

Santtu Kokko

**TILINPÄÄTÖSPETOSTEN HAVAITSEMINEN  
ALGORTIMIPOHJAISTEN MENETELMIEN AVULLA**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2024

# TIIVISTELMÄ

Kokko, Santtu

Tilinpäätöspetosten havaitseminen algoritmipohjaisten menetelmien avulla

Jyväskylä: Jyväskylän yliopisto, 2024, 38 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Riekkinen Janne

Tilinpäätöstietoja koskevat petokset ovat yksi talouspetosten merkittävimmistä ja huomattavimmista petosmuodoista. Tilinpäätöspetoksia voidaan pitää erityisen merkittävänä taloudellisen petoksen tyyppinä, koska niiden aiheuttamat taloudelliset tappiot ovat muihin taloudellisiin petoksiin verrattuna hyvin suuret, ja tilinpäätöspetokset aiheuttavat paljon negatiivisia vaikutuksia monille eri sidosryhmille. Tutkimukset ovat myös osoittaneet, että perinteisesti ihmisen tekemä tilinpäätösten tarkastaminen on epätarkkaa ja aikaa vievää sekä vain suhteellisen pieni osuus petoksista onnistutaan havaitsemaan. Lisääntyneen petosten määrän takia monet eri tahot ovat korostaneet tarvetta tehokkaaseen tilinpäätöspetosten havaitsemiseen. Tilinpäätöspetosten havaitsemiseksi onkin kehitetty erilaisia älykkäitä ja algoritmeihin pohjautuvia menetelmiä, joiden avulla pyritään tehostamaan tilinpäätöspetosten havaitsemista. Tässä kandidaatintutkielmassa tutkittiin, miten erilaiset algoritmeihin pohjautuvat menetelmät, kuten koneoppimisen, syväoppimisen ja datanlouhinnan menetelmät, suoriutuvat tilinpäätöspetosten havaitsemisessa ja kuinka menetelmien avulla yritykset voidaan luokitella tilinpäätöksien perusteella petollisiksi tai ei-petollisiksi. Tutkielman toteutettiin kuvailevana kirjallisuuskatsauksena. Tutkielmassa selvisi, että tilinpäätöspetosten havaitsemiseen käytetyillä petosindikaattoreilla on huomattava merkitys siihen, kuinka tarkasti yrityksiä voidaan luokitella oikein petollisiksi tilinpäätöksien perusteella. Huomioitavaa esimerkiksi oli, että taloudellisten tunnuslukujen käytön lisäksi oikeinluokittelun tarkkuutta lisää myös ei-taloudelliset tunnusluvut, kuten johdon kommentit liiketoiminnan suoriutumisesta. Yksittäisistä petoksentunnistusmenetelmistä tarkimmaksi osoittautuivat keinotekoisiiin neuroverkkoihin perustuvat syväoppimisen menetelmät, jotka saavuttivat useissa tutkimuksissa lähes 95 % oikeinluokittelutarkkuuden tilinpäätöspetoksissa.

Asiasanat: tilinpäätöspetos, algoritmi, koneoppiminen, syväoppiminen, datanlouhinta

## ABSTRACT

Kokko Santtu

Detecting financial statement fraud using algorithm-based methods

Jyväskylä: University of Jyväskylä, 2024, 38 pp.

Information systems, Bachelor's Thesis

Supervisor(s): Riekkinen, Janne

Financial statement fraud is one of the most significant and noteworthy forms of financial fraud. They can be considered highly significant due to their substantial costs compared to other types of financial fraud, and they cause numerous negative effects on various stakeholders. Additionally, studies have shown that traditionally human-based financial statement audits are inaccurate and time-consuming, with only a relatively small portion of frauds being successfully detected. As a result of the increased level of fraud, a number of stakeholders have highlighted the need for effective detection of financial statement fraud. Consequently, a range of intelligent and algorithm-based methods have been developed to improve the detection of financial statement fraud. This bachelor's thesis investigated how different algorithm-based methods, such as machine learning, deep learning and data mining methods, perform in detecting financial statement frauds and how these methods can be used to classify companies as fraudulent or non-fraudulent based on their financial statements. The thesis was conducted as a descriptive literature review. The study found that the fraud indicators used for detecting financial statement frauds significantly influence the accuracy of fraud detection. For example, it was noted that in addition to the use of financial indicators, the accuracy of the classification is also enhanced by non-financial indicators, such as management comments on business performance. Among the individual fraud detection methods, deep learning methods based on artificial neural networks proved to be the most accurate, achieving almost 95% accuracy in financial statement fraud classification in several studies.

Keywords: financial statement fraud, machine learning, deep learning, data mining

## KUVIOT JA TAULUKOT

KUVIO 1	Tilinpäätöspetosten, varojen väärinkäytön ja korruption mediaanitappiot vuosina 2020, 2022 ja 2024 .....	11
KUVIO 2	Tekoälyn, koneoppimisen ja syväoppimisen väliset suhteet sekä niiden keskeiset erot.....	17
TAULUKKO 1	Tutkielmassa esiteltyjen tutkimuksien tulokset taulukkona ....	27

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

KUVIOT JA TAULUKOT

1	JOHDANTO.....	6
2	TILINPÄÄTÖSPETOKSET .....	8
	2.1 Tilinpäätös ja tilintarkastus .....	8
	2.2 Tilinpäätöspetokset .....	9
3	ALGORITMEIHIN POHJAUTUVAT MENETELMÄT.....	12
	3.1 Algoritmit yleisesti .....	12
	3.2 Tekoäly .....	13
	3.2.1 Koneoppiminen .....	14
	3.2.2 Syväoppiminen.....	16
	3.3 Datanlouhinta.....	18
4	TILINPÄÄTÖSPETOSTEN HAVAITSEMINEN .....	19
	4.1 Petosten tunnistus koneoppimistekniikoiden avulla .....	19
	4.1.1 Petosten tunnistus taloudellisista tunnusluvuista .....	19
	4.1.2 Petosten tunnistus taloudellisista ja muista kuin taloudellisista tunnusluvuista .....	22
	4.2 Petosten tunnistus syväoppimistekniikoiden avulla .....	24
	4.3 Petosten tunnistus datanlouhintatekniikoiden avulla .....	25
5	POHDINTA .....	28
6	YHTEENVETO .....	32
	LÄHTEET .....	34

# 1 JOHDANTO

Tilinpäätöstietoja koskevat petokset ovat yleinen ja laajalti tiedostettu yritysten taloudelliseen raportointiin liittyvä ongelma. Nykyisin kaikki taloudelliset petosyritykset ovat lisääntyneet huomattavasti, mikä tekee petosten havaitsemisesta tärkeämpää kuin koskaan (Ashtiani & Raahemi, 2022). Samaan aikaan, kun tilinpäätöspetosten määrä on lisääntynyt, on myös petoksissa menetetyn rahan määrä kasvanut huomattavasti. Crajan, Kimin ja Lessmannin (2020) mukaan tilinpäätöspetosten mediaanitappiot vuonna 2020 olivat 945,000 dollaria per petostapaus. Aihetta voidaankin pitää merkittävänä petosten aiheuttamien suurien tappioiden takia. Tappioiden lisäksi tilinpäätöspetokset aiheuttavat paljon negatiivisia vaikutuksia moniin eri sidosryhmiin, kuten asianomaisiin yrityksiin ja sijoittajiin (Craja ym., 2020).

Samalla kun tilinpäätöstietoja koskevien petosten määrä on lisääntynyt, on myös jatkuvasti pyritty kehittämään tekniikoita niiden havaitsemiseksi. Koska tilinpäätöstä koskevien petosten määrä on lisääntynyt paljon, sijoittajat, akateemiset tutkijat, media, rahoitusyhteisö ja sääntelyviranomaiset ovat korostaneet tarvetta tehokkaaseen tilinpäätöspetosten havaitsemiseen (Wong & Venkatraman, 2015). Tilintarkastusta voidaankin pitää merkittävänä alana, jossa tekoäly ja muut algoritmipohjaiset menetelmät voivat tehostaa ihmisen tekemiä työtehtäviä merkittävästi tulevaisuudessa. Algoritmipohjaiset menetelmät voivat tarjota uuden lähestymistavan havaita erilaisia petoksia, mukaan lukien tilinpäätöspetoksia.

Tämän tutkielman tavoitteena on selvittää, miten algoritmeihin perustuvien menetelmien avulla voidaan havaita tilinpäätöstä koskevia petoksia. Vuonna 2020 laaditussa raportissa on käynyt ilmi, että vain suhteellisen pieni osuus tilinpäätöspetoksista havaittiin, minkä vuoksi on entistä enemmän alettu kiinnittämään huomiota automaattisiin järjestelmiin niiden havaitsemiseksi (Craja ym., 2020). Koska vain suhteellisen pieni määrä tilinpäätöspetoksista havaitaan, olisi tärkeää ymmärtää minkälaisilla tekniikoilla ja algoritmeihin perustuvilla menetelmillä petosten havainnointia voitaisiin tehostaa ja näin ollen muun muassa pienentää petoksissa menetetyn rahan määrää. Aikaisemmat tutkimukset aiheesta keskittyvät usein hyvin yleiselle tasolle taloudellisten petos-

ten havainnointiin, ja vain suhteellisen pieni osuus kirjallisuudesta käsittelee nimenomaan tilinpäätöstä koskevia petoksia. Sen vuoksi onkin hyödyllistä koota ja yhdistää kirjallisuudessa esiintyneiden algoritmipohjaisten menetelmien avulla esiin nousseita seikkoja ja samalla tarjota päivitetty kuva siitä, kuinka tilinpäätöspetoksia voidaan havaita uusilla tekniikoilla. Tutkielma pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

- Millaisilla algoritmipohjaisilla menetelmillä tilinpäätöspetoksia voidaan havaita?
- Mitkä ovat tehokkaimmat lähestymistavat ja menetelmät tilinpäätöspetosten havaitsemiseen?

Tutkielma on toteutettu kuvailevana kirjallisuuskatsauksena aiheesta löytyvän aiempien tutkimuksien ja lähdekirjallisuuden perusteella. Tutkielmassa käytetty lähdekirjallisuus on etsitty Google Scholar, IEEE Xplore, Scopus sekä Science Direct tietokannoista. Pääasiallisina hakutermeinä tiedonhankintaan on käytetty termejä "financial statement fraud", "detection", "artificial intelligence", "machine learning", "deep learning" ja "data mining". Lähdekirjallisuuden valinnassa on kiinnitetty erityistä huomiota julkaisuvuoteen, viittausten määrään sekä lähteen alkuperäiseen julkaisijaan. Ajantasaisen kuvan saamiseksi tutkimuksien tuloksia käsittelevissä artikkeleissa on pyritty valitsemaan vain suhteellisen uusia tutkimuksia, ja siksi tutkielmassa ei ole otettu huomioon yli 10 vuotta vanhoja tutkimuksia. Lähdekirjallisuudesta on myös lähtökohtaisesti karsittu pois ne lähteet, joiden viittauss määrä on hyvin pieni.

Tutkielma koostuu johdannosta, neljästä sisältöluvusta sekä yhteenvetoluvusta. Ensimmäisessä sisältöluvussa käsitellään, mitä tilinpäätökset ja tilinpäätöspetokset ovat. Luvussa määritellään kyseiset käsitteet, ja kerrotaan niiden yleisimmät piirteet. Toinen sisältöluke käsittelee algoritmipohjaisia menetelmiä, joita petosten havaitsemiseen voidaan käyttää. Tässä luvussa määritellään, mitä algoritmit ovat sekä käydään läpi koneoppimisen, syväoppimisen ja datanlouhinnan käytetyimmät menetelmät petosten tunnistukseen. Kolmannessa sisältöluvussa käydään läpi, miten kyseisillä menetelmillä petoksia voidaan havaita. Luvussa tarkastellaan valittujen tutkimuksien tuloksia ja tärkeimpiä esiin nousseita havaintoja. Neljäs sisältöluke on kirjoitettu pohdintalukuna, sisältäen pohdintaa esitetyistä tutkimuksista, lähdekritiikkiä tutkimuksia kohtaan ja jakotutkimusehdotuksia. Viimeinen luku on yhteenvetoluku, joka sisältää yhteenvedon aiheesta ja vastaa tutkimuskysymyksiin.

## 2 TILINPÄÄTÖSPETOKSET

Tässä luvussa tutustutaan tilintarkastukseen ja tilinpäätökseen yleisesti sekä käydään läpi, mitä tilinpäätöspetokset ovat, kuinka niitä tyypillisesti tehdään ja mitä niillä tavoitellaan. Ensimmäinen alaluku keskittyy käsitteiden määrittelyyn, toisessa alaluvussa käsitellään itse tilinpäätöspetoksia.

### 2.1 Tilinpäätös ja tilintarkastus

Tilinpäätöksellä (engl. financial statement) tarkoitetaan yrityksen julkaisemaan asiakirjaa tai laskelmaa, josta selviää yksityiskohtaista tietoa yrityksen taloudellisesta tilanteesta. Tilinpäätöksessä kuvataan esimerkiksi yrityksen tulot, menot, lainat, tuotot sekä myöhemmin mahdollisesti esiin tulevat ongelmat ja johdon kommentit liiketoiminnan suoriutumisesta (West & Bhattacharya, 2016). Tilinpäätöstä voidaan kutsua laajaksi ja kattavaksi kuvaukseksi yrityksen taloudellisesta tilanteesta, ja tilinpäätöksiä käytetään indikoimaan yrityksen suorituskykyä. Tilinpäätökset julkaistaan tyypillisesti vuosittain tai neljännesvuosittain (Ashtiani & Raahemi, 2022).

