

Dirbtinio intelekto sistemų inžinerija

Projektų inžinerijos vaidmuo kuriant DI sprendimus. Projekto tikslų ir sėkmės kriterijų apibrėžimas

Artificial Intelligence System Engineering

The role of project engineering in AI development. Defining project objectives and success criteria

- **Projektų inžinerija (Project Engineering)** apima visus gamybos ar apdorojimo projektavimo etapus – tiek naujų įrenginių/sistemų kūrimą, tiek esamų įrenginių/sistemų modifikacijas ir plėtrą.
- **Projektų inžinerija (Project engineering)** – tai sudėtingų išteklių ir procesų valdymas, ypač tada, kai fiziniai ir techniniai elementai yra lemiami projekto rezultatams (*PMBOK® Guide (2021)*)
- **Dirbtinio intelekto (DI) projektų inžinerija (Project engineering in AI)** apima inžinerijos principų integravimą su projektų valdymu, siekiant užtikrinti sėkmingą DI sprendimų kūrimą ir diegimą.

Dirbtinio intelekto ir mašininio mokymo projektų sudėtingumas

- Nerealistiški lūkesčiai, ką DI gali padaryti - greičiausias būdas projektui žlugti yra tada, kai vadovai ar verslo lyderiai mano, kad DI gali išspręsti visas jų problemas
- Technologinės žinios nereiškia, kad jūsų projektas pavyks;
- Nustatyta, kad 85% ML projektų žlunga (Gartner 2018), ir vis labiau kyla susirūpinimas, kad ML yra taikoma neetiškai ir žalingai (Bender et-al 2021)
- Visi mato tik galutinius rezultatus. Tačiau ML projektais yra daug sudėtingesni nei įprasti sistemų inžinerijos projektai ir jų valdymas (o jie patys savaime jau yra sudėtingi)

Atskirti ir suprasti DI bei ML

DI (Dirbtinis intelektas)

Mašininis mokymas (ML)

ML yra DI taikymas, leidžiantis išgauti žinias iš duomenų ir savarankiškai iš jų mokytis

DI yra platesnis terminas, juo nusakoma, kad mašinai ar sistema jaučia, mąsto, veikia ar prisitaiko kaip žmogus

DI (Dirbtinis intelektas)

ML

NLP

Deep
Learning

Robotika

Kita

Atskirti ir suprasti DI bei ML

- Dirbtinis intelektas yra plati sritis, kuri reiškia technologijų naudojimą kuriant mašinas ir kompiuterius, turinčius gebėjimą imituoti su žmogaus intelektu susijusias pažinimo funkcijas, tokias kaip gebėjimas matyti, suprasti ir reaguoti į kalba žodžiu ar raštu, analizuoti duomenis, teikti rekomendacijas ir dar daugiau.
- Nors dirbtinis intelektas dažnai laikomas sistema savaimė, tai yra technologijų rinkinys, integruotas į sistemą, kad ji galėtų mąstyti, mokytis ir veikti spręsdama sudėtingas problemas.

Atskirti ir suprasti DI bei ML

- Mašininis mokymasis yra dirbtinio intelekto poaibis, kuris suteikia mašinai ar sistemai gebėjimą mokytis ir tobulėti remiantis patirtimi. Vietoj programavimo, mašininis mokymasis naudoja algoritmus didelių duomenų kiekių analizei, ižvalgų išmokimui ir tada sprendimų priėmimui.
- Mašininio mokymosi algoritmai tobulina našumą laikui bégant, kai yra mokomi – paveikiami daugiau duomenų. Mašininio mokymosi modeliai yra rezultatas arba tai, ko programa išmoksta vykdydama algoritmą su mokymo duomenimis. Kuo daugiau duomenų naudojama, tuo geresnis bus modelis

