



Tekoälyn hyödyntäminen luotto- prosesseissa

Milla Maatraiva

2019 Laurea



Laurea-ammattikorkeakoulu

Tekoälyn hyödyntäminen luottoprosesseissa

Milla Maatraiva
Liiketalous
Opinnäytetyö
12/2019, 2019

Milla Maatraiva

Tekoölyn hyödyntäminen luottoprosesseissa

2019

2019

Sivumäärä

54 + liitteet

Opinnäytetyön tavoitteena on selvittää finanssialalla toimivalle yritykselle, kuinka yritykset soveltavat ja hyödyntävät tekoölyä luoton tai yritysten arviointiprosesseissa. Toimeksiantajayrityksen tavoitteena on kehittää luottoprosessia tekoölyn avulla ja soveltaa tekoölyn tuomia mahdollisuuksia yrityksen omassa luottoprosessissa. Aihe on muodostunut sen perusteella, että toimeksiantajayritys ei tällä hetkellä hyödynnä suoranaisesti tekoölyä luottoprosessissa ja käyttöönotto vaatii aiheen tutkimista sekä esimerkkejä. Tekoölyn hyödyntäminen luottoprosessissa on kuitenkin tullut ajankohtaiseksi toimeksiantajayritykselle ja sen mahdollisuuksia on alettu selvittämään.

Opinnäytetyön tietoperusta sisältää teoriaa tekoölystä ja sen mahdollisuuksista liiketoiminnassa sekä finanssiteknologiasta ja siihen liittyvistä ajankohtaisista aiheista. Lisäksi opinnäytetyön tietoperusta sisältää teoriaa luottokelpoisuuden arvioinnista ja luottokelpoisuuteen vaikuttavista tärkeimmistä tekijöistä. Tutkimusmenetelmän tietoperusta koostuu teemahaastatteluiden sekä benchmarkingin teoriasta.

Opinnäytetyö on laadullinen tutkimus, jossa tekoölyn mahdollisuudet luottoprosessissa pyritään selvittämään asiantuntijahaastatteluiden sekä sähköisten lähteiden avulla. Lisäksi tässä opinnäytetyössä pyritään selvittämään benchmarking-menetelmän avulla, kuinka eri yritykset hyödyntävät tekoölyä luottojen tai yritysten arviointiprosessissa ja minkälaiset tekniikat yrityksillä on käytössä. Asiantuntijahaastattelut ovat pääosin teknologia-alan ammattilaisten kanssa käytyjä teemahaastatteluita. Myös benchmarkingissa vertaillaan pääosin haastattelujen avulla eri yritysten tapaa soveltaa tekoölyä yritysten arviointiprosesseissa.

Asiantuntijahaastatteluiden sekä benchmarkingin avulla saatujen tuloksien avulla toimeksiantajayritys pystyy vertailemaan tekoölyn käyttömahdollisuuksia yrityksen sisällä sekä erilaisia yhteistyömahdollisuuksia sellaisien yritysten kanssa, jotka käyttävät tekoölyä yritysarviointeissa. Näiden tulosten avulla toimeksiantajayritys pystyy määrittelemään yrityksen omalle prosessille kannattavimmat tekoölyn käyttömahdollisuudet ja prosessista ne kohdat, joissa tekoölystä voidaan saada lisäarvoa. Lisäksi toimeksiantajayritys saa hyödyllistä tietoa tekoölystä ja sen tulevaisuuden näkymistä rahoitusallalla.

Asiasanat: Finanssiteknologia, tekoöly, koneoppiminen, neuroverkot, luottoprosessi, data-tiede, data-analytiikka, luonnollisen kielen käsittely

Milla Maatraiva

Utilizing Artificial Intelligence in Credit Processes

2019

2019

Pages

54 + attachments

The purpose of this thesis project was to find out how companies apply and utilize artificial intelligence in their credit processes or business valuation process. The aim of the commissioning company is to develop the credit process with the help of artificial intelligence and to apply the opportunities provided by artificial intelligence in the company's own credit process. The topic is based on the fact that the commissioning company does not currently use artificial intelligence directly in the credit process and implementation requires research and examples. However, the use of artificial intelligence in the credit process has become topical for the commissioning company and it has begun to be explored.

The theoretical part of this thesis includes a discussion of artificial intelligence and its possibilities for business as well as financial technology and related topical topics. In addition, the theoretical part of the thesis reviews the subject of credit worthiness assessment and the most important factors affecting credit worthiness. Research methods used in the thesis project, such as the semi-structured interview and benchmarking, are also presented.

The thesis is a qualitative research, whose aim is to examine the possibilities or utilising artificial intelligence in the credit process through expert interviews and online sources. In addition, the aim of this thesis project is to explore through benchmarking how companies utilize artificial intelligence in the credit process or business valuation process, and the kinds of techniques companies are using. Expert interviews mainly consisted of semi-structured interviews with technology professionals. In this thesis project, benchmarking was used as a method to compare different companies' ways of using artificial intelligence.

With the results of the expert interviews and benchmarking, the commissioning company is able to find out the possibilities of artificial intelligence which can be used in their own business and opportunities for collaboration with other companies who provides different kinds of artificial intelligence services to value businesses. These results allow the commissioning company to identify the most profitable uses for artificial intelligence for the company's own process and the points in the process where artificial intelligence can bring more value for the company. In addition, the commissioning company will receive useful information on artificial intelligence and its future prospects in the financial industry.

Keywords: Financial Technology, Artificial intelligence, Machine Learning, Neural Networks, Credit Process, Data Science, Natural Language Processing, Credit process

Sisällysluettelo

1	Johdanto	6
2	Tekoäly.....	7
2.1	Koneoppiminen	7
2.2	Neuroverkot ja syväoppiminen.....	10
2.3	Luonnollisen kielen käsittely (NLP)	15
2.4	Datatiede ja -analytiikka	16
2.5	Ohjelmistorobotiikka (RPA)	18
2.6	Etiikka ja regulaatio	18
3	Finanssiteknologia.....	20
3.1	API-Ohjelmointirajapinnat.....	21
3.2	PSD2 - Toinen maksupalvelu direktiivi.....	22
4	Luottokelpoisuuden arviointi	24
4.1	Tilinpäätöstiedot yritysrahoituksessa	26
4.2	Tunnusluvut.....	27
4.3	Yrityksen kassavirta- ja rahoituslaskelma.....	27
4.4	Vakuudet	28
4.5	Luottokelpoisuuden arviointi toimeksiantajayrityksessä	29
5	Tutkimusmenetelmät ja toteutus.....	31
5.1	Asiantuntijahaastattelut	32
5.2	Benchmarking.....	33
6	Asiantuntijahaastatteluiden tulokset	33
6.1	Tekoälyn hyödyntäminen luottoprosesseissa	34
6.2	Datan hankinta	37
6.3	Ihmisen rooli	38
6.4	Ensimmäiset askeleet	40
7	Benchmarking tulokset	41
7.1	Yritys Z	41
7.2	Yritys Y	42
7.3	Yritys X	42
7.4	Yritys W.....	44
7.5	Valuatum Oy	46
8	Johtopäätökset	50
	Lähteet	54
	Kuviot.....	57

1 Johdanto

Tekoälyä ja edistynyttä analytiikkaa vahvasti hyödyntävät yritykset ovat lähes poikkeuksetta alansa suurimpia menestyjiä ja sen myötä aiheutunut kilpailu sekä tehokkuusvaatimukset lisäävät yrityksille painetta hyödyntää tekoälyä. Pääsääntöisesti tekoäly toimii tukijärjestelmänä liiketoiminnassa, jossa tekoälyjärjestelmiä käytetään apuna esimerkiksi päätöksen teossa. Suomessa otetaan uusia teknologioita käyttöön melko hitaasti ja suomalaiset yrityspäätäjät vaativat usein lukuisia esimerkkejä onnistuneista projekteista ennen kuin teknologia otetaan käyttöön omaan liiketoimintaan. Tekoällyn ja edistyneen analytiikan hyödyt ovat usein niin merkittäviä, että sen tuoma kilpailuetu voi olla ratkaiseva yritystoiminnassa ja pian se ei ole enää pelkästään kilpailuetu, vaan siitä tulee yrityksille enemmänkin selviytymisen edellytys. Tekoälyllä voidaan esimerkiksi automatisoida, nopeuttaa ja tehostaa myös ajattelua vaativia työvaiheita. Lisäksi tekoäly mahdollistaa täysin uudenlaisten liiketoimintamallien toteuttamisen kustannustehokkaasti ja vähentää ihmisten tekemiä virheitä sekä nopeuttaa prosesseja. Tänä päivänä käytössä oleva tekoäly osaa tehdä ihmiselle haastavia tai mahdottomia tehtäviä, kuten tunnistaa vakuutuspetoksia tai ennakoita asiakkaan elämäntilanteen muutoksia, mutta se ei kuitenkaan pysty tekemään muita tehtäviä, kuin mitä sille on opetettu. (Hyödynnä tekoälyä liiketoiminnassa 2019)

Tekoällyn tuoma automaatio vaikuttaa arvonluontiin, työtulojen ja voittojen muodostumiseen sekä verojen lähteisiin. Kaikki tekoällyn tuomat automaatiot vaikuttavat taloudelliseen arvoon kahta eri kautta. Ensimmäisessä vaikutuksessa tuottavuus kasvaa merkittävästi, jolloin kansantuotteen kasvu lisääntyy siitä huolimatta, että esimerkiksi Suomessa työikäisen väestön määrä vähenee. Toisessa vaikutuksessa automaatoratkaisuja tarjoavat palvelu- ja tietotyö - yritykset voivat saavuttaa huikat tulot tuottamansa hyödyn avulla. Tekoällyn menestys riippuu siis siitä, kuinka paljon ja millaista lisäarvoa se tuo (Ailisto H, Helaakoski H, Dufva M, Tuikka T, 2017). Tekoällyn on ennustettu lisäävän Suomen bruttokansantuotetta miljardeilla euroilla sekä kasvattavan yritysten tuottavuutta ja kysyntää, mutta näiden toteuttaminen vaatii tekoällyn laajaa hyödyntämistä sekä merkittäviä investointeja, eikä suurin osa suomalaisista organisaatioista ole vielä tehnyt näitä merkittäviä investointeja tekoälykokeiluihin. Microsoftin teettämän selvityksen mukaan merkityksekkäimmät käytössä olevat tekoälyratkaisut ovat tuottaneet vastaajien mukaan lyhyessä ajassa jopa 2-5 kertaisen tuoton. Kyseisestä selvityksestä käy ilmi, että tekoälystä saatu hyöty seuraavan viiden vuoden aikana uskotaan muodostuvan henkilöstön omien kyvykkyyksien tukena olevasta tekoälystä sekä automaation tuomista tuottavuushyödyistä. (Alila 2019)

Tämän opinnäytetyön tarkoitus on selvittää, miten tekoälyä hyödyntämällä pystytään optimoimaan luottoprosesseja finanssialalla. Tuloksilla pyritään selvittämään, kuinka opinnäytetyön toimeksiantajayritys voi hyödyntää tekoälyä omassa luottoprosessissa lainahakemuksen analysoinnin näkökulmasta, jossa prosessi alkaa siitä, kun lainahakemus

vastaanotetaan, minkä jälkeen päätökseen vaikuttavat tekijät arvioidaan, ja loppuu tehtyyn päätökseen. Työstä on rajattu pois päätöksen teon jälkeinen osa lainaprosessia, kuten esimerkiksi perintä ja asiakaspalvelu sekä muut hallinnointiin liittyvät prosessit. Lisäksi tästä opinnäytetyöstä on rajattu pois lainahakemusta edeltävät tekijät, kuten asiakashankintaan liittyvät prosessit. Työssä keskitytään ainoastaan yritysluottojen arviointiprosessiin eli yrityksen vieraan pääoman ehtoiseen rahoitukseen ja toimeksiantajayrityksen arvioinnin kannalta oleellisiin tekijöihin. Arvioinnissa oleellista ja osa kokonaisuutta on myös vastuuhenkilöiden tausta sekä henkilötakaukset, joten myös henkilöiden luottokelpoisuuden arviointia tekoälyn avulla on siltä osin otettu työssä huomioon. Myös mahdollisten vakuuksien arviointi voi olla luottokelpoisuutta arvioitaessa oleellinen tekijä, joten opinnäytetyössä on pyritty etsimään tekoälyn tuomia ratkaisuja myös vakuuksien arviointiin. Aiheen rajaukseen on vaikuttanut halu syventyä tähän aiheeseen tarkemmin sekä opinnäytetyöhön käytössä oleva aika. Toimeksiantajayrityksessä tekoälyn tuomien ratkaisujen odotetaan helpottavan ja nopeuttavan työtehtäviä, auttavan ennusteiden ja arviointien tekemisessä sekä hoitavan rutiininomaisia työtehtäviä. Opinnäytetyön aihe on ajankohtainen ja opinnäytetyön tekijälle henkilökohtaisesti mielenkiintoinen, sillä tekoäly on merkittävässä roolissa liiketoiminnassa ja erityisesti myös nopeasti kehittyvällä rahoitusallalla.

2 Tekoäly

Tekoälyn tarkoitus on jäljitellä ihmisen tajuntaa sekä suorittaa ja tehostaa tehtäviä kuten ihminen. Esimerkiksi puheen ymmärtäminen, näköhavainnointi ja päätöksenteko ovat tehtäviä, joista tekoäly pystyy suoriutumaan myös ilman käyttäjän avustusta. Finanssialalla tekoälyn hyödyntäminen on tulevaisuudessa tärkeä taito ja datan hyödyntäminen tulee korostumaan tulevaisuudessa entistä enemmän (Finanssialalle, Tekoäly). Tekoäly voidaan jakaa heikkoon ja vahvaan tekoälyyn, mutta tällä hetkellä käytännössä kaikki käytössä oleva tekoäly on heikkoa tekoälyä. Heikolle tekoälylle on opetettu tietynlainen tehtävä eli syötetty ennalta käsky, jonka se kykenee sille annetulla logiikalla ratkaisemaan. Heikko tekoäly päättyy ratkaisuun sen logiikan mukaan, mitä sille on ohjelmoitu, eikä pysty laajentamaan osaamistaan itsenäisesti muille alueille. Vahva tekoäly kykenee ajattelemaan eli saavuttamaan tietoisuuden ja itsenäisen ajattelun kuten ihminen. Vahvaa tekoälyä ei kuitenkaan ole vielä kehitetty (Vahva- ja heikkotekoäly). Useimmiten nykypäivänä käytössä olevat tekoälyjärjestelmät pystyvät korvaamaan ihmisen analyysien- ja päätöksenteossa korkeammalla matemaattisella tarkkuudella ja jatkuvasti kasvavalla nopeudella ja tehokkuudella, mutta heikkotekoäly ei pysty ymmärtämään ihmisen aivoja. (Turunen 2019)

2.1 Koneoppiminen

Koneoppiminen on tekoälyn yksi osa-alueista, jossa on yhdistetty useita eri tieteen osa-alueita kuten tilastotieteitä, filosofiaa ja psykologiaa. Teknologian kehittyttyä myös koneoppiminen on

kehittynyt. Sen tarkoituksena on tuottaa algoritmeja, minkä jälkeen kerättyä dataa syötetään koneelle. Datan ja algoritmien yhteistyöllä koneet pystyvät tuottamaan vastauksia ja oppimaan aina yhä enemmän. Koneoppimisen avulla pystytään erottelmaan yhteyksiä valtavista tiedonmääristä (Finanssialalle, tekoäly). Koneoppiminen pohjautuu tilastotieteeseen, jossa tieto kerätään datasta. Erityisesti tekniikat, kuten esimerkiksi lineaarinen regressio ovat vanhoja menetelmiä, joita edelleen käytetään koneoppimisen menetelminä. Koneoppimisessa on erilaisia aloja, jotka voidaan perinteisesti jakaa ohjattuun oppimiseen, puoliohjattuun oppimiseen, ohjaamattomaan oppimiseen ja vahvistusoppimiseen. Käytettävä koneoppimisen ala valitaan ratkaistavasta ongelmasta riippuen (Elements of AI). Koneoppimisen malleille toimintaa ei ole ohjelmoitu valmiiksi jokaista eri tilannetta varten, vaan ne käyttävät dataa oppimiseen ja luokitteluun. Siinä mallille on annettu tavoite ja malli itse päättää, miten tavoitteeseen päästään, kun taas perinteinen ohjelmistokehitys perustuu logiikkaan, joka on kirjoitettu ohjelmistoon. Koneoppimisen mallit pystyvät ennustamaan lopputuloksia ja kehittymään jatkuvasti käyttäen algoritmeja, jotka oppivat käytössä olevasta datasta jatkuvasti enemmän. Mitä enemmän dataa mallilla on käytössä, sitä tarkempi on lopputulos. Data on usein jaettu kahteen eri dataan, joita ovat opetusdata ja testidata. Opetusdatalla opetetaan ennustamaan tiettyä lopputulosta ja testidata tarkistaa kuinka hyvin opetus onnistui (Merilehto 2018, 27-30). Opetusdataa hyödynnetään koneen opetusvaiheessa, minkä perusteella algoritmi ennustaa oikean vastauksen mille tahansa syötteelle. Kun halutaan tietää, kuinka tarkka mallin tekemä ennuste oli, on testidata tärkeässä roolissa, sillä sen avulla voidaan arvioida mallin tekemän ennusteen tarkkuutta. (Elements of AI)

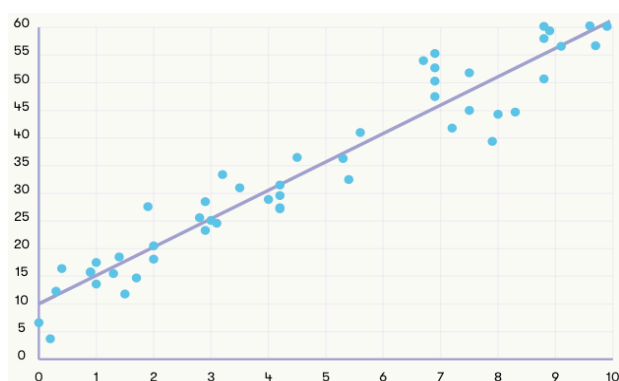
Suurin osa nykyisistä koneoppimisen sovelluksista hyödyntää menetelmänä ohjattua oppimista, jossa ennalta haluttu lopputulos tiedetään ja se annetaan opetusdatan syötteenä ja vasteena mallille (Merilehto 2018, 28). Ideana on, että mallia opetetaan esimerkkiaineiston avulla ja sitä käytetään tyypillisesti luokittelutehtäviin, joissa datanäytteet ovat luokiteltu muutamiin valittuihin luokkiin (Ylén, Bäck 2018). Esimerkki tällaisesta on valokuva liikennemerkistä ja tehtävä on tunnistaa, mikä liikennemerkki on kyseessä eli tunnistaa oikea luokka, kuten esimerkiksi nopeusrajoitus, stop-merkki jne. (Elements of AI)

Ohjaamaton oppiminen on taas ohjatun oppimisen vastakohta, jossa kone päättää itsenäisesti ilman ihmisen avustusta datan perusteella millaisia tuloksia tulisi saada säännönmukaisuuksien ja suhteiden pohjalta, kun taas ohjatussa oppimisessa algoritmeja ohjataan (Merilehto 2018, 27-30). Ohjaamattomassa oppimisessä oikeita luokkia ei ole, kuten ohjatussa. Ohjaamattoman oppimisen menetelmässä datasta on löydettävä jonkinlainen rakenne, kuten samankaltaisten tapausten ryhmiä tai muoto, jossa data pystytään esittämään tärkeimmän muuttujan tai ulottuvuuden avulla. Puoliohjattu oppiminen sijoittuu ohjatun ja ohjaamattoman koneoppimisen väliin. Sitä voidaan käyttää, jos eri kategorioiden väliset rajat ovat jollakin

tavalla epämääräisiä, eikä yksittäistä koneoppimisongelmaa voida helposti lokeroida. (Elements of AI)

Vahvistusoppimisen mallissa koneelle annetaan positiivista tai negatiivista palautetta siitä, kuinka se toimii eri tilanteissa ja sen avulla ohjataan konetta tekemään oikeita päätöksiä (Merilehto 2018, 27-30). Esimerkki tällaisesta on itseoppiva auto, jonka on pystyttävä operoimaan monimutkaisessa ympäristössä, jossa palaute siitä oliko jokin ratkaisu oikea vai väärä, tulee viiveellä. Malli sopii myös muihin sellaisiin tilanteisiin, joissa lopputulos ratkeaa viiveellä, kuten esimerkiksi pelit. (Elements of AI)

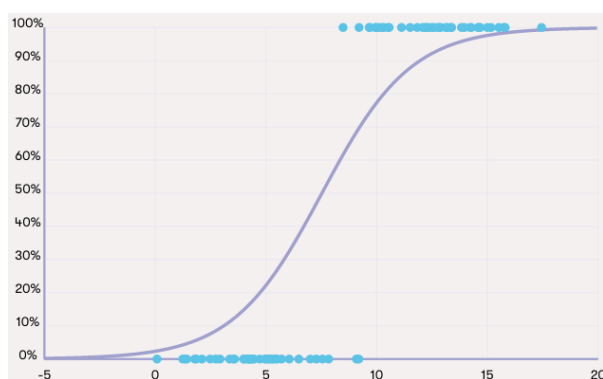
Lineaariregressio on tavanomainen regressioanalyysi, jossa lasketaan vaikutukset yhteen eri syötearvoista ja ne antavat vastaukseksi lukuarvon, jonka ei tarvitse olla kokonaisluku. Lineaarisuus tarkoittaa, että jos jotakin syötettä on kasvatettu tietyllä määrällä, tulosteena saatu arvo kasvaa aina saman verran. Annetusta datasta ja kertoimista lasketaan lineaarikombinaatio. Esimerkiksi jos perunoiden kilohinta on 2 euroa ja ostat 2,5 kiloa perunoita, tulee kokonaishinnaksi $2,5 \times 2$ eli 5 euroa. Perunoiden määrä on tässä tapauksessa datasta saatu syöte ja hinta painokerroin. Tällaisessa esimerkissä sekä kertoimet että syötedata on tunnettu. Jos käytössä on syötedata ja vasteet (oikeat vastaukset), valitaan kertoimet siten, että lineaariregressio ennustaa oikeat vastaukset mahdollisimman tarkasti. Tällaiseen tehtävään soveltuu ohjattu koneoppinen. Esimerkiksi jos ostoskorin hinta on yhteensä 20 euroa ja se sisältää eri tuotteita, emme voi päätellä tuotteiden hintoja, mutta riittävän monen ostoskorin perusteella voimme kuitenkin yleensä ratkaista sen. Kuviossa 1 on esitetty sekä datapisteet että lineaarinen malli.



Kuvio 1: Elements Of AI, Lineaarinen regressio kuvina

Logistinen regressioanalyysi on lineaariregression lähisukulainen, jota käytetään tapauksissa, jossa selitettävä muuttuja voi saada kaksi eri arvoa. Päätöksenteon apuna ollessa, se ei tarvitse suurta määrää dataa, sillä sitä käytetään, kun ennustuksen kohteena on kategorinen muuttuja.

