osalta muutama kymmenen näytettä tuottaa tarkkoja ja vakaita ennusteita - alle 10 prosentin ennustevirhe ja alle 2 prosentin keskihajonta. Verrattuna nykyiseen tekniikkaan WHAT (a) tarvitsee 2-10 kertaa vähemmän näytteitä saavuttaakseen samanlaisen ennustustarkkuuden ja (b) sen ennusteet ovat vakaampia (eli niiden keskihajonta on pienempi). Lisäksi osoitamme, että optimoijat voivat käyttää WHAT:n tuottamia ennustemalleja löytääkseen järjestelmäkonfiguraatioita, jotka lähestyvät optimaalista suorituskykyä.

**Tulos**

Nopeampien järjestelmäkokoonpanojen nopeampi löytäminen spektrioppimisen avulla

**Esimerkki 2.1352**

Useat harvuusrajoitteiset algoritmit, kuten Orthogonal Matching Pursuit tai Frank-Wolfe-algoritmi harvuusrajoitteineen, toimivat valitsemalla iteratiivisesti uuden atomin lisättäväksi nykyiseen nollasta poikkeavaan muuttujajoukkoon. Tämä valintavaihe suoritetaan yleensä laskemalla gradientti ja etsimällä sitten gradientin komponentti, jonka absoluuttinen arvo on suurin. Tämä vaihe voi olla laskennallisesti kallis erityisesti laajamittaisissa ja korkea-ulotteisissa aineistoissa. Tässä työssä pyrimme nopeuttamaan näitä harvuusrajoitteisia optimointialgoritmeja hyödyntämällä keskeistä havaintoa, jonka mukaan näiden algoritmien toimimiseen tarvitaan vain gradientin ylimmän tulon koordinaatti. Esittelemme siis ahneisiin menetelmiin ja satunnaistamismenetelmiin perustuvia algoritmeja, joilla pyritään arvioimaan gradientti ja sen ylin tulo edullisesti. Toinen panoksemme on se, että parhaan gradienttimerkinnän löytämisen ongelma esitetään monikätisen bandit-ongelman parhaan käsivarren tunnistamisena. Tämän uudenlaisen oivalluksen ansiosta pystymme tarjoamaan bandiittipohjaisen algoritmin, joka estimoi suoraan huippuarvot erittäin tehokkaasti. Lisäksi esitetään teoreettisia havaintoja, joiden mukaan tuloksena saadut epätarkat FrankWolfe- tai Orthogonal Matching Pursuit -algoritmit toimivat suurella todennäköisyydellä samalla tavalla kuin niiden tarkat versiot. Olemme tehneet useita kokeita, jotka osoittavat, että ehdottamillamme ahneilla deterministisillä ja bandit-menetelmillä voidaan saavuttaa kertaluokkaa suurempi nopeus ja samalla yhtä tehokas kuin tarkka gradientti, kun niitä käytetään OMP:n, Frank-Wolfen tai CoSaMP:n kaltaisissa algoritmeissa.

**Tulos**

Ahneet menetelmät, satunnaistamislähestymistavat ja monihaaraiset bandit-algoritmit tehokkaaseen harvuusrajoitteiseen optimointiin

**Esimerkki 2.1353**

Tarkastelemme kolmea eri algoritmia, joiden avulla [1]:n törmäystodistusmenetelmää voidaan käyttää keskitetyn monirobottiryhmän tapauksessa. Hyödyntämällä symmetriaa monirobottiryhmien konfiguraatioavaruudessa menetelmämme voivat merkittävästi vähentää törmäystarkastusten määrää verrattuna sekä [1] että tavallisiin törmäystarkastustoteutuksiin.

**Tulos**

Nopea törmäystarkastelu: Yksittäisistä roboteista monirobottiryhmiin

**Esimerkki 2.1354**

Ennustava koodaus (PDC) on viime aikoina herättänyt huomiota neurotieteiden ja tietojenkäsittelyn piirissä ehdokkaana yhdistäväksi paradigmaksi neuronaalisiin tutkimuksiin ja keinotekoisten neuroverkkojen toteutuksiin, jotka on suunnattu erityisesti valvomattomiin oppimisjärjestelmiin. Mismatch-negatiivisuutta (Mismatch Negativity, MMN) on myös hiljattain tutkittu suhteessa PC:hen, ja sen on todettu olevan hyödyllinen ainesosa neuraalisissa ennakoivissa koodausjärjestelmissä. Elävien organismien käyttäytymisen perusteella tällaiset verkot ovat erityisen käyttökelpoisia muodostettaessa spatio- ja temporaalisia siirtymiä ja muuttumattomia representaatioita syötemaailmasta. Useimmissa neurojärjestelmissä ei kuitenkaan edelleenkään oteta huomioon suurta määrää synapseja, vaikka muutamat koneoppimisen tutkijat ovat osoittaneet sen olevan tehokas ja erittäin tärkeä osa mitä tahansa neurojärjestelmää, jos tällaisen järjestelmän halutaan käyttäytyvän oikein. Tärkein pointtimme tässä on, että PDC-järjestelmät, joissa on MMN-ilmiö suuren synapsien määrän lisäksi, voivat parantaa huomattavasti minkä tahansa neuraalisen oppimisjärjestelmän suorituskykyä ja kykyä tehdä päätöksiä konemaailmassa. Tässä asiakirjassa ehdotamme uudenlaista biomimeettistä laskennallisen älykkyyden algoritmia - Deviant Learning Algorithm - joka on saanut innoituksensa näistä viimeaikaisten aivokognitiivisten löydösten ja teorioiden keskeisistä ajatuksista ja toiminnallisista ominaisuuksista. Osoitamme myös teoreettisten oivallusten ohjaamilla numeerisilla kokeilla, miten keksimämme biomimeettinen algoritmi voi saavuttaa kilpailukykyisiä ennusteita jopa hyvin pienillä ongelmakohtaisilla tiedoilla.

**Tulos**

Poikkeava oppimisalgoritmi: Harvojen epäsuhtaisten representaatioiden oppiminen ajassa ja avaruudessa

**Esimerkki 2.1355**

Vaikka monissa koneoppimisalgoritmeissa opitaan osa-avaruuksia, joilla on tiettyjä ominaisuuksia, voi olla haastavaa optimoida parametrimatriisia, joka on rajoitettu edustamaan osa-avaruutta. Yksi ratkaisu on käyttää Riemannin optimointimenetelmiä, jotka pakottavat tällaiset rajoitukset implisiittisesti hyödyntämällä sitä, että toteutettavissa olevat parametriarvot muodostavat moninaisuuden. Vaikka Riemannin menetelmiä on olemassa joihinkin erityisongelmiin, kuten yhden aliavaruuden oppimiseen, on olemassa yleisempiä aliavaruusrajoituksia, jotka tarjoavat lisää joustavuutta optimointiongelman asettamisessa, mutta joita ei ole muotoiltu moninaisuutena. Ehdotamme partitioidun aliavaruuden (Partitioned Subspace, PS) moninaisuutta sellaisten matriisien optimoimiseksi, jotka on rajoitettu edustamaan yhtä tai useampaa aliavaruutta. Jokainen piste moninaisuudella määrittelee syöttöavaruuden jakamisen toisiinsa nähden ortogonaalisiin aliavaruuksiin, joissa jakojen lukumäärä ja koko on käyttäjän määriteltävissä. Tämän seurauksena voidaan oppia erilaisia ominaisuuksien ryhmiä määrittelemällä kullekin osiolle erilaiset tavoitefunktiot. Havainnollistamme moninaisuuden ominaisuuksia kokeilemalla useiden tietokokonaisuuksien analysointia ja toimialueen mukauttamista.

**Tulos**

Monimuotoinen lähestymistapa vastavuoroisesti ortogonaalisten aliavaruuksien oppimiseen

**Esimerkki 2.1356**

Suurten tai jatkuvien tila-avaruuksien hallitseminen on ollut vahvistusoppimisen pitkäaikainen haaste. Ajallinen abstraktio on jossain määrin tehnyt tämän mahdolliseksi, mutta tehokas suunnittelu ajallisen abstraktion avulla on edelleen ongelma. Lisäksi tilallisten abstraktioiden käyttäminen eri tilanteita koskevien toimintatapojen oppimiseen samanaikaisesti, kun käytetään ajallisia abstraktiomalleja, on avoin ongelma. Ehdotamme tässä tehokasta algoritmia, joka on konvergentti lineaarisen funktion approksimaation avulla, kun suunnittelussa käytetään ajallisesti abstrakteja toimintoja. Näytämme, miten tätä algoritmia voidaan käyttää yhdessä satunnaisesti generoitujen vaihtoehtomallien kanssa useilla aikaskaaloilla suunnittelemaan agentteja, joiden on toimittava reaaliaikaisesti. Näiden satunnaisesti luotujen vaihtoehtomallien käytön useilla aikaskaaloilla osoitetaan vähentävän annetun tehtävän ratkaisemiseen tarvittavien päätösepookkien määrää, mikä vähentää tehokkaasti harkintaan tarvittavaa aikaa. ar X iv :1 70 3. 06 47 1v 1 [ cs .A I] 1 9 M ar 2 01 7

**Tulos**

Multi-Timescale, Gradient Descent, Temporal Difference Learning lineaarisilla vaihtoehdoilla.

**Esimerkki 2.1357**

Automaattinen puheentunnistus (ASR) mahdollistaa luonnollisen ja intuitiivisen käyttöliittymän lapsille tarkoitettuihin robottiopetussovelluksiin. On kuitenkin ratkaistava useita haasteita, jotta tällainen käyttöliittymä toimisi vakaasti realistisissa ympäristöissä, kuten lasten puheen tunnistamisen luontaiset vaikeudet ja luokkahuoneissa usein esiintyvä suuri taustamelu. Osana EU:n EASEL-hanketta olemme antaneet useita panoksia näiden haasteiden ratkaisemiseksi ja toteuttaneet oman ASR-moduulin käytettäväksi robotiikkasovelluksissa. Käytimme viimeisimpiä syviä neuroverkkoalgoritmeja, joiden suorituskyky on huomattavasti parempi kuin perinteisen GMM-menetelmän, ja käytimme tietojen lisäämiseen tarkoitettuja menetelmiä, joilla parannetaan häiriön ja puhujan vaihtelun kestävyyttä. ASR-moduulin ja muun dialogijärjestelmän välillä on tiivis integraatio, jonka ansiosta ASR voi vastaanottaa reaaliaikaisesti dialogin nykyiseen osaan liittyvät kielimallit, mikä parantaa tarkkuutta huomattavasti. Integroimme ASR-moduulin interaktiiviseen multimodaaliseen järjestelmään, jossa käytetään pientä humanoidirobottia ja jonka avulla lapset voivat oppia liikunnasta ja energiasta. Järjestelmä asennettiin julkiseen museotapahtumaan osana tutkimustutkimusta, jossa 320 lasta (3-14-vuotiaita) oli vuorovaikutuksessa robotin kanssa, ja ASR-järjestelmämme saavutti 90 prosentin tarkkuuden sujuvan ja lähes sujuvan puheen osalta.

**Tulos**

Lasten puheen automaattinen tunnistaminen meluisissa ympäristöissä käytettäviä robottisovelluksia varten

**Esimerkki 2.1358**

Esittelemme simplify-defun-työkalun, joka muuttaa tietyn funktion määritelmän uuden funktion yksinkertaistetuksi määritelmäksi ja tarjoaa ACL2:lla tarkistetun todisteen siitä, että vanha ja uusi funktio ovat ekvivalentteja. Tarvittaessa se tuottaa myös terminointi- ja vartijatodistukset uudelle funktiolle. Selitämme, miten työkalu on suunniteltu niin, että nämä todistukset onnistuvat. Esimerkit havainnollistavat sen hyödyllisyyttä erityisesti ohjelmien muuntamisessa synteesissä ja verifioinnissa.

**Tulos**

Monipuolinen, järkevä työkalu määritelmien yksinkertaistamiseen

**Esimerkki 2.1359**

Tässä artikkelissa tutkimme uskomusten päivitysoperaattoreiden luokkaa, jossa operaattorin määritelmä on kompositionaalinen lisättävän lauseen suhteen. Tavoitteena on tarjota päivitysoperaattori, joka on intuitiivinen, koska sen määritelmä perustuu päivityslauseen rakenteen rekursiiviseen hajotukseen, ja joka voidaan toteuttaa järkevästi. Päivitystä käsiteltäessä esitetään ensin määritelmä, joka on ilmaistu tietopohjan mallien avulla. Vaikka tämä operaattori tyydyttää keskeisen osan Katsuno-Mendelzonin päivityspostulaateista, kaikki postulaatit eivät täyty. Muut Katsuno-Mendelzonin postulaatit voidaan saavuttaa rajoittamalla sopivasti päivitettävän lauseen syntaktista muotoa, kuten osoitamme. Rajoittamalla päivityslauseen syntaktista muotoa saamme myös päivitysoperaattoreiden hierarkian, jossa Winslettin vakiosemantiikka on mielenkiintoisin peruslähestymistapa. Yleisessä tapauksessa algoritmi on eksponentiaalinen, mutta joidenkin kohtuullisten oletusten avulla saadaan algoritmi, joka on lineaarinen tietopohjan koon suhteen. Näin ollen tuloksena syntyvällä lähestymistavalla on joissakin tilanteissa paljon paremmat monimutkaisuusominaisuudet kuin muilla operaattoreilla. Tutkimme myös muita kompositionaalisia uskomusmuutosoperaattoreita: pyyhkiminen kehitetään päivityksen kaksoisoperaattoriksi, osoitamme, että unohdusoperaattori on määriteltävissä päivityksen avulla, ja annamme määritelmän kompositionaaliselle tarkistusoperaattorille. Saamme selville, että luonnollisimman määritelmän mukaan kompositionaalinen tarkistus tuottaa Satohin tarkistusoperaattorin.

**Tulos**

Kokoonpanon uskomusten päivitys

**Esimerkki 2.1360**

Tässä artikkelissa ehdotamme uutta neurologista lähestymistapaa parafraasien tuottamiseen. Perinteiset parafraasien generointimenetelmät hyödyntävät joko käsin kirjoitettuja sääntöjä ja sanastoon perustuvia linjauksia tai käyttävät tilastollisen koneoppimisen periaatteita. Tietojemme mukaan tämä työ on ensimmäinen, jossa tutkitaan syväoppimismalleja parafraasien generointiin. Ensisijainen panoksemme on pinottu jäännös-LSTM-verkko, jossa lisäämme jäännösyhteyksiä LSTM-kerrosten välille. Tämä mahdollistaa syvien LSTM-verkkojen tehokkaan harjoittelun. Kokeilemme malliamme ja muita huipputason syväoppimismalleja kolmella eri tietokokonaisuudella: PPDB, WikiAnswers ja MSCOCO. Arviointitulokset osoittavat, että mallimme päihittää sekvenssistä sekvenssiin -, huomiopohjaiset ja kaksisuuntaiset LSTM-mallit BLEU-, METEOR- ja TER-mittauksissa sekä sulautumiseen perustuvassa lauseiden samankaltaisuusmittarissa.

**Tulos**

Neuraalinen parafraasien tuottaminen pinotuilla jäännös-LSTM-verkoilla

**Esimerkki 2.1361**

Tässä artikkelissa ehdotetaan uutta menetelmää K-armed dueling bandit -ongelmaan, joka on muunnelma tavallisesta K-armed bandit -ongelmasta, joka tarjoaa vain suhteellista palautetta asepareista. Lähestymistapamme laajentaa Upper Confidence Bound -algoritmin suhteelliseen asetelmaan käyttämällä pareittaisten todennäköisyyksien estimaatteja lupaavan käsivarren valitsemiseen ja soveltamalla Upper Confidence Bound -algoritmia voittajan kanssa vertailukohtana. Todistamme, että katumuksen rajoitus on äärellisen pitkäaikainen ja on luokkaa O(log t). Lisäksi empiiriset tuloksemme, joissa käytämme todellisia tietoja tiedonhakusovelluksesta, osoittavat, että se on huomattavasti parempi kuin nykyinen tekniikka.

**Tulos**

Suhteellinen ylempi luottamusraja K-Armed Dueling Bandit -ongelmassa

**Esimerkki 2.1362**

Complete Tang Poems (CTP) on tärkein lähde Tang-runojen tutkimiseen. Tarkastelemme CTP:tä laskennallisin työkaluin tietyistä kielitieteellisistä näkökulmista, kuten distributiivisesta semantiikasta ja kollokaatioanalyysistä. Tällaisista kvantitatiivisista näkökulmista vertailemme "tuulen" ja "kuun" käyttöä Li Bain (李白) ja Du Fun (杜甫) runoissa. Värit toimivat runoissa kuten äänet elokuvissa, ja niillä on ratkaiseva rooli runojen mielikuvissa. Näin ollen tutkitaan värejä kuvaavia sanoja, ja "白" (bai2, valkoinen) on pääpaino, koska se on yleisin väri CTP:ssä. Tutkimme myös joitakin tapauksia, joissa värillisiä sanoja on käytetty vastakohtapareissa (對仗), jotka olivat keskeisiä runojen mielikuvien edistämisessä. CTP sisältää myös hyödyllistä historiallista tietoa, ja poimimme CTP:stä henkilöiden nimiä tutkiaksemme Tang-runoilijoiden sosiaalisia verkostoja. Tällaiset tiedot voidaan sitten yhdistää Harvardin yliopiston China Biographical Database -tietokantaan.

**Tulos**

Värien estetiikka ja sosiaaliset verkostot täydellisissä Tang-runoissa: Tang Tang: Explorations and Discoveries

**Esimerkki 2.1363**

Harva koodaus on yleinen tapa oppia paikallisia piirteitä objektien tunnistusta varten. Viime aikoina on alettu kiinnostua yhä enemmän ominaisuuksien oppimisesta spatio-ajallisesta, binokulaarisesta tai muusta monihavaintoaineistosta, jossa tavoitteena on koodata pikemminkin kuvien välisiä suhteita kuin yksittäisen kuvan sisältöä. Esitämme analyysin monikatselukuvaominaisuuksien oppimisesta, joka osoittaa, että piilomuuttujat koodaavat muunnoksia havaitsemalla kiertokulmia useiden kuvien poimujen jakamissa ominaistiloissa. Analyysimme auttaa selittämään viimeaikaisia kokeellisia tuloksia, jotka osoittavat, että muunnoskohtaisia piirteitä syntyy, kun monimutkaisia solumalleja koulutetaan videoilla. Analyysimme osoittaa myös, että muunnosvariantit piirteet voivat syntyä muunnosten representaatioiden oppimisen sivutuotteena.