Tarkemmassa tarkastelussa tilinpäätös koostuu neljästä osasta: tuloslaskelmasta, taseesta, rahavirtalaskelmasta ja liitetiedoista. Tuloslaskelmassa painotetaan yrityksen kuluja ja tuloja tietyltä ajanjaksolta, ja osiossa käy ilmi yrityksen nettotulos eli yrityksen tuloista vähennettävät menot. Tase puolestaan antaa päivitetyn kuvan yrityksen sen hetkisestä velkatilanteesta, varoista ja omasta pääomasta. Rahavirtalaskelmassa mitataan sitä, kuinka yritys onnistuu hankkimaan rahaa toimintakulujen rahoittamiseen, investointien rahoittamiseen ja velkasitoumusten maksamiseen. Selittävät liitetiedot ovat lisätietoja, jotka selventävät ja antavat lisätietoja tietyistä yrityksen tilinpäätöksessä julkaisuista eristä kattaen esimerkiksi myöhempiä tapahtumia ja omaisuuserien pois-toja. (Ashtiani & Raahemi, 2022.)

Tilintarkastuksella tarkoitetaan puolestaan tilinpäätöstietojen tarkastamista ja itse tilinpäätöksen lainmukaisuuden tarkastamista. Tilintarkastusta voi-



daan pitää monimutkaisena ja erittäin erikoistuneena prosessina (Werner, Wiese & Maas, 2021). Craja ym. (2020) mukaan tilintarkastuksen on perinteisesti tehnyt tilintarkastukseen erikoistunut ammattilainen käyden manuaalisesti läpi yrityksen tilinpäätöstietoja. Lisäksi he mainitsevat, että tilintarkastuksen yksi tärkeimmistä tehtävistä on havaita tilinpäätöksessä poikkeavuuksia, mutta olemassa olevista ohjeista huolimatta petosindikaattoreiden havaitseminen voi olla haastavaa. Useat tutkimukset ovat osoittaneet, että tilintarkastajat pystyvät havaitsemaan vain vähän kaikista tapahtuneista petoksista, minkä vuoksi on yhä enemmän keskitetty siirtymään kohti automaattisia järjestelmiä tilinpäätöspetosten havaitsemiseksi (Craja ym., 2020). Tilinpäätöstietojen manuaalista tarkastamista ja muita perinteisiä petosten havaitsemismenetelmiä on aiemmassa kirjallisuudessa kuvailtu paitsi kalliiksi, mutta myös epätarkoiksi, aikaa vieviksi ja epäkäytännöllisiksi (West & Bhattacharya, 2016).

## 2.2 Tilinpäätöspetokset

Tilinpäätöspetokset ovat yksi merkittävimmistä petostyypeistä koskien kaikkia taloudellisia petoksia. Tilinpäätöspetos voidaan määritellä olennaiseksi laiminlyönniksi tai vääristelyksi, joka johtuu siitä, että taloudellisia tietoja ei ole tarkoituksellisesti raportoitu yleisesti hyväksytyjen kirjanpitoperiaatteiden mukaisesti (Hajek & Henriques, 2017). Tilinpäätöspetoksia voidaan sanoa merkittäväksi petostyyppiä esimerkiksi sen takia, että suhteessa muihin petostyyppiin tilinpäätöspetoksissa menetetty rahan määrä on keskimäärin huomattavan suuri. Kuvio 1 havainnollistaa eri petostyyppien aiheuttamien tappioiden eroja.

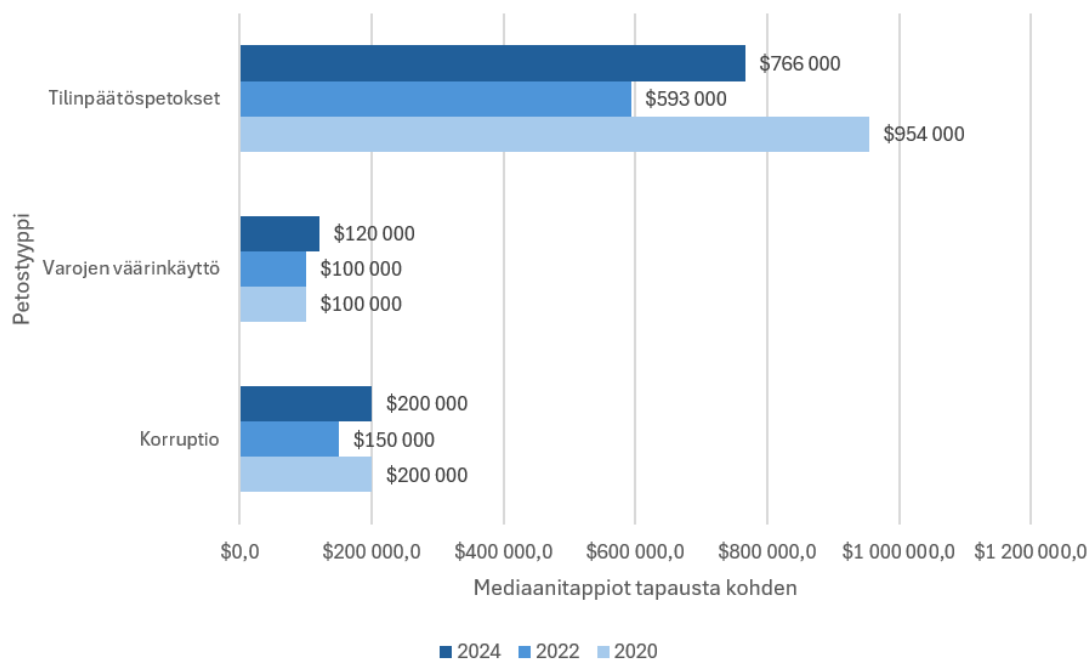
Tilinpäätöspetoksissa tyypillisintä on tilinpäätöksessä ilmoitettujen lukujen väärentäminen esimerkiksi liioittelemalla lukuja tai jättämällä jotakin laskuista pois. Tilinpäätöstietojen väärentämisellä yritykset pyrkivät saamaan taloudellista etua. Petoksen tarkoituksena voi olla esimerkiksi luoda vaikutelma siitä, että yritys olisi kannattavampi kuin se todellisuudessa on, ja tällä tavoin nostaa yrityksen osakekurssia tai helpottaa lainan saamista pankista (Ashtiani & Raahemi, 2022). Tavallisesti tilinpäätöstietojen väärentämisellä todellisuutta parempaan suuntaan tavoitellaan siis sitä, että sijoittajat kiinnostuvat yrityksestä hyvän tilinpäätöksen vuoksi tai että yritys saa edullisemmilla ehdoilla lainaa pankista. Sijoittajien keskuudessa onkin tavallista seurata yrityksiä julkaisemia tilinpäätöksiä ja tulosjulkistuksia, joiden perusteella sijoituspäätöksiä voidaan tehdä. Tilinpäätöstietojen väärentämisellä todellisuutta huonompaan suuntaan yritykset pyrkivät pääsääntöisesti saamaan veroetuja yritykselle tavoitellen matalamman verotuksen avulla suurempia säästöjä yritykselle. Yrityksen taloudellisen edun saavuttamisen lisäksi tilinpäätöspetoksiin syyllistytään useista syistä, kuten henkilökohtaisesta hyödyntämisestä, tarpeesta täyttää lyhyen aikavälin taloudelliset odotukset ja halusta salata huonoja uutisia (Hajek & Henriques, 2017). Esimerkkinä henkilökohtaisen hyödyn tavoittelusta voidaan käyttää sitä, jos

yrittäjien johtajat manipuloivat tilinpäätöksiä henkilökohtaista etua ajatellen, koska johdon palkkiot voivat usein olla sidottuja osakekurssin kehitykseen.

Tilinpäätöksiä koskevilla petoksilla on vaikutuksia moniin sidosryhmiin, jotka ovat yhteydessä petoksen tehneeseen yritykseen. Tilinpäätöspetoksisista aiheutuvat tappiot eivät kohdistu vain kyseiseen yritykseen vaan myös sijoittajiin, jotka menettävät osakkeidensa arvon ja työntekijöihin, jotka menettävät työpaikkansa petoksen seurauksena (du Toit, 2023). Lisäksi on huomioitavaa, että tilinpäätöspetoksilla voi olla laajempia vaikutuksia jopa yhteiskunnallisella tasolla, jos yritys onnistuu vähentämään omaa verotaakkaansa tilinpäätöspetoksen avulla. Edellä mainittujen seikkojen perusteella voidaankin todeta, että tilinpäätöspetoksilla on merkittäviäkin yhteiskunnallisia vaikutuksia, ja tilinpäätöspetosten seuraukset ulottuvat laajalle joukolle sidosryhmiä, joihin petoksen tehnyt yritys on kytköksissä. Nykyisessä kirjallisuudessa viitataan, että tilinpäätöspetokset ovat vakava huolenaihe sijoittajille, tilintarkastusyhtiöille, viranomaisille sekä muille pääomamarkkinoiden sidosryhmille (Hajek & Henriques, 2017).

Hajekin ja Henriquesin (2017) mukaan tilinpäätöspetokseen syyllistyneistä yrityksistä on voitu raportoida seuraavia tietoja: Petoksen tehneen yrityksen keskimääräinen osakekurssi laski 16,7 % sen jälkeen, kun lehdistössä oli ensimmäisen kerran julkistettu väitetty petos. Lisäksi 28 % petokseen syyllistyneistä yrityksistä ajautui konkurssiin tai selvitystilaan kahden vuoden kuluessa, ja 47 % petokseen syyllistyneistä yrityksistä poistettiin kansallisesta pörssistä petoksen seurauksena (Hajek & Henriques, 2017). Edellä mainitun perusteella voidaan todeta olevan hyvin todennäköistä, että tilinpäätöspetokseen syyllistynyt yritys kärsii huomattavasti taloudellisesti jäädessään kiinni petoksesta.

Tutkimuksessaan du Toit (2023) linjaa, että vaikka tilinpäätöspetokset edustavatkin kohtuullisen pientä määrää kaikista taloudellisista petostapauksista, ne aiheuttivat suurimmat kokonaistappiot tapauksista kohden. Otoksessa, jossa tutkittiin vuosien 2016-2017 aikana tehtyä 2690 petostapausta, tilinpäätöspetosten osuus oli 10% (du Toit, 2023). Edellä mainittu 10 % edustaa vain niitä tapauksia, joissa tilinpäätöspetos onnistuttiin havaitsemaan, joten summa, joka todellisuudessa vastaa menetetyt rahan määrää, voi olla vielä huomattavasti suurempi. Association of Certified Fraud Examiners:in vuonna 2020, 2022 ja 2024 julkaisemissa raporteissa taloudelliset petokset ovat jaettu kolmeen ryhmään: korruptioon, varojen väärinkäyttöön ja tilinpäätöspetoksiin. Raporteissa esitetyistä tiedoista erityisen huomioitavana tämän tutkielman kannalta voidaan pitää sitä, että raporttien mukana tilinpäätöspetos on petostyyppi, joka aiheuttaa suurimmat tappiot verrattuna muihin petostyyppeihin (Association of Certified Fraud Examiners, 2020, 2022 & 2024). Kuvista 1 voidaan havaita, että tilinpäätöspetosten aiheuttamat tappiot tapauksista kohden ovat huomattavasti suuremmat kuin kahden muun petostyyppin kohdalla.



KUVIO 1 Tilinpäätöspetosten, varojen väärinkäytön ja korruption mediaanitappiot vuosina 2020, 2022 ja 2024 (Association of Certified Fraud Examiners, 2020, 2022 ja 2024)

## 3 ALGORITMEIHIN POHJAUTUVAT MENETELMÄT

Tässä sisältöluvussa käsitellään erilaisia algoritmeihin pohjautuvia menetelmiä ja niiden ominaisuuksia, joiden avulla tilinpäätöspetoksia voitaisiin pyrkiä havaitsemaan. Kyseiset menetelmät on valittu sen perusteella, mitkä eniten esiintyvät nykyisissä tutkimuksissa tilinpäätöspetosten havaitsemiseksi. Ensimmäisessä alaluvussa määritellään mitä algoritmit ovat ja millaisia ominaisuuksia niillä on. Ensimmäisen alaluvun jälkeen seuraavat alaluvut käsittelevät valittuja menetelmiä ja niiden ominaisuuksia.

### 3.1 Algoritmit yleisesti

Epävirallisesti määriteltynä algoritmilla tarkoitetaan mitä tahansa hyvin määriteltäviä laskennallista toimenpidettä, joka ottaa syötteenä jonkin arvon tai arvojoukon ja tuottaa tulosteeksi jonkin arvon tai arvojoukon rajallisessa ajassa. Algoritmia voidaan myös kutsua eräänlaiseksi laskentavaiheiden sarjaksi, joka muuntaa syötteen tulokseksi, ja algoritmeja voidaan myös tarkastella välineinä, joilla ratkaistaan tarkoin määriteltäviä laskennallisia ongelmia. Laskennallisten ongelmien lausekkeessa määritellään yleisesti haluttu syöte/tuloste -suhde ongelmainstansseille, jotka ovat tyypillisesti hyvin suuria. Algoritmi kuvaa tietyn laskentamenettelyn, jossa saavutetaan tämä syöte/tuloste -suhde kaikkien ongelmainstanssien osalta. Laskennallista ongelmaa varten algoritmi on oikeanlainen, jos jokainen syötteenä annettu ongelmainstanssi pysähtyy, lopettaa laskennan rajallisessa ajassa ja antaa oikean ratkaisun ongelmainstanssiin. (Cormen ym., 2022.) Algoritmien merkittävimmäksi eduksi voidaan katsoa olevan se, että kun algoritmi on oppinut, mitä tietojen ja datan kanssa pitää tehdä, algoritmi voi tehdä työnsä täysin automaattisesti (Mahesh, 2019).

