



Markus Helander

Vuokrausliiketoiminnan liikevaihdon arvioinnin tarkkuuden kehittäminen hyödyntäen (Big) data analyysia

Metropolia Ammattikorkeakoulu

Tradenomi

Liiketalouden tutkinto-ohjelma

Opinnäytetyö

Toukokuu 2022

Tiivistelmä

Tekijä(t):	Markus Helander
Otsikko:	Vuokrausliiketoiminnan liikevaihdon arvioinnin tarkkuuden kehittäminen hyödyntäen (Big) data analyysia
Sivumäärä:	44 sivua
Aika:	Toukokuu 2022
Tutkinto:	Tradenomi
Tutkinto-ohjelma:	Liiketalous
Suuntautumisvaihtoehto:	Laskentatoimi ja rahoitus
Ohjaaja(t):	Lehtori Leena Aalto

Tämän opinnäytetyön tarkoituksena oli kuvata, kuinka Ramirent-konsernin vuokrausliiketoiminnan liikevaihdon arvioinnin tarkkuutta voidaan kehittää hyödyntäen (Big) data analyysia. Ramirentillä on toiminnassa päivittäisen vuokrausliiketoiminnan raportointi toimipisteittäin ja siinä on huomattu selviä haasteita. Laskuja ja tavaran palautuksia ei täsmällisesti tiedetä. Epätietoisuus asiakkaalle annettavista alennuksista on olemassa ja sen takia arviot tehdään pelkästään sopimusten pohjalta.

Kehittämishanke koostui historian tutkimisesta, jonka jälkeen käytettiin erilaisia algoritmeja ja kertoimia, joilla parannettiin tarkkuutta arvioissa luotettavien tuloksien mahdollistamiseksi. Tutkimus keskittyi pääosin Ramirent-konsernin Ruotsin liikevaihtoon ja vuokratuottoihin.

Tutkimuksen lisäksi keskityttiin data-analyysiin ja siitä saataviin hyötyihin vuokrausliiketoiminnan kehittämisessä. Tarkastelun apuna hyödynnettiin materiaalia Ramirentiltä ja aiheeseen liittyvää kirjallisuutta. Opinnäytetyössä kehitettiin tietynlainen ennustemalli vuokrausliiketoiminnan arvioinnin tarkkuuden kehittämiseksi, jota pyritään hyödyntämään lähitulevaisuudessa mahdollisimman pian.

Opinnäytetyön tutkimuksen perusteella ennustemalli on tärkeä, luotettava ja kannattava osa pitkäjänteistä liiketoimintaa. Aiheen tutkimisessa hyödynnettiin erilaisia tietotekniikan sovelluksia, tärkeimpänä Excel, joka mahdollisti luotettavan ennustemallin luomisen.

Avainsanat: data-analyysi, liikevaihto, ennustemalli, vuokrausliiketoiminta, arvioinnin tarkkuus

Abstract

Author(s): Markus Helander
Title: Development of the accuracy of the rental business revenue estimation using (Big) data analysis
Number of Pages: 44 pages
Date: May 2022
Degree: Bachelor of Business Administration
Degree Programme: Economics and Business Administration
Specialisation option: Accounting and Finance
Instructor(s): Leena Aalto, Lecturer

The purpose of this thesis was to describe how the accuracy of estimating the revenue of Ramirent Group's rental business can be developed using (Big) data analysis. Ramirent operates a daily reporting system in its rental business and has identified clear challenges. Invoices and returns of goods are not known exactly. There is uncertainty about the discounts to be given to the customer and that is why the estimates are made on the basis of contracts alone.

The development project consists of a study of history, after which various algorithms and coefficients are used to improve the accuracy of the estimates to enable reliable results. The study mainly focuses on the key figures of Ramirent Group's net sales and rental income in Sweden.

In addition to the research, the focus was on data analysis and its benefits in developing the rental business. Material from Ramirent and related literature was used to assist in the review. In the thesis, a certain type of forecasting model was developed for the development of the accuracy of the rental business evaluation, which aims to be utilized as soon as possible in the near future.

Based on the research of the thesis, the forecast model is an important, reliable, and profitable part of long-term business. Various IT applications were used to study the topic, most importantly Excel, enabling the creation of a reliable forecasting model.

Keywords: data analysis, revenue, forecast model, rental business, valuation accuracy

Sisällys

1	Johdanto	1
1.1	Aihe ja tutkimusongelma	1
1.2	Työn rajaaminen	2
1.3	Tutkimusmenetelmät	3
1.4	Tutkimusaineisto	3
2	Big data ja data-analyysi	4
2.1	Tarkoitus ja käyttömenetelmät	5
2.2	Big datan vaikutus markkinoihin	6
2.3	Big datan potentiaaliset riskit ja haasteet	7
2.4	Big datan hyödyntäminen vuokrausliiketoiminnassa	8
3	Data-analyysiin vaikuttavat tekijät	9
3.1	Liikevaihto	9
3.2	Liikevaihdon tunnusluvut	10
3.3	Kausiluontoiset tekijät	10
3.4	Sovellukset ja toiminnot	11
4	Ennustemallin käyttöönotto	11
4.1	Ennustuksen vaiheet	12
4.2	Kartoittaminen	14
4.3	Mallin rakennusosat	15
4.4	Ennustemallin käyttömahdollisuudet ja haasteet	17
4.5	Ennustemallin suoriutuminen	18
5	Tutkimus	18
5.1	Kehityssuunta ja analyysimallit	19
5.2	Ongelmat ja niihin reagointi	22
5.3	Ruotsin toimipisteet	23
5.4	Datan liikkuminen ja käyttäytyminen	24
5.5	Toteutus ja mahdolliset toimenpiteet	26
5.6	Muut tutkimukset ja ennustemallien vertailu	28
6	Tulokset	31

7	Johtopäätökset ja opinnäytetyöprosessi	34
7.1	Ennustemallien käyttö nykypäivänä	36
7.2	Päätelmät ennustemallin heikkouksista ja hyödyistä	37
8	Yhteenveto	39
8.1	Data-analytiikan reagoiminen teknologian jatkuvaan kasvuun	39
8.2	Ennusteet datan luotettavuuden tueksi	40
8.3	Validiteetti ja reliabiliteetti	41
	Lähteet	42

1 Johdanto

Nykypäivän teknologian kehittyessä yrityksillä on entistä enemmän dataa käytössään. Merkittävän informaation ja datan hyödyntäminen on muuttunut käyttäjäystävällisemmäksi yhdessä data-analytiikan jatkuvan kehityksen kanssa. Tällä hetkellä monet yritykset pystyvät jo hyödyntämään dataa tehokkaasti eri käyttötarkoituksiin organisaation sisä- ja ulkopuolella. (TechTarget 2021.)

Suurta määrää dataa hyödynnetään tälläkin hetkellä ympäri maailmaa eri aloilla. Kilpailu on kovaa ja yritysjohtajat arvioivat jatkuvasti yrityksensä informaation laatua, ja sitä, onko kilpailijoilla luotettavampaa ja ajankohtaisempaa tietoa. Tätä olen nähnyt myös henkilökohtaisen työurani varrella, ja tämä sai kiinnostukseni heräämään tutkimaan enemmän (Big) data-analytiikan hyötyjä vuokrausliiketoiminnan näkökulmasta.

Monet maailman yleisemmistä hakukoneista kuten Google, Yahoo tai Bing keräävät jatkuvasti tietoa asiakkaistaan erilaisten evästeiden kautta. Näin he voivat kohdentaa markkinointia näille henkilöille tiettyjen hakutulosten perusteella. Samalla tavalla yritykset analysoivat omaa dataansa saadakseen paremman käsityksen asiakkaasta ja omasta kannattavuudestaan pitkällä tähtäimellä hyödyntäen erilaisia tähän tehtävään soveltuvia tietotekniikan sovelluksia. (Analytics Insight 2021.)

1.1 Aihe ja tutkimusongelma

Tämän opinnäytetyön tavoitteena on kuvata ratkaisuja vuokrausliiketoiminnan liikevaihdon arvioinnin tarkkuuden kehittämiseen data-analytiikkaa hyödyntäen. Ilmiö havainnollistetaan toimeksiantajayrityksen Ramirent-konsernin kanssa yhteistyössä tehdyn ennustemallin avulla useaan otteeseen. Ramirent-konserni yrityksenä on yksi Euroopan johtavista konevuokrauskonserneista. Opinnäytetyö pyrkii selvittämään hyötyjä datan luotettavuudessa ja tehokkaassa käytössä

ennustemallin avulla. Lopputulemana opinnäytetyö pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

-Mitä sovelluksia voidaan hyödyntää päivittäisen datan analysoimiseksi?

-Mistä johtuu heitto ja epätarkkuus liikevaihdon arvioinnissa?

-Miten implementoida saatu data uusiin malleihin/kuvioihin, joita hyödynnetään tulevaisuudessa?

-Millainen on ennusteprosessi, kun verrataan myynnin alkuhetkeä lopulliseen tilanteeseen?

1.2 Työn rajaaminen

Opinnäytetyö rajataan Ramirent-konsernin Ruotsin liikevaihdon arvioinnin tarkkuuden kehittämiseen kuitenkin kansainvälisellä yhteistyöllä. Ruotsin liikevaihtoon kuuluu laitteiden vuokraus ja myynti ulkoisille asiakkaille. Rajaus on suoritettu Ruotsin vuokrausliiketoiminnan ympärille suurien heittojen takia kuukausittaisissa liikevaihdon luvuissa. Data-analyyseissa käytetään omaa ennustemallia ja tarvittaessa hyödynnetään myös ulkoisia malleja ja algoritmeja.

Data-analyysiä esitellään keskittyen liiketoimintaan ja sen kehittämiseen. Ennustemallista käydään läpi myös sen haitat, hyödyt ja käyttötarkoitus lähitulevaisuudessa yrityksen sisällä. Ennustemallin koodia ei käydä tarkemmin läpi vaan tarkastellaan sen tuomia mahdollisuuksia ja tehokkuutta helpottavia käyttömenetelmiä. Opinnäytetyön painopiste tulee olemaan Ruotsin vuokrausliiketoiminnassa, jonka osalta perehdytään data-analyyseihin, tehokkuuteen, luotettavuuteen ja toiminnan kehittämiseen koko konsernissa.

Tämän opinnäytetyön rakenne koostuu johdannosta, viitekehyksestä ja empiirisestä tutkimuksesta. Nämä kaikki osiot tiivistetään ja vedetään yhteen päätännössä. Pohjustava tietopohja käsitellään johdannossa ja viitekehyyksessä. Painopiste viitekehyyksessä tulee olemaan data-analytiikassa, ennustemallissa ja

vuokrausliiketoiminnassa. Empiirinen tutkimus kohdennetaan Ruotsin vuokrausliiketoimintaan ja tämä esitetään opinnäytetyössä viitekehyksen jälkeen.

1.3 Tutkimusmenetelmät

Opinnäytetyö toteutettiin kehittämishankkeena ja se oli tyypiltään kvantitatiivinen analyysi. Laskennallisia ja tilastollisia menetelmiä, algoritmeja ja kertoimia hyödynnettiin luotettavasti tutkittavan asian ratkaisemiseksi. Kvantitatiivisessa analyysissä eli määrällisessä analyysissä asiaa tarkastellaan ja tulkitaan tilastojen ja numeroiden avulla (Koppa 2015). Opinnäytetyön tekoprosessi sisälsi paljon erilaisia taulukoita, tilastoja ja analyyseja liittyen dataan ja näin ollen painoarvo kvantitatiiviselle tutkimukselle oli suurempi.

Opinnäytetyön empiirisessä osiossa tutkitaan Ruotsin liikevaihdon arvioinnin tarkkuutta ja ennustamista ennustemallin valossa. Pää tavoitteena on tuoda esiin mahdollisimman paljon tekijöitä, joiden takia luvut eivät ole täsmänneet alkuperäisiin arvoihin ja lähteä purkamaan näitä erilaisten kuvioiden ja algoritmien avulla lukijaystävälliseen muotoon.

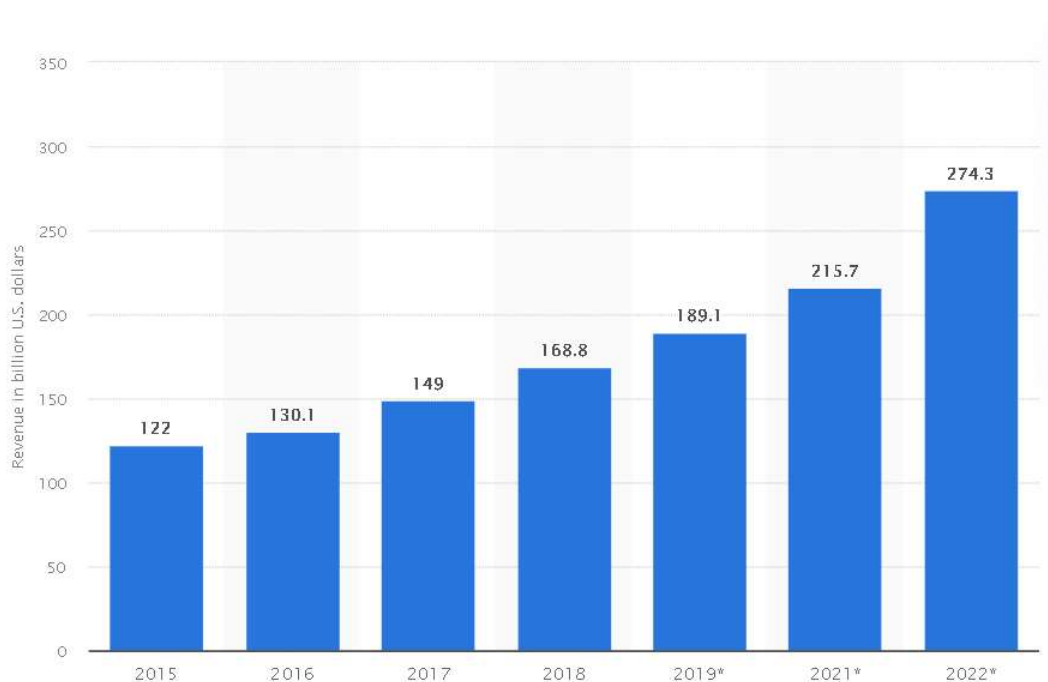
1.4 Tutkimusaineisto

Tämän opinnäytetyön tutkimusaineisto koostui toimeksiantajalta saaduista materiaaleista, ennustemallista ja data-analyyseista. Tämän lisäksi tarkasteltiin muita tutkimuksia liittyen liikevaihdon ennustamiseen ja data-analyysien tarkkaisuun. Tutkimusaineiston pohjalta luotiin Excelillä erilaisia taulukoita, tilastoja, kuvia ja algoritmeja, joilla kuvattiin ratkaisuja ja tärkeitä tunnuslukuja analyyseissa.

Tutkimus sisälsi materiaalia myös Ruotsin liiketoiminnan kuukausittaisesta liikevaihdon ennustamisesta. Ramirent-konsernilla on päivittäinen liikevaihdon raportointi suoraan Ranskaan mutta suurinta vaihtelua vertailuarvoissa on ollut nähtävissä nimenomaan Ruotsin piirissä.

