



Yrityksen maksuhäiriöitä ennustavan datapohjaisen työkalun kehitys

The development of a data-based tool for predicting corporate defaults

Lappeenrannan–Lahden teknillinen yliopisto LUT

Kauppatieteiden kandidaatintutkielma

2021

Viola Valtonen

Tarkastaja: Tutkijatohtori Jyrki Savolainen

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Kauppatieteet

Viola Valtonen

Yrityksen maksuhäiriöitä ennustavan datapohjaisen työkalun kehitys

Kauppatieteiden kandidaatintutkielma

2021

32 sivua, 3 kuvaa, 1 taulukko

Tarkastaja: Tutkijatohtori Jyrki Savolainen

Avainsanat: maksuhäiriö, luottoriski, riskimallinnus, logistinen regressio

Tässä työssä käsitellään yritysten maksuhäiriötekijöitä ja tutkitaan yritys X:n riskimallinnuksen työkalun kehitysmahdollisuuksia. Yritys X:n työkalu on riskimalli, joka on kehitetty logistisella regressiolla. Riskimallin muuttujia verrataan kirjallisuudessa esiintyviin maksuhäiriötekijöihin, ja mallin kehitysmahdollisuuksia tutkitaan lisättävien muuttujien kautta. Laadullinen tutkimus kysynnästä ja mielenkiinnosta riskimallin räätälöintiin toteutetaan puolistrukturoidun haastattelun kautta.

Tämän tutkimuksen mukaan kirjallisuuden yleisimmät maksuhäiriötekijät perustuvat pitkälti taloudelliseen ja määrälliseen tietoon. Yrityksen maksuhäiriöihin vaikuttavat muun muassa yrityksen kannattavuus, likviditeetti, pääomarakenne, toimiala, yrityksen koko ja ikä. Nämä tekijät ovat kiitettävästi mukana yritys X:n riskimallissa, mutta kehityksen kohteeksi jää esimerkiksi maksukäyttäytyminen ja maksuhäiriöhistoria.

Logistista regressiota tutkitaan menetelmänä maksuhäiriön ennustamisessa, ja samalla kartoitetaan yritys X:n riskimallin päivittämismahdollisuuksia uuden muuttujan lisäämisen suhteen. Kirjallisuus yritysriskimallien päivittämisestä on vähäistä, mutta lääketiede tarjoaa vaihtoehtoisia ratkaisuja. Päivittämismenetelmät riippuvat lisättävän datan rakenteesta. Lähtökohtaisesti mallin päivittämiseen suositellaan mallin uudelleenestimointia, mutta uuden datajoukon ollessa pieni, hillityt päivittämismenetelmät vähentävät mallin ylisovittamista. Yritys X:n riskimallin päivittäminen voi olla mahdollinen esimerkiksi mallin kalibroinnin, valikoivan uudelleenestimoinnin ja laajentamisen kautta.

Haastattelusta paljastuu kysyntää ja mielenkiintoa yritys X:n riskimallin kehittämiseen. Kaksi yritys X:n työntekijää kertovat asiakkaiden olevan kiinnostuneita esimerkiksi maksuviiveen ja maksuhäiriöhistorian muuttujista, ja he ehdottavat malliin enemmän reaaliaikaisuutta.

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

School of Business and Management

Business Administration

Viola Valtonen

The development of a data-based tool for predicting corporate defaults

Bachelor's thesis

2021

32 pages, 3 figures, 1 table

Examiner: Postdoctoral researcher Jyrki Savolainen

Keywords: default, credit risk, risk modeling, logistic regression

In this Bachelor thesis, corporate default factors are studied, and the development possibilities of case company X's risk modeling tool are investigated. The company X tool in question is a risk model that is built on a logistic regression. The variables of the risk model are compared to company default factors found in current literature. Development possibilities of the risk model are studied by investigating methods of adding new variables. Qualitative research on possible demand of the developed company X risk model is conducted through a semi-structured interview.

According to current literature, corporate default factors are largely based on financial and quantitative information. Examples of corporate default factors are profitability, liquidity, capital structure, industry, company size and age. These factors compare very well to the company X risk model's factors. Factors missing in the company X model include payment behavior and default history, and these factors remain as potential development.

Logistic regression is studied as a statistical method to predict corporate defaults. The possibility of updating the company X risk model with a new variable is investigated. The literature on updating corporate credit risk models is limited but medicine offers alternative solutions. Update methods depend on the structure of the data to be added. Re-estimation of the model is generally recommended, but when the new data is small, parsimonious methods prevent overfitting in the model. It may be possible to update the company X's risk model for example through calibration, selective re-estimation and extension of the model.

The interviews conducted reveal interest and customer demand for the developed risk model. Two employees from company X report that customers would be interested in variables such as payment behavior and default history. There is also a suggestion of more real-time risk modeling.

SISÄLLYSLUETTELO

1.	JOHDANTO.....	1
1.1	Tutkimuksen taustaa	1
1.2	Aikaisempi tutkimus	2
1.3	Tutkimuksen tavoitteet, tutkimuskysymykset ja rajaukset	2
2.	LUOTTORISKI JA MAKSUHÄIRIÖT	4
2.1	Luotto ja luottoriski yrityksissä	4
2.2	Maksuhäiriö ja maksuhäiriön todennäköisyys.....	5
2.2.1	Maksuhäiriön ennustamisen työkaluja.....	6
3.	LOGISTINEN REGRESSIO JA SEN PÄIVITTÄMINEN	8
3.1	Logistinen regressio	8
3.1.1	Opettamis- ja validaatiodata	9
3.2	Aiempi tutkimus logistisen regression päivittämisestä	10
3.2.1	Mallin kalibrointi, korjaaminen ja uudelleen estimointi	11
4.	MAKSUHÄIRIÖTEKIJÄT YRITYKSISSÄ.....	13
4.1	Yritysten luokittelu	13
4.2	Taloudelliset maksuhäiriötekijät.....	14
4.3	Muut maksuhäiriötekijät.....	15
4.4	Yritys X:n riskimallin maksuhäiriötekijät.....	18
5.	YRITYS X:N RISKIMALLIN PÄIVITTÄMINEN	19
5.1	Riskimallinnuksen räätälöinti ja yritys X.....	19
5.2	Otannan ja muuttujien vaikutus päivittämismenetelmän valintaan	20
5.3	Päivittämisvaihtoehdot	21
5.3.1	Mallin laajentaminen	22
5.4	Yritys X:n vaihtoehdot	24
6.	MIELENKIINTO YRITYS X:N RISKIMALLIN KEHITTÄMISEEN	26
6.1	Tutkimusmenetelmä ja aineisto	26
6.2	Tutkimuksen tulokset	27
7.	JOHTOPÄÄTÖKSET.....	30
	LÄHDELUETTELO.....	33

Kuvaluettelo

Kuva 1. Tutkimuksen rakenne

Kuva 2. Logistisen regression päivittämismenetelmät

Kuva 3. Riskinmallinnuksen kolmiportaisuus

Taulukkoluetelo

Taulukko 1. Yrityksen maksuhäiriön todennäköisyyteen vaikuttavat tekijät

1. JOHDANTO

1.1 Tutkimuksen taustaa

Yrityksen liiketoimintaan liittyy aina riskejä, jotka on otettava huomioon yrityksen toiminnassa. Luottoriski on tärkeä osa-alue riskienhallinnassa taloudellisten tappioiden välttämiseksi. Muun muassa yrityksen likviditeetti eli maksuvalmius voi olla asiakkaiden maksuista riippuvainen. Tämä työ käsittelee luottoriskiä ja maksuhäiriöitä yleisesti sekä yritys X:n tapauksessa. Tässä työssä luottoriski määritellään todennäköisyytenä siitä, että luottosopimuksen vastaanottava osapuoli ei kykene noudattamaan maksu- tai muita ehtoja. Yritysten luottoriskin hallintaan kuuluu olennaisesti maksuhäiriöiden ennustaminen, jota toteutetaan eri menetelmin. Tässä työssä paneudutaan yritys X:n kaupalliseen riskimallinnuksen työkaluun, joka ennustaa maksuhäiriön todennäköisyyttä. Kyseinen riskimalli pohjautuu logistiseen regressioon, joka on tilastollinen menetelmä, jossa tutkitaan tiettyjen tekijöiden vaikutusta tutkittavaan ilmiöön. Logistisessa regressiossa tutkittava ilmiö voi saada vain kaksi arvoa - tässä tapauksessa joko maksuhäiriö tai ei maksuhäiriötä.

Tässä työssä paneudutaan maksuhäiriöiden tärkeimpiin tekijöihin ja tutkitaan yritys X:n riskimallin kehittämismahdollisuuksia arvioimalla riskimallin päivittämismenetelmiä sekä haastatteleamalla yritys X:n työntekijöitä. Mahdollisiin kehittämistoimiin sisältyy riskimallin päivittäminen asiakastarpeisiin. Luottoriskien hallinta ja etenkin maksuhäiriöiden ennustaminen on tärkeää jokaisen yrityksen toiminnan kannalta. Yrityksen asiakkaiden maksuhäiriöt vaikuttavat yrityksen rahavirtaan, likviditeettiin ja täten myös heidän tunnuslukuihinsa. Maksuhäiriötekijöiden ymmärtäminen auttaa yritysjohtoa muun muassa kartoittamaan asiakkaitaan sekä suunnittelemaan omaa liiketoimintaansa. Logistisen regressiomallin ja sen päivittämismenetelmien tutkiminen auttaa kyseistä mallia käyttäviä tahoja ymmärtämään päivittämisen mahdollisuuksia. Jos mallin uudelleenestimointi eli uudelleen rakentaminen ei osoittaudu tarpeelliseksi, voi tämä tutkimus luoda edellytyksiä muiden päivittämismenetelmien testaamiseksi. Logistisen regressiomallin käyttö voi parhaassa tapauksessa tehostua ja räätälöityä käyttäjille.

1.2 Aikaisempi tutkimus

Luottoriskiä on tutkittu paljon tiedekirjallisuudessa. Ensimmäiset empiiriset tutkimukset konkurssien ennustamisesta ja maksukyvyttömyydestä ajoittuvat 1960-luvulle, jolloin Altman kehitti Z-mallin, jota sovelletaan vielä tänäkin päivänä (Salimi 2015, 233). Maksukyvyttömyys onkin vahvasti sidoksissa yrityskonkursseihin (Koulafetis 2017, 2). Maksuhäiriötekijöitä on tutkittu kiitettävästi kirjallisuudessa, ja tämä luo hyvän tutkimuspohjan kirjallisuuden tekijöille sekä mahdollistaa tekijöiden vertailun yritys X:n riskimalliin. Logistinen regressio on myös tilastollisena menetelmänä paljon käytetty ja tutkittu. Kuitenkin, logistisen regression yritysriskimallin päivittäminen asiakastarpeisiin on tämän tutkimuksen teon hetkellä uusi tutkimusaihe. Logistisen regressioon päivittämisestä löytyy lääketieteen kirjallisuutta, jossa tyypillisesti ennustetaan sairauksien todennäköisyyksiä. Luottoriskin hallinnassa regressiomallin päivittämistä ei kirjallisuudessa ole käsitelty laajemmin, ja tämä luo tutkimusaukon, jota hyödynnetään tässä työssä. Suppean kirjallisuuden takia aiheen tutkiminen tapahtuu teoreettisella tasolla, ja mahdollisia johtopäätöksiä tehdään hyvin varovaisesti. Kirjallisuuskatsaus aiheeseen sisältää sekä lääketieteen että luottoriskin tutkimuksia ja esimerkkejä, ja soveltamista alojen välillä tehdään kriittisesti. Vähäinen kirjallisuus tekee aiheesta kuitenkin hyvin tärkeän ja mielenkiintoisen tutkimuskohteen.