Vaikka algoritmeja voidaan määrittelynsä puolesta pitää laskennallisina ja matemaattisina ohjeina, voidaan algoritmeja käyttää moneen muuhun tarkoitukseen matemaattisten toimenpiteiden suorittamisen lisäksi. Algoritmeja käytetäänkin erilaisissa käyttötarkoituksissa monien eri alojen erilaisissa tehtävissä.

Perinteisesti algoritmeja hyödynnetään esimerkiksi tiettyjen haluttujen kohteiden tunnistamiseen suuresta datamäärästä, tietojen tallentamiseen tietokantoihin, työkalujen kehittämiseen tietojen analysointia varten, suurten tietomäärän hallitsemiseen ja käsittelemiseen, aiempaan dataan perustuvaan ennustamiseen sekä resurssien optimoimisen arvioimiseen (Cormen ym., 2022). Edellä mainitun perusteella voidaan todeta, että algoritmeja käytetään hyvin moneen tarkoituksen ja ne ovat väistämättä käytössä nykyaikana valtaosan ihmisten elämässä, tiedostetusti tai tiedostamatta. Algoritmeja koskevassa kirjallisuudessa kuitenkin korostetaan sitä, että ei ole olemassa yhtä ainoaa algoritmityyppiä, joka soveltuu parhaiten kaikkien ongelmien ratkaisemiseen vaan käytettävä algoritmi riippuu muun muassa siitä, millainen ongelma halutaan ratkaista ja kuinka montaa muuttuja ongelman ratkaisuun käytetään (Mahesh, 2019).

## 3.2 Tekoäly

Käsitteenä tekoälyä (engl. Artificial intelligence) voidaan kuvailla hyvin laajaksi ja moniulotteiseksi johtuen osittain esimerkiksi siitä, että pelkästään termin "älykkyys" liitettävät määritelmät eivät ole täysin yksiselitteisiä. Tekoälyä voidaan kuvailla monella eri tavalla ja hyvinkin erilaisista näkökulmista. Nykyisessä kirjallisuudessa linjataankin, että tekoälylle ei ole yleisesti hyväksyttyä määritelmää, ja näin ollen termiä "tekoäly" on käytetty monissa eri merkityksissä sekä alan sisällä että sen ulkopuolella (Wang, 2019). Tekoälyä on kuitenkin yleisesti esimerkiksi määritelty alaksi, joka käsittelee kaikkia näkökohtia, jotka liittyvät kognitiivisten toimintojen jäljittelyyn reaali maailman ongelmaratkaisussa ja sellaisten järjestelmien rakentamiseen, jotka oppivat ja ajattelevat kuin ihmiset (Holzinger, Langs, Denk, Zatloukal & Müller, 2019). Määriteltäessä tekoälyn käsitettä sen ominaisuuksia ja kykyä toimia vertaillaan usein ihmisälyyn. Kirjallisuudessa tekoälyä on myös kuvailtu tutkimuksena siitä, miten saada tietokoneet tekemään asioita, joissa ihmiset ovat tällä hetkellä parempia (Ertel, 2018, s.119). Edellä mainittua kuvausta tekoälystä voidaan pitää jopa ajattomana: lähtökohtaisesti tekoälyn luonne ei ole vuosikymmenten aikana muuttunut, koska sen perimmäinen tavoite suorituskyvyn parantamisesta on pysynyt samana. Tämän tutkielman kannalta oleellisena määritelmänä voidaan pitää Tchelidzen (2019) määritelmää, jossa tekoälyä kuvataan älykkääksi ohjelmistoksi, jolla on kyky suunnitella ja ratkaista ongelmia, sekä tehokkaimmaksi algoritmiksi, joka muistuttaa ihmisen älyä ja käyttäytymistä. Tässä tutkielmassa älykkyydellä viitataan kykyyn tehdä oikeita päätöksiä tilanteissa, joissa päätöksen voi selkeästi luokitella joko oikeaksi tai vääräksi. Seuraavissa kahdessa alaluvuissa käsitellään koneoppimista ja syväoppimista, jotka ovat tekoälyn alakäsitteitä.

### 3.2.1 Koneoppiminen

Koneoppiminen (engl. Machine learning) on yksi tekoälyn tärkeimmistä osa-alueista, ja sen voidaan katsoa määrittävän koko tekoälyn käsitettä paljon. Koneoppiminen voidaan määritellä tieteelliseksi tutkimukseksi algoritmeista ja tilastollisista malleista, joita tietokonejärjestelmät käyttävät suorittaakseen tietyn tehtävän ilman, että niitä on erikseen siihen ohjelmoitu (Mahesh, 2019). Koneoppimisen tutkimus käsittelee siis sitä, kuinka järjestelmä pystyy kehittymään automaattisesti aiemmin tapahtuneen kokemuksen avulla. Greenerin, Kandathilin, Moffatin ja Jonesin (2022) mukaan koneoppimista voidaan myös kuvailla laajaksi prosessiksi, jossa dataan sovitetaan ennustavia malleja tai tunnustetaan informatiivisia ryhmittymiä datasta. Lisäksi heidän mukaansa koneoppimisen alalla pyritään jäljittelemään ihmisen kykyä tunnistaa tietynlaisia malleja datasta objektiivisesti laskennan avulla. Koneoppimista pidetään erityisen hyödyllisinä tilanteissa, joissa analysoitava tietokokonaisuus on liian suuri tai monimutkainen inhimilliseen analyysiin ja/tai kun halutaan automatisoida datan analysointiprosessi toistettavan ja ajallisesti tehokkaan prosessin luomiseksi (Greener ym., 2022). Lukuisat tekoälyjärjestelmien kehittäjät ovat huomioineet, että monissa sovelluksissa voi olla paljon helpompaa kouluttaa järjestelmää näyttämällä sille esimerkkejä halutuista syöte- ja tulostekäyttäytymisestä kuin ohjelmoida se manuaalisesti ennakoimalla haluttu vastaus kaikille mahdollisille syötteille (Jordan & Mitchell, 2015). Edeltävästä huomiosta voidaan hyvin havaita, kuinka suuri merkitys koneoppimisella on tekoälyn suorituskyvyn parantamiseen, mikä olikin aiemmin mainittu yhdeksi tekoälyn perimmäisiksi tavoitteiksi. Koneoppimisen merkityksestä osana tekoälyä kertoo myös se, että nykyään useimmat tekoälyominaisuuksia tarjoavat älykkäät järjestelmät perustuvat nimenomaan koneoppimiseen (Janiesch Zschech & Heinrich, 2021).

Jordanin ja Mitchellin (2015) mukaan voi siis olla helpompaa kouluttaa järjestelmää näyttämällä sille esimerkkejä halutuista syöte- ja tulostekäyttäytymisestä kuin ohjelmoimalla järjestelmä manuaalisesti tuottamaan haluttuja vastauksia. Kyseinen koneoppimiseen liitettävä ilmiö voidaan havaita hyvin tilinpäätöspestosten havaitsemisen yhteydessä. Aihetta käsittelevissä tutkimuksissa on hyvin tyypillistä, että tilinpäätöksistä muodostettu otanta jaetaan kahteen osaan, joista toisen avulla petoksenhavaitsemismallia koulutetaan näyttämällä sillä tietty määrä petollisia ja ei-petollisia tilinpäätöksiä. Kouluttamisen jälkeen jäljelle jäänyttä osuutta otannasta käytetään mallin testaamiseen eli itse tilinpäätöspestosten havaitsemiseen. Toisin sanoen mallia opetetaan näyttämällä sille esimerkkejä halutuista syötteistä ja tulosteista sen sijaan, että malli manuaalisesti ohjelmoidaan saavuttamaan halutut tulokset.

Koneoppimisen alaisuuteen kuuluu monia erilaisia menetelmiä, joista tämän tutkielman kannalta tärkeitä ovat logistinen regressio, Mahalanobis-etäisyys, ensemble-oppiminen, karkea joukkoteoria sekä tukivektorikone. Seuraavat kappaleet sisältävät kyseisistä menetelmistä lyhyet kuvaukset.

Logistinen regressio on tilastollinen tekniikka, jolla arvioidaan todennäköisyyttä binääriselle tulokselle, määräytyen useiden tekijöiden perusteella (Saleh Hussein, Salah Khairy, Mohamed Najeeb & Alrikabi, 2021). Tosin sanoen

logistinen regressio on luokittelualgoritmi, jota käytetään ennustamaan arvoja tietyssä joukossa muuttujia. Petosten havaitsemisen yhteydessä on kuvailtu, että logistinen regressio on eräänlainen todennäköisyysperusteinen tilastollinen luokittelumalli, ja siinä käytetään niin sanottua logistista käyrää petosten havaitsemiseen (Patil, Nemade & Soni, 2018). Logistisen regression eduiksi voidaan katsoa kuuluvan sen erittäin tehokkaan kouluttamisen mahdollisuus ja sen tehokas kyky luokitella tuntemattomia tietueita (Alenzi & O, 2020).

Lokananin, Tranin ja Vuongin (2019) mukaan Mahalanobis-etäisyys on yksi yksikertaisimmista menetelmistä poikkeavien havaintojen suodattamiseen. Käsitteellisesti Mahalanobis-etäisyys mittaa Lokananin ym. (2019) mukaan datapisteen läheisyyttä jakauman keskipisteeseen, ja sitä voidaan pitää suoranaisena yleistyksenä keskihajonnasta. Petosten tunnistusta käsittelevissä tutkimuksissa Mahalanobis-etäisyyttä käytetään tavallisesti päätettäessä, onko tarkasteltava havainto poikkeava vai ei (Lokanan ym., 2019).

Ensemble-oppiminen on yläkäsite menetelmille, jotka yhdistävät useita tekijöitä päätöksen tekemiseksi, ja joiden lähtökohtana on, että useiden mallien yhdistämisen avulla voidaan todennäköisesti kompensoida yksittäisten mallien virheet (Sagi & Rokach, 2018). Aiemmat tutkimukset ovatkin osoittaneet, että ensemble-menetelmät suoriutuvat yleensä paremmin kuin mitkään yksittäiset ennustemenetelmät (Bao, Ke, Li, Yu & Zhang 2020). Segin ja Rokachin (2018) mukaan tarkemmassa tarkastelussa ensemble-menetelmät perustuvat algoritmiin, joka ottaa syötteenä joukon ennalta merkittyjä esimerkkejä ja tuottaa mallin, esimerkiksi luokittelijan, joka yleistää nämä esimerkit. Käyttämällä tätä tuotettua mallia, ennusteita voidaan tehdä uusille merkitsemättömille esimerkeille (Sagi & Rokach, 2018).

Karkea joukkoteoria (engl. Rough Set Theory) on koneoppimismenetelmä, joka tarjoaa tehokkaan matemaattisen työkalun epävarmuuden ja epävarman tiedon käsittelyyn (Zhang ym., 2021). Karkeaa joukkoteoriaa sovelletaan yleensä tietojen pelkistämiseen, sääntöjen poimimiseen ja tietojen louhintaan (Yeh, Chi, Lin & Chiu, 2016). Data-analytiikan näkökulmasta karkean joukkoteorian tärkeimmiksi eduiksi voidaan mainita sen kyky toimia ilman minkäänlaista ennakkotietoa aineistosta, arvioida datan merkityksellisyyttä ja itse generoida päätössäännöt datan perusteella (Zhang ym., 2021).

Cervantesin, Garcia-Lamontin, Rodríguez-Mazahuan ja Lopezin (2018) mukaan tukivektorikoneella tarkoitetaan luokittelumenetelmää, jota voidaan pitää yhtenä tunnetuimmista tekniikoista odotetun ratkaisun optimoimiseksi. Tukivektoritietokoneen poikkeuksellinen yleistämiskyky sekä sen optimaalinen ratkaisu- ja erottelukyky ovat herättäneet niin datanlouhinnan kuin koneoppimisen yhteisöjen huomion (Cervantes ym., 2020). Tukivektorikoneet voidaan katsoa kuuluvan sekä datanlouhinnan kuin koneoppimisen alaisuuteen. Tukivektoritietokoneen hyvien teoreettisten perusteidensa ja hyvän yleistämiskykynsä ansiosta siitä on tullut yksi käytetyimmistä luokittelumenetelmistä (Cervantes ym., 2020).

### 3.2.2 Syväoppiminen

Janieschin ym. (2021) mukaan viimeisen vuosikymmenten aikana koneoppimisen avulla on tullut esiin monia merkittäviä edistysaskeleita kehittyneissä oppimisalgoritmeissa. Yksi näistä edistysaskeleista oli keinotekoisien neuroverkkojen (engl. Artificial Neural Network, ANN) kehittyminen kohti yhä syvempiä ja paremmat oppimisominaisuudet omaavia neuroverkkorakenteita, joita kutsutaan syväoppimiseksi (Janiesch ym., 2021). Syväoppiminen (engl. Deep learning) on siis koneoppimisen käsite, joka perustuu keinotekoisiiin neuroverkkoihin ja joka jäljittelee aivojen hermoverkkojen rakennetta. Mohmandin, Sabharwalin ja Echizenin (2021) mukaan syväoppimisen avulla saavutetaan suurempi teho ja joustavuus, koska menetelmä kykenee käsittelemään suuren määrän erilaisia piirteitä epäjäseneltyä dataa käsitellessään. Syväoppimisalgoritmi ohjaa tiedot useiden kerrosten läpi, joista jokainen pystyy poimimaan piirteitä asteittain ja siirtämään ne seuraavalle kerrokselle (Mohmand ym., 2021). Syamin ja Sharmanin (2018) mukaan neuroverkkoja pidetään tehokkaimpina tilanteissa, joissa käsitellään epäselvää ja monimutkaista dataa, koska ne pystyvät poimimaan datasta kuvioita ja trendejä, joita ihmiset ja perinteiset tietokoneohjelmat eivät pysty havaitsemaan. Simuloimalla ihmisten aivoja neuroverkot ratkaisevat ongelmia esimerkkien avulla, kuten ihmisetkin, eikä niitä ole ohjelmoitu suorittamaan vain jokin tiettyä tehtävää (Syam & Sharma, 2018). Tämän tutkielman kannalta tärkeimmät syväoppimisen menetelmät ovat hierarkkinen huomioverkko, äärimmäinen gradientin vahvistaminen ja pitkäkestoinen lyhytkestomuisti. Seuraavat kappaleet sisältävät kolmesta edellä mainitusta menetelmästä lyhyet kuvaukset.