2 Big data ja data-analyysi

Tietojenkäsittelyn ja teknologian kasvu jatkaa kehittymistään samaan aikaan, kun taas datan määrä kasvaa ja sen käyttömuodot ja -mahdollisuudet muuttuvat. Tämä määrä big dataa on jo sellaisella tasolla, että sen suuruutta on vaikea ymmärtää ilman tekoälyä ja automatisaatiota. Tämä nykypäivänä nähtävä kehityssuunta on haasteellinen kuitenkin jatkuvasti kehittyvälle tietojenkäsittelylle ja sitä hyödyntäville ohjelmistoille mahdollistaen uusia liiketoimintamahdollisuuksia. Esimerkiksi Walmart, yksi maailman suurimmista vähittäiskauppaketjuista käsittelee yli miljoona maksutapahtumaa tunnissa kerryttäen datavarastoa tässä ajassa jopa 2,5 petabitin edestä. (Marr 2015.)



Kuvio 1. Big datan ja liiketoiminta-analytiikan tuotto havainnollistettuna (Statista).

Statistan kuviossa 1 on nähtävissä arvio big datan ja liiketoiminta-analytiikan tuotosta maailmanlaajuisesti vuosina 2015–2022. Olennainen osa kuviossa on jatkuva tuoton kasvu, mikä havainnollistaa teknologian jatkuvaa kehitystä. Big dataa hyödyntää tänä päivänä suurin osa keski- ja suurikokoisista yrityksistä ja näin ollen myös arvio vuodelle 2022 on 274,3 miljardia dollaria. (Statista.)

2.1 Tarkoitus ja käyttömenetelmät

Big data on ollut jo pitkään läsnä liiketoiminnallisissa keskusteluissa mutta sillä ei ole yleistä kansainvälistä määritettä. Suurta määrää järjestelmätöntä tietoa on pitkään kutsuttu big dataksi. Mitä tämä big data sitten oikeastaan tarkoittaa? Monilla teknologian asiantuntijoilla on hyvin erilaisia määritelmiä big datalle ja sitä yhtä ja oikeaa vastausta ei varmasti ole olemassa vielä tänä päivänäkään. (Big data framework 2019.)

Big dataa voidaan näin ollen pitää prosessina, joka tuottaa erilaisia päätöksenteon oivalluksia. Dataa analysoimaan käytetään ihmisiä ja teknologiaa, jotta voidaan tuottaa mahdollisimman nopeasti erityyppistä dataa eri lähteistä. Tällä pyritään mahdollistamaan niin sanottu virta sisältäen toimivan tietolähteen, jota voidaan hyödyntää jatkotutkimuksissa. (Kalyvas & Overly 2014.)

Yksinkertaisuudessaan big dataksi luokitellaan tietojoukot, joiden koko on suurempi kuin esimerkiksi tietokantaohjelmistotyökalujen kyky kaapata, tallentaa, havainnoida ja analysoida. Massiivisten tietojoukkojen, jota kutsutaan big dataksi, arvo liikkuu yleisemmin terabittejen ja petabittien välillä. (Marr 2015.)

Big datalla on lukuisia käyttömenetelmiä mutta useimmiten yrityksissä haetaan vastausta johonkin kysymykseen tai pyritään etsimään uusia tekijöitä, joiden avulla voidaan kehittää jotain tiettyä asiakohtaa organisaation sisä- tai ulkopuolella. Big dataa voidaan siis hyödyntää esimerkiksi tuotekehityksessä, kuten nyt tässä tapauksessa liikevaihdon arvioinnin tarkkuudessa tai halutessaan saada tietoa asiakkaasta, markkinoista ja kilpailijoista. Näin ollen voidaan sanoa, että big datan kaupalliset hyödyt ovat selvillä mutta nykypäivänä yritysjohtajien on oltava hyvin tarkkana riskin rajoittamisessa. Big datan käytöllä täytyy olla tietty liiketoiminnallinen tarkoitus ja sitä analysoivien ohjelmistojen tulee noudattaa tätä tarkoitusta. Big dataan liittyviä oikeudenkäyntejä on nähty useaan otteeseen viime vuosina, ja näissä suurimpina yhdistävänä tekijänä on ollut yksityisyyden loukkaus. Suuret yritykset, joille big data ja sen jatkuva analysointi on

toistuvaa, voivat pienentää tätä ja muita riskejä selvittämällä heidän liiketoimintamallinsa kaupallisen ja julkisen arvon. (Kalyvas & Overly 2014.)

2.2 Big datan vaikutus markkinoihin

Big data -analytiikkaa on nykyään nähtävissä melkein päivittäisessä elämässä. Yritys, kuten Google, pyrkii hyödyntämään big dataa päivittäin ja saamaan asiakkaastaan kaiken mahdollisen tiedon irti. Tämän tekee helpoksi yritykselle se, että harvemmin ihmiset vaivautuvat lukemaan kymmeniä sivuja pienellä kirjoitettuja käyttöehtoja, vaan valitsevat helpomman ratkaisun ja hyväksyvät ne säästämällä aikaa ja vaivaa. Tämän jälkeen esimerkiksi Google pystyy arvioimaan hakutuloksia, ennustamaan Internet-liikenteen käyttöä ja palvelemaan asiakkaita suoraan heistä saadun big data -raportin avulla. Mainonta taas kohdistetaan käyttäjien mieltymyksiin, evästeiden, historian ja monen muun seikan mukaan niin, että mainokset tuntuvat läheisemmiltä ja houkuttelevimmilta asiakkaan näkökulmasta. (Ohlhorst 2021.)

Big datan avulla voidaan toteuttaa monia erilaisia farmaseuttisia ja lääketieteellisiä tutkimuksia ihmisten hyvinvoinnin kehittämiseksi. Oireista ja taudeista kerätään tärkeää tietoa, jotta voidaan analyysien avulla kehittää vasta-aineita kyseisille taudeille ja kiputiloille. Yksi tärkeimmistä big dataa hyödyntävistä tahoista ovat meteorologit ja näin ollen sään ennustaminen. Ilman big dataa ja sen tarjoamia hyötyjä meillä ei olisi samanlaista käsitystä esimerkiksi maanjäristysten muodostumisesta, globaalista ilmastosta tai ylipäätään sääennustuksista. Ennustamista voidaan pitää yhtenä tärkeimpänä avainsanana, jonka big data on mahdollistanut ihmiskunnalle viime vuosikymmenien saatossa. (Ohlhorst 2021.)

Teknologian kehityksen myötä big datan vaikutus markkinoihin on hurja, mutta kuluttaja ei näe vaikutusta samalla lailla. Yhä vaikeampaa, monimutkaista ja suurta määrää dataa pystytään analysoimaan monilla eri ohjelmistoilla robotiikan avulla. Yritykset pyrkivät näin ollen tuottamaan erilaisia palveluita, tuotteita ja elektroniikkaa sen mukaan, mikä nykypäivänä asiakkaita kiinnostaa ja mikä on nähtävissä valtavirran mieltymysten mukaisessa käytössä. (Ohlhorst 2021.)

2.3 Big datan potentiaaliset riskit ja haasteet

Teknologian kehityksen myötä on saatu aikaiseksi yhä enemmän mielenkiintoisia innovaatioita. Monet näistä ovat olleet ihmiskunnan kannalta erittäin hyödyllisiä, tehokkaita tai tarpeita tyydyttäviä, kuten tekoälyn ja robotiikan nykyinen taso. Big data ja ylipäättään datan analysoiminen voidaan laittaa saman kategorian alle kehittämällä tehokkuutta ja luotettavuutta monissa erilaisissa prosesseissa eri toimialoilla. Monet näistä seikoista sisältävät myös riskejä ja haavoittuvuuksia, joita voidaan tai ei voida välttää joissakin tapauksissa. Big dataan liittyviä riskejä on lukuisia mutta tärkeimpiin nykypäivänä luokitellaan turvallisuusongelmat, eettiset ongelmat ja datan väärinkäyttö, kulujen hallinta ja varastoinnin ylläpitäminen. (Tilastokeskus 2020.)

Monet maailman suurimmista yrityksistä käsittelevät erittäin paljon arkaluonteista tietoa, kuten henkilökohtaisia asiakastietoja tai strategisia asiakirjoja. Datamäärä nykypäivänä on hyvin suurta ja tietomurrot ovat olleet kasvussa. Tämän takia syytökset yksityisyydenloukkauksista ovat olleet hyvin yleisiä talouden kohdistuvassa big datassa. Mediassa on ollut myös viimeaikoina paljon puhetta siitä, että ihmisten henkilökohtaisia tietoja olisi myyty eteenpäin eri yritysten välillä. Asiakas ei todellisuudessa koskaan tiedä, mitä yritys mahdollisesti saamallaan datalla tekee. (Forbes 2020.)

Eettisyyden nimissä big dataa on väärinkäytetty paljon muun muassa vakuutusyhtiöissä, joissa kolarit laskevat automaattisesti asiakkaan bonuksia. Vakuutusyhtiöt pystyvät myös näin ennustamaan tilanteita heille parhaaksi näkemällään tavalla. Big data aiheuttaa myös paljon kuluja yrityksille. Analysoinnit, raportit, arkistoinnit ja tallennukset maksavat huikkeen määrän rahaa ja tämän takia monet pienet ja keskikokoiset yritykset eivät usein hyödynnä big dataa. Näin suurta dataa varten tarvitaan isot kapasiteetit varastoida sitä ja sen jatkuva ylläpitäminen tietää automaattisesti kuluja organisaatiossa. (Hillier 2021.)

Nykypäivänä yritysten tulee pystyä mahdollistamaan datan jatkuva monitorointi ja automatisoida mahdollisimman monta eri prosessia riskien minimoimiseksi.

Jatkuva monitorointi mahdollistaa sen, että ongelman tapahtuessa yritys pystyy reagoimaan tähän heti ja alkaa tutkimaan mahdollista keinoa korjata kyseistä ongelmaa. Big dataa käsittelevät sovellukset pyritään rakentamaan niin käyttäjäväliseksi kuin on mahdollista, jotta isoimmilta virheiltiltä dataa käsitellessä pystytään välttymään. Isoimmat yritykset käsittelevät ja varastoivat dataa pilvipalveluun ja siihen suositellaan kaikkien yritysten siirtyvän, jos ei näin ole jo tapahtunut. Nykypäivänä pelkkä tietokone tai serveri itsessään ei pysty täyttämään samoja vaatimuksia kuin pilvipalvelut, kun puhutaan kaikista big dataan liittyvistä riskeistä liiketoiminnassa. (Talend 2021.)

2.4 Big datan hyödyntäminen vuokrausliiketoiminnassa

Ramirent-konsernin vuokrausliiketoiminta on yritystoimintaa, jossa koneita ja laitteita vuokrataan yksityisesti tai yritysten välillä sovittua korvausta vastaan. Vuokrattavien laitteiden kysyntä on kasvanut merkittävästi varsinkin Suomessa. Varsinaisten laitteiden ja koneiden ylläpito, säilyttäminen ja muut huoltokulut tulevat yksinkertaisesti kalliimmaksi kuin vuokrata tarvittavat laitteet esimerkiksi tiettyä projektia varten. Ramirent-konserni on yksi Euroopan suurimmista koneiden ja laitteiden vuokraajista ja kasvavan kysynnän myötä myös muutos on nähtävissä kasvavasta liikevaihdosta. Suurin osa konevuokraamojen asiakkaista on rakennusteollisuudessa ja yleisimmät vuokraukset liittyvät työmaalaitteistoihin. (Ramirent 2021.)

Useimmissa tapauksissa big dataa hyödynnetään jonkin tietyn asian tai prosessin kehittämisessä. Monissa tapauksissa markkinointiin ja sen kohdistamiseen hyödynnetään dataa, joka on saatu yrityksessä asioivilta asiakkailta. Ramirentin tapauksessa big dataa on kuitenkin hyödynnetty yrityksen sisällä, mikä on mahdollistanut entistä luotettavampia ja tehokkaammin saadut arviot kuukausittaisessa liikevaihdossa. Esimerkiksi aikajana-analyysit, regressiotestaukset, simulointi, koneoppiminen ja tekoäly pyörivät big data -sovelluksilla huomattavasti tehokkaammin verrattuna muuhun tietotekniikkaan. Big datan avulla valtavaa tietomäärää voidaan nyt halutessa analysoida monesta eri näkökulmasta, mikä on ollut aikaisemmin mahdotonta. (Talouselämä 2013.) Ramirent-konserni

hyödyntää monia näistä analyysimenetelmistä kehittääkseen arvioinnin tarkkuuden luotettavuutta, tehokkuutta ja tarkkuutta pienemmällä resurssien käytöllä.

3 Data-analyysiin vaikuttavat tekijät

Tässä luvussa perehdytään data-analyysiin vaikuttaviin tekijöihin, joita hyödynnetään ja käytetään ennustemallissa. Luvussa 3 tarkastellaan yleisimpiä käsitteitä ja tarkemmat yksityiskohtaisemmat käsitteet käydään läpi seuraavassa osiossa ennustemallia käsiteltäessä.

3.1 Liikevaihto

Liikevaihto käsitteenä tarkoittaa yrityksen tuotteista ja palveluista saatuja myyntituottoja ilman, että hintoihin on lisätty arvonlisäveroa. Liikevaihdosta vähennetään arvonlisäveron lisäksi myös mahdolliset alennukset, avustukset ja myynnin määrän mukaan muodostuneet verot. Liikevaihto on yksi hyvä esimerkki yrityksen laajuudesta, ja se kertoo, kuinka hyvin yrityksen tuotteet tai palvelut ovat tehneet myyntiä vuoden aikana. (Tilastokeskus.)

Arvonlisävero ei kuulu yritykselle vaan tämä raha tilitetään eteenpäin verottajalle. Näin ollen myöskään liikevaihto ei pelkästään kerro yrityksen kannattavuudesta ennen kuin siitä on vähennetty kaikki yritykselle vuoden aikana kertyneet kulut. Liikevoitto eli liiketoiminnan tulos ennen veroja ja rahoituseriä on nähtävissä, kun liikevaihdosta vähennetään yrityksen kulut ja poistot. (Yrityksen perustaminen.net.)

Suurin osa Ramirent-konsernin liikevaihdosta koostuu konevuokrauksista ja eteenpäin myydyistä koneista tai laitteistoista suoraan asiakkaalle. Tämän takia ennustemalli on pyritty rakentamaan niin, että sitä voidaan hyödyntää eri toimipisteiden ja eri maiden liikevaihdon tarkasteluissa.

Tämän opinnäytetyön data-analyysin ennustemalli keskittyy liikevaihdon ympärille pääosin kuukauden mittaisilta aikajaksoilta. Kuukauden aikana syntyneitä lukuja tarkastellaan ja analysoidaan erilaisia kuvioita ja algoritmeja hyödyntäen

luotettavaan analyysiin vedoten. Liikevaihdon ennustaminen yritykselle on hyvin tarkkaa, sillä pieleen mennyt ennuste voi tarkoittaa ylimääräisiä kuluja, jos esimerkiksi yritys on varautunut liian suureen kysyntään. Liikevaihtoa ennustaessa mallit yleensä perustuvat historialliseen dataan, jotka hyödyntävät aikasarja-analyyseja tarvittaessa.