Dirbtinis intelektas

- DI leidžia mašinai imituoti žmogaus intelekta problemoms spręsti
- Tikslas yra sukurti sistemą, kuri gali atlikti sudėtingas užduotis
- Mes kuriame sistemas, kurios gali spręsti sudėtingas užduotis kaip žmogus
- DI turi platų taikymo spektrą
- DI naudoja technologijas sistemoje, kad ji imituotų žmogaus sprendimų priėmimą
- DI dirba su visų tipų duomenimis: struktūruotais, pusiau struktūruotais ir nestruktūruotais
- DI sistemos naudoja logiką ir sprendimų medžius mokytis, samprotauti ir taisytis pačioms pasitaisyti

Mašininis mokymasis

- Mašininis mokymas leidžia mašinai savarankiškai mokytis iš ankstesnių duomenų
- Tikslas yra sukurti mašinas, kurios gali mokytis iš duomenų, kad padidintų rezultato tikslumą
- Norime kurti tokius sprendimus, kad mašinos būtų mokomos atlikti specifines užduotis ir teiktų tikslius rezultatus
- Mašininis mokymasis turi ribotą taikymo spektrą
- Naudoja savarankiško mokymosi algoritmus, kad sukurtų prognozavimo modelius
- Mašininis mokymasis gali naudoti tik struktūruotus ir pusiau struktūruotus duomenis
- Tokios sistemos remiasi statistiniais modeliais ir gali taisytis pačios, kai pateikiami nauji duomenys

Projekto tikslai

Projekto tikslai

- Verslo problemas reikia išversti į tyrimo klausimus, kurie atitinka išmatuojamus verslo rodiklius;
- DI projekto tikslai turėtų būti glaudžiai susieti su verslo tikslais, o ne tik su techniniais rodikliais, tokiais kaip RMSE, kurie gali tiesiogiai nenurodyti verslo sėkmės;

Pagrindiniai vaidmenys DI projektuose

- Sistemos architektūrų projektavimas;
- Duomenų integracija, apdorojimas ir plečiamumas;
- Daugiafunkcių komandų valdymas (pvz., duomenų mokslininkų, inžinierių, suinteresuotujų šalių);
- Atitikties reguliavimas ir etikos standartų užtikrinimas;.

- Tikslai nustato projekto paskirtį ir suderina techninę plėtrą su verslo poreikiais
- Gerai apibrėžti tikslai nustati kryptį dirbtinio intelekto modelio dizainui, funkcijų pasirinkimui ir vertinimo procesams.
- Tikslai turi būti - Konkretūs, Išmatuojami, Pasiekiami, Aktualūs ir Apibrėžti laiko atžvilgiu (SMART).
- **Specific, Measurable, Achievable, Relevant, and Time-bound (SMART)**
- Bendradarbiavimas su suinteresuotomis šalimis yra būtinus, kad užtikrintumėte, jog projekto tikslai atspindi organizacijos tikslus.

Sėkmės kriterijų nustatymas

- Sėkmės kriterijai yra standartai, naudojami projekto sėkmei įvertinti pagal apibrėžtus tikslus.
- **Pagrindiniai sėkmės kriterijų tipai dirbtinio intelekto projektuose:**
 - **Strateginis suderinamumas:** Sėkmė priklauso nuo projekto sederinumo su organizacijos tikslais ir išmatuojamos vertės sukūrimo.
 - **Verslo poveikio rodikliai:** Pajamų augimas, klientų įsitraukimas, operacinis efektyvumas, sumažėjęs klientų praradimas, padidėjęs pardavimas, klientų pasitenkinimas, produktyvumo pagerinimas, sąnaudų mažinimas.
 - **Apibrėžti KPI (angl. *Key Performance Indicator*)** (pagrindiniai veiklos rodikliai): Tendencijas prognozuojantys rodikliai (leading indicators) ir rezultato rodikliai (lagging indicators), tokie kaip laiku atlikti darbai, biudžeto laikymasis ir projekto etapų įvykdymo rodikliai.