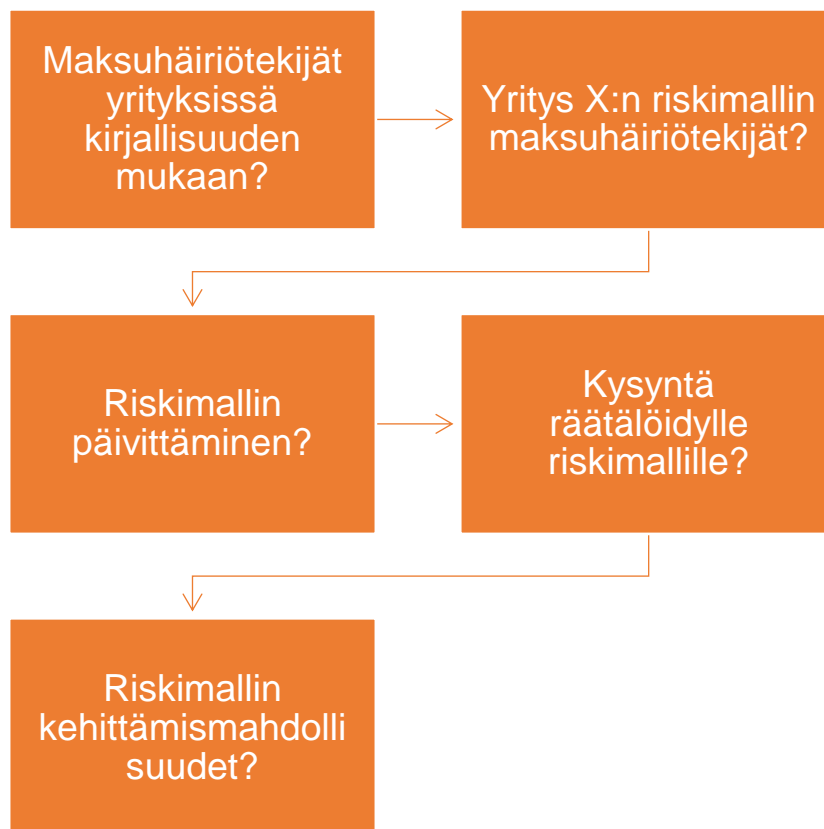
1.3 Tutkimuksen tavoitteet, tutkimuskysymykset ja rajaukset

Tämän tutkimuksen tavoitteena on selvittää yritysten tärkeimmät maksuhäiriötekijät kirjallisuuden avulla, vertailla maksuhäiriötekijöitä yritys X:n riskimalliin, tutkia logistista regressiota ja sen päivittämismahdollisuuksia riskienmallinnuksessa sekä selvittää mielenkiintoa yritys X:n riskimallin räätälöintiin. Näistä tavoitteista muodostuivat seuraavat tutkimuskysymykset:

- 1. Mitkä tekijät vaikuttavat yrityksissä maksuhäiriöihin kirjallisuuden mukaan?*
- 2. Onko yritys X:n riskimallissa huomioitu kirjallisuudesta löytyneet maksuhäiriöihin vaikuttavat tekijät?*
- 3. Miten on mahdollista lisätä uutta dataa toimivaan logistiseen regressiomalliin?*

4. ”Millaista mielenkiintoa ja kysyntää on odotettavissa yritys X:n riskimallin räätälöintiin?”

Tutkimus rajautuu luottoriskien ja maksuhäiriöennustamisen tutkimiseen. Logistista regressiota tutkitaan riskienhallinnan työkaluna. Markkinarajaus on yritysmarkkina, ja kuluttajamarkkinoiden luottoriskien ei perehdytä. Maantieteellinen rajaus on Suomi. Yritys X, joka on suomalainen yritys, toimii esimerkkinä tapauksessa, jossa yritysriskimallin päivittämistä pohditaan. Tutkimuksen aihe on tärkeä, sillä tutkimusta aiheesta on tehty hyvin vähän. Tuloksia voidaan mahdollisesti hyödyntää luottoriskien hallinnan riskimalleja kehitettäessä. Tutkimuksen rakenne esitetään kuvassa 1.



Kuva 1. Tutkimuksen rakenne

2. LUOTTORISKI JA MAKSUHÄIRIÖT

Tässä osuudessa käydään läpi yritysten luottoriskejä ja maksuhäiriöitä tutkimalla aikaisempaa kirjallisuutta ja tutkimuksia aiheista. Lisäksi olennaiset termit ja teoriat määritellään.

2.1 Luotto ja luottoriski yrityksissä

Luotto viittaa sekä rahan että hyödykkeiden luotolla ostamiseen. Elämme luottoyhteiskunnassa, jossa luotonanto ja luotonotto mahdollistavat allokoinnin kautta useimmat hankinnat ja investoinnit (Niemi 2014, 1). Koulafetis (2017, 1) toteaa, että luotto voi harkitusti käsitellynä rakentaa taloutta, allokoida pääomaa tehokkaasti ja synnyttää varallisuutta. Luotolla myynti on yleistynyt vuoden 2008 finanssikriisin jälkeen yritysten kamppaillessa niukkojen kassavarojen kanssa, jolloin useat pk-yritykset turvautuivat luottomaksuihin selviytyäkseen (Wang, Lan, Zhan & Xing 2018, 5843). Samalla huomattiin luottomyyntiin myös tukevan yrityssuhteita asiakkaiden kanssa (Siekelova, Kollar & Weissowa 2015, 325).

Luotot ovat suurin luottoriskin lähde, mutta luottoriskiä voi syntyä myös muun muassa joukkovelkakirjalainoista ja johdannaisopimuksista (Finanssivalvonta 2018, 5). Luottoriski (credit risk) on maksuhäiriön todennäköisyyden yläkäsite ja tarkoittaa yleisimmin riskiä vastapuolen maksukyvyttömyydestä. Población García (2017, 201) määrittelee luottoriskin todennäköisyytenä tappiosta, joka johtuu kumman tahansa osapuolen epäonnistuesssa osittain tai täysin luottosopimuksen ehtojen noudattamisessa. Ehtojen epäonnistuminen voi liittyä maksun määrään tai määräaikaan. Koulafetis (2017, 2) määrittää luottoriskin hyvin samankaltaisesti riskinä taloudellisesta tappiosta, joka johtuu luotonantajan tai sen vastapuolen kykenemättömyydestä noudattaen maksuehtoja. Doumpos, Lemonakis, Niklis ja Zopounidis (2019, 1) määrittelevät luottoriskin todennäköisyytenä siitä, että veloitettu osapuoli ei noudata velkasitoumustaan ehtojen mukaisesti. Jälkimmäinen luottoriskin määritelmä ottaa siis huomioon vain velkavastuullisen osapuolen riskin. Tämän tutkimuksen kannalta käytämme kyseistä määritelmää, sillä tässä työssä käsitellään vain velkavastuullisen riskiä. Finanssivalvonnan ohjeissa (2018, 5) ilmenee, että luottoriski voidaan mieltää myös vastapuoliriskiksi, vaikka vastapuoliriski liitetään tyypillisemmin keskusteluun rahoitusinstrumenteista. Voidaan todeta, että luottoriskin määritelmien välillä on eroja

kirjallisuudessa, mutta perusidea on selkeä – luottoriski kuvaa todennäköisyyttä siitä, että luottokaupan maksuehtoja ei noudateta.

Luottoriskien hallinnan tavoitteena on löytää hyväksyttävä riskitaso luottokaupan tuottoihin nähden (Siekelova et al. 2015, 327). Luottoriskien hallinta on väistämättä laajentunut kasvavan luotonoton myötä (Doumpos et al. 2019, 2-3). Luottoriskien hallinnan analyyttiset työkalut kiinnostavat luonnollisesti rahoituslaitoksia eniten, lainojen sisältäessä pitkäaikaisia riskejä. Luottoriskin arviointi on elintärkeää niiden toiminnan kannalta, koska lainanantopäätös pohjautuu arvioon takaisinmaksukyvyistä. Pankkiasiakkaiden luottoriskiä onkin tutkittu kirjallisuudessa suhteellisen paljon sekä kuluttaja- että yritysasiakkaissa. (Población García 2017, 13). Tässä työssä keskitytään kuitenkin vain yritysten luottoriskiä.

Población García (2017, 206) sekä Doumpos et al. (2019, 7) toteavat, että luottoriskien hallinta jakautuu kolmeen tarkasteltavaan luokkaan:

- 1) maksuhäiriön todennäköisyys (default probability)
- 2) luottotappion määrä prosentteissa (loss given default)
- 3) altistumisriski (exposure at default) eli luoton määrä silloin, kun maksuhäiriö tapahtuu

Tämän tutkimuksen aiheajaukseen nojaten seuraavaksi paneudutaan maksuhäiriön todennäköisyyteen, joka on myös kyseisistä luokista käytetyin.

2.2 Maksuhäiriö ja maksuhäiriön todennäköisyys

Maksuhäiriö on dikotominen ilmiö tarkoittaen sitä, että se joko tapahtuu tai ei. Maksuhäiriön todennäköisyys voidaan määrittellä maksun epäonnistumisen todennäköisyytenä tarkasteluvälin aikana. Tarkasteluvälin pidentyessä maksuhäiriöriski kasvaa. Tyypillisin tarkasteluväli on yksi vuosi, joka on linjassa taloudellisen raportoinnin kanssa. (Doumpos et al. 2019, 7). Yrityksen maksukyvyttömyys on vahvasti liitännäinen maksuhäiriöihin, ja Koulafetis (2017, 2) määrittelee maksukyvyttömyyden (insolvency) tilanteena, jossa veloitettun osapuolen velat ylittävät varat ja täten se ei kykene maksamaan velkojaan.

Euroopassa 25% yrityskonkursseista johtuu asiakkaiden maksuhäiriöistä (Gustin, 2014). Monilla tuotannollisilla aloilla, joilla maksu suoritetaan tilauksen yhteydessä maksuhäiriöriski ei ole oleellisin toimintariski, mutta maksukäytännön ulottuessa esimerkiksi 90 päivän

päähän maksuhäiriöriski kasvaa merkittävästi (Población García 2017, 34). Saatavien ajoituksella on siis tärkeä rooli maksuliikenteen sujuvuuden suhteen. Todistettu korrelaatio yrityksen maksukyvyttömyyden ja sen taseen saatavien määrän välillä korostaa maksuhäiriörisin hallinnan tärkeyttä (Siekelova et al. 2015, 327).

Yrityksen maksuhäiriön tarkka ennustaminen on monimutkaista tiedon asymmetrisyyden takia (Población García 2017, 35). Yritys on usein hyvin tietoinen omasta maksukyvystään, kun taas ulkopuolinen arvioija voi käyttää vain saatavilla olevaa tietoa, joka ei välttämättä ole aina ajan tasalla. Maksuhäiriön ennustamisen estimointitapoja on kuitenkin monia, ja ne voidaan jakaa karkeasti kolmeen luokkaan. Estimointitapoja ovat arvioiva lähestymistapa (judgmental approach), empiiriset mallit (empirical models) sekä rahoitusmallit (financial models). Näiden erot ilmenevät käytetyissä tiedonlähteissä.

Arvioiva lähestymistapa käyttää laadullisia ja määrällisiä lähteitä, jonka arvioiva asiantuntija yhdistää arvioksi. Kansainvälisesti tunnetuimmat luottoluokituslaitokset ovat Standard and Poor's (S&P), Moody's sekä Fitch, suuruusjärjestyksessä. Luottoluokituksessa yritykset asetetaan maksukykyyn liitännäisten ominaisuuksiensa perusteella eri luokkiin, luottokelpoisimmasta AAA-luokasta riskisimpään D-luokkaan. Näiden raja-arvojen välisiä luokkia on eri määrä mallista riippuen. Arvioiva lähestymistapa tyypillisesti vaatii kuitenkin enemmän resursseja. Empiiriset mallit perustuvat yrityksen historialliseen dataan, jonka kautta analysoidaan tarkasteltavan osapuolen maksukyvyttömyyttä. Tässä menetelmässä kootaan olennaiset maksuhäiriörisin tekijät ja niitä käytetään muuttujina empiirisessä mallissa. Muun muassa tilastolliset mallit ja koneoppiminen ovat laajalti käytettyjä empiirisiä malleja. Rahoitusmallit perustuvat teoriaan ja käyttävät yrityksen rahoitusinstrumentteja arvion lähteenä. Rahoitusinstrumentit heijastavat yrityksen suoriutumiskykyä ja maksuhäiriörisiä esimerkiksi osakehintadatan avulla. (Población García 2017, 207-211; Doumpos et al. 2019, 10-14.) Tässä työssä keskitytään maksuhäiriön todennäköisyyteen empiiristen mallien kautta, sillä logistinen regressio kuuluu tilastollisena mallina tähän luokkaan.

2.2.1 Maksuhäiriön ennustamisen työkaluja

Historian ensimmäinen ja myös tunnetuin maksukyvyttömyyden ennustamisen malli on Altmanin Z-malli, joka kehitettiin vuonna 1968 ennustamaan yrityskonkurssseja. Mallia on sovellettu paljon maksuhäiriöiden ennustamiseen. Z-malli pohjautuu taloudelliseen tietoon, mikä on tyypillistä maksuhäiriön ennustamisessa. (Koulafetis 2017, 28-30.)

Z-mallin kehityksen myötä luottoluokitus- ja luottopisteytysmallit syntyivät. Luottoluokitus (credit rating) on helposti sekoitettavissa luottopisteytykseen (credit scoring), joka tarkoittaa yleisemmin luottokelpoisuutta tietyin kriteerein numeerisesti pisteytettynä. Luottopisteytystä käytetään tyypillisemmin kuluttajamarkkinoilla, kun taas luottoluokitusta yritysmarkkinoilla (Koulafetis 2017, 5). Kuitenkin, molemmat ennustavat luottokelpoisuutta maksuhäiriön todennäköisyyden kautta. Logistinen regressio on ollut vuonna 2009 ammatinharjoittajien kesken yleisimmin käytetty menetelmä maksukyvyttömyyden ennustamiseen (Premachandra, Gurmeet & Toshiyuki 2009, 419).