3.2 Liikevaihdon tunnusluvut

Liikevaihtoon liittyy myös tunnuslukuja ja tässä osiossa käydään läpi niiden merkitys. Tärkeimpiä toimintaedellytyksiä yritykselle tänä päivänä ovat kannattavuus, vakavaraisuus ja maksuvalmius. Yrityksen toiminnalle näiden ylläpitäminen ja jatkuva monitorointi on hyvin tärkeää, mutta kun katsotaan liikevaihtoa ja sen analyysiä, on tarkasteltava aivan toisia katemittareita.

Liikevaihtoon suhtautetut katemittarit ovat: Myyntikate-%, käyttökate-%, liiketulos-%, rahoitustulos-%, nettotulos-% ja viimeisenä kokonaistulos-%. Näistä tärkeimpänä tämän opinnäytetyön kannalta voidaan pitää nettotuloksen, kokonaistuloksen, EBIT ja EBITDA prosenttiarvoja. Nettotulos käsitteenä tarkoittaa samaa kuin yrityksen toiminnan tulos tilikauden aikana. Tämä on se omistajille jäävä osuus, kun nettotuloksesta on poistettu rahoituserät ja verot. Kokonaistulos saadaan lisäämällä tai vähentämällä satunnaiset erät nettotuloksesta. Satunnaisia eriä voivat olla esimerkiksi pysyvien vastaavien kertaluontoiset myyntivoitot tai tappiot tai annetut ja saadut konserniavustukset. Prosenttiarvot nettotulokselle ja kokonaistulokselle saadaan jakamalla nämä liikevaihdolla, jonka jälkeen arvo kerrotaan sadalla ja lisätään %-merkki, jolloin tulos muutetaan prosenteiksi. EBIT-luvulla tarkoitetaan yritysarvon ja liiketuloksen suhdetta, kuten taas EBITDA-luvulla tarkoitetaan yrityksen kahdentoista kuukauden tulosta vähennettynä muuttuvat ja kiinteät kustannukset. (Yrityksen perustaminen.net.)

3.3 Kausiluontoiset tekijät

Data-analyysiä tehdessä kausiluontoisten tekijöiden rooli on suuri valmistellessa ennusteita liikevaihdosta. Ravintoloissa ja yökerhoissa kysyntää palveluiden ja

tuotteiden myynnistä on eniten varmasti viikonloppuisin, kun taas laskettelukeskuksien asiakaskysyntä on pääosin talviaikaan.

Ennustettavaa dataa liikevaihtoa analysoitaessa on otettu huomioon työpanos, katastrofitekijät ja mahdolliset pyhäpäivät. Viikonloppuisin ei ole myyntiä ollenkaan ja jonain päivinä myyntiä on yleisesti enemmän kuin jonain toisena arkipäivänä. Katastrofitekijöitä voivat olla muun muassa sään vaihtelut, kuten kovat pakkaset Suomessa tai muut yrityksen liiketoimintaa vaikeuttaneet ulkoiset tekijät. Kausiluontoisia tekijöitä hyödynnetään regressioita, algoritmeja ja kuvioita muodostaessa. Tämän avulla pystytään tekemään ennuste jo tulevien kuukausien liikevaihdoista samaa mallia noudattaen. (Ramirent 2021.)

3.4 Sovellukset ja toiminnot

Tässä opinnäytetyössä hyödynnettiin muutamia eri tietotekniikan sovelluksia datan analysoimiseksi. Pääsääntöisesti data ja sen käsittely tapahtui Microsoft Excel -taulukkolaskentaohjelmassa. Ennustemallia hyödynnettiin Excelissä ja sen avulla mahdollistetaan erilaisia algoritmeja, kuvioita ja malleja datan havainnollistamiseksi helppolukuisempaan muotoon.

Alkuperäinen koodi ennustemallin ympärille on tehty käyttäen ohjelmointikieltä nimeltä Python. Näin ollen ennustemalli pystytään lataamaan tiettyyn Excel-tiedostoon ja sen avulla voidaan tarkastella haluttuja tunnuslukuja analyysia varten. Tässä opinnäytetyössä ennustemallin koodia ei käydä tarkemmin läpi vaan tarkastellaan sen tuomia mahdollisuuksia ja tehokkuutta helpottavia käyttömenetelmiä.

4 Ennustemallin käyttöönotto

Tässä luvussa keskitytään itse ennustemalliin, jota hyödynnetään tutkimuksessa liittyen Ruotsin toimipisteiden liikevaihtoon. Luvussa 4 käydään tarkasti läpi ennustemallin rakentaminen, käyttö ja suoriutuminen todellisessa

tilanteessa. Ennustemalli on hyvä ja oleellinen osa liiketoiminnan tarkkuuden arvioinnissa ja luotettavuudessa pitkän tähtäimen liikevaihdon ennustamisessa.

4.1 Ennustuksen vaiheet

Yritykset tänä päivänä elävät sellaisessa ympäristössä, missä muutokset ja epävarmuus tulevasta ovat hyvin oleellisia. Ennusteet tulevasta ovat yleensä yrityksiä tukena pitkän aikavälin päätöksissä ja operatiivisissa päätöksissä. (Hanke & Wichern 2014.) Ramirent-konsernin kohdalla liikevaihto ja vuokratuottojen määrän arvio on mahdollista muodostaa hyödyntäen prosessoitua dataa teknologian kehityksen myötä.

Itse ennustusprosessi on hyvin monivaiheinen ja jokainen kohta on yhtä tärkeä mahdollisimman tarkan ja paikkansapitävän ennustuksen mahdollistamiseksi. Ennustusprosessi voi sisältää lukuisia määriä välivaiheita mutta yleisemmin se voidaan jakaa viiteen eri vaiheeseen: 1) Ongelman määrittäminen, 2) Datan manipuloiminen ja siivoaminen, 3) Mallin rakentaminen ja sen hyödyntäminen, 4) Mallin käyttö, 5) Ennusteiden arviointi ja tulokset (Hanke & Wichern 2014).



Kuvio 2. Liiketoiminnan ennustusprosessi havainnollistettuna (Hanke & Wichern 2014).

Hanke & Wichernin kuviossa 2 on nähtävissä viiteen eri osaan jaettu liiketoiminnan ennustusprosessi. Vaiheet toistavat hyvin samanlaista kaavaa, kuin aikaisemmin mainittu viiden vaiheen ennusteprosessi. Ennusteprosessia seuraamalla yritykset saavat paremmat käsityksen siitä, kuinka lähteä ratkaisemaan kyseistä ongelmaa ja miten hyödyntää lopullisia tuloksia välivaiheiden jälkeen.

Seuraavat ennusteprosessin vaiheet käydään toimeksiantajayrityksen näkökulmasta samalla tavalla, kun tätä kyseistä ennustemallia hyödynnettiin. Datan

kerääminen ja ongelman määrittäminen menevät hyvin käsikädessä, sillä tätä ennustemallia hyödynnettäessä datan tulee olla jo valmiina ja manuaalisesti syötettynä sovellukseen. Datatasaaminen ja koonti on yleensä aikaa vievää ja haastavaa ja se onkin yleensä ennustamisen hitain mutta tärkein välivaihe. Tässä vaiheessa, kun ongelma on jo entuudestaan tuttu, pitää alkaa miettimään sitä, kuka tulee ennusteita hyödyntämään, mihin tarkoitukseen ennusteita käytetään ja miten näistä hyödytään pidemmälläkin tähtäimellä liiketoiminnan tueksi (Hanke & Wichern 2014).

Seuraava eli toinen vaihe on ennustemallin kannalta yksi tärkeimmistä vaiheista ennusteprosessissa. Dataa pitää pystyä tarvittaessa karsimaan, lisäämään tai jopa muokkaamaan. Datasta voi puuttua tärkeitä arvoja, jotka vaikuttavat suoraan ennustemallin toimintaan. Tässä vaiheessa datasta voidaan visualisoida erilaisia trendejä ja muuttujia mutta kuitenkin pitäen datan aina relevanttina alusta loppuun saakka. Esimerkiksi dataa kauempaa menneisyydestä ei ole hyödyllistä käyttää vertailun kohteena, jos tiedämme tilanteen muuttuneen nyky päivänä aivan pääläelleen (Hanke & Wichern 2014).

Kolmannessa vaiheessa valitaan ennustemalli, joka täyttää tarvittavat vaatimukset ja sopii parhaiten käsittelemään tietynlaista dataa. Malleja voidaan kokeilla useampia sen parhaiten sopivan löytämiseksi kuitenkin niin, että malli on tarpeeksi yksinkertainen käyttää ja sitä käyttävät ymmärtävät miten malli tuottaa tietyt ennusteet. Ennustettu data ei varmasti ole tarpeeksi luotettavaa, jos malli on tuntematon sitä käyttäville henkilöille. Mallin mukaisesta ennustuksesta ei näin ollen voida lähteä tekemään jatkotoimenpiteitä, jos sitä käyttävät henkilöt eivät ole tarpeeksi hyvin koulutettuja näihin toimenpiteisiin. (Hanke & Wichern 2014.)

Neljäs vaihe sisältää mallin käyttöönoton ja tämän avulla aletaan tekemään ennusteita ja analyyskejä datasta. Historiallista dataa hyödyntämällä tehdään ennusteita ja verrataan niitä tunnettuihin kysynnän arvoihin, jotta saadaan tarkkuutta ja oikeudenmukaisuutta ennusteille. Datasta analysoidaan näin ollen kaikki tarpeellinen, jotta sitä voidaan hyödyntää viimeisessä vaiheessa tiettyyn

tarkoitukseen. Viimeisessä vaiheessa ennusteet ovat valmiina ja alkaa ennusteiden läpikäynti. Siinä vaiheessa, kun historiallista dataa on analysoitu tarpeeksi ja mallin arvoista on tarpeeksi tietoa, voidaan aloittaa mallin parantelu käyttäjien ja mallin tekijän toimesta. Ennusteilla on näin ollen iso vaikutus yrityksessä toteutettaviin seuraaviin toimenpiteisiin. Huonon ennusteen sattuessa yritysjohdon on tehtävä kaikkensa, jottei ennuste tulisi toteen. Hyvän ennusteen sattuessa tehtävänä on tukea ennustetta kaikilla mahdollisissa keinoilla tämän toteutumiseksi kuitenkin niin, että liiketoiminnan kannattavuus ja yrityksen tuloksen tekokyky on jatkuvan monitoroinnin alla. (Hanke & Wichern 2014.)

4.2 Kartoittaminen

Tämä kyseinen ennustemalli on historiaan pohjautuva malli, jota hyödynnetään tämän opinnäytetyön empiirisessä tutkimuksessa. Malli tarvitsee pohjalleen datan ja tämän kautta se pystyy ymmärtämään trendit ja syklisyyden tietoja tutkiessa. Kyseinen malli on otettu käyttöön vuoden 2021 syksyllä ja nyt datan määrän kasvaessa tätä voidaan hyödyntää entistä enemmän koko vuoden syklin identifioimiseksi.

Nykypäivänä datan historiasta voimme nähdä selkeitä laskevia ja nousevia trendejä ja oletuksena voimme ajatella niiden toistavan itseään. Tämä ei kuitenkaan pidä paikkaansa ja ajatuksena on, että pidemmällä aikavälillä saataisiin malliin enemmän historiaa tunnistamaan pidemmän tähtäimen trendejä. Mallia on kehitetty niin, että se pystyy tunnistamaan muuttuvia muuttujia, kuten juhlapyyhiä keskellä viikkoa tai viikon lopussa. Näin ollen se pystyy laskemaan mahdolliset nousut tai laskut vuokratuotoissa tietyille päiville verraten standardiin lukemaan.

Edellisessä luvussa 3.4 mainitut kausiluontoiset tekijät on otettu uutena osana mukaan ennustemalliin. Esimerkiksi talviaikaan tuotteet, kuten työmaille menevät lämmittimet ja valaisimet menevät hyvin kaupaksi, kun taas suuri lumen määrä voi mahdollisesti estää tiettyjen tuotteiden tai laitteiden käytön. Näihin positiivisiin ja negatiivisiin tekijöihin pystytään luomaan ennusteita ja tulevaisuudessa olisi tarkoitus pystyä hyödyntämään myös sääennusteista saatua dataa

suoraan ennustemalliin. Ennustemalli on tyyliltään hyvin mekaaninen ja se tekee standardin käsittelyn jälkeen arvionsa.

Pääsääntöisesti ennustemallin tehtävä on ennustaa tulevien kuukausien liikevaihto tai mahdolliset vuokratuotot, joissa hyödynnetään jo syötettyä dataa ja huomioidaan vaihtelevat muuttujat. Näiden tekijöiden pohjalta muodostetaan tietty noudatettava sykli. Malli noudattaa tiettyjä standardeja ja siihen voidaan tehdä muutoksia tarvittaessa käsin tarkemman arvion saavuttamiseksi.

4.3 Mallin rakennusosat

Tämän opinnäytetyön empiirisessä tutkimuksessa hyödynnettävä ennustemalli on rakennettu pääsääntöisesti sen datan ympärille, jota käsitellään siinä hetkessä. Jotta malli toimii, se tarvitsee seuraavat kohdat Excelissä: Viikonpäivät, kausiluontoiset tekijät ja mahdolliset katastrofitekijät, jotta se voi laskea uuden ennusteen.

Oheisessa taulukossa on nähtävissä esimerkki viikonpäivistä maanantaista sunnuntaihin ja jokaisen arkipäivän kohdalla numero yksi lukuun ottamatta viikonloppua. Kausiluontoisista tekijöistä johtuen todellinen arvo muuttuu hieman ylös tai alas standardista arvosta yksi. Esimerkissä on tiedossa hirmumyrsky kuun lopussa, ja todelliset arvot tippuvat näin ollen hieman alemmas oletetusta arvosta.

Taulukko 1. Ennustemallin perustekijät havainnollistettuna.

Viikonpäivä	Päivittäinen	Kausiluontoinen tekijä (Viikonpäivä)	Katastrofitekijä	
1,00	1,00	0,95	28.10.2021	0,2
2,00	1,00	1,05	29.10.2021	0,7
3,00	1,00	0,98	30.10.2021	0,3
4,00	1,00	1,02	31.10.2021	0,1
5,00	1,00	0,8		
6,00	0,00	0		
7,00	0,00	0		

Taulukossa 1 havainnollistetaan ennusteen laskeminen seuraavalle viikolle hyödyntäen ennustemallia. Ennusteessa on huomioitu kasvava budjetti alkuviikolta

ja jo tiedossa olevat ennusteet. Tässä esimerkissä on nähtävissä katastrofiennuste ja sen vaikutus lopulliseen ennusteeseen seuraavan viikon budjetista.

Taulukko 2. Budjetin laskeminen seuraavalle viikolle havainnollistettuna.