Hierarkkista huomioverkkoa (Hierarchical Attention Network) voidaan pitää yhtenä edistyneimmistä syväoppimisen lähestymistavoista käsitellessä kielen hierarkkisia rakenteita, kuten sanojen, lauseiden ja dokumenttien välistä hierarkkia. Hierarkkisen huomioverkon avulla pyritään käsittelemään tekstin luokitteluprosessissa esiintyvää ilmiötä, jossa sanan tai lauseen merkitys voi vaihdella dokumentin ja konteksti mukaan. Hierarkkinen huomioverkko havaitsee ensin ne sanat, jotka ovat tärkeitä lauseessa, ja sen jälkeen ne lauseet, jotka ovat tärkeitä koko dokumentissa ottaen huomioon kyseessä olevan kontekstin. Malli voi ottaa kontekstin huomioon esimerkiksi tunnistamalla, että jonkin sanan esiintyminen voi olla merkittävää, kun se on tietyssä lauseessa, kun taas sanan esiintyminen ei välttämättä ole tärkeä toisessa lauseessa. (Craja ym., 2020.) Tilinpäätöspetosten tunnistamisen yhteydessä hierarkkisen huomioverkon merkitys korostuu, kun tarkastelussa on tilinpäätöksen kirjallinen osuus, esimerkiksi johdon kommentit liiketoiminnan suoriutumisesta.

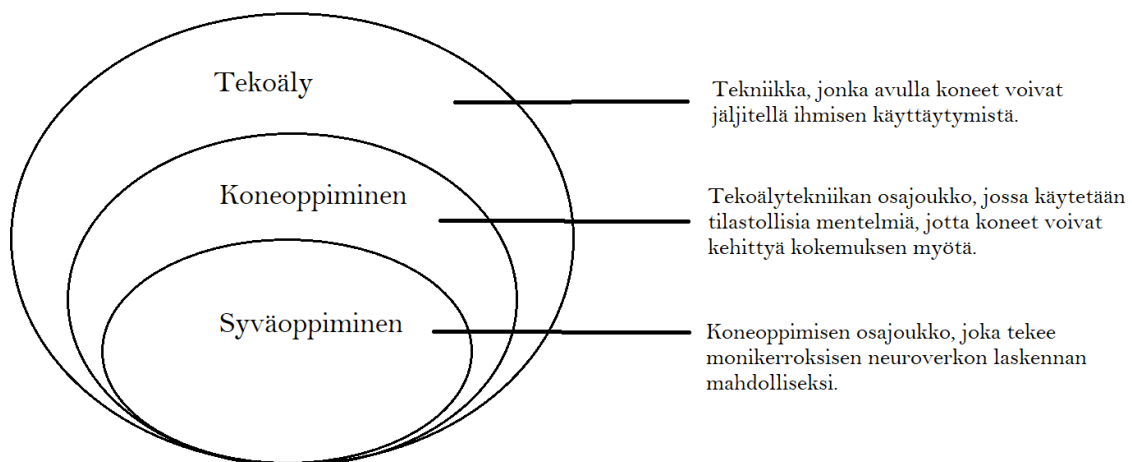
Carmonan, Climentin ja Momparlerin (2019) mukaan äärimmäinen gradientin vahvistaminen (engl. Extreme Gradient Boosting) on tekniikka, joka perustuu gradientin tehostamiseen päätöspuun avulla. Lisäksi heidän mukaansa matemaattisesti gradientti laskee, kuinka nopeasti tietty funktio muuttuu tietyn muuttujan suhteen, ja yksinkertaisiin päätöspuihin verrattuna äärimmäinen gradientin vahvistaminen tuottaa mallin, jossa on satoja tai jopa tuhansia päätöspuita. Äärimmäisen gradientin malleja voidaan pitää monimutkaisina,



mutta niiden ennustusvoima on suurempi kuin useimpien perinteisten menetelmien (Carmona ym., 2019). Kappale 3.3 sisältää lyhyen kuvauksen päätöspuista.

Janin (2021) mukaan pitkäkestoinen lyhytkestomuisti (engl. Long Short-term Memory, LSTM) on niin sanotun takaisinkytketyn neuroverkkojen (engl. Recurrent Neural Network) laajennus, jolla on pyritty ratkaisemaan sen pitkäaikaisen muistin mahdollinen heikko suorituskyky. Lisäksi hänen mukaansa takaisinkytketty neuroverkko on yksi suosituimmista syväoppimismalleista, jota käytetään usein luonteeltaan peräkkäisten tietojen käsittelyyn ja joka tuottaa ennusteita peräkkäisistä tiedoista suuremmalla tarkkuudella. Takaisinkytkettyyn neuroverkkoon verrattuna pitkäkestoinen lyhytkestomuisti on innovatiivisempi, koska se lisää neuroneihin enemmän vaiheita, mikä parantaa pitkäkestoinen muistin suorituskykyä huomattavasti (Jan, 2021). Pitkäkestoinen lyhytkestomuistin arkkitehtuuria kuvaillaan monimutkaisempana kuin takaisinkytketyn neuroverkon, mutta toisin kuin takaisinkytketty neuroverkko, pitkäkestoinen lyhytkestomuisti voi muistaa syötetyt tiedot pitkältä aikaväliltä (Laghrissi, Douzi, S., Douzi, K. & Hssina, 2021).

Kuvio 2 havainnollistaa 3.2 luvussa esiteltyjen tekoälyn, koneoppimisen ja syväoppimisen välisiä suhteita toisiinsa ja linjaa niiden keskeisimmät piirteet lyhyesti.



KUVIO 2 Tekoälyn, koneoppimisen ja syväoppimisen väliset suhteet sekä niiden keskeiset piirteet (Aggarwalin ym., 2022, s. 116 mukaan)

### 3.3 Datanlouhinta

Kerätyn datan määrä erilaisiin tietokantoihin on nykyisin valtava, kun dataa kerätään ja varastoidaan jatkuvasti monesta eri lähteestä. Garcían Luengon ja Herreran (2015, s. 16) mukaan tätä valtavaa datamäärää, jota kutustaan raakadataksi, on pääosin vaikea käsitellä ihmisen tai manuaalisten sovellusten avulla. Siitä lähtien, kun tehokkaat tallennus- ja yhteysvälineet ovat kehittyneet, monet eri järjestelmät, kuten World Wide Web (WWW), yritystoimintaan liittyvät palvelut, tieteen ja tekniikan eri sovellukset sekä verkot ovat tuottaneet ja tuottavat jatkuvasti eksponentiaalisesti lisää dataa (García ym., 2015). Roigerin (2017) mukaan tämä nopeasti ja eksponentiaalisesti kasvanut datan määrä ei ole mahdollistanut hyödyllisen tiedon tai organisoidun tiedon automaattista poimimista ja ymmärtämistä. Lisäksi hänen mukaansa tämä on johtanut datanlouhinnan (engl. Data mining) syntymiseen, joka on nykyään yleisesti tunnettu käytäntö monella eri tieteenalalla. Datanlouhinta voidaan määritellä yhtä tai useampaa algoritmia hyödyntäväksi prosessiksi, jossa datasta etsitään kiinnostavia rakenteita, jotka voivat olla monessa eri muodossa kuten sääntöjoukkoina, graafeina tai yhtenä tai useampana yhtälönä (Roiger, 2017). Karkeasti voidaan myös määritellä, että datanlouhinta on ongelmien ratkaisemista analysoimalla tietokannoissa olevia tietoja ja löytämällä tiedoista jo olemassa olevia, mutta tuntemattomia kuvioita (García ym., 2015, s. 16). Edellä mainitun perustella voidaankin tiivistää, että datanlouhinnan tavoite on poimia datasta ja informaatiosta hyödyllistä tietoa, ja johtaa siis jo olemassa olevasta aineistosta jäsenneltyä ja ymmärrettävää tietoa. Datanlouhinnan määritelmässä voidaankin edellä mainitulla tavalla myös painottaa sitä, kuinka prosessin avulla suuresta määrästä dataa löydetään mielenkiintoisia malleja ja tietoa eli kuinka raakadataa saadaan muuttettua hyödylliseksi tiedoksi (Han, Pei & Tong 2022, s. 2).

Tämän tutkielman kannalta oleellisin datanlouhinnan menetelmä on luokittelu- ja regressiopuu (engl. Classification and Regression Tree, CART). Chenin (2016) mukaan luokittelu- ja regressiopuu on datanlouhinta-algoritmi, joka perustuu binääriseen päätöksentekopuun menetelmään. Tyypillisesti luokittelu- ja regressiopuuta käytetään jatkuvien tai ei-parametristen tietojen luokitteluun (Chen, 2016). Janin (2018) mukaan tarkemmin tarkasteltuna luokittelu- ja regressiopuu eli päätöspuu jakaa datan kahteen osaan ennalta määritettyjen kriteerien perusteella. Kun data on jaettu kahteen osaan, syntyy kaksi osajoukkoa, jotka niin ikään jaetaan kahteen osaan uusilla jakamishdoilla. Tätä prosessia jatketaan, kunnes tiedot eivät enää ole jaettavissa, ja jakamisen seurauksena on muodostunut niin sanottu päätöspuu (Jan, 2018). Kyseistä päätöspuuta voidaan pitää tehokkaana tekniikkana luokitusongelmien käsittelyyn, koska se kykenee rakentamaan monimutkaisia puita monimutkaisten ongelmien ratkaisemiseksi (Pham, Prakash & Tien, 2018).

## 4 TILINPÄÄTÖSPETOSTEN HAVAITSEMINEN

Tässä luvussa käsitellään sitä, kuinka edellisessä luvussa esiteltyjen algoritmi-pohjaisten menetelmien avulla voidaan havaita petoksia tilinpäätöksissä. Ensimmäinen alaluku käsittelee koneoppimistekniikoiden hyödyntämistä tilinpäätöspetosten havaitsemiseen. 4.2 luku käsittelee petosten havaitsemista syväoppimistekniikoiden avulla ja 4.3 luku käsittelee petosten tunnistusta datanlouhintatekniikoiden avulla. Luvun lopussa esitetään taulukon muodossa läpikäytyjen tutkimuksien tulokset (taulukko 1).

### 4.1 Petosten tunnistus koneoppimistekniikoiden avulla

Petosten tunnistus koneoppimistekniikoiden avulla voidaan jakaa olemassa olevien tutkimuksien perustella karkeasti kahteen ryhmään. Ensimmäinen ryhmä pitää sisällään ne tutkimukset, jotka ovat käyttäneet koneoppimistekniikoita tilinpäätöspetosten tunnistukseen analysoimalla yritysten tilinpäätöksistä laskettuja taloudellisia tunnuslukuja. Toiseen ryhmään kuuluvat puolestaan ne tutkimukset, jotka ovat huomioineet petosten tunnistuksessa ennustemuuttujina sekä taloudellisia että muita kuin taloudellisia tunnuslukuja. Kappale 4.1.1 käsittelee petosten tunnistusta koneoppimistekniikoiden avulla taloudellisten tunnuslukujen perustella, ja kappale 4.1.2 käsittelee petosten tunnistusta sellaisten tutkimuksien perusteella, joissa on myös huomioitu ei-taloudelliset tunnusluvut.

#### 4.1.1 Petosten tunnistus taloudellisista tunnusluvuista

Nykyisessä kirjallisuudessa esiintyy verrattain vähän uusia tutkimuksia siitä, miten koneoppimistekniikoiden avulla voidaan havaita tilinpäätöspetoksia. Näistä tutkimuksista osa keskittyy siihen, kuinka petoksia voidaan tunnistaa vain taloudellisten tunnuslukujen avulla. Taloudelliset tunnusluvut ovat lukuja, jotka johdetaan itse tilinpäätöksessä olevista luvuista, ja ne ilmaisevat kahden tai useamman arvon suhteellista suuruutta. Esimerkiksi osakekohtainen tulos

(EPS) kuvaa yrityksen voiton määrää yhtä osaketta kohti. Raakadataksi tilinpäätöksien kohdalla voidaan kutsua tilinpäätöksessä ilmoitettuja yksittäisiä lukuja.

Bao ym. (2020) mukaan vain tunnuslukujen avulla petosten havaitsemista on perusteltu esimerkiksi sillä, että tunnusluvut ovat helposti saatavilla olevia taloudellisia tietoja ja niihin perustuvia petosten ennustumalleja voidaan soveltaa mihin tahansa julkisesti noteerattuun yritykseen pienin kustannuksin. Sen lisäksi tunnuslukujen käyttöä on perusteltu sillä, että ennustemallit perustuvat myös julkisesti saatavilla oleviin taloudellisiin tietoihin, jolloin uusien petosennustusmallien suorituskykyä voidaan suoraan verrata jo olemassa olevien mallien suorituskykyyn (Bao ym., 2020). Tunnuslukuina petosten havaitsemiseen on esimerkiksi käytetty kannattavuuslukuja, kuten myynnin ja investointien tuottoa, maksuvalmiuden tunnuslukuja, vakavaraisuussuhdetta kuvaavia tunnuslukuja sekä yrityksen rakennetta kuvaavia suhdelukuja, esimerkiksi lyhytaikaisen velan suhdetta kokonaisvelkoihin (Kanapickienė & Grundienė, 2015).

