

Joe Hyvärinen

Ehdottavan analytiikan metodit ja haasteet

Tietotekniikan kandidaatintutkielma

29. huhtikuuta 2022

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Joe Hyvärinen

Yhteystiedot: jojohyva@student.jyu.fi

Ohjaaja: Tuomo Rossi

Työn nimi: Ehdottavan analytiikan metodit ja haasteet

Title in English: Prescriptive analytics methods and challenges

Työ: Kandidaatintutkielma

Opintosuunta: Tietotekniikka

Sivumäärä: 18+0

Tiivistelmä: Ehdottava analytiikka pyrkii ehdottamaan parhaimman mahdollisen ratkaisuvaihtoehdon hyödyntäen suuria määriä dataa. Täten ehdottava analytiikka onkin tehokas tapa parantaa sekä tukea liiketoiminnan päätöksentekoa. Tässä tutkielmassa tutkitaan ehdottavan analytiikan metodeja ja haasteita. Tutkimus toteutettiin kirjallisuuskatsauksena. Tutkielman lopussa käydään läpi merkittävimpiä tuloksia sekä pohditaan monitavoiteoptimoinnin hyödyntämistä ehdottavan analytiikan yhteydessä.

Avainsanat: Ehdottava analytiikka, data-analytiikka, liiketoiminta-analytiikka, big data

Abstract: Prescriptive analytics aims to prescribe the best possible solution utilizing large amounts of data. Thus prescriptive analytics is a powerful way to improve and support decision making in businesses. This thesis explores prescriptive analytics methods and challenges. The study was conducted as a literature review. The key findings of the study are examined in the end of the thesis. There is also discussion about utilizing multiobjective optimization with prescriptive analytics.

Keywords: Prescriptive analytics, data analytics, business analytics, big data

Kuviot

Kuvio 1. Datasta päätöksentekoon (University of Jyväskylä 2021).....	11
--	----

Sisällys

1	JOHDANTO	1
2	KÄSITTEET	2
3	EHDOTTAVAN ANALYTIIKAN METODIT	3
3.1	Matemaattiset optimointimenetelmät.....	3
3.2	Simulointi ja kustannus-hyötyanalyysi.....	5
3.3	Logiikkamallit, tiedonlouhinta ja koneoppiminen	6
4	EHDOTTAVAN ANALYTIIKAN HAASTEET	7
4.1	Dataan liittyvät haasteet.....	7
4.2	Ehdottavan analytiikan ja ihmisen välisen suhteen haasteet	8
5	POHDINTA	10
	LÄHTEET	12

1 Johdanto

Yrityksissä pyritään jatkuvasti hyödyntämään massadataa (big data) ja liiketoiminta-analytiikkaa, jotta saataisiin etua kilpailijoihin. Ehdottavan analytiikan kehittymisen myötä yritykset pääsevät käyttämään hyväkseen big dataa uudesta näkökulmasta. Tämän tutkielman tavoite on selvittää ehdottavan analytiikan metodeja ja haasteita. Tutkielma on toteutettu kirjallisuuskatsauksena.

Liiketoiminta-analytiikalla pyritään parantamaan liiketoimintojen päätöksentekoa. Suurin osa tutkimuksista on tehty kuvailevasta ja ennustavasta analytiikasta. Ehdottavan analytiikan kehittymisen myötä, myös tutkimuksia on alettu toteuttamaan siihen liittyen. Ehdottavalla analytiikalla on tavoitteena parantaa yritysten suorituskykyä antamalla optimoituja ratkaisuvaihtoehtoja päätösten tueksi. Tässä tutkielmassa keskitytään nimenomaan ehdottavaan analytiikkaan. Tutkielmassa vastataan seuraaviin tutkimuskysymyksiin: ”Mitä ovat ehdottavan analytiikan menetit?” ja ”Mitä haasteita liittyy ehdottavaan analytiikkaan?”. Aineistohakuun on käytetty Google Scholaria. Aineistohaku suoritettiin seuraavilla hakulauseilla: ”prescriptive analytics”, ”prescriptive analytics methods” ja ”prescriptive analytics challenges”.

Tutkielman seuraavassa luvussa käydään läpi yleisiä termejä liittyen ehdottavaan analytiikkaan. Kolmannessa luvussa käsitellään ehdottavan analytiikan metodeja ja pyritään saamaan ymmärrys niistä. Kolmas luku vastaa tutkimuskysymykseen: ”Mitä ovat ehdottavan analytiikan menetit?”. Neljännessä kappaleessa käsitellään haasteita liittyen ehdottavaan analytiikkaan ja siinä myös vastataan tutkimuskysymykseen: ”Mitä haasteita liittyy ehdottavaan analytiikkaan?”. Molempiin tutkimuskysymyksiin löydettiin vastaukset aineistohausta saaduilla aineistoilla.