Viikko 1	8.10.2021					Uusi budjetti
Ennuste	100 000	105 000,00	Uusi ennuste	Tiedossa olevat ennusteet	Katastrofiennuste	99 450,79
1.10.2021	5,00	1,00		4160,00	0,00	
2.10.2021	6,00	0,00		0,00	0,00	
3.10.2021	7,00	0,00		0,00	0,00	
4.10.2021	1,00	1,00		4940,00	0,00	
5.10.2021	2,00	1,00		5670,00	0,00	
6.10.2021	3,00	1,00		4949,00	0,00	
7.10.2021	4,00	1,00		5406,00	0,00	
8.10.2021	5,00	0,80	4203,95		4203,95	79 875,00
9.10.2021	6,00	0,00	0,00		0,00	
10.10.2021	7,00	0,00	0,00		0,00	
11.10.2021	1,00	0,95	4992,19		4992,19	
12.10.2021	2,00	1,05	5517,68		5517,68	
13.10.2021	3,00	0,98	5149,84		5149,84	
14.10.2021	4,00	1,02	5360,03		5360,03	
15.10.2021	5,00	0,80	4203,95		4203,95	
16.10.2021	6,00	0,00	0,00		0,00	
17.10.2021	7,00	0,00	0,00		0,00	
18.10.2021	1,00	0,95	4992,19		4992,19	
19.10.2021	2,00	1,05	5517,68		5517,68	
20.10.2021	3,00	0,98	5149,84		5149,84	
21.10.2021	4,00	1,02	5360,03		5360,03	
22.10.2021	5,00	0,80	4203,95		4203,95	
23.10.2021	6,00	0,00	0,00		0,00	
24.10.2021	7,00	0,00	0,00		0,00	
25.10.2021	1,00	0,95	4992,19		4992,19	
26.10.2021	2,00	1,05	5517,68		5517,68	
27.10.2021	3,00	0,98	5149,84		5149,84	
28.10.2021	4,00	1,02	5360,03	0,20	1072,01	4288,03
29.10.2021	5,00	0,80	4203,95	0,70	2942,76	1261,18
30.10.2021	6,00	0,00	0,00	0,30	0,00	
31.10.2021	7,00	0,00	0,00	0,10	0,00	

Taulukossa 2 on nähtävissä viikonpäivät numeroituna yhdestä seitsemään. Tämän oikealla puolella on nähtävissä kausiluontoisten tekijöiden vaikutus alkaen päivästä 8.10.2021. Ennuste päivästä 8.10.2021 eteenpäin loppukuuhun asti saadaan laskettua kaavalla:

$$(a - b) * c / d , \quad (1)$$

missä a on alkuperäinen budjetti, b on jo tiedossa olevat ennusteet, c on kausiluontoisen tekijän arvo 8.10.2021 ja d on kaikkien kausiluontoisten tekijöiden summa aikaväliltä 8.-31.10.2021. Laskukaava lasketaan siis seuraavasti:

$$(105000 - 25125) * 0,8 / 15,20 \quad (2)$$

Tästä laskutoimituksesta saamme tulokseksi 4 203,95. Tämän solun lukitsemalla saamme ennusteet myös kuun loppuun asti. Katastrofi on merkitty tapahtuvan päiville 28.-31.10.2021 ja näin ollen katastrofikertoimia hyödyntäen saamme näille päiville omat ennusteet poissalukien viikonlopun 30.-31.10.2021. Budjetti miinus tiedossa olevien ennusteiden summa täsmää nyt 8.-31.10.2021 uuden ennusteen kanssa. Uusi budjetti seuraavalle viikolle on näin ollen: Budjetti (105 000 euroa) – Katastrofitekijät (5 549,21 euroa) = 99 450,79 euroa.

4.4 Ennustemallin käyttömahdollisuudet ja haasteet

Ennustemallin käyttömahdollisuudet ovat erittäin laajat ja vain mielikuvitus on rajana. Mitä enemmän malliin syötetään tietoa, sitä paremmin se pystyy lähitulevaisuudessa ennustamaan myös monimutkaisempiakin asioita, kuten sääolosuhteiden muutoksia ja pidempiaikaisia trendejä. Mallia on jo hyödynnetty muun muassa regressioanalyysieihin, jotta saadaan enemmän tietoa tulevien kuukausien liikevaihtojen ennusteista. Regressioanalyysit ovat monimuuttujamenetelmiä, joilla mitataan muuttujien välisiä riippuvuussuhteita. Malliin voidaan syöttää erilaisia algoritmitmeja ja ne ovat niin sanottuja kuvauksia prosessien tai tehtävien suorittamisesta niin, että niitä seuraamalla voidaan ratkaista tietty ongelma. Algoritmit toimivat ainoastaan säännönmukaisesti ja tavoitteiden mukaan. Tärkein hyöty algoritmista on ongelmanratkaisukyky ja erilaisia algoritmeja voidaan automatisoida tarpeiden mukaan hyödyntäen tekoälyä, jotta toiminnan tehokkuutta voidaan kehittää. (Peda 2021.)

Mallin haasteena on se, että se itsessään ei käytä harkintaa, eikä se sisällä tekoälyä. Tekoälyllä tarkoitetaan päättelyä, oppimista, suunnittelua ja luomista, jotka ovat ihmisen älyyn tarkoitettuja taitoja. Tekoäly pystyy näin ollen toimimaan itsenäisesti ja analysoimaan aiempia muuttujia ja vaikutuksia. (Euroopan parlamentti 2021.) Malli ei sisällä itseoppivia algoritmeja vaan toiminnot tapahtuvat historian ja syötetyn datan mukaan. Ennustemalli hyödyntää regressiomallia, joka olettaa kehityksen jatkuvan samanlaisena kuitenkin ottaen huomioon kausiluontoiset tekijät ja katastrofitekijät. Katastrofitekijöitä on kuitenkin päivitettävä malliin omatoimisesti ja malli itsessään ei pysty näitä tunnistamaan

tekoälyn puuttumisen takia. Ennuste itsessään on vain ennuste eikä lopullinen totuus. Ennusteet ovat aina jonkinlainen arvaus siitä, miten tapahtumat voisivat kehittyä tiettyjen asetusten ollessa voimassa. Ennustuksiin liittyy aina epävarmuutta ja varsinkin ennustaessa pidemmälle tulevaisuuteen.

4.5 Ennustemallin suoriutuminen

Tähän mennessä ennustemallia on käytetty pelkästään simulaatioharjoituksissa. Malli on osoittautunut hyvin toimivaksi, ja sen toimivuudesta ja käyttömahdollisuuksista on jo saatavissa osviittaa historialliseen dataan katsottuna.

Malli on toistaiseksi vielä kehitysvaiheessa, sillä historiallista dataa ei ole tarpeeksi käytettävissä. Lähitulevaisuudessa malliin on tarkoitus syöttää enemmän tietoa yleisistä trendeistä esimerkiksi liittyen kysyntään sen tehokkuuden maksimoimiseksi. Seuraavassa kappaleessa liittyen tämän opinnäytetyön empiiriseen tutkimukseen käydään tarkemmin läpi ennustemallin toimivuutta ja käyttömahdollisuuksia todellisessa tilanteessa.

Suurimmat ja hyödyllisimmät käyttömahdollisuudet ennustemallille lähitulevaisuudessa ovat kustannussäästöt, tehokkuus ja tavoite saada tietoa tulevaisuudesta erilaisin tilastollisin menetelmin. Inhimillisen virheen marginaali on aina olemassa, kun ennustemallia ja sen dataa käsittelee ihminen mutta tällä hetkellä se nähdään parhaaksi ratkaisuksi ennen kuin resursseja ja koulutuksia siirrettäisiin automatisointiin ja tekoälyn hyödyntämiseen.

5 Tutkimus

Tässä opinnäytetyön osiossa toteutin opinnäytetyöni empiirisen tutkimuksen. Tutkimuksessa tarkasteltiin rakennetun ennustemallin suoriutumista rajattuna Ramirent-konsernin Ruotsin toimipisteiden välille. Ennustemallin kasvua ja kehitystä seurattiin noin alle puolen vuoden mittaiselta ajanjaksolta lokakuusta 2021 helmikuuhun 2022 asti. Tutkimus keskittyi ennustemallin tehokkuuden, luotettavuuden ja näitä yhdistävien tekijöiden arviointiin ja tämän pohjalta myös sen

heikkouksiin ja mahdollisiin epäluotettaviin tekijöihin. Tämän tutkimuksen tavoitteena oli saada vastaus myös siihen, kuinka paljon ennustemalli on pystynyt kehittämään arvioinnin tarkkuutta konsernitasolla.

Ennustemallin ominaisuuksien vertailun lisäksi tutkimusvaiheessa tarkastelin yleisesti muita ennustemalleja ja artikkeleita. Nämä tutkimukset käsittelevät ennustemallin hyödyntämistä eri toimialoilla ja erilaisissa talouteen liittyvissä toiminnoissa, kuten pankkien ja vakuutusyhtiöiden hyödyntämissä ennustusmekanismeissa. Näillä ulkopuolisilla tutkimuksilla pyrittiin luomaan laajempi näkökulma, jotta lopullisista tuloksista voitiin tehdä tarkempia päätelmiä ja johtopäätöksiä.

Tutkimuksen toimialana tutkimusvaiheessa on finanssiala, johon kuuluu tässä opinnäytetyössä rahoitusala, vakuutusala ja organisaation sisäinen talouden hallinta. Organisaation sisäisen talouden hallinnalla tässä tutkimuksessa tarkoitetaan kuitenkin tärkeimpien talouden tunnuslukujen hallintaa ja liikevaihdon nousun tai laskun ymmärtämistä. Tarkastelin eri yhtiöiden sivuja, joista sain paljon paremman käsityksen yhtiöiden taloudellisista tavoitteista, kuin mitä olisi nähtävissä ilman perinpohjaista tutkiskelua. Ennustemallista on hyötyä monessa eri asiassa ja tässä tutkimuksessa selvitin mahdollisimman paljon näistä tekijöistä liittyen liiketoimintaan ja sen hallintaan.

5.1 Kehityssuunta ja analyysimallit

Tämän kehittämishankkeen ja opinnäytetyön ideana oli rakentaa sellainen ennustemalli, jonka avulla liikevaihdon ja muiden olennaisten tunnuslukujen arvioinnin tarkkuus kehittyisi paremmalle ja luotettavammalle tasolle. Ennustemallin rakentamisen jälkeen sitä on hyödynnetty useamman kuukauden ajan ja sen avulla on saatu monia tärkeitä arvioita tarkemmiksi, kuten nyt tärkeimmäksi noussut budjetointi.

Yleisesti ottaen malli on noussut yhdeksi tärkeimmistä elementeistä arvioita ja analyysieja tarkastellessa. Mikään malli ei kuitenkaan ole itsessään täydellinen ja tämän projektin edetessä myös malli itsessään on tuottanut poikkeamia ja

ongelmia matkan varrella. Suurimman haitat yleensä koostuivat vääränlaisesta datasta tai muista ongelmista liittyen dataan, jonka ratkaisemiseen kului ylimääräisiä resursseja ja ongelmanratkaisukykyä tämän yli pääsemiseksi. Nämä haitat eivät sinänsä ole verrattavissa suoraan mallin toimintakykyyn vaan malli ei ole pystynyt tunnistamaan poikkeamia datassa ja korjaukset on täytynyt tehdä käsin pohjiin, jotka sisältävät tuhansia rivejä tärkeää dataa. Malli ei sisällä automatiikkaa tai tekoälyä ja pääsääntöisesti kaikki muokkaukset tehdään käsin alkuperäiseen koodiin tai itse dataan, joka on prosessoitavana Excel -sovelluksessa.

Ennustemalli on käytännössä lineaarinen regressiomalli ja sen avulla datasta oli mahdollista tehdä paljon erilaisia analyysseja, kun dataa tarkasteltiin sen käytyä ennustemallin läpi. Data sisältää tärkeitä kohtia, jotka on hyvä tarkistaa ennen lopullisia analyysseja. Näitä tarkasteltavia kohtia ovat muun muassa: R2-luku, jakelut, pythonin koodi, N on suurempi kuin 30 ja mahdolliset muut tekijät regressioanalyysistä. R2-luku tarkoittaa regressiomallin selitysosuutta. Se pyrkii kertomaan kuinka suuren osuuden selitettävän muuttujan vaihtelusta regressionanalyysin selittävät muuttujat pystyvät selittämään. R2-luku vaihtelee nollan ja yhden välillä ja se saadaan laskemalla mallin tuottamien ennustearvojen ja muuttujan arvojen korrelaation neliö. Mitä suurempi luku on niin sitä paremmin pystyy kertomaan selitettävän muuttujan vaihtelusta. (KvantiMOTV 2008.) Jakelulla tarkoitettiin ennustemallista saatujen yksittäisten toimipisteiden myynnin nousu- tai laskusuhdannetta.

Analyysia tehdessä eri toimipisteitä tarkasteltiin samalla pohtien johtopäätöksiä ja syitä miksi kyseiset toimipisteet erottuivat joukosta niin selvästi. Pythonin koodilla tarkoitettiin itse koodia, jonka avulla ennustemalli toimii ja pystyy ymmärtämään muuttuvan datan ja siihen tehdyt muutokset. Koodi oli hyvä tarkastaa aina ennen itse datan käsittelyä mahdolliset virheiden minimoimiseksi alkuperäisessä koodissa. N on suurempi kuin 30 tarkoittaa sitä, että yleisen ohjeistuksen mukaan yli kolmenkymmenen yksikön tai tässä tapauksessa päivien otannalla on yleensä vakaa normaalijakauma riippumatta muuttujan jakauman muodosta

perusjoukossa. Jos yksikkö tai näyte näyttää erikoiselta tai populaatio on yksinkertaisesti vinoutunut, on turvallisempaa käyttää suurempia näytteitä (Lumen).

Tutkimuksessa vuokratulojen ja liikevaihdon ennustamista suoritettiin koko kuukauden mittaisilta ajanjaksoilta, jolloin otanta oli tarpeeksi suuri ja ennusteita voitiin käsitellä luotettavasti ennustaen budjettia jo seuraavalle kuukaudelle.

Taulukko 3. Analyysin regressiotilastot.

Tarkasteltava asia	Arvo
R	0,730286847
R Neliö	0,572163189
Korjattu R Neliö	0,568504534
Standardivirhe	7,025326233
Havainnot	272

Taulukossa 3 on nähtävissä esimerkki regressiotilastoista. Tarkasteltavana analyysissä on ollut 272 vuorokautta vuonna 2021. Oheinen taulukko kuvastaa mallin selitysstettä. Mallin toimivuutta eli yhteensopivuutta otoksen kanssa voi tarkastella mallin selityssteen avulla. Selitysstete luetaan siis R Neliö sarakkeelta. Tämä muutetaan yleensä prosenteiksi eli mallin selitysstete on 57,2 prosenttia. Korjattu selitysstete R Neliö huomioi kaikkien selitettävien muuttujien määrän ja sitä onkin yleensä parempi tarkastella selityssteen sijasta. Arvoja on muutettu toimeksiantajayrityksen yksityisyyden turvaamiseksi. (KvantiMOTV 2014.)

Taulukko 4. Analyysin ANOVA -taulukko.