**Tulos**

Monen näkymän ominaisuuksien oppiminen

**Esimerkki 2.1364**

Tässä asiakirjassa kuvataan Odin-kehys, joka on toimialariippumaton alusta sääntöpohjaisten tapahtumien louhintamallien kehittämiseen. Odinin tavoitteena on olla tehokas (sääntökieli mahdollistaa monimutkaisten syntaktisten rakenteiden mallintamisen) ja vankka (syntaktisista jäsennysvirheistä toipumiseksi syntaktisia malleja voidaan vapaasti sekoittaa pintapuolisiin, merkkipohjaisiin malleihin), mutta samalla se on yksinkertainen (jotkin verkkotunnuskieliopit voidaan ottaa käyttöön muutamassa minuutissa) ja nopea (Odin käsittelee yli 100 lausetta sekunnissa reaalimaailman verkkotunnuksessa, jossa on yli 200 sääntöä). Tässä on Odin-sääntökielen perusteellinen määritelmä sekä kuvaus Odinin Scala-rajapinnasta, jonka avulla näitä sääntöjä voidaan soveltaa mielivaltaisiin teksteihin. 1 ar X iv :1 50 9. 07 51 3v 1 [ cs .C L ] 2 4 Se p 20 15

**Tulos**

Odin Event Extraction Frameworkin ja sääntökielen kuvaus

**Esimerkki 2.1365**

Laajennamme aiempaa työtä lineaaristen mallien tehokkaan harjoittelun parissa soveltamalla stokastisia päivityksiä vain nollasta poikkeaviin ominaisuuksiin ja tuomalla painot tarpeen mukaan ajan tasalle. Tähän mennessä on kuvattu vain suljetun muodon päivitykset l1-, l∞- ja harvoin käytetylle l2-normille. Laajennamme tätä työtä osoittamalla oikeat suljetun muodon päivitykset suosituille l22- ja elastisen verkon regularisoidulle mallille. Esitämme dynaamisen ohjelmointialgoritmin, jolla lasketaan oikea elastisen verkon päivitys, jossa on vain yksi vakioaikainen aliongelman laskenta päivitystä kohti. Algoritmimme käsittelee sekä kiinteitä että laskevia oppimisnopeuksia, ja johdamme tuloksen sekä stokastiselle gradienttilaskeutumiselle (SGD) että eteenpäin-taaksepäin jakamiselle (FoBoS). Validoimme algoritmin empiirisesti osoittamalla, että sanasäkkitietokokonaisuudessa, jossa on 260 941 piirrettä ja keskimäärin 88 nollasta poikkeavaa piirrettä esimerkkiä kohti, menetelmämme kouluttaa logistisen regressioluokittimen elastisella verkon regularisoinnilla 612 kertaa nopeammin kuin muuten identtinen toteutus tiheillä päivityksillä.

**Tulos**

Tehokas elastinen verkkosäännöstely harvoille lineaarisille malleille

**Esimerkki 2.1366**

Kysymyksiin vastaaminen (QA) on ollut viime vuosina nousussa. Tämä on johtanut siihen, että sekä yritykset että tutkimuslaitokset ovat kehittäneet lukuisia kyselyvastauksen (QA) järjestelmiä. Vaikka joitakin laadunvarmistusjärjestelmien komponentteja käytetään uudelleen eri toteutuksissa, useimmissa järjestelmissä ei hyödynnetä komponenttien uudelleenkäytön kaikkia mahdollisuuksia. Näin ollen laadunvarmistusjärjestelmien kehittäminen on tällä hetkellä edelleen työläs ja aikaa vievä prosessi. Me vastaamme haasteeseen nopeuttaa uusien tai räätälöityjen laadunvarmistusjärjestelmien luomista esittelemällä konseptin, joka perustuu itseohjautuvaan lähestymistapaan laadunvarmistusjärjestelmien kokoamiseen. Lähestymistapamme mahdollistaa olemassa olevien, verkkopohjaisten laadunvarmistusjärjestelmien tai -moduulien uudelleenkäytön samalla kun kehitetään uusia laadunvarmistusalustoja. Tätä varten se perustuu siihen, että laadunvarmistusmoduulit kuvataan web-ontologiakielellä. Näiden kuvausten perusteella lähestymistapamme pystyy automaattisesti koostamaan laadunvarmistusjärjestelmiä käyttämällä tietoon perustuvaa lähestymistapaa.

**Tulos**

Itseohjautuvat kysymysten vastausjärjestelmät

**Esimerkki 2.1367**

Jokaisen CI-lausekeparin (CI=conditional indepen dence) semigrafinen sulkeuma on stokastinen CI-malli. Tämän tuloksen seurauksena osoitetaan, että kaikki CI-mallien todennäköisyysajattelun kannalta järkevät päättelysäännöt, joilla on korkeintaan kaksi etumerkkiä, ovat johdettavissa semigrafisista päättelysäännöistä. Tämä oikeuttaa semigrafoidien käytön stokastisten CI-mallien approksimaatioina probabilistisessa päättelyssä. Sivutuotteena annetaan luettelo kaikista 19:stä mainitun semigrafoidisen sulkeuman potentiaalisesta dominoivasta elementistä.

**Tulos**

Puoligrafoidit ovat stokastisten ehdollisen riippumattomuuden mallien kaksi-antiedeentaalisia approksimaatioita.

**Esimerkki 2.1368**

Tutkimme luokittelijoiden oppimisen ongelmaa oikeudenmukaisuusrajoituksen kanssa ja pyrimme kolmella pääasiallisella panoksella määrittelemään ongelman luontaiset kompromissit. Ensinnäkin suhteutamme kaksi olemassa olevaa oikeudenmukaisuusmittaa kustannusherkkiin riskeihin. Toiseksi osoitamme, että kustannustietoiselle luokittelulle ja oikeudenmukaisuusmittauksille optimaalinen luokittelija on luokan todennäköisyysfunktion instanssista riippuvainen kynnysarvo. Kolmanneksi osoitamme, miten tarkkuuden ja oikeudenmukaisuuden välinen kompromissi määräytyy kohteen ja herkkien ominaisuuksien luokkatodennäköisyyksien välisen linjauksen perusteella. Analyysimme perustana on yleinen kehys, jonka mukaan oikeudenmukaisuuden vaatimuksen mukaisen oppimisen ongelmana on kahden tilastollisen riskin erotuksen minimointi.

**Tulos**

Oikeudenmukaisuuden kustannukset luokittelussa

**Esimerkki 2.1369**

Kun klassista korkeamman kertaluvun logiikkaa käytetään metalogiikkana, johon voidaan pinnallisesti upottaa erilaisia muita (klassisia ja ei-klassisia) logiikoita, se soveltuu hyvin universaalin loogisen päättelytavan toteuttamiseen. Universaaliloginen päättely puolestaan, kuten jo Leibniz visioi, voi tukea koneiden rationaalisten argumenttien tiukkaa formalisointia ja syvällistä loogista analyysia. Seuraavassa kuvataan vastaava universaalilogiikan päättelykehys ja keskustellaan useista esimerkkisovelluksista. Tulevaisuudessa universaaliloginen päättely yhdistettynä rationaalisen argumentoinnin sopiviin, hallittuihin muotoihin voi toimia kommunikaatiokerroksena ihmisten ja älykkäiden koneiden välillä. 1 Rationaalinen argumentointi - ihmisen ja koneiden välinen kommunikaatiorajapinta Pyrkimyksellä ymmärtää, mallintaa ja toteuttaa rationaalista argumentointia ja universaalia loogista päättelyä ihmisaivoista riippumatta on pitkät perinteet ihmiskunnan historiassa. Se ulottuu ainakin Aristotelesin merkittävään syllogististen argumenttien tutkimukseen. Nykyään, kun älykäs tietotekniikka on lisääntynyt, kysymys on ajankohtaisempi kuin koskaan: jos ihmisten ja älykkäiden koneiden on tarkoitus elää sovinnollisesti rinnakkain, olla vuorovaikutuksessa ja tehdä yhteistyötä, tarvitaan asianmukaisia viestintätapoja niiden välille. Koneiden pitäisi esimerkiksi pystyä kuvaamaan, arvioimaan ja puolustamaan toimintaansa ja päätöksiään (toimintavaihtoehtojaan) sellaisessa muodossa, että ne ovat ihmisen ymmärrettävissä ja arvioitavissa. Tämä on ratkaisevan tärkeää, jotta älykkäät koneet voidaan sovittaa yhteen ja integroida sosiaalisesti hyväksyttävästi jokapäiväiseen (ihmis)elämään. Koneiden ja ihmisten välisten viestintävälineiden olisi mieluiten perustuttava inhimilliseen, rationaaliseen argumentointiin, joka on iät ja ajat ollut yhteiskunnallisten, oikeudellisten ja tieteellisten prosessiemme perusta. Tekoälyn nykyinen kehitys sen sijaan keskittyy vahvasti tilastolliseen tietoon, koneoppimiseen ja subsymbolisiin esityksiin, jotka kaikki ovat melko kaukana ihmisen rationaalisesta selittämisestä, ymmärtämisestä ja arvostelukyvystä. Haasteena on siis täydentää ja parantaa näitä nykyisissä tekoälyjärjestelmissä käytettäviä ihmiselle epäystävällisiä päättely- ja tiedon esittämistapoja sopivilla, ihmisen kognition kannalta sopivilla selityksillä eli rationaalisilla argumenteilla. Rationaalisten perustelujen vaihdon avulla ihmisen ja älykkäiden koneiden välille voidaan lopulta taata paljon kaivattu keskinäinen ymmärrys ja hyväksyntä ihmisten ja älykkäiden koneiden välillä. Tämä on erityisen tärkeää arvioitaessa koneiden toimintaa ihmisten esittämien oikeudellisten, eettisten, moraalisten, sosiaalisten ja kulttuuristen normien kannalta. Mutta millaisia formalismeja on käytettävissä, jotka voisivat toimia yleisimpänä perustana ihmisen tason rationaalisten argumenttien mallintamiselle koneissa?

**Tulos**

Universaali päättely, rationaalinen argumentointi ja ihmisen ja koneen vuorovaikutus

**Esimerkki 2.1370**

Normalisoidut satunnaismitat (NRM) muodostavat laajan luokan diskreettejä satunnaismittoja, joita käytetään usein Bayesin ei-parametristen mallien prioreina. Dirichlet-prosessi on tunnettu esimerkki NRM:stä. Useimmat NRM-seosmallien posterioriset päättelymenetelmät perustuvat MCMC-menetelmiin, koska ne on helppo toteuttaa ja niiden konvergenssi on hyvin tutkittu. MCMC kärsii kuitenkin usein hitaasta konvergenssista, kun hyväksymisaste on alhainen. Puupohjainen päättely on vaihtoehtoinen deterministinen posteriorinen päättelymenetelmä, jossa Bayesin hierarkkinen klusterointi (Bayesian hierarchical clustering, BHC) tai inkrementaalinen Bayesin hierarkkinen klusterointi (incremental Bayesian hierarchical clustering, IBHC) on kehitetty DP- tai NRM-seosmalleille (NRMM). Vaikka IBHC on lupaava menetelmä NRMM-mallien posterioriseen päättelyyn sen tehokkuuden ja sovellettavuuden vuoksi online päättelyyn, sen konvergenssi ei ole taattu, koska se käyttää heuristiikkaa, joka yksinkertaisesti valitsee parhaan ratkaisun useiden kokeilujen jälkeen. Tässä artikkelissa esitellään NRMM-mallien hybridinen päättelyalgoritmi, jossa yhdistyvät sekä MCMC:n että IBHC:n edut. IBHC:n rakentamat puut hahmottelevat datan osioita, jotka ohjaavat Metropolis-Hastings-menettelyä käyttämään sopivia ehdotuksia. MCMC:n luonteen periytyvä puuohjattu MCMC (tgMCMC) konvergoi taatusti, ja sen konvergenssi on nopeaa puiden ohjaamien tehokkaiden ehdotusten ansiosta. Kokeet sekä synteettisillä että reaalimaailman aineistoilla osoittavat menetelmämme edut.

**Tulos**

Puu-ohjattu MCMC-päätelmä normalisoiduille satunnaismittaisille sekoitusmalleille (Tree-Guided MCMC Inference for Normalized Random Measure Mixture Models)

**Esimerkki 2.1371**

Kilpailukykyisimpien CDCL (Conflict-Driven Clause Learning) SAT-ratkaisujen käyttämä päätösmuuttujan valintapolitiikka on joko VSIDS (Variable State Independent Decaying Sum) tai sen muunnokset, kuten eksponentiaalinen versio EVSIDS. VSIDS:lle ja sen muunnoksille on yhteistä se, että ratkaisuprosessissa hyödynnetään tilastotietoa, mutta ongelman rakennetieto jätetään huomiotta. Tästä syystä tässä asiakirjassa muutetaan päätöksentekomuuttujien valintakäytäntöä ja esitellään BCD:hen (Blocked Clause Decomposition) perustuva SAT-ratkaisutekniikka. Sen perusajatuksena on, että osa päätöksentekomuuttujista valitaan VSIDS-heuristiikalla ja osa päätöksentekomuuttujista valitaan BCD:n avulla saatujen lohkottujen joukkojen avulla. Verrattuna nykyiseen BCD-pohjaiseen tekniikkaan tekniikkamme on yksinkertainen, eikä sen tarvitse koodata CNF-kaavoja uudelleen. Sertifioitua UNSAT-ratkaisua varten tarkoitetut SAT-ratkaisijat voivat soveltaa myös BCD-pohjaista tekniikkaamme. Sovelluskokeilumme osoittavat, että uusi BCD:hen perustuva muuttujien valintakäytäntö voi parantaa SAT-ratkaisijoiden, kuten abcdSAT:n, suorituskykyä. BCD:tä käyttävä ratkaisija ratkaisi SAT Race 2015 -kilpailun instanssin, jota yksikään ratkaisija ei ollut tähän mennessä ratkaissut. Tämä osoittaa, että joissakin tapauksissa rakennetietoon perustuva heuristiikka on tehokkaampi kuin tilastotietoon perustuva heuristiikka.

**Tulos**

SAT-ratkaisijoiden parantaminen lohkottujen lausekkeiden purkamisen avulla

**Esimerkki 2.1372**

Tässä asiakirjassa esitellään alustava työ lähestymistavasta, jolla pyritään täyttämään logiikkaan perustuvan argumentoinnin ja abstraktien argumentointikehysten dynamiikkaa käsittelevien lukuisten lähestymistapojen välinen aukko. Ajatuksemme on, että vaikka argumentit ja hyökkäykset määritellään loogisen uskomuspohjan avulla, voi olla epävarmuutta siitä, kuinka tarkka argumentin sisältö on, ja siten myös sitä koskevien hyökkäysten olemassaolosta (tai puuttumisesta). Käytämme enthymemejä havainnollistamaan tätä argumenttien ja hyökkäysten epävarmuuden käsitettä. Kuten kirjallisuudessa väitetään, todelliset argumentit ovat usein enthymemejä eivätkä täysin määriteltyjä deduktiivisia argumentteja. Tämä tarkoittaa sitä, että jotkut osat parista (tuki, väite) voivat puuttua, koska niiden oletetaan kuuluvan "yleiseen tietoon", jolloin enthymemin vastaanottavan agentin pitäisi päätellä ne. Mutta agenttien käsitys yhteisestä tiedosta voi olla väärä, ja silloin ensimmäinen agentti voi esittää enthymemin, jota hänen vastustajansa ei pysty purkamaan täsmällisesti. On todennäköistä, että agentin tekemä enthymemin dekoodaus johtaa virheellisiin hyökkäyksiin tämän uuden argumentin ja olemassa olevien argumenttien välillä. Tällöin agentti voi saada tietoa hyökkäyksistä tai argumenttien hyväksymisstatuksista, jotka ovat ristiriidassa hänen argumentointikehyksensä kanssa. Esittelemme esimerkin tavasta, jolla tämä uusi tieto voidaan sisällyttää abstraktien argumentaatiokehysten dynamiikkaa käsittelevien nykyisten teosten avulla.

**Tulos**

Loogisen argumentaation ja abstraktien argumentaatiokehysten tarkistamisen välisen kuilun täyttäminen enthymemien avulla

**Esimerkki 2.1373**

Nimettyjen entiteettien tunnistusmallit (NER) kielelle L koulutetaan tyypillisesti käyttäen kyseisen kielen kommentoitua dataa. Tutkimme monikielistä NER-mallia, jossa L-kielinen NER-malli koulutetaan lähdekielellä (tai useammalla lähdekielellä). Esittelemme kielestä riippumattoman NER-menetelmän, joka perustuu monikieliseen wikifiointiin, tekniikkaan, jossa ei-englanninkielisen tekstin sanat ja lauseet perustetaan englanninkielisiin Wikipedia-tietueisiin. Näin minkä tahansa kielen tekstissä olevat maininnat voidaan kuvata käyttämällä kategorioita ja FreeBase-tyyppejä, jolloin saadaan, kuten osoitamme, vahvoja kielestä riippumattomia piirteitä. Tämän näkemyksen perusteella ehdotamme NER-mallia, jota voidaan soveltaa kaikkiin Wikipedian kieliin. Kun mallimme on koulutettu englanninkielellä, se päihittää vertailukelpoiset lähestymistavat CoNLL:n vakiomuotoisissa tietokannoissa (espanja, saksa ja hollanti), ja se toimii erittäin hyvin myös matalaresurssisilla kielillä (turkki, tagalog, yoruba, bengali ja tamil), joiden Wikipedia on huomattavasti pienempi. Lisäksi menetelmiemme avulla voimme harjoitella useilla lähdekielillä, mikä yleensä parantaa NER-tuloksia kohdekielillä. Lopuksi osoitamme, että kieliriippumattomia piirteitämme voidaan käyttää myös yksikielisten NER-järjestelmien parantamiseen, jolloin tulokset paranevat kaikilla yhdeksällä kielellä.

**Tulos**

Monikielinen nimettyjen entiteettien tunnistus Wikificationin avulla

**Esimerkki 2.1374**

Sekvenssimallin oppimisalgoritmit maksimoivat yleensä log-likelihoodin miinus mallin normi (tai minimoivat Hammingin häviön + normi). POS-merkinnöissä (part-ofspeech tagging) kohdekielen harjoitusaineistomme koostuu lauseiden sarjoista, joissa on sanakohtaiset merkinnät, jotka on projisoitu k kielen käännöksistä, joista meillä on merkittyjä tietoja, sanojen kohdistusten avulla. Koulutusdatamme on siis hyvin kohinaista, ja jos Rademacherin monimutkaisuus on suuri, oppimisalgoritmit ovat alttiita ylisovittamiselle. Normipohjaisessa regularisoinnissa oletetaan, että priorin leveys on vakio ja keskiarvo nolla. Sen sijaan ehdotamme, että käytämme k lähdekielen mallia estimoida Gaussin priorin parametrit uusien POS-taggerien oppimista varten. Tämä johtaa huomattavasti parempaan suorituskykyyn monilähteisissä siirtoasetelmissa. Esittelemme myös drop-out-version, jossa online-oppimisen aikana syötetään (empiiristä) Gaussin kohinaa. Lopuksi toteamme, että empiiristen Gaussin prioreiden käyttö johtaa paljon pienempään Rademacher-kompleksisuuteen ja on parempi kuin optimaalisesti painotettu mallien interpolointi. 1 Sekvenssimallien kieltenvälinen siirto-oppiminen Maailman ihmiset puhuvat noin 6900 eri kieltä. Avoimen lähdekoodin luonnollisen kielen prosessoinnin (NLP) työkalupakit, kuten OpenNLP1 ja CoreNLP2 , kattavat vain 6-7 kieltä, ja meillä on riittävästi merkittyä harjoitusdataa noin 20-30 kielen mallien tuottamiseen. Toisin sanoen valvotut sekvenssioppimisalgoritmit eivät riitä POS-mallien tuottamiseen kuin pienelle vähemmistölle maailman kielistä. Mitä voimme tehdä kaikille niille kielille, joista ei ole saatavilla koulutusdataa? Valvomattomilla POS-induktioalgoritmeilla on metodologisia ongelmia (otoksen sisäinen arviointi, yhteisön laajuinen hyperparametrien viritys jne.), ja suorituskyky on esteenä jatkokäyttösovelluksille. Joissakin valvomattoman POS-tunnisteiden induktioon liittyvissä töissä on oletettu muita resursseja, kuten tunnistesanakirjoja [Li et al., 2012], mutta myös tällaisia resursseja on saatavilla vain rajoitetulle määrälle kieliä. Kokeissamme oletamme, että koulutusdataa tai tunnistesanakirjoja ei ole saatavilla. Ainoa oletuksemme on useille kielille käännetty tekstinpätkä, tarkemmin sanottuna raamatunkatkelmia. Käytämme Raamatun dataa annotaatioiden projisointiin sekä kieltenvälisten sanasulkeumien oppimiseen (§3). Epävalvottu oppiminen typologisesti informoitujen prioreiden avulla [Naseem et al., 2010] on mielenkiintoinen lähestymistapa epävalvottuun POS-induktioon, jota voidaan soveltaa paremmin vähäisten resurssien kieliin. Työmme liittyy tähän työhön, mutta opimme tietoon perustuvia ennakkoarvioita sen sijaan, että olisimme määritelleet niitä, ja yhdistämme nämä ennakkoarvot annotaatioprojisointiin (oppiminen kohinaisista merkinnöistä) pikemminkin kuin valvomattomaan oppimiseen. https://opennlp.apache.org/ http://nlp.stanford.edu/software/corenlp.shtml