Muita maksuhäiriön ennustamisen työkaluja ovat muun muassa diskriminanttianalyysi, joka on tilastollinen malli sekä päätöspuut ja neuroverkot, jotka ovat koneoppimisen malleja. Diskriminanttianalyysi on ollut aiemmin hyvin suosittu luottopisteytyksessä (Bolton 2009, 19). Koneoppimisen mallit ovat datalähtöisiä malleja, jotka eivät tee tilastollisia oletuksia. Esimerkiksi neuroverkot kasvattavat liiketoiminnassa suosiotaan vahvan ennustavuutensa takia. (Doumpos et al. 2019, 46-55). Bolton (2009, 27) kommentoi, että neuroverkkoja käytetään tosin yleisemmin tilanteissa, joissa datan rakenne on heikosti ymmärretty. Bolton toteaa, että kun data ja tutkittava ongelma ymmärretään hyvin, voidaan tätä ymmärrystä hyödyntävien testimenetelmien odottaa suoriutuvan paremmin. Tämä selittää esimerkiksi logistisen regression pitkähistoriaista suosiota maksuhäiriön ennustamisessa, sillä yritysten maksuhäiriötekijät tunnetaan melko hyvin. Merkittävät parannukset voivat Boltonin mukaan ilmetä uusia indikaattoreita lisäämällä, mikä tukee tämän tutkimuksen tekoa.

3. LOGISTINEN REGRESSIO JA SEN PÄIVITTÄMINEN

Seuraavaksi tutustumme logistiseen regressioon tilastollisena menetelmänä. Tässä osiossa tutkitaan myös logistisen regressiomallin päivittämisen vaihtoehtoja.

3.1 Logistinen regressio

Logistinen regressio on tilastollinen menetelmä, jossa testattava ilmiö voi saada vain kaksi arvoa. Logistisessa regressiossa tutkitaan selittävien muuttujien vaikutusta selitettävään ilmiöön ja luodulla mallilla voidaan ennustaa kyseisen ilmiön todennäköisyyttä. Logistinen regressio ei ennusta tiettyjä tuloksia tai määriä kuten lineaarinen regressio, vaan todennäköisyyttä ennustevalilla 0-1. Verratessa lineaariseen regressioon, logistinen regressio eroaa myös olennaisesti selitettävän muuttujan suhteen. Logistisen regression selitettävä muuttuja on dikotominen eli kaksiarvoinen muuttuja, jossa vaihtoehtoisina tuloksina ovat 0 ja 1, toisin sanoen jokin ilmiö joko tapahtuu tai ei. Linearisessa regressiossa selitettävän muuttujan ominaisuuksiin kuuluu taas laaja muuttujaväli, esimerkiksi suhdelukuasteikollinen ansiotulojen määrä. (Stoltzfus 2011, 1099.) Logistinen regressio on tilastollinen menetelmä, joka kuuluu myös koneoppimiseen. Koneoppiminen jakautuu ohjattuun ja ohjaamattomaan oppimiseen ja logistinen regressio kuuluu ohjattuun luokkaan, jossa mallin data esitetään luokiteltuna ja nimettynä (Berry Azlinah Bee 2020, 4). Esimerkiksi yritys X:n riskimallin data on luokiteltu ja nimetty maksuhäiriötekijöinä.

Olennainen termi, vetosuhde (odds ratio) viittaa regressiossa kahden vaihtoehtoisen tapahtuman "vedon" väliseen suhteeseen (Kaakinen & Ellonen 2021). Regressiolaskenta tapahtuu logit- tai probit-mallin avulla. Mallien erot johtuvat datan tilastollisten jakaumien eroista. Probit-mallissa muuttujadatan jakauma oletetaan normaalijakautuneiksi ja logit-mallissa logistisiksi. Logit-mallia käytetään tilanteessa, jossa muuttuja ei ole normaalijakautunut. Tällöin muuttujan jakauma muutetaan logistiseksi logaritmin avulla, ja näin voidaan ohittaa oletus datan normaalijakaumasta. (Población García 2017, 209.) Selittävien muuttujien on myös oltava itsenäisiä, tarkoittaen riippumattomuutta muista muuttujista (Kaakinen & Ellonen 2021). Tästä syystä usein yksinkertainen malli on paras. Logistinen regressio rakennetaan otannan avulla. Otannan selitettävän muuttujan on sisällettävä sekä 0 että 1 arvoja, jotta ilmiötä ennustava malli voidaan rakentaa huomioiden pohjalta.

$$Y = \alpha + \beta_i * x_i \quad (1)$$

$$P(Y) = \frac{e^{\alpha + \beta_i * x_i}}{1 + e^{\alpha + \beta_i * x_i}} \quad (2)$$

Logistisen regression tehtävänä on etsiä parhaat estimaatit vakiotermeille (alfa α) ja regressiokerroille (beta β). Regressiokerroin kuvaa selittävän ja selitettävän muuttujan yhteyttä. Termi i kuvastaa muuttujien määrää. Logistinen regressio etsii toistuvasti vahvinta muuttujien lineaarikombinaatiota eli termijoukkoa, jolla saavutettaisiin korkein todennäköisyys ennustaa oikea lopputulos. (Stoltzfus 2011, 1099). Lineaarinen malli Y sisällytetään logistisen regressiomallin kaavaan, jolla saadaan tuloksen todennäköisyys $P(Y)$. Jos tuloksen todennäköisyys ennustaisi esimerkiksi yrityksen maksuhäiriötä, ennusteväli olisi 1 = Maksuhäiriö ja 0 = Ei maksuhäiriötä. Tässä tapauksessa pienempi tulos kuvastaa siis yrityksen parempaa luottokelpoisuutta.

Logistisen regression suorittamisen jälkeen tutkitaan muuttujien suhteiden merkitsevyyttä mallissa. Tarkastellaan siis selittävien muuttujien suhdetta selitettävään muuttujaan. Regressiokerroimen suuruus indikoi muuttujan selityskykyä tutkittavassa ilmiössä. Muuttujan parvosta nähdään selittävien muuttujien tilastollinen merkitsevyys halutulla riskitasolla. Yleisesti käytetty riskitaso on 5 prosenttia. On olemassa kuitenkin myös muita tapoja arvioida mallia. Devianssi (deviance) tutkii, kuinka selittäviä mallin ennusteet ovat. Uskottavuusosamäärä (likelihood-ratio test) testaa paraneeko mallin selittävyys muuttujia pudottamalla. Se testaa, voisiko regressiokerroimia kohdella nollina, ja tavoitteena on tehdä mallista yksinkertaisempi muuttujia pudottamalla. Wald-testi testaa regressiokerroimien tilastollista merkitsevyyttä. Isojen regressiokerroimien kohdalla on kuitenkin havaittu väärää negatiivisia eli tyyppin 2 virheitä. Score test testaa muuttujan merkitsevyyttä jakaumien ja derivaattojen avulla. (Bolton, 2009, 44-45.)

3.1.1 Opettamis- ja validaatiodata

Koneoppimismallien, kuten logistisen regression, kehittämiseen liittyy myös olennaisesti datan jakaminen eri joukkoihin mallin kehittämistä, validoimista ja testaamista varten. Datan jakaminen eri joukkoihin perustuu yksinkertaisesti siihen, että jos kehitettyä mallia halutaan validoida tai testata, emme voi käyttää samaa dataa, jota käytettiin mallin kehittämiseen.

Tämä johtaisi mallin yliarviointiin. Yksinkertaisella jaolla otanta voidaan jakaa opettamisdataan ja validaatiodataan. Opettamisdatalla kehitetään malli, ja mallin parametrit sovitetaan kyseiseen dataan. Validaatiodatalla arvioidaan mallin suoriutumiskykyä, ja validaatiodata voidaan jakaa sisäiseen ja ulkoiseen dataan. Sisäinen validaatiodata arvioi muun muassa mallin luokittelun tarkkuutta. Ulkoinen validaatiodata eli otanta uudesta, mutta samankaltaisesta ympäristöstä arvioi mallin ulkoista suoriutumista, mikä on tärkeää mallin sovellettavuuden arvioinnin kannalta. Testausdata, kolmantena luokkana, on tuttu koneoppimisen malleissa, ja sillä arvioidaan myös mallin suoriutumiskykyä. Testausdataan viitataan usein holdout- tai ulkoisen validaatiodatan termeillä, ja kirjallisuudessa on paljon termien sekalaista käyttöä. Datan jakaminen riippuu myös otannan koosta. Tässä työssä käsitellään otantaa opettamisdatan ja validaatiodatan kautta, sillä logistisessa regressiossa tämä jako on tyypillisin. (Doumpos et al. 2019, 34; Stoltzfus 2011, 1102.)

3.2 Aiempi tutkimus logistisen regression päivittämisestä

Kun malli luodaan tietystä otannasta, malli heijastaa tällöin kyseistä tapausta. Mikä tahansa tilastollinen malli siis tuottaa päteviä ennusteita alkuperäisen testijoukon pohjalta. Ihanteellisesti malli pystyy soveltamaan ennustuskkyä muihin tapauksiin, ja tähän voidaan vaikuttaa luomalla malli oikeiden vaiheiden kautta. Ennustemallin päivittäminen tulee kuitenkin tarpeeseen, jos mallin data vanhenee, muuttujien ennustamiskyky heikkenee tai uuden datan lisäämisen tarve esiintyy. Tässä työssä motiivi löytyy halusta lisätä uutta dataa yrityksen riskimalliin. Mallin päivittämiseen on useampia vaihtoehtoja. (Steyerberg 2019, 399.)

Logistisen regression luottoriskimallin päivittämisestä on yleisesti ottaen tiedekirjallisuudessa todella suppeasti tietoa. Doumpos et al. (2019, 37) mukaan luottopisteytys- ja luotto-luokitusmallit päivitetään joko kalibroimalla malli tai uudelleen estimoimalla malli. Analyttiset ennustemallit käyttävät tyypillisesti vain kourallista muuttujia varmistaakseen sen, että malleja on helppo ylläpitää ja päivittää. Muuttujien pieni määrä vähentää myös esimerkiksi multikollinearisuuden riskiä. Muuttujamäärän kasvaessa ylisovittamisen (overfitting) riski kasvaa, jolloin mallin suorituskyky heikkenee. Ylisovittaminen perustuu usein päällekkäiseen tietoon mallissa, ja tämä on yleistä, kun muuttujia on liian monta. (Doumpos et al. 2019, 37.) Doumpos et al. eivät mainitse tarkemmin mallin laajentamisesta uusien muuttujien avulla. He puhuvat mieluummin pienen muuttujamäärän puolesta mallissa. Logistisen regressiomallin päivittämistä on kuitenkin tutkittu lääketieteessä. Lääketieteessä käytetään paljon ennustavia malleja, jotka hyödyntävät eri indikaattoreita sairauksien ennustamiseen.

Logistinen regressio toimii hyvin sairauksien ennustamisessa, joissa haetaan dikotomista tulosta eli kyllä- tai ei-ennustetta. Lääketieteessä logistista regressiota päivitetään tyypillisemmin potilasotantaa laajentamalla kuin yksittäisillä muuttujalisäyksillä, mutta molemmista päivittämismenetelmistä löytyy kirjallisuutta. Seuraavaksi paneudumme logistisen regressiomallin päivittämismenetelmiin.

3.2.1 Mallin kalibrointi, korjaaminen ja uudelleen estimointi

Mallin päivittämisen syy voi olla vaikkapa uusi otanta tai uuden muuttujan lisääminen malliin. Steyerberg (2019, 400-402) toteaa lääketieteen tutkimuksessa, että kolme tapaa päivittää ennustemalli on

- 1) *mallin kalibrointi* (recalibration)
- 2) *mallin korjaaminen* (model revision)
- 3) *mallin laajentaminen* (model extension)

Mallin kalibroinnissa muutetaan regressiomallin vakiota ja mahdollisesti regressiokertoimia. *Mallin korjaamisessa* estimoidaan uudelleen joitain tai kaikkia regressiokertoimia ja *mallin laajentamisessa* laajennetaan mallia selittäville muuttujille ja estimoidaan koko malli uudelleen. Yleinen syy mallin huonolle suoriutumiseen uudessa datajoukossa on ero perusriskissä. Tällaiseen tilanteeseen soveltuu *mallin kalibrointi*, joka jakautuu joko vakiotermin tai vakiotermin ja regressiokertoimen kalibrointiin. Mallin vakiotermi ohjaa pitkälti keskiarvon tarkkuutta, joten käytännössä pelkällä vakiotermin kalibroinnilla voidaan vaikuttaa mallin ennustamiskykyyn paljon (Shipe, Deppen, Farjah & Grogan 2019, 582). Steyerberg (2019, 399) mukaan ”calibration-in-the-large” perustuu siihen, että mallin ennustaman keskiarvon tulisi olla sama kuin uuden datan keskiarvon. Vakiotermi voidaan kalibroida opettamisdatan ja uuden datan keskiarvojen erolla tai suorittamalla logistinen regressio uuteen dataan, jossa vain vakiotermi kalibroidaan ja loppuosa mallista pysyy kiinteänä. Vakiotermin kalibrointi perustuu oletukseen siitä, että selittävien tekijöiden suhteellinen ennustamisvahvuus on lähestulkoon samanlainen uudessa datajoukossa. Toimiva vakiotermin kalibrointi viittaa siihen, että datajoukkojen keskiarvojen erot johtuivat pitkälti mallin ulkopuolella olevista tekijöistä. Logistinen kalibrointi kalibroi vakiotermin lisäksi selittäville tekijöille yleisen regressiokertoimen b . Tämä tapahtuu suorittamalla logistinen regressio uuteen dataan, jossa alkuperäisen mallin yhtälö toimii kovariaattina. Kovariaatti on apumuuttuja, joka vaihtelee jonkin toisen muuttujan kanssa (Sarna, 2012). Kalibrointimenetelmissä ongelmana voi olla

yksittäisten regressiokertoimien vääristymät (Steyerberg, Harrell, Borsboom, Eijkemans, Vergouwe & Habbema 2001, 780). *Mallin korjaamisessa* estimoidaan useampi parametri. Estimoinnin kohteena voi olla vakiotermin lisäksi joko osa tai kaikki regressiokertoimista. Jos halutaan estimoida osa regressiokertoimista, testataan niiden merkitsevyyttä uudessa datassa. Estimoidamalla kaikki parametrit luodaan uusi malli uuden datan pohjalta. *Mallin laajentamisessa* malliin sisällytetään uutta dataa muuttujien muodossa. Malli voidaan päivittää muun muassa kalibroimalla vakio-termi ja yleinen regressiokerroin sekä testaamalla sekä alkuperäisten että uusien muuttujien merkitsevyyttä. Vaihtoehtoisesti mallin laajentamisessa voidaan estimoida koko malli uudelleen sekä alkuperäisellä että uudella datalla. (Steyerberg 2019, 399-402.)