	Neliöiden summa	df	Neliön keskiarvo	F	Sig.
Regressio	12820	1	8002,12	528,49	3,22
Jäännös	9580	112	53,87		
Yhteensä	22400	113			

Taulukko 4 on myös yksi tarpeellisista taulukoista regressioanalyysin tarkastuksen kannalta. Regression neliöiden summa 12820 voidaan jakaa neliöiden summan yhteismäärällä 22400 ja saamme tulokseksi 0,572321, joka täsmää ja vastaa mallin selitysstetta. Regression neliöiden summalla tarkoitetaan mallin selittämää vaihtelua ja neliöiden yhteissummalla mallin kokonaisvaihtelua. Regressioanalyysillä saadaan siis karkea osoitus mallin pätevyydestä ja toimivuudesta analysoidessa tietyn tyyppistä dataa. Tämän tyyppinen regressiomalli on hyvin yleinen lineaarisessa regressiomallissa ja tässä opinnäytetyön tutkimuksessa regressioanalyysia tehtiin pääsääntöisesti toimeksiantajayrityksen Ruotsin toimipisteiden vuokratulojen dataa analysoidessa.

5.2 Ongelmat ja niihin reagointi

Heti kartoituksen alussa pyrittiin toimeksiantajan kanssa tunnistamaan ongelmat ja hakemaan niihin ratkaisua tutkimusmielessä. Ongelmia etsittiin liittyen datan ennustettavuuteen, liikevaihdon täsmäyttämiseen ja ylipäätään datan käsittelyyn liittyen. Ongelmia nousi esille lukuisia mutta vain tärkeimmät ja pikaisia toimenpiteitä vaativat seikat nostettiin esille tarkempaa analyysia varten. Data ja siinä jo valmiiksi nähtävät poikkeamat ja virheet toistuivat useaan otteeseen. Ratkaisua oli haettava yhä syvemältä tämän välttämiseksi, jotta tietoihin pystyttäisiin tulevaisuudessa saamaan täysi luottamus.

Dataa ja sen poikkeavuutta verrattuna todellisiin arvoihin lähdettiin tutkimaan kohta kerrallaan datan synnystä sen lopulliseen tilaan asti. Tätä hyödynnettiin analyyseissa arvioinnin tarkkuuden kehittämiseksi sellaiseen tilaan, jossa sitä voitiin pitää luotettavana. Data, joka syötettiin suoraan Exceeliin, sisälsi lukuisia poikkeamia. Kausiluontoisia tietoja, pyhiä tai muuta suoranaisesti dataan vaikuttavia tekijöitä ei huomioitu ollenkaan. Tutkimuksen myötä pyrittiin saamaan vastaus myös siihen minkälainen vaikutus tällä on ja miten ennustemallin avulla nämä poikkeamat ja virheet pystyttäisiin minimoimaan.

5.3 Ruotsin toimipisteet

Tässä opinnäytetyön empiirisessä kartoituksessa rajaus on suoritettu Ruotsin toimipisteiden ympärille. Datan alkuperää, sen käsittelyä ja jatkotoimipiteitä tutkittiin kvantitatiivisin menetelmin hyödyntäen tilastoja, taulukoita ja erityyppisiä analyyseja viitaten ennustemallin tuomiin hyötyihin.

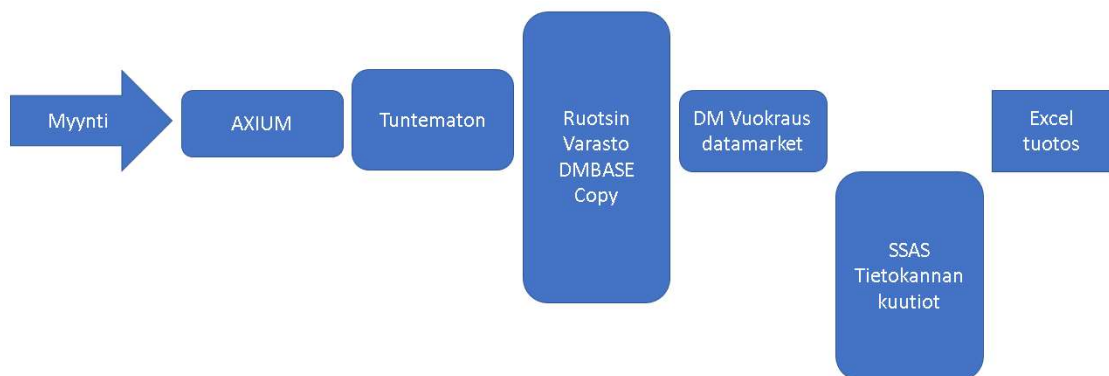
Ennustemalli rakennettiin alun perin nimenomaan Ruotsin toimipisteiden eroavien arvioiden pohjalta kehittääkseen luotettavuutta myös laajemmalla tasolla. Tässä kartoituksessa tutkittiin ennustemallin tehokkuutta, luotettavuutta ja sen yksinkertaisuutta liikevaihdon ja vuokratulojen arvioiden ennustamisessa kuukausitasolla. Data ja sen raportointi tapahtuu kaikilta toimipisteiltä ja data tallentuu myös automaattisesti pilvipalveluun nimeltä Snowflake. Tämän jälkeen ennustemallin koodi syötetään Exceeliin ja sitä kautta data siirretään myös Snowflakesta suoraan Exceeliin. Dataa voi olla tuhansia rivejä tai reilu kolmekymmentä riippuen suoraan otannan määrästä. Tässä vaiheessa poikkeamat ja epäloogisuudet datasta pyritään karsimaan ja saamaan se sellaiseen kuntoon, jotta siitä pystyttäisiin tekemään ennustetta ja analyysia pidemmälle aikavälille.

Rajaus Ruotsin toimipisteiden ympärille oli empiirisen kartoituksen kannalta onnistunut. Datassa ja sen käyttäytymisessä oli nähtävissä eroavaisuuksia esimerkiksi muiden pohjoismaiden tietoihin ja arvioihin verrattaessa. Kartoituksen tueksi luotiin taulukoita, kaavioita ja regressiotilastoja analyysin avuksi hahmotamaan todellista tilannehyötyä uusien käyttöön liittyvien muutoksien kautta.

Kartoituksessa tutkittiin myös yksittäisestä myynnistä saatua dataa, sen kasvua ja käyttäytymistä Exceliin päätymiseen saakka. Data ja sen syntyminen sisälsi erilaisia ja monipuolisia prosesseja, joista saatiin kasaan johdonmukainen polku mahdollistaen datan monipuolisen analysoimisen rakentaen monipuolista ja tehokasta tulevaisuutta organisaatiossa.

5.4 Datan liikkuminen ja käyttäytyminen

Toisena osuutena tässä opinnäytetyön empiirisessä kartoituksessa keskityttiin datan liikkumiseen sen synty pisteestä päättyen Exceliin eli viimeiseen vaiheeseen analyysia varten. Tässä osiossa tarkasteltiin pääsääntöisesti datan liikkumisen ja kehittymisen eri vaiheita, kuten myös sen käyttäytymistä ja mahdollisia muutoksia välivaiheiden aikana.



Kuvio 3. Datan prosessointikuvio välivaiheittain havainnollistettuna Axiom tietovarastosta.

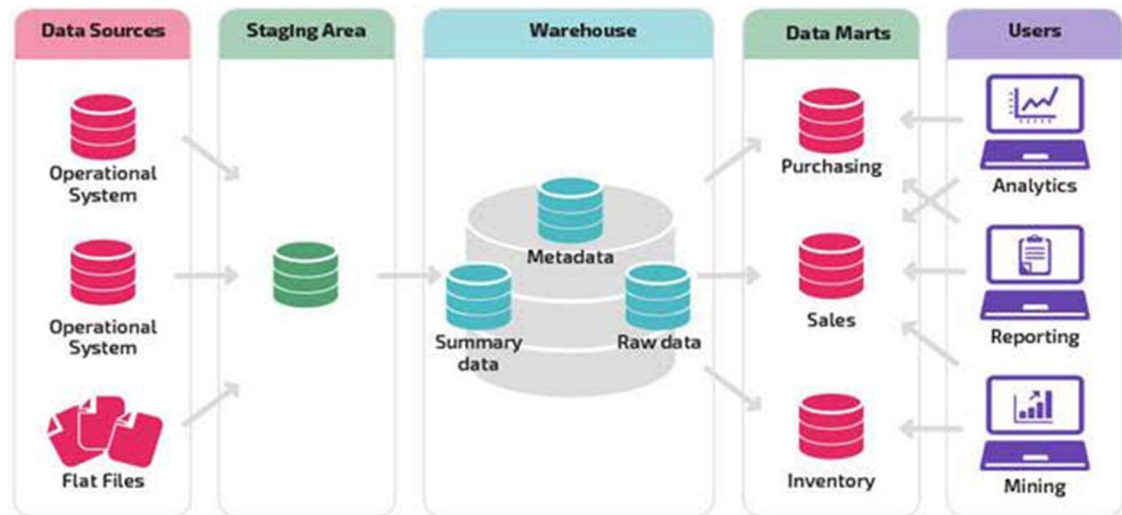
Kuviossa 3 on nähtävissä datan kulku alkuperäisestä myynti- tai vuokraustapahtumasta lopulliseen Excel tuotokseen asti. Prosessia tarkasteltiin ja välivaiheita kirjattiin tarkasti ylös kartoituksen ohessa, jotta saatiin parempi käsitys siitä mitä välivaiheita datan kulku sisältää ja mistä lähteä purkamaan mahdollisia ongelmatilanteita.

Yksittäisestä myynnistä saatu data ladataan automaattisesti erilliseen pilvipalveluun myyntitapahtuman jälkeen. Tiedostoja voidaan näin ollen myös ladata,

hallita ja siirtää toisesta sijainnista toiseen internetin välityksellä. Tässä tapahtumassa hyödynnettiin prosessia nimeltä FTP. Prosessin kansainvälinen nimi on *File transfer protocol* ja suomennettuna protokolla tiedostojen siirtoon. Tämä protokolla mahdollistaa tiedostojen siirron edestakaisin tietokoneiden välillä tai pilven kautta. (Investopedia.) Käyttäjät tarvitsevat vain Internet-yhteyden FTP-siirtojen suorittamiseen. Tässä vaiheessa data on siirtynyt jo automaattisesti pilveen ja henkilökunta pystyy käsittelemään sitä myös kauempaa ja hyödyntäen pelkästään etäyhteyttä. Kuvion 3 kohdassa Tuntematon hyödynnetään muutamia tiettyjä variaatioita kyseisenä päivänä. Lavastusalueita hyödynnetään mahdollisia muutoksia varten datassa ennen sen liikkumista tästä vaiheesta eteenpäin.

Mahdollisia hälytyksiä laukaisevia tekijöitä voidaan syöttää järjestelmään niin, että käyttäjä on tietoinen datan vääränlaisuudesta tai muusta poikkeavasta tiedosta. Tämän jälkeen yleensä muodostetaan FTP-yhteys eli muodostetaan tiedostonsiirtokanava kahden tietokoneen välille. Tämän välivaiheen jälkeen data siirtyy myös Ruotsin tietovarastoille nähtäväksi, jossa myös itse metadataa säilytetään ja käsitellään. Kopioita datasta otetaan kahden tunnin välein ja näin ollen pystytään aina palaamaan edelliseen versioon tarvittaessa.

Tässä vaiheessa raaka data liikkuu toimipisteiltä eteenpäin kolmessa eri kategoriassa, joita ovat: ostot, myynnit ja inventaario. Tärkeimpänä kohtana tässä vaiheessa on nähtävissä arvio päivittäisestä liikevaihdosta ja raportointi siitä eteenpäin tietoa käsitteleville käyttäjille. SSAS tietokannan kuutio lisätään tässä vaiheessa arvioon ja tämän avulla voidaan tehdä erilaisia mittauksia myynneistä ja muista arvoista.



Kuvio 4. Datan alkutilanne ja lopputilanne laajemmasta näkökulmasta.

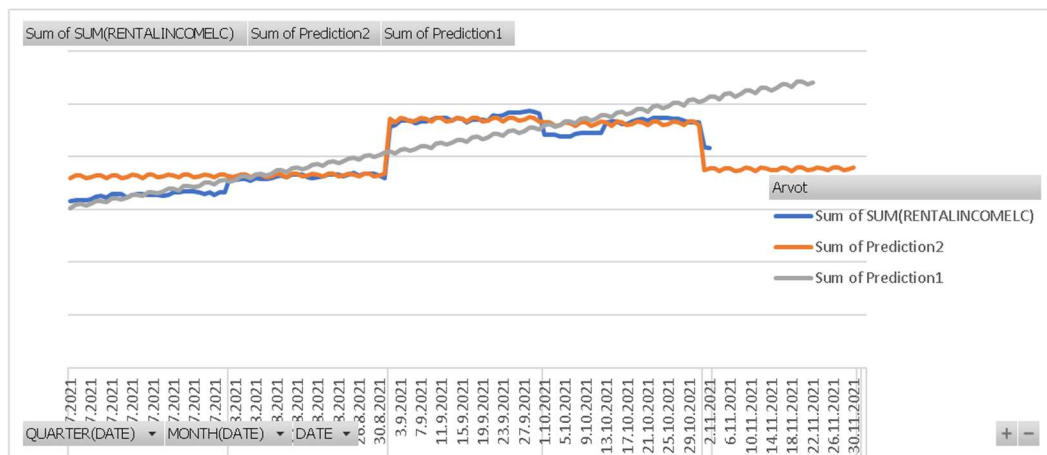
Kuviosta 4 on nähtävissä otsikoissa vasemmalta oikealle datan tietolähteet, laavastusalueet, varasto, datamarkkinat ja käyttäjät. Dataa on siis saatavissa operatiivisista järjestelmistä tai pelkästään erillisistä tiedostoista tarvittaessa. Itse datan liikkumisen jälkeen se siirtyy Exceliin analysoitavaksi ja läpikäytäväksi mahdollisesti ennustemallilla, riippuen siitä minkälaisesta datasta on tässä tilanteessa kyse. Osto- ja myyntitilastoja voidaan analysoida, kun taas liikevaihdon raportointi tehdään vain yrityksen myyntien perusteella.

5.5 Toteutus ja mahdolliset toimenpiteet

Projektin edistymisen ja jatkuvan kehityksen myötä suunnittelin ennustemallille erilaisia toteutuskuvioita, mahdollisia lisäyksiä ja toimintatapoja. Ennustemalli ja sen tuomat mahdollisuudet taloudessa ja sen hallinnassa on nähty mullistavana, tehokkaana ja ennen kaikkiaan luotettavana tekijänä. Ennustemalli rakennettiin aluksi vain vuokratuottojen arvioiden ennustamiseen. Kartoituksen myötä ennustemallia ja sen käyttömahdollisuuksia voidaan hyödyntää helposti myös budjetin, liikevaihdon, kassavirran ja monien muiden tekijöiden luotettavampaan ennustearvioon. Itse käytän myös kyseistä ennustemallia oman talouteni muutoksien ennustamiseen. Tämä kyseinen ennustemalli tunnistaa kausiluontoiset tekijät ja siihen voi syöttää jo tiedossa olevat kuukausittain toistuvat tekijät

muistiin niin, että suurin osa arvio- ja ennustekijöistä on valmiiksi tiedossa ennustemallin algoritmissa.