**Tulos**

Empiiriset Gaussin priorit kieltenväliseen siirto-oppimiseen

**Esimerkki 2.1375**

Tang-dynastia (618-907 jKr.) ja Song-dynastia (960-1279) ovat kaksi erittäin tärkeää ajanjaksoa kiinalaisen kirjallisuuden kehityksessä. Tangin ja Songin aikana runouden enemmistömuodot olivat vastaavasti Shi (詩) ja Ci (詞). Tang Shi ja Song Ci loivat kiinalaisen kirjallisuuden keskeiset perusteet, ja niiden vaikutus sekä kirjallisiin teoksiin että kiinalaisten yhteisöjen jokapäiväiseen elämään jatkuu tähän päivään asti. Tunnustaessaan Tang Shin merkityksen Qing-dynastian (1644-1912) Kiinan keisari Kangxi määräsi kokoamaan Tang-runokokoelman Quan-Tang-Shi (QTS, 全唐詩). QTS sisältää lähes 50 tuhatta teosta noin 2200 runoilijalta. Samanlainen pyrkimys koota Song Ci -kokoelma yksityiseltä sektorilta alkoi Ming-dynastian aikana (1368-1644), ja se saavutettiin kokoelmassa nimeltä Quan-Song-Ci (QSC, 全宋詞) Kiinan varhaisella tasavaltalaiskaudella (n. 1937). QSC sisältää noin 20 tuhatta teosta noin 1330 runoilijalta. Tarkat tilastot QTS:stä ja QSC:stä voivat hieman vaihdella lähteistä riippuen. Viimeisen yli tuhannen vuoden aikana kirjallisuuden ja kielitieteen tutkijat ovat tehneet lukemattomia tutkimuksia Tang- ja Song-dynastioiden runoudesta. Näin ollen ei ole mahdollista eikä tavoitteemme tarkastella kirjallisuutta tässä tiivistelmässä. Perinteiset tutkijat ovat tutkineet ja vertailleet eri tekijöiden ja eri aikakausina tuottamia runoteoksia ja tuottaneet niistä oivaltavia ja korvaamattomia analyysejä ja kommentteja. Useimmiten tutkijat keskittyivät valittujen runoilijoiden runoihin. Vaikka tietotekniset apuvälineet ovat tulleet käyttöön, tiettyjen runoilijoiden runojen tutkiminen on edelleen tärkeä ja suosittu runouden tutkimustyyppi. Ohjelmistotyökalut helpottavat runouden analysointia panoraamanäkökulmasta, ja ne voivat johtaa sovelluksiin, jotka aiemmin olisivat olleet hyvin haastavia. Esimerkiksi Zhou kollegoineen analysoi kerättyjen paritelmien ja Tang-runojen sisältöä paritelmien luomista varten. Yan ja hänen

**Tulos**

Tang- ja Song-dynastioiden kiinalaisen runouden kvantitatiiviset analyysit: Muuttuvat värit ja innovatiiviset termit esimerkkeinä.

**Esimerkki 2.1376**

Konvoluutiohermoverkot ovat erinomaisia kuvantunnistustehtävissä, mutta niiden hintana on suuri laskennallinen ja muistin monimutkaisuus. Tämän ongelman ratkaisemiseksi [1] kehitti tensorifaktorointikehyksen, jonka avulla voidaan pakata täysin kytkettyjä kerroksia. Tässä artikkelissa keskitymme konvoluutiokerrosten pakkaamiseen. Osoitamme, että vaikka tensorikehyksen [1] suora soveltaminen konvoluution 4-ulotteiseen ytimeen pakkaa kerroksen, voimme tehdä paremman. Muokkaamme konvoluutioytimen korkeamman kertaluvun tensoriksi ja faktoroimme sen. Yhdistämme ehdotetun lähestymistavan aiempaan työhön, jotta voimme pakata verkon sekä konvoluutio- että täysin kytketyt kerrokset, ja saavutamme 80-kertaisen verkon pakkausnopeuden 1,1 prosentin tarkkuuden pudotuksella CIFAR-10-tietokannassa.

**Tulos**

Perimmäinen tensorointi: konvoluutio- ja FC-kerrosten pakkaaminen samalla tavalla.

**Esimerkki 2.1377**

Ehdotamme tekniikkaa, jolla voidaan havaita ja luoda malleja paikallisesti vuorovaikutuksessa olevien dynaamisten järjestelmien verkossa. Keskeinen osa lähestymistapaamme on uudenlainen spatiaalinen superpositiologiikka, jonka semantiikka määritellään osioituneen kuvan neliöpuun kautta. Osoitamme, että tämän logiikan kaavat voidaan tehokkaasti oppia useiden eri tyyppisten kuvioiden positiivisista ja negatiivisista esimerkeistä. Osoitamme myös, että mallintarkastusalgoritmina toteutettu kuvioiden havaitseminen toimii erittäin hyvin testidatajoukoissa, jotka eroavat oppimisjoukoista. Määrittelemme logiikalle kvantitatiivisen semantiikan ja integroimme mallintarkastusalgoritmin hiukkasparvioptimoinnin kanssa laskennalliseen kehykseen, jossa syntetisoidaan parametreja, jotka johtavat haluttuihin kuvioihin reaktiodiffuusiojärjestelmissä.

**Tulos**

Formaalien menetelmien lähestymistapa reaktiodiffuusiosysteemien kuvioiden synteesiin

**Esimerkki 2.1378**

Semanttiset segmentointitehtävät voidaan mallintaa hyvin Markovin satunnaiskentällä (MRF). Tässä artikkelissa käsitellään semanttista segmentointia sisällyttämällä MRF:ään korkeamman asteen suhteet ja merkintäkontekstien sekoitus. Toisin kuin aiemmissa töissä, joissa MRF:ää optimoitiin iteratiivisella algoritmilla, me ratkaisemme MRF:n ehdottamalla konvolutiivista neuroverkkoa (Convolutional Neural Network, CNN), nimittäin syvää jäsentelyverkkoa (Deep Parsing Network, DPN), joka mahdollistaa deterministisen päästä päähän -laskennan yhdellä eteenpäin -suorittamiskerralla. DPN:ssä laajennetaan nykyaikaista CNN:ää mallintamaan yksikäsitteisiä termejä ja kehitetään lisäkerroksia, joilla voidaan approksimoida keskiarvokentän (MF) algoritmia pareittain esiintyville termeille. Sillä on useita houkuttelevia ominaisuuksia. Ensinnäkin, toisin kuin viimeaikaisissa töissä, jotka vaativat useita MF-iteraatioita takaisinkulkeutumisen aikana, DPN pystyy saavuttamaan korkean suorituskyvyn approksimoimalla MF:n yhden iteraation. Toiseksi DPN edustaa erityyppisiä pareittaisia termejä, jolloin monet olemassa olevat mallit ovat sen erikoistapauksia. Lisäksi DPN:n pareittaiset termit tarjoavat yhtenäisen kehyksen, jonka avulla voidaan koodata runsaasti kontekstuaalista tietoa korkea-ulotteisessa datassa, kuten kuvissa ja videoissa. Kolmanneksi DPN helpottaa MF:n rinnakkaistamista ja nopeuttamista, mikä mahdollistaa tehokkaan päättelyn. DPN:ää on arvioitu perusteellisesti tavanomaisilla semanttisilla kuvan/videon segmentoinnin vertailuarvoilla, joissa yhdellä DPN-mallilla saavutetaan huippuluokan segmentointitarkkuus PASCAL VOC 2012, Cityscapes- ja CamVid-tietokannoissa.

**Tulos**

Syväoppiva Markovin satunnaiskenttä semanttista segmentointia varten

**Esimerkki 2.1379**

Proteiinien sekundäärirakenteen ennustaminen on perustavanlaatuinen ongelma proteiinien rakenteen ennustamisessa. Tässä esitellään uusi valvottu generatiiviseen stokastiseen verkkoon (GSN) perustuva menetelmä paikallisen sekundäärirakenteen ennustamiseksi syvien hierarkkisten esitysten avulla. GSN on hiljattain ehdotettu syväoppimistekniikka (Bengio & Thibodeau-Laufer, 2013) syvän generatiivisen mallin globaaliin kouluttamiseen. Esittelemme GSN:n valvotun laajennuksen, joka oppii Markovin ketjun otoksen ehdollisesta jakaumasta, ja sovellamme sitä proteiinirakenteen ennustamiseen. Jotta malli voidaan skaalata täysimittaiseen, korkea-ulotteiseen dataan, kuten proteiinisekvensseihin, joissa on satoja aminohappoja, otamme käyttöön konvoluutioarkkitehtuurin, joka mahdollistaa tehokkaan oppimisen useiden hierarkkisten esityskerrosten yli. Arkkitehtuurimme keskittyy ainutlaatuisella tavalla ennustamaan rakenteellisia matalan tason merkintöjä, jotka on ilmoitettu mallin oppimien sekä matalan että korkean tason representaatioiden avulla. Sovelluksessamme tämä vastaa kunkin aminohappojäännöksen sekundäärirakenteen tilan merkitsemistä. Koulutimme ja testasimme mallia erillisillä joukoilla ei-homologisia proteiineja, joilla on alle 30 prosentin sekvenssi-identiteetti. Mallimme saavuttaa 66,4 % Q8-tarkkuuden CB513-tietokannassa, mikä on parempi kuin aiemmin raportoitu paras tulos 64,9 % (Wang et al., 2011) tässä haastavassa sekundäärirakenteen ennustamisongelmassa.

**Tulos**

Syvä valvottu ja konvolutiivinen generatiivinen stokastinen verkko proteiinien sekundaarirakenteen ennustamiseen

**Esimerkki 2.1380**

Parametrisoidut algoritmit ovat tapa ratkaista vaikeita ongelmia tehokkaammin, jos tietty parametri on pieni. Tässä artikkelissa sovellamme tätä ajatusta vastausjoukkojen ohjelmoinnin (ASP) alalla. Tätä varten ehdotamme kahdenlaisia ohjelmien graafiesityksiä, joissa hyödynnetään ohjelmien puunleveyttä parametrina. Puuleveys mittaa karkeasti, missä määrin ohjelman sisäinen rakenne muistuttaa puuta. Tärkein panoksemme on sellaisten parametrisoitujen dynaamisen ohjelmoinnin algoritmien suunnittelu, jotka toimivat lineaarisessa ajassa, jos tietyn ohjelman puunleveys ja painot ovat rajattuja. Aiempiin töihin verrattuna algoritmimme käsittelevät ASP:n koko syntaksia. Lopuksi raportoimme empiirisestä arvioinnista, joka osoittaa hyvää ajoaikakäyttäytymistä vertailuesimerkeissä, joiden puunleveys on pieni, erityisesti vastausjoukkojen laskemisessa.

**Tulos**

Vastausjoukkojen ratkaiseminen rajoitetulla puunleveydellä uudelleen tarkasteltuna\*

**Esimerkki 2.1381**

Klusterointi on hyödyllinen tietojen etsintämenetelmä, jota voidaan soveltaa laajasti useilla eri aloilla. Tietojen klusterointi perustuu kuitenkin suuresti klusterikeskusten alustamiseen, joka voi johtaa suureen klusterin sisäiseen vaihteluun ja kuolleisiin keskuksiin, mikä johtaa epäoptimaalisiin ratkaisuihin. Tässä artikkelissa ehdotetaan uutta varianssiin perustuvaa versiota perinteisestä Moving K-Means (MKM) -algoritmista nimellä Variance Based Moving K-Means (VMKM), joka voi jakaa datan optimaalisiin homogeenisiin klustereihin riippumatta klusterin alustuksesta. Algoritmi käyttää uudenlaista etäisyysmetriikkaa ja ainutlaatuista dataelementin valintaperustetta valitun elementin siirtämiseksi klustereiden välillä, jotta saavutetaan alhainen klusterin sisäinen varianssi ja vältetään kuolleet keskukset. Kvantitatiivinen ja kvalitatiivinen vertailu eri klusterointitekniikoihin tehdään neljällä tietokokonaisuudella, jotka on valittu kuvankäsittelystä, bioinformatiikasta, kaukokartoituksesta ja osakemarkkinoista. Laajassa analyysissä korostetaan ehdotetun menetelmän ylivoimaista suorituskykyä muihin tekniikoihin verrattuna.

**Tulos**

Varianssiin perustuva liikkuva K-Means-algoritmi

**Esimerkki 2.1382**

Ehdotamme uutta moduulia, tarkastajamoduulia, jolla parannetaan koodaaja-dekooderi-oppimisjärjestelmää. Arviointimoduuli on yleinen, ja se voidaan liittää olemassa olevaan koodaaja-dekooderimalliin. Tarkistajamoduuli suorittaa useita tarkistusvaiheita, joissa kooderin piilotetut tilat tarkistetaan huomiomekanismin avulla, ja tuottaa jokaisen tarkistusvaiheen jälkeen tietovektorin; tietovektoreita käytetään dekooderin huomiomekanismin syötteenä. Osoitamme, että tavanomaiset koodaajat ja dekooderit ovat kehyksemme erikoistapaus. Empiirisesti osoitamme, että kehyksemme voi parantaa nykyaikaisia koodaaja-dekooderijärjestelmiä kuvien ja lähdekoodin otsikointitehtävissä.

**Tulos**

Koodaus, tarkastelu ja dekoodaus: Reviewer Module for Caption Generation

**Esimerkki 2.1383**

Lauseiden samankaltaisuutta pidetään perustana monissa luonnollisen kielen tehtävissä, kuten tiedonhaussa, kysymyksiin vastaamisessa ja tekstin tiivistämisessä. Vertailtavien tekstikatkelmien välinen semanttinen merkitys perustuu sanojen semanttisiin ominaisuuksiin ja niiden välisiin suhteisiin. Tässä artikkelissa tarkastellaan joukko sanojen ja lauseiden samankaltaisuusmittoja ja verrataan niitä vertailutietokantojen avulla. Tutkituissa tietokokonaisuuksissa tulokset osoittivat, että hybridiset semanttiset mittarit toimivat paremmin kuin sekä tietoon että korpuksiin perustuvat mittarit. Yleiset termit semanttinen samankaltaisuus, luonnollisen kielen prosessointi, laskennallinen kielitiede, tekstin samankaltaisuus.

**Tulos**

Sanojen ja lauseiden samankaltaisuusmittareiden kattava vertaileva tutkimus.

**Esimerkki 2.1384**

Tässä työssä esitellään todennäköisyyspohjainen malli binääriselle CSP:lle, joka tarjoaa hienojakoisen analyysin sen sisäisestä rakenteesta. Jos oletetaan, että CSP:ssä voi tapahtua toimialueen muutos, siinä osoitetaan, miten ilmaistaan ennakoivalla tavalla todennäköisyys sille, että toimialueen arvosta tulee epäjohdonmukainen, ja sen jälkeen ilmaistaan odotusarvo kaaren epäjohdonmukaisten arvojen lukumäärälle kussakin rajoitusverkon toimialueessa. Näin se ilmaisee koko rajoitusverkon kaaren epäjohdonmukaisten arvojen määrän odotuksen. Seuraavaksi annetaan rajat kullekin näistä kolmesta todennäköisyysindikaattorista. Lopuksi esitetään polytime-algoritmi, joka levittää todennäköisyystietoa.

**Tulos**

Todennäköisyyspohjainen malli binääriselle CSP:lle

**Esimerkki 2.1385**

Kielen syvälliseen laskennalliseen käsittelyyn tarvitaan todistusaineistoa, ja yksi tällainen todistusaineisto on korpus. Tässä asiakirjassa kuvataan tekstipohjaisen korpuksen kehittämistä bishnupriya manipurin kielelle. Korpusta pidetään kielenkäsittelytehtävien rakennusaineena. Muiden intialaisten kielten tapaan sitä ei tunneta, joten sitä tutkitaan myös harvemmin. Tämän seurauksena kielestä puuttuu edelleen hyvä korpus ja kielenkäsittelyn perustyökalut. Tietojemme mukaan tämä on ensimmäinen yritys kehittää korpus bishnupriya manipurin kielelle. Avainsanat Korpus, kieli, bishnupriya manipuri, sana-analyysi, sanataajuus, sanataajuus

**Tulos**

Kohti Bishnupriya Manipuri -korpuksen kehittämistä

**Esimerkki 2.1386**

Yhteistyöhön perustuviin turvallisuusaloitteisiin turvaudutaan yhä useammin uhkien torjunnan ajantasaisuuden ja tehokkuuden parantamiseksi. Näistä yhteistoiminnallisella ennakoivalla mustalla listalla<lb>(CPB) pyritään ennustamaan hyökkäyslähteitä useiden sellaisten organisaatioiden antamien hälytysten perusteella, jotka saattavat olla<lb> samankaltaisten hyökkäysten kohteena. Valitettavasti CPB-ehdotuksissa on tähän mennessä keskitytty vain<lb>osumien määrän parantamiseen, mutta unohdettu yhteistyön vaikutus vääriin positiivisiin ja negatiivisiin tuloksiin. Lisäksi uhkatiedon jakaminen aiheuttaa usein tärkeitä luottamuksellisuuteen, luottamuksellisuuteen ja vastuuseen liittyviä kysymyksiä. Tässä asiakirjassa<lb>tarjoamme ensin kattavan mittausanalyysin kahdesta<lb>kannattavimmasta CPB-järjestelmästä: toisessa käytetään luotettavaa keskus-<lb>trallista osapuolta keräämään hälytyksiä [Soldo et al.., Infocom'10] ja<lb>vertaisten välinen järjestelmä, joka luottaa hallittuun tietojen jakamiseen [Freudi-<lb>ger et al., DIMVA'15], ja jossa tutkitaan yhteistyön<lb>vaikutusta sekä oikeisiin että virheellisiin ennusteisiin. Sitten esittelemme<lb>uusin yksityisyyden suojaa kunnioittavan lähestymistavan, joka parantaa<lb>merkittävästi aiempia töitä ja saavuttaa paremman tasapainon oikeiden ja<lb>väärien positiivisten ennusteiden osuuksien välillä samalla kun minimoidaan tietojen paljastuminen.<lb>Viimeiseksi esittelemme laajennuksen, jonka avulla järjestelmämme<lb>skaalautuu hyvin suurille organisaatiolukumäärille.

**Tulos**

Yksityisyyttä suojaavien ennakoivien mustien listojen rakentaminen ja mittaaminen

**Esimerkki 2.1387**

Tässä esseessä tutkitaan, miten naivi Bayes-luokitin ja tukivektorikone vertailevat kykyään ennustaa Thaimaan pörssin kehitystä. Tutkitaan SVM:n ja naive Bayes -luokittimen taustalla olevaa teoriaa. Algoritmeja koulutetaan käyttämällä tammikuun 2010 tietoja, jotka on poimittu MarketWatch.com-sivustolta. Syöttöominaisuudet valitaan SET100-indeksiä koskevien aiempien tutkimusten perusteella. Weka 3 -ohjelmistoa käytetään mallien luomiseen merkityistä harjoitusaineistoista. Kahden algoritmin vertailussa käytetään keskimääräistä neliövirhettä ja oikein luokiteltujen tapausten osuutta sekä useita muita virhemittauksia. Tämä essee osoittaa, että nämä kaksi algoritmia eivät tällä hetkellä ole riittävän kehittyneitä pörssin tarkkaan mallintamiseen. Naiivi Bayes on kuitenkin parempi kuin tukivektorikone Thaimaan pörssin ennustamisessa.