Esimerkiksi tuleeko maksuhäiriöitä vai ei tai onko asunnon myyntiaika alle vai yli kuukauden ja mitkä tekijät vaikuttavat siihen. Siinä syötteeksi valitaan päätökseen vaikuttavat tekijät, joita esimerkiksi kiinteistön myyntiajan arvioinnissa voi olla koko, hinta per neliö, postinumeroalue ja muita vaikuttavia tekijöitä, minkä jälkeen valituille syötteille annetaan painoarvot, joita säädetään. Sen jälkeen kaikki annetut syöte- ja painoparit kerrotaan keskenään ja summataan yhteen, minkä jälkeen summa muunnetaan vasteeksi (Merilehto 2018, 27-30). Kuviossa 2 on esimerkki, jossa siniset pisteet ovat opiskelijoita ja alhaalla oleva linja kertoo kokeeseen valmistumiseen käytetyn ajan (tunnit). Ylhäällä olevat pisteet ovat oppilaita, jotka ovat läpäisseet kokeen ja alhaalla olevat eivät. Vasemman reunan linjalla voimme ilmaista todennäköisyyden, jolla oppilas läpäisee kokeen sen perusteella, kuinka monta tuntia valmistautumiseen on käytetty. Todennäköisyys on saatu logistisesta regressiomallista, jossa käyrä kuvaa tätä todennäköisyyttä, mikä perustuu samankaltaisiin kertoimiin kuin lineaariregressiossa käytetyt kertoimet. (Elements of AI)



Kuvio 2: Elements Of AI, Logistinen regressio

Koneoppiminen pystyy ennustamaan tulevaa huomioidessa jatkuvasti uutta dataa, josta sen arvo tulee liiketoiminnalle. Käytössä olevasta datasta haetaan jatkuvasti poikkeamia ja piilossa olevia malleja. Koneoppiminen on liiketoiminnassa yksi väline kilpailuedun saavuttamiseksi pohdittaessa keinoja, joilla dataa ja analytiikkaa voidaan käyttää ja hyödyntää. Päivittämällä uutta dataa olennaisista lähteistä, huolehditaan siitä, että käytettävissä on aina tuoreimmat ennusteet, jolloin ennuste on tarkka ja sitä hyväksi käyttäen voidaan toimia tehokkaasti. (Merilehto A 2018, 32)

2.2 Neuroverkot ja syväoppiminen

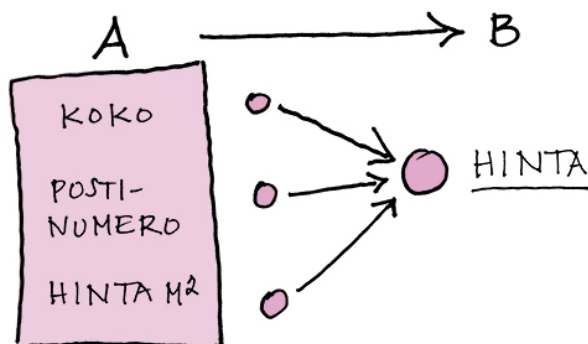
Keinotekoiset neuroverkot ovat yksi koneoppimisen menetelmistä, mutta myös ihmisen aivoista löytyy niin sanottuja oikeita biologisia neuroverkkoja, minkä takia aivoja tutkitaan, kun älykkäiden järjestelmien rakentamiseen haetaan inspiraatiota. Keinotekoisien neuroverkkojen

neuronit ovat yleisesti ottaen huomattavasti yksinkertaisimpia kuin oikeat hermosolut, sillä niissä ei huomioida oikeiden hermosolujen sisäisiä toimintamekanismeja. Keinotekoiset neuroverkot koostuvat matemaattisia yksiköitä, jotka oppivat havainnoimalla ja saavuttavat tavoitteita oppimansa kautta. Neuroverkko sisältää yhdessä toimivia neuroverkkokerroksia, joissa tapahtuu syväoppiminen. Näillä kaikilla neuroverkkokerroksilla on oma tehtävänsä: alimmat tunnistavat ääri viivoja käsin kirjoitetuista numeroista tai esimerkiksi jonkin tietyn eläimen pään muotoja. Pidemmälle edettäessä kerroksissa, lopulta neuroverkko pystyy tunnistamaan esineitä ja asioita. Neuronit tekevät kerralla vain yhtä asiaa ottaen sisään syötteen, prosessoivat sen ja antavat eteenpäin seuraavalle neuronille eli yksinkertaistettuna lähettävät toisilleen viestejä ja vastaanottavat niitä. Neuronit ovat solukeskuksista ja johdoista koostuvia tiedonkäsittelijöitä, joiden avulla ne kytkeytyneet toisiin neuroneihin. Viejähaarake voi kytkeytyä joko yhteen tai useampaan tuojahaarakkeeseen liittymäkohdissa eli synapseissa. Avain asemassa oppimisessa sekä muistin toiminnassa on toisiinsa kytkeytyneet neuronit, joissa jokainen neuroni reagoi tietyllä tavalla vastaanottamiinsa signaaleihin ja ajan mittaan mukautuu. Neuroverkot osaavat muun muassa kuvailla reaaliaikaisesti videolla tapahtuvia asioita, muuttaa puhetta tekstiksi ja kääntää ihmistä paremmin kieliä tietyissä kielipareissa sekä tunnistaa ja nimetä eläimiä valokuvista (Merilehto 2018, 45-55 & Elements of AI).

Myös neuroverkot ja etenkin syvät neuroverkot tarvitsevat suuren määrään dataa toimiakseen - mitä enemmän käytössä on dataa, sitä paremmin ne oppivat. Keinotekoinen neuroverkko koostuu toisiinsa kytkeytyneistä neuroneista, jotka harjoittelevat tiettyä toimintaa suurella datamäärällä (Merilehto A. 2018, 49-50). Neuroverkkoarkkitehtuuri mahdollistaa samanaikaisesti valtaviin datamääriin prosessoinnin, sillä jokainen neuroni pystyy prosessoimaan dataa toisistaan riippumatta, toisin kuin perinteisessä tietokonearkkitehtuurissa, jossa pystytään kyllä käsittelemään dataa, mutta ei pitkäkestoisella tallennuskapasiteetilla. Neuroverkkojen pitkäkestoinen muisti perustuu neuronien välisiin yhteyksiin, joihin tallentuu tietoa joka kerta, kun yhteydet ja painokertoimet mukautuvat (Elements of AI). Neuroverkkojen tehtävä onkin jäljitellä ihmisaivojen tiettyjä toimintamalleja ja sen kehitys tulee jatkumaan, sillä laskentateho kehittyä ja kapasiteetti kasvaa. Laaja sovellettavuus on neuroverkkojen etu verrattaessa moniin muihin koneoppimisen malleihin, sillä arkkitehtuuriltaan saman tyyliä verkkoja voidaan käyttää erilaisien ongelmien ratkaisuihin ja verkkoja erottavat toisistaan ainoastaan opetuksen seurauksena saadut painokertoimet. (Merilehto A. 2018, 49-50)

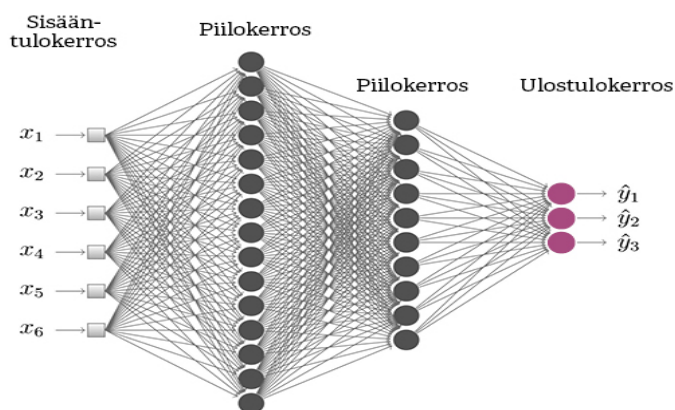
Neuroverkoille ei tarvitse erikseen kertoa mitkä asiat vaikuttavat mihinkin, sillä neuroverkoille annetaan vain syöte, jonka tuloksena on vaste. Esimerkiksi asunnon hinta saadaan neuroverkossa antamalla syötteenä asunnon koko ja postinumero, jolloin saadaan vasteksi asunnon hinta, kuten kuviossa 3 on esitetty. Mitä enemmän dataa on käytössä, sitä tarkemmin

neuroverkkoa voidaan kouluttaa saamaan tarkempi hinta. Neuroverkon harjoittamisessa datalla on merkittävä rooli, sillä vain riittävällä määrällä dataa neuroverkko pystyy toimimaan parhaimmillaan. Tässä tapauksessa riittävä data voi tarkoittaa jopa satojatuhansia esimerkkejä. Neuroverkkojen toimivuudessa on olennaisempaa data kuin sen käyttämä koodi. (Merilehto A 2018, 51)



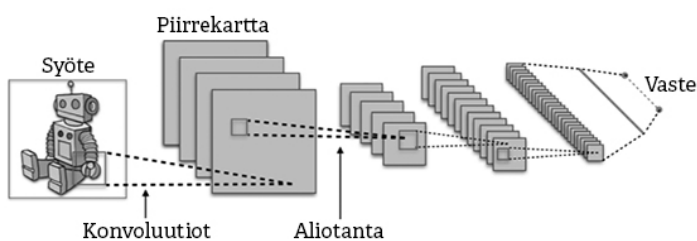
Kuvio 3: Syöte ja vaste neurovekoissa, Merilehto 2018, 51

Monikerroksinen perseptroniverkko koostuu yhteen tai useampaan kerrokseen sijoitetuista keinotekoisista neuroneista, jota voidaan käyttää yksinkertaisena luokittimena. Sen avulla kahteen eri luokkaan kuuluvia esimerkkejä voidaan erotella. Nämä kaksi luokkaa jaotellaan piilokerrokseen ja ulostulokerrokseen. Piilokerroksien neuronit saavat syötteensä muilta neuroneilta ja niiden tuloste toimii seuraavien neuroneiden syötteinä. Ulostulokerros on viimeisenä verkossa ja se tuottaa koko verkon tulosteen, eli yksittäisen kerroksen neuronit saavat syötteensä aiemman kerroksen neuroneilta ja antavat saamansa tulosteen seuraavalle kerrokselle. On olemassa myös kolmas kerros, nimeltään sisääntulokerros, jonka ainoa tehtävä on ottaa vastaan neuroverkkoon sisään tulevaa dataa ja välittää se eteenpäin ensimmäiselle piilokerrokselle. Tässä kerroksessa ei tapahdu oppimista, joten sitä ei pidetä neuronina. Esimerkki neuroverkon rakenteesta kuviossa 4. Perseptronineuronin painojen oppiminen datasta tapahtuu perseptronialgoritmin avulla. (Merilehto A, 2018, 51-52 & Elements of AI)



Kuvio 4: Neuroverkko, Merilehto 2018, 52

Konvoluutioneuroverkot ovat neuroverkkojen alalaji, jotka prosessoivat kuvista koostuvaa dataa. Niiden joidenkin kerroksien neuronin eivät ole yhteydessä kaikkien seuraavien kerroksien neuroneissa, kuten useissa muissa neuroverkoissa. Konvoluutioneuroverkoissa syötteenä käytetään kuvaa ja jokaisessa kerroksessa kuvaan käytetään filttereitä. Konvoluutioneuroverkot sopivat erityisen hyvin sellaisiin ongelmiin, mitkä ihmisen on helppo ratkaista, kuten esimerkiksi jonkin tietyn eläimen tai numeron tunnistaminen kuvasta. Tällaisella neuroverkolla läpikäydään esimerkiksi numeroita tai kasvoja sisältäviä datamassoja. (Merilehto A, 2018, 53)



Kuvio 5: Konvoluutioneuroverkko, Merilehto 2018, 53

Vastavirta-algoritmia käytetään neuroverkkojen opetusalgoritmina siten, että neuroverkon saaman syötteen käsittelyn loputtua, sen antamaa tulosta verrataan optimaaliseen lopputulokseen, jonka jälkeen neuronit säätävät toimintaansa valitun algoritmin mukaan. Näin seuraavalle kerrokselle tultua tulos on yhä lähempänä haluttua lopputulosta. Neuroverkko antaa jonkin todennäköisyyden sille, kuinka lähellä haluttua lopputulosta vaste on. Vastavirtausta käytetään tekniikkana ohjatussa oppimisessa, jossa ulostuloarvoja verrataan

haluttuihin arvoihin. Opetusdatasta saadusta tiedosta ja annetusta tuloksesta lasketaan neuroverkon tuottama virhe, joka lähtee takaisin ulostulokerroksesta sitä edeltävään kerrokseen. Jokainen kerros ja siellä oleva solmu saa virheestä palautteen, joka suhteutetaan kyseiseen kerrokseen eli solmun osuuteen ulostuloarvosta. Kaiken tämän tiedon perusteella painokertoimia päivitetään siten, että ulostuloarvo lähenee haluttua arvoa. Neuroverkko onnistuu löytämään kaavan toistoja lisäämällä, minkä avulla tulos on riittävä suhteessa annettuun tavoitteeseen. Neuroverkkokerroksia lisäämällä syntyy syviä neuroverkkoja. (Merilehto 2018, 54-55)

Syväoppimisessa puolestaan hyödynnetään usean kerroksen neuroverkkoja - mitä useampi kerros, sitä syvempi verkko. Verkon syvyys tekee oppimisesta monimutkaisempaa, joten se edellyttää suurta määrää dataa sekä enemmän laskentatehoa, mutta näiden edellytysten toteutuessa monimutkaisempien ilmiöiden oppiminen on mahdollista. Syväoppiminen perustuu graafisten näytönohjainten (GPU) tehojen lisääntymiseen (Merilehto 2018 & Elements of AI). Tekoäly prosessori GPU pystyy tekemään 320 biljoonaa laskutoimitusta per sekunti, mikä kuvastaa hyvin sitä, kuinka laskentateho mahdollistaa täysin uudenlaisen datan hyödyntämisen (Ojanperä 2017). Jos syväoppimiseen soveltuvalle järjestelmälle näytetään riittävä määrä etukäteen luokiteltuja kuvia esimerkiksi jostakin eläimestä, kykenee järjestelmä opettamaan itse itsensä tunnistamaan kuvassa olevan eläimen. Tällöin esimerkkikuvien tulee sisältää sekä oikeita että vääriä esimerkkejä eläimestä eli käytössä on oltava erilaista dataa tilanteesta. Liiketoiminnassa esimerkkinä voi olla henkilö, joka on saanut myönteisen luottopäätöksen ja henkilö, joka on saanut kielteisen luottopäätöksen. Näin saadaan dataa toimivasta prosessista sekä ei toimivista prosesseista. (Merilehto 2018, 56-59)

Syväoppimisessa järjestelmät voivat jo nyt opettaa itse itseään eli koneoppimisen malli, jossa avuksi on tarvittu koneoppimisen, datan ja laskentatehon optimoinnin osaavia ihmisiä, voi nyt opettaa itse itseään. Koneoppimiseen opetettu ihminen valitsee parhaan mallin kaikista käytössä olevista malleista, mutta syväoppimisen malli tekee juuri sen esityön itse. Automatisoitu koneoppiminen (AutoML) on jo nyt pystynyt opettamaan koneoppimisen ohjelmistolle, kuinka ohjelmoida koneoppimisen ohjelmistoja ja joissakin näistä tapauksissa se on suoriutunut ohjelmoinnista paremmin, kuin tutkijoiden suunnittelema tekoäly. Testissä, jossa ihmisen ja koneen ohjelmoimia kuvantunnistusjärjestelmiä verrattiin toisiinsa, onnistui koneoppimisen ohjelmoima järjestelmä 43 prosentissa tehtävistä, kun taas paras ihmisen ohjelmoima järjestelmä onnistui 39 prosentissa tehtävistä. Tänä päivänä AutoML pystyy kuitenkin ohjelmoimaan pääasiassa vain tavanomaisia tekoälyohjelmia, mutta tulevaisuuden tavoitteena on, että jatkossa ohjelmistot pystyvät suorittamaan toistuvista rutiinitehtäviä,

joihin ohjelmoijat käyttävät tänä päivänä paljon aikaa. Tutkijoiden arvion mukaan tämän asian kehitys nopeuttaisi tekoälyn kehitystä merkittävästi. (Merilehto A, 2018, 56-59)

2.3 Luonnollisen kielen käsittely (NLP)

Luonnollisen kielen käsittely (Natural Language Processing, NLP) on koneoppimisen osa-alue, joka käsittelee tietokoneiden ja ihmisten välistä vuorovaikutusta luonnollisen kielen avulla. Se pystytään saavuttamaan neuroverkkoihin perustuvalla syväoppimisella. Luonnollisen kielen käsittelyä käytetään luonnollisen tekstin ja puheen analysointiin sekä tuottamiseen. Ala sisältää konekääntämisen, automaattisen puheentunnistuksen, puhesynteesin, tekstintunnistuksen, älykkään tekstinsyötön ja puheen kääntämisen osa-alueet (Ailisto 2018). Luonnollisen kielen käsittely edellyttää algoritmien soveltamista luonnollisten kielisääntöjen tunnistamiseen ja purkamiseen siten, että järjestämätön data muunnetaan tietokoneiden ymmärtämään muotoon. Kun teksti on haettu, kone hyödyntää algoritmeja poimiakseen jokaisesta lauseesta merkityksen ja kokoaa niistä olennaiset tiedot. (Garbade 2018)

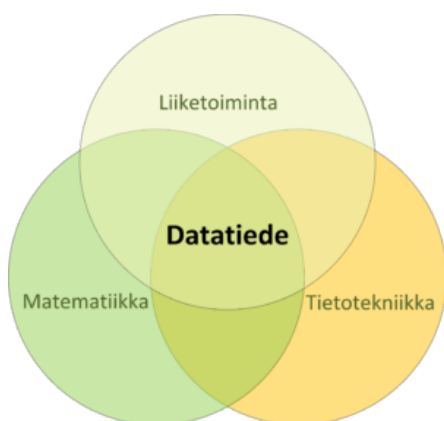
Luonnollisen kielen käsittelyllä voidaan analysoida strukturoimatonta eli järjestämätöntä dataa. Strukturoitua eli järjestettyä dataa löytyy esimerkiksi laskentataulukoista ja tilikirjoista, joita voidaan jo nyt analysoida kattavasti data-analytiikan ja automatiikan avulla. Nykyisin yli 80% datasta on kuitenkin järjestämättömässä muodossa, kuten sopimuksissa, sähköposteissa, PDF-tiedostoissa ja muissa dokumenteissa, joten on tärkeää kehittää digitaalisia assistentteja, jotka kykenevät lukemaan tätä dataa ja tunnistamaan sieltä merkittävät tiedot. Älykkäiden koneiden prosessointivoimaa hyödyntäen pystymme käyttämään korrelaatioteoriaa datan louhimiseen järjestämättömistä lähteistä. Yksi NLP-aplikaatio kykenee lukemaan tuhansia tai jopa miljoonia dokumentteja käyttäen siihen vain murto-osan siitä ajasta, mikä ihmisellä menee saman tehtävän tekemiseen. (Aaltonen 2019)

Luonnollisen kielen käsittelyn avulla pystytään poimimaan tekstimuodossa olevaa dataa, kuten hankintasopimuksia, uutisia ja yhtiöjärjestelmiä sekä sosiaalisessa mediassa olevaa tietoa. Usein tällaiset tiedot vaativat ihmiseltä käsityötä ja siksi niiden läpikäyminen on hidasta. Kehittynyt luonnollisen kielen käsittely on mahdollistanut tällaisen datan keräämisen osana koneoppimisprojektia. Yksi monista nykyään käytetyistä tekniikoista on tekstin ryhmittely, jolla tarkoitetaan objektien ryhmittämistä erilaisiin joukkoihin tai ”klustereihin”. Tekstin luokittelu, tyypillisesti konvoluutiohermoverkoilla, on ohjattu oppimismenetelmä, jossa oppiminen tapahtuu esimerkeistä ja niiden tunnisteista. ”Klusterointi” puolestaan on valvomaton menetelmä, missä malli oppii itsenäisesti ilman käyttäjä avustusta. Luonnollisen kielen käsittely auttaa hyödyntämään valtavaa määrää tekstiä, jota käytetään tyypillisesti laki- ja rahoitusallalla (Alanen 2019). Luonnollisen kielen käsittelyn avulla voidaan löytää ja kerätä tietoa dokumenteista ja jalostaa se sellaiseen muotoon, että koneoppimisjärjestelmät voivat

analysoida sen. Luonnollisen kielen käsittelyn täyden hyödyntämisen mahdollisuudet ovat vasta selviämässä. (Aaltonen 2019)

2.4 Datatiede ja -analytiikka

Datatiede mahdollistaa sen, että kaiken kokoisista tietomassoista löydetään riippuvuuksia ja säännönmukaisuuksia, joiden avulla luodaan ymmärrys menneestä, nykytilasta ja tulevaisuudesta. Data kerätään ja sen jälkeen käsitellään matemaattisilla ja tilastotieteellisillä menetelmillä. Datatieteen sovellukset vaihtelevat tilastollisesta analyysistä keinoälyä hyödyntäviin päätöksiä tekeviin robotteihin asti, mikä tekee datatieteestä erittäin skaalautuvan apuvälineen. Datatiede on yhdistelmä liiketoimintaa, tietotekniikkaa ja matematiikkaa (Kuvio 4). Lisäksi siihen voidaan yhdistää tietojenkäsittelytiedettä sekä tilastotiedettä, jossa voidaan hyödyntää muun muassa menetelmiä ja tekniikoita, kuten koneoppiminen, tiedonlouhinta, ennakoiva-analytiikka, todennäköisyyslaskenta, tietokannat, datan varastointi- ja kompressointi ja tekoäly. (Seppä K & Seppä A)



Kuvio 6: Datatieteen työkalupakki, Kristiina ja Anniina Seppä

Tyypillisesti ennusteita tuotetaan edistyneen analytiikan avulla, mikä perustuu käytännössä koneoppimisen algoritmeihin. Historiadata ja tilastotiede eivät yksinään riitä ennustamaan tulevaisuutta, mutta koneoppiminen työkaluna auttaa ennustamaan jopa yksittäisen asiakkaan käyttäytymisen (Opas tekoälyn hyötykäyttöön). Ennusteen tarkkuus riippuu siis ennustettavasta ilmiöstä sekä datasta, jonka laatua voidaan kuitenkin parantaa tilastotieteen menetelmin. Yrityksillä on usein hallussaan paljon dataa, mutta sitä käytetään vähän. Käyttökelpoista dataa on saatavilla myös julkisista lähteistä. On tärkeää, että yrityksillä on tietojenhallintasuunnitelma (Data Governance), jossa datan saatavuutta, yhdenmukaisuutta

sekä tietoturva on kuvattu. Tämä tietojenhallintasuunnitelma on myös tärkeä viedä käytännön tasolle. (Tekoäly Ostajan opas 2019)

Yleisnimitys valtaville datamäärille, joissa ei sovelleta perinteisiä datahallintotapoja on Big data tai massadata. Big dataan voidaan liittää kolme määrettä, joita ovat määrä (volume), nopeus (velocity) ja monimuotoisuus (variety). Määrä -termillä tarkoitetaan dataa, joka ei ole käsiteltävissä siedettävässä ajassa laitteilla tai ohjelmistoilla, jotka ovat yleisesti käytössä olevia. Nopeus -termillä tarkoitetaan sitä, että data kertyy ja kasaantuu nopeasti usein automaattisesti monista eri lähteistä ja monissa eri muodoissa. Monimuotoisuus -termillä tarkoitetaan dataa, jolla ei ole välttämättä ollenkaan rakennetta tai se on määritelty löyhästi, mistä johtuen sen analysointi sellaisenaan on hankalaa. (Analytiikka ja big data 2019)

Yleisesti ottaen rahoitus alalla on jo nyt suuri määrä dataa. Petoksien estämiseen sekä rahanpesua koskevien sääntöjen noudattamiseen useimmat rahoituslaitokset käyttävät Big Dataa. Sitä mukaan, kun rahoituslaitoksissa on siirrytty käyttämään tietokoneita ja matkapuhelimia petosten mahdollisuus on moninkertaistunut. Big dataa hyödyntämällä esimerkiksi vakuutusyhtiöiden matemaatikot pystyvät tarkemmin laskemaan kuinka hyvin vakuutuslaitoksen riskistö vastaa sen riskinottohalukkuutta, miten sen tuotteet käyvät kaupaksi, missä on parannamisen varaa ja millaiset trendit vallitsevat. Tiedonkäytössä ollaan siis siirtymässä ennakkontiin taannehtivasta kirjanpidosta. Big dataa hyödyntämällä, jotkut pankit ovat pystyneet kaksinkertaistamaan lainan annon ja vähentämään luottotappioiden määrää neljänneksellä. (Erinomaista liiketoimintaa 2013)