Ennustemalli on tällä hetkellä simulaatiokäytössä mutta empiirisen kartoituksen myötä on löytynyt uusia puolia ja käyttömahdollisuuksia tulevaisuutta ajatellen. Ennustemalli on jatkuvan kehityksen kohteena ja tarkoituksena olisi laajentaa sen käyttö simulaatiosta kansainväliseen toimintaan lähitulevaisuudessa mukaan lukien uudet käyttömahdollisuudet ja toiminnot. Uusia käyttömahdollisuuksia, joita tutkimuksessa tuli ilmi olivat: eri taloushallinnon tunnuslukujen arvioiden ennustaminen, kalenteripäivien poikkeamien/pyhien lisääminen suoraan algoritmiin ja kattavampien analyysien muodostaminen lopullisesta analyysistä.



Kuvio 5. Esimerkki vuokratuototon ennustuksesta.

Kuviossa 5 on nähtävissä sinisellä viivalla viiden kuukauden aikainen vuokratuottojen summaa kuvaava käyrä. Harmaa käyrä on nimeltään ennustus1 ja tämä ei ota huomioon käytännössä mitään poikkeavia tekijöitä, jonka takia ennustus on lineaarinen käyrä, joka noudattaa oletettua ja tasaista kasvua kuukauden loppuun asti. Oranssi käyrä on puolestaan nimeltään ennustus2 ja tässä on huomioituna suurin osa kausiluontoisista tekijöistä ja muista tekijöistä, joilla on mahdollisesti vaikutusta ennustuksen vaihteluun. Malli ei vielä pysty tunnistamaan kaikkia mahdollisia pieniä muutoksia liittyen ennustukseen, mutta kuvioista voimme nähdä oranssin käyrän noudattavan hyvin itse todellisen

summan tulosta. Ennustemalli on ottanut hyvin huomioon mahdolliset trendimuutokset ja suuren noususuhdanteen elokuun ja syyskuun vaihteessa vuonna 2021.

Ennustemallin toteutus kehittyi jatkuvasti kartoituksen yhteydessä. Ennustemalliin lisättiin erilaisia variaatioita, kuten erilaisten poikkeamien tai muiden dataa häiritsevien tekijöiden tunnistaminen. Ennustemallin on tiedostettu olevan jo sillä tasolla, että sitä voitaisiin käyttää todellisten arvioiden ennustamisessa jatkuvassa liiketoiminnassa. Ennustemalli ja sen kehittäminen on jatkuva prosessi ja muutoksia ja kehitystoimia tehdään jatkuvasti. Lähitulevaisuudessa ennustemallin on tarkoitus pystyä tunnistamaan erilaista dataa mahdollisesti eri lähteistä kuitenkin siten, että se pysyy helppokäyttöisenä ja tehokkaana välineenä talouden hallinnassa. Ennustemallia on testattu simulaatioharjoituksissa lisäten uusia variaatioita ja toimintoja sitä mukaa, kun eteen on tullut jonkinlainen ongelma. Haasteet ja hidasteet pyritään minimoimaan kehityksen myötä lähitulevaisuudessa.

5.6 Muut tutkimukset ja ennustemallien vertailu

Tämän opinnäytetyön empiiriseen kartoitukseen kuului myös erilaisten ennustemallien vertailu mukaan lukien omat tutkimukset niiden toimivuudesta ja riippuvuuksista yrityksen näkökulmasta. Finanssimaailmassa pankit ja rahoituslaitokset hyödyntävät erilaisia regressiomenetelmiä analyyseja varten, jotka perustuvat riippuvuuksien havainnointiin eri tapahtumien välillä. Tehokkain tapa analysoida dataa tapahtuu regressiomallien avulla, joka vaatii kuitenkin vahvuuksien ja heikkouksien oikeaoppisen tunnistamisen asiakkaan ja yrityksen välillä. Kausaalisuuksia eli syy-seuraussuhdetta kahden tapahtuman suhteesta on mahdollonta saada selville regressiomalleilla. Nämä keskittyvät ainoastaan riippuvuuksien löytämiseen avaten samalla mahdollisuuksia näiden ratkaisemiseen ja hyödyntämiseen analyysin näkökulmasta. (Dun & Bradstreet 2016.)

Kartoituksessa selvitettiin rahoitusallalla toimivia tekijöitä ja tällä toimialalla käytettäviä yleisimpiä ennustemalleja. Suomen Pankki hyödyntää erilaisia uusia

menetelmiä lähitulevaisuuden ennustamisessa ja talouden nykytilan seurannassa. Tutkimuksen tuloksena useat rahoitusalan toimijat mukaan lukien Suomen Pankki hyödyntävät lyhyen aikavälin ennustemallia, joita kutsutaan myös nimellä ”nowcasting” -malli. Näiden mallien tehtävänä on kerätä tietoa nopeasti muuttuvasta ympäristöstä ja lukuisista lähteistä muodostaen kattavan kokonaiskuvan talouden tilanteesta. Malli hyödyntää tilastollisia riippuvuussuhteita erilaisten muuttujien välillä ja tämän avulla pystytään ennustamaan bruttokansantuotteen kehitystä. Suomen Pankin malli on Suomen mittakaavassa suurin ja se käsittelee jopa 50 erilaista muuttujaa ja niiden yksityiskohtaisia muutoksia. (Euro ja Talous 2017.)

Finanssialalla käytetään hyvin erilaisia malleja, joista yleisempiä ovat lyhyen aikavälin ennusteen mahdollistavat ennustemallit eli toiselta nimeltään tunnetut nowcasting -mallit. Tärkeimpänä syynä tälle on ennusteen tarkkuus ja pienempi riski ennustettaessa lähitulevaisuuden arvoja. Mallit ovat kuitenkin suhteellisen samanlaisia ja pyrkivät yleensä samaan lopputulokseen tietoenkin ottaen huomioon toimialakohtaiset vaihtelut. Tässä opinnäytetyön yhteydessä rakennettu ennustemalli sulautuu hyvin jo valmiina olevien ennustemallien joukkoon. Suomen Pankin kehittämä ennustemalli on laaja ja jo lähes 50 muuttujan edestä dataa käsiteltäväksi on todella paljon. (Euro ja Talous 2017.)

Yhteinen ja toistuva tekijä lyhyen aikavälin ennusteissa on se, että ne pyrkivät etsimään muuttujien riippuvuuksia ja karsimaan kaiken sen pois, joka ei ole oleellista analyysia varten. Lyhyen aikavälin ennusteisiin on helpompi tehdä muutoksia ja korjauksia suoraan, kun taas pidemmän aikavälin ennusteisiin on yksinkertaisesti haastavampaa tehdä muutoksia jo niin pitkän ennustusvälin vuoksi. Pidemmän aikavälin ennustuksiin on mahdollista syöttää jo tiedossa olevia trendejä ja kausiluontoisia tekijöitä, jotka voivat mahdollisesti vaikuttaa datan muuttujiin ja arvoihin. Nämä yleiset trendit tai kausiluontoiset tekijät, kuten sesongit, eivät välttämättä aina käy toteen toivotulla tavalla ja näin ollen esimerkiksi jo kuusi kuukautta eteenpäin ennustettua dataa on vaikeampi lähteä etukäteen muokkaamaan ja korjaamaan. Lyhyen aikavälin ennusteet ovat yleensä

yhden kuukauden mittaisia ennusteita ja myös näin ollen yleensä tarkempia ja luotettavampia yrityksen ja asiakkaan näkökulmasta. (Euro ja talous 2017.)

Vakuutusalaan verrattuna ennustemallista löytyy samanlaisia piirteitä. Vakuutusosalalla etsitään jatkuvasti erilaisia muuttuvia riskejä, joihin reagoidaan asiakkaan näkökulmasta parhaalla mahdollisella ratkaisulla. Muuttuvia riskejä tällä hetkellä ovat megatrendit, joita pääosin ovat digitalisaatio ja kulutuskäyttäytymisen muutos. Nämä tekijät eivät ole suoranaisesti vaikutuksessa tutkimuksen ohessa tehtyyn ennustemalliin vaan siihen vaikuttaa yleensä kausiluontoiset tekijät tai muut riippuvuuksiin vaikuttavat muuttujat. Digitalisaatio voidaan nähdä mahdollisuutena tulevaisuuden vakuutusosalalle, jolloin robotiikka ja automaatio pystyy kehittämään asiakkaalle parhaan vakuutuksen tarpeiden täyttämiseksi. (Finanssialalle.)

Asiakkaan profilointia voidaan suorittaa tarkemmin myös big datan kehityksen myötä, joka tarkoittaa myös kuluttajalle riskinmukaisempaa hinnoittelua. Vakuutusyhtiöillä Suomessa on dominoivat markkinat, mutta tulevaisuudessa odotetaan nähtävän myös ulkomaalaisten vakuutusyhtiöiden tarjontaa korvaten osan suomalaisista vakuutusyhtiöistä. EU-integraatio kasvattaa vertaisvaikuttamisen osuutta vakuutusliiketoiminnassa ja myös sen myötä hallinnointi ja vakuutusyhtiöiden laajat rahoitukset hajautetaan. Vakuutusalaan vaikuttaa myös muita oleellisia tekijöitä, kuten talous, yhteiskunta ja politiikka. Vakuutusyhtiöt ovat kuitenkin suuria sijoittajia, kuten myös riskinhallinnan edistäjiä. Näin ollen muutokseen osataan varmasti reagoida ja adaptoitua etukäteen kuitenkin mahdolliset positiiviset ja negatiiviset tekijät huomioiden. Negatiivisena tekijänä tulevaisuudessa voidaan nähdä vakuutuksen saatavuus tai sen hinnan nousu liian suureksi asiakkaan näkökulmasta teknologian kehityksen myötä. Ennustemallin näkökulmasta näillä tekijöillä ei ole suoranaista vaikutusta arvioihin vaan sen tarkoituksena on ennustaa yrityksen liiketoiminnan kannalta parhaita muutoksia ja vaihtoehtoja tulevaisuuden kassavirran hallintaan. (Finanssialalle.)

Ennustemallien käyttöön liittyviä tärkeitä huomioita nousi ilmi kartoitusta tehdessä. Ennustekyvyn arvioiminen on yksi tärkeimmistä seikoista liittyen

ennustetarkkuuteen. Ennustevirheenä voidaan pitää ennusteen ja toteutuneen arvion erotusta ja paras tapa arvioida mallien ennustetarkkuutta on juuri vertailla syitä muutoksille näiden kahden arvon välillä. Hyvä ennustemalli on yleensä harhaton ja mahdollisimman tarkka. Tällöin malli ei poikkea suuresti todellisesta arvosta eikä se tuota systemaattisesti suurempia tai pienempiä ennusteita. Ennustemalli toimii yleensä paremmin mitä tuoreempaa dataa ja tietotekniikkaa on käytettävissä. Ennustemallia valitessa pitäisi näin ollen olla etukäteen jo selkeä tieto siitä minkälaista dataa on tarkoitus lähteä purkamaan ja analysoimaan, mitä kaikkia toimintoja ennustemallilta vaaditaan ja kuinka tehokkaasti tietyllä ennustemallilla voitaisiin saada ratkaisu kyseiseen ongelmaan tai kehityskohteeseen. (Euro ja Talous 2017.)

6 Tulokset

Tässä opinnäytetyön luvussa käydään läpi opinnäytetyöprosessin ja kartoituksen jälkeiset tulokset. Tulokset itsessään keskittyvät pääosin ennustemallin suoriutumiseen, mutta tuloksia oli myös saatavilla empiirisen kartoituksen ja muiden tutkimusten osalta.

Ennustemalli ja siltä vaaditut käyttömahdollisuudet toteutuivat hyvin. Malli kehitettiin siihen pisteeseen, että se pystyy käsittelemään tietynlaista dataa ottaen huomioon vaihtuvat muuttujat ja kausiluontoiset tekijät. Malli on kuitenkin vielä kehitysvaiheessa ja se ei yksinkertaisesti pysty vielä tunnistamaan erilaisia trendejä ja muita variaatioita mahdollisimman tarkan arvion tarkkuuden muodostamiseksi. Malli tarvitsee enemmän toistoja ja uutta dataa oppiakseen edellisistä ennusteista, joten tässä vaiheessa on vaikea sanoa mihin lopputulokseen mallilla olisi mahdollista päästä. Tämän osalta mallin halutut käyttömahdollisuudet toteutuivat tyydyttävästi kuitenkin ylläpitäen tietyn tason ennusteiden mahdollistamiseksi.

Data-analytiikan osalta ennustemalli osoittautui myös hyvin tehokkaaksi. Dataa pystyttiin analysoimaan lineaarisella regressiomallilla, havainnoimaan kausiluontoisten tekijöiden vaikutusta ja saamaan enemmän tietoa irti itse datasta.

Regression avulla oli nähtävissä useita tärkeitä talouden tunnuslukuja, kuten p-arvo ja standardipoikkeamat. P-arvo tarkoittaa tilastollista mittausta hypoteesin ja todellisten havaittujen arvojen välillä. Todennäköisyyttä mitataan olettaen, että nollahypoteesi on totta. Sitä suurempi on havaitun eron tilastollinen merkisevyys mitä pienempi p-arvo saadaan analyysistä. Kaiken kaikkiaan ennustemalli täytti kaikki odotukset mitä odotettiin data-analytiikan puolelta ja jatkossa siihen on mahdollista myös implementoida erilaisia analytiikan työkaluja tarpeiden mukaisesti.

Seuraavaksi opinnäytetyössä käydään läpi alussa esitetyt tutkimuskysymykset. Tietoperusta jätettiin pienemmälle tarkoituksella ja kysymyksiin pyrittiin etsimään vastaukset empiirisen kartoituksen edetessä. Vastaukset ovat myös osa omia päätelmiäni ja havainnoiteja koko opinnäytetyöprojektin mittaiselta ajankaksolta.

Mitä sovelluksia voidaan hyödyntää päivittäisen datan analysoimiseksi? Ennustemallin käyttöalustana toimii sovellus nimeltä Excel. Data varastoidaan pilvipalveluun, joka mahdollistaa sen siirron suoraan Exceliin. Mallin koodi on rakennettu ohjelmointikielellä nimeltä Python. Koodi sisältää kaikki mallin kannalta oleelliset toiminnot ja tämän avulla koodi pystytään implementoimaan ja lataamaan suoraan Exceliin datan analysoimiseksi. Yksinkertaisuudessa koodilla tarkoitetaan merkkien ja sisältöjen välisten kytkentöjen järjestelmää. Koodia pystytään muokkaamaan vaihtelevien trendien ja muiden muuttujien mukaan niin, että malli pystyy lähitulevaisuudessa olemaan entistä tarkempi ja luotettavampi todelliseen arvioon nähden.

Mistä johtuu heitto ja epätarkkuus liikevaihdon arvioinnissa? Liikevaihdon arviointi ja varsinkaan ennustaminen ei ole koskaan täysin tarkkaa ja etene oletusten mukaisesti. Projektin alussa raportoinnista saadut arviot eivät täsmänneet täysin todellisten arvojen kanssa. Ongelmaa haluttiin lähteä tutkimaan ja kehittämään pysyvämpää ratkaisua. Nopeasti ongelmaa tutkiessa huomattiin, että arviot eivät huomioi monia erilaisia tekijöitä, jotka vaikuttavat suoraan datan ulostuloon. Epätarkkuuteen liittyviä tekijöitä on varmasti olemassa lukuisia

mutta tässä rajatussa projektin aikajaksossa kirjattiin yleisimmät suoranaisesti dataan vaikuttavat negatiiviset tekijät ylös. Ensimmäisenä nostettiin esille kausiluontoisten tekijöiden ja trendien tuntemattomuus alussa. Tätä lähdettiin muokkaamaan suoraan malliin siten, että se tunnistaa nämä kausiluontoiset tekijät ja pyhät suoraan kalenterista.