**Tulos**

Miten naiivi Bayes-luokittelija ja tukivektorikone vertailevat kykyään ennustaa Thaimaan pörssin kehitystä?

**Esimerkki 2.1388**

Kehitämme kehyksen mallin valinnan jälkeistä päättelyä varten lineaarisen regression marginaaliseulonnan avulla. Tämän kehyksen ytimessä on tulos, joka luonnehtii vastauksen y lineaaristen funktioiden tarkkaa jakaumaa, joka riippuu valitusta mallista ("valintatilanteen" kehys). Näin voimme rakentaa validi luottamusvälit ja hypoteesitestit regressiokertoimille, joissa otetaan huomioon valintamenettely. Toisin kuin korkea-ulotteisen tilastoinnin alalla viime aikoina tehdyt työt, tuloksemme ovat täsmällisiä (ei-asymptoottisia), eivätkä ne edellytä ominaisarvojen kaltaisia oletuksia suunnittelumatriisista X. Lisäksi marginaalisen regression, luottamusvälien rakentamisen ja hypoteesien testaamisen laskentakustannukset ovat mitättömät verrattuna lineaarisen regression kustannuksiin, minkä vuoksi menetelmämme soveltuvat erityisen hyvin erittäin suurille aineistoille. Vaikka keskitymme marginaaliseen seulontaan havainnollistaaksemme valintaehdon sovellettavuutta, tämä kehys on paljon laajemmin sovellettavissa. Näytämme, miten ehdotettua kehystä voidaan soveltaa useisiin muihin valintamenettelyihin, kuten ortogonaaliseen matching pursuitiin, ei-negatiivisiin pienimpiin neliöihin ja marginaaliseen seulontaan + Lasso.

**Tulos**

Tarkka jälkikäteinen mallinvalinnan päättely marginaalista seulontaa varten

**Esimerkki 2.1389**

Tässä artikkelissa kuvaamme syväverkkoarkkitehtuurin, joka kartoittaa visuaalisen syötteen ohjaustoiminnoiksi robottirobottien tasokurottajatehtävässä 100-prosenttisella luotettavuudella reaalimaailman kokeissa. Verkko koulutetaan simulaatiossa ja hienosäädetään rajallisella määrällä reaalimaailman kuvia. Toimintatapojen hakua ohjaa kinematiikkaan perustuva ohjain (K-GPS), joka toimii tehokkaammin ja tuloksellisemmin kuin ε-Greedy. Ratkaiseva oivallus järjestelmässämme on, että havainto- ja ohjausverkkojen väliin on lisättävä pullonkaula ja että näitä verkkoja on aluksi koulutettava toisistaan riippumatta.

**Tulos**

Näköön perustuva tavoittaminen modulaaristen syväverkkojen avulla: simulaatiosta reaalimaailmaan

**Esimerkki 2.1390**

Automaattinen kuvamerkintä on yksi haastavimmista ongelmista konenäköalalla. Tehtävän tavoitteena on ennustaa avainsanojen määrä automaattisesti reaaliaineistosta otetuille kuville. Monet menetelmät perustuvat visuaalisiin ominaisuuksiin, joiden avulla lasketaan kuvanäytteiden välisiä samankaltaisuuksia. Näiden lähestymistapojen laskentakustannukset ovat kuitenkin hyvin korkeat. Nämä menetelmät edellyttävät monien harjoitusnäytteiden tallentamista muistiin. Tämän taakan vähentämiseksi on kehitetty useita tekniikoita, joilla voidaan vähentää piirteiden määrää tietokannassa. Manifold-oppiminen on suosittu lähestymistapa epälineaariseen dimensioiden pienentämiseen. Tässä artikkelissa tutkitaan diffuusiokarttojen moninaisuusoppimismenetelmää verkkokuvien automaattista merkintätehtävää varten. Diffuusiokarttojen moninaisuusoppimismenetelmää käytetään joidenkin visuaalisten piirteiden ulottuvuuden vähentämiseen. Laajat kokeet ja analyysit NUS-WIDE-LITE-verkkokuvatietokannassa, jossa on erilaisia visuaalisia piirteitä, osoittavat, miten tätä moninaisuusoppimismenetelmää voidaan soveltaa tehokkaasti kuvien annotoinnissa.

**Tulos**

VERKKOKUVIEN MERKINTÄ DIFFUUSIOKARTTOJEN MONINAISEN OPPIMISALGORITMIN AVULLA

**Esimerkki 2.1391**

syväoppimisen onnistuminen riippuu siitä, että löydetään tehtävään sopiva arkkitehtuuri. Kun syväoppiminen on laajentunut haastavampiin tehtäviin, arkkitehtuurien suunnittelu käsin on käynyt vaikeaksi. Tässä artikkelissa ehdotetaan automaattista CoDeepNEAT-menetelmää, jolla syväoppimisarkkitehtuurit voidaan optimoida evoluution avulla. Laajentamalla nykyisiä neuroevoluutiomenetelmiä topologiaan, komponentteihin ja hyperparametreihin tämä menetelmä saavuttaa tuloksia, jotka ovat verrattavissa parhaisiin ihmisen tekemiin suunnitelmiin tavanomaisissa vertailuarvoissa, jotka liittyvät objektien tunnistamiseen ja kielen mallintamiseen. Se tukee myös reaalimaailman sovelluksen rakentamista, joka koskee automaattista kuvatekstitystä aikakauslehden verkkosivustolla. Kun otetaan huomioon käytettävissä olevan laskentatehon odotettavissa oleva kasvu, syväverkkojen evoluutio on lupaava lähestymistapa syväoppimissovellusten rakentamiseen tulevaisuudessa.

**Tulos**

Kehittyvät syvät neuroverkot

**Esimerkki 2.1392**

Tässä artikkelissa tarkastelemme potilaan samankaltaisuusongelmaa yli 220 000 potilaan syöpäkohortissa. Lähestymistapamme hyödyntää ensin Word2Vec-kehystä ICD-koodien upottamiseksi vektoriarvoiseen esitykseen. Sen jälkeen ehdotamme peräkkäistä algoritmia tapaus-vertailuvertailua varten tässä diagnoosikoodien esitysavaruudessa. Uudenlainen käytäntö soveltaa peräkkäistä täsmäytystä vektoriesitykseen paransi täsmäytystarkkuutta useiden kliinisten tulosten avulla mitattuna. Raportoimme tulokset laajamittaisesta tietokokonaisuudesta osoittaaksemme menetelmämme tehokkuuden. Uusi menetelmä on erityisen merkityksellinen näin suuressa tietokokonaisuudessa, jossa suurin osa kliinisistä tiedoista on koodattu.

**Tulos**

Ohjauksen yhteensovittaminen purkauskoodijaksojen avulla

**Esimerkki 2.1393**

Entiteettien ja relaatioiden yhteinen louhinta on tärkeä tehtävä tiedon louhinnassa. Tämän ongelman ratkaisemiseksi ehdotamme aluksi uutta merkintäjärjestelmää, jonka avulla yhteinen uuttamistehtävä voidaan muuttaa merkintäongelmaksi. Tämän jälkeen tutkimme merkintäjärjestelmään perustuen erilaisia päästä päähän -malleja, joiden avulla entiteetit ja niiden suhteet voidaan poimia suoraan ilman entiteettien ja suhteiden erillistä tunnistamista. Teemme kokeita julkisella tietokokonaisuudella, joka on tuotettu etävalvontamenetelmällä, ja kokeelliset tulokset osoittavat, että merkintöihin perustuvat menetelmät ovat parempia kuin useimmat nykyiset putkijohdetut ja yhteiset oppimismenetelmät. Lisäksi tässä artikkelissa ehdotettu kokonaisvaltainen malli saavuttaa parhaat tulokset julkisessa tietokokonaisuudessa.

**Tulos**

Entiteettien ja suhteiden yhteinen uuttaminen uudenlaisen merkintäjärjestelmän perusteella

**Esimerkki 2.1394**

Esittelemme tehokkaan rekursiiviseen neuroverkkoon perustuvan menetelmän uutuuksien havaitsemiseen sovelluksessa, jossa havaitaan radioanomalioita. Tämä lähestymistapa lupaa parantaa merkittävästi naiivin poikkeamien havaitsemisen kykyä havaita pieniä poikkeamia erittäin monimutkaisilla ja monimutkaisilla monikäyttäjäradioalueilla. Osoitamme tämän lähestymistavan tehokkuuden useilla yleisillä todellisilla, kiinnostavilla radioviestintäkaistoilla ja määrittelemme havaitsemistehokkuuden havaitsemistodennäköisyytenä ja väärien hälytysten määränä eri häiriöiden ja kaistan tehosuhteiden välillä ja vertaamme sitä perusmenetelmiin.

**Tulos**

Rekurrenssiaikaista neuraalista radiotaajuuksien poikkeamien havaitseminen

**Esimerkki 2.1395**

Syvien neuroverkkojen (Deep Neural Networks, DNN) aktivointifunktio on muuttunut paljon viime vuosikymmeninä. Sen jälkeen, kun tunnettu ei-kyllästetty suoristettu lineaarinen yksikkö (ReLU) tuli markkinoille, monet ovat yrittäneet parantaa verkkojen suorituskykyä entistä monimutkaisemmilla funktioilla. Esimerkkejä ovat vuotava ReLU (LReLU) nollakaltevuuksien poistamiseksi ja eksponentiaalinen lineaarinen yksikkö (ELU) bias-siirtymän vähentämiseksi. Tässä artikkelissa esitellään parametrinen ELU (PELU), adaptiivinen aktivointifunktio, jonka avulla DNN:t voivat omaksua erilaisia epälineaarisia käyttäytymismalleja koko koulutusvaiheen ajan. Olemme mukana kolmella tavalla: (1) osoitamme, että PELU lisää verkon joustavuutta häviävän gradientin torjumiseksi, (2) tarjoamme gradienttipohjaisen optimointikehyksen funktion parametrien oppimiseksi ja (3) suoritamme useita kokeita MNIST-, CIFAR-10/100- ja ImageNet-verkoilla erilaisilla verkkoarkkitehtuureilla, kuten NiN, Overfeat, All-CNN, ResNet ja Vgg, osoittaaksemme lähestymistavan yleisen sovellettavuuden. Ehdotettu PELU on parantanut suhteellista virhettä 4,45 % ja 5,68 % CIFAR10- ja 100-testissä ja jopa 7,28 % ImageNet-testissä vain 0,0003 %:n parametrin lisäyksellä sekä nopeuttanut konvergenssinopeutta lähes kaikissa testiskenaarioissa. Havaitsimme myös, että PELU:ta käyttävällä Vgg:llä oli taipumus suosia aktivointeja, jotka kyllästyivät lähelle nollaa, kuten ReLU:ssa, lukuun ottamatta viimeistä kerrosta, joka kyllästyi lähelle -2. Nämä tulokset viittaavat siihen, että aktivointien muodon vaihtelu harjoittelun aikana yhdessä muiden parametrien kanssa auttaa hallitsemaan katoavia gradientteja ja harhasiirtymiä, mikä helpottaa oppimista.

**Tulos**

Parametrinen eksponentiaalinen lineaarinen yksikkö syviä konvolutiivisia neuroverkkoja varten

**Esimerkki 2.1396**

Ehdokkaiden peräkkäinen poistaminen on usein keino tehdä manipuloinnista vaikeasti laskettava. Osoitamme, että ehdokkaiden poistaminen ei välttämättä lisää manipuloinnin laskennallista monimutkaisuutta. Monien käytännössä käytettyjen äänestyssääntöjen laskennallinen monimutkaisuus kuitenkin kasvaa. Esimerkiksi jo tiedetään, että on NP-vaikeaa laskea, miten yksittäinen äänestäjä voi manipuloida yksittäisen siirrettävän äänestyksen (moniäänestyksen eliminointiversio) tulosta. Näytämme tässä, että on NP-vaikeaa laskea, miten yksittäinen äänestäjä voi manipuloida veto-äänestyksen eliminointiversion, siihen läheisesti liittyvän Coombsin säännön ja yleisen luokan pisteytyssääntöjen eliminointiversioiden tulosta.

**Tulos**

Heikoimman lenkin poistaminen: Manipulointi vaikeasti hallittavaksi?

**Esimerkki 2.1397**

Verkon sulauttamisella (NE) on tärkeä rooli verkkoanalyysissä, koska se pystyy esittämään kärkipisteet tehokkailla matalaulotteisilla sulauttamisvektoreilla. Nykyiset NE-mallit pyrkivät kuitenkin oppimaan kiinteän kontekstittoman upotuksen kullekin verteksille ja jättävät huomiotta niiden erilaiset roolit vuorovaikutuksessa muiden verteksien kanssa. Tässä artikkelissa oletamme, että yksi huippu näyttää yleensä eri näkökohtia vuorovaikutuksessa eri naapuripisteiden kanssa, ja sen pitäisi omistaa vastaavasti eri upotuksia. Tämän vuoksi esitellään ContextAware Network Embedding (CANE), joka on uusi NE-malli tämän ongelman ratkaisemiseksi. CANE oppii kontekstitietoisia upotuksia kärkipisteille keskinäisen huomion mekanismin avulla, ja sen odotetaan mallintavan kärkipisteiden välisiä semanttisia suhteita tarkemmin. Kokeissa vertaamme malliamme olemassa oleviin NE-malleihin kolmella todellisella tietokokonaisuudella. Kokeelliset tulokset osoittavat, että CANE parantaa merkittävästi nykyaikaisia menetelmiä linkkien ennustamisessa ja saavuttaa vertailukelpoisen suorituskyvyn vertikaalien luokittelussa.

**Tulos**

CANE: Kontekstitietoinen verkon sulauttaminen suhteiden mallintamiseen

**Esimerkki 2.1398**

Kuvaamme kysymysten vastaamisen mallin, jota voidaan soveltaa sekä kuviin että strukturoituihin tietopankkeihin.Malli käyttää luonnollisen kielen merkkijonoja kootakseen automaattisesti neuroverkkoja yhdistettävien moduulien kokoelmasta. Näiden moduulien parametrit opitaan yhdessä verkon kokoonpanoparametrien kanssa vahvistusoppimisen avulla, ja valvonnassa käytetään vain (maailma, kysymys, vastaus) -kolmioita. Lähestymistapamme, jota kutsumme dynaamiseksi neuraalimoduuliverkoksi, tuottaa huipputuloksia sekä visuaalisten että strukturoitujen alojen vertailutietoaineistoissa.

**Tulos**

Neuraaliverkkojen kokoamisen oppiminen kysymyksiin vastaamista varten

**Esimerkki 2.1399**

Rekursiivista neuroverkkoa (RNN) on sovellettu laajasti luonnollisen kielen käsittelyyn (NLP). Tällainen neuroverkko on suunniteltu peräkkäisten tietojen mallintamiseen, ja sen on todettu olevan varsin tehokas peräkkäisissä merkintätehtävissä. Tässä asiakirjassa ehdotetaan kaksisuuntaisen RNN:n käyttöä pitkän lyhytkestoisen muistin (LSTM) yksiköillä kiinalaisten sanojen segmentointiin, joka on tärkeä esikäsittelytehtävä kiinalaisten lauseiden ja artikkelien mallintamisessa. Klassiset menetelmät keskittyvät suunnittelemaan ja yhdistelemään käsityönä valmistettuja piirteitä kontekstista, kun taas kaksisuuntainen LSTM-verkko (BLSTM) ei tarvitse ennakkotietoa tai esisuunnittelua, ja se on asiantuntija säilyttämään kontekstuaalista tietoa molempiin suuntiin. Kokeiden tulokset osoittavat, että lähestymistapamme saavuttaa huippuluokan suorituskyvyn sanojen segmentoinnissa sekä perinteisen kiinan että yksinkertaistetun kiinan tietokokonaisuuksissa.

**Tulos**

Kaksisuuntainen LSTM-rekursiivinen neuroverkko kiinan sanojen segmentointiin

**Esimerkki 2.1400**

Käyttäjien yksityisyyteen liittyvien huolenaiheiden vuoksi suunnittelemme steganografisen järjestelmän ("stegosystem"), jonka avulla kaksi käyttäjää voi vaihtaa salattuja viestejä ilman, että vastustaja havaitsee, että tällainen vaihto tapahtuu. Ehdotamme uutta kielellistä stegosysteemiä, joka perustuu LSTM-neuraaliverkkoon (Long ShortTerm Memory). Esittelemme lähestymistapamme Twitter- ja Enron-sähköpostitietokannoissa ja osoitamme, että sen avulla saadaan laadukasta steganografista tekstiä ja samalla parannetaan kapasiteettia (salattuja bittejä sanaa kohti) huomattavasti verrattuna uusimpaan tekniikkaan.

**Tulos**

Steganografisen tekstin tuottaminen LSTM:n avulla

**Esimerkki 2.1401**

Sveitsin lumivyörytiedote laaditaan kahdesti päivässä neljällä kielellä. Koska manuaaliseen kääntämiseen ei ole riittävästi aikaa, käytetään täysin automaattista käännösjärjestelmää, joka perustuu ennalta määriteltyjen lauseiden luetteloon ja ennalta määriteltyihin sääntöihin siitä, miten nämä lauseet voidaan yhdistää lauseiksi. Koska tämä lausekeluettelo rajoittuu pieneen alakieleen, järjestelmä pystyy automaattisesti kääntämään tällaisia lauseita saksasta kohdekielille ranska, italia ja englanti ilman myöhempää oikolukua tai korjausta. Koska järjestelmä on ollut toiminnassa kahden talvikauden ajan, arvioimme tässä tuotettujen tekstien laatua kahden erilaisen tutkimuksen perusteella, joissa osallistujat arvioivat tekstejä, jotka on laadittu todellisista lumivyörytiedotteista ja jotka ovat peräisin molemmista lähteistä, lausekeluettelosta ja manuaalisesti kirjoitetuista ja käännetyistä teksteistä. Käyttäjät eivät juurikaan erota näitä kahta tekstityyppiä toisistaan 55 prosentin keskimääräisellä tunnistamisasteella ja antavat hyvin samanlaisia arvioita niiden kielellisestä laadusta. Kaiken kaikkiaan luettelointijärjestelmän tulosta voidaan pitää käytännössä samanlaisena kuin lumivyöryennustajien kirjoittamaa ja sitten ammattikääntäjien käsin kääntämää tekstiä. Lisäksi ennustajat ilmoittivat, että järjestelmä kuvasi kaikki olennaiset tilanteet riittävällä tarkkuudella. Ennustajien työtaakka ei muuttunut luettelon käyttöönoton myötä: vastaavien lauseiden etsimiseen kuluva ylimääräinen aika kompensoituu sillä, että heidän ei enää tarvitse tarkistaa käsin käännettyjä tekstejä. Päivittäisten käännöskustannusten vähenemisen odotetaan korvaavan alkuperäiset kehityskustannukset muutamassa vuodessa.

**Tulos**

Täysin automaattinen monikielinen käännös, jossa on lausekeluettelo - menestyksekäs käyttö Sveitsin lumivyörytiedotteessa.

