

Riku Leino

SYSTEMAATTISEN JA TILASTOLLISEN LÄHESTYMISTAVAN  
KEHITTÄMINEN SJOITTAMISEEN

Liiketalous

2021

Tekijä(t) Leino, Riku	Julkaisun laji Opinnäytetyö, AMK	Päivämäärä 03/2021
	Sivumäärä 78	Julkaisun kieli Suomi
Julkaisun nimi <b>Systemaattisen ja tilastollisen lähestymistavan kehittäminen sijoittamiseen</b>		
Tutkinto-ohjelma Liiketalous, tradenomi		
Tiivistelmä  Opinnäytetyön aiheena on systemaattisen ja tilastollisen lähestymistavan kehittäminen sijoittamiseen. Se tehtiin omaan käyttöön tarpeiden mukaiseksi. Opinnäytetyötä tehdessä käytettiin omaa havainnointia, kirjallisuutta, internettiä ja tutkimuksia. Teoriaosuudessa käsitellään rahoitusmarkkinoiden toimintaa, portfolioteoriaa ja CAPM-mallia. Sijoitusprosessi ja teoriaa mukaileva päätöksenteko muodostaa yhtenäisen kokonaisuuden. Kokonaisuus on ohjelmoitu Pythonilla, jonka apuna on käytetty Exceliä.  Opinnäytetyön tavoitteena on kehittää yhtenäinen menetelmien kokonaisuus kvantitatiiviseen sijoittamiseen. Tuoton ennustamisessa käytetään historiallisia tilinpäätöstietoja ja osakekurssin kehitystä. Näiden avulla muodostetaan 21 aikajakson ennusteet tuotollisen perusteella, mikä olisi ainakin aikaisemmin kuvannut toteutunutta tuottoa hyvin. Yhtiöiden tuottoennusteet kootaan yhteen, minkä jälkeen niiden avulla muodostetaan simulointia varten mahdollisimman hyvään lopputulokseen johtava sijoitussuunnitelma.  Simulointien lopputuloksena syntyy 18 erilaista menetelmien yhdistelmää. Vertailutietona käytetään tasahajautusta yhtiöihin vuodesta 2002 lähtien. Arvioinnissa on käytetty keskimääräistä tuottoa, tuottojen keskihajontaa ja Sharpen lukua. Kaikilla menetelmillä oli joko korkeampi keskimääräinen tuotto tai pienempi tuottojen keskihajonta kuin tasahajautuksella. Kahdeksalla menetelmällä saavutettiin korkeampi keskimääräinen tuotto ja pienempi tuottojen keskihajonta kuin tasahajautuksella, jolloin myös Sharpen luku on korkeampi.		
<u>Asiasanat</u> Sijoittaminen, pörssi		

Author(s) Leino, Riku	Type of Publication Bachelor's thesis	Date 03/2021
	Number of pages 78	Language of publication: Finnish
Title of publication <b>Developing a systematic and statistical approach to investing</b>		
Degree program Bachelor of Business Administration		
Abstract  <p>The topic of the thesis is developing a systematic and statistical approach to investing. It was done for own use according to needs. During the thesis one's own observation, literature, internet and research were used. The theoretical part deals with the operation of financial markets, portfolio theory and the Capital Asset Pricing Model. The investment process and decision-making in accordance with the theory form a unified whole. The entity is programmed with Python, which has been assisted by Excel.</p> <p>The aim of the thesis is to develop a unified set of methods for quantitative investing. Historical financial statement data and share price developments are used to predict returns. These are used to generate predictions for 21 time periods based on what would have at least previously described actual returns well. Predicted returns of the companies are compiled together, after which they are used to form an investment portfolio that leads to the best possible outcome in the simulation.</p> <p>The simulations result in 18 different combinations of methods. Even diversification to companies since 2002 has been used as a benchmark. Average return, standard deviation and Sharpe ration have been used in the evaluation. All methods had either a higher average return or a lower standard deviation of the returns than even diversification. Eight methods achieved a higher average return and a lower standard deviation of the returns than even diversification, resulting in a higher Sharpe ratio as well.</p>		
<u>Key words</u> Investing, stock exchange		

## SISÄLLYS

1	JOHDANTO.....	5
1.1	Viitekehys .....	6
1.2	Tavoite .....	9
2	TIETOJEN MALLINTAMINEN.....	10
2.1	Perustiedot.....	12
2.2	Tiedon hyödyntäminen .....	15
2.3	Testaus .....	18
2.4	Sijoitusprosessi .....	19
3	KEHITTÄMINEN.....	22
3.1	Toteutustapa ja lähestyminen.....	23
3.2	Tutkimusmenetelmät.....	26
3.3	Jatkuva parantaminen.....	27
3.4	Automaatio.....	31
4	ENNUSTAMINEN JA SIMULOINTI .....	33
4.1	Aikajakso .....	33
4.2	Vastaavuudet ja ennustaminen.....	35
4.3	Simulointimenetelmät .....	37
4.4	Painoarvo .....	40
4.5	Simulointi.....	42
5	HALLINTA.....	45
5.1	Tunnusluvut ja toteutunut tuotto.....	46
5.2	Ennusteajakajakso.....	50
5.3	Ryhmittely ja iterointi .....	51
5.4	Luokittelu .....	54
5.5	Raportti .....	59
6	TEKNINEN TOTEUTUS .....	61
6.1	Excel .....	61
6.2	Ohjelmointi .....	63
7	LOPPUTULOS .....	66
7.1	Validiteetti.....	66
7.2	Reliabiliteetti.....	68
7.3	Loppusanat.....	72
	LÄHTEET.....	74

# 1 JOHDANTO

Olen sijoittanut muutaman vuoden ja sijoitusprosessin kehittäminen systemaattisemmaksi on ollut kiinnostava ja haasteellinen asia. Opinnäytetyön tekeminen oli mahdollista liittää sijoitusprosessin kehittämiseen, jossa ongelmana on perusteltujen päätösten tekemisen vaikeus ja tiedon suuri määrä. Kiinnostavien ostokohteiden ja tilanteiden löytämiseksi on toimittava järjestelmällisesti. Prosessissa voidaan hyödyntää nopeasti yleistynyttä ja ilmaiseksi saatavilla olevaa kehittyntä teknologiaa. Yhtiöiden historiallisia tilinpäätöstietoja ja osakekurssitietoa on saatavissa paljon, jota lähdetään hyödyntämään sijoituspäätösten tekemiseksi.

Opinnäytetyön menetelmillä yritetään selvittää, mihin osakkeisiin kannattaisi tällä hetkellä sijoittaa ja kuinka paljon. On tunnistettava, mitkä ovat olennaisia sijoituskohteen tuottoon vaikuttavia tietoja ja minkälaiseen tuottoon ne ovat historiallisesti johtaneet. Sijoituspäätöksen tukena olevien tunnuslukujen yhteisvaikutusta on vaikea arvioida yksinkertaisilla menetelmillä. Hyödyntämällä tilastollisia menetelmiä voidaan ainakin selvittää, mitä on aikaisemmin tapahtunut vastaavissa tilanteissa.