Päivittämismenetelmän valinta riippuu päivittämisen tarpeesta ja siitä, että lisätäänkö uusia muuttujia. Päivittämismenetelmä riippuu sopivista syistä soveltaa ennustemallia uuteen ympäristöön, ennustemallin muuttujien määrästä ja otannan koosta. Tämän lisäksi malli ei saa olla selkeästi ylisovitettu, ja sen pitäisi olla kohtuullisin keinoin kehitetty otannan kokoon nähden. Esimerkiksi pienestä otannasta kehitetty malli saattaa vaatia suurempia korjauksia mallia päivitettäessä. (Steyerberg, 2019, 413.) Kuvassa 2 esitetään päivittämissvaihtoehdot tilanteessa, jossa 8 muuttujan lääketieteen ennustemalli päivitetään, kun 8 uuden muuttujan data on saatavilla.

Nr	Label	Notation	Predictors considered	Parameters tested	Parameters estimated
No updating					
1	Apply original prediction model	–	8	0	0
Recalibration					
2	Update intercept	α	8	0	1
3	Recalibration of intercept and slope	$\alpha + \text{calibration slope } \beta_{\text{overall}}$	8	0	2
Model revision					
4	Recalibration + selective reestimation	$\alpha + \beta_{\text{overall}} + \gamma_{1..8}$ $p \leq 0.05$	8	8	2–9
5	Reestimation	$\alpha + \beta_{1..8}$	8	0	9
Model extension					
6	Recalibration + selective reestimation + selective extension	$\alpha + \beta_{\text{overall}} + \gamma_{1..8}$ $p \leq 0.05 + \beta_{9..16}$ $p \leq 0.05$	16	16	2–17
7	Reestimation + selective extension	$\alpha + \beta_{1..8} + \beta_{9..16}$ $p \leq 0.05$	16	8	9–17
8	Reestimation + extension	$\alpha + \beta_{1..16}$	16	0	17

Kuva 2. Logistisen regression päivittämissmenetelmät. (Steyerberg 2019, 402)

4. MAKSUHÄIRIÖTEKIJÄT YRITYKSISSÄ

Seuraavaksi paneudutaan aiempiin tutkimuksiin yritysten maksuhäiriöistä ja tutkitaan kirjallisuuden merkittävimpiä maksuhäiriörisiin liitettyjä tekijöitä. Indikaattoreita maksuhäiriöihin ja maksukyvyttömyyteen löytyy laajalti kirjallisuudesta. Tässä osiossa vastataan tutkimuskysymyksiin 1. *”Mitkä tekijät vaikuttavat yrityksissä maksuhäiriöihin kirjallisuuden mukaan?”* ja 2. *”Onko yritys X:n riskimallisissa huomioitu kirjallisuudesta löytyneet maksuhäiriöihin vaikuttavat tekijät?”*.

4.1 Yritysten luokittelu

Lähtökohtana maksukyvykkyyden arvioinnissa on se, että yrityksiä tarkastellessa arvioidaan niiden käytössä olevat resurssit. Tämän lisäksi tulevaisuuden tuloja ja kassavirtaa arvioidaan toiminnan ja investointien kautta. (Koulafetis 2017, 120.) Tyypillisesti maksuhäiriötekijöitä kuvataan suhdelukuina ja tunnuslukuina, ja tilinpäätökset ovat avainasemassa etenkin luottoriskin arvioinnissa (Doumpos et al. 2019, 30). Arvioitaessa suurta yritysjoukkoa on tärkeää kategorisoida yritykset ensin ominaisuuksiensa mukaan. Kirjallisuudessa esiintyneitä maksuhäiriöön liitännäisiä yritysominaisuuksia ovat yrityksen koko, toimiala ja ikä. Ne eivät itsessään selitä maksuhäiriöitä, mutta kyseisten muuttujien skaalassa on eroja maksukyvyyn suhteen.

Toimiala on yhteydessä maksuhäiriörisiin toimialan elinvoimaisuuden sekä vakauden kautta, ja näihin vaikuttavia tekijöitä ovat muun muassa kasvuodotukset, teknologiset muutokset, riskisyys ja kilpailutaso. Suuret erot toimialojen kulurakenteissa luovat eroja yritysten maksuhäiriörisin vertailussa. Esimerkiksi jokin taloudellinen tunnusluku voi olla yliedustettu tietyn toimialan kulurakenteen yrityksellä. Toimialoissa on siis eroja maksukyvyyn suhteen. (Bonfim 2009, 298; Koulafetis 2017, 121; Población García 2017, 35.)

Yrityksen koko on yleisesti ottaen kirjallisuuden mukaan maksukykyä vahvistava tekijä. Tätä mittaavat liikevaihto ja henkilöstömäärä. Pienemmät yritykset ovat suuremman maksuhäiriörisin alaisia (Bonfim 2009, 285-286). Aiheeseen lisäten, Dwyer, Kocagil & Stein (2004, 13) mukaan suurempien yritysten maksuhäiriöt ovat harvinaisempia. Kuitenkin, isossa skaalassa niiden suuremmat luottotappiot hävittävät yrityskoon positiivista vaikutusta

maksukykyyn. Yrityksen koon lisäksi konserniin kuuluvat yritykset voivat olla maksukyvykkyydeltään vahvempia, sillä konsernin skaalaetujen lisäksi ne tyypillisesti saavat tarvittaessa konsernilta taloudellista tukea esimerkiksi konserniavustusten muodossa. Konserniavustuksella voidaan kasvattaa saavan osapuolen yhteisön tuloa tai pienentää tappiota (Vero, 2021). Yrityksen ikä on myös maksukyvykkyyteen positiivisesti vaikuttava tekijä. Uusilla yrityksillä on enemmän vaikeuksia saada rahoitusta, sillä niihin liittyy enemmän epävarmuutta. Nuorten yritysten vertailuvuosien datan puuttuessa luotettavuutta on vaikea arvioida. Tämän lisäksi ne ovat haavoittuvaisempia epäsuotuisille markkinaolosuhteille. Pidempiin markkinoilla toimineet yritykset vastaanottavat rahoitusta usein paremmilla ehdoilla ja alhaisemmilla koroilla. (Doumpos et al. 2019, 31.)

4.2 Taloudelliset maksuhäiriötekijät

Edellä mainittujen tekijöiden lisäksi taloudelliset tunnusluvut ovat suuressa roolissa maksuhäiriöiden ennustamisessa. Esimerkiksi Altmanin Z-malli käyttää suhdelukuja, jotka arvioivat yrityksen likviditeettiä, kannattavuutta, pääomaa ja suorituskykyä. Myös Koulafetis (2017, 120) toteaa tärkeimpien taloudellisten maksuhäiriötekijöiden liittyvän yrityksen vipuvaikutukseen, likviditeettiin, kannattavuuteen, kassavirran vakauteen ja pääomarakenteseen. Kyseiset tunnusluvut ovat tyypillisesti laskettu 12kk aikaväliä käyttäen.

Kannattavuus on hyvin tärkeä taloudellinen maksukyvykkyyden indikaattori. Kannattavuus kuvaa yrityksen suorituskykyä ja toimintaedellytyksiä, ja se mainitaankin kirjallisuudessa usein ensimmäisenä taloudellisena tekijänä maksuhäiriöitä tutkittaessa. Korkea kannattavuus vähentää maksuhäiriöriskiä. Kannattavuuden tunnuslukuja ovat muun muassa liikevoittoprosentti, käyttökateprosentti, oman pääoman tuottoprosentti (ROE) ja kokonaispääoman tuottoprosentti (ROA). (Koulafetis 2017, 124; Doumpos 2019, 81.)

Likviditeetti on olennainen maksuhäiriötekijä, sillä se kuvaa nimenomaan maksuvalmiutta. Se indikoi kykyä muuttaa omaisuuserät rahaksi sekä kykyä selvitä lyhytaikaisista veloista. Likviditeetin tunnuslukuja ovat tyypillisesti quick ratio ja current ratio. Current ratio vertaa likvidien erien suhdetta lyhytaikaisiin velkoihin. Quick ratio kuvaa myös samaa suhdetta, mutta jättää vaihto-omaisuuden eli varastot ja ennakkomaksut huomiotta, jolloin se kuvaa vielä ”nopeampaa” likviditeettiä. (Averkamp 2021; Doumpos et al. 2019, 30.)

Pääomarakenteen vaikutuksesta yrityksen suorituskykyyn on ristiriitaista tutkimusta. Osa tutkimuksista väittää velkavivun parantavan suorituskykyä ja jotkut taas painottavat negatiivista yhteyttä nousevien korkojen takia. Myös toimiala vaikuttaa osaltaan pääomarakenteen luonteeseen. Li, Niskanen & Niskanen (2019, 594) tutkimuksen mukaan matalan luottoriskin pk-yritysten yhteys velkavipuun on suorituskykyä heikentävä. Tätä selittää se, jos yrityksellä on riittävää tulo-rahoitusta, suosii se tätä vieraan pääoman yli. Kyseistä teesiä vahvistaen, Doumpos et al. (2019, 81) mukaan velan määrällä on positiivinen yhteys maksuhäiriöihin. Pääomarakenne ja luottoriski vaikuttavat myös toisiinsa vähintäänkin niin, että korkea velan määrä nostaa yrityksen luottoriskiä ja korkea luottoriski nostaa velan korkokustannuksia. Maksuhäiriön ennustamista varten pääomarakenteen ja vakavaraisuuden kuvaavia tunnuslukuja ovat omavaraisuusaste, velkaantumisaste ja velkasuhde.

4.3 Muut maksuhäiriötekijät

Taloudellisen tiedon hyödyntäminen on käytännöllistä ja tehokasta maksuhäiriön ennustamisessa, sillä tilinpäätöstiedot ovat suhteellisen helposti saatavilla. Kuitenkin, on olemassa maksuhäiriötä selittäviä tekijöitä, jotka eivät perustu tilinpäätöstietoihin. Muun muassa kvalitatiivinen eli laadullinen tieto on tutkimusten mukaan arvokasta yrityksen maksukykyä arvioitaessa.

Esimerkiksi hallinnointikoodi (corporate governance) ja johtamisen tehokkuus ovat yhteydessä yrityksen taloudelliseen vakauteen. Yrityksen etiikka, hallintorakenne, päätöksenteon läpinäkyvyys ja sisäinen auditointi ovat hallinnointikoodin tietoa, jota ei ole täysin saatavilla tilinpäätöksistä. Johtamisen tehokkuutta voi mitata muun muassa varastojen, saatavien ja ostovelkojen erinäisillä suhdeluvuilla, näistä esimerkkinä vaihto-omaisuuden kiertoaika. (Doumpos et al. 2019, 30-32; Koulafetis 2017, 125.)

Yrityksen maksuhäiriöhistoria on vahva maksuhäiriöindikaattori. Yritykset, jotka ovat saaneet maksuhäiriömerkinnän aikaisemmin, ovat suuremman riskin alaisia uuden maksuhäiriön suhteen. (Bonfim 2009, 298.) Virallisia maksuhäiriömerkintöjä voi olla haastavaa seurata reaaliaikaisesti ja pitkällä aikavälillä, sillä maksuhäiriömerkinnät ilmestyvät rekisteriin viiveellä maksuhäiriön ajankohdasta ja yrityksen yksittäinen maksuhäiriö pysyy rekisterissä korkeintaan neljä vuotta (Dun & Bradstreet, 2021). Myös maksukäyttäytyminen on hyödyllinen tieto maksukyvyyn arvioinnissa (Siekelova et al. 2015, 327). Intrum (2018) tutkii Euroopan yritysten maksutapaa vuosittain. Intrumin vuoden 2018 European Payment Report

analysoi 9 607 eurooppalaisen yrityksen maksukäyttäytymisen piirteitä. Pohjoismaissa noudatetaan lyhyempiä maksuehtoja, ja luottotappion määrä on pieni verrattuna esimerkiksi Etelä-Eurooppaan. Suomen keskimääräinen luottokaupan maksuajankohta yritysotannassa oli 25 päivää vuonna 2018, Euroopan keskiarvon ollessa 34. Suomessa yritysten maksukäyttäytymiseen vaikuttaa lainsäädännön lisäksi kunnioitus kirjoittamattomia sääntöjä kohtaan. (Intrum, 2018.) Päähuomiona on kuitenkin se, että maksuviive voi toimia tekijänä yrityksen maksuhäiriöennustamisessa, sillä yrityksen jatkuva viivästyminen maksuissa voi indikoida taloudellisia vaikeuksia.