**Esimerkki 2.1402**

Niin sanottujen rajoitusongelmien ratkaisemiseen tarkoitetut ohjelmat ovat monimutkaisia ohjelmistoja, jotka edellyttävät, että toteuttaja tekee monia suunnittelupäätöksiä enemmän tai vähemmän mielivaltaisesti. Nämä päätökset vaikuttavat merkittävästi valmiin ratkaisijan suorituskykyyn [13]. Kun suunnittelupäätös on kerran tehty, sitä ei voi helposti perua, vaikka toisenlainen päätös saattaakin olla sopivampi tiettyyn ongelmaan. Tutkimme koneoppimisen käyttöä näiden päätösten tekemiseen automaattisesti ratkaistavan ongelman mukaan esimerkkinä alldifferent-rajoitteesta. Järjestelmämme kykenee tekemään ei-triviaaleja, monitasoisia päätöksiä, jotka ovat parempia kuin se, että tehdään aina oletusvalinta.

**Tulos**

Koneoppimisen käyttäminen rajoitusratkaisijan toteuttamista koskevien päätösten tekemiseen

**Esimerkki 2.1403**

Tässä artikkelissa kehitämme kehyksen tietoteoreettiselle oppimiselle, joka perustuu äärettömän jaollisiin matriiseihin. Muotoilemme entropian kaltaisen funktionaalin positiivisesti definiittisille matriiseille Renyin aksiomaattisen entropian määritelmän pohjalta ja tutkimme joitakin tämän funktionaalin keskeisiä ominaisuuksia, jotka johtavat äärettömän jaettavuuden käsitteeseen. Ehdotettu muotoilu välttää tiheyden estimoinnin pistokkeen ja tuo mukanaan toistavien Hilbert-avaruuksien ytimien esityskyvyn. Sovellusesimerkkinä johdamme valvotun metrisen oppimisalgoritmin, jossa käytetään matriisipohjaista analogiaa ehdolliselle entropialle ja saavutetaan nykyiseen tekniikkaan verrattavia tuloksia.

**Tulos**

Informaatioteoreettinen oppiminen äärettömän jaettavissa olevilla ytimillä

**Esimerkki 2.1404**

Erikoistarkoituksiin tarkoitetut rajoitusten etenemisalgoritmit käyttävät usein implisiittisesti lyhyitä tukia - tutkimalla muuttujien osajoukkoa ne voivat päätellä tukia (perustelu sille, että muuttuja-arvopari voi edelleen olla osa rajoituksen täyttävää tehtävää) kaikille muille muuttujille ja arvoille ja säästää huomattavan paljon työtä - mutta lyhyitä tukia ei ole tutkittu sellaisenaan. Tämän artikkelin kaksi tärkeintä saavutusta ovat lyhyiden tukien tunnistaminen tärkeiksi rajoitteiden etenemisen kannalta ja HaggisGAC:n esittely, joka on tehokas ja toimiva yleiskäyttöinen etenemisalgoritmi lyhyiden tukien hyödyntämiseen. Koska HaggisGAC on monimutkainen, esitämme sen optimoituna versiona yksinkertaisemmasta ShortGAC-algoritmista. Vaikka kokeet osoittavat ShortGAC:n tehokkuuden verrattuna muihin yleiskäyttöisiin etenemisalgoritmeihin, kun käytettävissä on kompakti joukko lyhyitä tukia, osoitamme teoreettisesti ja kokeellisesti, että HaggisGAC on vielä parempi. HaggisGAC toimii myös paremmin kuin GAC-Schema täyspitkillä tuilla. Esittelemme myös muunnelma-algoritmin HaggisGACStable, joka on mukautettu välttämään takapolkujen tekemistä ja joka voi joissakin tapauksissa olla nopeampi ja vähentää merkittävästi muistin käyttöä. Kaikki ehdotetut algoritmit soveltuvat erinomaisesti rajoitteiden disjunktioiden etenemiseen. Kaikissa disjunktioita koskevissa kokeissa havaitsimme algoritmiemme olevan nopeampia kuin Constructive Or ja GAC-Schema vähintään kertaluokkaa ja jopa kolme kertaluokkaa.

**Tulos**

Lyhyet ja pitkät tuet rajoitusten levittämistä varten

**Esimerkki 2.1405**

Toiminnallisella neurokuvauksella voidaan mitata aivojen vastetta ulkoiseen ärsykkeeseen. Sitä käytetään aivojen kartoittamiseen: näiden havaintojen perusteella tunnistetaan asianomaiset aivoalueet. Tämä ongelma voidaan muuntaa lineaariseksi valvotuksi oppimistehtäväksi, jossa neurokuvantamistietoja käytetään ärsykkeen ennustajina. Aivojen kartoitus nähdään tällöin tuen talteenotto-ongelmana. Toiminnallisen magneettikuvauksen (fMRI) aineistolla tämä ongelma on erityisen haastava, koska i) näytteiden määrä on pieni rajallisen hankinta-ajan vuoksi ja ii) muuttujat korreloivat voimakkaasti. Ehdotamme näiden vaikeuksien voittamista käyttämällä harvoja regressiomalleja uusille muuttujille, jotka saadaan alkuperäisten muuttujien klusteroimalla. Satunnaistamistekniikoiden, esim. bootstrap-näytteiden, ja muuttujien klusteroinnin käyttö parantaa harvalukuisten menetelmien palautusominaisuuksia. Osoitamme lähestymistapamme hyödyn laajassa simulaatiotutkimuksessa sekä kahdessa fMRI-tietoaineistossa.

**Tulos**

Pienten näytteiden aivokartoitus: harva talteenotto alueellisesti korreloituneissa malleissa satunnaistamalla ja klusteroimalla.

**Esimerkki 2.1406**

Sanakirja määrittelee sanat toisten sanojen avulla. Määritelmät voivat kertoa sinulle tuntemattomien sanojen merkitykset, mutta vain jos tiedät määrittelevien sanojen merkitykset. Kuinka monta sanaa sinun on tiedettävä (ja mitkä niistä), jotta voit oppia kaikki loput määritelmistä? Pelkistimme sanakirjat niiden "perusytimiin" (GK), noin 10 prosenttiin sanakirjasta, joista kaikki muut sanat voidaan määritellä. GK-sanoilla osoittautui olevan psykolingvistisiä korrelaatioita: ne opittiin varhaisemmassa iässä ja konkreettisemmin kuin muu osa sanakirjasta. Mutta voidaan tiivistää vielä enemmän: GK:lla osoittautuu olevan sisäinen rakenne, jossa on vahvasti kytkeytynyt "ydinydin" (KC) ja sitä ympäröivä kerros, josta voidaan johtaa määrittelyetäisyyksien hierarkia aina koko sanakirjan periferiaan asti. Myös nämä määritelmäetäisyydet korreloivat psykolingvististen muuttujien kanssa (oppimisikä, konkreettisuus, kuvallisuus, suullinen ja kirjallinen esiintymistiheys) ja näin ollen ehkäpä jokaisen päässämme olevan "mentaalisen sanaston" kanssa.

**Tulos**

Hierarkiat sanakirjassa Määritelmä Space

**Esimerkki 2.1407**

Nansonin ja Baldwinin äänestyssäännöt valitsevat voittajan poistamalla peräkkäin ehdokkaat, joilla on alhainen Borda-pistemäärä. Osoitamme, että näillä säännöillä on useita toivottavia laskennallisia ominaisuuksia. Erityisesti painottamattomilla äänillä on NP-vaikeaa manipuloida kumpaakin sääntöä yhdellä manipulaattorilla, kun taas painotetuilla äänillä on NP-vaikeaa manipuloida kumpaakin sääntöä pienellä määrällä ehdokkaita ja manipulaattoreiden koalitiolla. Koska vain parin muun äänestyssäännön tiedetään olevan NP-vaikea manipuloitavissa yhdellä manipulaattorilla, Nansonin ja Baldwinin säännöt vaikuttavat teoreettisesta näkökulmasta erityisen resistenteiltä manipuloinnille. Ehdotamme myös useita approksimaatiomenetelmiä näiden kahden säännön manipuloimiseksi. Kokeet osoittavat, että molempia sääntöjä on usein vaikea manipuloida käytännössä. Nämä tulokset viittaavat siihen, että eliminointityylin äänestyssääntöjä on syytä tutkia lisää.

**Tulos**

Nansonin ja Baldwinin sääntöjen manipulointi

**Esimerkki 2.1408**

<lb>Tutkimme stokastisen optimoinnin ongelmaa syväoppimista varten paral-<lb>lel-laskentaympäristössä viestintärajoitusten alaisena. Tässä ympäristössä ehdotetaan uutta algoritmia<lb>, jossa kommunikaatio ja työn<lb>koordinointi rinnakkaisten prosessien (paikalliset työläiset) välillä perustuu elastiseen voimaan, joka<lb>liittää heidän laskemansa parametrit parametri<lb>palvelimen (master) tallentamaan keskusmuuttujaan. Algoritmi antaa paikallisille työläisille mahdollisuuden suorittaa enemmän explo-<lb>raatiota, eli algoritmi antaa paikallisten muuttujien vaihdella kauempana<lb>keskusmuuttujasta vähentämällä paikallisten työläisten<lb>ja masterin välisen viestinnän määrää. Osoitamme empiirisesti, että syväoppimisympäristössä, johtuen<lb>monien paikallisten optimaalien olemassaolosta, suuremman eksploraation salliminen voi johtaa<lb>parempaan suorituskykyyn. Ehdotamme synkronisia ja asynkronisia muunnelmia<lb>uusesta algoritmista. Esitämme asynkronisen muunnoksen<lb>vakausanalyysin round-robin -järjestelmässä ja vertaamme sitä yleisempään rinnakkaistettuun<lb>menetelmään ADMM. Osoitamme, että EASGD:n stabiilisuus on taattu, kun yksinkertainen<lb>stabiilisuusehto täyttyy, mikä ei päde ADMM:n tapauksessa. Lisäksi<lb>esitämme algoritmimme momentumiin perustuvan version, jota voidaan soveltaa sekä<lb>synkronisissa että asynkronisissa asetuksissa. Algoritmin<lb>asynkronista muunnosta<lb>sovelletaan konvoluutiohermoverkkojen kouluttamiseen kuvien luokittelua varten<lb>CIFAR- ja ImageNet-tietokannoissa. Kokeet osoittavat, että uusi algoritmi<lb>kiihdyttää syväarkkitehtuurien kouluttamista verrattuna DOWNPOURiin ja muihin<lb>yleisiin peruslähestymistapoihin ja on lisäksi erittäin viestintätehokas.

**Tulos**

Syväoppiminen elastisella keskiarvotuksella SGD:llä

**Esimerkki 2.1409**

FTL-algoritmin (follow the leader), joka on ehkä yksinkertaisin kaikista verkko-oppimisalgoritmeista, tiedetään suoriutuvan hyvin, kun sen käyttämät häviöfunktiot ovat kupera ja positiivisesti kaareva. Tässä artikkelissa kysymme, onko olemassa muita "onnekkaita" asetuksia, joissa FTL saavuttaa epälineaarisen, "pienen" katumuksen. Tutkimme erityisesti lineaarisen ennustamisen perusongelmaa ei-tyhjällä, konveksilla ja kompaktilla alueella. Muiden tulosten ohella todistamme, että toimialueen rajan kaarevuus voi toimia ikään kuin häviöt olisivat kaarevia: Tällöin todistamme, että niin kauan kuin tappiovektoreiden keskiarvoilla on positiivisia pituuksia, jotka ovat nollasta poispäin rajoitettuja, FTL:n katumuksen kasvuvauhti on logaritminen, kun taas esimerkiksi polytooppialueilla ja stokastisilla tiedoilla sen odotettu katumus on äärellinen. Aikaisemmin tunnetun meta-algoritmin pohjalta saadaan myös algoritmi, joka nauttii samanaikaisesti pahimman tapauksen takuista ja FTL:n käytettävissä olevasta rajoituksesta.

**Tulos**

Johtajan seuraaminen ja nopeat nopeudet lineaarisessa ennusteessa: Kaarevat rajoitusjoukot ja muut säännönmukaisuudet∗∗.

**Esimerkki 2.1410**

Valvotut (lineaariset) upotusmallit, kuten Wsabie [5] ja PSI [1], ovat osoittautuneet menestyksekkäiksi luokittelu-, suosittelu- ja merkintätehtävissä. Vaikka ne ovat skaalautuvia suurille tietokokonaisuuksille, ne eivät kuitenkaan hyödynnä ylimääräistä dataa täysimääräisesti lineaarisen luonteensa vuoksi, ja ne ovat tyypillisesti alimitoitettuja. Ehdotamme uutta mallien luokkaa, jonka tavoitteena on parantaa suorituskykyä ja säilyttää samalla monet nykyisten upotusmallien luokkien eduista. Uusi lähestymistapamme toimii oppimalla iteratiivisesti lineaarinen sulautusmalli, jossa seuraavan iteraation piirteet ja merkinnät painotetaan uudelleen edellisen iteraation funktiona. Kuvaamme useita mallin muunnelmia ja esitämme joitakin alustavia tuloksia. 1 (Valvotut) lineaariset upotusmallit Tavalliset lineaariset upotusmallit ovat muotoa: f(x, y) = xUV y = ∑ (Valvotut) lineaariset upotusmallit.

**Tulos**

Affiniteettipainotettu sulauttaminen

**Esimerkki 2.1411**

Ehdotamme menetelmää, jossa yhdistetään relaatiologiset esitykset ja neuroverkko-oppiminen. Yleinen nostettu arkkitehtuuri, joka mahdollisesti heijastaa jonkinlaista taustatietoa, kuvataan relaatiosääntöjen avulla, jotka voidaan laatia käsin tai oppia. Relaatiosääntöjoukko toimii mallina, jonka avulla voidaan kehittää mahdollisesti syviä neuroverkkoja, joiden rakenteet heijastavat myös annettujen harjoitus- tai testausesimerkkien rakenteita. Eri esimerkkejä vastaavat eri verkot jakavat painonsa, jotka kehittyvät yhdessä koulutuksen aikana stokastisen gradienttilaskualgoritmin avulla. Kehys mahdollistaa hierarkkiset relaatiomallinnusrakenteet ja latenttien relaatiokäsitteiden oppimisen sääntöjä vastaavien jaettujen piilokerrosten painojen avulla. Merkittävien relaatiokäsitteiden löytäminen ja kokeet 78:lla relaatio-oppimisen vertailuarvolla osoittavat menetelmän suotuisan suorituskyvyn.

**Tulos**

Nostetut relationaaliset neuroverkot

**Esimerkki 2.1412**

Tässä artikkelissa käsitellään kysymystä siitä, miten kielenkäyttö vaikuttaa yhteisön reaktioihin kommentteihin verkkokeskustelufoorumeilla ja viestin ja viestinviejän suhteelliseen merkitykseen. Tutkimuksessa ehdotetaan uutta kommenttien luokittelutehtävää, joka perustuu Reddit-keskusteluissa käytettyyn yhteisön kommentoituun karmaan ja jossa otetaan huomioon kommenttien aihe ja ajoitus. Kuuden subredditin keskusteluketjujen kanssa tehty kokeellinen työ osoittaa, että erityyppisten kielellisten piirteiden merkitys vaihtelee kiinnostuksen kohteena olevan yhteisön mukaan.

**Tulos**

Puhuu yleisölle: Mihin ihmiset reagoivat verkkokeskusteluissa?

**Esimerkki 2.1413**

Tekstit esittävät yhtenäisiä tarinoita, joilla on tietty teema tai yleinen asetelma, esimerkiksi tieteiskirjallisuus tai lännenelokuva. Tässä artikkelissa esittelemme tekstien luomismenetelmän, jota kutsutaan uudelleenkirjoittamiseksi ja jolla muokataan olemassa olevia ihmisen kirjoittamia kertomuksia niiden teeman muuttamiseksi muuttamatta taustalla olevaa tarinaa. Sovellamme lähestymistapaa matemaattisiin sanatehtäviin, joissa se voi auttaa oppilaita pysymään sitoutuneempina, kun kaikki kotitehtävät voidaan nopeasti muuttaa suosikkielokuvan teeman mukaisiksi muuttamatta opetettavia matematiikan käsitteitä. Uudelleenkirjoitusmenetelmämme käyttää kaksivaiheista dekoodausprosessia, jossa ehdotetaan uusia sanoja kohdeteemasta ja pisteytetään tuloksena syntyneet tarinat useiden syntaktisen, semanttisen ja temaattisen johdonmukaisuuden määrittelevien tekijöiden mukaan. Kokeet osoittavat, että lopulliset tarinat edustavat tyypillisesti hyvin uutta teemaa ja testaavat samalla alkuperäisiä matematiikan käsitteitä, ja ne ovat parempia kuin useat vertailutarinat. Julkaisemme myös uuden tietokokonaisuuden, joka sisältää ihmisen kirjoittamia matematiikan sanatehtävien uudelleenkirjoituksia useista eri teemoista.

**Tulos**

Teemojen uudelleenkirjoittaminen algebran sanaongelmien luomiseksi

**Esimerkki 2.1414**

Suuren mittakaavan monirelaation upottamisella viitataan tehtävään, joka liittyy olioiden ja suhteiden latenttien representaatioiden oppimiseen suurissa tietämysgraafeissa. Tehokkaan ja skaalautuvan ratkaisun löytäminen tähän ongelmaan on ratkaisevan tärkeää, jotta tietoon perustuva päättely todella onnistuisi monenlaisissa sovelluksissa. Tässä artikkelissa ehdotetaan uutta kehystä latenttien representaatioiden optimoimiseksi upotettujen entiteettien ja relaatioiden analogisten ominaisuuksien suhteen. Muotoilemalla oppimistavoite differentioituvalla tavalla mallimme nauttii sekä teoreettisesta voimasta että laskennallisesta skaalautuvuudesta, ja se päihitti merkittävästi suuren määrän edustavia perusmenetelmiä vertailutietokannoissa. Lisäksi malli tarjoaa tyylikkään yhdistelmän useille tunnetuille monirelaation upottamisen menetelmille, joiden voidaan osoittaa olevan kehyksemme erityisiä instansseja.

**Tulos**

Analoginen päättely monisuhteisia sulautumia varten

**Esimerkki 2.1415**

Monilla aloilla on toivottavaa arvioida käyttäjien mieltymyksiä pikemminkin laadullisesti kuin määrällisesti. Tällaiset kvalitatiivisten preferenssijärjestysten esitykset ovat tärkeä osa automaattisia päätöksentekovälineitä. Ehdotamme mieltymysten graafista esitystä, joka heijastaa mieltymyslausumien ehdollista riippuvuutta ja riippumattomuutta ceteris paribus -tulkinnan (kaikki muut ovat yhtä suuria) mukaisesti. Tällainen esitys on usein kompakti ja luultavasti luonnollinen. Kuvaamme useita tähän esitystapaan perustuvia hakualgoritmeja dominointitestaukseen; nämä algoritmit ovat varsin tehokkaita erityisesti erityisten verkkotopologioiden, kuten ketju- ja puustruktuuristen verkkojen sekä polytrees-verkkojen osalta.

**Tulos**

Päättely ehdollisilla Ceteris Paribus -preferenssilausumilla

**Esimerkki 2.1416**

Tekstien semanttisen samankaltaisuuden määrittäminen on luonnollisen kielen käsittelyn keskeinen tutkimusaihe. Koska vektoripohjaiset mallit lauseiden esittämiseen käyttävät usein pinnallista tietoa, tarkan semantiikan kuvaaminen on vaikeaa. Sitä vastoin loogiset semanttiset esitystavat kuvaavat lauseen semantiikan syvempiä tasoja, mutta niiden symbolinen luonne ei tarjoa tekstin samankaltaisuuden asteittaisia käsitteitä. Ehdotamme menetelmää semanttisen tekstuaalisen samankaltaisuuden määrittämiseksi yhdistämällä pinnallisia piirteitä luonnollisen päättelyn todistuksista poimittuihin piirteisiin, jotka osoittavat, että lauseparien välillä on kaksisuuntaisia seuraussuhteita. Luonnollisen päättelyn todistuksia varten käytämme ccg2lambda-järjestelmää, joka on korkeamman asteen automaattinen päättelyjärjestelmä, joka muuntaa Combinatory Categorial Grammar (CCG) -derivaatiopuut semanttisiksi esityksiksi ja suorittaa luonnollisen päättelyn todistuksia. Kokeet osoittavat, että järjestelmämme pystyi päihittämään muut logiikkapohjaiset järjestelmät ja että todisteista johdetut piirteet ovat tehokkaita tekstuaalisen samankaltaisuuden oppimisessa.