2 Käsitteet

Ehdottavan analytiikan ymmärtämiseksi on tärkeää tietää sitä edeltävät liiketoiminta-analytiikan vaiheet. Šikšnys ym. (2016) mukaan liiketoiminta-analytiikka voidaan jakaa kolmeen eri vaiheeseen, niiden vaikeuden, arvon ja älykkyyden perusteella. Heidän mukaansa nämä kolme vaihetta ovat kuvaileva analytiikka, jossa vastataan kysymykseen: ”Mitä tapahtui”, ennustava analytiikka, jossa vastataan kysymykseen: ”Mitä tulee tapahtumaan?” ja ehdottava analytiikka, jossa vastataan kysymykseen: ”Kuinka saada jotain tapahtumaan?”.

Ehdottava analytiikka on liiketoiminta-analytiikan kehittynein vaihe ja voi tuoda yrityksille eniten arvoa ja älykkyyttä (Šikšnys ym. 2016). Šikšnys ym. (2016) mainitsivat, että ehdottavan analytiikan tehtävät ovat parhaimpien mahdollisten ratkaisuvaihtoehtojen ehdottaminen ja näiden ehdotusten seurauksien havainnollistaminen. Myös Frazzetto ym. (2019) mukaan, ehdottava analytiikka mahdollistaa parhaimman ratkaisuvaihtoehdon ehdottamisen tiettyjen tavoitteiden mukaisesti. Basu (2013) taas huomautti, että ehdottavan analytiikan mallien tulee pystyä hyödyntämään sekä järjesteltyä että järjestelemätöntä dataa, jotta voidaan tehdä paras mahdollinen päätös.

Tutkimuksen kannalta on myös oleellista tietää, mitä on big data tai massadata. Roy ym. mukaan big data määritellään suurivolyymiseksi, nopeaksi tietoresurssiksi, joka koostuu järjestelemättömistä teksti-, ääni- ja videotiedostoista (Roy ym. 2016). Heidän mukaansa merkittäviä haasteita big datan käsittelyyn liittyen ovat tietotyyppien monimuotoisuus, datan epävarmuus ja joskus tiedonkeruun nopeus.

3 Ehdottavan analytiikan metodit

Tässä luvussa vastataan tutkimuskysymykseen: ”Mitä ovat ehdottavan analytiikan metodit?”. Aluksi käydään läpi optimointiin liittyviä metodeja. Sen jälkeen esitellään tilanteita, missä on käytetty simulaatiota ja kustannus-hyötyanalyysia. Lopuksi käsitellään logiikkamalleja sekä tiedonlouhinnan ja koneoppimisen tekniikoita.

3.1 Matemaattiset optimointimenetelmät

Lineaariset optimointimenetelmät (engl. linear programming) ovat tällä hetkellä käytetyimpiä optimointimenetelmiä. Lineaarinen optimointi on matemaattinen menetelmä lineaarisen kohdefunktion maksimointiin tai minimointiin, johon voi vaikuttaa yhtä- ja epäsuuret lineaariset rajoitteet. Von Bischoffshausen ym. (2015) käyttivät lineaarista optimointia yrityksen myyntitulojen maksimointiin. Heidän ehdottavan analytiikan mallinsa suunnitteli yrityksen myyntityöntekijät maksimoidakseen myyntitulot.

Lineaarinen kokonaislukuoptimointi (engl. integer linear programming) eroaa lineaarisesta ohjelmoinnista siten, että sen muuttujat ovat rajoitettu kokonaislukuihin. Muun muassa Aref ym. (2015) ja Ghoniem ym. (2017) käyttivät lineaarista kokonaislukuoptimointia. Aref ym. (2015) uudelleen suunnittelivat yrityksen ohjelmistopinon (engl. software stack) yhtenäisen deklaraatiivisen ohjelmointimallin ympärille, jotta se tukisi matemaattista optimointia. Ghoniem ym. (2017) taas arvioivat Qatarin majoitusinfrastruktuurin maksimaalisen kapasiteetin FIFA World Cup 2022 -turnaukseen.