Big Datan hyödyntäminen antaa yrityksille mahdollisuuden hankkia uutta ymmärrystä, minkä pohjalta voidaan tehdä parempia päätöksiä. Internetpalveluihin ja sivustoille kertyy päivittäin valtava määrä uutta sisältöä, kuten tekstejä, kuvia ja videoita, mistä johtuen internet on tärkeä datan lähde. Näiden erilaisten tietojen yhdistämisellä eri datalähteistä, saadaan kerättyä suuri määrä dataa. Se on myös mahdollistanut laajojen opetusaineistojen kokoamisen, mikä on ollut merkittävä edistysaskel myös koneoppimisen kehityksessä (Ylén, Bäck ym. 2018). Maailmassa on valtava määrä digitaalista tietoa, jota yritykset, julkiset sektorit ja tavalliset kansalaisetkin tuottavat jatkuvasti lisää. Valtaosa maailmassa olevasta digitaalisesta datasta on luotu johonkin tarkoitukseen, mutta sen tarkoitusta ei ole tarkkaan määritelty, eikä tietoa ole linkitetty muuhun tietoon. Tällöin kyseessä on strukturoimaton eli järjestämätön data, jonka määrä kasvaa eksponentiaalisesti. Kuten edellisessäkin luvussa todettiin, arvioiden mukaan kaikesta digitaalisesta datasta yli 80 % on strukturoimatonta dataa. Strukturoimatonta dataa syntyy muun muassa sosiaalisessa mediassa julkaistuista valokuvista, chat-keskusteluista, mediassa tehdyistä kommentteista, Facebook-päivityksistä, yrityksiä tuottamista pdf-tiedostoista, videotallenteista, asiakaspalvelupuheluiden tallenteista sekä yhä enemmän teollisen internetin sensoritiedoista. Strukturoimaton data on usein hyödyntämätön resurssi, jota Googlen kaltaiset yritykset keräävät internetistä ja jalostavat siitä tekoälyn avulla

strukturoitua eli järjestettyä dataa. Tekoälyn avulla voidaan ymmärtää luonnollista kieltä, tunnistaa hahmoja sekä analysoida ja löytää merkityksiä valtavista tietomassoista. Aiempaa monipuolisempi ja nopeampi analyysi saadaan päätöksenteon pohjaksi, kun strukturoimattoman datan analyysit yhdistetään jo olemassa olevaan strukturoituun dataan (Ruuska & Ojanperä, 2018). Strukturoidulla datalla tarkoitetaan perinteisesti tietokantoihin tallennettua dataa, joka on helposti käsiteltävissä. Esimerkkejä strukturoidusta datasta ovat numerot, päivämäärät sekä sanaryhmät ja numerot (nk. merkkijonot), kuten nimet ja osoitteet. (Hurwiz, Nugent, Halper & Kaufman, 2013)

2.5 Ohjelmistorobotiikka (RPA)

Ohjelmistorobotiikka automatisoi manuaalisia ja rutiininomaisia toimintaprosesseja, joissa päätökset ovat kuvattavissa perinteisten ehtolauseiden avulla ja tarvittava data on digitaalisessa ja rakenteellisessa muodossa. Sitä sovelletaan siitä syystä usein yksinkertaisiin sääntöihin perustuviin työtehtäviin. Ohjelmistorobotin käyttö perustuu erilaisiin tietojärjestelmiin, joista se hakee tietoa hyödyntäen eri tietojärjestelmissä olevaa dataa ja muodostaa niiden pohjalta esimerkiksi raportteja normaalien käyttöliittymien kautta (Kääriäinen, Aihkisalo ym. 2018 & Tuomisto 2019). Robotit tekevät pääsääntöisesti sellaisia tehtäviä, joiden tekemiseen on selkeät säännöt, mutta niitä ei ole kannattavaa tai välttämätöntä teettää ihmisellä, kuten säännölliset ja tiettyä kaavaa noudattavat tukitoiminnot eli niin kutsutut Back office -toiminnot. Ohjelmistorobotti ei kuitenkaan pysty päättämään tai tunnistamaan syy-seurasuhteita. Toimiakseen optimaalisesti, tarvitsee robotti opetusta, ylläpitoa sekä ajoympäristön. Ohjelmistorobotilla, koneoppimisella sekä tekoälyllä on suuri merkitys prosessien optimoinnissa. Tekoäly ja robotiikka voidaan soveltaa mm. raporttien, manuaalisten tai esimerkiksi lakitekstien läpikäymiseen sekä niiden datasta poimituihin keskeisimpien asioiden tunnistamiseen. Tekoäly tekee päätelmiä robotin useista tietolähteistä löytämistä tiedoista. (Toivola 2017)

2.6 Etiikka ja regulaatio

Etiikalla tarkoitetaan sosiaalisia käytännesääntöjä, joita pidetään yleensä hyvänä tai oikeana. Vaikka kenellä tahansa voi olla yksilöllisiä henkilökohtaisia etiikoita, toimimme pääasiassa yhteiskunnan jäsenenä tai joinakin muina ihmisryhminä, joilla on jonkinlainen käsitys siitä mikä on hyvää ja oikeaa tai mikä on huonoa tai väärin. Usein nämä liitetään säädöksiin tai lakeihin (Turunen 2019). Tekoälyä pidetään usein itseisarvona erilaisten ongelmien ratkaisuisissa. Oleellisia kysymyksiä tekoälyn etiikkaa pohdittaessa on esimerkiksi kansalaisten turvallisuuden, yhteiskunnan luottamuksen säilyminen, palvelujen tasa-arvoinen

saavutettavuus, kuulluksi tulemisen mahdollisuus sekä ratkaisujen läpinäkyvyys ja oikeutus ihmisarvoon. (Leikas 2018)

Tekoälyjärjestelmät vähentävät manuaalisen työn määrää ja lisäävät johdonmukaisuutta päätöksentekoihin. Kun neuroverkkoihin perustuva tekoälyjärjestelmä tekee päätöksiä, perustuu päätös suureen datamäärään, mikä tarkoittaa sitä, että päätöksenteossa on ollut enemmän tietoa käytössä, kuin ihmisen toimiessa. Esimerkiksi luottopäätöksiä tehdessä käytössä voi olla neuroverkkoihin perustuvia tekoälypohjaisia sovelluksia luoton pisteytyksen määrittämiseen, jolloin järjestelmä käyttää päätöksenteossa tilastollisia menetelmiä ja yhdistää tietoja eri tietokannoista. Tekoälyn tekemästä päätöksestä voidaan todeta, että luottoa ei pidä myöntää, jolloin lainanhakijoiden pyytäessä lisätietoja, saavat he perusteluksi pelkästään järjestelmään syötetyistä tietojoukoista pohjautuvan epämääräisen perustelun. Loppujen lopuksi edes yrityksessä työskentelevä ihminen ei pysty perustelemaan kunnolla mikä oli luotonhylkäämisen todellinen syy, jolloin kukaan ei tiedä tekikö järjestelmä syrjivän päätöksen esimerkiksi rodun, kotikaupungin tai sukupuolen perusteella. Tätä kutsutaan ”black box” ongelmaksi, jossa neuroverkkoihin ja syväoppimiseen perustuvien tekoälyratkaisujen tekemiä päätöksiä on vaikea ymmärtää, sillä päätökset perustuvat valtavaan määrään dataa. Vaikka tällaiset järjestelmät voivat olla erittäin tarkkoja, eivät ne yleensä ole täysin läpinäkyviä, eikä aina edes asiantuntijat tai järjestelmän tekijät pysty selittämään syitä päätöksille. Mitä tärkeämpiä ja vaikuttavampia päätöksiä neuroverkkoihin perustuvat tekoälyjärjestelmät tekevät, sitä enemmän tulisi kiinnittää huomiota niiden eettisiin seurauksiin. Jos yhteiskunta ja ihmiset antavat yhä enemmän päätösvaltaa tällaisille tekoälyjärjestelmille, on tärkeää tietää, mitä mustan laatikon sisällä tapahtuu. (Turunen & Telakivi 2019)

Luotettava tekoäly ei koske pelkästään tekoälyjärjestelmän luotettavuutta, vaan myös kaikkien järjestelmän elinkaareissa osallisena olevien prosessien ja toimijoiden luotettavuutta. Euroopan komission perustaman tekoälyasiantuntijaryhmän tekemän ohjeistuksen mukaan luotettava tekoäly sisältää kolme osaa (2018):

- Sen tulee olla laillinen ja varmistaa kaikkien sovellettavien lakien ja asetusten noudattaminen. Oikeudellisia lähteitä ovat mm. EU:n primaarioikeus (Euroopan unionin perussopimukset ja sen perusoikeuskirja), EU:n johdettu oikeus (kuten yleinen tietosuoja-asetus, tuotevastuudirektiivi, asetus henkilötietojen vapaasta liikkuvuudesta, syrjinnän vastaiset direktiivit, kuluttajalaki ja työturvallisuutta ja työterveyttä koskevat direktiivit), YK:n ihmisoikeussopimukset ja Euroopan neuvoston

yleissopimukset (kuten Euroopan ihmisoikeussopimukset) sekä lukuisia muita EU:n jäsenvaltioiden lakeja.

- Sen tulee olla eettinen, osoittaa eettisten periaatteiden ja arvojen kunnioittaminen ja niiden noudattamisen varmistaminen. Lait eivät aina ole ajan tasalla tekniikan kehityksen kanssa. Ne voivat toisinaan olla eettisten normien mukaisia tai ne eivät yksinkertaisesti sovellu tiettyihin kysymyksiin. Jotta tekoälyjärjestelmät olisivat luotettavia, tulisi niiden myös olla yhdenmukaisia eettisten normien kanssa.
- Sen tulee olla vankka sekä teknisestä että sosiaalisesta näkökulmasta, sillä myös tekoälyjärjestelmät voivat aiheuttaa tahattomia haittoja. Tällaisten järjestelmien on toimittava turvallisella, luotettavalla ja varmallalla tavalla. Tahattomien vaikutusten estämiseksi on varmistettava suojaustoimenpiteet ja varmistettava, että tekoälyjärjestelmät ovat kestäviä. Täten on siis varmistettava, että tekoälyjärjestelmä on vankka.

3 Finanssiteknologia

Moderni yhteiskunta ja kehittynyt talous tarvitsevat kattavia finanssipalveluita. Terminä finanssipalvelut tarkoittavat kaikkea toimintaa mitä pankit, vakuutusyhtiöt, rahastoyhtiöt, sijoituspalveluyritykset ja muut rahoitusalan yritykset harjoittavat asiakkaiden kanssa. Kyse on siis pankki-, vakuutus- ja sijoituspalveluiden tarjoamisesta kansantaloudessa. Tästä toiminnasta muodostuvat finanssijärjestelmä ja finanssimarkkinat. Tekninen kehitys muokkaa finanssimarkkinoiden luonnetta, mikä tarkoittaa sitä, että tietojenkäsittelyyn sekä asiakkaan ja yrityksen väliseen kommunikaatioon liittyvä kustannuksien nopea vähentyminen on lisännyt uusien palveluiden käyttöönottoa ja tehostanut toiminnan kasvua. Suomessa finanssimarkkinoita on tehostanut ja monipuolistanut merkittävästi tekninen kehitys, mikä on otettu laajassa mittakaavassa käyttöön. Finanssisektori on yhteiskunnan elimellinen osa, eikä se elä erillään muusta yhteiskunnasta, joten finanssialan yritysten toiminnassa näkyvät monipuolisesti ne arvot, joita esiintyy suomalaisessa yhteiskunnassa muutoinkin. Näin ollen yhteiskunnan arvojen muuttuessa muuttuvat myös finanssialan yritykset. (Alhonsuo, Nisén ym. 2012, 13-25)

Finanssiteknologia eli FinTech on tällä hetkellä yksi maailman nopeimmin kasvavista toimialoista. Finanssimarkkinoille on muodostunut nopeasti useita finanssipalveluita tarjoavia yrityksiä, kun aikaisemmin finanssipalvelut ovat olleet lähes yksinoikeudella pankeilla. Pankit kuitenkin edelleen hallitsevat finanssialaa tarjottavissa finanssipalveluissa, kuten palveluita, jotka liittyvät talletuksiin, maksupalveluihin ja luottoihin, mutta markkinoille on myös tullut lukuisia uusia toimijoita. Esimerkiksi lainaa on saatavilla pankkien lisäksi muilta luotonantajilta, kuten verkkopalveluista, missä yritys tai yksilö kohtaa luotonantajat. FinTech yritykset voivat

tarjota sellaisia palveluita, joita aikaisemmin ei ole ollut tarjolla, kuten luoda uusia tapoja esimerkiksi lainaamiseen, maksamiseen, rahan lähettämiseen sekä investoimiseen. Pankkien ja FinTech yritysten yhteistyö voi kuitenkin luoda merkittäviä uusia palveluita markkinoille, mistä johtuen pankit ja FinTech yritykset tekevät nykypäivänä usein yhteistyötä. Pankeilla on usein hallussaan paljon tietoa asiakkaistaan, kun taas FinTech yrityksillä on teknologia hallussa. Yhteistyön mahdollistaa API-ohjelmointirajapinta, josta on tarkempaa tietoa seuraavassa luvussa. (Finanssiala Ry, FinTech)

Teknologisten innovaatioiden käyttöönoton arvioidaan muuttavan finanssimarkkinasektoria merkittävästi tulevaisuudessa. Finanssiteknologia mahdollistaa entistä tehokkaammin uusia rahoituksen välittämismuotoja, kuten esimerkiksi joukkorahoituksen. Näihin liittyen uskotaan muodostuvan yhä enemmän uutta yritystoimintaa tulevaisuudessa, arvioi valtionvarainministeriö (Finanssiteknologia mullistaa rahoitusmarkkinat 2016). Teknologian myötä uudenlaisia verkkopalveluita otetaan käyttöön ja yritysten toiminta sekä työkalut kehittyvät entisestään, jolloin perinteisten palveluiden rinnalle on kehitettävä uusien teknologiayritysten tarpeisiin sopivia ratkaisuja. Datan kerääminen ja hyödyntäminen on helpottunut erilaisten palveluiden siirryttyä verkkoon, minkä ansiosta asiakkaiden käyttäytymisestä ja tarpeista on mahdollista saada tarkempaa tietoa. (Finanssialan tulevaisuuden näkymät)

3.1 API-Ohjelmointirajapinnat

API on teknologia ja digitaalinen tuote, jota käytetään tekoälyn hyödyntämiseen sekä alustatalouteen. API-Ohjelmistorajapinta (Application programming interface) mahdollistaa tiedon jakamisen organisaation sisäisesti tai vaihtoehtoisesti eri organisaatioiden välillä, jolloin useampi ohjelma keskustelee keskenään. Rajapinta voi olla datarajapinta, jota kautta jonkin tietyn palvelun sisältämä data siirretään toisiin järjestelmiin. Rajapinta voi myös olla toiminnallinen rajapinta, joka lisäksi tarjoaa laskenta-algoritmeja tai mahdollisuuden muuttaa rajapinnan kautta järjestelmän tietoja (Finanssiala Ry, Finanssialalle, API-ohjelmistorajapinta)

API-Ohjelmointirajapinnat voidaan luokitella eri osiin: sisäinen API, jota käytetään yleensä yritysten omien palveluiden kehittämiseen, Avoimen datan API, jossa tietosisältö on avointa ja ilmasta ja lisensoitu avoimen datan lisenssillä sekä Avoin API, mikä sisältää sekä kumppanuus API:n että julkisen API:n. (Niinioja 2018, 57-99)

Usein yritysten API-talouteen liittyminen tapahtuu kumppani-API:n hyödyntämisellä eli avaamalla toiselle yritykselle pääsy keskeisiin prosesseihin. Kumppanirajapintoja voidaan hyödyntää kahdella eri strategialla: hyödynnetään kumppanin rajapintoja tai tuotetaan kumppanille suunnattu API tai useampi. Kumppani-API:t perustuvat sopimuksiin, jotka ovat

yleensä tuttuja strategisten kumppanien kanssa tehtyjä avauksia. Liiketoiminnassa APIen kautta tapahtuva automatisoitu tietojenkäsittely luo yritysten välille keskinäisriippuvuutta, mikä vahvistaa niiden välistä sidettä. APIen avulla kumppaniverkoston skaalautuvuus paranee, eikä lukitse kumppania esimerkiksi yhden ohjelmistokielen käyttöön, vaan omassa järjestelmässä voidaan käyttää yrityksen haluamaa toteutuskieltä. Yhteistyö APIen avulla ei siis ole tiivis integraatio vaan jättää liikkumavaraa sen suhteen, miten data esitetään ja prosessoidaan sekä mihin dataa käytetään. Sopimuksissa voidaan myös määrittää kumppanin velvollisuuksia sekä rajoittaa kumppanin oikeuksia datan käytön suhteen (Moilanen 2018, 83-92). Julkinen API, joka kumppani APIin lisäksi luokitellaan avointen APIen luokkaan mahdollistaa palveluiden ja sovellusten viestiä saumattomasti keskenään. Se mahdollistaa digitaalisen sisällön laajasti saavutettavaksi eri palveluissa. Julkisen APIen avulla organisaation on mahdollista saavuttaa laajempaa tietoisuutta omille tuotteilleen laajentamalla osaksi useita sovelluksia ja ekosysteemejä. (Honkanen 2018, 92-94)

3.2 PSD2 - Toinen maksupalveludirektiivi

PSD2 (Payment Service Directive 2) eli toinen maksupalveludirektiivi on Euroopan parlamentin hyväksymä maksupalveludirektiivi, jonka tavoitteena on saattaa entistä laajemmin erilaiset maksupalvelut sääntelyn piiriin sekä vastaamaan sääntely markkinoilla tapahtunutta kehitystä. Direktiivi on asetettu voimaan kahdessa osassa, joiden lakimuutokset ovat astuneet pääasiassa voimaan 13.1.2018. Ensimmäisessä maksupalvelua muutettiin lailla 898/2017 ja toisessa maksulaitoslakia muutettiin 890/2017. (Finanssivalvonta, PSD2, 2019)

Kolmannet palveluntarjoajat (Third Party Payment Service Provider, TPP) on tuotu sääntelyn ja valvonnan piiriin, mikä on yksi keskeisimmistä muutoksista maksupalvelulainsäädännössä. Maksutoimeksiantopalvelun tarjoajat (Payment Initiation Service Providers, PISP) sekä tilitietopalvelun tarjoajat (Account Information Service Providers, AISP) ovat uusia maksupalveluiden tarjoajia. Pankkien, joilla on tilit hallussaan, on mahdollistettava näille palveluntarjoajille pääsy asiakkaiden tileille, mikäli asiakas on siihen antanut suostumuksen. Muita keskeisiä muutoksia ovat vaatimus asiakkaan vahvasta tunnistautumisesta internetmaksamisessa ja tilin online-käytössä sekä korttipohjaisten maksuvälineiden liikkeeseenlaskijat (Third Party Payment Instrument Issuer). PSD2 täten edellyttää tilinpitäjäpankilta standardoitua, julkaistua rajapintaa palveluja varten (kuviot 7). (Finanssivalvonta, PSD2, 2019 & Finanssivalvonta, FinTech ja sääntely, 2019)

PSD2 mahdollistaa esimerkiksi sen, että pankeissa olevien tilien tiedot voivat olla yhdellä kertaa katsottavissa. Mikäli asiakas antaa palvelun tarjoajalle luvan käyttää tilitietoja, voidaan tilitietojen perusteella tarjota sijoitus- tai lainapalveluita (Kysymyksiä ja vastauksia toisesta maksupalveludirektiivistä 2019). API teknologian ja PSD2-direktiivin yhteistyö tulee vaikuttamaan pankkien ja teknologiayritysten väliseen toimintaan merkittävästi. PSD2-

direktiivin myötä pankkien tulee integroida API ohjelmistorajapintoja toimintaansa, jotta ne voivat säilyttää asemansa markkinoilla (Platform Banking). Platform bankingista on tulossa megatrendi finanssialalla, jossa API teknologiaa hyödyntämällä rakennetaan ekosysteemiä organisaatioiden välillä. (Finanssiala Ry, Finanssialalle, API-ohjelmistorajapinta)



Kuvio 7: API-Ohjelmistorajapinta, Finanssiala Ry

Käytännössä PSD2 direktiivi tarkoittaa räjähdysmäistä kasvua sovellusten ja uusien palveluiden käyttöönotossa. Direktiivin myötä pankeille on tulossa lisää uusia kilpailijoita, mutta samalla se myös mahdollistaa uudenlaiset yhteistyöt pankkien ja fintech yritysten välille. Lakiuudistuksen myötä pankit ovat rakentaneet rajapintoja omista järjestelmistä datan käyttöä varten yhteistyön mahdollistamiseksi. Lakimuutos nopeuttaa ja tehostaa markkinoita, jolloin innovaatioita syntyy enemmän ja niiden myötä myös uudenlaisia palveluita. (Turunen 2017)

Useat fintech-yritykset ovat jo rakentaneet alustan, josta on pääsy pankkien tietoihin. Helsinki Fintech farmin mukaan vielä ei kuitenkaan ole selkeytynyt, kuinka eri pankit tulevat käyttämään partneroitumisstrategioita tai nähty, miten Open Banking- ansaintalogiikat rakentuvat. Näiden kuitenkin uskotaan selkeytyvän tämän vuoden aikana. Avoimet rajapinnat eivät koske pelkästään pankkeja, vaan myös muut Open Data -hankkeet tulevat kehittymään, kuten esimerkiksi julkisen sektorin kanssa tehdyt alustat vaikuttavat finanssialan dynamiikkaan (Viisi Digital finance - ja fintech trendiä Suomessa 2019). Pankit ovat huolissaan PSD2 direktiivin myötä kilpailun vääristymisestä, sillä FinTech -yritykset pystyvät nyt tuotekehitysideoiden varassa kääntymään pääomasijoitusyritysten puoleen. Turvatakseen olemassa olonsa pankkien on kehitettävä aktiivisesti uusia liiketoimintamalleja rajapintojen avaamisen vuoksi. Muuten

niillä voi olla vaarana kutistua persoonattomiksi ja mauttomiksi transaktioiden varastoijiksi. (Niinioja 2018, 24)

4 Luottokelpoisuuden arviointi

Yritysten luottokelpoisuus terminä on yleinen, mutta sille ei ole olemassa yksiselitteistä määritelmää. Yleisesti luottokelpoisuuden arvioinnilla pyritään selvittämään, kuinka mahdollinen asiakas pystyy suoriutumaan velvoitteistaan tietyn ajan kuluessa. Se, kuinka eri organisaatiot selvittävät luottokelpoisuuden vaihtelee oleellisesti. Yleisesti ottaen tarkemmissa luottokelpoisuuden selvittämisprosesseissa selvitetään yrityksen maksutapaa, vastuuhenkilöiden taustoja ja taloutta kuvaavia tietoja. (Ijäs 2002, 59)

Yritysrahoituspäätöksiä tehdessä arvioidaan yrityksen maksukykyä ja riskejä, joihin voidaan käyttää yritystutkimusmenetelmiä. Hinnoitteluun vaikuttaa yrityksen aikaisempi toiminta, kassavirta ja tulos sekä tulevaisuuden näkymät ja tarjolla olevat vakuudet, jotka luoton antajan tulee arvioida. Tärkein lähtökohta on asiakasyrityksen tunteminen, jolloin luoton antajan on tunnettava yrityksen toiminta ja taloudellinen tilanne hyvin, jotta pystytään arvioimaan tarvittavia tekijöitä luottopäätöksen pohjaksi ja riskien hallitsemiseksi. Lisäksi luoton antajan on pystyttävä arvioimaan asiakasyrityksen tulevaisuuden kehitystä. Yritykseen tutustuminen edellyttää tiedon hankkimista esimerkiksi yrityksen liikeideasta, toimialasta ja sen markkinoista, toimintaympäristöstä, tuotannosta, markkinoinnista, tutkimus- ja tuotekehityksestä, laskentatoimesta, tilinpäätöksestä ja budjetoinnista, asiakaskunnasta sekä johdosta. Yritystutkimuksissa pyritään selvittämään täyttyvätkö rahoituspäätöksille asettamat edellytykset. Yritystutkimus koostuu kertaluonteisista selvityksistä ja tilinpäätösanalyseistä. Yritystutkimuksissa käytetään apuna tietotekniikkaa esimerkiksi ennustemallien tekemiseen. (Alhonsuo, Nisén ym. 2012, 250-251)