Trendejä oli mahdollista lisätä malliin myös mutta jokaista vaihtelevaa trendiä oli käytännössä mahdotonta ennustaa ja niitä lisättiin tarpeiden mukaan käsin. Toisena tekijänä voidaan pitää mahdollisia datan sisältämiä virheitä jo sitä järjestelmään ladattaessa. Tämä tarkoittaa suoraan sitä, että data tulisi käsiteltäväksi jo virheellisenä. Silloin virheelle on mahdollisuus, jos epäkohtia datassa ei huomioida tarpeeksi ajoissa. Kolmantena tekijänä voidaan pitää mahdollisia virheitä itse raportoinnissa. Datan olomuoto voi muuttua sitä siirrettäessä tai oleellisia tietoja voi kadota isojen tiedostojen siirrossa useaan kertaan. Liikevaihdon arvioon heittoon ja epätarkkuuteen saatiin oleellisia korjauksia ennustemallin rakentamisen myötä ja tällä hetkellä ongelmien karsimista ja minimointia suoritetaan jatkuvasti.

Miten implementoida saatu data uusiin malleihin/kuvioihin, joita hyödynnetään tulevaisuudessa? Lähitulevaisuudessa mallia on tarkoitus käyttää päivittäisen liikevaihdon ja vuokraustuottojen ennustamiseen. Kuvioita ja analyyseja on saatavilla regressioanalyysin avulla ennustemallista. Lähitulevaisuudessa tarkoituksena on kehittää kuvioita ja malleja tarpeiden mukaan. Tällä tarkoitetaan sitä, että malli oppii tunnistamaan datan, jotta se osaa myös muodostaa tarpeita tyydyttävät kuviot ja mallit analyysejä varten. Tällä hetkellä malli on ollut vain simulaatiokäytössä ja kehityksen alla liittyen erilaisten algoritmien ja kuvioiden käyttöön. Nykyään Excel -taulukkolaskentaohjelma sisältää kattavan osan hyödyllisistä analyysityökaluista ja niiden käyttö on tehty helpoksi, kun käsittelyssä on jo valmista dataa. Mallia olisi tarkoitus myös automatisoida halutulla tavalla lähitulevaisuudessa, jotta uusien mallien ja kuvioiden teko mahdollistettaisiin tekoälyä hyödyntäen.

Millainen on ennusteprosessi, kun verrataan myynnin alkuhetkeä lopulliseen tilanteeseen? Ennusteprosessi itsessään on hyvin laaja käsite ja se on esitelty opinnäytetyössä laajemmin luvussa 4 alakappaleessa 4.1. Yksittäisen myynnin kohdalla tapahtuma on hyvin samakulkuinen verrattuna laajempaan otantaan.

Alussa tärkeintä on ongelman tunnistaminen ja tieto siitä mitä halutaan lähteä tutkimaan ja ratkaisemaan. Tämän prosessin jälkeen data ”puhdistetaan” eli se käsitellään siihen muotoon, jotta sitä voidaan alkaa hyödyntämään mallin valinnassa. Malli on hyvä valita datan perusteella ja siinä vaiheessa, kun on tietoa minkälainen analyysi on tarpeen. Ennustemallia valittaessa on tärkeää tutkia samalla mahdollisia ratkaisuja, joita tämä kyseinen ennustemalli voi tuottaa tarpeiden täyttämiseksi. Tämän jälkeen mallia sovelletaan saatuun dataan ja tehdään johtopäätöksiä lopullisista arvioista. Dataa analysoidaan halutulla tavalla ja lopulliset tulokset ja arviot lähetetään eteenpäin johtoportaalille, jossa mahdolliset toimenpiteet tehdään. Lopullisena tilanteena voidaan pitää sitä, kun negatiivisiin arvioihin reagoidaan esimerkiksi resurssien kohdistamisella muualle tai positiivisella arviolla kapasiteetin kasvattaminen tietyillä toimialoilla.

7 Johtopäätökset ja opinnäytetyöprosessi

Tässä opinnäytetyössä tarkoituksena oli luoda ennustemalli hyödyntämällä big dataa ja sen avulla saamaan luotettavuutta ja tehokkuutta liikevaihdon arvioinnin tarkkuuden kehittämiseen. Tarkoituksena oli myös löytää ongelmatekijöitä ja selvittää miten näitä voitaisiin ehkäistä ja minimoida lähitulevaisuudessa hyödyntäen ennustemallia. Tämän tyylistä ennustemallia ei ole ollut toimeksiantajalla entuudestaan käytössä ja pääsin itse myös olemaan läsnä ennustemallin rakentamisen yhteydessä. Ennustemalli ja suurin osa sen sisältämistä komponenteista olivat entuudestaan minulle täysin tuntemattomia ja pidin myös siitä, kuinka minut perehdytettiin näiden asioiden ympärille ennustemallin kehityksen myötä.

Opinnäytetyön suunnittelu aloitettiin jo syksyllä 2021. Projektista ja siihen liittyvän tiimin kasauksesta oli keskustelua vuosien 2021 ja 2022 vaihteessa.

Ennustemalliin liittyvä projekti aloitettiin pikimmiten vuoden 2022 alussa ja malli saatiin rakennettua ja viimeistelyä kahden kuukauden sisällä. Mallin rakentaminen ja kehitys tuotettiin toimeksiantajayrityksen toimesta. Olin mukana projektissa ja seurasin projektia aktiivisesti tehden samalla muistiinpanoja jo siitä, miten ennustemallia voitaisiin hyödyntää lähitulevaisuudessa ja minkälaisia analyyseja siitä olisi mahdollista saada ulos. Projektiin osallistuneiden henkilöiden yhteinen ymmärrettävä kieli oli englanti ja se tuotti myös hieman haasteita. Taloushallinnon ja ohjelmistojen entuudestaan tuntemattomat termit olivat aluksi hieman haastavia ymmärtää, mutta niiden kanssa pääsi hyvin yhteisymmärrykseen projektin edetessä.

Kartoituksen edetessä ymmärsin entistä enemmän, kuinka tärkeää talouden ja kassavirran hallinta yritykselle on. Sen merkitys on läsnä jokapäiväisessä kanssakäymisessä yrityksen ja yritysten välillä. Kassavirran ja liikevaihdon käsittely ja hallinta on suoranaisesti linkitettyä myös kirjanpitoon. Ennustemallin avulla on mahdollista antaa kirjanpidolle arviot jo seuraavan kuukauden luvuista, jotta myös heidän työnsä helpottuisi ja saataisiin parempi käsitys lähiaikoina tapahtuvista mahdollisista muutoksista ja kehityskohteista.

Liikevaihdon arvion tarkkuuden kehittäminen aiheena oli minusta erittäin kiinnostava. Aihe on haastava ja laaja itsessään mutta halusin samalla oppia uutta ja kehittää omaa taitotasoa tänä aihepiirin parissa. Ennustemallin lopputulos on mielestäni hyvä, vaikka se vaatiikin vielä paljon kehitysaskeleita ennen kuin se on valmis käsittelemään monia muita tekijöitä ja erityyppistä dataa. Näen itse ennustemallin hyvin hyödyllisenä ja käyttökelpoisena liikevaihdon, vuokraustulojen ja budjetin ennustamiseen lähitulevaisuudessa. Ennustemallista saatavat analysointityökalut helpottavat kuukausittaista raportointia ja tulevat samalla säästämään paljon resursseja ja aikaa muihin tärkeisiin osa-alueisiin.

Oman arvioni pohjalta uskon, että toimeksiantaja tulee hyötymään ennustemallista. Projektin ohessa ennustemalli saatiin sellaiseen tilaan, että sitä pystyttiin jo kokeilemaan simulaatioharjoituksissa. Ennustemalli suoriutui näistä hyvin ja toimeksiantajan mukaan malli on tarkoitus ottaa käyttöön mahdollisimman pian

todellisen datan ennustamiseen huomioiden jatkuvat kehitysmahdollisuudet. Mallia on helppo päivittää mukauttamalla algoritmia niin, että se on mahdollisimman todenmukainen huomioiden jatkuvasti vaihtelevat trendit ja muuttujat. Malli pohjautuu toistaiseksi pääsääntöisesti datan ja sen sisältämän tiedon ympärille mutta lähitulevaisuudessa on tarkoitus automatisoida monia prosesseja tehokkuuden ja luotettavuuden maksimoimiseksi.

7.1 Ennustemallien käyttö nykypäivänä

Tämän päivän talouselämässä ennusteet ovat yleistyneet huomasti ja niitä tarjotaan yrityksille yhä useamman ulkopuolisen toimijan kautta. Ennustemalleja kaupitellaan paremman tulevaisuuden toivossa yrityksille luoden yhden vaadittavista lisäyksistä talouden ennustamiselle. Loppujen lopuksi ennustemallit ovat kuitenkin vain ennustuksia ja tämän pohjalta ei ole viisainta ryhtyä toimiin välittömästi ennusteen pohjalta, ennen kuin tarpeeksi tietoa ja dataa on pystytty käsittelemään luotettavasti.

Tehdessäni kartoitusta tuli mieleeni muutama tärkeä skenaario, jotka olisi hyvä rakentaa lopullisen ennusteen pohjalta. Nämä skenaariot ovat paras vaihtoehto, keskivertainen vaihtoehto ja huonoin vaihtoehto. Parhaassa tapauksessa kaikki odotukset toteutuvat ja silloin ollaan myös ihannetilanteessa tulevan kannalta. Keskivertainen vaihtoehto on nähtävissä silloin, kun parhaasta vaihtoehdosta poistetaan suurimman onnistumiset. Keskivertaisesta ennusteesta yleensä muodostetaan virallinen ja lopullinen ennuste. Huonoimmassa vaihtoehdossa lopullisesta ennusteesta otetaan pois myös niin sanotut normaalit ja varmat tapahtumat, jotta nähdään mikä tilanne voisi olla huonoimmillaan. Näiden kolmen skenaarion pohjalta voidaan luoda keskiarvo virallisesta ennusteesta, joka useimmiten peilaa keskivertaisen tilanteen kanssa. Lopullisen ja tarkan ennusteen saaminen lisää huomattavasti työntekijöiden osaamista ja motiivia tulevien ennusteiden muodostamiseen organisaatiossa.

Nykypäivän ennustemalleista suurin osa sisältää paljon tai jopa kokonaan automatisoituja prosesseja ja niiden pohjalta pyritään saamaan ennusteet yhä

luotettavimmiksi. Joissain tapauksissa automatisointi ei välttämättä aina toteudu oletetusti esimerkiksi datan muuttumisen myötä. Ennustemalli voi olla rakennettu tietynlaisen datan pohjalle ja seuraavan kuukauden datan sisältäessä muuttujia, voi ennustemalli ennustaa väärin tai jättää tiettyjä seikkoja huomioimatta. Ihmisen hallitsemaan ennustemalliin voidaan tehdä korjauksia samantien ongelman noustua ilmi, mutta robotti pyrkii saamaan ennustuksen läpi huomioimatta virhettä, ellei malliin ole integroitu virheenilmoitustyökalua. Näin ollen automatisaation kehityttyä ennustemalleista tulee tehokkaampia ja luotettavia mutta tällä hetkellä sillä on myös omat negatiiviset puolensa vaihtelevien muutujien kannalta.

7.2 Päätelmät ennustemallin heikkouksista ja hyödyistä

Ennustemallia itsessään sisältää lukuisia etuja ja heikkouksia. Heikkouksia on yleensä mahdollista minimoida tai eliminoida kokonaan. Kappaleessa 4.4. kävin jo muutamat oleelliset käyttömahdollisuudet ja haasteet läpi. Kuitenkin tutkimusta tehdessä näin todellisen kuvan siitä, miten ennustemalli käyttäytyi ja mitkä tekijät nousivat yleisemmiksi eduiksi ja heikkouksiksi.

Ensimmäisenä etuna halusin nostaa esille mahdollisuuden saada arvokasta tietoa. Ennustemalli ei tule aina osumaan täysin kohdalleen todellisen arvion kanssa ja osuessaan mukana on ollut hyvin paljon tuuria. Tämän takia ennuste pelkästään sisältää jo hyvin paljon tietoa yritykselle. Yritys saa kattavamman kuvan niistä seikoista, joilla on ollut vaikutusta yrityksen arvioon positiivisella tai negatiivisella panoksella. Tietty toimipiste on voinut esimerkiksi raportoida todella huonosta tuloksesta kuukaudelta ja ennusteen avulla on mahdollista nähdä myös, onko sen pitäminen tai resurssien sijoittaminen kannattavaa enää jatkossa. Ennusteen avulla totutaan myös katsomaan menneitä ja reaaliaikaista dataa tulevan kysynnän ennustamiseksi. Kysynät vaihtelut, kuten esimerkiksi sesongit, trendit tai muut tekijät pystytään ennakoimaan tehokkaammin. Tärkeimpänä seikkana voidaan pitää mahdollisuutta suunnata resursseja muualle tai tehdä konkreettisia muutoksia ennusteen avulla.

Toisena etuna näin kyvyn oppia aikaisemmista virheistä. Ennusteet parantuvat ja kehittyvät jatkuvasti vain ainoastaan niitä tekemällä. Edellisen kuukauden ennustetta tutkimalla pystytään karsimaan virheet ja mahdolliset haittatekijät seuraavan kuukauden ennustetta varten. Kolmantena etuna näin mahdollisuuden alentaa kuluja yrityksessä. Oikein tehtynä ennuste mahdollistaa prosessin muokkauksen tehokkuuden lisäämiseksi organisaatiossa. Monilla yrityksillä asiakkaat ovat kuitenkin pääprioriteetti, ja ennusteiden avulla on helpompi määrittää mitä ja milloin asiakkaat haluavat jotain. Tiedetyt tuotteet myyvät enemmän kuin toiset ja siksi esimerkiksi varastokapasiteettia on hyvä pienentää vähemmän myyvien tuotteiden osalta yleisen kannattavuuden kehittämiseksi.

Ennustemallin heikkoudet peilautuvat yleensä suoraan malliin. Mallin sisältämät komponentit ja toimintamahdollisuudet kertovat yleensä myös sen mahdollisista heikkouksista. Yksi asia, johon malli itsessään ei voi vaikuttaa on sen tarkkuus. Mikään malli ei suoranaisesti pysty vielä ennustamaan sadan prosentin tarkkuudella, mutta sitä lähemmäs on mahdollista päästä erilaisilla muutoksilla. Toisena heikkoutena, jonka halusin nostaa ennustemallista, on siihen kuuluva aika ja resurssikäyttö. Malli pitäisi pystyä tehdä niin käyttäjäystävälliseksi ja aikatehokkaaksi kuin mahdollista, jotta tältä pystyttäisiin välttymään. Tämä ei kuitenkaan ole aina mahdollista ja huomasin myös itse, että joidenkin analyysien tekemiseen ja datan kasaamiseen kului huikea määrä aikaa.