Sekoitetussa lineaarisessa kokonaislukuoptimoinnissa (engl. mixed integer linear programming) vain jotkin muuttujista ovat rajoitettu kokonaislukuihin ja muut voivat olla muita kuin kokonaislukuja. Esimerkiksi Kawas ym. (2013) käyttivät sekoitettua lineaarista kokonaislukuoptimointia optimoidakseen ison teknologiayrityksen B2B-myyntiprosessin. Optimointia varten he loivat mallin, joka tunnistaa ihanteellisen ihmismäärän per myyntiluokka, mikä mahdollistaa tuoton kannalta tuottavimman myyntiryhmän valitsemisen tulevaisuuden myyntimahdollisuuksiin jokaisessa asiakassegmentissä. Mainittakoon, että Kawas ym. (2013) malli hyödynsi myös ennustavaa analytiikkaa, joka teki ennustuksia mennei-

den myyntitietojen perusteella, jotta saataisiin myyntiryhmien koon ja koostumuksen sekä erityyppisiltä asiakkailta ja mahdollisuuksista saatujen tulojen käyttäytymissuhde.

Epälineaarinen optimointi (engl. nonlinear programming) eroaa lineaarisesta optimoinnista silloin, kun jotkin rajoitteet tai kohdefunktio on epälineaarinen. Esimerkiksi Goyal ym. (2016) hyödynsivät sekoitettua epälineaarista kokonaislukuoptimointia sähköverkon ylläpitoon. He loivat mallin, joka seurasi sähköverkon eri osien kuntoa, tunnisti niiden jäljellä olevan elinkaaren ja arvioi verkon vakauden. Malli antoi myös kokonaisvaltaisen arvion sähköverkon kunnosta sekä riskiarvion. Se myös ehdotti parhainta ylläpitostrategiaa budjettiin nähden ja minimoi ylläpidon kustannuksen keskiarvon hinnan käyttäen sekoitettua epälineaarista kokonaislukuoptimointia (engl. mixed integer nonlinear programming).

Ito ja Fujimaki (2017) käyttivät binääristä neliöllistä epälineaarista optimointia (engl. binary quadratic nonlinear programming) jonkin tuotteen hinnoittelustrategian optimointiin. Heillä oli käytössä 50 eri oluttuotetta sekä simuloitu tietojoukko. Heidän tuloksensa osoittivat, että mallin optimoima hinnoittelustrategia paransi tuotteen bruttovoittoa 8.2 prosentilla. Lisäksi Huang, Bergman ja Gopal (2019) tutkivat laajennuspaikkojen valintaa jälleenmyyjille, jotka myyvät lisätuotteita, joiden kysyntä on johdettu perustuotteesta. Lopulta he päätyivät binääriseen neliölliseen epälineaariseen optimointiongelmaan. Heidän tuloksensa viittasivat siihen, että valitsemalla laajennuspaikat heidän luomallaan mallilla, myyntiodotukset voivat nousta huomattavasti.

Stokastisilla optimointimeteodeilla (engl. stochastic optimization) pyritään minimoimaan tai maksimoimaan kohdefunktio, kun satunnaisuutta on mukana. Bertsimas ja Kallus (2020) totesivat, että stokastinen optimointi on ollut pitkään päätöksenteon tukena, kun päätöksentekijällä on epävarmuutta. Heidän mukaansa toinen huomionarvoinen lähestymistapa päätöksenteon tueksi epävarmuuden alaisuudessa on robusti optimointi (engl. robust optimization) ja sen datakäyttöiset variantit.

Vankassa optimoinnissa haetaan tiettyä määrää vankkuutta (engl. robustness) epävarmuutta vastaan, joka voidaan esittää deterministisenä vaihtelevuutena itse ongelman parametrien ja/tai sen ratkaisun arvossa. Esimerkiksi Berk ym. (2019) loivat henkilöstösuunnittelun mallin, joka käyttää robustia optimointia. He tutkivat henkilöstösuunnittelun päätöksiä yrityk-

sissä, joissa myytiin sopimusperusteisia konsultointiprojekteja. He totesivat, että sopimusten ja tuloennusteiden epävarmuuden takia konsultointiyrityksillä on vaikeuksia palkata oikeat henkilöt heidän projekteihinsa. Heidän ensimmäinen tavoitteensa oli tehdä dynaamisia päätöksiä palkkauksen suhteen, jotka maksimoivat tuoton. Toinen tavoite oli pysyä samanaikaisesti mahdollisimman joustavana palkkausprosessin yhteydessä.

3.2 Simulointi ja kustannus-hyötyanalyysi

Ehdottavassa analytiikassa on hyödynnetty myös simulointia. Muun muassa Marathe ym. (2014) käyttivät synteettistä informaatiota hyödykseen luomalla synteettisen väestön ja infrastruktuurin. Synteettisen informaation perusteella he simuloivat ydinaseiskun jälkiseurauksia Washington DC:ssä. He osoittivat, miten ehdottavan analytiikan, synteettisen informaation ja simulaatioiden avulla voidaan luoda vastatoiminen strategia, jolla palautetaan kaupungin viestintäinfrastruktuuri takaisin. Viestintäinfrastruktuurin palauttaminen nopeasti voi pelastaa useita tuhansia ihmisiä tällaisessa tilanteessa.