Opinnäytetyössä keskitytään menetelmien selvittämiseen ja yhteensovittamiseen, koska käytössä olevan aineiston kanssa ei ole vielä järkevää tehdä lopullisia päätöksiä parhaista menetelmistä. Päätöksentekoa perustellaan historiallisilla tapahtumilla, jossa samankaltaisten tapahtumien oletetaan toistuvan suunnilleen samankaltaisina.

Käytössä olevien tilinpäätöstietojen ja osakekurssitiedon avulla lasketaan kokonaisuutta kuvaavia tunnuslukuja. Ennustamisessa tunnusluville määritellään raja-arvot, joiden avulla tietokannasta etsitään vastaavia tilanteita. Vastaavuuksien perusteella määritellään tuottoennuste ja keskihajonta (riski). Pienillä raja-arvoilla annetaan enemmän painoarvoa tärkeimmille tunnusluville kuin suurilla raja-arvoilla. Lähitulevaisuus otetaan tilinpäätöstietojen osalta huomioon, koska se on jossakin määrin ennakoitavissa. Historiallisten tietojen osalta käytetään toteutuneita lukuja ikään kuin ne olisi pystytty ennakoimaan tarkasti. Tilinpäätöstietojen ennusteita on saatavilla nykyhetkeä varten yhtiötä seuraavien analyytikoiden analyyseistä tai ennusteet voidaan

myös laatia itse. Tulevaisuuden painoarvo pidetään kuitenkin maltillisena, koska sitä on hyvin vaikea ennakoida tarkasti pidemmällä aikavälillä.

Arviointia varten muodostetaan sijoitusportfolio, jonka sisältö muodostetaan tuottoenusteiden, keskihajontojen ja käytettävän sijoitussuunnitelman perusteella. Erilaisia sijoitussuunnitelmia vertaillaan simuloimalla sijoitusvarallisuuden kehitystä. Lopulliset valinnat riippuvat sijoittajan mieltymyksestä tuoton ja riskin suhteen.

## 1.1 Viitekehys

Rahoitusteoria keskittyy siihen, miten markkinat toimivat tiettyjen oletusten mukaan. Tehokkaiden markkinoiden teorian mukaan rahoitusmarkkinoiden tehtävä on allokoita resursseja yli- ja alijäämäsektorin välillä ja varat hinnoitellaan edullisimmalla tavalla kaikille markkinaosapuolille. Kaiken saatavissa oleva tiedon pitäisi sisältyä arvopapereiden hintoihin, jolloin ylituottoja ei pitäisi olla mahdollista saavuttaa systemaattisesti. Kaikkien sijoittajien ei kuitenkaan tarvitse toimia rationaalisesti tehokkaiden markkinoiden toteutumisista varten vaan riittää, että riittävän moni sijoittaja tekee sijoituspäätöksensä rationaalisesti. (Tyrylahti 2011, 2-7.)

Tehokkailla markkinoilla osakekurssin reaktio uuteen tietoon pitäisi tapahtua välittömästi oikealle tasolle, mutta sijoittajat käsittelevät tietoa yleensä oman aikansa, minkä jälkeen kurssi nousee hitaasti oikealle tasolle. Väärinymmärrykset vaikuttavat myös omalta osaltaan osakekurssiin. Osalla sijoittajista on mahdollisuus suurempiin tuottoihin pienemmällä riskillä, kun markkinat eivät toimi tehokkaasti. Heikosti tehokkailla markkinoilla kurssikehityksen perusteella ei pitäisi pystyä ennustamaan tulevaa kurssikehitystä. Puolivahvasti tehokkailla markkinoilla tilinpäätösanalyysin perusteella ei pitäisi pystyä ennustamaan tuottoja, koska tietojen pitäisi olla kaikkien osapuolien tiedossa. Vahvasti tehokkailla markkinoilla osakekurssiin sisältyy kaikki julkinen tieto sekä sisäpiirin tieto ennen kuin tiedot on edes julkistettu. Osakkeiden satunnainen poimiminen tai indeksisijoittamisen pitäisi olla paras strategia tehokkailla markkinoilla. (Puttonen 2018, 168-179.)

Behavioristisessa rahoitusteoriassa huomioidaan psykologisten tekijöiden näkökulma. Tutkimuksissa on todistettu ajattelutavoissa tapahtuvan systemaattisia virheitä. Tästä

syystä sijoitustoiminnassa syntyy yli- ja alireagoitteja. Behavioristisessa rahoitusteoriassa sijoittajien käyttäytyminen näkyy osakemarkkinoilla myös kurssiheiluntana. Rahoitusteoria ja behavioristinen teoria eivät kuitenkaan ole toisiaan pois sulkevia eikä niitä voida vertailla paremmuudessa. Teorioiden hyödyntäminen yhdessä antaa sijoittajalle lisää ymmärrystä ja mahdollisuuden hyödyntää tietoa toimiessaan osakemarkkinoilla. Epäratio-naalisen käyttäytymisen ollessa markkinoilla systemaattista, voidaan siitä myös syste-maattisesti hyötyä. (Tyrylahti 2011, 10-11.)

Nykyään iso osa kaupankäynnistä ei liity yhtiön perusteissa tapahtuviin muutoksiin, vaan ne tehdään muista syistä. Yhtiön markkina-arvon ollessa eriytynyt yrityksen to-dellisesta arvosta, rationaaliset sijoittajat eivät välttämättä pysty korjaamaan hinnoit-teluvirhettä. Laumakäyttäytyminen sijoitusmarkkinoilla on arvioiden mukaan yksi syy, mikä voimistaa osakekurssin nousuja ja laskuja. (Niskanen 2016, 170-171.)

Sijoittajille olennaista pitäisi olla yhtiön markkinahinta, jonka perusteena on tulevat kassavirrat. Periaatteessa arvonmääritys on yksikertainen toimenpide, jossa tulevat kassavirrat diskontataan valitulla tuottovaatimuksella nykyarvoon. Kassavirtoihin si-sältyvät ainakin osingot ja markkinahinnan muutoksesta johtuva myyntivoitto tai myyntitappio. Markkinahinta on herkkä oletettujen osinkojen ja tuottovaatimuksen muutoksille, koska yrityksen elinkaari saattaa olla ikuinen. Jo pelkästään osinkojen ennustaminen muutaman vuoden päähän on haastavaa kokeneellekin analyytikolle, jo-ten yhtiön tulevaisuuden näkymien ennustaminen saman aikaperiodin lopussa on lähes mahdotonta. (Puttonen 2018, 77-100.)