Tutkimukset, jotka ovat keskittyneet petosten havaitsemiseen vain taloudellisten tunnuslukujen avulla, on pääasiassa toteutettu niin, että jokaisessa tutkimuksessa erikseen määritellään tunnusluvut, joiden vaihtelua mitataan. Tutkimuksissa esiintyvät tunnusluvut ovat kuitenkin usein suurilta osin samoja, ja joissakin tutkimuksissa on suoraan jonkin hyväksi todetun aiemman tutkimuksen käyttämiä tunnuslukuja. Mittaamalla määriteltyjen tunnuslukujen vaihtelun suuruutta tilinpäätöstiedoista pyritään havaitsemaan poikkeamia, jotka voivat indikoida petoksesta. Tutkimuksissa on esimerkiksi oletettu tunnuslukujen vaihtelevan neljännesvuosittain tiettyjen keskiarvojen ympärillä, jolloin oletuksen avulla pystytään osoittamaan poikkeavat tiedot mittaamalla, ylittääkö kunkin datan etäisyys ”keskipisteestä” tietyn kynnyksarvon (Lokanan ym., 2019). Tutkimuksissa, joissa tilinpäätöspetoksia on havaittu vain taloudellisten tunnuslukujen avulla, on käytetty pääasiallisina koneoppimistekniikoina logistista regressiota, ensemble-menetelmiä ja tutkimuksissa kehiteltyjä omia koneoppimiseen perustuvia malleja.

Kanapickienėnin ja Grundienėnin (2015) tutkimuksessa käytettiin logistisen regression mallia petosten luokittelumallin luomisessa. Tutkimuksen otanta koostui 165 liettualaisen yrityksen tilinpäätöksestä vuosilta 1998–2009. Aluksi malliin määriteltiin tietyt vakiot, jonka jälkeen malleihin vähitellen lisättiin riippumattomia muuttujia eli taloudellisia suhdelukuja, joilla on vahva korrelaatio suhteeseen alkuperäiseen riippuvaan muuttujaan. Tutkimuksen empiirisessä osiossa havaittiin, että mallin avulla voidaan erotella tunnusluvut, joiden arvot voivat viitata petoksiin tilinpäätöksissä, ja 165 tilinpäätöksen suuruudesta otannasta malli pystyi luokittelemaan petosten perustella tapaukset oikein 84,8 % tarkkuudella. Tutkimuksessa linjattiin, että suunniteltua mallia voivat käyttää tilinpäätöstietojen ulkopuoliset käyttäjät tehdessä investointipäätöksiä ja arvioidessaan yrityksiä (Kanapickienė & Grundienė, 2015).

Lokananin ym. (2019) tilastollisten tunnuslukuihin perustuvan tutkimuksen otanta perustui 937 Vietnamin pörssissä noteerattujen yritysten neljännes-

vuosittaisiin ja vuosittaisiin tilinpäätöstietoihin vuosilta 2011–2016. Tutkimuksen data koostui yhteensä 22 488 havainnosta mainittujen vuosien ajalta. Tutkimuksessa hyödynnettiin Mahalanobis-etäisyyttä päättämään, ovatko tutkittavat havainnot poikkeavia. Tutkimuksessa laskettiin tietty kynnyksisarvo, johon Mahalanobis-etäisyyden avulla laskettua arvoa verrattiin, ja jos jokin etäisyyksistä oli suurempi kuin määritelty kynnyksisarvo, voitiin havaintoa pitää poikkeavana. Tutkimuksessa havaittiin, että A-luokan yrityksiksi määritellyillä yrityksistä huomattavan suuri osuus (68,89 %) oli sellaisia, joissa oli hyvin vähän poikkeamia, mitä pidettiin terveenä merkkinä siitä, että ne noudattavat oikeudellisia ja eettisiä normeja. Käytetyllä poikkeamienhavaitsemismallilla voitiin kuitenkin myös havaita, että 23,51 % luokittelemattomista yrityksistä oli merkittäviä poikkeamia tilinpäätöksissä, mikä osoittaa, että tilinpäätöksissä on todennäköisesti petoksia tai virheitä (Lokanan ym., 2019). Tutkimuksessa esitetyt tulokset viittaavat siihen, että käytetyn mallin avulla poikkeamat voidaan havaita hyvin, mutta tutkimuksessa käytetyn mallin avulla ei erikseen voitu jaotella poikkeavuuksia petoksiin ja virheisiin tilinpäätöksissä. Poikkeavuudella tarkoitettiin mitä tahansa poikkeuta tilinpäätöksessä eikä tutkimuksessa erikseen määritelty viittaako poikkeus automaattisesti petokseen.

Baon ym. (2020) tutkimuksen otanta koostui kaikista Yhdysvaltojen julkisesti noteeratuista yrityksistä vuosilta 1991–2008. Tutkimus eroaa useimmista tutkimuksista siinä, että tutkimuksessa käytetään raakoja kirjanpitolukuja toisin kuin aikaisemmat tutkimukset ovat käyttäneet taloudellisia suhdelukuja. Tutkimuksessa käytettiin yhtä tehokkaimmista koneoppimismenetelmistä, ensemble-oppimista uuden petosten ennustusmallin luomiseen. Tutkimuksessa verrattiin käytettyä ensemble-oppimismallia kahteen, aiemmissä tutkimuksissa paljon käytettyihin ennustusalgoritmeihin logistiseen regressioon ja tukivektorikoneeseen. Tutkimuksen empiirinen osio osoitti, että käytetty ensemble-oppimismalli on suoraan perempi kuin kaksi vertailumallia ja, että raakadatasta on mahdollista poimia enemmän hyödyllistä ennakoivaa tietoa, jos rakennetaan koneoppimismalli, joka on suunniteltu hyödyntämään juuri sellaista tietoa (Baon ym., 2020).

Kaikkien edellä mainittujen tutkimuksien voidaan katsoa sisältävän tiettyjä puutteita. Ensinnäkin, kuten todettua, jokaisen artikkelin petostentunnistumenetelmä on ottanut huomioon vain tilinpäätöstiedoista johdettavat tunnusluvut. Huomioon ei siis ole otettu tilinpäätöksen muita osia kuten johdon kommentteja liiketoiminnan suoriutumuksesta. Kyseinen rajausta vain tiettyyn osaan tilinpäätöstietoja voi olla merkittävä seikka, mikäli taloudellisten tunnuslukujen lisäksi petoksia voidaan havaita analysoimalla myös tilinpäätöstietojen muita osia. Näin ollen vaikka taloudellisten tunnuslukujen perusteella petosten havaitseminen voisikin olla kohtuulliseen tarkkaa, voi tilinpäätöspetoksia jäädä silti havaitsematta, kun petosten havaitseminen on kohdistettu vain taloudellisiin tunnuslukuihin. Sen lisäksi Kanapickienénin ja Grundienénin (2015) tutkimuksen otannan voidaan katsoa olevan hyvin suppea verrattuna Lokananin ym. (2019) ja Baon ym. (2020) artikkeleihin. Kyseisessä artikkelissa tutkimustuloksetkaan eivät tarjoa aiheelle hedelmällistä suuntaa jatkotutkimukselle eikä

tutkimuksessa saatuja tuloksia voida pitää tarkkuuden osalta erityisen merkittävänä. Lokananin ym. (2019) ja Baon ym. (2020) tutkimukset ovat otannan puolesta sen sijaan huomattavasti laajempia, mutta artikkelien yrityksistä ja niiden tilinpäätöksistä koostuvat otannat ovat maantieteellisesti rajattuja, kuten myös Kanapickienenin ja Grundienenin (2015) artikkelin otanta. Vaikka Lokananin ym. (2019) ja Baon ym. (2020) tutkimuksien tulokset osoittavat kohtuullista tarkkuutta poikkeamien havaitsemiseen, artikkelien tutkimukset eivät välttämättä anna täysin yleistettävää kuvaa koko tilinpäätösprosessin havaitsemisesta.

#### **4.1.2 Petosten tunnistus taloudellisista ja muista kuin taloudellisista tunnusluvuista**

On tehty muutamia tutkimuksia, joissa taloudellisten suhdelukujen lisäksi tilinpäätösprosessin ennustemuuttujiksi valitaan muitakin kuin taloudellisia suhdelukuja. Kuten aiemmin mainittiin, tilinpäätökset koostuvat numeeristen lukujen lisäksi esimerkiksi johdon kommentteista liiketoiminnan suoriutumisesta, joten kaikkea tilinpäätöksessä olevaa dataa ei voida analysoida lukuina ja näin ollen samalla tavalla. Kirjallisuudessa muiksi kuin taloudellisiksi suhdeluvuiksi on muun muassa määritelty hallituksen jäsenten lukumäärä, perheenjäsenten ja ulkopuolisten osuus hallituksen jäsenistä, perheenjäsenten omistamien osakkeiden osuus sekä omistuksien keskittyminen (Yeh ym., 2016). Muita kuin taloudellisia suhdelukuja ovat myös riippumattomien johtajien osuus, suurimpien osakkeenomistajien osakeomistussuhde, ulkopuolisten valvojen määrä sekä se, onko hallituksen puheenjohtaja ja toimitusjohtaja sama henkilö ja onko yhtiön tilintarkastuksen tehnyt jokin neljästä suuresta tilintarkastusyrietyksestä (Ashtiani & Raahemi, 2022). Aihetta on tutkittu myös suoraan kielellisten muuttujien avulla esimerkiksi tutkimalla, kuinka paljon tilinpäätöksen kirjalliset osuudet sisältävät negatiivisia ja epävarmoja sanoja, joista on voitu päätellä petosten tapahtumista (Hajek & Henriques, 2017).

Kirjallisuudessa on huomioitu, että vaikka on tärkeää tarkastella muitakin kuin taloudellisia tietoja, ei-taloudelliset tiedot ovat tutkimuksissa olleet paljon vähemmän tutkimuksen kohteena kuin taloudelliset muuttujat (Ashtiani & Raahemi, 2022). Esimerkiksi taloudellisten ja kielellisten tietojen yhdistämiseen tilinpäätösprosessin älykkäässä ennustamisessa on kiinnitetty hyvin vähän huomiota (Hajek & Henriques, 2017). Tutkimuksissa nimenomaan uskotaan, että muut kuin taloudelliset tunnusluvut ovat keskeisiä tekijöitä yritysten tilinpäätöksissä ja näin ollen tarjoavat korvaamatonta tietoa tilinpäätösprosessin havaitsemista varten (Yeh ym., 2016). Muiden kuin taloudellisten tunnuslukujen käyttämistä perustellaankin ennustustarkkuuden parantamisella eli ei-taloudellisten tunnuslukujen avulla voidaan mahdollistaa erittäin tarkka malli petollisten tilinpäätösraportoinnin havaitsemiseksi (Jan, 2018).

Yehin ym. (2016) tutkimuksessa petosten havaitsemiseen käytettiin 38 muuttujan joukkoa, joka koostui 23 taloudellisesta tunnusluvusta ja 15 muusta kuin taloudellisesta tunnusluvusta. Tutkimuksen otos koostui sadasta taiwanilaisesta pörssinoteeratusta yrityksestä, joiden tilinpäätöksiä tutkittiin vuosilta 1996–2007. Tutkimuksen ensimmäisessä vaiheessa hyödynnettiin karkeaa jouk-

koteoriaa tärkeiden muuttujien valitsemiseen petosten havaitsemista varten. Jo tässä vaiheessa voitiin päätellä, että muut kuin taloudelliset tunnusluvut tarjoavat arvokasta tietoa petosten havaitsemisessa. Tutkimuksen toisessa vaiheessa käytettiin tukivektorikonetta luokittelumallin muodostamiseen. Tutkimuksessa havaittiin, että edellä mainittu kaksivaiheinen menettely tarjosi merkittävästi tarkemman luokittelutodennäköisyyden, kun luokittelun otoksessa oli mukaan luettu myös ei-taloudelliset tunnusluvut. Vain taloudellisten tunnuslukujen avulla tehty luokittelu oli keskimäärin oikeassa 81 % tapauksissa ja taloudellisten sekä muiden kuin taloudellisten tunnuslukujen avulla tehty luokittelu oli keskimäärin oikeassa 88 % tapauksissa. Kyseinen malli teki myös vähiten luokitteluvirheitä muihin malleihin verrattuna: malli luokitteli vähiten ei-petollisia yrityksiä virheellisesti petollisiksi ja vähiten petollisia yrityksiä virheellisesti ei-petollisiksi. Tutkimuksen tuloksien perusteella voitiin päätellä, että karkeaa joukkoteoriaa ja tukivektorikonetta hyödyntävällä kaksivaiheisella petostenluokitteluamallilla saavutettiin paremmat luokittelutulokset kuin muilla esitellyillä malleilla ja että virheiden esiintyvyys kyseisellä menetelmällä oli alhaisin. Lisäksi tuloksista pystyttiin myös päättämään, että muiden kuin taloudellisten tunnuslukujen käyttö petostentunnistuksessa lisää tunnistuksen tarkkuutta jopa huomattavasti (Yeh ym., 2016).

Hajekin ja Henriquesin (2017) tutkimuksessa tutkittiin, voitaisiinko kehittää parannettu tilinpäätöspestosten havaitsemisjärjestelmä yhdistämällä yritysten taloudellisista tiedoista ja johdon kommentteista johdettuja erityispiirteitä. Tutkimuksen otos koostui 311 yhdysvaltalaisesta yrityksestä, joiden väitettiin syyllistyneen petolliseen tilinpäätösraportointiin vuosina 2005–2015. Koko aineisto koostui 622 yrityksestä eri toimialoilta, joista petollisia oli 311 ja ei-petollisia 311. Yrityksistä 47 % oli listattu New Yorkin pörssissä ja 53 % Nasdaqissa. Aluksi tutkimuksessa luotiin satunnaisesti 30 ositettua otosta aineistosta, jotka koostuivat taloudellisista ja kielellisistä muuttujista. Tutkimuksessa otoksien luokitteluun käytettiin 14 koneoppimistekniikkaa, joiden suorituskykyä arvioitiin erilaisilla luokittelumittareilla ja vertailtiin perinteisillä tilastollisilla testeillä. Tutkimuksen tuloksista selvisi, että koneoppimistekniikoista ensemble-menetelmät suoriutuivat muita menetelmiä paremmin petollisten yritysten tunnistamisessa petollisiksi. Huomioitavaa oli kuitenkin, että Bayesinverkko -niminen koneoppimistekniikka (engl. Bayesian belief network) saavutti korkeimman tarkkuuden ei-petollisten yritysten tunnistamisessa. Havainto viittasi siihen, että ei-petollisten yritysten oikea ennustaminen on vähemmän monimutkaista, kun taas petollisten yritysten havaitseminen vaatii monimutkaisempia ja vähemmän tulkinnanvaraisia koneoppimismenetelmiä (Hajek & Henriques, 2017).

