

Viikkoraportti 1

Työn edistyminen

Tällä viikolla olen aloittanut projektin kanssa tyoskentelyn. Aiheenani on koneoppimisella toteutettava nimigeneraattori, jonka kirjoitan Java-kielellä. Olen valinnut algoritmit ja tietojärjestelmät, kuten määrittelydokumentissa sanotaan. En ole vielä kirjoittanut riviäkään koodia, mutta olen luonut projektin, lisännyt sen githubiin ja tehnyt suunnitelmia projektin etenemiseksi.

Opin

Olen käyttänyt gittiä aikaisemminkin, mutta siitä on jo useita kuukausia, joten kesti hetken palautella kaikki mieleen. Opin kuitenkin, kuinka tärkeä datan määrä koneoppimisessa on, joten päädyin muuttamaan alkuperäisen ideani (motivaatiolausegenerointi) joksikin vähän yksinkertaisemmaksi, josta on helpompi löytää riittävästi dataa.

Epäselvää

Ei vielä mitään. Muutamia kysymyksiä on, mutta lähinnä ne liittyvät algoritmeihin ja tietojärjestelmiin ja tiedän, että saan niihin vastauksen kun jähden tutkiskelemaan syvemmin

Seuraavaksi

Ensi viikolla aion käydä läpi materiaalia valittuihin algoritmeihin ja tietojärjestelmiin liittyen, ja ryhtyä soveltamaan sitä aiheeseeni. Haluaisin saada jonkinikäisen rungon/kokonaiskuvan valmiiksi.

Viikkoraportti 2

Työn edistyminen

Tällä viikolla sain aloitettua projektin virallisesti. Olen saanut aikaiseksi ne luokat, joita uskon tarvitsevani:

Kirjain: säilyttää kirjaimen tiedot ja puun tapaukselle, jossa kyseinen kirjain on “ensimmäinen”

Kirjaimet: Taulukko, joka huolehtii kaikista kirjaimista ja siitä, mikä kirjain on milloinkin ensimmäinen tai sen hetkisen tutkimuksen alainen kirjain. Myös hoitaa tallentamisen ja tekstitiedoston lukemisen.

Nimeäjä: Luokka, joka loppupelissä luo nimen. Kyselee kirjaimet- luokalta tietoa todennäköisyyksistä ja arpoo ne. Arpomista ei ole vielä painotettu

Opettaja: Luokka, jossa käydään läpi data ja lasketaan todennäköisyydet kullekin kirjaimelle ja kirjainparin jälkeiselle kirjaimelle. Luokilla on perus metodit tehtynä (Get, Set, Tallenna, Lue jne). Alla näkyy muutama alustava kaavio ja pseudokoodi projektiin liittyen

Opin

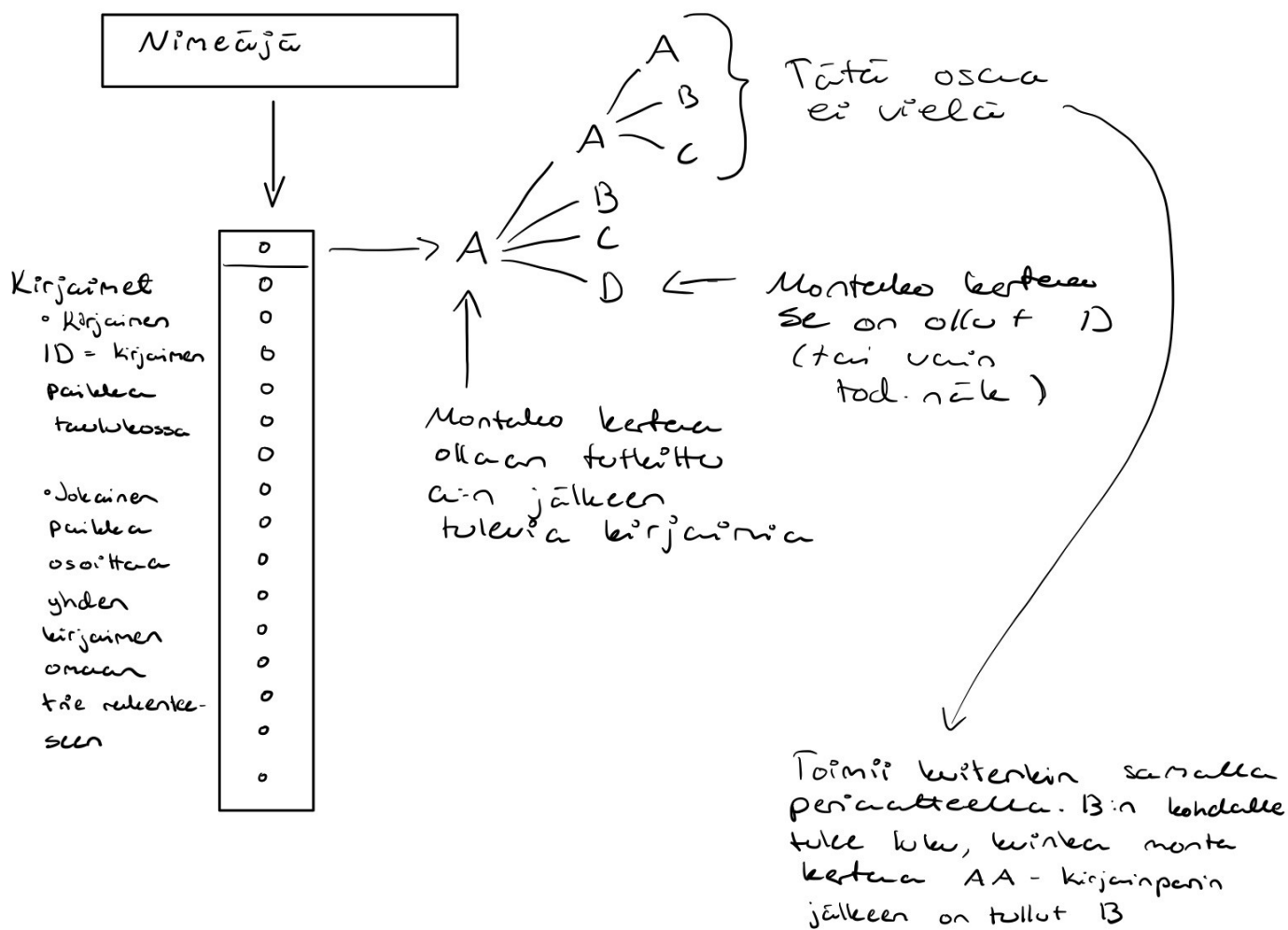
Opin, että Puu-rakennetta sanotaan myös trie-rakenteeksi. En ollut koskaan aikaisemmin kuullut termiä, joten se hämmensi hetken. Opin myös, että puurakennetta voi käyttää melko moneen asiaan.