Sijoituskohteet muodostavat sijoitusportfolion, jonka tuottoennuste on osakkeiden pai-notettu keskiarvo. Portfolioteoriassa riskittömän tuoton taso, odotettu tuotto ja odote-tun tuoton keskihajonta muodostavat yhtenäisen kokonaisuuden. Hajauttamalla use-ampaan osakkeeseen on mahdollista pienentää sijoitusportfolion tuoton keskihajontaa pienemmäksi kuin yksittäisen osakkeen keskihajonta. Hajautushyöty riippuu osakkei-den tuottojen korrelaatiosta. Sijoitusportfolio muodostetaan minimivarianssiportfolion yläpuolella olevalta tehokkaalta alueelta ja sen lopulliseen valintaan vaikuttaa sijoitta-jan mieltymykset tuoton ja riskin suhteen. (Yritysrahoitus, 171-188)

Mallin kehittämisessä teorioihin tutustuminen on vaikuttanut valittuihin menetelmiin, yhtenäisen kokonaisuuden luomisessa. Oikeaa yrityksen markkina-arvoa ei yritetä laskea, vaan yrityksen markkina-arvo on tietyn suuruinen ja se on matkalla kohti jotakin markkina-arvoa. Arvon oletetaan muuttuvan samalla tavalla kuin se on aikaisemminkin sijoittajien toimesta muuttunut vastaavassa tilanteessa. Tunnistamalla kurssinousun ja laskuun vaikuttavia tekijöitä pyritään saamaan ylituottoa verrattuna tasahajautettuun sijoitusportfolioon. Sijoitusportfolio muodostetaan lopuksi sen perusteella, mikä on aikaisemmin johtanut haluttuun lopputulokseen.

Opinnäytetyössä ei suoraan seurata portfolioteoriaa tai siitä edelleen kehitettyä CAPM-mallia (Capital Asset Pricing Model), mutta päätöksenteossa on mukailtu menetelmien päätöksentekoa. CAPM-mallissa sijoituksen tuottoennuste määritellään suhteessa koko markkinatuottoon. Edelleen kehitetyissä kolmen ja viiden faktorin mallissa tuottoa pystytään selittämään paremmin faktorien eli tunnuslukujen avulla. (Fama, French 1993; Fama, French 2013)

Tärkein hyödynnetty asia on alla havainnollistettu tuotto-keskihajonta-kuvio, jossa pystyakselilla on keskimääräinen tuotto ja vaaka-akselilla tuoton keskihajonta. Kuviossa on esitetty yhden yhtiön yhdelle päivälle lasketut tuottoennusteet ja keskihajonnat. Oikeanpuolisessa pisteessä ennusteajajakso on lyhin ja keskihajonta suurin. Aikajakson pidentyessä keskihajonta tyypillisesti laskee ja pisteet siirtyvät vasemmalle. Kuviossa voidaan esittää joko tuottoennuste tai toteutunut tuotto ja näiden keskihajonnat. Pystyakselilla suurempi tuottoennuste on parempi vaihtoehto kuin pienempi ja vaaka-akselilla pienempi keskihajonta on parempi kuin suurempi. Kuviota hyödynnetään tuottoennusteen valinnassa eri aikajaksoilla ja sijoitusportfolion muodostamisessa. Tuotto voi olla positiivista tai negatiivista, mutta keskihajonta on aina positiivinen. Behavioristinen lähestymistapa näkyy mallissa tuottoennusteen muodostuksessa. Toteutuneen kurssikehityksen ennustetaan olevan tuottoennusteen ja keskihajonnan muodostamisessa rajoissa.



Kuvio 1. Tuotto-keskihajonta-kuvio

Mallissa tehtävät valinnat perustuvat tuottoennusteen, keskihajontaan tai näiden suhdeluukuun, joka tunnetaan paremmin Sharpen lukuna. Mitä suurempi Sharpen luku, sitä parempi on tuotto-riski-suhde (Sharpe 1966). Kuviossa 1 punaisen viivan osoittaman ennustepisteen Sharpen luku on joukon korkein. Minimivarianssi olisi pistejoukon vasemmanpuoleisimmassa pisteessä, jonka yläpuolella sijaitsevat korkeimman Sharpen luvun piste ja suurimman tuottoennusteen piste. Riskitöntä korkoa kuviossa ei olla huomioitu, mutta huomioitaessa punaisen viivan vasemmanpuolista päätä nostettaisiin riskittömän koron verran.

## 1.2 Tavoite

Opinnäytetyötä on edeltänyt huomattavasti karkeampi malli, jonka pohjalta syntyi idea opinnäytetyön aiheesta. Aikaisempi malli tehtiin ensin muutamalle yhtiölle, jonka jälkeen sitä pystyttiin laajentamaan muihin yhtiöihin. Samanlaista lähestymistapaa pyritään hyödyntämään myös tässä, mutta laajennusvaihe ajoittuu opinnäytetyön valmistumisen jälkeiseen aikaan. Opinnäytetyön tulokset tulevat omaan käyttöön, joten toimeksiantajana toimin itse. Pieniä muutoksia ja yksityiskohtien hiomista jää tehtäväksi myös opinnäytetyön valmistumisen jälkeen.

Opinnäytetyön tavoitteena on kehittää yhtenäinen menetelmien joukko kvantitatiiviseen sijoittamiseen. Käytössä olevan historiallisen tiedon avulla muodostetaan

lähitulevaisuuden kurssikehitystä kuvaava tuottoennuste sen perusteella, mikä olisi ainakin aikaisemmin kuvannut tuottoa mahdollisimman hyvin. Yhtiöiden tuottoennusteet kootaan yhteen, minkä jälkeen niiden avulla muodostetaan mahdollisimman hyvään lopputulokseen johtava sijoitussuunnitelma sijoitusvarallisuuden allokoimiseksi. Aiheeseen liittyy monta osa-aluetta, joiden kehitystä on tarkoitus jatkaa tulevaisuudessa, mikäli lähestymistapa näyttää opinnäytetyön valmistuessa lupaavalta.

Menetelmissä hyödynnetään tekoälyssä käytettyjä menetelmiä kuten neuroverkkoa, regressiota, virhefunktiota, aktivointifunktiota ja vastavirta-algoritmia. Tekninen toteutus ei ole opinnäytetyön keskipisteessä, mutta se on kuitenkin erittäin tärkeä osa toteutusta. Koneoppimiseen ja menetelmiin sisältyy paljon laskentaa, minkä tehostaminen on tärkeää mallin myöhemmän laajentumisen kannalta.

Lopullisia tuloksia menetelmien käytöstä ei tavoitella, koska aineistoa ei vielä ole riittävästi. On tärkeää saada ymmärrystä erilaisista vaihtoehdoista ja käsitys aiheeseen liittyvästä teoriasta ja sen hyödyntämisestä. Esitetyt menetelmät on voitava kuitenkin sovittaa yhteen saumattomasti ja loogisesti. Eteen tulevissa valintatilanteissa mukailaan sijoittamisen taustalla olevaa teoriaa. Tehtävät valintapäätökset perustuvat tuloksiin, jolloin ne ovat paremmin perusteltuja kuin mielipiteet.

Alustavien tulosten avulla tavoitellaan parempaa tuottoa kuin hajauttamalla tasaisesti yhtiöihin. Parempi tuotto voi tarkoittaa korkeampaa keskimääräistä vuosituottoa, pienempää tuottojen keskihajontaa tai suurempaa Sharpen lukua. Tulokset riippuvat vahvasti tulevien tuottojen ennustamiskyvystä, jota arvioidaan erikseen.