**Tulos**

Semanttisen tekstuaalisen samankaltaisuuden määrittäminen luonnollisen deduktion todistusten avulla

**Esimerkki 2.1417**

Visuaalinen kysymysten vastaaminen (VQA) on edistynyt huomattavasti toukokuusta 2015 lähtien, sillä se on klassinen ongelma, jossa visuaaliset ja tekstimuotoiset tiedot yhdistetään järjestelmään. Monissa valaisevissa VQA-teoksissa tutkitaan syvällisesti kuvan ja kysymyksen koodauksia ja sulauttamismenetelmiä, joista huomio on tehokkain ja tehokkain mekanismi. Nykyiset huomioperusteiset menetelmät keskittyvät visuaalisten ja tekstuaalisten piirteiden riittävään fuusioimiseen, mutta niistä puuttuu huomio siihen, mihin ihmiset keskittyvät esittääkseen kysymyksiä kuvasta. Perinteiset tarkkaavaisuuteen perustuvat menetelmät kiinnittävät piirteelle yhden arvon kussakin avaruudellisessa sijainnissa, jolloin menetetään paljon hyödyllistä tietoa. Näiden ongelmien korjaamiseksi ehdotamme yleistä menetelmää, jolla voidaan suorittaa saliency-tyyppinen esivalinta päällekkäisille alueellisille piirteille kaksisuuntaisen LSTM:n (BiLSTM) keskinäisen korrelaation avulla, ja käytämme uudenlaista elementtiviisaaseen kertolaskuun perustuvaa huomiomenetelmää, jolla voidaan kaapata pätevämpää korrelaatiotietoa visuaalisten ja tekstuaalisten piirteiden välillä. Teemme kokeita laajamittaisella COCO-VQA-tietokokonaisuudella ja analysoimme mallimme tehokkuutta vahvoilla empiirisillä tuloksilla.

**Tulos**

Tehtävälähtöinen visuaalinen erottuvuus ja tarkkaavaisuuteen perustuva visuaalinen kysymyksiin vastaaminen

**Esimerkki 2.1418**

Tietokonenäön tai luonnollisen kielen käsittelyn kaltaisilla aloilla esiintyvän suuren monitulkintaisuuden käsittelemiseksi vankat ennustemenetelmät etsivät usein monipuolisen joukon korkealaatuisia ratkaisukandidaatteja tai ehdotuksia. Strukturoiduissa<lb>ennustusongelmissa tästä tulee pelottava tehtävä, koska ratkaisuavaruus (kuvien<lb>merkinnät, lauseenjäsennykset jne.) on eksponentiaalisen suuri. Tutkimme ahneita algo-<lb>ritmejä monipuolisen osajoukon ratkaisujen löytämiseksi strukturoiduissa tulosavaruuksissa<lb>nostamalla uusia yhteyksiä kombinatoristen elementti<lb>joukkojen yli olevien submodulaaristen funktioiden ja graafisia malleja varten tutkittujen korkea-asteisten potentiaalien (High-Order Potentials, HOP) välille. Erityisesti<lb>näytämme esimerkkien avulla, että kun submodulaaristen monimuotoisuusfunktioiden marginaaliset voitot<lb>sallivat strukturoidut esitykset, tämä mahdollistaa tehokkaan (sub-lineaarisessa ajassa tapahtuvan) approxi-<lb>mate-<lb>maksimoinnin pelkistämällä ahneen lisäysvaiheen päättelyksi<lb>tekijägraafissa, jossa on sopivasti konstruoidut HOPit. Keskustelemme hyödyistä, kompromisseista ja osoitamme, että rakenteemme johtavat huomattavasti parempiin ehdotuksiin.

**Tulos**

Submodulaarinen kohtaa strukturoidun: Monipuolisten osajoukkojen löytäminen eksponentiaalisesti suurista strukturoiduista kohdejoukoista.

**Esimerkki 2.1419**

Aikasarjat herättävät yhä enemmän huomiota tilastotieteen, koneoppimisen ja hahmontunnistuksen parissa, sillä ne ovat laajalti esillä sekä teollisuudessa että korkeakouluissa, mutta tehokkaassa aikasarjojen visualisoinnissa on saavutettu vain vähän edistystä niiden ajallisen ulottuvuuden ja monimutkaisen dynamiikan vuoksi. Viimeaikaiset pyrkimykset käyttää verkkomittareita aikasarjojen luonnehtimiseen luokittelua varten innoittamana esitämme lähestymistavan, jolla aikasarjat visualisoidaan monimutkaisina verkkoina, jotka perustuvat ensimmäisen asteen Markov-prosessiin ja ajalliseen järjestykseen. Toisin kuin klassiset pylväsdiagrammit, viivapiirrokset ja muut tilastoihin perustuvat kuvaajat, lähestymistapamme tarjoaa intuitiivisemman visualisoinnin, joka säilyttää paremmin sekä ajallisen riippuvuuden että taajuusrakenteet. Se tarjoaa luonnollisen käänteisoperaation, jolla kuvaaja voidaan kuvata takaisin aikasarjaan, jolloin on mahdollista käyttää kuvaajatilastoja aikasarjojen luonnehtimiseen parempaa visuaalista tarkastelua ja tilastollista analyysia varten. Kokeelliset tuloksemme osoittavat, että graafi on tehokas erilaisissa tehtävissä, kuten järjestelmien tunnistamisessa, luokittelussa ja poikkeamien havaitsemisessa sekä synteettisissä että todellisissa aikasarjadatassa.

**Tulos**

Ajallisen Markovin dynamiikan koodaaminen graafiin aikasarjojen visualisointia varten

**Esimerkki 2.1420**

Artikkelissa käsitellään metristen luokittelualgoritmien muuttamista. Siinä tutkitaan erityisesti k-Nearest Neighbours -algoritmia sen soveltamiseksi peräkkäisiin tietoihin. Lisäksi ehdotetaan metristen luokittelualgoritmien yleistämismenetelmää. Osana sitä on kehitetty algoritmi sekventiaalisen datan luokittelu- ja merkintäongelman ratkaisemiseksi. Artikkelissa käsitellään myös kehitetyn luokittelualgoritmin etuja verrattuna olemassa olevaan algoritmiin. Ehdotetun algoritmin tehokkuutta verrataan CRF-algoritmiin avoimen CoNLL2000-tietoaineiston pilkkomistehtävässä.

**Tulos**

Metristen luokittelualgoritmien yleistäminen sekvenssien luokittelua ja merkintöjä varten

**Esimerkki 2.1421**

Tämä on koneoppimissovellus, jossa käsitellään suurta dataa. Esittelemme erittäin tarkkoja ennustemenetelmiä harvinaisille tapahtumille puolistrukturoiduissa konelokitiedostoissa, joita NORC:n tietokoneavusteinen puhelinhaastatteluverkko (CATI) tuottaa suurella nopeudella ja määrällä kyselytutkimusten suorittamiseksi. Sovellamme harkitusti luonnollisen kielen käsittelytekniikoita (NLP) ja tiedonlouhintastrategioita kouluttaaksemme tehokkaita oppimis- ja ennustemalleja lokitiedostojen harvinaisten virheilmoitusten luokittelemiseksi - ilman pääsyä lähdekoodiin, päivitettyyn dokumentaatioon tai sanakirjoihin. Erityisesti yksinkertaista mutta tehokasta lähestymistapaamme, jossa ominaisuudet jaetaan etukäteen epätasapainoisesta datasta oppimiseen yhdessä naivien Bayes-luokittelijoiden kanssa, voidaan mahdollisesti yleistää valvottuun tai puolivalvottuun oppimiseen ja ennustemenetelmiin muita kriittisiä tapahtumia, kuten verkkohyökkäysten havaitsemista, varten.

**Tulos**

Koneellinen oppiminen CATI-verkon konedataa varten

**Esimerkki 2.1422**

Mielenterveysfoorumit ovat verkkoyhteisöjä, joissa ihmiset ilmaisevat asioitaan ja hakevat apua moderaattoreilta ja muilta käyttäjiltä. Tällaisilla foorumeilla on usein viestejä, joiden sisältö on vakava ja jotka osoittavat, että käyttäjä on akuutissa ahdingossa ja että on olemassa itsensä vahingoittamisyrityksen riski. Moderaattoreiden on reagoitava näihin vakaviin viesteihin ajoissa, jotta estetään mahdollinen itsensä vahingoittaminen. Päivittäin lähetettävän sisällön suuren määrän vuoksi moderaattoreiden on kuitenkin vaikea löytää näitä kriittisiä viestejä ja vastata niihin. Esitämme kehyksen, jonka avulla käyttäjien sisältö luokitellaan neljään vakavuusluokkaan, jotka on määritelty itsensä vahingoittamiseen viittaavien ajatusten perusteella. Mallimme perustuvat ominaisuuksiltaan monipuoliseen luokittelukehykseen, joka sisältää leksikaalisia, psykolingvistisiä, kontekstuaalisia ja aihepiirien mallintamiseen liittyviä ominaisuuksia. Lähestymistapamme parantavat huomattavasti mielenterveysfoorumien sisällön vakavuuden luokittelun nykytilaa (jopa 17 prosentin parannus F-1-pisteisiin verrattuna). Ehdotetun mallin avulla analysoimme käyttäjien mielentilaa ja osoitamme, että foorumin pitkäaikaiskäyttäjien riskin vakavuus vähenee ajan myötä. Analyysimme moderaattoreiden ja käyttäjien välisestä vuorovaikutuksesta osoittaa lisäksi, että ilman automaattista tapaa tunnistaa kriittinen sisältö moderaattoreiden on todellakin haastavaa vastata käyttäjille ajoissa.

**Tulos**

Sisällön vakavuuden luokittelu verkossa toimivilla mielenterveysfoorumeilla∗∗ .

**Esimerkki 2.1423**

Tässä artikkelissa esitellään multimodaalinen rekursiivinen neuroverkko (m-RNN), jonka avulla voidaan luoda uusia kuvatekstejä. Se mallintaa suoraan sanan tuottamisen todennäköisyysjakaumaa, kun otetaan huomioon aiemmat sanat ja kuva. Kuvatekstit luodaan tämän jakauman mukaisesti. Malli koostuu kahdesta aliverkosta: syvästä rekursiivisesta neuroverkosta lauseita varten ja syvästä konvoluutioverkosta kuvia varten. Nämä kaksi aliverkkoa ovat vuorovaikutuksessa toistensa kanssa multimodaalisessa kerroksessa muodostaen koko m-RNN-mallin. Mallimme tehokkuus validoidaan neljällä vertailutietoaineistolla (IAPR TC-12, Flickr 8K, Flickr 30K ja MS COCO). Mallimme on parempi kuin uusimmat menetelmät. Lisäksi sovellamme m-RNN-mallia kuvien tai lauseiden hakutehtäviin, ja saavutamme merkittävän suorituskyvyn parannuksen verrattuna uusimpiin menetelmiin, jotka optimoivat suoraan hakua varten luokittelun tavoitefunktion. Tämän työn projektisivu on: www.stat.ucla.edu/ ̃junhua.mao/m-RNN.html. 1

**Tulos**

SYVÄLLINEN TEKSTITYS MULTIMODAALISILLA TOISTUVILLA NEUROVERKOILLA (M-RNN)

**Esimerkki 2.1424**

Esittelemme itseopiskelevan lähestymistavan valvomattomaan riippuvuuksien jäsennykseen, jossa käytetään uudelleen olemassa olevia valvottuja ja valvomattomia jäsennysalgoritmeja. Lähestymistapamme on nimeltään "iterated reranking" (IR), ja se lähtee liikkeelle valvomattoman jäsentäjän tuottamista riippuvuuspuista ja parantaa näitä puita iteratiivisesti käyttämällä valvotussa jäsentelyssä käytettyjä monipuolisempia todennäköisyysmalleja, jotka puolestaan on koulutettu näillä puilla. Järjestelmämme tarkkuus on 1,8 prosenttia korkeampi kuin Spitkovskyn ym. (2013) stateof-the-art-jäsennyspalvelun tarkkuus WSJ-korpuksessa.

**Tulos**

Valvomaton riippuvuuksien jäsennys: Käytetään valvottuja jäsentäjiä

**Esimerkki 2.1425**

Tässä artikkelissa esitellään empiirinen tutkimus kahdesta laajalti käytetystä sekvenssin ennustamismallista, ehdollisista satunnaiskentistä (CRF) ja pitkistä lyhytkestoisen muistin verkoista (LSTM), kahdessa vietnaminkielisen tekstinkäsittelyn perustehtävässä, joihin kuuluvat puheosien merkitseminen ja nimettyjen entiteettien tunnistaminen. Osoitamme, että käyttämällä vain yksinkertaisia sanapohjaisia piirteitä ja minimaalista käsityönä tehtyä piirteiden suunnittelua voidaan saavuttaa vahva alaraja merkintätarkkuudelle, joka on 90,65 % ja 86,03 % näiden kahden tehtävän vakiotestisarjoissa. Erityisesti osoitamme empiirisesti sanojen upotusten yllättävän tehokkuuden molemmissa tehtävissä molemmilla malleilla. Huomautamme, että uusin LSTM-malli ei aina päihitä merkittävästi perinteistä CRF-mallia etenkään keskikokoisissa aineistoissa. Lopuksi annamme joitakin ehdotuksia ja keskusteluja sekvenssien merkintämallien tehokkaasta käytöstä käytännön sovelluksissa.

**Tulos**

Vietnamin kielen tekstinkäsittelyssä käytettävien erottelevien sekvenssien merkintämallien empiirinen tutkimus

**Esimerkki 2.1426**

Variatiivinen päättely tarjoaa tehokkaan työkalun likimääräiseen todennäköisyyspäätelmään monimutkaisissa, strukturoiduissa malleissa. Tyypilliset variationaalisen päättelyn menetelmät edellyttävät kuitenkin sellaisten päättelyverkkojen käyttöä, joiden todennäköisyystiheysfunktiot ovat laskennallisesti käsiteltävissä. Tämä rajoittaa suuresti variationaalisten päättelymenetelmien suunnittelua ja toteuttamista. Tarkastelemme villejä variationaalisia päättelymenetelmiä, jotka eivät edellytä päättelyverkkojen tiheysfunktioita, ja joita voidaan näin ollen soveltaa haastavammissa tapauksissa. Sovellusesimerkkinä käsittelemme stokastista gradientti-Langevin-dynamiikkaa (SGLD) päättelyverkkona ja käytämme menetelmiämme SGLD:n askelkoon automaattiseen säätämiseen, mikä tuottaa merkittävää parannusta käsin suunniteltuihin askelkokojärjestelmiin verrattuna.

**Tulos**

KAKSI MENETELMÄÄ VILLIEN MUUNNOSTEN PÄÄTTELYYN

**Esimerkki 2.1427**

Univaiheiden malli antaa tärkeitä vihjeitä unihäiriön diagnosoinnissa. Analysoimalla univaiheiden kuvioita ja poimimalla niiden piirteitä EEG-, EOG- ja EMG-signaaleista voimme luokitella univaiheet. Tässä tutkimuksessa esitellään uusi luokittelumalli, jolla ennustetaan univaiheita suurella tarkkuudella. Pääajatuksena on yhdistää Deep Belief Networkin (DBN) generatiivinen kyky ja LSTM:n (Long Short-term Memory) erottelukyky ja sekvenssikuvioiden tunnistuskyky. Käytämme DBN:ää, jota käsitellään automaattisena korkeamman tason ominaisuuksien luojana. DBN:n syötteenä on 28 "käsintehtyä" piirrettä, joita on käytetty aiemmissa univaihetutkimuksissa. Vertailimme menetelmäämme muihin tekniikoihin, jotka yhdistivät DBN:n ja Piilotetun Markovin mallin (Hidden Markov Model, HMM).Tässä tutkimuksessa hyödynnämme uniaineiston sekvenssi- tai aikasarjaominaisuuksia. Tietojemme mukaan suurin osa nykyisistä polysomnogrammista tehdyistä unianalyyseistä perustuu luokittelussa vain yksittäiseen instanssiin (ei-sekvenssi). Tässä tutkimuksessa käytimme kahta tietokokonaisuutta: avointa tietokokonaisuutta, jota käsitellään vertailukohtana, ja toista tietokokonaisuutta, joka on univaiheita koskeva tietokokonaisuutemme (ladattavissa), jolla tulokset voidaan todentaa tarkemmin. Kokeemme osoittivat, että DBN:n ja LSTM:n yhdistelmä antaa paremman kokonaistarkkuuden 98,75 % (Fscore=0,9875) vertailutietokannassa ja 98,94 % (Fscore=0,9894) MKG-tietokannassa. Tämä tulos on parempi kuin univaiheiden luokittelun nykytaso, joka oli 91,31 %i. Avainsanat - univaiheiden luokittelu, pitkä lyhytkestoinen muisti, syvä uskomusverkko, syväoppiminen.

**Tulos**

Generatiivisten ja diskriminoivien neuroverkkojen yhdistäminen univaiheiden luokittelussa

**Esimerkki 2.1428**

Valvotussa binäärihaketuksessa halutaan oppia funktio, joka kuvaa korkea-ulotteisen ominaisuusvektorin binäärikoodien vektoriksi, jotta sitä voidaan soveltaa nopeaan kuvien hakuun. Tämä johtaa tyypillisesti vaikeaan optimointiongelmaan, joka on epäkonveksinen ja epäsilevä, koska kyseessä ovat diskreetit muuttujat. Monissa töissä ongelmaa on yksinkertaisesti höllennetty harjoittelun aikana, ratkaistu jatkuva optimointi ja typistetty koodit jälkikäteen. Tämä antaa kohtuullisia tuloksia, mutta on melko epäoptimaalista. Viimeaikaisessa työssä on yritetty optimoida tavoite suoraan binäärikoodien yli ja saavutettu parempia tuloksia, mutta hash-funktio opittiin edelleen jälkikäteen, mikä on edelleen epäoptimaalista. Ehdotamme yleistä kehystä hash-funktioiden oppimiseen affiniteettipohjaisten häviöfunktioiden avulla, joissa käytetään apukoordinaatteja. Tämä sulkee silmukan ja optimoi yhdessä hash-funktioiden ja binäärikoodien yli niin, että ne vähitellen vastaavat toisiaan. Tuloksena syntyvä algoritmi voidaan nähdä korjattuna, iteroituna versiona menettelystä, jossa ensin optimoidaan koodit ja sitten opitaan hash-funktio. Tähän verrattuna optimoinnillamme saadaan taatusti parempia hash-funktioita, mutta se ei ole paljon hitaampi, kuten on kokeellisesti osoitettu erilaisilla valvotuilla tietokokonaisuuksilla. Lisäksi kehyksemme helpottaa optimointialgoritmien suunnittelua mielivaltaisille häviö- ja hash-funktiotyypeille.