Luottokelpoisuuden selvittäminen vaatii usean palasen yhteen kokoamista, jotta oikeanlainen päätelmä saadaan aikaiseksi. Joskus kuitenkin jo yksittäinen pala esimerkiksi selvitetystä tiedosta, jossa on tullut ilmi yrityksellä olevan useita maksuhäiriömerkintöjä, riittävät vakuuttamaan, että mahdollinen luotonanto on suuri riski. Toisaalta havainto siitä, että asiakas on noudattanut eräpäiviä eli positiivinen maksutapa ei pelkästään riitä vakuuttamaan, että asiakas on luottokelpoinen. Maksuviiveiden ohella yrityksen epäluotettavuutta luottoasiakkaana vahvistavat esimerkiksi alhainen maksuvalmius (quick ratio < 0,5 tai current ratio < 0,8), vastuuhenkilöiden aiemmat konkurssiyhteydet tai henkilökohtaiset maksuhäiriöt, yritykseen kohdistuneiden luottotietokyselyiden olennainen lisääntyminen lyhyellä aikavälillä,

yrittäjän poistaminen verottajan ennakoperintärekisteristä tai havainto epätavallisen täynnä tavaraa olevasta varastosta. (Ijäs 2002, 59-60)

Yrittäjän olemassa olevista asiakkaista tärkeää tietoa löytyy yrityksen omista järjestelmistä, sillä asiakkaan maksukäyttäytyminen on arvokasta tietoa luottoriskien hallinnan kannalta. On kuitenkin hyvä pitää mielessä, että yritykset priorisoivat velkojat ja aloittavat maksamisen niistä, joiden tyytyväisyys on tärkeää oman toiminnan jatkuvuuden kannalta, joten maksukyvyyn heikkeneminen saattaa paljastua vasta myöhemmin, mikäli maksukäyttäytymistä seurataan vain omasta järjestelmästä. Myös asiakasyrittäjän varasto voi kertoa ongelmista. Yrittäjissä pyritään pitämään varaston koko optimaalisena ja varastointiajat lyhyinä, sillä se sitoo merkittävän määrän pääomaa. Yrittäjän varaston ollessa liian täynnä, on se usein merkki huonosta menekistä. (Ijäs 2002, 62-66)

Yrittäjän taloudellinen analyysi perustuu yrityksen virallisiin tilinpäätöksiin, jossa keskeisimmät osat ovat tuloslaskelma ja tase liitetietoineen. Tuloslaskelmasta voidaan selvittää tilikauden jakokelpoinen voitto ja taseesta yrityksen omaisuus ja tapa, jolla yritystoiminta on rahoitettu. Yrittäjän taloudellista menestymistä voidaan tutkia erilaisten tilinpäätösanalyysitekniikoiden avulla. Tällaisia tekniikat voidaan luokitella prosenttilukumuotoiseen tilinpäätösten laadintaan, trendianalyysiin ja tunnuslukuanalyysiin. Prosenttilukumuotoinen tilinpäätös tarkoittaa nimensä mukaisesti lukujen esittämistä prosentteina liikevaihdosta (tuloslaskelma) ja taseen loppusummasta (taseen erät). Tällä tekniikalla tilinpäätösosien keskinäiset suhteet saman tilikauden sisällä saadaan selkeästi esiin. Trendianalyysissä prosenttilukumuotoiset tilinpäätökset laaditaan useilta peräkkäisiltä vuosilta vertailua varten ja tarkasteluperiodin ensimmäistä tilikautta pidetään perusvuotena. Tunnuslukuanalyysi perustuu tuloslaskelman ja taseen eristä koottuihin laskuihin, jotka ovat usein suhdelukumuotoisia tunnuslukuja. Ne mittaavat yrityksen menestymisestä eri osatekijöitä, kuten kannattavuutta, rahoitusrakennetta ja maksuvalmiutta. Analyysin vertailukohteiden perusteella tilinpäätösanalyysi voidaan jaotella tiettyä tilikautta koskevaan ajalliseen poikkileikkausanalyysiin ja peräkkäisten tilikausien aikasarja-analyysiin. Ajallisessa poikkileikkausanalyysissä yrityksen lukuja verrataan tietyn tilikauden ajalta muihin, usein saman toimialan yrityksiin. Aikasarjatarkastelussa yrityksen kehitystä seurataan sen omien lukujen perusteella pitkällä aikavälillä, jolloin johtopäätöksiä yrityksen taloudellisen tilan kehityssuunnasta voidaan tehdä. Nämä kaksi analyysityyppiä voidaan myös yhdistää, jolloin voidaan tehdä vertailuja muihin yrityksiin pitkällä aikavälillä. (Niskanen J & Niskanen M, 2010, 49-72)

Yrittäjän tilinpäätös-, maksutapa- ja maksuhäiriötietojen sekä vastuuhenkilöiden yrityskytkeiden ja henkilökohtaisten luottotietojen selvittäminen antaa luottokelpoisuuden selvittämiseksi hyvän pohjan, sillä tiedot avaavat yrityksen taloudellista tilaa, aikaisempaa

maksukäyttäytymistä, vastuuhenkilöiden henkilökohtaista asioiden hoitokykyä sekä menestymistä muussa liiketoiminnassa. (Ijäs 2002, 62-66)

Monissa yrityksissä luottokelpoisuuden arvioinnissa käytetään tukena luottotietoyhtiöitä. Luottotietoyhtiöistä on saatavilla yhdestä lähteestä laajaa informaatiota asiakkaista, sillä luottotietoyhtiöiden data on kerätty useammista eri lähteistä, joita ovat esimerkiksi kaupparekisteri, tilastokeskus, virallinen lehti, oikeusrekisterikeskus, kihlakunnan ulosotto-osastot, käräjäoikeudet, verovirastot sekä rahoitus- ja luottokorttiyhtiöt. Siitä huolimatta luottopäätöksen tekemiseen ei välttämättä kuitenkaan yksistään riitä luottoyhtiöiden tietokannoista löytyvä data. (Ijäs 2002, 66)

4.1 Tilinpäätöstiedot yritysrahoituksessa

Yrityksen kirjanpito kuvaa yrityksen rahaprosessia, minkä vuoksi sillä on vahva yhteys yritysrahoitukseen. Tilikauden aikana tapahtuneiden juoksevien kirjausten pohjalta laaditaan tilinpäätös, jossa tuloslaskelma ja tase liitetietoineen ovat keskeisiä osia. Tuloslaskelma antaa informaatiota yrityksen taloudellisesta suorituskyvystä tilikauden aikana ja siitä voidaan selvittää jakokelpoinen voitto tilikaudelta. Taseessa vastaavaa puolelta (debt) nähdään yrityksen omaisuus ja vastattavaa puolelta (credit) nähdään, miten yritystoiminta on rahoitettu. Tilinpäätökseen katsotaan lisäksi kuuluvan toimintakertomus, josta selviää oleelliset tiedot kirjanpitovelvollisen toiminnan kehittymisestä. Kirjanpitolaki ja kirjanpitoasetus säätelevät tilinpäätöksen laadintaa Suomessa. Tuloslaskelman ja taseen sisältö ja muoto on määrätty kirjanpitoasetuksessa. (Niskanen J & Niskanen M 2010, 45)

Tilinpäätöstiedoilla voidaan selvittää yrityksen taloudesta kolme avaintekijää; kannattavuus, maksuvalmius ja vakavaraisuus. Tilinpäätöstiedot ovat julkaisuhetkellä usein 7-8 kuukautta vanhoja, joten tilinpäätöstietoja moititaan usein siitä, että ne kertovat vain menneestä. Varman kuvan tilinpäätöstiedot kuitenkin antavat siitä, millaisesta tilanteesta yritys on lähtenyt kuluvalle tilikaudelle. Jos viime tilikaudenlopussa yrityksen taloudellinen tilanne oli heikko, tulisi selvittää, mitä kuluneen 7-8 kuukauden aikana on tapahtunut. Vahvistusta kuluneelle 7-8 kuukaudelle voidaan saada muun muassa viimeaikaisesta maksukäyttäytymisestä, maksutavasta, miten asiakkaan toimialalla menee yleisesti, mitkä ovat toimialan tulevaisuuden näkymät sekä selvittämällä, miten yrityksen edustajat kommentoivat viimeaikaista kehitystä. Toisaalta myöskään viime tilikauden vahva tulos ei yksinään kerro yrityksen hyvästä tulevaisuudesta. Etenkin pienemmissä yrityksissä tilanne saattaa muuttua nopeasti, ja yksikin huomattava ja vääräksi osoittautuva investointipäätös tai huomattava luottotappio voi kääntää kehityksen suunnan. (Ijäs 2002, 81)

Lainsäädännön mukaan kaikkien osakeyhtiöiden tulee toimittaa tilinpäätökset kaupparekisteriin kahden kuukauden sisällä tuloslaskelman ja taseen vahvistamisesta (Ijäs

2002, 81). Keskinäisen kiinteistöosakeyhtiön ja asunto-osakeyhtiön ei kirjanpitolain mukaan tarvitse ilmoittaa tilinpäätöstä kaupparekisteriin (PRH, osakeyhtiön tilinpäätösilmoitus, 2019). Henkilöyhtiöiden on julkaistava tiedot kuuden kuukauden sisällä tilikauden päättymisestä vain, jos vähintään kaksi seuraavista raja-arvoista ylittyy: liikevaihto 6,24 miljoonaa euroa, tase 3,125 miljoonaa euroa tai palveluksessa keskimäärin 50 henkilöä. Näillä raja-arvoilla suurin osa henkilöyhtiöistä jää informaatiovelvollisuuden ulkopuolelle. (Ijäs 2002, 82)

4.2 Tunnusluvut

Yrityksen taloudellista suorituskykyä voidaan mitata tilinpäätöksestä lasketuilla tunnusluvuilla. Taloudellinen suorituskyky muodostuu useasta osatekijästä, ja usein se jaetaan kannattavuuteen, vakavaraisuuteen ja maksuvalmiuteen eli likviditeettiin. (Niskanen J & Niskanen M 2010, 55)

Tunnuslukuja voidaan hyödyntää yrityksen takaisinmaksukykyä arvioitaessa. Tunnuslukuja tulkittaessa tulee kiinnittää huomiota absoluuttisten arvojen lisäksi niiden kehitykseen ja tasoon suhteessa koko toimialan lukuihin, joten vertailu aiempiin tilikausiin on päätöksen tekemisen kannalta oleellista. Vertailu edellisiin tilikausiin kertoo tunnuslukujen kehityksestä, mutta niiden tasosta saa käsityksen vasta vertailemalla niitä saman toimialan yritysten keskimääriäisiin lukuihin. Arvot ovat helposti selvitettävissä luottotietoyhtiöiden tietokannoista. Yrityksen liiketoiminnan volyyymi myös vaikuttaa tunnuslukujen tulkintaan - pientä liiketoimintaa harjoittavan yrityksen tunnusluvut voivat näyttää harhaanjohtavan hyviltä, sillä esimerkiksi 35 prosentin liikevoitto voi tarkoittaa rahassa tuhatta euroa, mikä ei riitä vahvistamaan käsitystä tuottoisasta toiminnasta. (Ijäs 2002, 83-84)

Kannattavuuden, maksuvalmiuden ja vakavaraisuuden selvittämiseen käytetyt tunnusluvut kuvaavat hyvin yrityksen taloudellista asemaa ja tuloksentekeykyä, mutta myös niiden väliset suhteet ovat tärkeitä. Jos esimerkiksi omavaraisuus on alhainen, tarvitaan hyvä kannattavuus, jotta selvittää vieraan pääoman lyhennyksistä ja koroista. Yleisen katetason pudotessa ja hintatason voimistuessa yritys ajautuu helposti taloudellisiin vaikeuksiin. Pääsääntö on, että vähintään yhden osa-alueen tulisi olla hyvässä kunnossa, jotta yrityksellä on toimintaedellytyksiä. Tärkein näistä kolmesta on vakavaraisuus, sillä se antaa aikaa esimerkiksi kannattavuusongelmien korjaamiselle. (Ijäs 2002, 85)

4.3 Yrityksen kassavirta- ja rahoituslaskelma

Rahoituslaskelma kertoo, mistä yritykseen tullut raha on peräisin ja mihin rahat on käytetty. Rahoituslaskelmasta nähdään siis konkreettisesti kassaan tullut ja kassasta poistunut raha. Siitä voidaan selvittää yrityksen rahavirrat, jotka ovat peräisin liiketoiminnasta, investoinneista sekä

rahoituksesta. Kassaperusteiset luvut kertovat tuloslaskelmaa tarkemmin yrityksen toimintaedellytyksistä. Rahavirrat, jotka ovat kertyneet liiketoiminnasta kertovat, kuinka yritys on selviytynyt myyntituloillaan juoksevista kuluista ja muista välittömistä menoista. Investointien rahavirroiksi luokitellaan esimerkiksi aineellisen ja aineettoman käyttöomaisuuden hankinnat sekä myynnit, joita voivat olla esimerkiksi yritysostot ja laitehankinnat. Yrityksien tulorahoitus ei kuitenkaan aina kata kaikkia investointeja, jolloin yhtiö päätyy hakemaan lisärahoitusta. Vieraan ja oman pääoman muutokset puolestaan voidaan selvittää rahoituksen rahavirroista, joista nähdään millä yritys on rahoittanut toimintansa ja kuinka paljon omistajille on maksettu osinkoja. Rahavirrat liiketoiminnassa kertovat yrityksen tehokkuudesta, kasvumahdollisuuksista sekä maksuvalmiudesta. Likvidien varojen merkitys korostuu, mikäli talouskasvu on hidasta. Rahavirrat vaihtelevat yrityksen eri tilikausilla sekä koon että toimialan suhteen, joten rahavirtoja tulee verrata eri tilikausiin sekä muihin samalla toimialalla toimiviin yrityksiin. (Felt 2008)

Yrityksen kassavirrat eivät ole suoraan luettavissa tilinpäätöksestä, sillä tilinpäätös laaditaan tulojen osalta suoriteperusteella ja menojen osalta suorite- ja käyttöperusteella. Kassavirta käsitteenä on yritysrahoituksessa tärkeä, koska kaikkien sijoituskohteiden arvonmääritys perustuu kassavirtoihin eikä tilinpäätösvoittoiin. Kassavirran ja tuloslaskelman ero on luonteeltaan ajoitusero, sillä laskuista tulevat suoritukset eivät yleensä kerry yrityksen kassaan tilikauden aikana vaan osa siirtyy maksettavaksi tilikauden jälkeiselle ajalle. Suoritusperusteinen liikevaihto ja liikevaihdon kassaanmaksut eroavat toisistaan, sillä laskut lasketaan mukaan liikevaihtoon, vaikka suoritusta ei olisi saatu tilikauden aikana. (Niskanen J & Niskanen M 2010, 66-67)

4.4 Vakuudet

Vakuuksien tarkoituksena on varmistaa luoton antajalle luoton takaisinmaksu. Se ei kuitenkaan ole yksinään taee luoton antamiselle, mutta se on yksi tekijä, jota luoton antaja arvioi maksukykyä ja riskiä harkitessa. Luoton antajan vastuulla on myös vakuuden arviointi. Vakuuksia on kahdenlaisia: reaalivakuudet ja henkilövakuudet. (Alhonsuo, Nisén ym. 2012, 250-251)

Reaalivakuus tarkoittaa sitä, että pantin haltijalla eli luoton antajalla, joka on myöntänyt luoton vakuutta vastaan, on oikeus saada saatavaansa vastaa summa vakuuden arvosta. Yleispanttauksessa pantattu omaisuus on vakuutena luoton saajan kaikista voimassaolevista tai myöhemmin syntyvistä vastuista ja veloista, jotka kyseinen luoton antaja on myöntänyt. Erityispanttauksessa puolestaan vakuus kohdistuu erikseen panttaussopimuksessa mainittuun

velkaan tai vastuuseen. Yleisimpiä reaalivakuuksia ovat asunto-osakkeet ja kiinteistöt, arvopaperit, obligaatiot ja talletustodistukset. (Alhonsuo, Nisén ym. 2012, 250-251)

Henkilövakuudessa eli takauksessa henkilö sitoutuu vastaamaan päävelallisen velasta. Tällöin takaaja on vastuussa velan maksamisesta koko omaisuudellaan. Sitoumus voidaan kuitenkin rajoittaa määrällisesti, jolloin takaaja on vastuussa velasta kyseisen määrän osalta. Takaus voi olla laillinen tai omavelkainen. Laillisessa takauksessa takaajan vastuu alkaa vasta siinä vaiheessa, kun päävelallinen todetaan maksukyvyttömäksi ulosmittauksessa. Omavelkaisessa takauksessa velan myöntänyt voi vaatia erääntyntä saatavaa myös ensisijaisesti takaajalta maksettavaksi. Yhdessä lainassa voi myös olla useampia takaajia, jolloin kukin vastaa yhteisvastuullisesti koko velasta ja koko erääntyntä saatavaa voidaan vaatia maksettavaksi keneltä tahansa takaajista. (Alhonsuo, Nisén ym. 2012, 250-251)

4.5 Luottokelpoisuuden arviointi toimeksiantajayrityksessä

Tässä kappaleessa oleva tieto on selvitetty haastattelemalla toimeksiantajayrityksen luottoanalyttikkoja, jotta on saatu tarkempi kuva lähtötilanteesta ja mitä asioita tulee selvittää luottokelpoisuuden arviointiprosessissa. Tiedon perusteella voidaan kartoittaa paremmin tilanteita, joihin tekoälyä voidaan ja kannattaa hyödyntää.

Lainahakemuksissa asiakkaalta pyydettäviin pakollisiin materiaaleihin kuuluvat tilinpäätös, tilitapahtumat maksuliikennetililtä viimeisimmältä kolmelta kuukaudelta, hallituksen päätös lainan hakemisesta sekä lainarahan käyttötarkoitus ja siihen liittyvät materiaalit. Koska tilinpäätöstiedot voivat olla useita kuukausia vanhoja, tarvitaan päätöksen tueksi myös kirjanpitoajot viimeisimmästä vahvistetusta tilinpäätöksestä tähän hetkeen saakka, josta selviää yrityksen tuoreimmat luvut ja kassavirrat.

Liiketoimintaa ja yrityksen tilannetta arvioidaan yrittäjän kertomuksen, tilinpäätösanalyysin ja tilitapahtumien kautta. Näiden tekijöiden tulee myös tukea toisiaan. Tilinpäätöksestä tarkastellaan yrityksen vakavaraisuutta, kannattavuutta, maksuvalmiutta, yrityksen kokoa ja rakennetta. Tilinpäätöslukuja tarkastellaan pääsääntöisesti aina neljän viimeksi vahvistetun tilikauden ajalta. Tilinpäätöksiä vertailemalla pystytään seuraamaan kehityksen trendiä. Tosin myös esimerkiksi Asiakastiedolta on saatavilla graafeja, josta yritysten tilinpäätöksiä kehitystä pystytään seuraamaan. Myös toimialalla on suuri merkitys tilinpäätöksiä analysoitaessa, koska esimerkiksi kannattavuusluvut ovat eri suuruisia eri toimialoilla. Toiset toimialat ovat myös pääomaintensiivisempiä kuin toiset. Lainaa hakevan yrityksen lukuja verrataan samalla toimialalla toimivien verrokkiyhtiöiden lukuihin. Tilitapahtumat maksuliikennetililtä ovat myös tärkeä osa kokonaisuutta, sillä niistä voidaan tutkia yrityksen reaalisia rahavirtoja. Tilitapahtumista pyritään havainnoimaan, keneltä yhtiö saa suorituksia ja kenelle yhtiö maksaa ja vastaavtko nämä volyymeiltaan annettuja tietoja. Asiakkaan toimittamissa erillisissä tiliote-tiedoista on

hankala todentaa helposti kokonaiskuvaa yhtiön viime kuukausien volyymeista sekä tilitapahtumien osapuolista. Tiedostomuotoja ja esitystapoja on useita. Viitesiirtoja ei ole välttämättä purettu auki yksittäisiksi tapahtumiksi maksa- tai vastaanottaja tietoineen. Tämä hankaloittaa automaatiota. Dokumentteja on myös kohtalaisen helppo väärentää tai muokata ja väärennöstä on hankala huomata. PSD2 tulee kuitenkin helpottamaan automaatiota, koska tiedostot ovat silloin aina samassa muodossa ja tiedot saadaan suoraan pankista, jolloin väärentäminen ja tietojen pois jääminen ei ole mahdollista.

Lainarahan käyttötarkoitukseen liittyvät materiaalit riippuvat paljon käyttötarkoituksesta. Esimerkiksi jos käyttötarkoituksena on rakennus, tarvitaan tuohon liittyvät vuokrasopimukset ja muut rakennukseen liittyvät sopimukset, jotka analyytikon on luettava ja löydettävä esimerkiksi kohtia, joista vielä tulee pyytää lisämateriaalia. Lisäksi yrittäjää haastatellaan avoimin kysymyksiin, josta saadaan yrittäjän kertomus tueksi päätöksentekoon. Yrityksen takaisinmaksukykyä arvioitaessa selvitetään yrityksen kassavirrat, jotka ovat oleellisin asia, sillä myös tilinpäätös voidaan oikaista kassavirran osalta. Lisäksi selvitetään, millaista tulosta yhtiö on tehnyt aikaisempina toimintavuosina ja kuinka realistinen investointisuunnitelma on. Investointisuunnitelma sisältää tyypillisesti kuvauksen toteutettavasta hankkeesta, sen vaikutuksesta yhtiön toimintaan sekä projektin kustannukset ja kannattavuuslaskelman. Investoinnin takaisinmaksuajan tulee olla järkevä ja vastata hankittavan kohteen taloudellista käyttöikä. Esimerkiksi kone- tai laiteinvestoinnin tulee tuottaa vapaata kassavirtaa vähintään sen verran, että sen hankintaan tarvittava laina korkoineen saadaan kuoletettua koneen käyttöaikana. Rakentamiseen liittyvissä hankkeissa täytyy arvioida rakentamiseen liittyvien operatiivisten riskien ja rakennuskustannusten lisäksi mahdolliseen myyntiin liittyvät riskit. Hankkeen takaisinmaksun kannalta on oleellista, että kohde saadaan myytyä kustannuksia korkeempaan hintaan, vaikka kustannukset ylityisivät ja hinta ei olisikaan tavoitehinnan mukainen. Jos rakennus rakennetaan omaan käyttöön, huomiodaan takaisinmaksukyvyssä vuokratkustannusten poisjänti. Investointien kannattavuutta, realistisuutta ja vaikutusta arvioidessa, tarkastetaan myös investointiin liittyvät sopimukset. Näiden palasien yhteen kokoamisen jälkeen voidaan arvioida investoinnin vaikutusta liiketoimintaan, kannattavuuteen ja vakavaraisuuteen.

Takaisinmaksukykyyn vaikuttaa myös yrityksen asiakaskunta, jolloin selvitetään, onko yritys riippuvainen esimerkiksi vain yhdestä tai kahdesta asiakkaastaan tai onko uusia sopimuksia tehty lähiaikoina. Joissakin tapauksissa yksi asiakassopimus saattaa kattaa ison osan liikevaihdosta, jolloin sillä on iso merkitys ja tällöin on syytä selvittää myös kyseisen asiakkaan taloudellinen asema. Yrityksen takaisinmaksukykyä arvioitaessa tutkitaan myös yrityksen kulurakennetta. Esimerkiksi liiketoiminta voi olla kannattavaa, mutta rahoituskulut ja pääomarakenne huonoja tai laskujen kiertonopeus hidasta. Yrityksen kulurakennetta saattaa keventää esimerkiksi automaatio. Yrityksen luotettavuutta ja uskottavuutta arvioidessa tarkistetaan myös yrityksen maksutapakäyttäytyminen, yrityksen toiminnan laajuus ja ikä,

yrittäjien maksuhäiriöt ja muut yritysytteudet sekä omistajien taloudellinen tilanne ja sitoutuminen yhtiöön, pystyykö tarvittaessa pääomittamaan yhtiötä.