Kahden edellä mainitun tutkimuksen perusteella voidaan päätellä, että myös muut kuin taloudelliset tunnusluvut ovat tärkeää ottaa huomioon tilinpäätöspestosten havaitsemisessa, ja ne voivat antaa merkityksellistä tietoa tilinpäätöksen petollisuudesta. Yehin ym. (2016) tutkimuksessa verrattiin tuloksia vain taloudellisten tunnuslukujen perustella tehdystä tutkimuksesta sekä toisaalta tutkimuksesta, jossa kaikki tunnusluvut otettiin huomioon. Jälkimmäisen

kohdalla petosten oikeinluokittelutarkkuus oli huomattavasti suurempi. Haje-kin ja Henriquesin (2017) tutkimuksen perustella voitiin määritellä, että ei-taloudelliset tunnusluvut ovat myös tärkeitä muuttujia petosten havaitsemisessa, koska ne tarkentavat petosten oikeinluokittelua. Tutkimuksissa saatiin siis vahvaa näyttöä siitä, että muutkin kuin taloudelliset tunnusluvut on syytä ottaa huomioon tilinpäätöspetosten havaitsemisessa.

## 4.2 Petosten tunnistus syväoppimistekniikoiden avulla

Uusia tutkimuksia, joissa käsitellään tilinpäätöspetosten havaitsemista syväoppimiseen perustuvilla tekniikoilla, on useita, mutta tässä katsauksessa huomioon otetaan neljä tutkimusta. Tutkimuksista karsittiin pois sellaiset, joissa on selviä puutteita esimerkiksi otannan koossa tai tuloksien esittelyssä. Luvussa esitellään ne tutkimukset, joiden tutkimusmenetelmät ovat olleet hyvin perusteltuja, otannat ovat pääasiassa suuria ja joiden tulokset ovat olleet vertailukelpoisia.

Omarin, Joharin ja Smithin (2017) tutkimuksessa kehitettiin keinotekoisien neuroverkkojen pohjalta matemaattinen malli. Mallin avulla verrattiin Malesian pienen markkina-arvon yhtiöiden joukosta yrityksiä, jotka olivat syyllistyneet tilinpäätöspetokseen, ei-petollisiin yrityksiin. Tutkimuksessa käytetty otanta koostui yhteensä 110 yrityksestä ja niiden vuosien 2009–2011 vuosittaisista tilinpäätöksistä, joten tarkasteltujen tilinpäätösten kokonaismäärä oli 550. Tutkimuksessa käytettiin kymmentä taloudellista tunnuslukua petosindikaattoreina, ja muita kuin taloudellisia muuttujia ei otettu huomioon. Tutkimuksessa tehdyt tulokset osoittivat, että petollisen tilinpäätösraportoinnin ennustemalli pystyi ennustamaan petollisen tilinpäätösraportoinnin oikein 94,87 % tarkkuudella (Omar ym., 2017).

Crajan ym. (2020) tutkimuksessa tarkasteltiin muutamien eri syväoppimistekniikoiden suorituskykyä tilinpäätöspetosten havaitsemiseksi yhdistämällä taloudellisten tunnuslukujen tiedot ja johdon kommentit yritysten vuorikertomuksissa. Tutkimuksen pääpaino oli selvittää, onko talous- ja tekstidatan uusi yhdistelmä informatiivisempi tietotyyppi petosten havaitsemiseksi kuin talous- ja tekstidatan käyttö erikseen. Kyseisen tutkimuksen otanta koostui 7757 yrityshavainnosta Yhdysvalloista vuosilta 1993–2019, joista 208 oli petollisia ja 7549 ei-petollisia. Tutkimuksessa käytettiin johdon keskustelua ja analyysiä liiketoiminnan tilasta ensisijaisen raakatekstietojen lähteenä, ja yhteensä tutkimuksessa hyödynnettiin yhdeksää kielellistä piirrettä ja 47 taloudellista muuttujaa ennustemuuttujina. Tutkimuksen empiirisissä tuloksissa selvisi, että syväoppimistekniikoista hierarkkisen huomioverkon avulla määritellyt johdon kommenttien tekstipiirteet tuottavat lupaavia luokittelutuloksia ja ne vahvistavat merkittävästi taloudellisia suhdelukuja. Tutkimuksessa esitellyistä syväoppimistekniikoista korkeimman oikeinluokitteluprosentin saavutti äärimmäisen gradientin tehostamisen menetelmä 90,83 % oikeinluokitteluasteella (Craja ym., 2020).



Xiuguon ja Shengyongin (2022) tutkimuksessa pyrittiin kehittämään paranneltu järjestelmä tilinpäätöspetosten havaitsemiseksi käyttämällä viimeisintä tekniikkaa hyödyntäviä syväoppimismalleja. Kyseiset mallit perustuvat numeerisiin piirteisiin, jotka puolestaan johdettiin tilinpäätöksen numeeristen ja tekstuaalisten tietojen yhdistelmästä. Tutkimuksen aineiston muodosti 5130 tilinpäätöstä, joista 244 tapausta olivat petollisia ja 4886 tapausta olivat ei-petollisia. Tutkimuksessa käytetyt tiedot olivat peräisin Kiinan osakemarkkinoiden ja kirjanpitoluokittelutietokannasta vuosilta 2016–2020. Petosten havaitsemiseksi oli käytetty yhteensä 74 muuttujaa, joista 58 oli taloudellisia muuttujia ja 16 muita kuin taloudellisia muuttujia. Tutkimuksen empiiriset tulokset osoittivat, että syväoppimismenetelmien suorituskyky oli merkittävästi parempi perinteisiin koneoppimismenetelmiin verrattuna. Syväoppimismenetelmistä parhaimman oikeinluokitteluprosentin saavutti pitkäkestoinen lyhytkestomuisti, jonka oikeinluokitteluprosentti oli 94,98 % (Xiuguo & Shengyong, 2022).

Edellä mainitun tutkimuksen kanssa on päästy hyvin vastaavanlaisiin tuloksiin toisessakin tutkimuksessa. Janin (2021) tutkimuksessa tutkittiin Taiwanin pörssissä listattujen yhtiöiden tilinpäätöksiä eri syväoppimismenetelmillä. Edellä mainittuun Xiuguonin ja Shengyongin (2022) tutkimukseen verrattuna Janin (2021) tutkimuksessa käytetyn otannan koko ja petosten tunnistukseen käytettyjen indikaattorien määrä olivat pienemmät, mutta muuttujat niin ikään sisälsivät niin taloudellisia kuin ei-taloudellisia muuttujia. Tutkimuksen empiiriset tulokset viittaavat siihen, että pitkäkestoinen lyhytkestomuistin -malli suoriutuu hyvin kaikissa suorituskykyarvoissa, ja se saavutti petollisten ja ei-petollisten tilinpäätöspetosten oikeinluokittelutarkkuudeksi 94,88 % (Jan, 2021).

Edellä mainittujen neljän tutkimuksen perusteella voidaan havaita, että syväoppimismenetelmillä saavutetut tulokset ovat parempia kuin koneoppimistekniikoilla saavutetut, varsinkin mallien oikeinluokitteluprosenttien kannalta. Koneoppimistekniikoiden oikeinluokitteluprosentit jäivät tutkimuksissa alle 90 %, kun taas syväoppimistekniikoiden avulla tutkimuksissa saavutettiin lähes 95 % oikeinluokittelu. Havaintoa voidaan tietyllä tapaa pitää ennustettavana, koska syväoppimista ja sen eri malleja pidetäänkin tekoälyn ja koneoppimisen yhtenä tärkeimmistä ja tehokkaimmista osa-alueista.

Edellä mainituissa tutkimuksissa esiintyi muutamia huomioimisen arvoisia hyviä piirteitä. Erityisen hyvää Carajan ym. (2020) ja Xiuguon ja Shengyongin (2022) tutkimuksissa olivat suuret otannat, jotka molemmat olivat lukumääräisesti useiden tuhansien suuruiset. Otannan suuren koon avulla tuloksien yleistämistä voidaan pitää enemmän luotettavana. Hyvää oli myös se, että tutkimuksissa oli otettu huomioon muutkin kuin taloudelliset tunnusluvut Omarin ym. (2017) tutkimusta lukuun ottamatta.

### **4.3 Petosten tunnistus datanlouhintatekniikoiden avulla**

Datanlouhintatekniikoiden käyttöä tilinpäätöspetosten havaitsemisessa on viimeisen vuosikymmenen aikana tutkittu verrattain vähän. Alle kymmenen vuot-

ta vanhoja tutkimuksia aiheesta on tehty reilu kymmenen kappaletta, mutta tässä tutkielmassa niistä käydään läpi kolme. Karsittuja tutkimuksia ei ole otettu huomioon, koska niissä esitetyt tulokset eivät ole saatavilla tai vertailukelpoisia muiden tutkimusten tuloksien kanssa tai tutkimuksien tekotavassa on merkittäviä puutteita, esimerkiksi käytetty otoskoko on erittäin pieni.

Chenin (2016) tutkimuksessa tarkasteltiin vuosien 2002–2013 aikana Taiwanin pörssissä listattuja yrityksiä kaksivaiheisen menettelyn avulla. Tutkimuksen otanta koostui 176 yrityksestä, joista 44 oli syyllistynyt tilinpäätöspetokseen. Ensimmäisessä vaiheessa muuttujien valintaan käytettiin kahta päätöspuualgoritmia, joiden avulla 30 muuttujan joukosta valikoitiin molemmille algoritmeille optimaalisimmat muuttujat petosten tunnistukseen. Toisessa vaiheessa hyödynnettiin viittä datanlouhintatekniikkaa petollisten tilinpäätösten havaitsemismallien muodostamiseksi. Yhteensä havaitsemismalleja muodostettiin siis 10, kun kahta muuttujien valintaan käytettyä algoritmia sovellettiin viiteen eri datanlouhintatekniikkaan. Malleista parhaimmaksi osoittautui CHAID&CART (Chi-squared Automatic Interaction Detection & Classification and Regression Trees) -yhdistelmä, jonka oikeinluokittelutarkkuus tilinpäätöspetosten tunnistamiseksi oli 87,97 % (Chen, 2016).

Myös Jan (2018) tutki Taiwanissa Taipeiin pörssissä listattuja yrityksiä, joita tutkimuksessa oli yhteensä 160, ja niistä 40 oli syyllistynyt tilinpäätöspetokseen. Tutkimus oli rajattu koskemaan vuosia 2004–2014 koskevia taloudellisia tietoja. Muuttujien valintaan hyödynnettiin keinotekoisia neuroverkkoja ja tukivektorikoneita, joiden avulla alkuperäiset 22 muuttujaa supistettiin merkittäviin muuttujiin. Tämän jälkeen tutkimuksessa hyödynnettiin neljää luokittelumallia, joista parhaan luokittelutarkkuuden saavutti ANN&CART-luokittelumalli (Artificial Neural Network & Classification and Regression Trees), jonka oikeinluokittelutarkkuus oli 90,83 % (C. Jan, 2018).

Linin, Chiun, Huangin ja Yenin (2015) tutkimuksen aineisto koostui niin ikään taiwanilaisista yrityksistä. Tutkimusaineisto oli vuosilta 1998–2010, ja siihen sisältyi yhteensä 576 yritystä, joista 129 oli syyllistynyt tilinpäätöspetokseen. Tutkimuksessa hyödynnettiin 32 muuttujaa tilinpäätöspetosten indikaattoreina, ja tutkimuksessa käytetyt datanlouhintamenetelmät olivat logistinen regressio, päätöspuut ja keinotekoiset neuroverkot. Tutkimuksen empiiriset tulokset osoittivat, että suurimman oikeinluokittelusta saavutti keinotekoiset neuroverkot, 92,8 % ja päätöspuut, 90,3 % (Lin ym., 2015).

Kolme edellä mainittua tutkimusta edustaa niiden tekotapojen ja tulosten perusteella merkittävimpiä tutkimuksia datanlouhintatekniikoiden osalta. Chenin (2016) ja Janin (2018) tutkimukset ovat hyödyntäneet kaksivaiheista menettelytapaa, joissa erillisen menetelmän avulla määritellään aluksi muuttujat ja sitten toisella menetelmällä suoritetaan itse luokittelu. Kyseinen kaksivaiheinen menettelytapa on huomioon otettava menettely, sillä muuttujien valitsemisessa voidaan katsoa olevan suuri merkitys lopullisten tulosten kannalta. Huomioitavaa kuitenkin on myös se, että edellä mainituista kolmesta tutkimuksesta suurin oikeinluokittelusta saavutettiin kuitenkin Linin ym. (2015) tutkimuksessa, jossa ei hyödynnetty keskivaiheista menettelytapaa.

Neljännessä kuvussa esiteltyjen tutkimuksien tulokset ovat esitetty taulukon muodossa (taulukko 1). Taulukosta käy muun muassa ilmi tutkimuksissa käytetyt tekniikat, otantojen koot ja oikeinluokitteluasteet.