**Tulos**

Affiniteettipohjaisen binäärihaketuksen optimointi apukoordinaattien avulla

**Esimerkki 2.1429**

Tiivistelmä Tässä artikkelissa tarkastellaan uudelleen hajamielisen kuljettajan ongelmaa, jossa hänen on valittava eri vaihtoehtoja, joilla on erilainen hyöty, kun muistaminen on puutteellista ja järjestelmän tuntemus vaihtelee. Klassinen hajamielisen kuljettajan ongelma edustaa tapausta, jossa tieto on rajallista, ja sillä on merkitystä viestinnän ja oppimisen, sosiaalisten valintojen, mekanismien suunnittelun, huutokauppojen, tiedon, uskomusten ja rationaalisen toimijuuden teorioiden yleisten alojen kannalta. Laajojen pelien puitteissa tällä ongelmalla on sovelluksia moniin tekoälyskenaarioihin. On selvää, että agentin suorituskyky paranee, kun käytettävissä oleva tieto lisääntyy. Osoitetaan, että peräkkäisten valintojen epäyhtenäinen jakostrategia menestyy paremmin kuin kiinteän todennäköisyyden strategia. Tarkastelemme sekä klassisia että kvanttilähestymistapoja ongelmaan. Väitämme, että kvanttipäätösten ylivoimaista suorituskykyä, jossa on pääsy kietoutuneisuuteen, ei voi verrata oikeudenmukaisesti klassiseen algoritmiin. Jos oletetaan, että agenttien kognitiivisilla järjestelmillä on pääsy kvanttiresursseihin tai että niillä on kvanttimekaaninen perusta, tämä voidaan hyödyntää ylivoimaiseen suorituskykyyn.

**Tulos**

Poissaolevan kuljettajan ongelma Redux Redux

**Esimerkki 2.1430**

Parsimony, mukaan lukien harvinaisuus ja matala arvojärjestys, on osoittautunut menestyksekkääksi datan mallintamiseksi lukuisissa koneoppimis- ja signaalinkäsittelytehtävissä. Perinteisesti tällaiset mallinnusmenetelmät perustuvat iteratiiviseen algoritmiin, joka minimoi tavoitefunktion, jossa on parsimoniaa edistäviä termejä. Iteratiivisen optimoinnin luontaisesti peräkkäinen rakenne ja datasta riippuva monimutkaisuus ja viive ovat merkittävä rajoitus monissa sovelluksissa, jotka vaativat reaaliaikaista suorituskykyä tai joihin liittyy suuria datamääriä. Toinen näiden mallinnustekniikoiden kohtaama rajoitus on vaikeus sisällyttää niitä diskriminoiviin oppimisskenaarioihin. Tässä työssä ehdotamme, että painopiste siirretään mallista etsintäalgoritmiin, ja kehitämme prosessikeskeisen näkemyksen yksinkertaisesta mallintamisesta, jossa iteratiivisen optimoinnin sijasta käytetään opittua determinististä kiinteän kompleksisuuden etsintäprosessia. Näytämme periaatteellisen tavan rakentaa opittavissa olevia pursuit-prosessiarkkitehtuureja strukturoiduille harvalukuisille ja kestäville matalarivisille malleille, jotka on johdettu proksimaalisen laskeutumisen algoritmien iteraatiosta. Nämä arkkitehtuurit oppivat lähestymään tarkkaa parsimonista esitystä murto-osalla tavanomaisten optimointimenetelmien monimutkaisuudesta. Osoitamme myös, että sopivilla koulutusjärjestelmillä voidaan luonnollisesti laajentaa parsimonious-malleja diskriminoiviin asetuksiin. Huipputason tuloksia osoitetaan useissa haastavissa kuvan- ja äänenkäsittelyn ongelmissa, joissa nopeutuminen on useita kertaluokkia tarkkoihin optimointialgoritmeihin verrattuna. ∗P. Sprechmann ja G. Sapiro työskentelevät sähkö- ja tietotekniikan laitoksella, Duken yliopisto, Durham 27708, Yhdysvallat. Sähköposti: pablo.sprechmann@duke.edu, guillermo.sapiro@duke.edu. †A. M. Bronsteind on School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel.Sähköposti: bron@eng.tau.ac.il. ‡Työtä ovat osittain tukeneet NSF, ONR, NGA, DARPA, AFOSR, ARO ja BSF. 1 ar X iv :1 21 2. 36 31 v1 [ cs .L G ] 1 4 D ec 2 01 2

**Tulos**

Tehokkaiden harvalukuisten ja matalalla sijalla olevien mallien oppiminen

**Esimerkki 2.1431**

Esittelemme käytännöllisen, eriytetysti yksityisen algoritmin, jolla voidaan vastata suureen määrään kyselyitä suurten ulottuvuuksien tietokokonaisuuksissa. Kuten kaikilla tähän tehtävään tarkoitetuilla algoritmeilla, myös meidän algoritmimme pahimman tapauksen monimutkaisuus on välttämättä eksponentiaalinen datan ulottuvuuden suhteen. Algoritmimme paketoi kuitenkin laskennallisesti vaikean vaiheen tiiviisti määritellyksi kokonaislukuohjelmaksi, joka voidaan ratkaista ei-yksityisesti tavallisilla ratkaisijoilla. Todistamme algoritmillemme tarkkuus- ja yksityisyysteoriat ja osoitamme sitten kokeellisesti, että algoritmimme toimii hyvin käytännössä. Algoritmimme pystyy esimerkiksi vastaamaan tehokkaasti ja tarkasti miljooniin kyselyihin Netflix-tietokannassa, jossa on yli 17 000 attribuuttia; tämä on moninkertainen parannus nykyiseen tekniikan tasoon verrattuna.

**Tulos**

Kaksoiskysely: Käytännöllinen yksityisen kyselyn julkaisu korkea-ulotteisille tiedoille

**Esimerkki 2.1432**

Esittelemme menetelmän, jolla voidaan luoda uusia dialogijärjestelmiä hyvin pienistä määristä kommentoimatonta dialogidataa, ja osoitamme, miten sanatason tutkimus käyttäen vahvistusoppimista (RL) yhdistettynä inkrementaaliseen ja semanttiseen kielioppiin Dynamic Syntax (DS) mahdollistaa järjestelmien löytämisen, tuottamisen ja ymmärtämisen monista uusista dialogivaihtoehdoista. Menetelmällä vältetään kalliiden ja aikaa vievien dialogitekojen annotaatioiden käyttö ja tuetaan luonnollisempia (inkrementaalisia) dialogeja kuin vuoropohjaiset järjestelmät. Tässä yhteydessä kielen tuottamista ja dialogin hallintaa käsitellään yhteisenä päätöksenteko-/optimointiongelmana, ja RL:n MDP-malli rakennetaan automaattisesti. Toteutetun järjestelmän avulla osoitamme, että tämä menetelmä mahdollistaa laajan valikoiman vuoropuheluvariaatioiden automaattisen kuvaamisen, vaikka järjestelmä koulutettaisiinkin vain yhdestä vuoropuhelusta. Vaihtoehtoihin kuuluvat kysymys-vastausparit, yli- ja alivastaaminen, itse- ja muukalaiskorjaukset, selventävä vuorovaikutus, jaetut lausahdukset ja ellipsit. Tämä yleistysominaisuus johtuu DS-kieliopin rakenteellisesta tietämyksestä ja rajoituksista, ja se tuo esiin joitakin rajoituksia, joita viimeaikaiset, pelkästään koneoppimistekniikoita käyttäen rakennetut järjestelmät ovat sisältäneet.

**Tulos**

Inkrementaalisten dialogijärjestelmien käynnistäminen: kielellisen tiedon käyttö minimaalisesta datasta oppimiseen.

**Esimerkki 2.1433**

Tässä asiakirjassa ehdotamme hajautettuja puun ytimiä (DTK) uudeksi menetelmäksi, jolla voidaan vähentää puun ytimien aika- ja tilakompleksisuutta. Käyttämällä lineaarisen monimutkaisuuden algoritmia puiden vektoreiden laskemiseen upotamme puunpätkien ominaisuusavaruudet matalaulotteisiin avaruuksiin, joissa ytimen laskenta tehdään suoraan pistetuotteella. Osoitamme, että DTK:t ovat nopeampia, korreloivat puun ytimien kanssa ja saavuttavat tilastollisesti samanlaisen suorituskyvyn kahdessa luonnollisen kielen käsittelytehtävässä.

**Tulos**

Hajautetut puuytimet

**Esimerkki 2.1434**

Potentiaalipelit ja hajautetut osittain havaittavissa olevat MDP:t (Dec-POMDP) ovat kaksi yleisesti käytettyä moniagenttisen vuorovaikutuksen mallia staattista optimointia ja peräkkäistä päätöksentekoa varten. Tässä artikkelissa esitellään suodatettu fiktiivinen peli toistuvien potentiaalipelien ratkaisemiseksi, joissa kunkin pelaajan havaintoja toisten pelaajien toimista häiritsee satunnainen kohina, ja käytetään tätä algoritmia rakentamaan online-oppimismenetelmä Dec-POMDP:iden ratkaisemiseksi. Erityisesti todistamme, että havaintojen kohina estää tavallista fiktiivistä peliä konvergoitumasta Nashin tasapainoon potentiaalisissa peleissä, mikä tekee fiktiivisestä pelistä epäkäytännöllisen myös Dec-POMDP-pelien ratkaisemisessa. Tämän ongelman ratkaisemiseksi johdamme suodatetun fiktiivisen pelin ja annamme ehdot, joiden vallitessa se konvergoi Nash-tasapainoon potentiaalisissa peleissä, joissa on kohinaisia havaintoja. Tämän jälkeen käytämme suodatettua fiktiivistä peliä Dec-POMDP-pelien ratkaisijan rakentamiseen ja osoitamme uuden algoritmimme suorituskyvyn laatikon työntämisongelmassa. Tuloksemme osoittavat, että olemme jatkuvasti suorituskykyisempiä kuin nykyaikainen Dec-POMDP-ratkaisija, keskimäärin 100 %:lla koko havaintofunktion kohinan vaihteluvälillä.

**Tulos**

Suodatettu fiktiivinen peli häiriintyneiden havaintopotentiaalipelien ja hajautettujen POMDP-pelien osalta.

**Esimerkki 2.1435**

Lifting pyrkii nopeuttamaan todennäköisyyspäätelmien tekemistä hyödyntämällä mallin symmetrisyyttä. Tarkat nostetut päättelymenetelmät, kuten niiden propositionaaliset vastineet, toimivat jakamalla malli ja ongelma rekursiivisesti. Propositionaalisessa tapauksessa on olemassa muodollisia rakenteita, kuten hajotuspuut (dtrees), jotka edustavat tällaista hajotusta ja joiden avulla voimme määrittää päättelyn monimutkaisuuden etukäteen. Tällä hetkellä ei kuitenkaan ole olemassa vastaavaa rakennetta eikä analogisia kompleksisuustuloksia nostettua päättelyä varten. Tässä artikkelissa esitellään FO-puut, jotka päivittävät propositionaaliset d-puut ensimmäisen kertaluvun tasolle. Näytämme, miten nämä puut voivat luonnehtia nostetun päättelyn ratkaisua todennäköisyyspohjaiselle loogiselle mallille (nostettujen operaatioiden sarjana), ja teemme teoreettisen analyysin nostetun päättelyn monimutkaisuudesta puun uuden nostetun leveyden käsitteen avulla.

**Tulos**

Ensimmäisen kertaluvun hajotuspuut

**Esimerkki 2.1436**

Esittelemme uudenlaisen kehyksen, jonka avulla voidaan palauttaa likimääräisesti datamatriisit, jotka ovat matalarivisiä graafeissa, otantamittauksista. Tällaisten matriisien rivit ja sarakkeet kuuluvat niiden rivien ja sarakkeiden välille muodostettujen graafien ensimmäisten ominaisvektoreiden alueeseen. Hyödynnämme tätä ominaisuutta, jotta voimme palauttaa epälineaariset matalarivirtaiset rakenteet tehokkaasti otantamittauksista pienin kustannuksin (lineaarisesti n:n suhteen). Ensin otetaan käyttöön RIP-ehto (Resrtricted Isometry Property) tällaisten matriisien rivien ja sarakkeiden tehokasta tasaista näytteenottoa varten, joka perustuu graafien ominaisvektoreiden kumulatiiviseen koherenssiin. Toiseksi ehdotetaan huipputason nopeaa matalan rankin palautusmenetelmää näytteeksi otetuille tiedoille. Lopuksi esitellään useita tehokkaita, rinnakkaisia ja parametrittomia dekoodereita sekä niiden teoreettinen analyysi matalarivi- ja klusteri-indikaattorien dekoodaamiseksi koko datamatriisin osalta. Tällä tavoin voitamme tavanomaisten lineaaristen low-rank-arvojen palautusmenetelmien laskennalliset rajoitukset suurissa tietokokonaisuuksissa. Menetelmäämme voidaan pitää myös merkittävänä askeleena kohti epälineaaristen low-rank-rakenteiden tehokasta talteenottoa. Menetelmämme nopeuttaa n × p kokoisen matriisin palauttamista yhden ytimen koneella p/k nopeammin kuin Robust Principal Component Analysis (RPCA), jossa k p on aliavaruuden ulottuvuus. Numeerisesti voimme palauttaa matalan rankin matriisin, jonka koko on 10304 × 1000, 100 kertaa nopeammin kuin Robust PCA.

**Tulos**

Kompressiivinen PCA matalien matriisien matriisien matriisien kuvaajissa (Compressive PCA)

**Esimerkki 2.1437**

Sovellamme maksimaalisen entropian periaatetta valitaksemme ainutlaatuisen yhteisen todennäköisyysjakauman kaikkien uskottavuusverkon määrittelemien yhteisten todennäköisyysjakaumien joukosta. Yksityiskohtaisesti aloitamme osoittamalla, että Bayes-puun ainutlaatuinen yhteinen jakauma on yhteneväinen sen ehdollisten jakaumien maksimaalisen entropian mallin kanssa. Tämä tulos ei kuitenkaan enää päde yleisille Bayes-verkoille. Esitämme siis uudenlaiset maksimaalisen entropian mallit, jotka lasketaan peräkkäin. Tämän jälkeen osoitamme, että kaikille yleisille Bayes-verkoille sekventiaalinen maksimi-entropiamalli on yhteneväinen ainoan yhteisen jakauman kanssa. Lisäksi sovellamme uutta peräkkäisen maksimietropian periaatetta in terval Bayesin verkkoihin ja yleisemmin uskottaviin verkkoihin. Osoitamme erityisesti, että tämä sovellus vastaa useita pieniä paikallisia entropian maksimointeja.

**Tulos**

Credal Networks under Maximum Entropy

**Esimerkki 2.1438**

Arvioimme merkkipohjaiseen sekvenssistä sekvenssiin -malliin perustuvaa semanttista jäsentäjää SemEval2017-tapahtuman yhteydessä, jossa käsiteltiin AMR-ohjelmien semanttista jäsentämistä. Tietojen lisäämisen, superhahmojen ja POS-taggauksen avulla saavutamme huomattavia parannuksia suorituskykyyn verrattuna perustason merkkitason malliin. Vaikka parannamme aiempia merkkipohjaisia neuraalisia semanttisia jäsennysmalleja, kokonaistarkkuus on edelleen alhaisempi kuin uusimmalla AMR-jäsennysmallilla. Kokonaisuus, jossa neurosemanttinen jäsentäjämme yhdistetään olemassa olevaan perinteiseen jäsentäjään, parantaa suorituskykyä hieman.

**Tulos**

The Meaning Factory SemEval-2017 -tapahtumassa Tehtävä 9: AMR:ien tuottaminen neuraalisen semanttisen jäsennyksen avulla

**Esimerkki 2.1439**

Halpernin ja Pearlin määrittelemä todellisen kausaalisuuden teoria ja sen määrällinen mittari - vastuun aste - on osoittautunut erittäin hyödylliseksi tietojenkäsittelytieteen eri osa-alueilla, koska sen tuottamat tulokset ja intuitiomme vastaavat hyvin toisiaan. Tässä artikkelissa kuvaan kausaalisuuden sovelluksia muodolliseen verifiointiin, nimittäin vastaesimerkkien selittämiseen, kattavuusmittareiden tarkentamiseen ja symbolisen liikeradan arviointiin. Käsittelen myös lyhyesti kausaalisuuden viimeaikaisia sovelluksia oikeudelliseen päättelyyn.

**Tulos**

Syy-yhteys ja vastuu formaalissa todentamisessa ja sen ulkopuolella

**Esimerkki 2.1440**

Elinikäinen vahvistusoppiminen tarjoaa lupaavan kehyksen sellaisten monipuolisten agenttien kehittämiseen, jotka voivat kerätä tietoa elinikäisen kokemuksen aikana ja oppia nopeasti uusia tehtäviä aikaisemman tiedon pohjalta. Nykyiset elinikäisen oppimisen menetelmät eivät kuitenkaan ole katuvia kokemuksen määrän kasvaessa, ja niihin sisältyy rajoituksia, jotka voivat johtaa epäoptimaalisiin tai vaarallisiin ohjauskäytäntöihin. Näiden ongelmien ratkaisemiseksi kehitämme elinikäisen politiikan gradienttioppijan, joka toimii vastakkaisessa ympäristössä oppiakseen useita tehtäviä verkossa ja asettaen samalla turvallisuusrajoituksia opituille politiikoille. Osoitamme ensimmäistä kertaa, että elinikäinen politiikan haku on epälineaarista, ja validoimme algoritmimme useilla dynaamisilla vertailujärjestelmillä ja sovelluksella, joka koskee nelikopterin ohjausta.

**Tulos**

Turvallinen toimintaperiaatteiden haku elinikäiseen vahvistusoppimiseen, jossa on epälineaarinen katumushäiriö

**Esimerkki 2.1441**

Vaikka syväoppiminen on johtanut huomattaviin edistysaskeliin monissa eri sovelluksissa, sillä on vaikeuksia aloilla, joilla tietojen jakauma muuttuu oppimisen aikana. Sen sijaan biologiset neuroverkot sopeutuvat jatkuvasti muuttuviin alueisiin mahdollisesti hyödyntämällä monimutkaisia molekyylikoneistoja monien tehtävien ratkaisemiseksi samanaikaisesti. Tässä tutkimuksessa esitellään älykkäitä synapseja, jotka tuovat osan tästä biologisesta monimutkaisuudesta keinotekoisiin neuroverkkoihin. Kukin synapsi kerää ajan mittaan tehtävään liittyvää tietoa ja hyödyntää tätä tietoa uusien muistojen nopeaan tallentamiseen unohtamatta vanhoja. Arvioimme lähestymistapaamme luokittelutehtävien jatkuvassa oppimisessa ja osoitamme, että se vähentää huomattavasti unohtumista säilyttäen samalla laskennallisen tehokkuuden.