5 Tutkimusmenetelmät ja toteutus

Tämä opinnäytetyö on toteutettu laadullisena tutkimuksena, jossa käytetään tutkimusmenetelmänä teemahaastatteluja eli puolistrukturoituja asiantuntijahaastatteluja sekä benchmarking-menetelmää. Lisäksi tutkimuksen kannalta oleellista tietoa on etsitty erilaisista sähköisistä lähteistä tai webinaareista, joissa kirjoittaja tai puhuja on aiheen asiantuntija. Tässä opinnäytetyössä tärkeimpinä tutkimuskysymyksinä on tekoälyn tuomat mahdollisuudet ja tekoälyn tuottama lisäarvo toimeksiantajayritykselle sekä miten muut yritykset hyödyntävät tekoälyä luottoprosesseissa tai yrityksiä arvioinnissa. Tyypillisiä piirteitä laadulliselle tutkimukselle ovat tiedonhankinta todellisista tilanteista sekä se, että saadun tiedon lähteenä on ollut ihminen, minkä perusteella voidaan määrittää aineistosta merkitykselliset asiat (Hirsjärvi, Remes, Sajavaara 2007). Haastatteluissa kysymykset ovat suunniteltu etukäteen, mutta myös avoimemmalle keskustelulle on jätetty tilaa, jotta haastatteluista on saatu oleellista tietoa myös asioista, joita ei alun perin ole huomioitu teemahaastattelun haastattelurungossa.

Kummassakin menetelmässä aineiston keruu toteutettiin puolistrukturoituina teemahaastatteluina. Yksityiskohtaisten kysymyksiä sijaan haastattelut etenevät tiettyjen aiheiden kannalta keskeisten teemojen varassa. Teemahaastattelut sopivat tilanteeseen, jossa haastattelija ei ohjaa liikaa vastaajia ja keskustelu on vapaampaa, vaikka kysymykset ovat pääsääntöisesti ennalta suunniteltuja. Teemahaastattelu ottaa huomioon sen, että haastateltavien tulkinnat asioista ja antamien tulkintojen merkitykset ovat keskeisiä, kuten myös se, että merkitykset syntyvät vuorovaikutuksessa. Teemahaastattelu on puolistrukturoituhaastattelu siksi, että haastattelun aihepiirit ja teema-alueet ovat kaikille samat, mutta kysymykset eivät ole tarkkamuotoisia lomakekysymyksiä eikä myöskään täysin vapaita niin kuin syvähaastattelut. (Hirsjärvi & Hurme, 2009)

Asiantuntijahaastattelut on valittu menetelmäksi sillä perusteella, että vahvin ja tuorein tieto tekoälyn hyödyntämisessä on teknologia-alan ammattilaisilla, joilla on myös tietoa tulevaisuuden näkymistä. Asiantuntijat haastatteluihin ovat löytyneet toimeksiantajayrityksen edustajien kautta. Benchmarking-menetelmä on valittu menetelmäksi siitä syystä, että sen avulla voidaan tarkemmin vertailla erilaisia tekoälyn tuomia mahdollisuuksia nimenomaan opinnäytetyön toimeksiantajayrityksen tarpeiden kannalta sekä saada selkeitä käytännön esimerkkejä toimeksiantajayrityksen tietoon. Myös näiden yritysten edustajilla, jotka on valikoitu haastateltaviksi, on vahvaa tietoa finanssialalla vallitsevista teknologioista. Myös nämä haastateltavat ovat löytyneet ja suostuneet haastateltaviksi toimeksiantajayrityksen

edustajien kautta. Yksi haastateltavista soveltui sekä asiantuntijahaastatteluun että benchmarkingiin.

Tätä opinnäytetyötä tehdessä asiantuntijahaastatteluita sekä benchmarkingin yhteydessä tehtyjä haastatteluita tehtiin yhteensä 8 kappaletta. Haastatteluista on kertynyt runsaasti materiaalia, joista on kirjoitettu muistiinpanoja haastattelun yhteydessä tai haastateltavan luvalla äänitettyjä. Haastatteluiden jälkeen kerätty aineisto on purettu ja tarkasteltu mahdollisimman pian haastattelun jälkeen, jotta aineisto on ollut vielä tuore ja helpommin dokumentoitavissa. Aineisto on litteroitu haastatteluiden jälkeen, minkä pohjalta oleelliset asiat ovat koostettu tekstiksi tutkimustuloksiin tähän opinnäytetyöhön. Asiantuntijahaastatteluista on pyritty löytämään yhdenmukaisuuksia ja poikkeavuuksia vertailemalla niitä keskenään. Yhdenmukaisuudet ovat vahvistaneet kyseisten asioiden olevan todennäköisemmin juuri niin kuin useimmat haastateltavat ovat esittäneet. Benchmarkingmenettelyn yhteydessä tehtyjä haastatteluita on myös verrattu keskenään ja siltä pohjalta pyritty löytämään toimeksiantajayritykselle toimivia kehitysehdotuksia. Lisäksi asiantuntijahaastatteluita sekä benchmarking-menettelmän avulla saatuja tuloksia on pyritty vertailemaan keskenään johtopäätöksissä, sillä asiantuntijoilla voi olla tiedossa ratkaisuja, joita benchmarkingin menetelmässä käytetyillä yrityksillä ei ole käytössä. Tästä syytä myös molemmat menetelmät ovat tämän opinnäytetyön hyödyn kannalta tärkeitä.

5.1 Asiantuntijahaastattelut

Tutkimuksessa haastateltiin neljää tekoälyyn perehtynyttä asiantuntijaa, jotka työskentelevät teknologia-alalla sekä yhtä koneoppimiseen suuntautunutta tohtorikoulutettavaa. Haastatteluiden avulla pyrittiin selvittämään finanssialalla toimivien yritysten toimintatapoja luottoprosesseissa, joissa hyödynnetään tekoälyä. Haastateltavilla teknologiayritysten asiantuntijoilla on vahva näkemys siitä, kuinka finanssialalla hyödynnetään tekoälyä yleisesti luottoprosesseissa ja mitkä ovat tekoälyn tulevaisuuden näkymät luottoprosesseissa. Tästä johtuen juuri heitä suositeltiin haastateltaviksi. Lisäksi heillä, kuten myös tohtorikoulutettavalla on tietoa eri tekniikoiden hyödyntämisestä. Aineiston kerääminen toteutettiin teemahaastatteluina eli puolistrukturoituina haastatteluina, joissa kysymykset olivat ennalta suunniteltu (liite 1). Haastattelujen yhteydessä kysyttiin kuitenkin tarvittaessa lisäkysymyksiä tarkennukseksi. Seuraavan haastattelun kysymykset saattoivat hieman muuttua edellisen haastattelun tuotua lisää tarkennuksia ja ajatuksia kysymyksiin. Haastattelun

aiheeseen ja haastateltaviin yrityksiin oli perehdytty tarkasti ennen varsinaisia haastatteluja.

5.2 Benchmarking

Benchmarking on tutkimusmenetelmä, jonka avulla vertaillaan eri toimijoiden toimintatapoja ja selvitetään miten toiset toimivat ja menestyvät. Sen tarkoituksena on tunnistaa heikkouksia omassa toiminnassa ja luoda niille kehitysideoita ja tavoitteita. Benchmarking vaatii ensin oman kehitettävän kohteen identifioinnin, jonka jälkeen selvitetään toisten organisaatioiden toimintatapoja. Toimintatapoja voidaan selvittää tutustumiskäynneillä ja monenlaisilla tiedonhankinnoilla, esimerkiksi internetistä. Benchmarkingin jälkeen saatuja tuloksia tutkitaan kriittisesti ja sovelletaan oman organisaation toimintaan. Vertailukohteeksi voidaan ottaa esimerkiksi toisten toimialojen organisaatiot, kilpailijat tai toimialan tilastolliset keskiarvot ja standardit. Organisaatioiden kulttuurierot voivat kuitenkin olla suuria, joten kaikki vertailukohteen hyvät puolet eivät välttämättä ole suoraan sovellettavissa oman organisaation toimintaan. Havaituista toimista voidaan kuitenkin oppia uutta sekä soveltaa toimintaa omaan toimintaan. (Ojasalo K, Moilanen T, Ritalahti J, 2015, 43-186)

Tässä opinnäytetyössä tutkittiin benchmarking-menetelmää hyödyntäen, kuinka yritykset, jotka jollakin tavalla arvioivat yrityksiä, soveltavat tekoälyä omissa prosesseissaan. Lisäksi selvitetään mitä hyötyä tai mahdollisesti huonoja puolia tekoäly sovelluksista on ollut yritykselle. Myös tässä menetelmässä tieto kerättiin teemahaastatteluiden eli puolistrukturoitujen haastattelujen avulla lukuun ottamatta yhtä yritystä, josta löytyi riittävästi tietoa yrityksen nettisivuilta. Haastateltavina oli Suomessa toimivien yritysten edustajia, joilla on tiedettävästi käytössään erilaisia teknologioita, kuin toimeksiantajayrityksellä. Yrityksillä on usein käytössään tekoälysovelluksia, joiden hyötyjä, toimivuutta ja soveltamista pyrittiin haastattelujen avulla selvittämään. Toimeksiantajayrityksen tavoitteena on ottaa omaan käyttöön tekoälysovelluksia lähitulevaisuudessa, joten vertailu tekoälyn soveltamistavoista on ajankohtainen. Benchmarking on käyttökelpoinen menetelmä tässä opinnäytetyössä, sillä sen avulla pyritään kehittämään organisaation laatua, tuottavuutta, toimintaprosesseja ja työtapoja, joita myös tämän opinnäytetyön toimeksiantajayritys pyrkii liiketoiminnassaan kehittämään.

6 Asiantuntijahaastatteluiden tulokset

Tässä luvussa on yhdistetty asiantuntijahaastatteluiden, sähköisten lähteiden sekä webinaarien avulla saatua tietoa liittyen tekoälyn hyödyntämiseen luottoprosesseissa. Sähköisistä lähteistä

ja webinaareista kerätty tieto, joissa kirjoittaja tai esiintyjä on alan asiantuntija, on liitetty tämän kappaleen tuloksiin.

6.1 Tekoälyn hyödyntäminen luottoprosesseissa

Tärkeää on ymmärtää mikä on koneoppimisen mahdollisuus verrattuna perinteiseen ohjelmointiin. Olennaista tekoälyn hyödyntämisessä on se, että on olemassa historiadataa siitä, miten arvioita on aiemmin tehty, mitkä kriteerit ovat olleet ja mikä on ollut lopputulema. Jos tällaista historiadataa löytyy, on mahdollista rakentaa koneoppimismalli helposti etenkin, jos samoja kriteereitä käytetään edelleen, eikä merkittäviä muutoksia ole tapahtunut esimerkiksi markkinoilla. Käytännössä koneoppimismalli tässä tapauksessa toimii niin, että koneelle syötetään kriteerit ja sen jälkeen kone antaa jonkin arvion (Haastattelu 5). Syötteenä voidaan käyttää esimerkiksi tilinpäätöstietoja, kassavirtalaskelmaa, tasetta sekä esimerkiksi luottotietoyhtiöiltä tai perintätoimistoilta hankittua dataa. Syötteenä voidaan käyttää myös tilitapahtumatietoja, joissa näkyy korttipääteellä ja verkkopankissa käyty rahaliikenne. Eri lähteistä voidaan kerätä tämän tyyppistä dataa ja rakentaa sen avulla koneoppimisen malli, joka oppii saadun datan perusteella. (Haastattelu 1 ja 2)

Tyypillisesti finanssialalla hyödynnetään myös luonnollisen kielen käsittelyä (NLP), jos analyytikon tehtävään kuuluu dokumenttien tai erilaisten sopimuksien lukeminen, joiden avulla on tarkoitus saada mahdollisimman nopeasti selville dokumentin sisältö ja tehdä siitä yhteenveto ja analyysi. Kaikista digitaalisessa muodossa olevista dokumenteista pystytään irrottamaan tietoa luonnollisen kielen avulla ja se tulee arvioiden mukaan yleistymään koko ajan enemmän ja enemmän (Haastattelu 5 ja 4). Tästä käytännön esimerkki on investointipankille tehty ratkaisu, jossa analyytikkojen tehtävänä oli lukea satoja sivuja pitkiä riskidokumentteja yrityksistä ja alleviivata sieltä riskit ja tehdä yhteenveto. Tekoälyn tuoma ratkaisu tähän oli, että dokumentteja sisältävään historiadataan oli merkitty riskejä, joiden pohjalta koneoppimisen malli koulutettiin. Käytännössä malli toimii siten, että analyytikon saadessa uuden dokumentin, analyytikko syöttää sen mallille ja malli käsittelee dokumentin automaattisesti ja näyttää suoraan minkälaisia riskejä mistäkin kappaleesta on löytynyt, minkä jälkeen analyytikko validoi, oliko mallin ehdottama riski oikea riski vai ei. Tämä nopeuttaa prosessia huomattavasti ja samalla mallia opetetaan koko ajan paremmaksi. Jos kone on tehnyt väärän arvion, voidaan mallille merkitä, että koneen ehdotus ei ollut riski. Tämä on aktiivista oppimista, jossa ihminen on mukana. (Haastattelu 5)

Tiedonhakuja voidaan myös automatisoida ilman tekoälyä, jos on tiedossa, miten kaikkea saatavilla olevaa tietoa voidaan järkevällä tavalla käyttää päätöksen tukena. Kaikki se mitä yrityksessä Googlataan analysoitaessa yrityksiä, on työ, minkä kone pystyy hoitamaan ihmisen puolesta. Prosessi on automatisoitavissa siten, että se mitä Google tekee, pystyy tekemään mikä tahansa muu kone helposti ja hakemaan tarvittavat tiedot, kunhan tiedetään mitä tietoa

tarvitaan. Google kattaa vain osan internetissä olevista tiedoista, mutta ei koko internetiä. Kaikki tällainen tieto on tuotavissa luottoprosesseihin niin sanotusti pureskeltuna. Tällaisia työkaluja on helposti rakennettavissa, eikä se vaadi tekoälyä. (Haastattelu 4)

Koneoppimisessa päätöksen teko perustuu opetusdataan, jota järjestelmälle on syötetty. Luottoprosesseissa käytetään lähes kaikissa tapauksissa ohjattua oppimista, jossa järjestelmälle on ennalta kerrottu haluttu lopputulos, sillä ohjatun oppisen malli on helposti kontrolloitavissa. Joissakin, mutta todennäköisesti ainakin toistaiseksi harvoissa tapauksissa saatetaan kuitenkin käyttää myös vahvistusoppimisen mallia. Luottoprosesseissa siinä kuitenkin voi olla riski, sillä malli oppii koko ajan lisää samalla, kun se tekee päätöksiä. Vahvistusoppimisen malli etsii jatkuvasti parempia ratkaisuja aiemmin tehtyjen päätösten pohjalta ja käyttää hyväkseen aiemmin oppimaansa tietoa. Esimerkiksi luottopäätöksissä vahvistusoppimisen malli saattaa tehdä huonoja päätöksiä vain sen pohjalta, että malli tarkistaa, että saman kaltaiselle yritykselle on jo kerran annettu myönteinen luottopäätös, joka päätyi luottotappioksi, mutta vahvistusoppimisen malli haluaa tietää, oliko se vain sattumaa ja tehdä uuden positiivisen luottopäätöksen ja oppia sen pohjalta lisää. Jos tällaista menetelmää halutaan käyttää luottopäätösten teossa, olisi hyvä olla käytössä simulaatioympäristö, jossa dataa on turvallista testata. (Haastattelu 1,2 ja 4)

Yhdessä haastattelussa haastateltava kertoi, että regressiomallit ja analyysit ovat osoittaneet parhaan toimivuuden luottopäätöksiä tehdessä. Hyvyys on siinä, että niiden toimintakykyä pystytään myöhemmin arvioimaan ja raportoimaan. Sellaisiin koneoppiviin malleihin, kuten neuroverkot, joissa muuttujat muuttuvat jatkuvasti on suhtauduttu varauksella, koska niitä ei pysty jälkikäteen selittämään. Tämä on niin kutsuttu ”black box”, jossa muuttujan vaikutusta ei pystytä jälkikäteen todentamaan. Tämän tyyppisiä malleja kuitenkin kokeillaan jatkuvasti ja käytetään yhä enemmän ja enemmän. Tietoaineistot, joita käytetään tällaisissa ennusteissa, laajenee koko ajan ja menetelmiä, joita käytetään, kokeillaan ja testataan jatkuvasti, mutta tällä hetkellä 99% maailmassa myönnytyistä luotoista taustalla on perinteiset tilastolliset analyysit eli regressiomallit. Toisessa haastattelussa haastateltavalta kysyttiin, mikä on jo vanhentunutta tekoälyn käyttöä. Sain vastaukseksi, että on vaikea nimetä mikä on uutta ja mikä vanhaa, sillä teknologia ei koskaan ole itseisarvo vaan lähtökohta on se, että voidaanko jokin tietty liiketoiminnan osa-alue ratkaista tekoälyn avulla. Tyypillisesti vanhana saatetaan pitää esimerkiksi joitakin regressiomalleja, mutta tiettyihin tilanteisiin ne voivat kuitenkin olla nimenomaan parhaita ratkaisuja, eikä suinkaan yhtään huonoja ratkaisuja. Monimutkaisempia menetelmiä, kuten neuroverkkoja voidaan kokeilla niissäkin ja tarkistaa, että paraneeko tulokset vai ei.

Kiinteistövakuuksien arvioinnissa voidaan hyödyntää saman kaltaisia malleja, joita kiinteistövälitysyritykset hyödyntävät. Kiinteistön arvoon vaikuttaa muun muassa alue, koko ja se missä kunnossa kiinteistö tai asunto on. Kiinteistöjen tai asuntojen arvioinnissa voidaan

käyttää apuna tekoälyjärjestelmää, jossa malli koulutetaan historiadataan avulla. Historiadata voi sisältää esimerkiksi tiedon aiemmista myyntihinnoista tai vuokrahinnoista, jolloin näiden ja muiden annettujen kriteereiden perusteella pystytään antamaan arvio siitä, mikä on kiinteistön tai asunnon optimaalinen hintahaarukka. Tässä tapauksessa ihminen syöttää järjestelmälle asunnon tai kiinteistön tiedot, minkä jäljkeen malli antaa ennusteen hintahaarukasta, mutta ihminen lopulta tässäkin tekee virallisen päätöksen. Prosessia voidaan tällä tavoin kuitenkin nopeuttaa huomattavasti. (Haastattelu 5)

Ohjelmistorobotiikan (RPA) ja koneoppimisen yhdistäminen tuo merkittävää arvoa toisilleen, sillä ne ovat toisiaan täydentäviä tekniikoita. Ohjelmistorobotiikkajärjestelmät kykenevät strukturoimaan ja hallitsemaan digitaalisen työn kulkua liiketoimintaprosesseissa, kun taas koneoppiminen ansiosta RPA-järjestelmät pystyvät tekemään monimutkaisempia päätöksiä. Tietyt koneoppimismallit mahdollistavat ohjelmistorobotiikan käsitellä strukturoimatonta dataa, kuten esimerkiksi kuvia tai tekstiä. Koneoppiminen yhdistettynä ohjelmistorobotiikkaan antaa myös RPA-järjestelmälle mahdollisuuden tehdä laajempia päätöksiä nopeammin ja tarkemmin, kuin yksin toimiva RPA-järjestelmä. Näiden tekniikoiden yhdistämisessä RPA-järjestelmän tehtävänä on hallita prosessin kulkua, kerätä tieto useammista tietojärjestelmistä ja mahdollisesti tehdä joitakin yksinkertaisia päätöksiä sääntöjen avulla. Kun RPA-järjestelmä ja koneoppimisjärjestelmä ovat yhdistetty toisiinsa, tyypillisesti RPA-järjestelmä on vuorovaikutuksessa koneoppimisjärjestelmän kanssa API-rajapintojen kautta. Koneoppimisen malli, jota on koulutettu datan avulla, ja joiden vasteet tunnetaan, käyttävät tietoja, joita RPA-järjestelmä on kerännyt transaktiojärjestelmästä. (Davenport 2019)

Koneoppimisen ja ohjelmistorobotiikan yhdistäminen on merkityksellistä aina, kun liiketoiminnan automatisointia harjoitetaan integroidulla ja strategisella tavalla. Näiden kahden tekniikan yhdistäminen on välttämätöntä, jotta liiketoimintaprosessit voidaan automatisoida tehokkaasti ja strategisesti. Lisäämällä ohjelmistorobotiikkaan älykkyyttä ja yhdistämällä koneoppimismahdollisuudet automaatioprosessiin, voidaan suunnitella edistyneen ohjelmistorobotiikan malli, joka pystyy analysoimaan, ymmärtämään ja tekemään johtopäätöksiä strukturoidusta sekä strukturoimattomasta datasta. Näiden kahden tekniikan yhdistäminen ei siis pelkästään pysty käsittelemään dataa vaan käyttämään sitä myös tehokkaasti. Tällainen älykäs ohjelmistorobotiikka analysoi dataa ennen toimimistaan, oppii jatkuvasti käyttämästään datasta, tulee ajan myötä älykkäämmäksi ja tekee älykkäitä päätöksiä aiemman oppimisen perusteella. Koneoppimisen ja ohjelmistorobotiikan yhdistäminen on kannattavaa silloin kun käsitellään, analysoidaan, vertaillaan ja jäsennellään valtava määrä dataa. Tätä teknologiaa voidaan käyttää esimerkiksi kuvan- ja puheentunnistukseen sekä asiakirjatietojen poimintaan. Yksinkertaisesti kiteytettynä koneoppiminen suorittaa ajattelun, kun ohjelmistorobotiikka suorittaa toiminnon (Crnoja & Col 2019). Ohjelmistorobotiikan avulla voidaan kerätä dataa ja automatisoida datan keräämistä. Ohjelmistorobotiikkaa kehitettäessä tulee luoda ohjelma, jota kautta data siirtyy

järjestelmästä toiseen. Ohjelmistorobotiikka on tärkeässä roolissa, kun luodaan automaattista alustaa, josta voidaan rakentaa koneoppimisjärjestelmä. Järjestelmä luo organisaatiolle arvoa itse kerätystä datasta ja sitä kautta moninkertaistaa datan arvon. (Ojanperä 2017)

6.2 Datan hankinta

Useammasta lähteestä ja eri muodoissa olevan datan kokoaminen yhdeksi kattavaksi kokonaisuudeksi on yksi keskeisimmistä haasteita ennen varsinaista analyysia. Tiedon ja tietolähteiden määrä on lisääntynyt digitalisoitumisen ansiosta, mutta eri tietolähteistä saatavan datan määrä vaihtelee. Se kuinka tietoja on eri lähteistä saavutettavissa, riippuu datan avoimuudesta, lisensoimisesta ja maksullisuudesta. Lisäksi datan saantia rajoittaa yksityisyyden suoja ja henkilötietoihin liittyvät ehdot, joita voitaisiin hyödyntää yrityksen vastuuhenkilöiden tai takaajien analysoinnissa. Haasteita datan hyödyntämiselle asettaa myös datan virheellisyys, monimuotoisuus ja monitulkintaisuus. Datan hyödyntäminen päätöksenteon apuna vaatii useamman tiedon kokoamista, suodattamista, esikäsittelyä, yhdistämistä sekä analysointia. (Ylén, Bäck ym. 2018)

Haastatteluissa on käynyt ilmi, että yrityslainoissa koneoppimisen menetelmänä käytetään ohjaukseen perustuvaa menetelmää. Ohjatussa oppimisessa malli oppii tunnistamaan riskejä, jolloin käytössä on oltava esimerkkejä eri tapauksista ja siihen on oltava syöte (input) eli parametrit, joilla päätöksiä tehdään ja vaste (output) halutusta tilanteesta. Tällöin käytössä on oltava nimenomaan tähän tiettyyn liiketoimintaprosessiin liittyvää dataa. Jos tällaista dataa löytyy, on yrityksellä edellytykset tehdä koneoppimisen malli jopa pelkästään omalla datalla. Datan lähteinä voidaan käyttää organisaation sisäisiä datan lähteitä, avoimia lähteitä tai ostettavissa olevaa dataa, jota on finanssisektorilla paljonkin. Jos dataa ei ole, se ei ole ongelma vaan enemmänkin hidaste, jolloin tapa kerätä dataa tulisi implementoida osaksi nykyistä prosessia. Kunnes dataa on riittävästi, tulee etsiä malli ja lähetään viemään kehitystä eteenpäin. (Haastattelu 5)

Markkinoilla on maksullisia toimijoita, jotka kokoavat ja sen jälkeen jalostavat dataa muille yrityksille hyödynnettäväksi, kuten esimerkiksi Tilastokeskus ja Suomen ympäristökeskus. Tilastokeskuksen kaupunki- ja seutuindikaattorit sisältävät kaupunkeihin liittyviä tietoja, joita ovat esimerkiksi aluetalous, asuminen, koulutus, kulttuuri, kunnallistalous, liikenne, rakentaminen, rikollisuus, sosiaalitoimi, terveydenhuolto, tulot, työmarkkinat, vaalit ja väestö. Tällaista alueille kohdistettua dataa on saatavilla ilmaiseksi sekä maksullisena Tilastokeskukselta. Tilastokeskuksen ilmaisia Paavo-aineistojen kautta on saatavissa avointa dataa postinumeroalueittain esimerkiksi asuntorakenteista, koulutusasteista, asukkaiden ja talouksien tuloista, talouksien koosta, elämänvaiheesta, rakennuksista, asunnoista sekä työpaikoista. Lisäksi Suomen Ympäristökeskuksen ja Tilastokeskuksen ympäristörakenteen (YKR) aineistosta on saatavilla vastaavanlaista tietoa ruutu-pohjaisesti, joka tarjoaa vieläkin

tarkempaa tietoa postinumeroalueittain. Dataa voidaan hankkia myös esimerkiksi Väestörekisterikeskukselta, jolla on erilaista tietoa henkilöistä, asunnoista, rakennuksista ja huoneistoista, mutta aineisto ei kuitenkaan ole saatavilla edes maksua vastaan, vaan niiden käyttöön tarvitaan erillinen käyttöluupa. Väestörekisterikeskuksen väestötietojärjestelmää sekä PRH:n yritys- ja yhteistyötietojärjestelmiä voidaan myös hyödyntää datan lähteinä.