Myös stokastista simulaatiota (engl. stochastic simulation) on käytetty. Stokastisessa simulaatiossa simuloidaan sellaista järjestelmää, jossa on muuttujia, jotka voivat muuttua satunnaisesti. Esimerkiksi Jank, Dölle ja Schuh (2018) hyödynsivät stokastista simulaatiota tuotevalikoiman voittojen maksimointiin yrityksen tavoitteiden mukaisesti. He totesivat, että tuotepäälliköt eivät pysty suunnittelemaan tuotevalikoimaa, jotka täsmäävät yrityksen tavoitteisiin. Suunnittelun lisäksi he eivät pysty mukauttamaan tuotevalikoimaa tehokkaasti. Syyt haasteisiin tuotevalikoimien kanssa johtuu sisäisten ja ulkoisten tekijöiden monimutkaisuudesta ja nopeasti muuttuvista riippuvuuksista. Haasteiden takia he loivat mallin, joka hyödynsi stokastista simulaatiota.

Todennäköisyysperusteista lähestymistapaa kustannus-hyötyanalyysiin käytettiin optimaalisen ratkaisun ehdottamiseen ja riskien arviointiin ehdottavassa ylläpidossa (Koops 2020). Koops (2020) luoma todennäköisyysperusteinen kustannus-hyötyanalyysiin perustuva malli antaa parhaiten tuottavan ratkaisun sekä kertoo millainen riski ja potentiaalinen kustannus ratkaisuun liittyy. Hänen mallinsa pystyi antamaan kahta eri ehdotusta, korjaus tai korvaaminen. Wiener prosessiin perustuvaa degradaatio mallia, joka pystyy toteuttamaan satunnaisia

vaikutuksia johtuen puutteellisesta korjauksesta, käytettiin Monte Carlo simulaation kanssa havainnollistamaan mallin arvoa. Simulaation tulokset osoittivat, että korjauksen ehdottaminen on usein tuottavampaa, vaikka korvaaminen oli joissain tapauksin suositeltavampi. Joissain tapauksissa myös väärän päätöksen valinnan riski ja siihen liittyvä hinta nousi huomattavasti, kun mukana oli useita epävarmuuden ja vaihtelevuuden aiheuttavia tekijöitä.

3.3 Logiikkamallit, tiedonlouhinta ja koneoppiminen

Logiikkamallit ovat oletettuja kuvauksia syiden ja seurausten tapahtumaketjuista. Esimerkiksi Matyas ym. (2017) hyödynsivät logiikkamallia. He esittivät uudenlaisen tietopohjaisen menettelyllisen lähestymistavan ehdottavan ylläpidon suunnittelun toteuttamiseen tuotantotekniikan alalle. Heidän mallinsa pohjautui matemaattisten funktioiden esittämiin sääntöihin kullekin koneen komponentille. Se otti myös huomioon ennusteen koneen komponenttien rasiuksen kestosta kyseisten olosuhteiden alla, olosuhteisiin perustuvan seurannan ja vaihtelun tuotteiden laadussa. Lisäksi Lee ym. (2014) hyödynsivät logiikkamallia. Heidän luoma tietojärjestelmä pystyi antamaan neuvoja sekä tutkimussuosituksia tutkijoille. Heidän järjestelmä perustui viiden W:n (5W1H) metodologiaan. Metodissa pyritään vastaamaan kuuteen kysymykseen: "Kuka?", "Mitä?", "Milloin?", "Missä?", "Miksi?" ja "Miten?".

Yhtä tiedonlouhinnan tekniikkaa myös hyödynnettiin ehdottavassa analytiikassa. Esimerkiksi Gröger, Schwarz ja Mitschang (2014) hyödynsivät binääripuita (engl. binary trees) päätöspuina (engl. decision tree) ja he myös johtivat sääntöjä päätöspuista suorittaakseen suosituksiin perustuvaa liiketoimintaprosessien optimointia kokonaisvaltaisen prosessivaraston päälle.

Koneoppimista käytettiin yhdessä tutkimuksessa. Padmanaban ja Mukesh (2020) esittivät k:n keskiarvon klusterointiin (engl. k-means clustering) perustuvan lähestymistavan sijoittaa dataa mukautuvasti hajautettujen solmujen (engl. distributed nodes) välillä ottaen huomioon datan herkkyden ja turvallisuuden. Heidän lähestymistapansa on laajennettavissa ja skaalautuva erilaisiin datan herkkyys- tai turvallisuusvaatimuksiin.