## 2 TIETOJEN MALLINTAMINEN

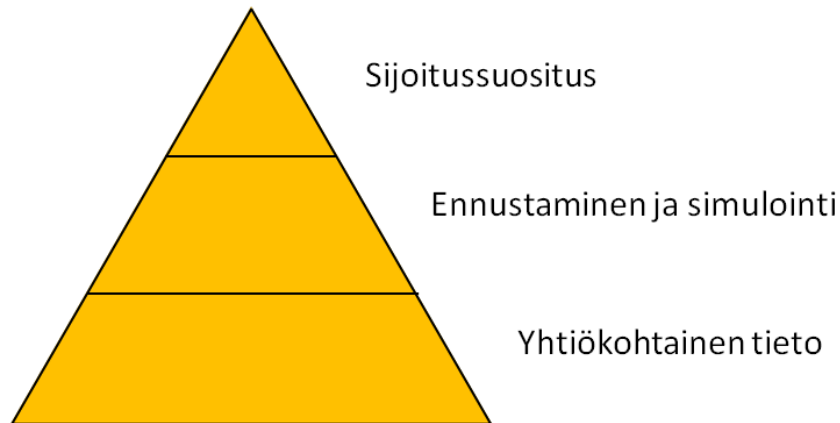
Menetelmien avulla pyritään ensiksi ennustamaan osakkeen tulevaa tuottoa, joka koostuu maksetusta osingosta ja osakekurssin muutoksesta. Toiseksi ennusteiden perusteella tehdään sijoituspäätös, joita arvioidaan simuloimalla varallisuuden kehitystä. Yhtiöiden arvonmäärittämiseen vaikuttava tieto yritetään tunnistaa, minkä jälkeen

tietoja käytetään ennusteiden määrittelyyn. Tärkeimpiä tietolähteitä ovat yhtiön julkaisemat tiedotteet, analyytikoiden ennusteet ja historiallinen kurssikehitys. Kaikkien lähteiden tiedot kerätään yhteen paikkaan tunnuslukujen laskentaa varten. Jokaisen päivän tunnusluvut ja oleelliset tiedot kootaan yhtenäiseen tietokantaan. Kun minkä tahansa päivän tiedot ovat tarpeeksi lähellä ennustettavan päivän tietoa, niin päivästä tulee ennustettavan päivän vastaavuus. Tietojen avulla muodostetaan jokaiselle päivälle vastaavuuksista koostuva tietyn suuruinen ryhmä tuottoennusteen laskemista varten. Ryhmän tietojen avulla lasketaan päivittäinen tuottoennuste ja keskihajonta, jota havainnollistetaan myöhemmin.

Jatkuvan parantamisen avulla mallia kehitetään ja laajennetaan vaiheittain. Sijoittajien käyttäytymisen oletetaan tulevaisuudessa pysyvään samanlaisena kuin aikaisemminkin. Ennustettavan päivän vastaavuuksien toteutuneiden tuottojen oletetaan toistuvan samankaltaisena tuottona myös tulevaisuudessa. Sijoittajien samankaltainen käyttäytyminen on välttämättömyys mallin toiminnalle, jota ilman tällaisen menetelmän käyttö olisi hyödytöntä.

Sijoitusportfolion kehityksen simulointi alkaa aikaisintaan vuoden 2002 ensimmäisestä kvartaalista eteenpäin. Tällöin tarvittavat perustiedot kahdelta edeltävältä vuodelta tunnuslukujen laskentaa varten on kerätty. Toteuttamiseen valitaan aluksi yhtiöitä suunnilleen samalta toimialalta, jolloin nähdään paremmin, voisiko tämän kaltaisen mallin käyttö olla järkevää. Simuloinnissa kaupankäynti osakkeilla suoritetaan päivän päätöskurssilla, jolloin osakkeita voidaan joko myydä tai ostaa. Osakkeiden lyhyeksi myynnin tai päivänsisäistä kaupankäynnin mahdollisuutta ei opinnäytetyössä hyödynnetä. Sijoituspäätökset perustuvat täysin käytössä olevaan historialliseen tietoon.

Suunnitteluvaiheessa määritellään toimintaympäristön keskeiset käsitteet, joiden on tarkoitus muodostua loogiseksi ja yhtenäiseksi kokonaisuudeksi. Mallia pyritään kuvailemaan mahdollisimman selkeästi ja ymmärrettävästi. Mukailleen Ari Hovin pyramidirakenteista mallia, tietoja voidaan kuvata kolmessa eri tasossa. (Oppia [www-sivut](http://www-sivut)).



Kuvio 2. Pyramidirakenne

Ylätaso suunnitellaan karkealla tavalla niin, että se on helppo esittää ja kuvaa keskeisimmät asiat. Ylätasolla esitetään sijoitussuositus mukana oleviin yhtiöihin. Sijoitussuositus perustuu keskitasolla tehtävään analytiikkaan ja simulointiin, jonka keskeiset tiedot kuten käytetyt tunnusluvut ja ennusteet esitetään ylätasolla. Keskitason osa-alueet suunnitellaan mahdollisimman tarkalla tasolla. Historiallisen tiedon pohjalta määriteltävä tuottoennuste ja keskihajonta muodostavat yhden osa-alueen. Toisella osa-alueella määritellään, kuinka paljon yhtiöihin kannattaisi sijoittaa ennusteiden perusteella. Keskitasolla tehdään suurin osa menetelmien kehitystoimenpiteistä, jotka voidaan ottaa käyttöön ylätasolla, jos suorituskyky on muutosten jälkeen parantunut. Alatasolla kuvataan yhtiökohtaisesti kaikki kerättävä tieto, joiden avulla lasketaan ylempillä tasoilla tarvittavat tunnusluvut. (Oppia [www-sivut](#))

## 2.1 Perustiedot

Tilinpäätösanalyysissä käytettäviä tietoja ja tunnuslukuja on listattu alla olevaan taulukkoon. Kaikkia tietoja ei yritetä hyödyntää heti, vaan suurin osa on tarkoitus ottaa osaksi mallia menetelmien vakiinnuttua. Tällöin kokonaisuuden olennaiset osat saadaan helpommin yhdistettyä toisiinsa. Tuloslaskelman tietojen osalta hyödynnetään esimerkiksi tilikauden tuloksen muutosta ja tuloksen suhdetta yhtiön markkina-arvoon. Markkina-arvo saadaan kertomalla osakkeiden lukumäärä osakkeen hinnalla. Tunnuslukujen laskentakaavat ovat pitkälti vakiintuneita ja niiden pitäisi löytyä helposti verkosta, mutta joskus laskennassa voi olla hieman toisistaan eroavia

laskutapoja. Esimerkiksi oman pääoman kohdalla voidaan käyttää joko keskimääräistä oman pääoman määrää tai kauden lopussa olevaa määrää, jolloin lopputulos on eri suuruinen. Eroavaisuuksilla ei kuitenkaan pitäisi olla merkittäviä vaikutuksia lopputulokseen. Tunnuslukujen ja perustietojen listaus auttaa hahmottamaan kuinka paljon erilaista tietoa on saatavilla jo pelkästään yhtiön tilinpäätöksistä. Tunnuslukuja voidaan nykyhetken lisäksi tarkastella menneiden ja ennustettujen lukujen näkökulmasta.