Listattujen yritysten kohdalla pörssidata voi olla arvokasta yrityksen nykyhetken ja tulevaisuudennäkymien tarkastelussa. Tämä kuuluu aiemmin mainittuun maksuhäiriöriskin rahoitusmalli -estimointitapaan, jossa suoritetaan rahoitusinstrumenttien tarkastelua. Muun muassa osakekehityksen lisäksi arvioidaan osakkeiden volatilitteettia luottoriskin kartoittamiseksi. (Doumpos et al. 2019, 31-32.)

Yrityksen markkina- ja kilpailutilanne ovat myös yhteydessä yrityksen suoriutumiskykyyn ja täten maksukykyyn. Kilpailutaso, markkinoiden avoimuus ja markkinoiden ennuste toimivat arvioitavina tekijöinä luottoriskin hallinnassa. (Doumpos et al. 2019, 31.) Boss, Fenz, Pann, Pühr, Schneider & Ubl (2009, 85) toteavat makroekonomisten tekijöiden ja luottoriskin olevan viimeaikaisen tutkimuksen kohteena. Kiinnostus aiheeseen on kasvanut etenkin finanssikriisin jälkeen. Bonfim (2009, 298) tutki makrotaloudellisten tekijöiden vaikutusta yritysten maksukyvykkyyteen, ja tutkimuksesta paljastui niiden selittävä yhteys maksuhäiriöriskiin. Makrotalouden kehityksen yhteyden lisäksi tutkimuksessa ilmeni uutena tietona, että vahvan talouskasvun aikana yrityksissä ilmeni ajoittain liiallista riskinottoa. Liiallinen riskinotto ilmeni myöhemmin taipumuksena maksukyvyttömyyteen. Makrotaloudelliset tekijät, esimerkkeinä BKT ja inflaatio, tutkimuksen mukaan selittävät maksuhäiriöriskiä.

Laajemmassa analyysissä on otettava huomioon eri laskentatoimen käytännöt, etenkin vertailtaessa kansainvälisiä yrityksiä. Kansainvälisessä vertailussa Koulafetisin (2017, 121) mukaan yrityksen yleiseen riskiasemaan vaikuttaa valtiokohtainen riski. Kyseinen riski sisältää esimerkiksi taloudellisia, valuuttakurssiin liittyviä ja poliittisia аспекteja. (Población García 2017, 35). Tässä työssä kansainväliseen vertailuun ei kuitenkaan paneuduta tämän enempää, tutkimuksen rajoituksessa maantieteellisesti Suomeen.

Taulukko 1. Yrityksen maksuhäiriön todennäköisyyteen vaikuttavat tekijät. +/- -sarake kuvastaa maksuhäiriötekijän suhdetta maksuhäiriöön

Default probability factor		Maksuhäiriötekijä	Esimerkkejä	+/-	Lähde
1	Industry	Toimiala		+/-	Bonfim (2009); Koulafetis (2017); Población García (2017)
2	Size	Koko	Liikevaihto	-	Bonfim (2009); Doumpos et al. (2019)
3	Age	Ikä		-	Doumpos et al. (2019)
4	Profitability	Kannattavuus	Liikevoitto%, ROE, ROA	-	Bonfim (2009); Doumpos et al. (2019); Koulafetis (2017)
5	Liquidity	Likviditeetti	Quick ratio, Current ratio	-	Bonfim (2009); Doumpos et al. (2019)
6	Capital Structure	Pääomarakenne	Omavaraisuusaste	-	Bonfim (2009); Doumpos et al. (2019); Li et al. (2019)
7	Management efficiency	Johtamisen tehokkuus	Vaihto-omaisuuden kiertoaika	+	Doumpos et al. (2019)
8	Corporate Governance	Hallinnointikoodi	Yrityksen hallintorakenne	+/-	Doumpos et al. (2019); Koulafetis (2017)
9	Default history	Maksuhäiriöhistoria	Maksuhäiriömerkintä	+	Bonfim (2009)
10	Payment behavior	Maksukäyttäytyminen	Maksuviive	+	Doumpos et al. (2019); Siekkelova et al. (2015)
11	Market conditions	Markkinatilanne	Kilpailutaso	+	Doumpos et al. (2019)
12	Stock market data	Pörssidata	Osakekehitys	-	Doumpos et al. (2019); Población García (2017)
13	Macroeconomic factors	Makrotaloudelliset tekijät	BKT, inflaatio	+/-	Bonfim (2009); Doumpos et al. (2019)

Taulukko 1 vastaa tutkimuskysymykseen 1. *”Mitkä tekijät vaikuttavat yrityksissä maksuhäiriöihin kirjallisuuden mukaan?”*. Taulukko sisältää lähteineen kirjallisuudesta löytyneet maksuhäiriötekijät. Osa tekijöistä vahvistaa ja osa lieventää maksuhäiriön todennäköisyyttä. Kaiken kaikkiaan, yritysten maksuhäiriöihin vaikuttaa kirjallisuuden mukaan määrällisiä ja laadullisia tekijöitä. Maksuhäiriötekijät ovat kuitenkin pääosin taloudellisia ja määrällisiä tekijöitä. Voidaan myös todeta, että tilinpäätöstieto on arvokasta maksuhäiriöriskin arvioinnissa, sillä tilinpäätökset sisältävät paljon taloudellista ja määrällistä tietoa.

4.4 Yritys X:n riskimallin maksuhäiriötekijät

Seuraavaksi vertaillaan kirjallisuuden maksuhäiriötekijöitä yritys X:n riskimallin tekijöihin. Vastataan siis samalla tutkimuskysymykseen 2. *”Onko yritys X:n riskimallissa huomioitu kirjallisuudesta löytyneet maksuhäiriöihin vaikuttavat tekijät?”*. Yritys X:n riskimallissa on maksuhäiriötekijöitä yhteensä yli 20, ja ne kattavat useimmat edellä mainituista maksuhäiriötekijöiden kategorioista. Toimiala, liikevaihto ja ikä ovat huomioitu riskimallissa, ja niiden arvot on jaoteltu eri luokkiin. Kannattavuus, likviditeetti ja pääomarakenne ovat huomioitu muun muassa nettotulosprosentin, sijoitetun pääoman tuottoprosentin, current ration, oma-varaisuusasteen ja vieraan pääoman takaisinmaksuajan tunnuslukujen kautta.

Markkinatilannetta kuvaa osittain toimiala, sillä kilpailutaso on toimialakohtainen. Makrotaloudellisista tekijöistä inflaation voidaan nähdä kytkeytyvän yritys X:n riskimallin liikevaihdon kasvuprosenttiin, mutta yhteys on oletettavasti pieni. Hallinnointikoodin, markkinatilanteen ja makrotaloudellisten tekijöiden vaikutuksen voidaan sanoa olevan vähän huomioitu. Yritys X:n riskimalli ei siis rajoitu tilinpäätöstietoihin, mikä on hyvä lähtökohta ottaen huomioon, että tyypillisin tapa ylläpitää luottoriskin hallintaa on tilinpäätöstietojen kautta.

Maksuhäiriöhistoriaa, maksukäyttäytymistä sekä pörssitietoja ei ole huomioitu mallissa. Etenkin aiemman maksuhäiriömerkinnän vaikutus maksuhäiriöriskiin oli kirjallisuuden mukaan suuri, joten sen lisäämistä riskimalliin voi harkita. Myös yrityksen maksukäyttäytyminen voisi olla hyvä Yritys X:n riskimalliin lisättävä muuttuja. Sen vaikutus maksuhäiriöriskiin on kirjallisuudessa todettu useassa tutkimuksessa. Pörssitiedon saanti rajoittuu listattuihin yrityksiin, ja osittaista tietoa voi olla haastava lisätä valmiiksi laajasti vertailtavaan riskimalliin.

5. YRITYS X:N RISKIMALLIN PÄIVITTÄMINEN

Logistisen regressiomallin päivittämismenetelmät käytiin läpi aiemmin, ja tässä osiossa käydään läpi vaihtoehtoiset menetelmät lisätä dataa yritys X:n riskimalliin. Samalla vastataan tutkimuskysymykseen 3. ”Miten on mahdollista lisätä uutta dataa toimivaan logistiseen regressiomalliin?”. Kysymystä tutkitaan teoreettisella tasolla.

5.1 Riskinmallinnuksen räätälöinti ja yritys X

Avataan ensin räätälöidyn riskinmallinnuksen termi ja sen tavoitteet. Yritys X:n riskimalli toimii logistisella regressiolla, joten yritys X toimii esimerkkinä tilanteessa, jossa logistisen regressiomallin päivittämistä voisi toteuttaa. Teoreettinen tutkimus logistisen regression päivittämisen vaihtoehtoista voi selkeyttää mahdollisia vaihtoehtoja asiakaskohtaisen datan käyttöönotossa, mikä voi mahdollistaa räätälöidyn riskinmallinnuksen.



Kuva 3. Riskinmallinnuksen kolmiportaisuus

Selkeytetään aihetta avaamalla vaiheen kolmiportaisuus. Kuvan 3 ensimmäisessä portaassa yritys X tarjoaa riskimallia omille asiakasyrityksille, ja toisessa portaassa nämä asiakasyritykset käyttävät sitä omien asiakasyritysten arvioinnissa. Kolmannen portaassa asiakasyritykset olisivat siis riskiarvioinnin kohteena, ja tässä työssä pohditaan heidän datansa käyttöönottoa riskimallissa. Tavoitteena on räätälöidä yritys X:n asiakasyrityksille riskinmallinnuksen riskimalli, joilla he voivat täten arvioida omia asiakkaitaan.

Riskinmallinnuksen räätälöintiä tutkitaan kirjallisuuden ja aiempien tutkimusten kautta. Tutkiminen tapahtuu teoreettisella tasolla. Myös tutkimuksen terminologia muuttuu hieman, sillä luottoriskin ja maksuhäiriön ennustamisen sijaan puhutaan nyt yritysten riskinmallinnuksesta. Tämä johtuu siitä, että tässä vaiheessa ei tarkasti määritellä mahdollisesti

lisättävän datan luonteen kriteerejä, ja täten mahdollisesti lisättävä data ei rajaudu maksuhäiriötekijöihin. Riskinmallinnus viittaa terminä paremmin aiheeseen.

Yrityksessä X logistisen regression riskimalli luokittelee yritykset 12 kuukauden maksuhäiriötodennäköisyytensä mukaan. Riskimalli muistuttaa luottoluokitusta (credit rating), joka arvioi yrityksen luottokelpoisuutta ennustamalla yrityksen maksuhäiriön todennäköisyyttä. Otanta, jolla yritys X:n logistinen regressio on kehitetty, on todella suuri. Malli on kehitetty otannalla, jonka ennusteita on seurattu ja jälkeenpäin näitä ennusteita on verrattu yritysten tapahtuneisiin maksuhäiriöihin. Kehittämisessä on käytetty myös opettamis- ja validaatiotietoa, joista puhuttiin osiossa 3.1.1. Mallissa on hyödynnetty yritysten tilinpäätöstietoja useammalta tilikaudelta ja näiden lisäksi käytetty muita tietoja. Yritys X:n logistinen regressiomalli on myös ammattilaisten kehittämä, mikä tekee mallista hyvin luotettavan.

5.2 Otannan ja muuttujien vaikutus päivittämismenetelmän valintaan

Ennustemallien, kuten logistisen regressiomallin päivittäminen tapahtuu tyypillisesti joko uuden ympäristön tai uuden muuttujan lisäämisen vuoksi. Tutkimusta uuteen ympäristöön päivittämisestä on olemassa enemmän, ja useimmiten yksinkertaiset menetelmät riittävät mallin päivittämiseen, jos on tarpeeksi todisteita mallin sopivuudesta uuteen ympäristöön. Esimerkkinä, yhdessä lääketieteen tutkimuksessa etsittiin parasta päivittämismenetelmää uuteen ympäristöön ja päädyttiin useasta vaihtoehdosta calibration in the large -menetelmään, jossa vain mallin vakiotermejä muutetaan. (Vergouwe, Nieboer, Oostenbrink, Debray, Murray, Kattan, Koffijberg, Moons, & Steyerberg 2017, 4529). Tässä työssä tutkitaan kuitenkin uuden muuttujan tai muuttujien lisäämistä malliin, ja tällöin on pohdittava tapauskohtaisia tekijöitä. Yhtenä tavoitteena on selvittää, onko riskimallin uudelleenestimointi välttämätön muuttujia lisätessä.