4 Ehdottavan analytiikan haasteet

Tässä luvussa vastataan tutkimuskysymykseen: ”Mitä ovat ehdottavan analytiikan haasteet?”. Alkuun käydään läpi dataan liittyviä haasteita, jota seuraa ihmisen ja ehdottavan analytiikan väliseen suhteeseen kohdistuvien ongelmien tarkastelu.

4.1 Dataan liittyvät haasteet

Ehdottavan analytiikan mallien luontiin liittyy haasteita, miten saadaan luotua malleja, jotka käyttävät big dataa. Lepenioti ym. (2020) mukaan big datan käyttö varsinkin päätöksenteon yhteydessä on yksi isoimmista haasteista ehdottavassa analytiikassa. Lepenioti ym. (2020) mukaan hajautettujen alustojen käyttö voi parantaa big datan prosessointia ehdottavan analytiikan malleissa. Haasteita kuitenkin tulee esille, kun aletaan kehittämään ehdottavan analytiikan mallien algoritmeja hajautetuille alustoille, sillä laskennallinen vaativuus nousee merkittävästi.

Päätöksenteon automatisointiin liittyy haasteita. Lepenioti ym. (2020) mainitsivat, että ehdottavan analytiikan epävarmuudesta johtuen, päätöksenteon automatisointi on erittäin haastavaa. Väärän päätöksen jatkuva automatisointi voi koitua erittäin kalliiksi yritykselle.

Datan saatavuus, etenkin data aiemmista tapahtumista, on usein suuri haaste ehdottavan analytiikan mallia kehittäessä (Sappelli ym. 2017). Sappelli ym. (2017) totesivat myös, että datan puute johtuu yleensä yksityisyydestä ja lainsäädännöllisistä asioista. On myös mahdollista, ettei aiempia tekoja seurata ja kirjata ylös, jolloin ei saada dataa aiemmista teoista. Sappelli ym. (2017) mukaan vähäisen tai puutteellisen datan käyttö voi johtaa simulaatiomalliin tai tietämykseen pohjautuvaan malliin, joka ei hyväksikäytä ehdottavaa analytiikkaa sen täyteen potentiaaliin. Sappelli ym. (2017) mainitsivat vielä yhden mahdollisen haasteen liittyen dataan. Kun syötetään aiempia tekoja mallille, päätöksenteko voi johtaa yhteen tiettyyn päätökseen. Tämä voi johtaa ehdottavan analytiikan algoritmin ajautumaan loputtomaan silmukkaan.

Matyas ym. (2017) totesivat, että yksi suuri haaste ehdottavan ylläpidon toteutuksessa on da-

tan keruu ja käsittely. Toisaalta Roy ym. (2016) totesivat merkittävän haasteen liittyen big data pohjaisen ylläpitojärjestelmän kehittämiseen. Heidän mukaansa big datan monimuotoisuus tekee big dataan pohjautuvista järjestelmistä monimutkaisia ja komplekseja.

Terveydenhuoltoalalla on myös havaittu ongelmia dataan liittyen. Oesterreich ym. (2020) listasivat useita dataan liittyviä ongelmia ehdottavassa analytiikassa terveydenhuollossa. Ensimmäisenä haasteena he havaitsivat datan heterogeenisyyden eli datan monimuotoisuuden. Heidän mukaansa monimuotoista dataa on vaikea käsitellä. He myös mainitsivat, että ohjelmistojen tuottajilla on haasteita integroida monimuotoista dataa heidän ohjelmiinsa. Toisena ongelmana he nostivat esille datan keräämisen terveydenhuollon tarkoituksiin. Heidän mukaansa tästä koituu etenkin terveydenhuollon ammattilaisille haasteita käyttää terveydenhuollon dataa yhdistelmäanalysejä (engl. aggregated analyses) varten. Kolmas haaste liittyen dataan oli datan saatavuus. Oesterreich ym. (2020) mainitsivat, että terveyteen liittyvää dataa kerätään jatkuvasti, mutta usein se data ei ole käytettävissä tai saatavilla myöhemmin. Viimeisimpänä dataan liittyvänä haasteena he mainitsivat datan laadun. Heidän mukaansa merkittävä ongelma on virheelliset diagnoosit, jotka johtuvat heikosta datan laadusta. Lopuksi Oesterreich ym. (2020) huomauttivat, että yksi tärkeä haaste terveydenhuollossa on yksityisyys ja eettiset syyt. Tämä johtaa siihen, että potilaiden yksityinen data ei saa joutua väärin käsiin, jottei sitä hyväksikäytetä. Datan yksityisyys johtaa vielä yhteen heidän mainitsemaan haasteeseen, joka on eturistiriidat. Oesterreich ym. (2020) mukaan esimerkiksi vakuutusyrietykset sekä työntekijät haluaisivat pääsyn yksityisiin terveystietoihin. Tämä voi taas johtaa yksittäisten henkilöiden epätasa-arvoiseen kohteluun.