Epäselvää

Ei ole suuria kysymyksiä tällä hetkellä. Yleisesti mietityttää, että olenko lähdössä ratkaisemaan tätä oikealla ja kaikista tehoikkaimmalla tavalla.

Seuraavaksi

Ensi viikolla työstän Opetus-luokkaa ja testailen kaikki muut luokat toimiviksi.



Pseudokoodi - runko

1. Ota kirjain (todennäköisyys ensimmäiseksi kirjaimeksi)
2. Mene kirjaimen puuhun ja arvo kirjain (todennäköisyydet painotettu)
3. Lisää kirjain string-jonoon
4. Aseta kirjain uudeksi puukirjaimeksi
5. Onko string yli 15 merkkiä pitkä (jos on, lopeta)
6. Katsotaan, lopetettuanko (todennäköisyydet vika kirjain ja nimen pituus)
7. Jatka kohdasta 2.

↳ Kun tarjotun nimen tarjotaan "Hyväksyntä, kuulostaa suomalaiselta" ja valmis nimi pyydetään jonkin hyväksymis-protokollan läpi

Pseudokoodi opetus

1. Ota sana
2. +1 alkukirjain pinoon alkukiinnetle
3. Seuraava kirjain +1 seuraavalle kirjaimelle kirjaimen puussa, +1 kokonaiskirjainten (tämän kirjaimen määrän)
4. Onko seuraava kirjainta, jos ei +1 viimeisen kirjain
ja +1 sanan pituudelle
5. Siirry seuraavaan ja jatka kohdasta 2

Viikkoraportti 3

Työn edistyminen

Tällä viikolla olin heiman jumissa trie-rakenteen kanssa, joten projekti eteni melko hitaasti. Joiduinmyös tekemään uudelleen muutamia luokkia ja oliaota. Nyt minulla on kuitenkin melko selvät sävelet siitä, minne ollaan menossa ja miten. Tällä viikolla keräsin myös opetusdatan kasaan ja muokkasin sen luettavaan muotoon

Opin

Opin trie-raknteesta ja siitä, miten sen avulla voidaan toteuttaa tämä projekti. Sain mahtavan ahaa-elämyksen.

Epäselvää

Ei mitää, tällä hetkellä kaikki on melkoisen selvää.

Seuraavaksi

Ensi viikolla testailen koodia ja haluaisin saada nimen luojaan toimimaa. Sen lisäksi tallenna- ja lue-metodit olisi hyvä saada toimimaan

Viikkoraportti 4

Työn edistyminen

Tällä viikolla olen kamppaillut toString() ja Parse() metodien kanssa, joita tarvitaan tallentamiseen ja takkennetun tiedon lukemiseen. Jouduin lisäämään jokaiselle solmulle oman ID-numeron, jonka avulla on mahdollista tunnistaa vanhempi ja lapset. Parse()-metodi jäy vielä vähän kesken, mutta senkin pitäisi toimia kohta. Tällä viikolla tein melkein kaikkiin metodeihin dokumentoinnin, ensiviikolla teen loput dokumentoinnit ja testit

Opin

Palauttelin mieleen tallenna ja lue-metodit aikaisemmasta projektista. En oppinut mitään suoranaisesti uutta, mutta oli hyvä kerrata valhoja asioita

Epäselvää

En ole ihan varma, miksi List<Solmu> ei toimi kunnolla, mutta olen melko varma että se on helppo saada selville tai vain vaihtaa se taulukoksi. Jos kuitenkin nopeasti Main:a katsomalla keksit, mikä mättää, niin kerro toki.

Seuraavaksi

Ensi viikolla aion saada dokumentoinnin, testauksen, tallentamisen ja tiedoston lukemisen toiminaan kunnolla.