Nieboer, Vergouwe, Ankerst, Roobol-Bouts & Steyerberg (2016, 1) mainitsevat, että uusia muuttujia lisättäessä malliin joudutaan usein käyttämään pientä päivittämisjoukkoa eli uutta dataa. Liian pieni datajoukko kuitenkin altistaa mallin ylisovittamiselle sekä alhaiselle luotettavuudelle ja täten mallin suorituskyky voi heikentyä. Tämä johtuu siitä, että mallin ennusteet ovat pienessä datajoukossa liian optimistisia. Uuden datajoukon koko vaikuttaa siis päivittämismenetelmän valintaan. Pienen datajoukon kohdalla suositellaan tyypillisesti hyllyttävä mallin muokkauksia, kun taas suuremman datajoukon kohdalla laajemmat mallin muokkaukset, kuten koko mallin uudelleenestimointi on saavutettavissa luotettavammin

tuloksin (Nieboer et al. 2016, 7-9). Myös yritys X:n tapauksessa pitää siis ottaa huomioon uuden datan koko ja sen suhde mallin opettamisdataan. Yritys X:n uusi data, jota käsitellään seuraavissa kappaleissa, olisi data yritys X:n asiakkaiden omista asiakasyritysjoukoista eli toisin sanoen data kolmannesta portaasta. Uuteen dataan sisältyisi olennaisesti myös riskimalliin lisättävä muuttuja tai muuttujat, ja tämä tukee räätälöidyn riskimallinnuksen ideaa. Yritys X:n tapauksessa alkuperäisen mallin opettamisdataa suurempi yritysjoukkodata ei olisi todennäköinen, sillä mallin opettamisdata on niin suuri, että ei ole todennäköistä, että se ylittäisi yhden yritysjoukon määrän.

Vaikka lisättävää dataa ole vielä tarkennettu, on sille kuitenkin asetettava jonkinlaiset kriteerit koon ja rakenteen suhteen. Ensinnäkin, logistinen regressio olettaa selittävien muuttujien normaalijakaumaa probit-mallissa. Jos lisättävä muuttuja ei ole normaalijakautunut, käytetään logit-mallia muuttamalla muuttuja logaritmfunktiolla logaritmiseksi. Muuttujan koon suhteen, yksi yleisohje on vähintään 10 näytettä jokaista mallin syötemuuttujaa kohden. Toisaalta, myös 20 näytteen rajasta on ollut puhetta kirjallisuudessa. Yleisesti ottaen alle 100 otannan malli johtaa usein ylisovittamiseen mallissa. Noin 500 otannasta kehitetty malli parantaa ennusteiden tarkkuutta. Nämä ovat kuitenkin vain yleisiä ohjeita ja todellisuudessa mallin suoriutuminen riippuu tapauksesta. (Bujang, Sa'at, Sidik & Joo 2018, 123.)

5.3 Päivittämisvaihtoehdot

Tutkimme seuraavaksi tutkimuskysymystä 3 yritys X:n riskimallin kautta, sillä se toimii esimerkkinä logistisesta regressiosta. Aiemmin läpikäytyjä logistisen regression päivittämismenetelmiä käydään läpi yritys X:n tapauksessa, arvioimalla niiden soveltuvuutta. Pääkategoriat päivittämisvaihtoehdoille olivat 1) mallin kalibrointi, 2) mallin korjaaminen ja 3) mallin laajentaminen.

Voimme alkuun karsia yhden menetelmän, joka kuuluu kategoriaan 1) mallin kalibrointi. Malli ei ota huomioon uutta muuttujaa vaan soveltuu pikemminkin uuteen ympäristöön päivittämiseen. Kuten luvussa 3.2.1 todettiin, pelkän vakiotermin kalibrointi perustuu oletukseen siitä, että selittävien tekijöiden suhteellinen ennustamisvahvuus on lähestulkoon samanlainen alkuperäisessä opettamisdatassa ja uudessa datassa. Lisätessä uutta selittävää muuttujaa yritys X:n riskimalliin, muuttujan ennustamisvahvuus olisi toivottavasti korkea. Tästä syystä emme voi yritys X:n tapauksessa hyödyntää tätä menetelmää. Siirrytään muiden vaihtoehtojen tutkimiseen.

Mallin korjaaminen eli päivittämissvaihtoehto 2) ei myöskään välttämättä ota uutta muuttujaa tarpeeksi tarkasti huomioon, joten jätämme sen yritys X:n esimerkin käsittelyn takia huomiotta. Tästä menetelmästä ei kirjallisuudesta myöskään löydy suosituksia uutta muuttujaa lisätessä. Mallin korjaaminen eli estimoiminen voisi tapahtua tilanteessa, jossa sen päivittämistä uuteen otantaan harkittaisiin. Täten riskimallin luokittelutarkkuutta voitaisiin arvioida verraten edelliseen testiin luokittelutarkkuudesta. Kuitenkin, uuden muuttujan lisäämiseen mallin korjaaminen ei tässä tapauksessa välttämättä sovellu. Jäljelle jää mallin laajentaminen, joka huomioi lisättävän muuttujan.

5.3.1 Mallin laajentaminen

Shipe et al. (2019, 582) tutkimuksessa tuodaan esiin suositus mallin laajentamisesta uuteen dataan, kun logistiseen regressioon lisätään uutta muuttujaa. Tämä pohjautuu siihen, että tavoitteena on kehittää logistinen regressio mahdollisimman suurella datajoukolla. Päivittämissvaihtoehto 3) eli mallin laajentaminen tuotiin esiin luvussa 3.2.1. Koska tässä työssä etsitään keinoa lisätä malliin dataa yritysjoukkokohtaisesti, mallin laajentaminen voisi tapahtua yhdistämällä yritysjoukon data osaksi yritys X:n alkuperäistä mallia. Yritysjoukon data sisältäisi alkuperäisten muuttujien lisäksi uuden lisättävän muuttujan datan. Tässä työssä käsittelemme vain yhden eikä useamman muuttujan lisäystä yksinkertaisuuden vuoksi. Yritysjoukon dataan viitataan uutena datana.

Steyerberg (2019, 402) mukaan mallin laajentaminen jakautuu kolmeen vaihtoehtoon:

- 1) *kalibrointi, valikoiva uudelleenestimointi ja valikoiva laajentaminen*
- 2) *uudelleenestimointi ja valikoiva laajentaminen*
- 3) *uudelleenestimointi ja laajentaminen*

Käsitellään ensin näistä ensimmäinen vaihtoehto.

$$Y = \alpha + \beta_i * x_i \quad (1)$$

$$Y_1 = \alpha_{uusi} + \beta_{yleinen} * Y \quad (3)$$

(Mukaillen Steyerberg 2019, 401-403)

Ensin suoritetaan logistinen kalibrointi eli päivitetään vakiotermin α sekä yleinen regressio-kerroin β suorittamalla logistinen regressio uudessa datassa (Y kovariaattina), jolloin saadaan yhtälö Y_1 . Alkuperäistä mallia kuvaa Y, esimerkkinä yritys X:n riskimalli.

Shipe et al. (2019, 582) mukaan yksittäisen muuttujan vahva ennustamiskyky ei takaa lisättyä ennustamiskykyä yhdessä toisen mallin kanssa. Seuraavaksi siis testataan kaikkien muuttujien tilastollista merkitsevyyttä uudessa datassa, ja valitaan malliin p-arvoltaan riskitasoa pienemmät tai yhtä suuret muuttujat. Riskitaso voi olla esimerkiksi 0,05 eli 5 prosenttia.

$$Y_2 = \alpha_{uusi} + \beta_{yleinen} * Y + \gamma_i * X_i + \gamma_z * Z \quad (4)$$

(Mukaillen Steyerberg 2019, 403)

Uudet regressiokertoimet γ saadaan ajamalla logistinen regressio uudessa datassa, jossa Y_1 toimii kiinteänä muuttujana (offset variable). Tällöin saadaan Y_2 , jossa X_i kuvastaa mallin merkitseviä ja alkuperäisiä muuttujia sekä Z merkitsevää, uutta muuttujaa.

Nieboer et al. (2016, 3-4) ehdottavat samanlaista menetelmää uuden merkkimuuttujan (marker variable) lisäämiseen. Menetelmä on nimeltään *kalibrointi ja laajentaminen*. Merkkimuuttuja on muuttuja, joka on vahvasti yhteydessä selitettävään muuttujaan, mutta ei teoreettisesti liity muihin selittäviin muuttujiin (Simmering, Fuller, Richardson, Ocal, & Atinc 2015, 474). Tällainen muuttuja voisi olla yritys X:n tapauksessa esimerkiksi maksuviive. Nieboer et al. tutkimuksen malli esitetään yksinkertaisemmin kuin Y_2 , ilman selittävien muuttujien tilastollisen merkitsevyyden testaamista.

$$Y_3 = \alpha_{uusi} + \beta_{yleinen} * Y + \gamma_z * Z \quad (5)$$

(Mukaillen Nieboer et al. 2016, 4)

Vaihtoehtoisesti mallin laajentamisen voisi toteuttaa vaihtoehdolla 2) eli *uudelleenestimoinnalla ja valikoiden laajentamalla* tai 3) *uudelleenestimoinnalla ja laajentamalla*. Molemmissa vaihtoehdoissa estimoidaan koko malli uuteen dataan, mutta 2) vaihtoehdossa malliin lisätään vain uudessa datassa tilastollisesti merkitsevät muuttujat.

$$Y_4 = \alpha_{uusi} + \beta_{uusi} * X_i * Z \quad (6)$$

(Mukaillen Steyerberg 2019, 403)

Yritys X:n malli olisi molemmissa tapauksissa Y_4 , sillä tässä tutkimuksessa tutkitaan yhden muuttujan lisäämistä. Ainoa ero menetelmissä on tilastollisen merkitsevyyden testaaminen. X kuvastaa alkuperäisen datan muuttujia, Z uutta muuttujaa ja β_{uusi} muuttujien regressioker-toimia. Steyerberg (2019, 399) toteaa, että suuria muokkauksia, kuten uudelleenestimointia pitäisi yrittää vain suurella validaatiojoukolla. Validaatiojoukolla hän viittaa uuteen datajouk-koon. Validaatiodatan väljä määrittely kirjallisuudessa näkyy ajoittain useiden kirjoittajien teksteissä. Steyerbergin mukaan yksi riski, joka liittyy mallin korjaamiseen tai laajentami- seen liian pienessä päivittämisdatabaasissa, on luotettavien estimaattorien korvautuminen epä- luotettavilla estimaattoreilla. Koska yritys X:n alkuperäinen malli on lähtökohtaisesti luotet- tava, mallin rakenne muuttuisi paljon, jos se uudelleenestimoidaan mahdollisesti pienestä datajoukosta. Uusi data, joka tässä teoreettisessa tapauksessa koostuu yhdestä yritysjou- kosta, on todennäköisesti moninkertaisesti pienempi kuin yritys X:n riskimallin opetta- misdata. Vedoten pienen datajoukon riskeihin, joita käsiteltiin luvussa 5.2, olisi Y_4 hyvin harkinnanvarainen menetelmä. Suuri otanta on hyvä lähtökohta uudelleenestimoinnille, jo- ten yritys X:n tapauksessa olisi harkittava tapauskohtaisesti Y_4 -yhtälöllä sovellettavia yri- tysjoukkojen riskimalleja.

Esimerkkinä toimivasta uudelleenestimoinnin tapauksesta, Mileris & Boguslauskas (2011, 129) tutkimuksessa perehdytään luottoriskimallin kehittämiseen. Tutkimuskohteena on lo- gistinen regressio, joka pohjautuu 100 liettualaisen yrityksen 20 tunnuslukuun ja viiteen tili- kauteen. Kun regressiomalliin haluttiin lisätä 100 uuden yrityksen ajankohtaisemmat tun- nusluvut, malli päivitettiin laajentamalla opettamisdata uuteen dataan ja estimoimalla malli uudestaan. Luokittelun tarkkuus kasvoi 82,2 prosentista 93,4 prosenttiin. Tässä tapauk- sessa uuden datajoukon koko oli suhteellisesti laaja. Estimoinnin mahdollistama luokittelun tarkkuuden parantuminen korostaa sitä, kuinka toimiva mallin uudelleenestimointi on koh- tuullisen datajoukon kohdalla.

5.4 Yritys X:n vaihtoehdot

Voidaan todeta, että suurempaa datajoukkoa lisätessä malliin, avautuu enemmän vaihto- ehtoja päivittämissvaihtoehtojen suhteen. Yksinkertaisuudessaan tämä tarkoittaa sitä, että jos meillä on puitteet estimoida logistinen regressio uudelleen, kannattaa se tehdä. Kaiken kaikkiaan, logistisen regression päivittämismenetelmä riippuu vahvasti lisättävästä datasta.

Pieni lisättävä datajoukko suosii vähemmän mallin muokkausta, ja isompi datajoukko enemmän.