TAULUKKO 1 Tutkielmassa esiteltyjen tutkimuksien tulokset taulukkona.

Lähde	Käytetty tekniikka	Otannon ajanjakso	Otannon alkuperämaa	Otannon koko	Käytettyjen muuttujien määrä	Oikeinluokitteluaste
Kanapickiené & Grundiené, 2015	Logistinen regressio	1998-2009	Liettua	165	51	84,80 %
Lokanan ym., 2019	Mahalabonis-etäisyys	2011-2016	Vietnam	22488	24	Ei saatavilla
Bao ym., 2020	Ensemble-oppiminen	1991-2008	Yhdysvallat	Kaikki listatut yritykset	14	Ei saatavilla
Yeh ym., 2016	Karkea joukkoteoria&Tukivektorikone	1996-2007	Taiwan	100	38	88 %
Hajek & Henriques, 2017	Ensemble-menetelmät	2005-2015	Yhdysvallat	622	30	Ei saatavilla
Omar ym., 2017	Keinotekoiset neuroverkot	2009-2011	Malesia	550	10	94,87 %
Craja ym., 2020	Äärimmäinen gradientin tehostaminen	1993-2019	Yhdysvallat	7757	56	90,83 %
Xugo & Shengyong 2022	Pitkäkestoinen lyhytkestomuisti	2016-2020	Kiina	5130	74	94,98 %
Jan 2021	Pitkäkestoinen lyhytkestomuisti	2001-2019	Taiwan	153	18	94,88 %
Chen 2016	CAHID&CART	2002-2013	Taiwan	176	30	87,97 %
Jan 2018	ANN&CART	2004-2014	Taiwan	160	22	90,83 %
Lin ym., 2015	Keinotekoiset neuroverkot	1998-2010	Taiwan	576	32	92,80 %

## 5 POHDINTA

Edellisessä luvussa on käyty läpi 12 tutkimusta ja niiden tuloksia sekä tärkeimpiä havaintoja petosten tunnistuksesta. Tämä luku sisältää pohdintaa edellisen luvun tutkimuksista mukaan lukien lähdekritiikkiä ja ehdotuksia tulevaisuuden tutkimusta varten aiheeseen liittyen.

Ensimmäinen huomioitava seikka, joka voidaan nostaa esille 12 esitellystä tutkimuksesta, on tutkimuksien maantieteellinen raja-  
aus. Jokaisen artikkelin otanta muodostuu yhden maan yrityksistä tai maan sisällä tietyssä pörssissä noteeratuista yrityksistä. Merkille pantavaa on esimerkiksi se, että 12 tutkimuksesta viisi sijoittuu Taiwaniin ja kahdeksan tutkimuksen otanta koostuu aasialaisen maan yrityksistä. Kolmen tutkimuksen otanta koostuu yhdysvaltalaisista yrityksistä, ja vain yhden tutkimuksen otanta koostuu eurooppalaisen maan, Liettuan yrityksistä. On selvästi havaittavissa, että tilinpäätös-  
petosten havaitsemista koskevan kirjallisuuden ja tutkimuksien pääpaino on Aasiassa. Yleisesti myös tutkimuksien otannat muodostetaan vain yhden maan yrityksistä, ei useamman maan yrityksistä tai muulla tavalla maantieteellisesti rajatun alueen yrityksistä, kuten vaikka eurooppalaisista yrityksistä.

Maantieteellisen rajauksen takia kuva muodostuu hyvin paikallisesti käsitellystä ilmiöstä. Esitetyt tulokset voivat korreloida vahvasti koko ilmiön kanssa, mutta täysin yleistettävää kuvaa tutkimuksista ei voida saada. On huomioitavaa, että eri maissa tilinpäätöksiä koskeva lainsäädäntö on erilainen. Esimerkiksi Vietnamin tehtyjä neljännesvuosiraportteja ei ole lain mukaan pakko tarkastaa, jolloin osa Lokananin ym. (2019) tutkimuksen tilinpäätöksistä on tarkastamattomia (Lokanan ym., 2019). Yhden maan yrityksistä muodostetuista otannoista voidaan luultavasti luotettavimpana pitää yhdysvaltalaisista yrityksistä muodostettua otantaa. Yhdysvaltalaiset yritykset muodostavat erittäin merkittävän osuuden koko maailman yritysten markkina-arvosta, joten on todennäköistä, että yleistettävimmän kuvan aiheesta voi saada, kun otanta muodostuu yhdysvaltalaisista yrityksistä.

Tutkimuksia vertaillessa maantieteellisen rajauksen lisäksi esiin nousee myös otantakokojen suuruuden vaihtelu sekä käytettyjen muuttujien määrät. Otannan suuruuden ja käytettyjen muuttujien määrän vaihtelu on esiteltyjen

tutkimusten välillä huomattavan suurta. Esimerkiksi tarkastelujen tilinpäätösten määrä vaihtelee tutkimuksien kesken reilusta sadasta jopa kymmeneen tuhansiin. Muuttujia petosten tunnistamiseksi on puolestaan vähimmillään ollut käytössä vain 10, ja enimmillään tutkimuksissa on kuitenkin käytetty 74 muuttujaa. Molemmissa kohdissa on havaittavissa kohtuullisen suuri ero tutkimusten kesken. Tutkimusten tuloksia, joissa otantakoko eli tutkittujen tilinpäätösten määrä on suuri, voidaan pitää erityisen merkityksellisenä, koska suuren otantakoon voidaan katsoa johtavan luotettavampiin tuloksiin. Esimerkiksi pienen otantakoon voidaan helpommin katsoa altistuvan suurelle satunnaisvaihtelulle eli toistettaessa tutkimusta samalla ja pienemmällä otantakolla, tulokset saattavat vaihdella suhteellisen paljon. Muuttujien valinnassa puolestaan on syytä korostaa sitä, että myös muut kuin taloudelliset tunnusluvut tulee ottaa huomioon tutkimuksissa, koska parhaimmat oikeinluokitteluasteet saavutettiin tutkimuksissa, joissa taloudellisten tunnuslukujen lisäksi oli analysoitu kirjallista dataa.

Esitellyistä tutkimuksista voidaan korostaa sitä, että vaikka parhaimmillaan käytetyt menetelmät saavuttivat hyvän, lähes 95 % luokitteluasteen, tutkimuksissa ei ole mainittu, että mallit pystyisivät korvaamaan ihmisen tekemän työn täysin. Tutkimukset eivät siis lähtökohtaisesti ota kantaa siihen, voiko tämänhetkisillä algoritmipohjaisilla menetelmillä korvata tilintarkastajien työn kokonaan. Näin ollen on pääteltävissä, että mikään malli ei ole saavuttanut niin absoluuttista oikeinluokittelua, että ihmisen tekemä työ olisi korvattavissa. Tutkimuksien tuloksien perustella voitaisiin mahdollisesti linjata, että algoritmipohjaisia menetelmiä voidaan käyttää hyödyllisenä työkaluna tilintarkastuksen yhteydessä, koska ne tarjoavat mahdollisuuden käydä läpi dataa nopeasti ja tehokkaasti, mutta ammattilaisen tekemä analyysi on vielä tarpeellista.

Tulevaisuuden uusia tutkimuksia ajatellen on muutamia asioita, joita tutkimuksissa voitaisiin ottaa huomioon. Ensinnäkin voisi olla tarpeellista tutkia aihetta kontekstissa, joka ei ole täysin maantieteellisesti rajattu. Tutkimuksen otanta voitaisiin muodostaa joko suuremmasta alueesta, kuten Pohjois-Amerikasta tai Euroopasta tai otanta voitaisiin yksinkertaisesti muodostaa useamman kuin yhden maan yrityksistä. Jatkotutkimuksissa voisi ylipäätään ottaa enemmän huomioon eurooppalaisia maita ja niiden yrityksiä, koska nykyinen aiheen tutkimus keskittyy hyvin vähän eurooppalaisiin yrityksiin. Lisäksi voitaisiin katsoa myös hyödylliseksi, että samalla kun maantieteellistä rajausta laajennettaisiin, myös mallien koulutukseen ja testaukseen käytetyn otannan suuruus kasvaisi samalla. Näin ollen mallin voitaisiin lopulta katsoa kykenevän tunnistamaan tilinpäätöspetoksia paremmin, kun otannan laajentamisen myötä mallia koulutettaisiin useamman kuin yhden maan yritysten tilinpäätöksien avulla.

Yksi seikka, joka voitaisiin myös ottaa huomioon jatkotutkimuksissa, on otannan muodostaminen uusista ja ei-tarkastetuista tilinpäätöksistä. Tehdyt tutkimukset aiheesta on toteutettu lähtökohtaisesti niin, että käytettyä petostenhavaitsemismallia sovelletaan jo tarkastettujen tilinpäätösten tutkimiseen. Tutkimuksissa petostenhavaitsemismalleja on siis käytetty jo ihmisten tarkas-

tamien tilinpäätösten luokitteluun. Voisi kuitenkin olla hyödyllistä tutkia, kuinka petostenhavaitsemismallit suotiutuvat uusien ja vielä tarkastamattomien tilinpäätöspetosten luokittelussa petollisiksi ja ei-petollisiksi. Esimerkiksi tässä tutkielmassa esitelty paras oikeinluokitteluprosentti 94,98 tarkoittaa käytännössä sitä, että tutkimuksessa käytetty malli on luokitellut 94,98 % tarkkuudella jo ihmisen luokittelemat tilinpäätöksen oikein. Voisi siis olla hyödyllistä tutkia, kuinka käytetyt mallit suoriutuvat tarkastamattomien tilinpäätösten luokittelussa, jonka jälkeen tilintarkastuksen ammattilainen kävisi läpi mallin tekemän luokittelun ja kuinka onnistuneesti luokittelu on tehty. Näin ollen saataisiin todennäköisesti paremmin selville, millaisessa roolissa algoritmeihin perustuvat menetelmät voivat olla tilintarkastuksen yhteydessä tulevaisuudessa. Kyseisellä menettelyllä voitaisiin myös mahdollisesti eritellä tarkemmin, millaisissa asioissa koneen tekemä työ on tehokkaampaa ihmisen tekemään työhön verrattuna, ja millaiset ihmisen kykyihin liitettävät piirteet ovat vielä konetta parempia ja tarvittavia tilintarkastuksessa.

Tämän lisäksi jatkotutkimuksissa otantakokojen suuruutta voisi lisätä ja paikoittain kiinnittää enemmän huomiota muuttujien valintaan. Otantakokoja suurentamalla tutkimukset tarjoaisivat aikaisempaa luotettavamman tulokset, ja otantakokoja voisikin lisätä juuri maantieteellistä rajausta laajentamalla. Muuttujia valittaessa tulisi jatkossakin huomioida muutkin kuin taloudelliset suhdeluvut, ja muuttujien valintaa tulisi tehdä perustellusti. Nykyiset tutkimukset osittain nojautuvat aiemmissä tutkimuksissa käytettyihin muuttujiin eli niissä käytetään täysin samoja muuttujia kuin jossakin aiemmassa tutkimuksessa. Muuttujien valinta tulisi kuitenkin olla tutkimuskohtaista, jolloin jokaisessa tutkimuksessa erikseen määritellään käytettävät muuttujat. Voisi myös perustellusti sanoa, että valittujen muuttujien lukumäärän tulisi olla runsas, jolloin petosten haviseminen kohdistuisi laajalti yrityksen eri tietoihin.

Yksittäisistä menetelmistä puhuttaessa tulvaisuudessa voisi olla perusteltua jatkaa syväoppimistekniikoiden kehittämistä, koska kyseiset tekniikat osoittautuivat tehokkaimmiksi. Yksi huomioitava menettelytapa, jota jatkotutkimuksissa voitaisiin soveltaa edelleen, on kaskivaiheinen menettely. Kaksivaiheinen menettely on ollut käytössä yhdessä koneoppimista ja kahdessa datanlouhintaa käsittelevässä tutkimuksessa. Näissä tutkimuksissa aluksi jonkin algoritmin avulla määriteltiin oleelliset muuttujat, ja sitten toinen algoritmi suoritti itse petosten havaitsemisen. Vaikka kyseiset tutkimukset eivät saavuttaneet oikeinluokitteluasteen perusteella parhaimpia mahdollisia tuloksia, voisi kaksivaiheista menettelyä perustellusti kuitenkin soveltaa myös muiden eri petostentunnistusmallien kanssa. Chenin (2016), Yehin (2016) ja Janin (2018) tutkimukset, joissa kaksivaiheista menettelyä hyödynnettiin, eivät edusta kaikkein uusimpia tutkimuksia aiheesta. Näin ollen kaksivaiheista menettelyä voisi sovittaa yhteen esimerkiksi juuri parhaimmat luokittelutulokset saavuttaneiden syväoppimistekniikoiden kanssa.

Janin (2018) tutkimus edusti oikeinluokitteluasteen perusteella parhaimpia tuloksia kaksivaiheisista menettelytavoista, kun tutkimuksen ANN&CART-luokittelumalli saavutti lähes 91 % oikeinluokitteluasteen. Xiuguon ja Shengy-

ongin (2022) tutkimuksessa puolestaan saavutettiin paras luokitteluaste yksittäisen menetelmän avulla, kun pitkäkestoinen lyhytkestomuisti saavutti 94,98 % oikeinluokitteluasteen. Yksi jatkotutkimusehdotus on, että muodostettaisiin kaksivaiheiden petostentunnistusmalli kahden edellä mainitun tutkimuksen parhaista osuuksista. Kaksivaiheiseksi malliksi muodostuisi ANN&LSTM (Artificial Neural Network&Long Short-term Memory). Kyseisessä mallissa keinotekoiset neuroverkot suorittaisivat menettelyn ensimmäisen vaiheen eli oleellisten muuttujien valinnan. Toisen vaiheen mallissa toteuttaisi pitkäkestoinen lyhytkestomuisti, joka suorittaisi tilinpäätöspetosten havaitsemisen ja yrityksien luokittelun petollisiksi ja ei-petollisiksi. Kyseinen muodostettu malli edustaisi tämänhetkisten tutkimuksien ja tässä tutkielmassa esiteltyjen tutkimuksien parhaimpia ja tehokkaimpia menetelmiä.