4.2 Ehdottavan analytiikan ja ihmisen välisen suhteen haasteet

Käyttäjän luottamus ehdottavan analytiikan mallin ratkaisuihin on yksi merkittävä haaste (Sappelli ym. 2017). Sappelli ym. (2017) mukaan ehdottavan analytiikan mallin tulisi tuottaa avoimia, helposti selitettäviä ja jäljitettäviä ratkaisuvaihtoehtoja, jotta käyttäjä voi luottaa, sekä hyväksyä nämä ratkaisuvaihtoehdot. Käyttäjän tulee olla kykeneväinen päättämään, onko annettu ratkaisu optimaalinen. Käyttäjä pystyy käyttämään omaa luovuuttaan, sekä sopeutumiskykyjään ennennäkemättömissä tilanteissa, mihin taas ehdottavan analytiikan malli ei ole kykeneväinen. Lopuksi Sappelli ym. (2017) huomauttivat, että käyttäjän ja mallin vä-

listen ääripäätilanteiden tasapainottaminen tulee olemaan merkittävä haaste jatkossa.

Terveydenhuollossa on tärkeä luoda luottamusta ehdottavan analytiikan ja potilaan välille. Oesterreich ym. (2020) mukaan luottamuksen tulee olla välttämätön ehto ehdottavan analytiikan käyttöön terveydenhuollossa. Moni potilas ei ole valmis jakamaan omia terveystietojaan, koska he eivät luota siihen, miten heidän terveystietojaan käsitellään. Oesterreich ym. (2020) mainitsivat, että luottamuksen puute voi myös johtua epäselvistä terveydenhuollon linjauksista. He myös totesivat, että sidosryhmillä, mitkä käsittelevät terveystietoja tulisi olla avoimesti saatavilla olevat linjaukset, mistä käy ilmi, miten ja mihin tarkoituksiin terveystietoja käytetään.

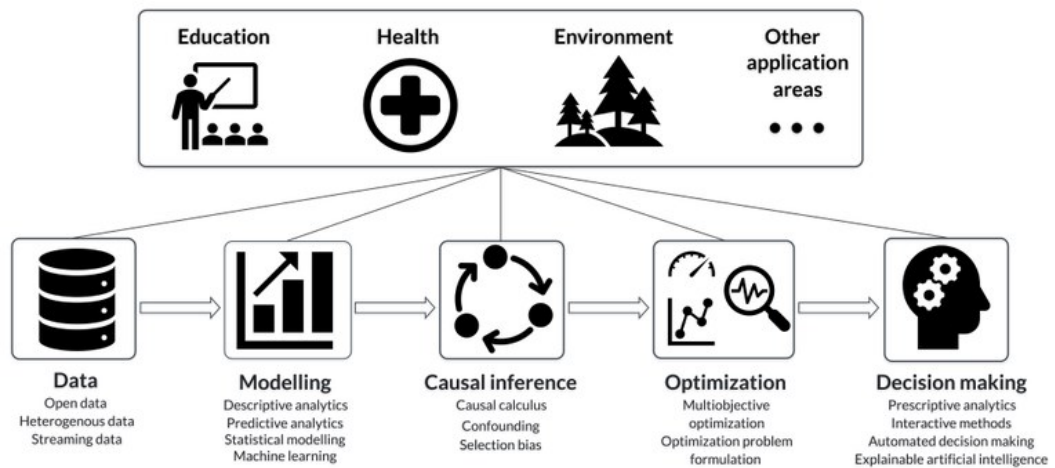
Ehdottavan analytiikan sovellusten käyttö vaatii koulutusta ja harjoittelua. Esimerkiksi terveydenhuollossa erilaiset sidosryhmät, kuten työntekijät ja muut hoitolaitokset, tarvitsevat koulutusta uuden sovellusten käyttöön (Oesterreich ym. 2020). Oesterreich ym. (2020) mukaan resurssien puute terveydenhuollossa on yksi keskeinen haaste, miksi ehdottavan analytiikan sovellukset eivät leviä terveydenhuollon piireissä.