Lähtökohtaisesti mallin uudelleenestimointi on suositeltu ja yksinkertainen päivittämismenetelmä. Kuitenkin, se vaatii kohtuullisen otannan tuottaakseen luotettavia tuloksia ja ennen kaikkea välttääkseen ylisovittamista. Jos käsitellään pieniä lisättäviä datajoukkoja, ovat hyllytyt päivittämismenetelmät todistettu kirjallisuudessa laajoja päivittämismenetelmiä paremmiksi ylisovittamisen riskin kannalta. Kaiken kaikkiaan, *kalibrointi*, *valikoiva uudelleenestimointi ja valikoiva laajentaminen* sekä *kalibrointi ja laajentaminen* ovat huomionarvoisia päivittämismenetelmiä yritys X:n tapauksessa. Näitä kuvastavat mallit Y_2 ja Y_3 . *Uudelleenestimointi ja valikoiva laajentaminen* eli malli Y_4 on harkinnanvarainen päivittämismenetelmä yritys X:n tapauksessa, sillä se vaatii riittävän suuren yritysjoukon datan luotettavia tuloksia varten. Tässä tutkimuksessa on myös syytä punnita logistisen regression päivittämismenetelmän soveltuvuutta lääketieteestä yritysriskimalliin. Teoriassa, esitetyt vaihtoehdot ovat päteviä myös yritysriskiä mallintaessa, mutta aiheesta ei ole vielä käytännön esimerkkejä tai kirjallisuutta.

6. MIELENKIINTO YRITYS X:N RISKIMALLIN KEHITTÄMISEEN

Seuraavaksi tarkastellaan tutkimuskysymystä 4. ”Millaista mielenkiintoa ja kysyntää on odotettavissa yritys X:n riskimallin räätälöintiin?”. Kysymystä tutkitaan laadullisen tutkimuksen kautta. Tavoitteena on tutkia yritys X:n riskimallin kehittämismahdollisuuksia kartoittamalla mahdollista kysyntää.

6.1 Tutkimusmenetelmä ja aineisto

Tässä osiossa esitellään lyhyesti tutkimuksessa käytetty laadullinen tutkimusmenetelmä ja -aineisto. Aineiston kerääminen ja aineiston analyysi kuvaillaan. Laadullinen tutkimus suoritetaan tässä työssä puolistrukturoidun haastattelun kautta. Laadullinen tutkimus ei ole tutkimusmenetelmänä tieteellinen eikä testaa hypoteeseja, mutta tavoittelee ymmärrystä siitä, miten osallistujat kokevat tutkimusaiheen. Vaikka haastattelun vastaukset ovat aina subjektiivisia, vastaukset voivat avata tutkimusaiheen merkitystä ja linkittymistä muihin aiheisiin. (Kvale 2007, 84-87.) Haastateltavina toimivat kaksi yrityksen työntekijää, joilla on käsitys riskimallista ja sen asiakkaista. Käytännön kokemukset ja näkemykset sopivat erinomaiseksi vastapainoksi työn teoreettiseen tarkasteluun, ja työntekijöiden omat huomiot ovat arvokkaita työn kehityksen kannalta. He osaavat tuoda esiin huomioita sekä omasta että asiakkaidensa puolesta.

Haastattelu toteutettiin puolistrukturoituna videohaastatteluna. Puolistrukturoidun haastattelun idea on laatia kysymykset ennakkoon ja esittää ne kutakuinkin samassa muodossa, mutta vastaaminen on vapaamuotoista. Ideana on antaa haastateltavien vastata kysymyksiin niin laajasti kuin on tarve. Käsiteltävät aiheet on listattu ennakkoon, mutta haastattelussa on avoimuutta muutoksiin. (Kvale 2007, 60-63.)

Haastattelukysymykset ovat seuraavanlaiset:

Onko asiakkailla ollut puhetta tai mielipiteitä riskimallin kehityksestä?

Millaisia uusia muuttujia voisi hyödyntää riskiluokituksessa?

Mitä kehityskohtia riskimallissa on?

Mitkä ovat yritys X:n riskimallin vahvuudet?

6.2 Tutkimuksen tulokset

Tässä osiossa käydään läpi haastattelun tuloksia. Haastattelukysymyksiä oli 4, ja ne esitettiin haastateltaville videohaastattelussa. Haastattelu tallennettiin äänittämällä ja tämän jälkeen tallennus litterointiin haastattelukysymysten ja niiden vastauksien osalta. Litterointi tarkoittaa puheen ja toiminnan purkamista kirjoitettuun muotoon. Litterointi on keskeinen osa laadullisen aineiston haltuunottoa ja analyysiprosessia. Tämän jälkeen suoritettiin laadullinen sisällönanalyysi. Testissä esiintyvät numerot 1 ja 2 viittaavat haastateltaviin eli yritys X:n työntekijöihin. Ensimmäinen kysymys tiedustelee mahdollista asiakaskysyntää räätälöidylle riskimallinnukselle.

”Onko asiakkailta ollut puhetta tai mielipiteitä riskimallin kehityksestä?”

Tämän kysymyksen avulla koitetaan paneutua aiheeseen asiakkaiden kautta – tavoitteena kartoittaa riskimallin kehittämismahdollisuutta. Haastateltavien vastaus kysymykseen on kyllä, ja he tuovat esiin, että myös riskimallin potentiaaliset asiakkaat ovat esittäneet toiveita riskimallin kehityksestä. Vastauksesta voidaan tulkita, että riskimallin kehittäminen hyödyttäisi nykyisten asiakkaiden lisäksi tulevia asiakkaita. Vastauksista on tulkittavissa riskimallin kehitysvara, jota tutkitaan tarkemmin seuraavien kysymysten kautta.

”Millaisia uusia muuttujia voisi hyödyntää riskiluokituksessa?”

2: *”Maksuviivekanta ja maksuhäiriötiedot ovat ehkä ne kaikkein oleellisimmat oman näkökseni mukaan tai ainakin sellaiset, mitä asiakkaat kommentoivat.”*

1: *”Olen ainakin samaa mieltä siitä, että nämä tekijät asiakkaat pääasiassa haluaisivat riskiluokitukseen mukaan.”*

Maksuviive ja maksuhäiriöt ovat yritys X:n riskimallista puuttuvia tekijöitä, jotka esiintyvät kirjallisuudessa vahvasti maksuhäiriötekijöinä. Maksuviive sisältyy tässä työssä

maksukäyttäytymisen kategoriaan maksuhäiriöriskitekijöiden taulukossa. Vastauksista käy ilmi, että niille olisi kysyntää yritys X:n riskimallissa. Maksuviive ja maksuhäiriöt nousevatkin esiin haastattelun aikana useamman kerran. Haastateltavat mainitsevat myös, että jotkin asiakkaat saattavat seurata omin menetelmin asiakkaidensa maksuviiveitä. Lähtökohtaisesti hyvä riskimalli toimii niin, että asiakkaan on itse haettava muualta mahdollisimman vähän tietoa. Nämä tekijät voitaisiin haastattelun perusteella sisällyttää yritys X:n riskimalliin.

2: *”Toinen mitä on maksuhäiriöiden ja maksuviiveen lisäksi on kysytty, on se, että olisi vastuuhenkilötaustaa. Siinä olisi yksi tekijä, miltä osin mallin osumatarkkuutta voisi parantaa.”*

Kolmas mahdollisesti lisättävä tekijä, vastuuhenkilötausta, viittaa yrityksen vastuuhenkilön historiaan esimerkiksi aikaisempien yritysten menestyksen kautta. Useampi aiempi konkurssi voi viitata nykyisen yrityksen mahdollisesti heikkoon suoritukseen. Kaiken kaikkiaan, kolme haastattelussa esiintyvää mielenkiintoista muuttujaa ovat maksuviive, maksuhäiriöhistoria ja vastuuhenkilötausta, joista kaksi ensimmäistä esiintyvät tärkeimpinä tekijöinä. Näitä voisi olla syytä hyödyntää yritys X:n riskimallissa. Kyseisten muuttujien sisällyttäminen riskimalliin voi riippua muuttujien tilastollisesta merkitsevyydestä logistisessa regressiossa.

”Mitä kehityskohtia riskimallissa on?”

Koska yritys X:n riskimalli perustuu pitkälti tilinpäätösdataan, on malli myös ajallisesti sidottu tilinpäätösten julkaisuajankohtiin. Haastateltavat esittävät mahdollisena kehityskohteena riskimallin reagoitakyvyn, joka voitaisiin heidän mukaansa saavuttaa maksuviive- ja maksuhäiriöiden lisäämisen lisäksi yrityksen tilaa tarkkailemalla. Saneeraus- ja muita yrityksen tilan muutoksia voitaisiin seurata mallissa, jolloin riskimalli olisi ennakoivampi.

1: *”Monet asiakkaat, varsinkin talousjohtajat tai controllerit kysyvät sitä, että miten nopeasti tämä riskiluokka reagoi muutoksiin ja miten reaaliaikainen se on.”*

Esiin tuodaan yleinen huoli riippuvaisuudesta tilinpäätöksiin. Pieni yritys, jolla ei ole tilinpäätöstä jää ulos riskimallista, sillä riskimallin muuttujat ovat pitkälti tilinpäätöstietoa. Tähän voisi toimia ratkaisuna jonkinlainen kompromissi – osa tiedosta voisi sisältyä malliin,

esimerkiksi maksuviive- ja maksuhäiriötieto. Näillä tekijöillä on sekä kirjallisuuden että haastateltavien mukaan suuri vaikutus maksuhäiriöriskiin.

Kehitysehdotusten lisäksi haastattelussa esiin tulee myös mahdollinen uudelleenbrändäyksen tarve, mikä olisikin luontainen prosessi tuotteen kehittämisessä. Jos malliin sisällytetään lisää muuttujia ja päivitetään muitakin ominaisuuksia, on syytä kommunikoida kehityksestä sekä nykyisille että potentiaalisille asiakkaille. Uudelleenbrändäystä voisi pohtia tuotteen kehityksen yhteydessä tai jälkeen.

”Mitkä ovat yritys X:n riskimallin vahvuudet?”

Riskimallin vahvuuksina tuodaan esiin sen yksinkertaisuus ja helppokäyttöisyys. Riskimalli ja sen käyttöympäristö ovat visuaalisesti miellyttäviä, ja tuotteen yleisilmettä on keuhuttu paljon. Riskien mallintaminen on kokonaisuutena melko monimutkainen prosessi, ja myös yritys X:n riskimallissa riskitekijöitä on monia. On tärkeää osata tiivistää kerätty tieto yksinkertaiseen muotoon. Tämä on osattu tehdä hyvin yritys X:n riskimallissa haastateltavien vastauksien mukaan.

Haastattelun tulokset voidaan tiivistää niin, että haastateltavien mukaan kehitysvara on olemassa, ja uusia muuttujia voitaisiin lisätä malliin. Kysyntää kehitetylle riskimallille on odotettavissa. Riskitekijöiden ei tarvitse rajautua tilinpäätösdataan vaan pikemminkin muuta dataa kannattaisi hyödyntää enemmän. Mielenkiintoista on haastattelun tuloksien vertailtavuus kirjallisuuden tekijöihin – reaaliaikaisuus korostuu enemmän haastateltavien vastauksissa. Muuttujat, jotka lisäävät riskimallin reagoitokykyä yrityksen riskitason muutokseen, lisäävät riskimallin kysyntää. Asiakaskontaktit ovat tuoneet esiin toiveen reagoitokyvystä, ja tähän toiveeseen voidaan vastata usean muuttujan avulla. Maksuviive, maksuhäiriöhistoria, vastuuhenkilötausta ja yrityksen tila ovat pitkälti tilinpäätöksen ulkopuolista tietoa, jotka mahdollisesti viestivät riskitasoa tehokkaasti ja reaaliaikaisesti. Kyseiset tekijät, tai osta niistä, voisivat lähtökohtaisesti toimia lisättävinä muuttujina riskimalliin. Mitä tärkeintä, toivottu käsitys asiakaskysynnästä vahvistui haastattelun kautta. Itse tuotteen päivittäminen vaatii varmuuden siitä, että päivitetylle tuotteelle on kysyntää. Haastattelun perusteella voidaan vahvistaa käsitystä siitä, että riskimalliin lisättäville muuttujille olisi mahdollisesti kysyntää, sillä keskustelua kehityskohdista on jo käyty asiakkaiden kanssa.

7. JOHTOPÄÄTÖKSET

Tässä tutkimuksessa käsiteltiin yritysten maksuhäiriötekijöitä ja tutkittiin yritys X:n riskimallinnuksen työkalun kehittämismahdollisuuksia. Kirjallisuudesta selvitetty maksuhäiriötekijät perustuvat pitkälti taloudelliseen ja määrälliseen tietoon, ja tilinpäätöstieto on kirjallisuuden mukaan avainasemassa yrityksen maksuhäiriön todennäköisyyden määrittämisessä. Maksuhäiriöihin positiivisesti tai negatiivisesti vaikuttavia tekijöitä ovat muun muassa yrityksen toimiala, koko, ikä, kannattavuus, likviditeetti ja pääomarakenne. Muualta kuin tilinpäätöksistä saatavia tekijöitä ovat maksukäyttäytyminen, maksuhäiriöhistoria, johtamisen tehokkuus, hallinnointikoodi, markkinatilanne, pörssidata ja makrotaloudelliset tekijät. Yritys X:n riskimallista jäivät huomiotta esimerkiksi maksukäyttäytyminen ja maksuhäiriöhistoria.