Useammissa haastatteluissa on käynyt ilmi kysymyksen ”miten data hankintaa” -kohdalla, että datan hankinnassa teknologiana on käytetty hyvin usein API-rajapintoja, jotta data saadaan eri lähteistä omiin järjestelmiin. Osa datasta, kuten esimerkiksi tilitiedot voidaan hankkia pankeilta. Nordea on julkisesti ilmoittanut rajapintojen avauksen olevan heille itselleen liiketoimintamahdollisuus ja haluavansa tehdä sen mahdollisimman hyvin, minkä takia Nordea on lähtenyt hakemaan kehittäjäyhteisöjä. Jarkko Turunen Nordealta kommentoi asiaa Api talous 101 -kirjan kirjoittajille:

”Nordean lähestymistapa PSD2:een on ollut hyvin proaktiivinen. Nordea avasi Open Banking -portaalinsa kehittäjille keväällä 2017, jonka jälkeen jo yli 1600 kehittäjää on rekisteröitynyt kokeilemaan ohjelmistorajapintoja testiympäristössä. Pilotivaiheessa valikoidut ulkopuliset palveluntarjoajat rakentavat sovelluksia rajapintojen päälle ja vahvistavat yhdessä pilottiasiakkaiden kanssa, että Open Banking -ratkaisu toimii kaikilta osin luotettavasti ja odotetulla tavalla. Pilotissa käytetään aluksia vain suomalaisten asiakkaiden tietoja, jonka jälkeen mukaan saadaan asiakastiedot myös muista Pohjoismaista. Nordeassa uskotaan PSD2:n tarjoavan mahdollisuuksia uusiin liiketoimintamalleihin, joissa pankki sekä sen asiakkaat ja kumppanit voivat hyötyä uusista innovaatioista ja tarjonnasta. Puhutaan tulevaisuuden maksu- ja sovellusekosysteemeistä. API-alustat toimivat jatkossa pankkipalvelujen uutena jakelukanavana. Pakollisten PSD2-rajapintojen lisäksi Nordea näkee Open Banking -ratkaisussa mahdollisuuden tulovirtoihin esimerkiksi maksullisten Premium API:n kautta.” (Turunen 2018)

6.3 Ihmisen rooli

Koneoppimismallit eivät korvaa ihmisen tekemää työtä täysin, vaan tyypillisesti kyseessä on raskas ja aikaa vievä prosessi, jossa on käytössä useita työntekijöitä, jolloin prosessiin halutaan apua koneoppimismalleilta. Koneoppimista hyödyntämällä prosessista tulee aikaisempaa nopeampi ja ihminen voi keskittyä niin sanotusti järkevämpiin asioihin. Tällöin kone ei tee lopullista päätöstä itse, vaan antaa vain ehdotuksen. Ihminen aina tarkistaa työn. (Haastattelu 5)

Myös tekoäly oppii ihmiseltä samalla, kun se auttaa tehostamaan ihmisen tekemää työtä. Kone suorittaa toiminnon, kysyy asiantuntijalta sisään syötettäviä tietoja ja oppii saamastaan vastauksesta, minkä ihmisasiantuntija on antanut. Ihanteellinen vuorovaikutusprosessi ei pelkästään auta parantamaan asiantuntijoiden työtä, vaan myös kone pystyy hyödyntämään

järjestelmän kanssa vuorovaikutuksessa olevan asiantuntijan tietoa ja ammattitaitoa. Tällä tavalla asiantuntijoilla olevaa hiljaista tietoa voidaan hyödyntää ottamalla se osaksi järjestelmää, kuten kuviossa 5 on esitetty. (Hulkko 2018)



Kuvio 8: Human in the loop AI, Silo AI 2018.

Konkreettisessa esimerkissä asiantuntijat selaavat päivittäin lukuisia asiakirjoja etsiessään erilaisia havaintoja. Vaikka tekoäly voidaankin kouluttaa havaitsemaan riskejä käyttämällä esimerkiksi erilaisia luonnollisen kielen käsittelymenetelmiä (NLP), tekstin tulkitseminen ja prosessin luotettavuus riippuu edelleen suuresti ihmisen korkean tason asiantuntemuksesta. Tekoälyjärjestelmät tuottavat asiakirjoista löydetyistä tiedoista ehdotuksia riskejä osoittavista lauseista, minkä jälkeen ihminen tarkastelee tuloksia antaen jokaiselle ehdotukselle palautetta siitä, onko järjestelmän tuottama havainto hyvä tai huono. Tekoälyjärjestelmä oppii ihmisen antamasta palautteesta ja tällä tavalla oppii parantamaan suorituskykyään. Ihmisen suorittaman valvonnan ansiosta aikaisempi laatu säilyy, mutta se vain suoritetaan nopeammin jokaisella palautekierröksellä. Tätä tapaa voidaan soveltaa lähes kaikessa liiketoimintaprosesseissa, jossa ihmisasiantuntijat soveltavat hiljaista tietoa tehtävien suorittamiseen, esimerkiksi rahoituksessa petosten havaitsemiseen. Tällaista järjestelmää voidaan käyttää myös vähemmän kokeneiden työntekijöiden kouluttamiseen. Järjestelmä on alun perin koulutettu korkean tason asiantuntijoiden kanssa, jolloin järjestelmä voi tehdä ehdotuksia ja löytää poikkeavuuksia myös muille työntekijöille, mitkä olisivat muuten saattaneet jäädä huomiotta vähäisemmän kokemuksen takia. Tekoälyn ei siis ole tarkoitus korvata ihmisasiantuntijoita, vaan lisätä organisaatiossa asiantuntijatietoa ja jakaa sitä. Tekoälyjärjestelmät oppivat saamastaan palautteesta ja auttavat ihmisasiantuntijoita parantamaan laatua ja suorittamaan toimintoja nopeammin (Hulkko 2018). Esimerkiksi luonnollisen kielen käsittelyssä ihmisen on tarkistettava dokumentteja ja tukevatko tosiseikat tarjolla olevaa dataa ja kuinka relevanttia data on. Luonnollisen kielen käsittelyssä

dokumentteja pystytään lukemaan koneellisesti ja tuottamaan niistä digitaalisia tiedostoja. Ihminen kuitenkin pystyy keskittymään niihin, mitkä tarvitsevat kriittisempää käsittelyä ja enemmän aikaa. (Aaltonen 2019)

Ihmisen työ tulee kaiken kaikkiaan muuttamaan muotoaan. Tiedon hakuun ja tekstin analysointiin liittyvät prosessit tulevat automatisoitumaan sen sijaan, että ihminen käyttää aikaa tiedon hakuun ja haetun tiedon analysointiin, vaan tulevaisuudessa entistä enemmän koneälyn ehdotuksien pohjalta tehdään päätöksiä, mutta ihminen kuitenkin edelleen tekee sen päätöksen, mutta vain nopeammin. (Haastattelu 5)

Tekoälyjärjestelmät tekevät kuitenkin ihmistä vähemmän virheitä, vaikka myös virheitä voi tapahtua esimerkiksi huonosti toteutetusta tekniikasta johtuen. Tekoälyn tekemiksi virheiksi voidaan enneminkin luokitella niin sanotut eettiset virheet, jossa esimerkiksi päätöksen tekoon on vaikuttanut asuinpaikka, ihon väri tai sukupuoli. Toisaalta myös pahimmassa tapauksessa ihminen voi tehdä samat virheet, sillä ihmisen päätökset voivat perustua tunteisiin. Virheitä on todennäköisesti sattunut puolin ja toisin, mutta haastattelujen perusteella ihminen tekee useammin virheitä päätöksen teossa. (Haastattelu 1,2 ja 3)

6.4 Ensimmäiset askeleet

Yritysten ottaessa käyttöön koneoppimista omassa liiketoiminnassa, tulee yritysten pitää huoli, että työntekijöillä on perusymmärrys koneoppimisesta ja sen aiheuttamasta muutoksesta. Jo pienemmälläkin materiaalilla voidaan saavuttaa työntekijöille huomattavasti aikaisempaa parempi tietoisuus koneoppimisesta ja sen hyödyntämisestä. Koneoppimisen käyttö vaatii suuren määrän dataa, joten organisaatioissa on määriteltävä datastrategia. Lisäksi on mietittävä, mitä dataa organisaatio myös tulevaisuudessa tarvitsee, jotta tekoälyjärjestelmä voidaan tulevaisuudessa opettaa sellaiseksi, jota kilpailukyky edellyttää. Jo alkuvaiheessa on siis mietittävä, minkälaista dataa tarvitaan esimerkiksi kolmen vuoden kuluttua. (Siilasmaa 2017)

Yritys tarvitsee kyvykkyyttä järjestelmien luomiseen siinä vaiheessa, kun tekoälyn käyttö tulee ajankohtaiseksi. On siis oltava mahdollisuus siihen, että testijärjestelmä on rakennettavissa, kun herää kysymys, että voiko halutun ongelman ratkaista tekoälyjärjestelmällä. Uutta kilpailukykyä ja älykkyyttä tulee luoda tuotteisiin ja palveluihin koneoppimisen kautta, joten yrityksissä, jotka haluavat tekoälyn avulla parantaa kilpailukykyä, kannattaa alkaa toimia konkreettisesti jo nyt kun vielä on aikaa, sillä muutos tulee myöhemmin olemaan vielä nopeampaa. Ihmiset tulevat edelleen jatkamaan tietokoneiden, laskentakapasiteetin ja algoritmien kehittämistä, joten jonakin päivänä järjestelmät saavuttavat todellisen älykkyyden. Tällä hetkellä ei kuitenkaan ole mitään näkymää siihen, miten se on mahdollista,

sillä nykyiset järjestelmät eivät koskaan tule skaalautumaan sille tasolle, vaan tarvitaan jokin täysin erilainen keksintö ennen kuin se on mahdollista. (Siilasmaa 2017)

Yrityksien, jotka haluavat hyödyntää tekoälyä liiketoimintaprosesseissa kannatta etsiä hyvä tekoäly-yritys kumppaniksi, jonka kanssa käydään läpi tärkeimmät prosessit, mitkä prosessit vievät kaikista eniten aikaa ja mitkä toistuvat. Näistä yhdessä tekoäly-yrityksen kanssa katsotaan merkittävimmät kohdat ja identifioidaan ne, mihin koneoppimista voidaan hyödyntää ja mihin sillä olisi merkittävä vaikutus liiketoiminnassa. Sitä kautta rakennetaan toimintasuunnitelma niiden ympärille, jolloin tekoäly-yritys pureutuu organisaation prosesseihin ja selvittää mihin on järkevä hyödyntää tekoälyä. Kaikkea voi tehdä, mutta kannattaako, on pääpointti. On parempi, että aluksi koneoppimisen malli analysoi ja antaa arvion, jolloin jossain vaiheessa kone voi myös automaattisesti hyväksyä ainakin osan hakemuksista. Aluksi ei kuitenkaan kannata pyrkiä täydelliseen automaatioon, vaan antaa ihmisen tehdä ne päätökset, mutta vain aikaisempaa nopeammin ja tarkemmin. (Haastattelu 5)

7 Benchmarking tulokset

Benchmarking-menetelmässä on käytetty viittä eri yritystä, jotka kaikki jollain tavalla analysoivat yrityksiä liiketoimintaprosesseissaan. Tulokset on kerätty haastattelemalla neljän eri yrityksen edustajaa sekä yhden yrityksen nettisivuilta löytyneestä tiedosta. Viidennelle yritykselle ei ole tehty haastattelua. Haastattelut ovat koostuneet liitteessä 2 olevasta haastattelurungosta, joihin saatuja vastauksia on avattu tässä luvussa.

7.1 Yritys Z

Yrityksen käytössä ei ole varsinaista tekoälyjärjestelmää vaan ohjelmistorobotiikkajärjestelmä, joka tekee saadun datan perusteella ehdotuksen annettujen ehtolauseiden perusteella, myönnetäänkö lainaa vai ei. Ihminen tekee kuitenkin aina lopullisen päätöksen. Lainahakemukset, jotka ovat maksimissaan 100.000 euroa pyritään tekemään ohjelmistorobotiikan ehdotuksen mukaisesti. Selvät tapaukset, joissa lainaa ei myönnetä pyritään myös ensisijaisesti tekemään ohjelmistorobotiikan ehdotuksen mukaisesti. Suuremmat hakemukset käsittelee aina ihmisasantuntija. Myös mahdollisten vakuuksien analysoinnin tekee ihmisasantuntija, eikä niiden arviointiin käytetä tekoälyä tai muuta automatiikkaa. Maksullisina datan lähteinä käytetään asiakastietoa, josta mm. tilinpäätöstiedot ja taseet haetaan sekä Bisnodea ja perintäyhtiö Lindorfia. Maksuttomana datan lähteenä käytetään YTJ:tä, josta hankitaan yrityksen perustietoja. Lainaa hakevista yrityksistä kerätään dataa, minkä jälkeen ohjelmistorobotiikkajärjestelmä tekee ehdotuksen, myönnetäänkö lainaa vai ei, mutta ihminen tekee kuitenkin aina lopullisen päätöksen. Yritys ei käytä ainakaan tällä hetkellä

koneoppimisjärjestelmää eli järjestelmä ei opi uutta sille syötetystä datasta, eikä käytä historiadataa ehdotuksia tehdessään, vaan tekee ehdotuksen sille syötettyjen ehtolauseiden perusteella.

7.2 Yritys Y

Yritys ei itse suoranaisesti käytä tekoälyä luotonmyönnössä, mutta dataa kerätään automaattisesti eri lähteistä. Data kerätään rajapintojen kautta järjestelmään muun muassa Bisnodelta, perintätoimistoilta ja Risk Ratelta. Riskrate perustuu reaaliaikaiseen ja todelliseen yrityksen maksukäyttäytymiseen, jossa hyödynnetään kolmansia tietolähteitä ja tekoälyä. Yritys Y:n järjestelmään tulee suoraan Riskraten ”liikennevalot”, sen mukaan miten yrityksen reaaliaikaisen maksukäyttäytymiseen on ennustettu vaikuttavan riskiin. Yritys valmistuu luotonmyöntöön myös pitkälti perintätoimistoilta tulevan datan avulla. Haastateltava kuitenkin mainitsee, että jos tekoälyn käyttöön olisi nyt täysin vapaat kädet niin yritysanalyysia haluttaisiin kokeilla erilaisilla pehmeillä mittareilla, mitkä eivät perustu pelkästään taloudellisiin lukuihin, jolla voitaisiin selvittää miten yritys, hallitus tai vastuuhenkilöt käyttäytyvät esimerkiksi sosiaalisessa mediassa tai miten esimerkiksi yrityksen toimipaikka vaikuttaa riskeihin.

7.3 Yritys X

Yritys X hyödyntää yrityksen arvioinnissa tilastollisiamenetelmiä ja malleja, jossa käytetään algoritmeja laskemaan todennäköisyyksiä, mutta haastateltava mainitsee, että tekoäly sanana on kuitenkin paljon sanottu. Yritys käyttää erilaisten mallien rakentamiseen datana kaikkea käytössä olevaa tietoa, kuten yrityksen nimeä, osoitetta, toimipaikkaa, vastuuhenkilöitä ja niiden linkkejä muihin yrityksiin, taloudellisia tunnuslukuja, konsernirakenteita, taseita tai tytä- ja emoyhtiöiden tietoja. Myös maksukäyttäytymisestä on saatavilla erilaista dataa esimerkiksi viranomaislähteistä, josta tulee tieto konkurseista ja muista maksuhäiriöistä, kuten trattatiedot. Maksukäyttäytymistietoa kerätään mallien tueksi myös eri asiakkaiden tai kumppaneiden myyntireskontrasta, josta laskuissa olevaa relevanttia tietoa irrotetaan aineistokantaan, minkä jälkeen analytiikkaosasto rakentaa tilastollisia malleja. Käytännössä tällaisessa tilanteessa on jokin vastemuuttuja, jota vastaan ennustetaan. Rahoituslalla tyypillisesti ennustetaan, millä todennäköisyydellä yritys saa maksuhäiriön esimerkiksi seuraavan 12 tai 36 kuukauden aikana, kuinka todennäköisesti yritys ajautuu konkurssiin tai kuinka suurella todennäköisyydellä yritys kasvaa, kuinka paljon ja missä ajassa. Kaikissa tapauksissa ennustus perustuu siihen, että dataa kerätään niin paljon, kun pystytään laadullisesti kohdistamaan yrityksen y-tunnukseen. Mallinnustyön lopputuloksena syntyy osajoukko niistä valituista tiedoista, mitä yritys X:llä on käytössä. Tilastollisessa analyysissa puhutaan siitä, että eri muuttujilla on eri painoarvot. On löydettävä ne tekijät, mitkä aidosti vaikuttavat yrityksen

tilanteeseen. Mitä laadukkaampi aineisto ja mitä paremmat aikasarjat, sitä paremmat algoritmit sen päälle rakennetaan ja sitä tarkempi on lopputulos. Luokitusmallien toimivuutta tarkistetaan ja arvioidaan erilaisten ristiintarkastus mallien kautta, mitkä kertovat, kuinka hyvin mallit oikeasti toimivat. Näitä malleja päivitetään jatkuvasti.

Luonnollisen kielen prosessointia (NLP) yritys X hyödyntää muun muassa siten, että yritys kerää tietoaineistoa myös suomalaisien yritysten nettisivuilta. Sivuilta imuroidaan kaikki data, kuten linkit Facebookiin, Twitteriin, LinkedIn:iin tai Youtubeen. Näistä tiedoista kerätään ja strukturoidaan rakenteellinen tieto yritys X:n tietokantaan, minkä jälkeen pystytään kertomaan millä yrityksellä on esimerkiksi digitaalista jalanjälkeä. Myös näitä tietoja tarkastellaan jatkuvasti ja selvitetään, onko näillä muuttujilla ollut vaikutusta luokitusmallien performanssiin. Kun ennustetaan esimerkiksi millä todennäköisyydellä yritys kasvaa yli 10% tai 5%, löytyy internetistä tietoa myös mm. siitä, onko yritys rekrytoimassa ja minkä tyyppistä aktiivisuutta internetistä löytyy. Nämä kaikki tiedot ovat olleet tilastollisesti merkittäviä muuttujia, kun ennustetaan yrityksen kasvukykyä.

Toinen käytännön esimerkki, jossa yritys X hyödyntää luonnollisen kielen prosessointia, on tilinpäätöstiedot ja niiden liitetiedot, joista on automaattisesti luettu ne tekijät, joilla tilinpäätöstä voidaan esimerkiksi leasingvastuiden osalta oikaista automaattisesti. Tilinpäätösaineiston toimintakertomuksesta ja liitetiedoista voidaan nähdä myös esimerkiksi, jos jossakin yrityksessä on otettu toiminnanohjauskäyttöön ja sen käyttöönotossa on ollut haasteita, mikä on vaikuttanut liiketoimintaan ja useita investointeja on tehty, jolloin tulevaisuuden näkymät ovat vakaat. Tällaista tekstiä löytyy toimintakertomuksista ja liitetiedoista, jossa voidaan luonnollisen kielen prosessointia hyödyntää.

Suomessa esimerkiksi kaikki tilinpäätösaineistot ovat ainakin toistaiseksi pdf -muodossa ja yritys X lukee ne digitaaliseen muotoon koneilla. 70-80% niistä saadaan onnistumaan täysin automaattisesti, vaikka tiedot tilinpäätöksissä voivat olla miten sattuu, mutta koneet pystyvät lukemaan niitä digitaalisesti ja keräämään tiedot oikeisiin paikkoihin. Loput 20-30% täytyy tarkistaa käsin, että tiedot ovat oikeissa paikoissa. Tämän haastateltava sanoo olevan kuitenkin normaalia arkipäivää ja digitaalista tietojenkäsittelyä, eikä luonnollisen kielen prosessointia. Luonnollisen kielen prosessoinnilla pystytään esimerkiksi irrottamaan relevanttia tietoa kuvailevasta tekstistä, mikä ei ole tilinpäätösluku tai strukturoitu tieto.

Tekoälyä hyödynnetään yritys X:ssä myös yrityksen todellisen toimialan selvittämiseen. Yritykset ilmoittavat aina itse oman toimialansa, mutta tyypillisesti toimiala ei ole täysin se, minkä yritys on ilmoittanut itsestään, jolloin yritys saattaa todellisuudessa tehdä jotakin muuta. Kaikella aineistolla, jota yrityksistä kerätään, voidaan oikaista yrityksen virallista toimialaa siten, että kaikesta kuvailevasta tekstistä, jota on pystytty keräämään internetistä tai muista teksteistä, on voitu selvittää, että yritys tekee todellisuudessa jotakin muuta. Näitä

testaillaan yritys X:ssä ja käytetään erilaisia työkaluja, miten ennustettavuus olisi automaattista, eli kokeiluja on, mutta haastateltavan mukaan ei vielä ”mainstreamia”.

Vakuuksien arvioinnissa yritys X käyttää myös tilastollisiamalleja, mutta haastateltava sanoo, että se, onko ne perinteisiä, itseoppivia vai tekoälymalleja, on menetelmä kysymys. Yritys X arvioi asuntojen ja kiinteistöjen kehittymistä, mikä perustuu tilastolliseen mallinnustyöhön, jossa otetaan koko aineisto käyttöön: kaikki toteutuneet kauppahinnat, toteutuneet kohteet postinumerotasolla ja mikä missäkin osoitteessa olevan kiinteistön arvo nyt ja tulevaisuudessa. Tämän haastateltava sanoo kuitenkin olevan raakaa mallinnustyötä.

Erlaisia yrityksien hankkeita arvioitaessa perinteisesti pankit ovat tehneet tilinpäätösanalyysia ja asiakkaalla on ollut käytössä ennustemalleja, jossa ennustetaan, että millä tavalla jokin tietty investointi vaikuttaa tulevien tilikausien tuloslaskelmaan ja taseeseen. Jos investointi vaikuttaa positiivisesti liiketoimintaan, niin miltä tuloslaskelma ja tase näyttää investoinnin jälkeen, mutta haastateltava sanoo tekoälyn olevan kaukana tällaisessa ennusteessa, sillä se on raakaa arviota, jossa on automaattiset laskentamallit. Esimerkiksi jos yritys tarvitsee 1 miljoonan euron investoinnin, tulee se syöttää laskentamoduuliin, jolloin voidaan saada arvio siitä, miten miljoonan euron investointi vaikuttaa tuleviin kassavirtoihin. Haastateltava sanoo, että jos tässä halutaan käyttää tekoälyä, pitää pystyä osoittamaan, että toimiiko sellainen malli paremmin, kuin entinen tai nykyinen.