<b>Tilinpäätösanalyysi</b>	
<b>Perustiedot</b>	<b>Tilinpäätösanalyysi</b>
<ul style="list-style-type: none"> <li>• Liikevaihto</li> <li>• Liikevoitto</li> <li>• Voitto ennen veroja</li> <li>• Tilikauden voitto</li> <li>• Pitkäaikaiset varat</li> <li>• Liikearvo</li> <li>• Lyhytaikaiset varat</li> <li>• Vaihto-omaisuus</li> <li>• Myyntisaamiset</li> <li>• Taseen loppusumma</li> <li>• Oma pääoma</li> <li>• Pääomalainat</li> <li>• Pitkäaikaiset velat</li> <li>• Lyhytaikaiset velat</li> <li>• Ostovelat</li> <li>• Korolliset velat</li> <li>• Liiketoiminnan rahavirta</li> <li>• Investointien rahavirta</li> <li>• Rahoituksen rahavirta</li> <li>• Rahavarat</li> <li>• Korollinen vieras pääoma</li> </ul>	<ul style="list-style-type: none"> <li>• Myyntikate-%</li> <li>• Käyttökate-%</li> <li>• Liiketulos-%</li> <li>• Nettotulos-%</li> <li>• Rahoitustulos-%</li> <li>• Kokonaispääoman tuotto-%</li> <li>• Sijoitetun pääoman tuotto-%</li> <li>• Oman pääoman tuotto-%</li> <li>• Omavaraisuusaste</li> <li>• Velka-%</li> <li>• Nettovelka-%</li> <li>• Nettovelkaantumisaste-%</li> <li>• Käyttöpääoma-%</li> <li>• Myyntisaamisten kiertoaika</li> <li>• Ostovelkojen kiertoaika</li> <li>• Nettokäyttöpääoma-%</li> <li>• Quick ratio</li> <li>• Current ratio</li> <li>• Lainojen hoitokate</li> <li>• Liikevaihdon muutos-%</li> <li>• Nettorahoituskulut, %</li> </ul>

Kuvio 3. Tilinpäätösanalyysin tietoja (Heikinmatti, Jahkonen, Kanervisto, Kekki, Marjomaa, Ruusulaakso & Toivio 2017)

Sijoittajan kannalta yhtiön markkina-arvo on erittäin olennainen osa sijoituspäätöksen tekoa. Markkina-arvoa verrataan moniin yhtiön tilinpäätöstietoihin. Näiden

suhdelukujen perusteella moni sijoittajista etsii tukea sijoituspäätökselle. Riittävän monen sijoittajan seurattessa samoja tunnuslukuja pitäisi niiden myös vaikuttaa tuotoihin samalla tavalla eri ajanjaksoina. Sijoittajien yleisesti käyttämien tunnuslukujen kirjainlyhennyksiä on selitetty tietojen ja tunnuslukujen yhteydessä alla olevassa kuviossa.

<b>Pörssin tiedot ja tunnusluvut</b>	
<b>Selitykset</b>	<b>Tiedot ja tunnusluvut</b>
<ul style="list-style-type: none"> <li>• P = Markkina-arvo</li> <li>• E = Tilikauden tulos</li> <li>• CF = Liiketoiminnan kassavirta</li> <li>• EV = Yritysarvo</li> <li>• EBIT = Liikevoitto</li> <li>• EBITDA = Käyttökate</li> <li>• S = Liikevaihto</li> <li>• B = Oma pääoma</li> </ul>	<ul style="list-style-type: none"> <li>• Osakekurssi</li> <li>• Osakkeiden kokonaismäärä</li> <li>• Markkina-arvo</li> <li>• Yritysarvo</li> <li>• P/E-luku</li> <li>• P/CF-luku</li> <li>• EV/EBIT-luku</li> <li>• EV/EBITDA-luku</li> <li>• P/S-luku</li> <li>• P/B-luku</li> <li>• Tulostuotto-%</li> <li>• Osinkotuotto-%</li> <li>• Osakekohtainen tulos (EPS)</li> <li>• Osakekohtainen osinko</li> <li>• Osakekohtainen oma pääoma</li> <li>• Osinkosuhde</li> </ul>

Kuvio 4. Pörssin arvostusluvut (Alma Talentin www-sivut 2021)

Esimerkeissä on käytetty muutamia tunnuslukuja, joista P/E-luku tarkoittaa yhtiön markkina-arvoa jaettuna 12 kuukauden juoksevalla tuloksella. Vaihtoehtoisesti P/E-luku voidaan laskea jakamalla osakurssi 12 kuukauden juoksevalla osakekohtaisella tuloksella. P/B-luku saadaan jakamalla markkina-arvo oman pääoman määrällä. Tuotto on osakekurssin muutosprosentti lisätynä osinkotuotolla. Liikevoiton ja tuloksen muutoksella mitataan, kuinka paljon juoksevan 12 kuukauden tulos muuttuu yhden

kvartaalin aikana. Kvartaalipäivän avulla kuvataan sitä, kuinka monta päivää viimeisimmän taloudellisen raportin julkaisusta on kulunut.

Tunnusluvut eivät aina ole aivan ongelmattomia, koska perustiedot voivat olla negatiivisia tai positiivisia, mikä vaikeuttaa niiden laskemista. Tunnusluvulle voi tulla esimerkiksi väärä etumerkki laskettaessa muutosta. Tilanteet pitää tunnistaa ja laskea tunnusluku oikealla etumerkillä.

Lukuarvon nimittäjän ollessa lähellä nollaa, muutosprosentit voivat olla suuria, mikä aiheuttaa myös ongelman. Ratkaisuna voidaan määrittää suurin ja pienin arvo, joka tunnusluvulle annetaan. Osakkeiden lukumäärä on kerätty vuosineljänneksittäin, mikä aiheuttaa ongelmia, koska lukumäärä voi muuttua osakeannin takia vuosineljänneksen aikana. Ongelman takia on tunnistettava tunnusluvut ja päivät, jolloin muutokset osakkeiden lukumäärässä otetaan huomioon.

Osakkeen kurssikehityksen perusteella voidaan laskea kymmeniä teknisen analyysin tunnuslukuja, joita voidaan käyttää apuna tuoton ennustamisessa. Opinnäytetyössä on tässä vaiheessa hyödynnetty teknisen analyysin osalta viimeisen 91 kaupankäyntipäivän tuottoa. Momentum-ilmiössä tuleva tuotto on samansuuntainen kuin aikaisemmin toteutunut tuotto, jolloin on mahdollista saada parempaa tuottoa sijoitukselle (Dung 2013, 48). Menneen tuoton aikaväli valittiin regressioanalyysin suurimman regressio-kertoimen perusteella. 91 kaupankäyntipäivää oli pisin tarkasteltava ajanjakso, joten pidempi aikajakso voisi olla parempi valinta. Useita aikajaksoja olisi mahdollista käyttää yhdessä, jolloin voitaisiin saavuttaa parempi lopputulos. Regressioanalyysin tuloksia on tarkemmin esitettyinä kuviossa 24.