**Tulos**

Jatkuva oppiminen synaptisen älykkyyden avulla

**Esimerkki 2.1442**

Vaikka havaintotehtävissä, kuten kuvien luokittelussa, on tapahtunut edistystä, tietokoneet suoriutuvat edelleen huonosti kognitiivisista tehtävistä, kuten kuvien kuvaamisesta ja kysymyksiin vastaamisesta. Kognitio on keskeistä tehtävissä, joissa ei ole kyse vain visuaalisen maailman tunnistamisesta vaan myös sen päättelystä. Kognitiivisissa tehtävissä käytettävät mallit, joita käytetään kuvien rikkaan sisällön käsittelyyn, koulutetaan kuitenkin edelleen samoilla tietokokonaisuuksilla, jotka on suunniteltu havaintotehtäviä varten. Jotta mallit menestyisivät kognitiivisissa tehtävissä, niiden on ymmärrettävä Ranjay Krishna Stanford University, Stanford, CA, USA E-mail: ranjaykrishna@cs.stanford.edu Yuke Zhu Stanford University, Stanford, CA, USA Oliver Groth Dresden University of Technology, Dresden, Saksa Justin Johnson Stanford University, Stanford, CA, USA Kenji Hata Stanford University, Stanford, CA, USA Joshua Kravitz Stanford University, Stanford, CA, USA Stephanie Chen Stanford University, Stanford, CA, USA Yannis Kalantidis Yahoo Inc., San Francisco, CA, USA Li-Jia Li Snapchat Inc., Los Angeles, CA, USA David A. Shamma Yahoo Inc., San Francisco, CA, USA Michael S. Bernstein Stanford University, Stanford, CA, USA Li Fei-Fei Stanford University, Stanford, CA, USA kuva. Kun kysytään "Millä ajoneuvolla henkilö ajaa?", tietokoneiden on tunnistettava kuvassa olevat kohteet sekä suhteet ratsastus(mies, vaunut) ja veto(hevonen, vaunut), jotta ne voivat vastata oikein, että "henkilö ajaa hevosvetoisilla vaunuilla". Tässä artikkelissa esittelemme Visual Genome -tietokannan, jonka avulla tällaisten suhteiden mallintaminen on mahdollista. Keräämme tiheitä merkintöjä objekteista, attribuuteista ja suhteista kussakin kuvassa näiden mallien oppimista varten. Tietokantamme sisältää yli 100 000 kuvaa, joissa kussakin kuvassa on keskimäärin 21 objektia, 18 attribuuttia ja 18 pareittaista objektien välistä suhdetta. Kanonisoimme kohteet, attribuutit, suhteet ja substantiivilauseet aluekuvauksissa ja kysymysten vastauspareissa WordNetin synseteiksi. Yhdessä nämä merkinnät edustavat tiheintä ja laajinta kuvien kuvausten, objektien, attribuuttien, suhteiden ja kysymysten vastausten aineistoa.

**Tulos**

Kielen ja näön yhdistäminen käyttämällä joukkoistettuja tiheitä kuvamerkintöjä (Crowdsourced Dense Image Annotations)

**Esimerkki 2.1443**

Neuraalisen konekääntämisen tavoitteena on rakentaa yksi suuri neuroverkko, joka voidaan kouluttaa maksimoimaan käännössuorituskyky. Huomiomekanismilla varustetun koodaaja-dekooderiarkkitehtuurin avulla saavutetaan nykyisiin huipputason fraasipohjaisiin järjestelmiin verrattavissa oleva käännössuorituskyky. Suuren sanaston käytöstä tulee kuitenkin pullonkaula sekä koulutuksessa että suorituskyvyn parantamisessa. Tässä artikkelissa ehdotetaan uutta arkkitehtuuria, joka oppii morfologiaa käyttämällä kahta toistuvaa verkkoa ja hierarkkista dekooderia, joka kääntää merkkitasolla. Näin saadaan aikaan syvä merkkitason malli, joka koostuu kuudesta toistuvasta verkosta. Tällaisella syvällä mallilla on kaksi suurta etua. Se välttää suuren sanavaraston ongelman radikaalisti, ja samalla se on tehokkaampi koulutuksessa kuin sanapohjaiset mallit. Mallimme saa korkeamman BLEU-pistemäärän kuin bpe-pohjainen malli yhden epookin harjoittelun jälkeen En-Fr- ja En-Cs-käännöstehtävissä. Lisäanalyysit osoittavat, että mallimme pystyy oppimaan morfologiaa.

**Tulos**

SYVÄ MERKKITASON NEURAALINEN KONEKÄÄNTÄMINEN OPPIMALLA MORFOLOGIAA

**Esimerkki 2.1444**

Optunity on ilmainen ohjelmistopaketti, joka on omistettu hyperparametrin optimointiin. Se sisältää erityyppisiä ratkaisijoita, jotka vaihtelevat suuntaamattomista menetelmistä suorahakuun, hiukkasparvi- ja evoluutiooptimointiin. Suunnittelussa keskitytään helppokäyttöisyyteen, joustavuuteen, koodin selkeyteen ja yhteentoimivuuteen olemassa olevien ohjelmistojen kanssa kaikissa koneoppimisympäristöissä. Optunity on kirjoitettu Python-kielellä, ja se sisältää rajapintoja ympäristöihin, kuten R:ään ja MATLABiin. Optunity käyttää BSD-lisenssiä ja on vapaasti saatavilla verkossa osoitteessa http://www.optunity.net.

**Tulos**

Helppo hyperparametrin haku Optunityn avulla

**Esimerkki 2.1445**

Monien valvotun oppimisen algoritmien, kuten neuroverkkojen, huonona puolena on se, että ne tarvitsevat luonnostaan suuren määrän harjoitusdataa (Benediktsson et al., 1993). Vaikka suurista datamääristä puhutaankin paljon, luokittelun tekeminen pienestä datamäärästä on edelleen ongelma. Vaikka muut menetelmät, kuten tukivektorikoneet, pystyvät käsittelemään vähän näytteitä, ne ovat luonnostaan binääriluokittelijoita (Cortes ja Vapnik, 1995), ja ne tarvitsevat oppimisstrategioita, kuten One vs All (yksi vastaan kaikki), kun kyseessä on moniluokitus. Kun luokkia on paljon, tämä voi muodostua ongelmalliseksi. Tässä artikkelissa esitellään uusi lähestymistapa valvottuun oppimiseen klusterointimenetelmän avulla. Toisin kuin perinteiset menetelmät, kuten K-Means (MacQueen, 1967), Gravitational Clustering ei vaadi klusterien alkuperäistä määrää, vaan se muodostaa klusterit automaattisesti, yksittäisiä näytteitä voidaan painottaa mielivaltaisesti ja se vaatii vain vähän näytteitä ja on samalla joustava ylisovitusta vastaan. Avainsanat: koneoppiminen, luokittelu, klusterointi.

**Tulos**

Johdatus gravitaatioklustereihin

**Esimerkki 2.1446**

Inferotemporaalisen aivokuoren (IT) neuronipopulaatiot ylläpitävät eksplisiittistä koodia kohteen identiteetille, joka sietää myös kohteen ulkonäön muutoksia, kuten sijaintia, mittakaavaa ja katselukulmaa [1, 2, 3]. Vaikka oppimissääntöjä ei tunneta, viimeaikaiset tulokset [4, 5, 6] viittaavat siihen, että käytössä on ajallisiin assosiaatioihin perustuva valvomaton menetelmä, esim. Foldiakin jäljityssääntö [7]. Tällaiset menetelmät hyödyntävät visuaalisen maailman ajallista jatkuvuutta olettamalla, että lyhyiden aikajaksojen visuaaliset kokemukset ovat yleensä identiteetiltään muuttumattomia. Näin ollen yhdistämällä lähekkäisten aikojen kuvien representaatioita voidaan saavuttaa representaatio, joka sietää kaikki videossa tapahtuneet muutokset. Monet aiemmat tutkimukset ovat osoittaneet, että tällaiset säännöt voivat toimia yksinkertaisissa tilanteissa, joissa ei ole taustahäiriöitä, mutta visuaalisen häiriön esiintyminen on ollut ongelmallista tälle lähestymistavalle. Tässä osoitamme, että suuriin luokkakohtaisiin suodattimiin (malleihin) perustuvalla ajallisella yhdistämisellä vältetään häiriöongelma. Järjestelmämme oppii valvomattomasti internetistä kerätyistä luonnollisista videoista ja pystyy suorittamaan vaikean kasvontunnistustehtävän luonnollisissa kuvissa ilman rajoituksia: Labeled Faces in the Wild [8]. Tätä työtä on tukenut Center for Brains, Minds and Machines (CBMM), jota rahoittaa NSF STC -palkinto CCF 1231216. ar X iv :1 40 9. 38 79 v2 [ cs .C V ] 2 4 A pr 2 01 5

**Tulos**

Sotkua kestävien visuaalisten representaatioiden valvomaton oppiminen luonnollisista videoista

**Esimerkki 2.1447**

Hienojakoinen oliotyyppien luokittelu (FETC) on tehtävä, jossa olio luokitellaan laajaan joukkoon oliotyyppejä. Etävalvontaparadigmaa käytetään laajalti tämän tehtävän harjoitusaineiston tuottamiseen. Luodussa harjoitusaineistossa annetaan kuitenkin sama joukko merkintöjä jokaiselle entiteetin maininnalle ottamatta huomioon sen paikallista kontekstia. Nykyisissä FETC-järjestelmissä on kaksi merkittävää haittaa: oletetaan, että harjoitusaineisto on kohinatonta, ja käytetään käsityönä laadittuja piirteitä. Meidän työmme voittaa molemmat haitat. Ehdotamme neuroverkkomallia, joka oppii yhdessä entiteettien maininnat ja niiden kontekstin esittämisen, jotta käsinmuotoiltuja piirteitä ei enää käytettäisi. Mallimme käsittelee harjoitusdataa kohinaisena ja käyttää saranahäviöfunktion ei-parametrista muunnosta. Kokeet osoittavat, että ehdotettu malli päihittää aiemmat nykyaikaiset menetelmät kahdessa julkisesti saatavilla olevassa tietokokonaisuudessa, nimittäin FIGER(GOLD) ja BBN, sillä mikro-F1-pistemäärä paranee keskimäärin 2,69 prosenttia. Mallimme yhdessä tietokokonaisuudessa oppima tieto voidaan siirtää muihin tietokokonaisuuksiin, kun käytetään samaa mallia tai muita FETC-järjestelmiä. Nämä tiedonsiirtomenetelmät parantavat entisestään vastaavien mallien suorituskykyä.

**Tulos**

Hienojakoinen entiteettityyppien luokittelu oppimalla yhdessä representaatioita ja tarrojen sulautuksia.

**Esimerkki 2.1448**

Tässä artikkelissa ehdotetaan tehokasta algoritmia (HOLRR) sellaisten regressiotehtävien käsittelyyn, joissa ulostuloilla on tensorirakenne. Muotoilemme regressio-ongelman pienimmän neliön kriteerin minimoimiseksi monilineaarisen rank-rajoituksen alaisena, mikä on vaikea ei-konveksinen ongelma. HOLRR laskee tehokkaasti likimääräisen ratkaisun tähän ongelmaan, ja sillä on vankat teoreettiset takuut. Lisäksi esitellään kernel-laajennus. Synteettisillä ja todellisilla aineistoilla tehdyt kokeet osoittavat, että HOLRR päihittää monimuuttuja- ja monilineaariset regressiomenetelmät ja on huomattavasti nopeampi kuin nykyiset tensorimenetelmät.

**Tulos**

Korkeamman asteen matalaraajainen regressio

**Esimerkki 2.1449**

Tässä artikkelissa tutkimme uutta sovellusta, 3D-kohtausten ymmärtämistä, joka on todennäköisesti merkittävä mobiilialalla lähitulevaisuudessa. Tutkimuksen tavoitteena on lyhentää suoritusaikaa ja samalla täyttää tuloksen laatua koskevat tavoitteemme. Aiemmassa työssä osoitimme ensimmäistä kertaa, että tämä sovellus on mahdollista toteuttaa sulautetuissa järjestelmissä, joissa on tehorajoituksia, ja korostimme, että algoritmisella suunnittelutasolla tehdyillä päätöksillä on suurin vaikutus. Koska algoritmisuunnitteluavaruus on liian suuri, jotta sitä voitaisiin arvioida tyhjentävästi, käytämme hyvien algoritmisuunnitelmien löytämiseen aiemmin esiteltyä monitavoitteista satunnaismetsä-aktiivista oppimisennustuskehystä, jota kutsutaan HyperMapperiksi. Osoitamme, että HyperMapper yleistyy viimeisimmän huippuluokan 3D-kohtausten ymmärtämiseen tarkoitetulla algoritmilla ja nykyaikaisella GPU-pohjaisella tietokonearkkitehtuurilla. HyperMapper pystyy päihittämään asiantuntijan, joka virittää algoritmiparametrit automaattisesti ja käsin tässä työssä tarkasteltavien tietokonenäkösovellusten luokassa. Lisäksi käytämme joukkoistamista 3D-kohtausten ymmärtämiseen tarkoitetun Android-sovelluksen avulla osoittaaksemme, että sulautetussa järjestelmässä saatua Pareto-rintamaa voidaan käyttää saman sovelluksen nopeuttamiseen kaikissa 83 älypuhelimessa ja tabletissa, jolloin nopeus kasvaa 2x-12x.

**Tulos**

Algoritmin suorituskyvyn ja tarkkuuden välinen kompromissi 3D-näkymäsovelluksissa HyperMapperin avulla

**Esimerkki 2.1450**

Poikkeavuuksien havaitseminen on tärkeä tehtävä monissa reaalimaailman sovelluksissa, kuten petosten havaitsemisessa, epäilyttävän toiminnan havaitsemisessa, terveydenhuollon seurannassa jne. Tässä artikkelissa käsittelemme tätä ongelmaa valvotun oppimisen näkökulmasta verkko-oppimisympäristössä. Maksimoimme tunnetun Gmean-mittarin luokan epätasapainon oppimiseen verkko-oppimisympäristössä. Erityisesti osoitamme, että Gmean-arvon maksimointi vastaa konveksisen korvaavan häviöfunktion minimointia, ja ehdotamme sen perusteella uutta online-oppimisalgoritmia poikkeamien havaitsemiseen. Tämän jälkeen osoitamme laajoilla kokeilla, että ehdotetun algoritmin suorituskyky summamittarin suhteen on yhtä hyvä kuin hiljattain ehdotettu Cost-Sensitive Online Classification (CSOC) -algoritmi luokkien epätasapainon oppimiseen erilaisissa vertailuanalyyseissä ja että sen suoritusaika on lähellä havainto-algoritmia. Toinen johtopäätöksemme on, että muut kilpailukykyiset online-algoritmit eivät suoriudu johdonmukaisesti erikokoisista tietokokonaisuuksista. Tämä osoittaa ehdotetun lähestymistapamme mahdollisen sovellettavuuden.

**Tulos**

Verkkopoikkeamien havaitseminen luokan epätasapainon oppimisen avulla

**Esimerkki 2.1451**

Huomiomekanismin avulla syötetystä datasta ehdollistetut toistuvat sekvenssigeneraattorit ovat viime aikoina osoittaneet erittäin hyvää suorituskykyä useissa tehtävissä, kuten konekääntämisessä, käsialan synteesissä [1, 2] ja kuvatekstien luomisessa [3]. Laajennamme huomiomekanismia puheentunnistuksessa tarvittavilla ominaisuuksilla. Osoitamme, että vaikka konekääntämisessä [2] käytetyn mallin mukautus saavuttaa kilpailukykyisen 18,7 %:n foneemivirheprosentin (PER) TIMIT-foneemintunnistustehtävässä, sitä voidaan soveltaa vain lausumiin, jotka ovat suunnilleen yhtä pitkiä kuin ne, joihin se on koulutettu. Tarjoamme laadullisen selityksen tälle epäonnistumiselle ja ehdotamme uudenlaista ja yleistä menetelmää, jolla huomiomekanismiin lisätään sijaintitietoisuus, jotta ongelma voidaan ratkaista. Uusi menetelmä tuottaa mallin, joka on kestävä pitkille syötteille ja saavuttaa 18 % PER yksittäisillä lausumilla ja 20 % 10 kertaa pidemmillä (toistuvilla) lausumilla. Lopuksi ehdotamme tarkkaavaisuusmekanismiin muutosta, joka estää sitä keskittymästä liikaa yksittäisiin ruutuihin, mikä edelleen vähentää PER:n 17,6 %:iin.

**Tulos**

Huomioon perustuvat mallit puheentunnistusta varten

**Esimerkki 2.1452**

Keinotekoinen yleinen älykkyys (Artificial General Intelligence, AGI) on tutkimusala, jonka tavoitteena on selvittää älykkyyden periaatteet, jotka toimivat tietystä ongelma-alueesta tai ennalta määritellystä kontekstista riippumatta, ja hyödyntää näitä periaatteita syntetisoidakseen järjestelmiä, jotka kykenevät suorittamaan mitä tahansa älyllistä tehtävää, johon ihminen kykenee, ja lopulta menemään sitä pidemmälle. Vaikka "kapea-alainen" tekoäly, joka keskittyy tiettyjen ongelmien ratkaisemiseen, kuten puheentunnistukseen, tekstin ymmärtämiseen, visuaalisten kuvioiden tunnistamiseen, robottien liikkeisiin jne. on viime aikoina saavuttanut vaikuttavia läpimurtoja, yleisen älykkyyden ymmärtäminen on edelleen vaikeasti saavutettavissa. Tässä artikkelissa tarjoamme uudenlaisen teoreettisen lähestymistavan yleisen älykkyyden ymmärtämiseen. Aluksi esitellään lyhyesti nykyinen käsitteellinen lähestymistapa. Kritiikkimme paljastaa useita vakavia rajoituksia, jotka juontavat juurensa älykkyyden käsitteen ontologisiin juuriin. Sitten ehdotamme paradigman muutosta älykkyydestä, joka käsitetään yksittäisten agenttien pätevyytenä, joka määritellään suhteessa a priori annettuun ongelma-alueeseen tai tavoitteeseen, älykkyyteen, joka käsitetään itseorganisoitumisen prosessina, jonka avulla älykkäät agentit yksilöityvät. Kutsumme tätä prosessia avoimeksi älykkyydeksi. Tämä paradigmaattinen muutos laajentaa älykkyyden käsitettä merkittävästi sen nykyisiä perinteisiä määritelmiä pidemmälle ja voittaa kritiikissä esiin tulleet vaikeudet. Avoin älykkyys kehitetään kognitiivisen kehityksen prosessin abstraktioksi, jotta sen soveltaminen voidaan ulottaa koskemaan yleisiä agentteja ja järjestelmiä. Esittelemme ja keskustelemme ajatuksen kolmesta näkökulmasta: yksilöitymisen filosofisesta käsitteestä, järjen muodostamisesta - objektien ja suhteiden maailman synnyttämisestä - ja yleisten kognitiivisten agenttien yksilöitymisestä kognition enaktiivisen lähestymistavan ja kokoamisteorian valossa. Tutkimme näitä selvittääksemme, missä mielessä formatiiviset yksilöitymisprosessit ovat todellakin älyllisiä ja miksi ne ovat avoimia. Lisäksi osoitamme, miten avoin älykkyys voidaan muotoilla vuorovaikutuksessa olevien elementtien hajautetun, itseorganisoituvan verkoston (eli monimutkaisen adaptiivisen järjestelmän) avulla ja miten tällainen prosessi on skaalautuva. Kehys tuo esiin tärkeän suhteen koordinoinnin ja älykkyyden välillä sekä uudenlaisen käsityksen arvoista. Lopuksi esitetään joukko kysymyksiä tulevaa tutkimusta varten.

**Tulos**

Älykkäiden agenttien yksilöinti

**Esimerkki 2.1453**

Ehdotamme käsitteellisesti yksinkertaista ja kevyttä kehystä syvälle vahvistusoppimiselle, jossa käytetään asynkronista gradienttilaskeutumista syvien neuroverkkojen ohjainten optimointiin. Esittelemme neljän tavanomaisen vahvistusoppimisalgoritmin asynkroniset muunnelmat ja osoitamme, että rinnakkaisilla toimijaoppijoilla on vakauttava vaikutus koulutukseen, minkä ansiosta kaikilla neljällä menetelmällä voidaan menestyksekkäästi kouluttaa neuroverkon ohjaimia. Parhaiten suoriutuva menetelmä, actor-criticin asynkroninen muunnos, ylittää Atari-verkkotunnuksen nykyisen huipputason, kun sitä koulutetaan puolet lyhyemmässä ajassa yhdellä moniydinsuorittimella GPU:n sijaan. Lisäksi osoitamme, että asynkroninen actor-critic onnistuu monissa jatkuvissa moottorinohjausongelmissa sekä uudessa tehtävässä, jossa etsitään palkintoja satunnaisista 3D-suosituksista visuaalista syötettä käyttäen.

**Tulos**

Asynkroniset menetelmät syvä vahvistusoppimista varten