Haastattelun lopussa haastateltava tarkentaa vielä: *”Puhutaan paljon tekoälystä ja koneoppimisesta ja on hyvä, että niitä nostetaan. Se tuo kysyntää myös meidän palveluille ja voitais jotain näistä meidänkin tuotteista markkinoida tekoälytuotteina jos me haluttais, mutta tosi asia on se, että aika paljon se on kuitenkin automaattista tietojenkäsittelyä eli ATK:ta. Enemmän on tarvetta tiedolle, joka on hyvin analysoitu.. ja teknologia, miten tämä tempu toteutetaan, on paljon kustannustehokkaampaa, kun aikaisemmin. Mekin pystytään tekemään kokeiluja nopeasti, että toimiiko joku vai ei, koska teknologian hankintakustannus on merkittävästi pienempi kuin 10 vuotta sitten.”*

7.4 Yritys W

Yritys W on automatisoinut henkilölainojen luottopäätökset, mutta yrityksen luottotiimi kuitenkin käsittelee joissakin tapauksissa päätökset manuaalisesti loppuun. Aina kuitenkin lopullista päätöstä ei tee ihminen vaan automattinen luottopäätös. Tietosuoja-asetuksen myötä asiakkaalla on oikeus olla joutumatta niin sanotusti automaattisen päätöksen tekemäksi, jolloin asiakkaan vaatiessa hakemus voidaan ottaa uudelleen manuaaliseen käsittelyyn. Tekoälyä yritys ei kuitenkaan tällä hetkellä juurikaan hyödynnä automaattisessa luottopäätöksessä, mutta yrityksestä kuitenkin löytyy datan louhinta ja tekoälyosaamista. Kuitenkaan

automaattiseen luottopäätökseen sitä ei hyödynnetä. Haastateltava sanoo itse uskovansa, että tekoäly soveltuu enemmän vastauksien etsimiseen ja datan pyörittämiseen, jolloin varisnaiseen automaattipäätökseen ei kuitenkaan käytettäisi tekoälyä nykyisestä regulaatiosta johtuen.

Yritys hyödyntää datan lähteinä omaa dataa sekä julkisista lähteistä saatavaa dataa, mutta myös asiakkaalta pyydettävät tiedot ovat osa automaattipäätöstä. Automaattisen luottopäätöksen taustalla on ihmisen tekemää koodia, mikä on perustunut käytössä olevaan dataan tai yritys W:n omaan riskipolitiikkaan tai viranomaisilta tuleviin linjauksiin. Automaattisessa luottopäätöksessä ei siis ole käytössä tekoälyä tai robotteja.

Yrityksellä on myös vakuudellisia tuotteita, jotka ovat automaattisen luottopäätöksen piirissä, jolloin myös vakuuksia arvioidaan automaattipäätännässä, mutta myöskään sitä ei luokitella tekoälyksi. Yritys saa joko omista tai yhteistyökumppanin rekistereistä arvion kohteen vakuusarvosta, jota hyödynnetään yhtenä osana automaattista päätöstä. Yritys W:llä on kuitenkin myös omia vakuusarvomalleja, mistä tulee joidenkin muuttujien perusteella arvioitu vakuuden arvo. Siinä mielessä vakuuden arviointi on siis myös automatisoitu, että automaattipäätös voidaan tehdä ilman, että vakuuskohdetta arvioi ihminen vaan sen sijaan vakuuskohde syötetään mallille, joka antaa vakuudelele arvon, jota sen jälkeen hyödynnetään automaattipäätöksessä. Erilaisten hankkeiden arviointi on kuitenkin manuaalista, eikä ainakaan haastateltavalla ole tiedossa, että miten niitä pystyttäisiin automatisoimaan.

Haastateltava arvioi, ettei varsinainen automatiikka ole tehnyt virheitä luottopäätöksissä, mutta automatiikan ja datan takana on kuitenkin ihminen, jolloin datan olessa vanhentunutta tai jossakin määrin virheellistä, on siinä tapauksessa mahdollista, että joissakin tilanteissa tulee myös virheellisiä päätöksiä. Juurikin siitä syystä osa päätöksistä siirretään ihmiselle, mikäli päätöksestä ollaan epävarmoja tai jos tilanne on se, että julkisista tai omista rekistereistä tuleva data on jostakin syystä vanhentunutta tai virheellistä.

Haastateltava sanoo, että tällaista automaattista luottopäätäntää on tehty jo toistakymmentä vuotta, mutta digitalisaation rooli tulee entisestään korostumaan ja yhä useampaa tuotealuetta pyritään automatisoimaan. Ihmisen työnkuva on siis varmasti muuttunut ja tulee myös tulevaisuudessa muuttumaan, kun datan lähteitä tulee yhä enemmän, tekniikka kehittyy ja tekoäly otetaan mukaan prosesseihin.

Haastateltava arvioi, että tekoälyä voitaisiin yrityksessä hyödyntää tulevaisuudessa enemmän vastauksien etsimiseen ja datan louhimiseen, kun datan laatu paranee. Yritys W:ssä hyödynnetään luonnollisen kielen prosessointia jossakin määrin kehitykseen, mikä liittyy luoton kokonaisprosessiin, mutta ei kuitenkaan automaattiseen päätökseen. Haastateltavan arvion mukaan luonnollisen kielen prosessoinnin rooli tulee kasvamaan tulevaisuudessa enemmän esimerkiksi kehittämisessä ja pitkien asiakirjojen jäsentelyssä, mikä nopeuttaa tietojen poimimista. Tekoäly ja luonnollisen kielen prosessointi tuo siis haastateltavan mukaan uusia

ulottuvuuksia, sillä dataa on niin paljon, että sen analysointiin tullaan käyttämään tulevaisuudessa entistä tehokkaampia välineitä myös siitä syystä, että datan louhiminen on nykyisin aikaisempaa halvempaa.

7.5 Valuatum Oy

Tämä on ainoa benchmarkingissa käytetty yritys, jolle ei ole tehty haastattelua tätä opinnäytetyötä tehdessä, joten liitteen 2 kysymykset eivät liity tähän yritykseen. Yrityksen nettisivuilta on löytynyt kattavasti tietoa siitä, kuinka tekoälyä hyödynnetään arvioinneissa. Valuatum Oy arvioi automaattisesti muun muassa yritysten konkurssiriskiä. Tavoitteena on arvioida, kuinka todennäköisesti yritys menee konkurssiin lähitulevaisuudessa. Ennusteet perustuvat yritysten historiallisiin tilinpäätöstietoihin sekä useisiin muihin erilaisiin indikaattoreihin, kuten esimerkiksi kannattavuuteen, velkasuhteeseen ja maksuvalmiuteen. Malleissa käytetään syöteinä historiallisista tilinpäätöksistä laskettuja suhteita, joiden avulla voidaan laskea todennäköisyys siitä, kuinka todennäköisesti yritys menee konkurssiin seuraavien vuosien aikana. Malli perustuu tilastolliseen arvioon, joten on hyvä pitää mielessä, että myös yritys, jolle on ennustettu alhaista konkurssin todennäköisyyttä lähivuosina, saattaa mennä konkurssiin, sillä yritysten tilanteet voivat muuttua nopeasti. Sen sijaan, että luku otetaan absoluuttiseksi totuudeksi, tulisi arvioitua todennäköisyyttä tarkastella suhteessa muihin samankaltaisiin yrityksiin.

Jotta tällainen malli pystyy toimimaan hyvin, on sille annettava suuri määrä esimerkkejä konkurssiin menneistä yrityksistä sekä yrityksistä, jotka eivät ole menneet konkurssiin. Tällä tavalla mallia pystytään kouluttamaan. Valuatumin järjestelmässä malli koulutetaan historiallisista tilinpäätöksistä saaduilla tiedoilla, jossa näitä tietoja käytetään syöteinä. Tuloksena on tieto siitä, menikö yritys konkurssiin vai ei. Tätä kutsutaan opetusdataksi, mikä luodaan käyttämällä useiden, jopa kymmenien satojen tuhansien yritysten tietoja. Konkurssiriskin ennustamisongelmaa kutsutaan luokitteluongelmaksi, koska tulokset ovat binaarisia indikaattoreita siitä, menikö yritys konkurssiin vai ei. Tuloksena olevia malleja kutsutaan binaariluokittimiksi.

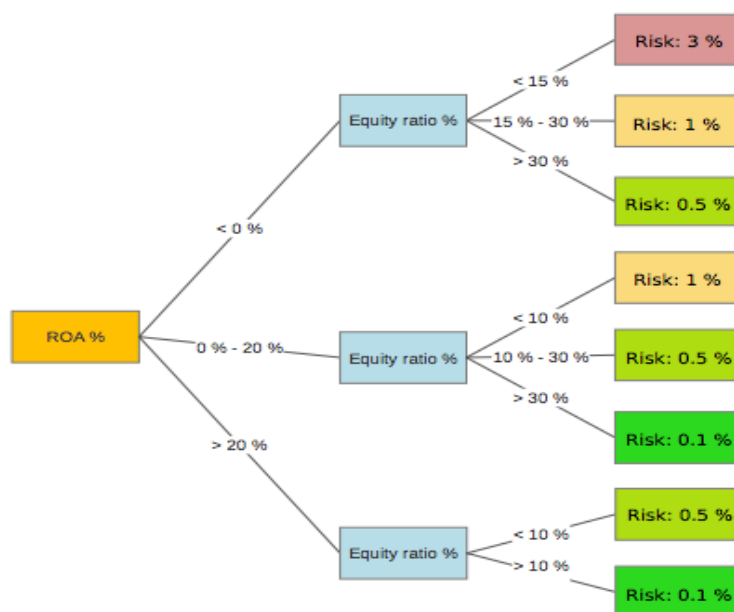
Kun mallia koulutetaan opetusdatalla, mallin parametreja muutetaan automaattisesti virhelaskujen minimoimiseksi. Virhe voidaan laskea monin eri tavoin, mutta perusajatuksena on laskea virhe siten, että virhe pienenee ennusteiden ollessa tarkempia. Mallin parametrit asetetaan siten, että arvioitu todennäköisyys on maksimoitu (mahdollisimman lähelle 100 %) niillä yrityksillä, jotka todella menivät konkurssiin ja ennustettu konkurssin todennäköisyys on lähellä 0 % yrityksille, jotka eivät menneet konkurssiin. Ei kuitenkaan ole mahdollista saavuttaa tarkalleen arvioituja todennäköisyyksiä (0 % ja 100 %), sillä kahdesta yrityksestä, joilla on samanlaiset syötearvot, toinen saattaa mennä konkurssiin ja toinen ei. Konkurssi voi johtua esimerkiksi yrityksessä tapahtuneista radikaaleista muutoksista. Tavoitteena on siis saada

keskimäärin ennustetut todennäköisyydet mahdollisimman lähelle todellisia tuloksia. Tällainen opetusprosessi suoritetaan automaattisesti tilastollisissa ohjelmistopaketeissa toteutettujen tilastollisten menetelmien avulla. Tämän jälkeen mallin suorituskykyä vahvistetaan käyttämällä erillisiä datajoukkoja, jota kutsutaan testidataksi. Nämä tiedot ovat otettu samasta lähteestä kuin opetusdata, mutta datajoukot eivät esimerkiksi sisällä samoja datapisteitä. Kyseessä on siis malli, jossa syötteenä on käytetty yrityksen tilinpäätöstä ja vasteena todennäköisyys konkurssiin menemiselle mallin opetusdatan mukaisesti.

Logistinen regressio on tilastollinen menetelmä, joka pystyy antamaan todennäköisyyksiä tapahtumille, joille on annettu numeraaliset syötteen. Siitä syystä se soveltuu hyvin konkurssiriskin arviointiin. Logistinen regressio on eniten käytetty menetelmä konkurssiriskinarviointiin. Logistisessa regressiossa, kuten myös lineaarisessa regressiossa jokainen syöte kerrotaan painokertoimella, jonka jälkeen tulokset summataan yhteen. Jotta tulosta voidaan tulkita todennäköisyytenä, saatu summa kulkee logistisen funktion läpi, joka antaa luvun välillä 0 ja 1, jolloin konkurssiriski on 0-100%. Logistinen regressio antaa kertoimen jokaiselle syöttelelle, josta voidaan nähdä kuinka paljon mikäkin syöte myötävaikuttaa yrityksen konkurssiriskiin (White box). Toisaalta logistinen regressiomenetelmä on melko rajallinen, sillä siinä oletetaan, että jokainen syöte myötävaikuttaa lineaarisesti konkurssiriskiin ja, että syötteen ovat riippumattomia toisistaan. Mallia voidaan laajentaa ottamaan nämä huomioon, mutta se tekee mallista monimutkaisemman ja vaikeammin tulkittavan. Valuatum Oy:lla on kuitenkin havaittu, että toisenlaiset mallit antavat parempia tuloksia.

Satunnaismetsämalli (random forest) rakennetaan päätöksentekopuiden (decision tree) avulla. Päätöspuu on vuokaavion kaltainen rakenne, jossa jokainen solmu koostuu yhdestä sisääntuloattribuutille tehdystä testistä. Esimerkiksi jos yrityksen kokonaispääoman tuotto prosentti (ROA) on joko positiivinen tai negatiivinen. Jokainen kaaviohaara edustaa testin

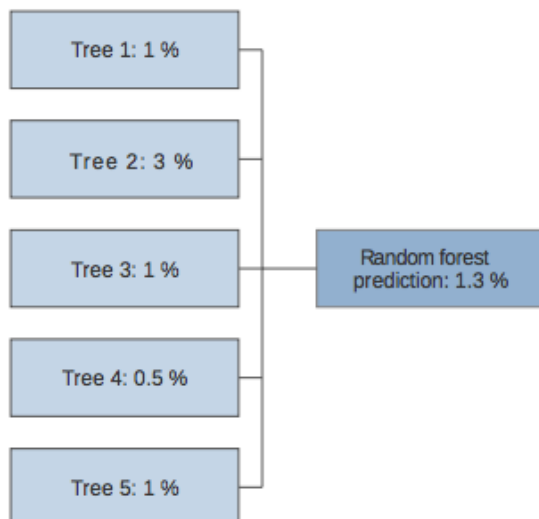
tulosta. Jokaisen haaran lopussa on päätös, joka Valuatumin tapauksessa on ennuste todennäköisyydelle, että yritys menee konkurssiin (kuvio 6).



Kuvio 9: Päätöspuu, Valuatium Oy

Päätöspuuta kouluttaessa, jokainen solmu on asetettu siten, että arvioidut konkurssit vastaavat todellisia konkurseja mahdollisimman tarkasti. Satunnaismetsä koostuu monista, jopa sadoista tai tuhansista päätöspuista, jotka ovat generoitu opetusdatan eri osilla. Tuloksena on kokoelma päätöksentekopuita, joilla on erilaiset rakenteet. Satunnaismetsämallin vaste on vain kunkin

yksittäisen päätöspuun vasteen keskiarvo ja vaste voidaan sen vuoksi tulkita todennäköisyytenä (kuvio 7).



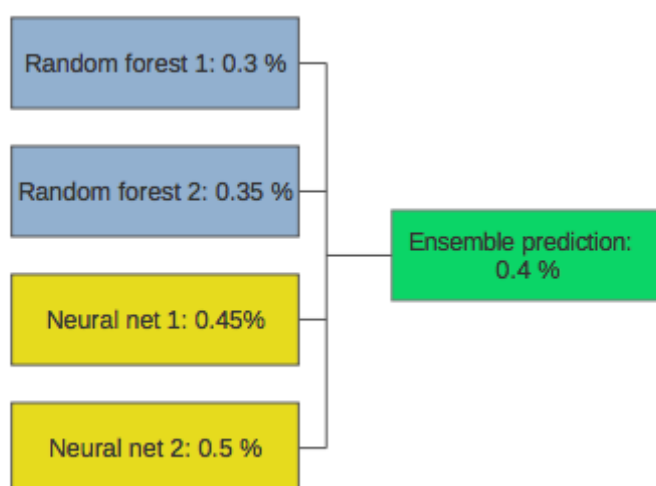
Kuvio 10: Viidestä päätöspuusta koostuva satunnaismetsä, Valuatum Oy

Satunnaismetsämalli pystyy mallintamaan monimutkaisempia suhteita datasta, mikä voi johtaa parempaan suorituskkyyn ongelma-alueen ollessa monimukainen. Satunnaismetsämalli laskee myös ominaisuuden tärkeyden jokaiselle syötemuuttujalle ja siitä syystä malli antaa tietoa siitä, mitkä muuttujat ovat parhaiten ennustettavissa. Haittapuolena on se, että mallin parametreja ei ole helppo tulkita. Lopullinen todennäköisyys voidaan laskea käyttämällä useita satoja tai tuhansia erilaisia päätöksentekopuita, mitkä tekevät erittäin vaikeaksi nähdä suoraan, kuinka eri syötemuuttujat vaikuttavat tulokseen. Tästä syystä tätä mallia voidaan kutsua mustaksi laatikoksi (black box).

Kun konkurssia ennustetaan neuroverkoilla, vastaanottaa jokainen neuroni syötteenä yhden tai useamman numeron, minkä jälkeen neuroni punnitsee nämä syötteet, summaa ne ja välittää summan epälineaarisen aktivointifunktion kautta, minkä avulla neuroverkko voi mallintaa epälineaarisia riippuvuuksia syötteiden ja vasteiden välillä. Neuroverkoissa vaste voidaan tulkita todennäköisyyksinä niin kauan, kuin aktivointifunktio on logistinen funktio. Neuroverkkoa koulutetaan säätämällä painokertoimet jokaiseen neuroniin siten, että ennustettu konkurssin todennäköisyys on mahdollisimman lähellä yrityksen todellista tilaa. Tämä tehdään yleensä vastavirta-algoritmia käyttäen. Neuroverkot voivat oppia erittäin monimutkaisia suhteita syöte- ja vastemuuttujien välillä, mutta tuloksena oleva malli on myös

monimutkainen. Siitä syystä ei ole mahdollista nähdä, kuinka erilaiset syötteet vaikuttavat saatuun tulokseen. Myös tätä mallia voidaan tästä syystä kutsua mustaksi laatikoksi (black box).

Kaikki nämä edellä mainitut menetelmät laskevat todennäköisyydet eri tavoin. Näin ollen tuloksena olevat mallit ovat erilaisia myös silloin, kun niitä koulutetaan samalla opetusdatalla. Kaikilla menetelmillä on myös omat heikkoudet ja vahvuudet, ja siksi ne tekevät usein erilaisia virheitä. Jos mustalaatikko hyväksytään, voidaan saavuttaa parhaat tulokset yhdistämällä erilaisia malleja sen sijaan, että käytössä olisi ainoastaan yksi edellä mainituista menetelmistä. On siis mahdollista ensin kouluttaa useita satunnaismetsämalleja ja useita neuroverkkomalleja, minkä jälkeen ne voidaan kerätä yhteen, jolloin malli tuottaa kaikkien mallien ennusteiden keskiarvon, kuten kuviossa 8.



Kuvio 11: Esimerkki ryhmämallista, Valuatum Oy

Tällaisen menetelmän heikkoutena on se, että todennäköisyyden laskeminen voi viedä hieman enemmän aikaa, kuin ainoastaan yhtä mallia käyttäessä, sillä kokonaisuus koostuu monista eri malleista. Myös tämä malli luokitellaan mustaksi laatikoksi, koska olemassa ei ole selkeää tapaa nähdä, kuinka tiettyhen syötteiden muutokset vaikuttavat tulokseen.

8 Johtopäätökset

Tänä päivänä tekoälyä hyödynnetään useimmiten apuälyänä ja resurssitukena. Tekoälyä ei suurissa määrin vielä käytetä täysin korvaamaan ihmisten tekemiä päätöksiä vaan antamaan ehdotuksia ja nopeuttamaan prosesseja. Tekoäly ja osaava työntekijä ovat paras yhdistelmä myös luottoprosessin kannalta - ihminen oppii tekoälyltä ja tekoäly oppii ihmiseltä. Tärkein ensiaskel tekoälyn käyttöönotossa on identifioida luottoprosessista se, missä tekoälyä

kannattaa hyödyntää ja pitää myös mielessä, että joitakin prosessin osia voidaan automatisoida ilman varsinaista tekoälyä. Tekoälyn hyödyntäminen ja tekoäly -sanana käyttö näkyy kuitenkin selkeästi myös ihmisten ja jopa ammattilaisten tulkintaeroina. Liiketoiminnan kannalta on tärkeää, että yritykset, joiden kuuluu saada lainaa, myös saavat sitä, eikä pelkästään pyrkii hylkämään lainoja, joissa todetaan suuri riski. Kun tekoälyä hyödynnetään tiedon keräämiseen sekä ennusteiden tekemiseen, voidaan saavuttaa tarkkoja lopputuloksia, joissa näkyy sekä tekoälyn että ihmisen panos.

Tämän hetkessä prosessissa toimeksiantajayritys ei hyödynnä kaikkea käytettävissä olevaa strukturoimatonta dataa, jota tekoälyn avulla voidaan hyödyntää luottoprosessissa, jotta ennusteista saadaan aiempaa tarkempia. Tiedon keräämiseen tarvitaan siis aiempaa enemmän automaatiota päätöksiä tueksi ja prosessien nopeuttamiseksi, mihin on löytynyt ratkaisuja tätä opinnäytetyötä tehdessä. Luonnollisen kielen käsittely on noussut yhdeksi merkittäväksi tekijäksi, mitä hyödynnetään tyypillisesti rahoitusallalla tiedon keräämiseen internetistä sekä asiakirjojen tai erilaisien sopimuksien lukemiseen ja riskien löytämiseen. Tämän hetkessä prosessissa analyttikko käy erilaisia sopimuksia sekä lainarahan käyttötarkoitukseen liittyviä muita materiaaleja manuaalisesti läpi. Apuna prosessissa voidaan käyttää luonnollisen kielen käsittelyä löytämään materiaaleista riskejä nopeasti, kunhan järjestelmälle on ensin opetettu riskit. Luonnollisen kielen käsittelyn avulla voidaan saada usein hyödyntämätön strukturoimaton data hyödynnetyksi, jos sitä hyödynnetään keräämään erilaista tietoa internetistä, jolloin voidaan selvittää esimerkiksi yrityksen todellinen toimiala, kasvunäkymiä tai muita internetistä löytyviä luottopäätöksen kannalta merkittäviä tekijöitä. Todellinen toimiala on merkittävä tekijä myös toimeksiantajayrityksen arviointiprosessissa, sillä tilinpäätösanalyysit perustuvat osittain myös toimialaan, jolloin tilinpäätöstietoja voidaan myös verrata oikeaan toimialaan, eikä välttämättä siihen, minkä yritys on itsestään ilmoittanut. Tällaisia palveluita on myös jo saatavilla maksullisena, joten yhteistyö tällaisen palveluntarjoajan kanssa voi tuoda huomattavaa hyötyä toimeksiantajayritykselle.

PSD2 avaa mahdollisuuden selvittää arvioinnissa olevien yritysten tilitapahtumia pidemmältä ajalta, kuin nykyisessä prosessissa, jossa arvioidaan 3 kuukauden tilitapahtumat. Luonnollisen kielen käsittelyä voidaan hyödyntää etsimään tilitapahtumista esimerkiksi riskejä ilman, että analyttikon tarvitsee lukea tilitapahtumat rivi kerrallaan. Lisäksi PSD2 ansiosta tilitapahtumista ei voida poistaa asiakkaan toimesta mitään tietoa, mikä taas on mahdollista kun käytössä on eri muodoissa lähetetyt tiedostot, kuten esimerkiksi PDF-tiedostot. Koska luonnollisen kielen käsittelyllä voidaan etsiä erilaisista asiakirjoista tärkeimmät kohdat ja niissä olevat riskit, voidaan kokeneempien työntekijöiden asiantuntemusta jakaa myös vähemmän kokeneille työntekijöille, sillä tällaiset järjestelmät ovat koulutettu kokeneiden asiantuntijoiden avulla, jolloin järjestelmät pystyvät neuvomaan myös vähemmän kokeneita työntekijöitä löytämään riskejä ja kertomaan esimerkiksi sen, mitä lisätietoja yrityksiltä tarvitaan, jotta päätös voidaan tehdä. Arvioiden mukaan luonnollisen kielen käsittely tulee näkymään tulevaisuudessa entistä

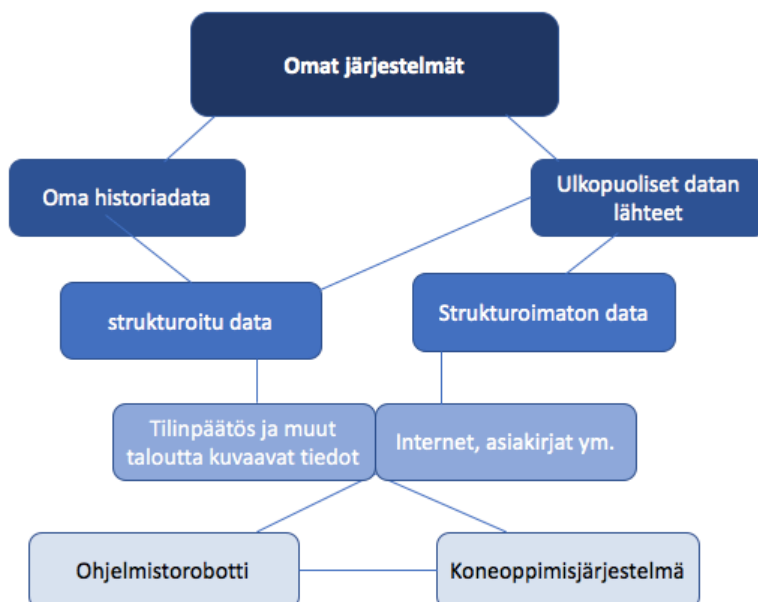
enemmän, sillä strukturoimaton data on tällä hetkellä usein käyttämätöntä, vaikka ylivoimaisesti suurin osa datasta on strukturoimatonta.