## 6 YHTEENVETO

Tässä kandidaatintutkielmassa tutkittiin, kuinka algoritmipohjaisia menetelmiä voidaan hyödyntää tilinpäätöspetosten havaitsemiseen. Viimeisien vuosikymmenien aikana algoritmipohjaiset menetelmät, kuten tekoälyn osajoukot, ovat kehittyneet huomattavasti mahdollistaen tehokkaan datan analysoimisen. Aihetta voidaan pitää merkittävänä, koska aiemmat tutkimukset ovat osoittaneet, että vain pieni osuus tapahtuneista tilinpäätöspetoksista havaitaan. Myös tilinpäätöspetoksissa menetetyn rahan määrä on ollut huomattavan suuri muihin petostyypppeihin verrattuna, ja tilinpäätöspetosten on katsottu aiheuttavan monia negatiivisia vaikutuksia useille eri sidosryhmille. Tutkielman tavoitteena oli muodostaa ajantasainen kuva tilinpäätöspetosten havaitsemisesta ja koota yhteen aiempien tutkimuksien merkittävimpiä huomiota ja tuloksia. Ajantasaisen kuvan muodostamiseksi tutkielmassa ei otettu huomioon yli 10 vanhoja tutkimuksia. Tutkielma on toteutettu kuvailevana kirjallisuuskatsauksena. Tarkastellun lähdeaineiston pohjalta tutkielmassa pyrittiin vastaamaan kahteen tutkimuskysymykseen:

- Millaisilla eri algoritmipohjaisilla menetelmillä tilinpäätöspetoksia voidaan havaita?
- Mitkä ovat tehokkaimmat lähestymistavat ja menetelmät tilinpäätöspetosten havaitsemiseen?

Tutkielmassa kävi ilmi, että algoritmeihin perustuvia menetelmiä petosten havaitsemiseksi on useita. Pääasiallisina menetelminä aiemmissa tutkimuksissa on hyödynnetty koneoppimiseen, syväoppimiseen ja datanlouhintaan perustuvia menetelmiä. Nämä päämenetelmät sisältävät useita yksittäisiä menetelmiä, jotka perustuvat algoritmeihin sekä matemaattisiin malleihin ja kaavoihin. Lähelläkohtaisesti yksittäisiä menetelmiä voitiin katsoa yhdistävän niiden kyky käsitellä dataa tehokkaasti ja järjestelmällisesti sekä niiden kyky tunnistaa suuresta datamäärästä oleellista ja haluttua tietoa.

Tutkielmassa selvisi myös, että valittaessa lähestymistapaa petosten havaitsemiseksi muuttujien valinnalla oli merkittävä rooli. Tutkimuksien perus-



teella osoittautui, että tilinpäätöspetoksia havaittaessa on tärkeää ottaa huomioon taloudellisten suhdelukujen lisäksi myös kielellisiä muuttujia, jotka voivat antaa tärkeää tietoa tilinpäätöspetosten havaitsemista varten. Sen lisäksi tutkielmassa selvisi, että petosten oikeinluokitteluasteen perusteella yksittäisistä menetelmistä tehokkaimmaksi osoittautuivat syväoppimiseen perustuvat menetelmät. Parhaimmillaan syväoppimiseen perustuvat menetelmät saavuttivat 94,98 % oikeinluokitteluasteen luokiteltaessa yrityksiä petollisiksi ja ei-petollisiksi niiden tilinpäätöksien perusteella.

Tuloksista voitiin päätellä, että tietyt algoritmipohjaiset menetelmät suoriutuvat kohtuullisen tarkasti havaitsemaan tilinpäätöspetoksia ja petoksesta indikoivia poikkeamia. Tutkimuksissa saaduista tuloksista voidaan johtopäätöksenä sanoa, että algoritmipohjaiset menetelmät tarjoavat uuden näkökulman ja suhteellisen tehokkaan lähestymistavan tilinpäätöspetosten havaitsemiseen. Esitellyissä tutkimuksissa ei kuitenkaan korostettu sitä, että tilintarkastajien tekemää työtä ei tarvittaisi enää ollenkaan, vaan menetelmien voidaan katsoa toimivan hyödyllisinä työkaluina tilintarkastajille.

Jatkotutkimuksia ajatellen voisi olla tarpeellista ja perusteltua tutkia tilinpäätöspetosten havaitsemista laajemmalla mittakaavalla. Tilinpäätöksistä koostuvan otannan voisi muodostaa useammasta maasta yhden maan sijaan, ja otantaan voisi sisällyttää esimerkiksi eurooppalaisia maita. Nykyinen aiheen tutkimus on rajoittunut tutkimaan aihetta vain harvojen maiden kontekstissa, ja aiheen tutkimus on maantiellisesti painottunut muun muassa aasialaisiin maihin. Lisäksi jatkotutkimuksissa otannan suuruutta sekä valittavien muuttujien määrää voisi lisätä. Tutkimuksissa tarkastelun kohteena ollut otanta oli paikoittain kohtuullisen pieni, ja muuttujien määrä vähäinen. Otannan kokoa kasvatamalla ja käytettyjen muuttujien määrää lisäämällä jatkotutkimuksien tuloksista voidaan saada luotettavampia ja yleistettävämpiä.

Jatkotutkimuksissa voisi ottaa huomioon myös kaksivaiheisen menettelyn, jota oli käytetty muutamissa tutkimuksissa. Kyseinen kaksivaiheinen menettely on ollut hyvin vähän esillä nykyisissä tutkimuksissa. Sen avulla voitaisiin kehittää useita eri yhdistelmämalleja, joiden suoriutumista tilinpäätöspetosten havaitsemiseen ei vielä ole kokeiltu. Jatkotutkimuksia varten tutkielmassa esiteltiin ehdotus kaksivaiheista menettelyä hyödyntävästä mallista ANN&LSTM, joka edustaisi tutkielmassa esiteltyjen tutkimuksien parhaimpia menetelmiä yhdistettynä.

## LÄHTEET

- Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics*, 115–123. <https://doi.org/10.52866/ijcsm.2022.01.01.013>
- Alenzi, H. Z., & O, N. (2020). Fraud Detection in Credit Cards using Logistic Regression. *International Journal of Advanced Computer Science and Applications*, 11(12). <https://doi.org/10.14569/IJACSA.2020.0111265>
- Ashtiani, M. N., & Raahemi, B. (2022). Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review. *IEEE Access*, 10, 72504–72525. <https://doi.org/10.1109/ACCESS.2021.3096799>
- Association of Certified Fraud Examiners. (2020). Report to the nations 2020 global study on occupational fraud and abuse. <https://acfe-public.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf>
- Association of Certified Fraud Examiners. (2022). Occupational Fraud 2022: A report to the nations. <https://acfe-public.s3-us-west-2.amazonaws.com/2022+Report+to+the+Nations.pdf>
- Association of Certified Fraud Examiners. (2024). Occupational Fraud 2024: A report to the nations. <https://www.acfe.com/-/media/files/acfe/pdfs/rtnn/2024/2024-report-to-the-nations.pdf>
- Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach. *Journal of Accounting Research*, 58(1), 199–235. <https://doi.org/10.1111/1475-679X.12292>
- Carmona, P., Climent, F., & Momparler, A. (2019). Predicting failure in the U.S. banking sector: An extreme gradient boosting approach. *International Review of Economics & Finance*, 61, 304–323. <https://doi.org/10.1016/j.iref.2018.03.008>

- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. *SpringerPlus*, 5(1), 89. <https://doi.org/10.1186/s40064-016-1707-6>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). *Introduction to Algorithms, fourth edition*. MIT Press.
- Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decision Support Systems*, 139, 113421. <https://doi.org/10.1016/j.dss.2020.113421>
- du, T. E. (2023). The red flags of financial statement fraud: A case study. *Journal of Financial Crime, ahead-of-print*(ahead-of-print). <https://doi.org/10.1108/JFC-02-2023-0028>
- Ertel, W. (2018). *Introduction to Artificial Intelligence*. Springer.
- García, S., Luengo, J., & Herrera, F. (2015). *Data Preprocessing in Data Mining*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-10247-4>
- Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23(1), 40–55. <https://doi.org/10.1038/s41580-021-00407-0>
- Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods. *Knowledge-Based Systems*, 128, 139–152. <https://doi.org/10.1016/j.knosys.2017.05.001>
- Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining and Knowledge Discovery*, 9(4), e1312. <https://doi.org/10.1002/widm.1312>

- Jan, C.-L. (2018). An Effective Financial Statements Fraud Detection Model for the Sustainable Development of Financial Markets: Evidence from Taiwan. *Sustainability*, 10(2), Article 2.  
<https://doi.org/10.3390/su10020513>
- Jan, C.-L. (2021). Detection of Financial Statement Fraud Using Deep Learning for Sustainable Development of Capital Markets under Information Asymmetry. *Sustainability*, 13(17), Article 17.  
<https://doi.org/10.3390/su13179879>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695.  
<https://doi.org/10.1007/s12525-021-00475-2>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349, 255–260.  
<https://doi.org/10.1126/science.aaa8415>
- Kanapickienė, R., & Grundienė, Ž. (2015). The Model of Fraud Detection in Financial Statements by Means of Financial Ratios. *Procedia - Social and Behavioral Sciences*, 213, 321–327.  
<https://doi.org/10.1016/j.sbspro.2015.11.545>
- Laghrissi, F., Douzi, S., Douzi, K., & Hssina, B. (2021). Intrusion detection systems using long short-term memory (LSTM). *Journal of Big Data*, 8(1), 65. <https://doi.org/10.1186/s40537-021-00448-4>
- Lin, C.-C., Chiu, A.-A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowledge-Based Systems*, 89, 459–470.  
<https://doi.org/10.1016/j.knosys.2015.08.011>
- Lokanan, M., Tran, V., & Vuong, N. H. (2019). Detecting anomalies in financial statements using machine learning algorithm: The case of Vietnamese listed firms. *Asian Journal of Accounting Research*, 4(2), 181–201.  
<https://doi.org/10.1108/AJAR-09-2018-0032>
- Mahesh, B. (2019). Machine Learning Algorithms -A Review. *International Journal of Science and Research*, 9(1), 381–386.  
[3.https://doi.org/10.21275/ART20203995](https://doi.org/10.21275/ART20203995)
- Mohmand, A., Sabharwal, M., & Echizen, I. (2021). An Assessment Study of Gait Biometric Recognition Using Machine Learning (s. 571–585).  
[https://doi.org/10.1007/978-981-15-3383-9\\_52](https://doi.org/10.1007/978-981-15-3383-9_52)

- Omar, N., Johari, Z. 'Amirah, & Smith, M. (2017). Predicting fraudulent financial reporting using artificial neural network. *Journal of Financial Crime*, 24(2), 362–387. <https://doi.org/10.1108/JFC-11-2015-0061>
- Patil, S., Nemade, V., & Soni, P. K. (2018). Predictive Modelling For Credit Card Fraud Detection Using Data Analytics. *Procedia Computer Science*, 132, 385–395. <https://doi.org/10.1016/j.procs.2018.05.199>
- Pham, B. T., Prakash, I., & Tien Bui, D. (2018). Spatial prediction of landslides using a hybrid machine learning approach based on Random Subspace and Classification and Regression Trees. *Geomorphology*, 303, 256–270. <https://doi.org/10.1016/j.geomorph.2017.12.008>
- Roiger, R. J. (2017). *Data Mining: A Tutorial-Based Primer, Second Edition* (2. p.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315382586>
- Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4), e1249. <https://doi.org/10.1002/widm.1249>
- Saleh Hussein, A., Salah Khairy, R., Mohamed Najeeb, S. M., & Alrikabi, H. Th. S. (2021). Credit Card Fraud Detection Using Fuzzy Rough Nearest Neighbor and Sequential Minimal Optimization with Logistic Regression. *International Journal of Interactive Mobile Technologies (ijIM)*, 15(05), 24. <https://doi.org/10.3991/ijim.v15i05.17173>
- Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, 69, 135–146. <https://doi.org/10.1016/j.indmarman.2017.12.019>
- Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1–37. <https://doi.org/10.2478/jagi-2019-0002>
- Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, 41, 100514. <https://doi.org/10.1016/j.accinf.2021.100514>
- West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66. <https://doi.org/10.1016/j.cose.2015.09.005>
- Wong, S., & Venkatraman, S. (2015). Financial Accounting Fraud Detection Using Business Intelligence. *Asian Economic and Financial Review*, 5(11), 1187–1207. <https://doi.org/10.18488/journal.aefr/2015.5.11/102.11.1187.1207>

- Xiuguo, W., & Shengyong, D. (2022). An Analysis on Financial Statement Fraud Detection for Chinese Listed Companies Using Deep Learning. *IEEE Access*, *10*, 22516–22532. <https://doi.org/10.1109/ACCESS.2022.3153478>
- Yeh, C.-C., Chi, D.-J., Lin, T.-Y., & Chiu, S.-H. (2016). A Hybrid Detecting Fraudulent Financial Statements Model Using Rough Set Theory and Support Vector Machines. *Cybernetics and Systems*, *47*(4), 261–276. <https://doi.org/10.1080/01969722.2016.1158553>
- Zhang, P., Li, T., Wang, G., Luo, C., Chen, H., Zhang, J., Wang, D., & Yu, Z. (2021). Multi-source information fusion based on rough set theory: A review. *Information Fusion*, *68*, 85–117. <https://doi.org/10.1016/j.inffus.2020.11.004>