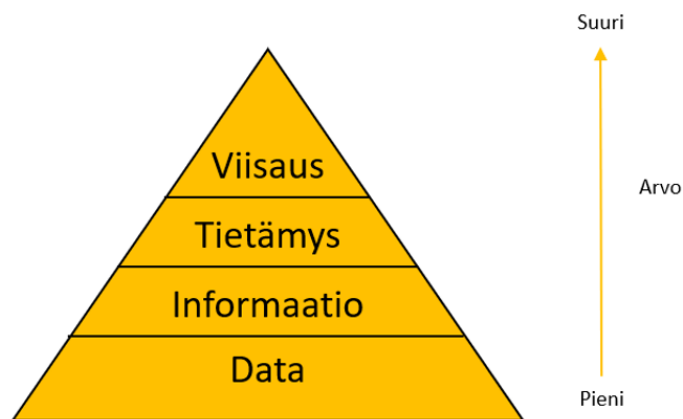
## 2.2 Tiedon hyödyntäminen

Yhtiö on taloudellinen kokonaisuus, josta kerättävät tiedot muodostavat tarinaa numeroiden muodossa. Lähtötilanteessa tiedot ovat hajallaan ulkoisissa tietolähteissä yhtiöiden ja muiden toimijoiden verkkosivuilla. Tiedot haetaan manuaalisesti tai automaattisesti ja tallennetaan myöhempää prosessointia ja analysointia varten. Historiantiedot ovat olennaisia aikaperspektiivin luomiseksi. Uuteen tietoon voidaan reagoida

nopeammin, kun tiedetään, mitä aikaisemmin on tapahtunut ja kuinka nopeasti. Väärä tai väärin ymmärretty tieto voi toisaalta johtaa väriin johtopäätöksiin. Tunnusluvut lasketaan aina samalla kaavalla yhden totuuden periaatteiden mukaisesti, mikä pienentää virheiden mahdollisuutta. (Hovi, Hervonen & Koistinen 2009, 4-16)

Käytössä hyödynnetään ainoastaan strukturoitua tietoa kuten päivämääriä, kurssitietoa ja tilinpäätöstietoja. Yhtiöiden tilinpäätöksistä ja osavuosikatsauksista hyödynnetään tuloslaskelmasta, taseesta ja rahoituslaskelmasta löytyviä tietoja. Strukturoimattomia tietoja olisi esimerkiksi erilaiset irralliset tekstit, kuvat tai videot. Tilinpäätöstiedot ja kaupankäyntitiedot muodostavat jalostusketjun jälkeen operatiivisen tietokannan, jota hyödynnetään analysoinnissa. Päätöksenteko perustuu ainoastaan näihin tietoihin. (Hovi, Hervonen & Koistinen 2009, 18-30)

DIKW(Data, information, knowledge, wisdom)- pyramidi kuvaa tilannetta, jossa datan merkityksen kasvaessa siitä muodostuu ensin informaatiota. Merkityksen edelleen lisääntyessä informaatiosta syntyy tietämystä ja lopulta viisautta. Datasta saatava arvo kasvaa matkalla kohti viisautta. (Sandholm 2020, 13)



Kuvio 5. DIKW- pyramidi (Sandholm 2020, 13)

Data voi olla esimerkiksi tilinpäätöksen tieto liikevaihdosta, josta saadaan informaatiota laskemalla muutos aikaisemman ajankohdan liikevaihtoon verrattuna. Tietämystä syntyy yhdistämällä informaatiota kuten esimerkiksi liikevaihdon 5 prosentin kasvu ja sijoituksen tuoton 10 % kasvu. Tapahtumien jälkeen tiedetään liikevaihdon muutoksen

todennäköinen vaikutus tuottoon. Viisautta syntyy, kun havaintojen ja matemaattisen mallin perusteella ennustetaan miksi ja kuinka paljon yhtiöstä saatava tuotto käyttäytyy tulevaisuudessa. (Figueroa, 2019)

Business Intelligencen avulla vastataan tiedon jalostamisesta, analysoimisesta ja hyödyntämisestä. Apuna käytetään ennalta määriteltyjä suorituskyvyn mittareita, kuten ennustevirhettä, regressiokerrointa, korrelaatiota, tuottoa ja keskihajontaa. Tavoitteena on päätöksenteon nopeuttaminen, parantaminen, oikea-aikainen vastaus tietotarpeeseen, strategian toteuttaminen ja tehokkuuden parantaminen. Menneestä ajasta voidaan laatia raportti, seurata nykyhetkeä tai ennakoida tulevaa. Menetelmien tunnuspiirteitä ovat analyttisyys, kvantitatiivisyys ja liiketoimintalähtöisyys. (Hovi, Hervonen & Koistinen 2009, 74-82)

Kaikki tiedot eivät ole saatavilla täysin ajantasaisena. Kurssitieto on mahdollista hakea pörssin aukioloaikana, mutta tilinpäätösinformaatio julkaistaan yleensä selvästi viiveellä. Toisinaan ennakoitua olennaista tietoa julkaistaan jo ennen tilikauden päättymistä. Analyysipalveluiden analyytikot yrittävät ennakoida tulevaa kehitystä, joten näitä tietoja voidaan käyttää apuna lähitulevaisuuden ennakoimiseen. Tiedot jaetaan osiin, jotka päivittyvät päivittäin tai kvartaaleittain. Harvemmin kerättävä tieto kohdistetaan päivittäisiin tietoihin päivälle, jolloin se on julkaistu. (Hovi, Hervonen & Koistinen 2009, 29-30)

Tietovaraston suunnittelun lähtökohtana on sen laajuus, käyttötarpeet ja saatavilla olevat tiedot. Tiedot kannattaa kerätä alusta asti, jos niitä voidaan hyödyntää tulevaisuudessa. Tiedon tarkkuus vaikuttaa suoraan tietovaraston laajuuteen. Kerätessä tietoa yksityiskohtaisemmalla tasolla, tietovaraston laajuus kasvaa samassa suhteessa. Käsittelemättömät tiedot tallennetaan omiin tauluihin, mistä niitä lähdetään jalostamaan tarpeen mukaan erilaisten kaavojen avulla yhteenvetotauluihin, joissa ovat laskettuna kaikki hyödynnettävät tunnusluvut. Laskemisen jälkeen kaikkien yhtiöiden tiedot kootaan yhteen. (Hovi, Hervonen & Koistinen 2009, 31-47)

Jokaiselle päivälle voidaan laskea esimerkiksi 100 tunnuslukua tai muuta tietoa, joiden lisäksi lisätään 3 ennustetietoa jokaista 21 aikajaksoa kohti. Vuodessa on noin 250 kaupankäyntipäivää, jolloin jokaista yhtiötä kohti löytyy n. 40 000 tietoa yhtä vuotta

kohden. Tulevaisuudessa tietoa voisi olla kerättynä 50 yhtiöltä 10 vuoden ajalta, jolloin yksittäisten tietojen määrä olisi 20 miljoonaa. Tietomäärän pitäisi olla edelleen helposti hallittavissa. Yksittäisiä rivejä olisi tässä vaiheessa suunnilleen 125 000.