Selkeästi hylättävissä olevat lainahakemukset voidaan automatisoida myös ilman varsinaista tekoälyä. Yksinkertaisien päätösten tekemiseen voidaan käyttää esimerkiksi ohjelmistorobottia, jolle on opetettu minkälaiset hakemukset ovat ehdottomasti hylättävä, mutta myös tällöin järjestelmä tekee ainoastaan ehdotuksen, jonka jälkeen ihminen tekee vielä lopullisen päätöksen. Tällöin työntekijöiden aikaa ei kulu selkeiden tapausten tutkimiseen, vaan he voivat keskittyä monimutkaisempien hakemusten käsittelyyn ja arviointiin. Ohjelmistorobotti ja koneoppimisjärjestelmä voidaan myös yhdistää, ja sillä on useiden lähteiden mukaan suuri merkitys toimivan tekoälyn kannalta. Jos ohjelmistorobotti on aluksi tehnyt yksinkertaisia päätöksiä, voidaan koneoppimisjärjestelmän avulla myös laajentaa ohjelmistorobotin osaamistaan tekemään monimutkaisempia päätöksiä. Kun ohjelmistorobottiin yhdistetään koneoppimismalli, voidaan tehdä johtopäätöksiä sekä strukturoidusta että strukturoimattomasta datasta. Sekä strukturoidun että strukturoimattoman datan hyödyntämisellä on merkitystä tarkkojen ennusteiden lopputulokseen.

Kiinteistövakuuksien arvioinnissa voidaan myös hyödyntää koneoppimisjärjestelmää, jossa hyödyksi voidaan käyttää erilaisia datan lähteitä kiinteistöjen hinnan ja myyntiajan arviointiin. Välttämätöntä ei ole rakentaa mallia itse, vaan tarkistaa myös voiko prosessia optimoida sellaisen yhteistyökumppanin kautta, jolla on tällainen malli valmiina. Vaikka toimeksiantajayrityksen liiketoiminta ei perustu kiinteistöihin tai asuntoihin, minkä takia siihen liittyvä dataa ei välttämättä löydy riittävästi yritykseltä entuudestaan, voidaan arvioinnin kannalta oleellista dataa hankkia avoimista tai maksullisista datalähteistä, joita on mm. tämän opinnäytetyön kappalleessa 6.2. lueteltu. Datan pohjalta voidaan rakentaa koneoppimisjärjestelmä, joka arvioi myös vakuuksia.

Koneoppimismalli voidaan rakentaa yrityksen oman datan pohjalta, mutta vaikka omaa dataa ei olisi riittävästi käytössä, voidaan koneoppimisen malli kuitenkin rakentaa ja hyödyntää avoimia lähteitä tai maksullisia lähteitä, joita finanssisektorilla on paljonkin. Markkinoilla on useita maksullisia toimijoita, joiden kautta on mahdollista saada jalostettuna dataa hyödynnettäväksi omaan prosessiin. Oman datan pohjalta voidaan tarkistaa, onko esimerkiksi luottotappiolainoissa usein joitakin yhdenmukaisuuksia tai vastaavasti, onko ajallaan maksetuissa lainoissa joitakin yhdenmukaisuuksia. Useamman vuoden tilinpäätöstietojen lisäksi tilikauden jälkeisen ajan kirjanpitoajoilla voidaan selvittää yrityksen tuoreimmat luvut ja kassavirrat, jotka myös voidaan ottaa osaksi koneoppimisjärjestelmää, joka historiadatan perusteella pystyy selvittämään minkälainen merkitys lopputuloksen kannalta on ollut tuoreilla luvuilla ja kassavirroilla suhteessa tilinpäätökseen. Koska ohjattu oppiminen on luottopäätöksissä usein käytetty tekniikka, jossa datanäytteet on luokiteltu muutamiiin eri

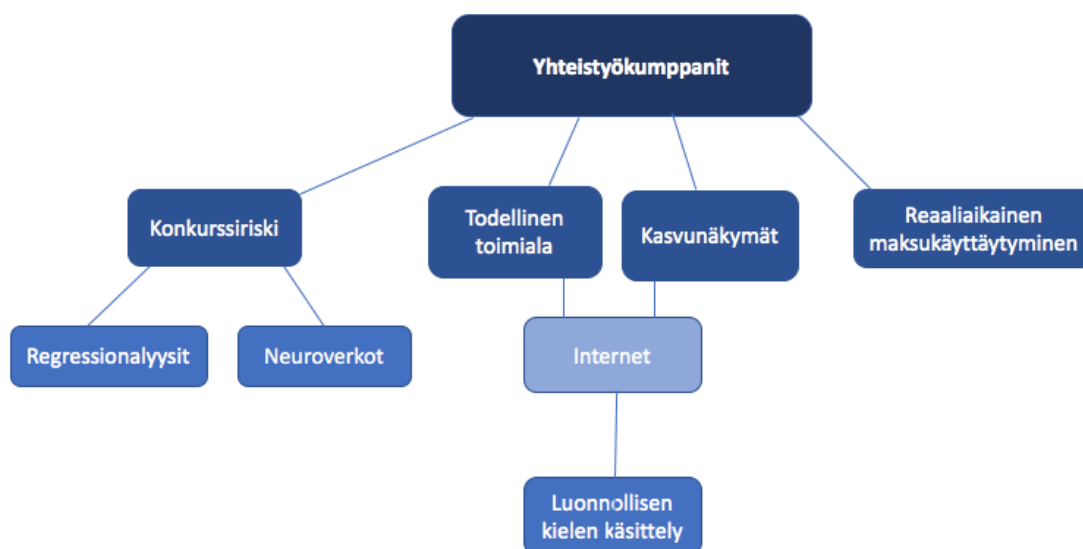
luokkiin, kuten esimerkiksi luottotappioksi kirjatut lainat voivat olla yksi oma luokkansa, ajallaan takaisinmaksetut lainat oma luokkansa ja myöhässä takaisin maksetut oma luokkansa, joiden historiadata perustella järjestelmä pystyy löytämään yhdenmukaisuuksia esimerkiksi kirjanpitoajoista ja tilitapahtumista, jotka ovat osana järjestelmän tekemää ehdostusta tai päätöstä. Koska toimeksiantajayrityksen prosessissa on tärkeää, että eri tekijät tukevat toisiaan, on koneoppimisjärjestelmästä juuri myös siksi hyötyä, sillä sen avulla voidaan selvittää mitkä tekijät tukevat mitään tekijöitä ja löytää sen pohjalta syy-seuraussuhteita. Takaisinmaksukykyä arvioitaessa toimeksiantajayrityksessä tutkitaan yrityksen kulurakennetta, josta voi selvittää esimerkiksi liiketoiminnan olevan kannattavaa, mutta rahoituskulujen ja pääomarakenteen huonoja. Koneoppimisjärjestelmä pystyy historiadata perusteella löytämään myös näiden tekijöiden vaikutuksia ja suhteita lopputulokseen. Kuviossa 12 on kuvattu, miten yritys pystyy hyödyntämään tekoälyä, jos yritys rakentaa oman tekoälyjärjestelmän tuomaan lisäarvoa prosessiin.



Kuvio 12: Omat järjestelmät

Markkinoilla on myös yrityksiä, jotka jo hyödyntävät tekoälyä arvioimaan yrityksen konkurssia tai maksukäyttäytymistä erilaisilla menetelmillä, joissa hyödynnetään historiadataa ja vertaillaan yrityksen kehitystä aiempiin vuosiin sekä saman toimialan yrityksiin, joten yhteistyötä tällaisten yritysten kanssa tulisi harkita. On kuitenkin myös tärkeä ymmärtää, minkälaisiin menetelmiin perustuu yhteistyökumppaniyrityksen arviot ja analyysit yrityksistä. Opinnäytetyötä tehdessä on selvinnyt, että luottopäätöksissä, joissa pääsääntöisesti arvioidaan konkurssiriskiä, käytetään ainakin toistaiseksi suurimmaksi osaksi perinteisempiä regressiomalleja, kuten logistista regressiomallia niiden läpinäkyvyyden takia, jolloin voidaan

jälkikäteen selvittää, mitkä tekijät vaikuttivat koneen tekemään päätökseen tai ehdotukseen. Neuroverkkoihin perustuvat mallit, joissa eri muuttujat neuroverkon eri kerroksissa muuttuvat jatkuvasti, voivat olla tarkempia, mutta niiden tekemiä päätöksiä tai ehdotuksia voi olla vaikea ihmisen perustella. Siitä huolimatta neuroverkkojen käyttö tulee todennäköisesti yleistymään tulevaisuudessa. Jos ”black box” hyväksytään yrityksen prosesseissa, voi neuroverkkojen hyödyntäminen luottojen arvioinneissa olla tarkempi, kuin esimerkiksi logistinen regressio. On siis yritys kohtaista, kuinka tärkeänä pidetään sitä, että syyt järjestelmien tekemille päätöksille tai ehdotuksille tunnetaan. Konkurssiriskiä arvioivia palveluntarjoajia on jo olemassa, jotka tiedettävästi hyödyntävätä tekoälyä ja sen tuomaa lisäarvoa, eli ei välttämättä ole kannattavaa rakentaa samaan dataan pohjautuvaa järjestelmää itse vaan selvittää yhteistyömahdollisuuksia. Optimaalinen hyöty ja tarkempi ennuste voidaan kuitenkin saavuttaa hyödyntämällä näitä valmiita tekoälyä hyödyntäviä palveluita esimerkiksi konkurssiriskin ja maksukäyttäytymisen arviointiin, mutta sen lisäksi yrityksellä voi olla käytössä oman datan pohjalta rakennettu koneoppimisjärjestelmä, jonka avulla pystytään selvittämään löytyykö esimerkiksi luottotappioiksi kirjatuista lainoista yhtäläisyyksiä tai vastaavasti, onko ajallaan maksetuilla ja parhaiten tuottaneilla lainoilla yhtäläisyyksiä. Tällöin tekoälyn antama hyöty ja erilaiset datan lähteet tulevat arvioinnissa hyödynnetyiksi. Myös luonnollisen kielen käsittelyä voidaan hyödyntää yhteistyökumppaneiden avulla, kuten tämän luvun toisessa kappaleessa on kuvattu. Yhteistyökumppaneiden hyödyntäminen voi olla kannattavaa myös siksi, että sen aloittaminen on vaivatonta. Kuviossa 13 on kuvattu yhteistyökumppaneiden tai erilaisten tekoälyä hyödyntävien palveluiden käyttömahdollisuuksia.



Kuvio 13: Yhteistyökumppanit

Lähteet

Painetut:

Merilehto A. 2018. Tekoäly matkaopas johtajille. Balto print, Liettua.

Moilanen J, Niinioja M, Seppänen M, Honkanen M. 2018. API talous 101. Balto print, Liettua.

Niskanen J, Niskanen M. 2010. Yritysrahoitus. Edita Prima Oy, Helsinki.

Ijäs S. 2002. Luottoriskien hallinta tuloksen tekijänä. Gummerus kirjapaino OY, Jyväskylä.

Ojasalo K, Moilanen T, Ritalahti J. 2015. Kehittämistyön menetelmät, uudenlaista osaamista liiketoimintaan. Sanoma Pro Oy.

Alhonsuo S, Nisén A, Nousiainen S, Pellikka T, Sundberg S. 2012. Finanssitoiminnan käsikirja. Bookwell Oy, Jyväskylä.

Sähköiset:

Hurwitz, Nugent, Halper, Kaufman. 2013. Big Data for Dummies (e-kirja). <https://proquest.safaribooksonline.com/book/databases/business-intelligence/9781118644171>

Teknologian tutkimuskeskus VTT, Tuottoa ja tehokkuutta Suomeen tekoälyllä, 2017. Viitattu 9/2019. Saatavilla <https://www.vtt.fi/inf/pdf/policybrief/2017/PB1-2017.pdf>

Microsoft, kuinka Suomi on valmistautunut tekoälyn tuloon, 2019. Viitattu 9/2019. saatavilla <https://pulse.microsoft.com/fi-fi/business-leadership-fi-fi/na/fa2-tutkimus-uncovering-ai-in-finland/>

CGI, Hyödynnä tekoälyä liiketoiminnassa, 2019. Viitattu 9/2019. saatavilla <https://www.cgi.fi/fi/mita-on-tekoaly>

Tekoäly.info, vahva- ja heikkotekoäly. Viitattu 9/2019. Saatavilla https://xn--tekoaly-eua.info/mita_tekoaly_on/

Finanssialalle, Tekoäly. Viitattu 9/2019. Saatavilla <http://www.finanssialalle.fi/opintomateriaalit/finanssialan-perusteet/innovaatiot/tekoaly.html>

Elements of AI verkkokurssi. Viitattu 10/2019. Saatavilla <https://www.elementsofai.com/fi>

Datatiede, mikä datatiede. Viitattu 9/2019. Saatavilla <https://www.datatiede.fi/mika-datatie/>

e-Craft Business Insight, Opas tekoälyn hyötykäyttöön. Viitattu 9/2019. Ladattavissa https://businessinsight.ecraft.com/tekoalyopas_2019?utm_campaign=Opas%3A%20Teko%3%A4lyopas%202019&utm_source=web

CGI, Tekoälyn ostajan opas. 2019. Viitattu 9/2019. Ladattavissa <https://www.cgi.fi/fi/laataa/tekoaly-ostajan-opas>

Solveto, Analytiikka ja big data. 2019. Viitattu 9/2019. Saatavilla <https://www.sovelto.fi/ratkaisut/ict-ja-uudet-teknologiat/big-data/>

CGI, Erinomaista liiketoimintaa big datan avulla. 2013. Viitattu 9/2019. Saatavilla https://www.cgi.fi/sites/default/files/files_fi/white-papers/white_paper_erinomaista_liiketoimintaa_big_datan_avulla.pdf

Finanssiala Ry, Finanssialalle, FinTech. Viitattu 9/2019. Saatavilla <http://www.finanssialalle.fi/opintomateriaalit/tulevaisuuden-finanssiala/tulevaisuuden-pankki/fintech.html>

Finanssiala Ry, Finanssialalle, API-ohjelmistorajapinta. Viitattu 9/2019. Saatavilla <http://www.finanssialalle.fi/opintomateriaalit/tulevaisuuden-finanssiala/tulevaisuuden-pankki/api-ohjelmistorajapinta.html>

Salkunrakentaja, Finanssiteknologia mullistaa rahoitusmarkkinat. 2016. Viitattu 9/2019. Saatavilla <https://www.salkunrakentaja.fi/2016/11/finanssiteknologia-mullistaa-rahoitusmarkkinat/>

Finanssivalvonta, PSD2. 2019. Viitattu 9/2019. Saatavilla <https://www.finanssivalvonta.fi/saantely/saantelykokonaisuudet/psd2/>.

Finanssivalvonta, FinTech ja sääntely: miten maksupalveludirektiivi muuttaa maailmaa. 2019. Viitattu 9/2019. Saatavilla <https://www.slideshare.net/Finanssivalvonta/slush-side-eventsntely301116>

Finanssiala Ry, Finanssialalle, Kysymyksiä ja vastauksia toisesta maksupalveludirektiivistä PSD2. 2019. Viitattu 9/2019. Saatavilla <http://www.finanssiala.fi/uutismajakka/Sivut/QA-Toinen-maksupalveludirektiivi.aspx>

Ethics Guidelines for Trustworthy AI, Independent high-level expert group on artificial intelligence set up by the European commission 2018. Viitattu 10/2019. Saatavilla: <https://ec.europa.eu/futurium/en/ai-alliance-consultation>

Felt E, Talouselämä - rahavirrat on raaka mittari. 2008. Viitattu 10/2019. Saatavilla <https://www.talouselama.fi/uutiset/rahavirta-on-raaka-mittari/bb760ab6-3d52-35ea-8ec7-b707e29e2784>

Hulkko V, Most value from AI with human-in-the-loop solutions 2018. Viitattu 10/2019. Saatavilla <https://silo.ai/most-value-human-in-the-loop-ai/>

Turunen J, Fintech Industry everywhere. 2017. Viitattu 10/2019. Saatavilla <https://www.youtube.com/watch?v=hHvVqGv8Ahc&list=PL1BvsZKUJkdrCmPeA9-gCxp8z5z0PzdBi&index=6>

Helsinki Fintech, Viisi Digital finance- ja fintech-trendiä Suomessa 2019. 2019. Viitattu 10/2019. Saatavilla <https://www.helsinkifintech.fi/news/viisi-digital-finance-ja-fintech-trendia-suomessa-2019/>

KASKI17 - Risto Siilasmaa: Kotitehtävä: Opi ymmärtämään tekoälyä vuonna 2017. Viitattu 10/2019. Saatavilla https://www.youtube.com/watch?v=WO1wND-_0Uc

Crnoja & Col, Infusing RPA With Machine Learning. 2019. Viitattu 10/2019. Saatavilla: <https://www.digitalistmag.com/cio-knowledge/2019/07/22/infusing-rpa-with-machine-learning-powerful-medley-part-5-06199588>

Davenport, Combining Robotic Process Automation and Machine learning. 2019. Viitattu 10/2019. Saatavilla: <https://www.sas.com/content/dam/SAS/documents/marketing-whitepapers-ebooks/third-party-whitepapers/en/combining-robotic-process-automation-machine-learning-110369.pdf>

Ojanperä, Value of Fintech data. 2017. Viitattu 10/2019. Saatavilla: <https://www.youtube.com/watch?v=3rVP5Fhy88w&list=PL1BvsZKUJkdrCmPeA9-gCxp8z5z0PzdBi&index=8>

Garbade, Simple introduction to Natural Language processing. 2018. Viitattu 10/2019. Saatavilla: <https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32>

Alanen, NLP & Textual data: how text clustering provides business insights. Viitattu 10/2019. Saatavilla: <https://silo.ai/text-clustering-for-business-insights/>

Aaltonen, Mitä NLP tarkoittaa? 2019. Viitattu 10/2019. Saatavilla: <https://home.kpmg/fi/fi/home/Pinnalla/2019/09/digiajan-dynaaminen-tilintarkastus-nlp.html>

Leikas, Tekoälyn etiikka - mistä ylipäättään puhutaan. 2018. Viitattu 10/2019. Saatavilla: <https://vttblog.com/2018/12/11/tekoalyn-etiikka-mista-ylipaataan-puhutaan/>

Valuatum Oy, Automated estimation of bankruptcy risk. Viitattu 11/2019. Saatavilla: https://s3-eu-west-1.amazonaws.com/valu-wordpress/wp-content/uploads/sites/10/2019/05/27123643/Estimation_of_bankruptcy_risk_20180614.pdf

Kuviot

Kuvio 1: Elements Of AI, Lineraarinen regressio kuvina	9
Kuvio 2: Elements Of AI, Logistinen regressio	10
Kuvio 3: Syöte ja vaste neurovekoissa, Merilehto 2018, 51	12
Kuvio 4: Neuroverkko, Merilehto 2018, 52	13
Kuvio 5: Konvoluutioneuroverkko, Merilehto 2018, 53.....	13
Kuvio 6: Datatieteen työkalupakki, Kristiina ja Anniina Seppä	16
Kuvio 7: API-Ohjelmistorajapinta, Finanssiala Ry.....	23
Kuvio 8: Human in the loop AI, Silo AI 2018.	39
Kuvio 9: Päättöspuu, Valuatum Oy	48
Kuvio 10: Viidestä päätöspuusta koostuva satunnaismetsä, Valuatum Oy	49
Kuvio 11: Esimerkki ryhmämallista, Valuatum Oy	50
Kuvio 12: Omat järjestelmät.....	53
Kuvio 13: Yhteistyökumppanit	54

Liite 1: Haastattelurunko (Asiantuntijat)

Teema 1: Yleistä

1. Mikä on sinun taustasi?
2. Kuinka kauan olet työskennellyt tai toiminut aiheen parissa?
3. Millä tavalla olet työskennellyt aiheen parissa?

Teema 2: Tekoälyn hyödyntäminen yritysluottoprosesseissa

1. Miten tekoälyä ja koneoppimista voidaan hyödyntää luottoprosessien optimointiin? Eli miten luottoprosesseista saadaan mahdollisimman tehokas tekoälyn avulla?
2. Miten koneoppimista ja tekoälyä voidaan hyödyntää luottoriskien arvioimisessa nimenomaan yrityslainoissa? riski koostuu usein hankkeesta, yrityksen luvuista (myös tuoreista), vastuuhenkilöistä ja vakuuksista sekä näiden kaikkien luomasta kokonaisuudesta.
3. Miten tekoäly on päätenyt tiettyihin päätöksiin?
4. Mikä on jo vanhaa tietoa tekoälyn käytössä ja mikä on huomisen käyttöä?
5. Mitä virheitä tekoäly on tehnyt luottoprosesseissa? Onko tekoäly tehnyt virheitä, jonka ihminen olisi helposti löytänyt? Jos niin minkälaisissa tilanteissa?
6. Mistä ja miten data hankitaan? Mistä saada aina uusin data? Käyttävätkö yritykset omaa dataa vai yhteistyökumppanin? Miten tiedonhankintaa voidaan parantaa?

Teema 3: Ensimmäiset askeleet tekoälyn käyttöönotossa

1. Miten paljon on käytännön esimerkkejä olemassa tekoälystä yritysluottojen arvioinnissa ja miten paljon on vasta suunnitteilla?
2. Miten pankit tai muut FinTech yritykset pääsivät alkuun tekoälyn käytössä? Mitä tulisi ottaa huomioon?
3. Tuleeko mieleen jotakin muuta aiheeseen liittyvää?

Liite 2: Haastattelurunko (Benchmarking)

Teema 1: Yleistä

1. Mikä on sinun taustasi?
2. Kuinka kauan olet työskennellyt tai toiminut aiheen parissa?
3. Millä tavalla olet työskennellyt aiheen parissa?

Teema 2: Tekoälyn hyödyntäminen

1. Hyödyntääkö yritys jollakin tavalla tekoälyä yritysarvioinneissa? Jos niin miten? Mitä kaikkea tällöin arvioidaan ja miten?
2. Mistä hankitte kaiken datan?
3. Hyödynnättekö tekoälyä vakuuksien tai yrityksiä vastuuhenkilöiden arvioinnissa, miten?
4. Hyödynnetäänkö tekoälyä joidenkin hankkeiden arvioinnissa? Luotonanto perustuu usein erilaisiin hankkeisiin eli miten tekoälyä voidaan hyödyntää hankkeiden arvioinnissa?
5. Mikä on mielestänne jo vanhentunutta tekoälyä ja mikä vasta tulevaa?

Teema 3: Muuta

1. Miten pääsitte alkuun?
2. Tiedättekö kuinka paljon on käytännön esimerkkejä olemassa tekoälystä yritysluottojen arvioinnissa ja miten paljon on vasta suunnitteilla?
3. Tuleeko mieleen jotakin muuta aiheeseen liittyvää?