Logistista regressiota tutkittiin maksuhäiriön ennustamisen työkaluna, ja samalla tutkittiin logistisen regressiomallin päivittämisvaihtoehtoja. Lääketieteen kirjallisuutta hyödynnettiin vaihtoehtojen kartoittamisessa. Päivittämisvaihtoehtoja on kalibrointi, estimointi ja laajentaminen, joista tutkittiin mallin laajentamista yritys X:n riskimallin tapauksessa. Tavoitteena oli arvioida keinoja lisätä uutta dataa yritys X:n riskimalliin. Lähtökohtaisesti mallin uudelleenestimointi on suositeltu päivittämismenetelmä, mutta uudelleenestimointi vaatii suhteellisen suuren määrän uutta dataa, jotta malli voidaan kehittää luotettavaksi. Kirjallisuuden mukaan 500 suuruisen otannan malli, jossa on vähintään 10, ja mieluummin yli 20, näytettä muuttujaa kohden on suositeltu ohje mallin rakentamiseen. Tämä parantaa ennusteiden tarkkuutta. Päivitettäessä mallia pienellä datajoukolla, hillityt päivittämismenetelmät vähentävät mallissa ylisovittamista eli liian optimistisia ennusteita. *Mallin kalibrointi, valikoiva uudelleenestimointi ja laajentaminen* eli Y_2 tai *mallin kalibrointi ja laajentaminen* eli Y_3 todettiin mahdollisesti toimiviksi päivittämismenetelmiksi yritys X:n tapauksessa. *Uudelleenestimointi ja valikoiva laajentaminen* eli Y_4 voisi olla harkinnanvarainen menetelmä, mutta lisätävän datajoukon täytyy olla riittävän suuri. Vasta mallin päivittämisen myötä voidaan arvioida sen muuttujien tilastollista merkitsevyyttä, mallin luokittelutarkkuutta ja muita ominaisuuksia. Tässä vaiheessa voidaan siis vain todeta, että kyseinen päivittämismenetelmä on teoriassa mahdollinen yritysriskimallissa.

Haastattelun mukaan yritys X:n riskimallin kehittämiseksi olisi odotettavissa mielenkiintoa ja kysyntää. Haastateltavat suosivat itse riskimallin kehittämistä ja päivittämistä, ja osoittivat myös esimerkein yritys X:n asiakkaiden mielenkiintoa kehittämiseen. Kehityskohtia olisi

riskimallin reagointikyky, reaaliaikaisuus ja uusien tekijöiden lisääminen. Haastateltavien toive oli lisätä tilinpäätösten ulkopuolista dataa, jotta riskimalli saavuttaisi ketteryyttä riskiluokituksen arviointiin. Esiin tuotuja lisättäviä muuttujia olivat maksuviive, maksuhäiriöhistoria, vastuuhenkilötausta ja yrityksen tila.

Tässä työssä mainitut kirjallisuuden maksuhäiriötekijät ovat hyvin luotettavia, sillä kyseiset tekijät esiintyvät useissa tieteellisissä lähteissä ja jokaisen tekijän vaikutus yritysten maksuhäiriöihin on testattu. Toisaalta, toiset tekijät esiintyvät kirjallisuudessa enemmän kuin toiset – esimerkiksi toimialaa mainittiin pidemmällä aikavälillä enemmän kuin maksukäyttäytymistä. Kuitenkin, tässä työssä taulukoidut yritysten maksuhäiriötekijät ovat hyvin vertailtavissa nykyiseen kirjallisuuteen. Voidaan todeta, että tuloksiin voidaan luottaa ja tietoa yleisesti hyödyntää. Laadullisen tutkimuksen haastattelun tuloksia voi hyödyntää tässä tapauksessa, pitäen mielessä tulosten subjektiivisuuden. Yritys X:n työntekijöiden vastaukset tukevat näkemystä mahdollisesta asiakaskysynnästä päivitettyyn riskimalliin. Haastatelluista ei kuitenkaan voida tehdä suoria johtopäätöksiä, sillä laadullisessa tutkimuksessa ei testata hypoteeseja. Käytännön kokemus ja haastateltavien näkemykset kuitenkin syventävät ymmärrystä aiheesta yritys X:n tapauksessa.

Logistisen regression päivittämismenetelmien arvioiminen on haasteellisempaa, sillä mallin päivittäminen on tapauskohtaista. Logistisen regression päivittämisestä on melko vähän kirjallisuutta, mikä ehdottomasti vaikuttaa tulosten yleistettävyyteen. Suuri osa tutkimuksista löytyy lääketieteestä, jossa ennustetaan todennäköisyyksiä esimerkiksi erilaisiin sairauksiin. Perusidea on sama – tiettyä riskiä mallinnetaan usean selittävän tekijän kautta. Kuitenkin, tutkiminen tapahtuu kriittisellä silmällä ja tutkimusta logistisen regression yritysriskimallin päivittämisestä on tehty teoreettisella tasolla. Löydettyjen päivittämismenetelmien soveltaminen riskimalliin on harkinnanvaraista, ja riippuu paljon lisättävästä datasta. Päivittämismenetelmiin liittyviä kriteerejä ja ominaisuuksia on vertailtu kriittisesti yritys X:n tapauksessa. Tärkeinä lähteinä toimineet Steyerberg (2019) sekä Nieboer et al. (2016) tutkimukset tukevat tässä työssä esitettyjen riskimallin päivittämismenetelmien toimivuutta. Päivittämismenetelmien soveltaminen yritysriskimalleissa jää kuitenkin testattavaksi käytännössä.

Tämän tutkimuksen tulosten vertailu kirjallisuuden muihin vastaaviin tuloksiin on haastavaa kirjallisuuden vähäisyyden takia. Vähäinen kirjallisuus logistisen regression päivittämisestä on toisaalta iso syy tämän tutkimuksen toteuttamiselle. Tässä työssä on löydetty teoriassa sovellettavia päivittämismenetelmiä yritysriskimalliin. Teoreettisesti toimivaa

päivittämismenetelmää on syytä koettaa soveltaa yritysriskimallin tapauksessa. Päivittämissä on huomioitava lisättävän datajoukon koko ja rakenne. Lisättäviä muuttujia ovat muun muassa maksuviive ja maksuhäiriöhistoria. Jatkotutkimusta aiheesta olisi ehdottomasti tässä tutkimuksessa mainittujen logistisen regression päivittämismenetelmien soveltaminen yritysriskimalleissa.

LÄHDELUETTELO

Averkamp, H. (2021) Accounting Coach: What is the difference between the current ratio and the quick ratio?

[Verkkolähde] [Viitattu 4.11.2021] Saatavissa:

[What is the difference between the current ratio and the quick ratio \(accountingcoach.com\)](https://www.accountingcoach.com)

Berry, M., Azlinah, M., Bee, W. (2020) *Supervised and Unsupervised Learning for Data Science*. 1st ed. 2020. Cham: Springer International Publishing. s.4

Bolton, C. (2009). Logistic regression and its application in credit scoring. s. 19-45

Bonfim, D. (2009) Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of banking & finance*. [Online] 33 (2), s. 285–298

Boss, M., Fenz, G., Pann, J., Pühr, C., Schneider, M., & Ubl, E. (2009). Modeling credit risk through the Austrian business cycle: An update of the OeNb model. *Financial Stability Report*, 17, s.85

Bujang, M., Sa'at, N., Sidik, T., & Joo, L. (2018). Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. *The Malaysian journal of medical sciences: MJMS*, 25(4), s.123

DOUMPOS, M., LEMONAKIS, C., NIKLIS, D. and ZOPOUNIDIS, C., 2019. *Analytical Techniques in the Assessment of Credit Risk An Overview of Methodologies and Applications*. 1 Edition. Cham: Springer International Publishing. s.1-81

Dun & Bradstreet (2021) Maksuhäiriömerkinnän poistuminen

[Verkkolähde] [Viitattu 4.11.2021] Saatavissa:

[Maksuhäiriömerkinnän poistuminen \(dnb.com\)](https://www.dnb.com)

Dwyer, D., Kocagil, A., Stein, R., 2004. The Moody's KMV EDF Riskcalc v3.1 Model. Moody's KMV Company s.13

Finanssivalvonta, 2018. Määräykset ja ohjeet 4/2018: Luottoriskien hallinta ja luottokelpoisuuden arviointi rahoitussektorin valvottavissa. s.5

[verkkodokumentti] [viitattu 27.10.2021] Saatavissa:

[Luottoriskien hallinta ja luottokelpoisuuden arviointi rahoitussektorin valvottavissa \(finanssivalvonta.fi\)](https://www.finanssivalvonta.fi/luottoriskien-hallinta-ja-luottokelpoisuuden-arviointi-rahoitussektorin-valvottavissa)

Gustin, D. (2014) Even sellers must manage credit risk like banks

[Verkkolähde] [Viitattu 11.10.2021] Saatavissa:

[Even sellers must manage credit risk like banks \(spendmatters.com\)](https://www.spendmatters.com/even-sellers-must-manage-credit-risk-like-banks)

Intrum, 2018, European Payment Report 2018.

[Verkkolähde] [Viitattu 11.10.2021] Saatavissa:

[European Payment Report 2018 \(www.intrum.com\)](https://www.intrum.com/european-payment-report-2018)

Kaakinen, M., Ellonen, M. Logistinen regressio. Teoksessa *Kvantitatiivisen tutkimuksen verkkokäsikirja*. Tampere: Yhteiskuntatieteellinen tietoarkisto [Viitattu 11.11.2021.] Saatavissa:

[Logistinen regressio \(fsd.tuni.fi\)](https://fsd.tuni.fi/logistinen-regressio)

KOULAFETIS, P., 2017. *Modern Credit Risk Management Theory and Practice*. London: Palgrave Macmillan UK. s.2-125

Kvale, S. 2007 *Doing Interviews*. SAGE Publications Ltd s.60-87

Li, K., Niskanen, J. & Niskanen, M., 2019. Capital structure and firm performance in European SMEs: Does credit risk make a difference? *Managerial finance*, **45**(5), s. 594.

Mileris, R., & Boguslauskas, V. (2011). Credit risk estimation model development process: Main steps and model improvement. *Inžinerinė ekonomika*, **22**(2), s.129.

Nieboer, D., Vergouwe, Y., Ankerst, D., Roobol-Bouts, M., & Steyerberg, E. (2016) Improving prediction models with new markers: A comparison of updating strategies. *BMC medical research methodology*. [Online] **16** (1), s.1-9.

Niemi, M. 2014. Luotto-oikeus: luottotyypit, perintäprosessit ja takaisinsaanti. s.1

POBLACIÓN GARCÍA, F.J., 2017. *Financial Risk Management Identification, Measurement and Management*. Cham: Springer International Publishing. s.13-211

Premachandra I., Gurmeet B., Toshiyuki S., 2009, p.419 *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique* s.419

SALIMI, A., 2015. Validity of Altmans Z-score model in predicting bankruptcy in recent years. *Academy of Accounting and Financial Studies journal*, 19(2), s. 233.

Shipe, M., Deppen, S., Farjah, F., & Grogan, E. (2019). Developing prediction models for clinical use using logistic regression: an overview. *Journal of thoracic disease*, 11(Suppl 4), s.582.

Siekelova A., Kollar B., Weisssova I., 2015, Impact of credit risk management s.325-327

Simmering, M., Fuller, C., Richardson, H., Ocal, Y., & Atinc, G. (2015). Marker Variable Choice, Reporting, and Interpretation in the Detection of Common Method Variance: A Review and Demonstration, *Organizational Research Methods*, 18(3), s. 474

Steyerberg, E., 2019. *Clinical Prediction Models A Practical Approach to Development, Validation, and Updating*. 2 edn. Cham: Springer International Publishing. s.399-413

Steyerberg, E., Harrell F., Borsboom, G., Eijkemans, M., Vergouwe, Y., Habbema, J. (2001) Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*. [Online] 54 (8), s.780

STOLTZFUS, J.C., 2011. Logistic Regression: A Brief Primer. *Academic Emergency Medicine; Acad Emerg Med*, 18(10), s.1099.

Vergouwe, Y., Nieboer, D., Oostenbrink, R., Debray, T. , Murray, G., Kattan, M., Koffijberg, H., Moons, K., & Steyerberg, E. (2017) A closed testing procedure to select an appropriate method for updating prediction models. *Statistics in medicine*. [Online] 36 (28), s.4529

Vero, 2021, Konserniavustus

[Verkkolähde] [Viitattu 9.11.2021] Saatavissa:

[Konserniavustus \(vero.fi\)](https://vero.fi/konserniavustus)

WANG, J., LAN, Y., ZHAO, R. and XING, T., 2018. The optimal payment policy for a firm: cash sale versus credit sale. *Soft computing (Berlin, Germany)*, 22(17), s.5843.