### 2.3 Testaus

Testauksen avulla tavoitellaan ennustamisessa tunnuslukujen optimaalista hyödyntämistä, halutulla vastaavuuksien määrällä. Lisättäessä uuden yhtiön tiedot kasvanut tietokanta vaikuttaa myös saataviin tuloksiin. Tulokset tulevat alkuvaiheessa elämään voimakkaammin, kun tietojen määrä kasvaa nopeasti. Lisääntyneen tiedon pitäisi parantaa luottamusta menetelmiin ja parantaa ennustavuutta, kun voidaan käyttää pienempiä raja-arvoja ja niiden muutosnopeuksia. Huonosti hyödynnetyt tunnusluvut eivät paranna mallin toimintaa, vaan rajaavat tietokannasta käytettävien tietojen määrää ilman ennustavuuden parantumista. On kuitenkin epäselvää, kuinka monta vastaavuutta olisi mallin hyödyntämisen kannalta optimaalista.

Tuottoennusteen ja toteutuneen tuoton korrelaatiota käytetään optimoinnissa arvioinnin mittarina tuottoennusteelle. Arviointi voidaan kohdistaa kaikille ennusteille tai luokitella aineisto yhtiön, tappiollisuuden, voitollisuuden, kvartaalin tai jonkin muun tunnusluvun mukaan. Tuottoennuste määritellään 21 valitulle päivälle seuraavan 4,5 kuukauden aikajaksolla, joista valitaan yksi jollakin myöhemmin esitettävällä valintaperusteella. Korrelaatio lasketaan ensin kaikilla aikajaksoilla, minkä jälkeen niiden keskiarvoa käytetään vertailutietona. Korrelaatioiden arvoista voidaan tehdä havainnot ja mallin kehittämiseksi, kuten esimerkiksi, että tappiollisten päivien korrelaatio on erisuuruinen kuin yhtiön ollessa voitollinen.

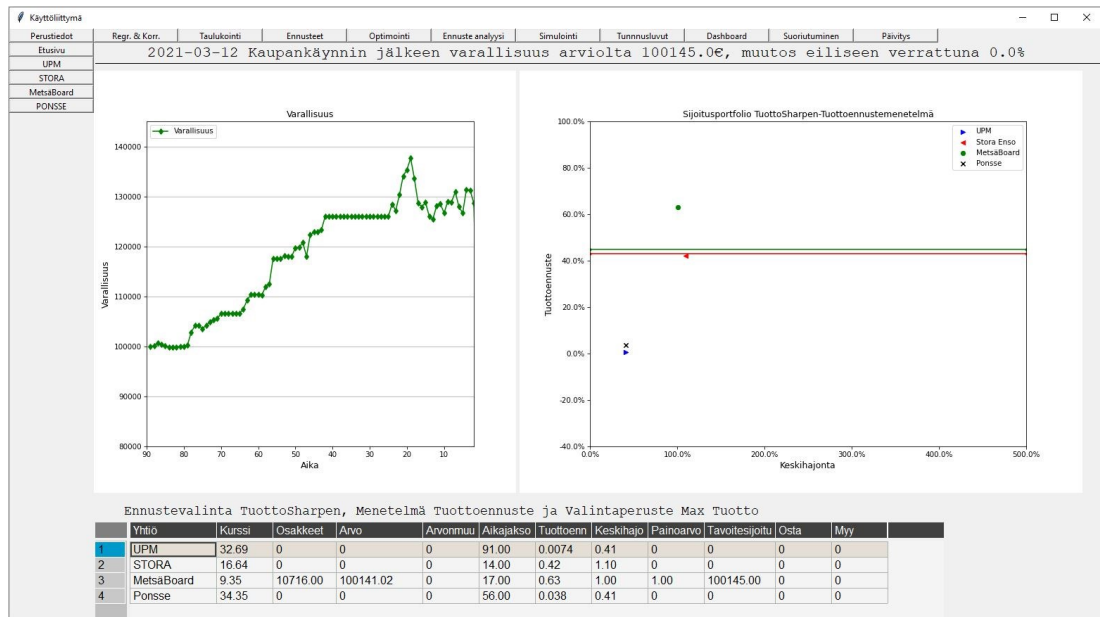
Sijoitussuunnitelmia testataan simuloimalla sijoitusvarallisuuden kehitystä. Sijoitussuunnitelmat muodostuvat erilaisista päivittäisten painoarvojen määrittelyistä yhtiöille. Ennusteaikajakson valintaperuste ja painoarvon määrittelyssä käytetty menetelmä valitaan simuloinnin tulosten perusteella. Lopputuloksen valinnassa käytetään apuna keskimääräistä toteutunutta tuottoa ja tuottojen keskihajontaa samoilla periaatteilla kuin ennusteaikajakson valinnassa. Simuloinnin jälkeen tiedetään mitä menetelmiä varsinaisessa sijoitustoiminnassa käytetään.

Kaikki päätökset perustuvat perustietoihin, joten niiden pitäisi olla oikaistuina sen suuruisia kuin sijoittajat ovat ne ymmärtäneet. Tiedon oikeellisuuden toteaminen on työllä, koska tiedot pitää tarkistaa yhtiön tiedotteista. Tietoihin voi liittyä kertaluonteisia eriä, joita pitää mahdollisesti oikaista. Epävarmoissa tilanteissa on mahdollista vertailla ennustevirhettä ilman oikaisua ja oikaisun kanssa. Sijoittajat ovat tulkinneet lukua tavalla, joka antaa pienemmän korrelaation. Tyypillisiä oikaisuja tuloslaskelmassa ovat kertaluonteiset arvonmuutokset tai uudelleenjärjestelykulut. Taseen puolella omaan pääomaan luettavat oman pääoman ehtoiset lainat voivat vääristää oman pääoman määrää. Tämä vaikuttaa oman pääoman lisäksi korollisten velkojen määrään ja näihin liittyviin tunnuslukuihin. Kirjanpitoikäntäällä voi myös olla merkittävä vaikutus raportoituihin lukuihin.