- Pagrindiniai sėkmės kriterijų tipai dirbtinio intelekto projektuose:
 - **Suderinamumas su suinteresuotomis šalimis:** kaip gerai suinteresuotujų šalių apibrėžtus tikslus.
 - **Atitiktis ir rizikos mažinimas:** Rodikliai, susiję su reglamentų laikymusi arba rizikos mažinimu.
 - **Modelio našumo rodikliai:** Tikslumas (accuracy), recall, precizija (precision), F1 balas (F1 score).
 - **Modelio panaudojimas ir plečiamumas:** Sėkmė apima modelio praktinį poveikį, jo plečiamumą ir prisitaikymą prie realių duomenų
 - **Kokybės užtikrinimas ir stebėjimas:** Palaikyti dirbtinio intelekto našumą laikui bėgant naudojant kokybės rodiklius, tokius kaip tikslumas, precizija ir jautrumas (recall). Atitikti apibrėžtus kokybės standartus ir užtikrinti sistemos patikimumą bei plečiamumą ilgalaikiam naudojimui.

Projekto tikslų ir sėkmės kriterijų suderinimas su verslo tikslais

- **Iššūkiai dirbtinio intelekto projektuose:**
 - Užtikrinti, kad tikslai būtų įgyvendinami ir susieti su konkrečiais verslo rezultatai.
 - Vengti tik techninių rodiklių (pvz., modelio tikslumo), kurie gali neatspindėti verslo vertės
- **Strategijos:**
 - Itraukti verslo suinteresuotąsias šalis anksti, kad būtų nustatyti aktualūs sėkmės rodikliai.
 - Naudoti vertės teikimo sistemą (angl. **Value Delivery Framework**), kad dirbtinio intelekto rezultatai būtų susieti su verslo verte.
 - Reguliariai atnaujinti tikslus ir sėkmės kriterijus, kad atspindėtų kintančius tikslus ir projektą.

Projektų inžinerija dirbtinio intelekto sėkmėi

- **Projektų inžinerijos etapai kuriant dirbtinį intelektą:**
 - **Planavimo etapas:** Apibrėžti aiškius tikslus, reikalavimus ir išteklius.
 - **Vykdymo (execution) etapas:** Įgyvendinti griežtą dirbtinio intelekto modelio mokymo, testavimo ir našumo stebėjimą.
 - **Vertinimo etapas:** Naudoti nustatytais kriterijus įvertinti, ar dirbtinio intelekto modeliai atitinka sėkmės slenkstinius kriterijus (threshold).
 - **Inžinerijos indėlis:**
 - Architektūrinis dizainas plečiamumui užtikrinti.
 - Procesų valdymas, skirtas iteracijoms modelio kūrimui vystyti
 - Rizikos vertinimas ir mažinimas, ypač dirbtinio intelekto specifinėms rizikoms (pvz., modelių veiklos pablogėjimo procesas laikui bėgant, šališkumas)

Nuolatinis suderinamumas ir pritaikomumas

- **Nuolatinis vertinimas:** Dirbtinio intelekto projektuose tiek tikslai, tiek sėkmės kriterijai turi būti reguliariai pervertinami dėl duomenų ir rinkos dinamikos.
- **Projektų inžinerijos vaidmuo pritaikomume:**
 - Užtikrinti suderinamumą su verslo strategija, įvertinant sėkmės kriterijus
 - Efektyviai valdyti išteklius ir terminus, kad būtų įtraukti iteraciniai pakeitimai
 - Reguliariai įtraukti suinteresuotąsias šalis, kad būtų nuolat išlaikoma projekto kryptis, kuomet įgyjamos naujos ižvalgos

Literatūra

- Veljko Krunic - "Succeeding with AI"
- PMBOK® Guide (2021)
- Klaus Haller - "Managing AI in the Enterprise"
- Chip Huyen - "Designing Machine Learning Systems"