Omasta mielestäni datan käsittelyä ja analysointia varten pitää olla siihen määrätty tiimi, jotta resursseja ja aikaa kuluu mahdollisimman vähän. Koordinaatio on yksi tärkeimmistä seikoista analyysia tehdessä, jotta kaikki pysyvät samalla polulla ja ylimääräisiltä virheiltä pystytään välttymään. Heikkoutena voidaan pitää myös malliin keskitettyjä kuluja. Tiimin palkkaaminen ja parempiin sovelluksiin sijoittaminen tarkoittaa automaattisesti enemmän kuluja. Tämä voi tuntua vaikealta sijoitukselta mutta yleensä ennustemalli pystyy kattamaan kulut pidemmällä tähtäimellä. Tähän ei kuitenkaan ole olemassa mitään takuuta sijoitusta tehdessä ja se on yksinkertaisesti vain johdon päätettävissä, kannattaako riskiä ottaa.

8 Yhteenveto

Tämän opinnäytetyön tarkoituksena oli löytää keinoja kehittää liikevaihdon arvioinnin tarkkuutta. Tarkkuuden kehittämiseen hyödynnettiin big data analyysia. Luotettavuuden, tehokkuuden ja tarkkuuden parantamiseksi rakennettiin ennustemalli, joka painottui pääsääntöisesti liikevaihdon ja vuokratuottojen ennustamiseen kuukausitasolla. Projekti ennustemallin ympärille osoittautui onnistuneeksi ja ennustemalli on tarkoitus ottaa käyttöön lähitulevaisuudessa, jotta tietotaso tulevien kuukausien arvioista olisi luotettavampaa ja jo entuudestaan tiedossa. Seuraavissa alakappaleissa käydään läpi data-analytiikan ja ennusteiden osuutta tänä päivänä sitä hyödyntävien yritysten perspektiivistä.

8.1 Data-analytiikan reagoiminen teknologian jatkuvaan kasvuun

Teknologian kehitys on tällä hetkellä suurimmassa nousussa mitä sen ollut koskaan aikaisemmin historiassa. Kehitys voidaan nähdä positiivisena sekä negatiivisena tekijänä lähitulevaisuudessa. Tekoälyn, robottien ja älylaitteiden lisääntyessä on kehityksessä mukana pysyminen entistä haastavampaa. Teknologian kehitystä on nähtävissä jatkuvana paineena vaikuttaen yhteiskuntaan, opiskeluun, talouteen ja työelämään. Ihmisen rooli teknologian kehityksen myötä työelämässä on ollut pitkään pohdinnan alla useiden tutkijoiden keskuudessa. Tuleeko automatisaatio ja robotit korvaamaan ihmisen kokonaan ja minkälaisiin työtehtäviin nämä muutokset pohjautuvat eniten? Teknologian kehityksen myötä talouden hallinta ja varsinkin data-analytiikka tulee muuttumaan radikaalisti sovellusten ja työtehtävien parissa. (Kaleva 2019.)

Tutkimuksen myötä näen data-analytiikan tulevaisuuden hyvin valoisena. Sovellukset kehittyvät, kyberturvallisuus on entistä luotettavampaa ja tietoa saadaan eteenpäin entistä nopeammin ja resurssitehokkaasti. Tietyt manuaalista työtä vaativat työtehtävät voidaan korvata tekoälyllä ja automatisaatiolla ja ihmiselle jää itse datan analysoiminen. Päätöksientekoon en näe tekoälyn vielä kykenevän ja loppujen lopuksi tekoälyllä ei vielä tänä päivänä ole älyllistä taitoa tehdä päätöksiä sisältäen riskejä. Eettisyys on myös yksi tärkeä tekijä

lähitulevaisuudessa. Tekoälyllä ei ole empatiaa tai muutakaan kykyä asettua toisen asemaan ja tiedostamaan mikä on oikein ja väärin missäkin tilanteessa. Mielestäni teknologian kehityksessä on oltava valppaana jatkuvasti ja pyrittävä kouluttamaan työvoimaa uusiin tai vaihtuviin työtehtäviin niin, että teknologian kehitys ei tule yllätyksenä missään vaiheessa. Näitä vaiheita seuraten teknologian kehitys myös talousalalla on nähtävissä kaiken kaikkiaan positiivisena ja motivoivana tekijänä työtehtävien muuttuessa resurssien jakamisen ja tekoälyn myötä.

8.2 Ennusteet datan luotettavuuden tueksi

Ennusteet itsessään eivät ole uusi käsite, sillä esimerkiksi bruttokansantuotteen vaihtelun ennustamista ja sääennusteita on tehty jo pidemmältä ajalta. Talouden hallinnassa ennusteet ovat kuitenkin uudempi käsite ja yhä yleistymässä. Ennustemalleja on otettu käyttöön yhä useammassa yrityksessä lähinnä turvana tulevalle. Yritykset haluavat olla tietoisempia tulevasta ja reagoida muutoksiin hyvissä ajoin. Ennusteisiin ei voi luottaa tietenkään täysin, mutta ne antavat yleensä oikein tehtynä hyvän kuvan lähitulevaisuuden arvioista.

Tämän takia luotettavuus on ollut yksi tärkeimmistä kohdista, kun tarkastellaan ennustemallin suoriutumista. Luotettavuudelle tänä päivänä ei voida osoittaa liikaa painoarvoa mutta tutkimuksen edetessä ennustemallin luotettavuudesta sai paremman käsityksen. Mallin sisältäessä tarpeelliset toiminnot ja komponentit, voidaan sitä pitää entistä luotettavampana. Tutkimuksen aikana ennustemallissa ajettu data poikkesi vain vähän todellisista arvoista. Ennustemalli otti huomioon kausiluontoiset tekijät ja muut vaihtelevat muuttujat. Mahdolliset raportointivirheet ja datan poikkeavuudet oli korjattavissa käsin siihen muotoon, että tietoa oli mahdollista analysoida.

Yleisesti ottaen big data -analytiikka on hyvin luotettavaa, sillä data perustuu asiakkaiden tai organisaation toiminnoista kerättyyn tietoon. Käyttömieltyyksistä ja entistä paremmin myyvistä tuotteista saadaan olennaista tietoa, johon on helpompi reagoida. Big dataa kannattaakin jokaisen yrityksen hyödyntää,

jotta yritys ja sen toiminnot olisivat myös asiakkaan näkökulmasta mieluisemat. Ennustemalli taas on enemmän hyödyllinen yrityksen sisäiselle toiminnalle, kuten talouden hallinnalle ja monitoroinnille. Mielestäni tämän avulla yritys on kokoajan askeleen edellä todellista tilannetta ja pystyy tekemään päätöksiä estääkseen katastrofin tai kohentaakseen asemaansa tietyillä toimialoilla.

8.3 Validiteetti ja reliabiliteetti

Tämä opinnäytetyö on pyritty toteuttamaan mahdollisimman tarkasti ja luotettavasti. Tässä alaluvussa arvioidaan tämän opinnäytetyön validiutta ja reliabiliteettiä.

Validiteetilla tarkoitetaan tutkimuksessa käytetyn mittausmenetelmän pätevyyttä. Sen avulla saadaan selville, onko tutkimuksessa tutkittu tarkoituksen mukaisia haluttuja asioita. Reliabiliteetti kuvastaa taas tutkimuksen luotettavuutta. Kun tutkimus voidaan toteuttaa toistuvasti kahden eri henkilön toimesta, voidaan puhua tutkimuksen reliabiliteetin olevan kunnossa. (Tilastokeskus.)

Tässä opinnäytetyössä esitellyt toimenpiteet kehityshankkeen mukaisesti on esitelty tarkasti, jotta niistä saatiin mahdollisimmat luotettavat tulokset. Ensimmäisenä tutkimusmenetelmänä hyödynnettiin laskennallisia ja tilastollisia menetelmiä, algoritmeja ja kertoimia tutkittavan asian ratkaisemiseksi. Kvantitatiivisella analyysillä on yleensä korkea reliabiliteetti ja vastaukset eivät perustu mielipiteisiin. Laskennallisten menetelmien tutkimisen tavoitteena oli löytää ratkaisuja epätarkkuuksien ja heittojen minimoimiseksi liikevaihdon arvioinnissa.

Toisena menetelmänä hyödynnettiin opinnäytetyöprojektin aikana toimeksiantajayrityksen kanssa yhteistyössä rakennettua ennustemallia. Tämän mallin avulla saatiin enemmän vastauksia jo aikaisemmin vaivanneisiin kysymyksiin ja arviot saatiin lähemmäs totuudenmukaisia arvoja. Näiden tutkimusmenetelmien käyttö yhdessä on taannut luotettavat tutkimustulokset. Ennustemalli täsmensi ensimmäistä menetelmää ja yhdessä tuloksista saatiin luotettavia. Opinnäytetyössä on selvitetty oikeita ja oleellisia asioita aiheeseen liittyen.

Lähteet

Beers, Brian 2022. P-Value. Investopedia. Päivitetty maaliskuu 2022. <https://www.investopedia.com/terms/p/p-value.asp>. Viitattu 9.4.2022.

Chowdhury, Madhurjya 2021. Big Data Analytics Benefits: How To Analyse Big Data? Analytics Insight. Päivitetty 27.7.2021. <https://www.analyticsinsight.net/big-data-analytics-benefits-how-to-analyse-big-data/>. Viitattu 20.11.2021.

Dun & Bradstreet 2016. Luottoriskien ennustaminen datan täyttämässä maailmassa. <https://www.dnb.com/fi-fi/syvenna-osaamistasi/ajatuksiamme/luottoriskien-ennustaminen-datan-tayttamassa-maailmassa/>. Viitattu 25.3.2022.

Enterprise big data framework 2019. Where does “Big data” come from? <https://www.bigdataframework.org/short-history-of-big-data/>. Viitattu 8.1.2022.

Euro ja talous 2017. Kansantalouden seurantaan uusia välineitä. Päivitetty 18.12.2017. <https://www.eurojatalous.fi/fi/2017/5/kansantalouden-seurantaan-uusia-valineita/>. Viitattu 23.3.2022.

Euroopan parlamentti. Mitä tekoäly on ja mihin sitä käytetään? Päivitetty 29.3.2021. <https://www.europarl.europa.eu/news/fi/headlines/society/20200827STO85804/mita-tekoaly-on-ja-mihin-sita-kaytetaan>. Viitattu 5.1.2022.

Finanssialalle. Vakuutusalan tulevaisuus. <https://www.finanssialalle.fi/opintomateriaalit/finanssialan-perusteet/vakuuttaminen/vakuutusalan-tulevaisuus.html>. Viitattu 25.3.2022.

Forbes 2020. Your big data responsibility: The Rise in data ethics. <https://www.forbes.com/sites/forbestechcouncil/2020/06/08/your-big-data-responsibility-the-rise-in-data-ethics/?sh=6d26953672f7>. Viitattu 10.1.2022.

Hanke, John & Wichern, Dean 2014, Business forecasting, 9. painos, Pearson Education Limited, Harlow, England. Sivü 510.

Hillier, Will 2021. Careerfoundry. Is big data dangerous? <https://www.estuate.com/company/blog/content/are-you-fighting-5-biggest-risks-big-data>. Viitattu 3.12.2021.

Investopedia. File Transfer Protocol (FTP). Päivitetty marraskuu 2021. <https://www.investopedia.com/terms/f/ftp-file-transfer-protocol.asp>. Viitattu 17.3.2022.

Kalyvas, James & Overly, Michael 2014. Big data. A business and legal guide. Auerbach Publications 2014, Incorporated 2014, The United States of America.

Kvantitatiivisen tutkimuksen verkkokäsikirja. Regressioanalyysi. Päivitetty joulukuun 2008. <https://www.fsd.tuni.fi/menetelmaopetus/regressio/analyysi.html>. Viitattu 5.2.2022.

Kvantitatiivisin tutkimuksen verkkokäsikirja. Regressioanalyysi. SPSS-harjoitus 1. Päivitetty lokakuun 2014. <https://www.fsd.tuni.fi/menetelmaopetus/regressio/harjoitus1.html>. Viitattu 5.2.2022.

Limnell, Jarno 2019. Kaleva. Teknologian radikaali kehitys haastaa ihmistä ja luottamusta. <https://www.kaleva.fi/teknologian-radikaali-kehitys-haastaa-ihmista-ja-l/1705164>. Viitattu 12.4.2022.

Lumen. Module 10: Inference for Means. Distribution of Sample Means. <https://courses.lumenlearning.com/wmopen-concepts-statistics/chapter/distribution-of-sample-means-3-of-4/>. Viitattu 5.2.2022.

Lähdemäki, Sakari 2021. Toimialakohtainen ennustaminen. Labore. <https://labore.fi/t&y/toimialakohtainen-ennustaminen/>. Viitattu 23.3.2022.

Marr, Bernard 2015. Big data. John Wiley & Sons, Incorporated.

Ohlhorst, Frank J. 2012. Big data analytics. Turning big data into big money. John Wiley & Sons, Incorporated 2012, The United States of America.

Peda 2021. Ohjelmointi perusopetuksessa. Mitä on algoritmien ajattelu? <https://peda.net/jyu/it/koulutusteknologia/op/keos-2017/aa/moaa>. Viitattu 6.1.2022.

Ramirent 2021. Yritys. Ramirent Suomessa. <https://www.ramirent.fi/tietoa-ramirentista/yritys>. Viitattu 22.12.2021.

Statista 2021. Revenue from big data and business analytics worldwide from 2015 to 2022. <https://www.statista.com/statistics/551501/worldwide-big-data-business-analytics-revenue/>. Viitattu 7.12.2021.

Talend 2021. Big Data and Privacy: What Companies Need to Know to Ensure Trusted Data. <https://www.talend.com/resources/big-data-privacy/>. Viitattu 5.12.2021.

Talouselämä 2013. Big data muuttaa maailmaa. <https://www.talouselama.fi/kumppaniblogit/big-data-muuttaa-maailmaa/6e3988d0-e07e-35ea-b52c-dc3e31a91394>. Viitattu 22.12.2021.

TechTarget 2021. Big data analytics. Päivitetty helmikuu 2021. <https://search-businessanalytics.techtarget.com/definition/big-data-analytics>. Viitattu 20.11.2021.

Tilastokeskus 2020. 4. Big data. Päivitetty 3.12.2020. https://www.stat.fi/til/ichte/2020/ichte_2020_2020-12-03_kat_004_fi.html. Viitattu 3.12.2021.

Tilastokeskus. Liikevaihto. <https://www.stat.fi/meta/kas/lvaihto.html#tab1>. Viitattu 3.1.2022.

Tilastokeskus. Reliabiliteetti. <https://www.stat.fi/meta/kas/reliabiliteetti.html>. Viitattu 23.4.2022

Tilastokeskus. Validiteetti. <https://www.stat.fi/meta/kas/validiteetti.html>. Viitattu 23.4.2022.

Yrityksen-perustaminen.net. Yrityksen liikevaihto. <https://yrityksen-perustaminen.net/liikevaihto/>. Viitattu 3.1.2022.