5 Pohdinta

Tutkielmassa käytiin läpi ehdottavan analytiikan metodeja ja haasteita. Tutkielman perusteella huomataan, että ehdottava analytiikka on jatkuvasti kasvava ja kehittyvä aihe. Löydettyjä metodeja avataan, jotta niistä saataisiin parempi ymmärrys. Huomattavasti suosituimpia metodeja olivat matemaattiset optimointimenetelmät. Ehdottavaan analytiikkaan liittyi myös useita haasteita, joista merkittävimmät olivat dataan liittyviä. Datan käyttö ja sen saatavuus ehdottavan analytiikan yhteydessä koitui merkittäväksi haasteeksi. Toinen huomionarvoinen haaste liittyi käyttäjän ja ehdottavan analytiikan väliseen suhteeseen. Käyttäjillä oli vaikeuksia luottaa ehdottavan analytiikan luomiin ratkaisuvaihtoehtoihin.

Yksi huomionarvoinen menetelmä, jota ei tullut esille tutkimuksessa on monitavoiteoptimointi (engl. multiobjective optimization) ja sen hyödyntäminen ehdottavassa analytiikassa. Monitavoiteoptimoinnissa pyritään lineaarisen optimoinnin tapaan maksimoimaan tai minimoimaan tavoitefunktio, mutta monitavoiteoptimoinnissa on useita tavoitefunktioita (Miettinen 2021). Usean tavoitefunktion optimoinnin takia monitavoiteoptimointi voisi olla hyvä lähestymistapa tukemaan päätöksentekoa oikean elämän erilaisiin ongelmiin, koska oikean elämän ongelmissa on usein useita eri tavoitteita, jotka ovat ristiriidassa keskenään. Yleensä, kun on useita tavoitefunktioita, niin jostain pitää leikata, jotta voidaan saada hyötyä muualta (Miettinen 2021). Monitavoiteoptimoinnin tavoitteena onkin auttaa päätöksentekijää löytämään sopiva tasapaino näiden leikkausten ja hyötyjen väliltä hänen omien tavoitteiden mukaisesti (Miettinen 2021).

Mielenkiintoinen lähestymistapa datavetoiseen päätöksentekoon, jossa käytetään monitavoiteoptimointia ja ehdottavaa analytiikkaa on DEMO (Decision analytics utilizing causal models and multiobjective Optimization). DEMO on temaattinen tutkimusalue Jyväskylän yliopistossa (University of Jyväskylä 2021). DEMO:a hyödynnetään useilla eri aloilla, kuten terveys- ja hyvinvointi-, koulutus-, energia- ja metsäalalla (University of Jyväskylä 2021). DEMO:n datasta päätöksentekoon prosessissa hyödynnetään useita eri menetelmiä, joista keskeisimpiä ovat ennustava ja ehdottava analytiikka, tilastollinen mallintaminen, syyseurauspäätely (engl. causal inference) ja monitavoiteoptimointi (University of Jyväskylä 2021).



Kuvio 1. DEMO:n (Decision analytics utilizing causal models and multiobjective Optimization) hyödyntämä konsepti datasta päätöksentekoon (University of Jyväskylä 2021).

Tulevaisuuden kannalta ehdottavan analytiikan sovellukset tulee olla luotettavia, jotta ne voivat tuoda merkittävää arvoa yrityksille päätöksenteon tueksi. Luottamusta voidaan parantaa hyödyntämällä kaikkea saatavilla olevaa dataa ehdottavan analytiikan sovelluksissa. Ilman kaikkea tarjolla olevaa materiaalia ehdottavan analytiikan ratkaisuvaihtoehdot eivät tule olemaan yhtään sen luotettavampia, mitä ne nyt ovat.

Lähteet

Aref, Molham, Balder ten Cate, Todd J Green, Benny Kimelfeld, Dan Olteanu, Emir Pasalic, Todd L Veldhuizen ja Geoffrey Washburn. 2015. “Design and implementation of the Logic-Blox system”. Teoksessa *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, 1371–1382.

Basu, ATANU. 2013. “Five pillars of prescriptive analytics success”. *Analytics magazine* 8:12.

Berk, Lauren, Dimitris Bertsimas, Alexander M Weinstein ja Julia Yan. 2019. “Prescriptive analytics for human resource planning in the professional services industry”. *European Journal of Operational Research* 272 (2): 636–641.

Bertsimas, Dimitris, ja Nathan Kallus. 2020. “From predictive to prescriptive analytics”. *Management Science* 66 (3): 1025–1044.

Frazzetto, Davide, Thomas Dyhre Nielsen, Torben Bach Pedersen ja Laurynas Šikšnys. 2019. “Prescriptive analytics: a survey of emerging trends and technologies”. *The VLDB Journal* 28 (4): 575–595.

Ghoniem, Ahmed, Agha Iqbal Ali, Mohammed Al-Salem ja Wael Khallouli. 2017. “Prescriptive analytics for FIFA World Cup lodging capacity planning”. *Journal of the Operational Research Society* 68 (10): 1183–1194.

Goyal, Aanchal, E Aprilia, G Janssen, Y Kim, Tarun Kumar, Richard Mueller, D Phan, Abhishek Raman, J Schuddebeurs, J Xiong ym. 2016. “Asset health management using predictive and prescriptive analytics for the electric power grid”. *IBM Journal of Research and Development* 60 (1): 4–1.

Gröger, Christoph, Holger Schwarz ja Bernhard Mitschang. 2014. “Prescriptive analytics for recommendation-based business process optimization”. Teoksessa *International Conference on Business Information Systems*, 25–37. Springer.

- Huang, Teng, David Bergman ja Ram Gopal. 2019. “Predictive and prescriptive analytics for location selection of add-on retail products”. *Production and Operations Management* 28 (7): 1858–1877.
- Ito, Shinji, ja Ryohei Fujimaki. 2017. “Optimization beyond prediction: Prescriptive price optimization”. Teoksessa *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 1833–1841.
- Jank, Merle-Hendrikje, Christian Dölle ja Günther Schuh. 2018. “Product portfolio design using prescriptive analytics”. Teoksessa *Congress of the German academic association for production technology*, 584–593. Springer.
- Kawas, Ban, Mark S Squillante, Dharmashankar Subramanian ja Kush R Varshney. 2013. “Prescriptive analytics for allocating sales teams to opportunities”. Teoksessa *2013 IEEE 13th International Conference on Data Mining Workshops*, 211–218. IEEE.
- Koops, Lily Geraldine. 2020. “Optimized Maintenance Decision-Making—A Simulation-supported Prescriptive Analytics Approach based on Probabilistic Cost-Benefit Analysis”. Teoksessa *PHM Society European Conference*, 5:14–14. 1.
- Lee, Mikyoung, Minhee Cho, Jangwon Gim, Do-Heon Jeong ja Hanmin Jung. 2014. “Prescriptive analytics system for scholar research performance enhancement”. Teoksessa *International conference on human-computer interaction*, 186–190. Springer.
- Lepenioti, Katerina, Alexandros Bousdekis, Dimitris Apostolou ja Gregoris Mentzas. 2020. “Prescriptive analytics: Literature review and research challenges”. *International Journal of Information Management* 50:57–70.
- Marathe, Madhav V, Henning S Mortveit, Nidhi Parikh ja Samarth Swarup. 2014. “Prescriptive analytics using synthetic information”. Teoksessa *Emerging Methods in Predictive Analytics: Risk Management and Decision-Making*, 1–19. IGI Global.
- Matyas, Kurt, Tanja Nemeth, Klaudia Kovacs ja Robert Glawar. 2017. “A procedural approach for realizing prescriptive maintenance planning in manufacturing industries”. *CIRP Annals* 66 (1): 461–464.

Miettinen, Kaisa. 2021. “DEMO Tutorial on (Interactive) Multiobjective Optimization”. YouTube. https://www.youtube.com/watch?v=JK_jJBK5Zgc.

Oesterreich, Thuy Duong, Christian Fitte, Alina Behne ja Frank Teuteberg. 2020. “Understanding the role of predictive and prescriptive analytics in healthcare: A multi-stakeholder approach”.

Padmanaban, Revathy, ja Rajeswari Mukesh. 2020. “HadoopSec: Sensitivity-aware Secure Data Placement Strategy for Big Data/Hadoop Platform using Prescriptive Analytics”. *GSTF Journal on Computing (JoC)* 5 (3).

Roy, Rajkumar, Rainer Stark, Kirsten Tracht, Shozo Takata ja Masahiko Mori. 2016. “Continuous maintenance and the future—Foundations and technological challenges”. *Cirp Annals* 65 (2): 667–688.

Sappelli, Maya, Maaïke HT de Boer, Selmar K Smit ja Freek Bomhof. 2017. “A vision on Prescriptive Analytics”. *ALLDATA 2017*, 54.

Šikšnys, L, TB Pedersen, L Liu ja M Özsu. 2016. “Prescriptive analytics”. *Encyclopedia of Database Systems*, 1–2.

University of Jyväskylä, JYU. 2021. “Decision analytics utilizing causal models and multiobjective Optimization (DEMO)”. <https://www.jyu.fi/demo>.

Von Bischhoffshausen, Johannes Kunze, Markus Paatsch, Melanie Reuter, Gerhard Satzger ja Hansjoerg Fromm. 2015. “An information system for sales team assignments utilizing predictive and prescriptive analytics”. *Teoksessa 2015 IEEE 17th Conference on Business Informatics*, 1:68–76. IEEE.