## 2.4 Sijoitusprosessi

Valittujen menetelmien käyttö yllätasolla alkaa hakemalla yhtiökohtaisesti osakekursseja verkosta. Viimeisin tilinpäätöstiedote on tiedossa, kuten on myös analyytikoiden viimeisimmät ennusteet lähitulevaisuuden kehityksen osalta. Tietojen avulla voidaan laskea tarvittavat tunnusluvut ennustamista varten. Nykytilanteen tunnistamisen jälkeen tietokannasta etsitään tavoiteltu määrä vastaavuuksia. Vastaavuuksien toteutuneiden tuottojen avulla lasketaan tuottoennuste ja keskihajonta eri aikajaksoilla sekä valitaan paras ennusteajajakso sijoittamisen kannalta. Sijoitusvarallisuus allokoidaan simuloinnin perusteella valitun sijoitussuunnitelman mukaisesti.

Dashboard:n avulla annetaan yleiskatsaus tämän hetken varallisuudesta ja sijoituskohteista. Yhtiöiden tuottoennuste, keskihajonta ja valittu aikajakso ovat kiinnostavimpia tietoja, joiden lisäksi voidaan esittää keskeisimmät tunnusluvut. Vastaavuudet ennusteiden taustalla esitetään omilla välilehdillä. Sijoitussuunnitelman mukaisesti tehtävät toimenpiteet on tärkein esitettävä tieto tulevaisuuden osalta. Toteutuneet toimenpiteet on kirjattava manuaalisesti seuraavaa kaupankäyntipäivää varten.



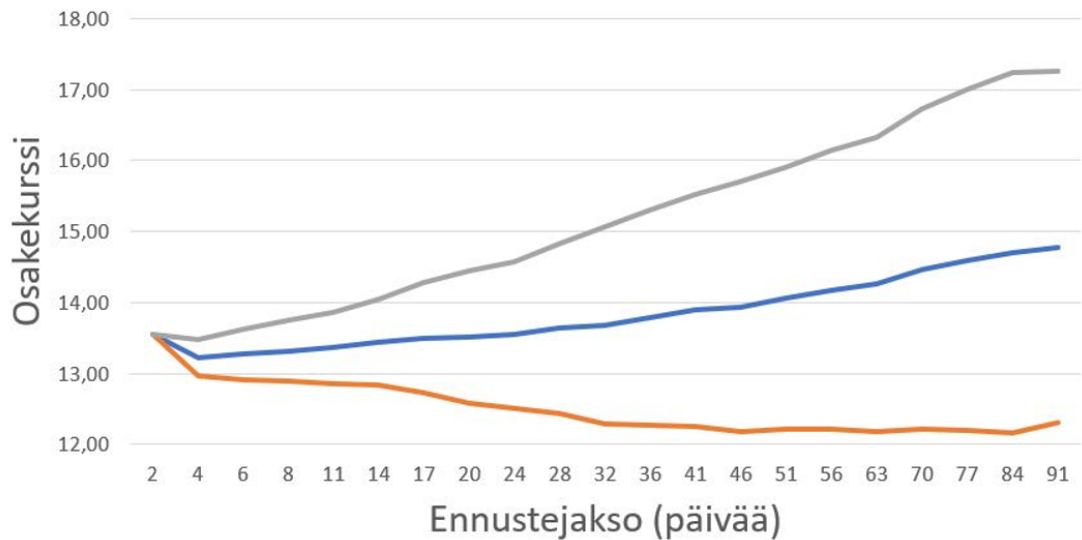
Kuvio 6. Dashboardin alustava näkymä

Vastaavuudet esitetään yhtiökohtaisesti, jolloin saadaan parempi kuva ennusteen taustalla olevista tiedoista. Esimerkiksi yhdessä analyysissä 13.4.2007 Ponssen ennusteen taustalla oli 216 päivän tiedot, joista 49 (23 %) oli Ponssen ja näistä 14 (6 %) oli samalta kuukaudelta kuin ennustettava päivä. Ponssen samalta kuukaudelta olevat tiedot vääristävät jonkin verran todellista suorituskykyä. Metsä Board:n osuus on 33 (15 %) päivää, Stora Enson osuus on 68 (31 %) päivää ja UPM-Kymmenen osuus on 66 (31 %) päivää. 216 riviä sisältävä taulukko olisi vaikea esittää, joten tiedot esitetään alla olevassa kuvassa yksinkertaisemmassa muodossa.

13.4.2007 Ponsen vastaavuudet			
Yhtiö	Vuosi	Kuukausi	Vastaavuudet
Metsä Board	2014	4	13
Metsä Board	2014	9	5
Metsä Board	2014	10	1
Metsä Board	2017	6	2
Metsä Board	2017	7	12
Ponsse	2007	4	14
Ponsse	2007	6	1
Ponsse	2015	6	7
Ponsse	2015	7	1
Ponsse	2015	9	2
Ponsse	2015	10	10
Ponsse	2017	12	14
Stora Enso	2004	9	3
Stora Enso	2004	10	5
Stora Enso	2007	6	7
Stora Enso	2007	7	8
Stora Enso	2014	12	11
Stora Enso	2015	1	5
Stora Enso	2016	6	2
Stora Enso	2016	7	4
Stora Enso	2017	9	10
Stora Enso	2017	10	4
Stora Enso	2017	12	6
Stora Enso	2018	1	3
UPM-kymmene	2016	6	4
UPM-kymmene	2016	7	6
UPM-kymmene	2016	9	10
UPM-kymmene	2016	10	3
UPM-kymmene	2017	12	9
UPM-kymmene	2018	1	4
UPM-kymmene	2018	3	4
UPM-kymmene	2018	4	2
UPM-kymmene	2018	6	1
UPM-kymmene	2018	7	8
UPM-kymmene	2018	9	12
UPM-kymmene	2018	10	3

Kuvio 7. 13.4.2007 Ponsen tuottoennusteeseen sisältyvien tietojen yhteenveto

Kaikista toteutuneiden tuottojen tiedoista laaditaan havainnollistava kuvio, jossa keskimäinen viiva kuvaa tuottoennustetta. Ylimmässä viivassa tuottoennusteeseen on lisätty keskihajonta ja alimmassa viivassa tuottoennusteesta on vähennetty keskihajonta. Suunnilleen 68 % toteutuneista tuotoista pitäisi olla ylimmän ja alimman viivan välissä, eli alle yhden keskihajonnan etäisyydellä tuottoennusteesta (Härmä, 2010).



Kuvio 8. Tuottoennuste ja keskihajonta osakekurssin muutoksena

### 3 KEHITTÄMINEN

Tiedolla johtamisella pyritään oikeaan tietoon perustuvaan automaattiseen operatiiviseen päätöksentekoon. Analytiikan avulla opitaan aineistosta ja ennustetaan tapahtumien vaikutuksia tarkasteltavan ongelman ratkaisemiseksi. Useimmiten vaikeinta on ymmärtää toimintaa riittävän hyvin ja määrittää tavoitteet. Päätöksenteko ei saa perustua arvaukseen tai mielipiteisiin, vaan datasta saatuun informaatioon. Automatiikka tehostaa prosessia ja luo uudenlaisia mahdollisuuksia. Tavoitteena on mahdollisimman tarkka ja luotettava tieto esimerkiksi sijoituksen tulevaisuuden tuotosta. (Advian www-sivut 2021)

Opinnäytetyön alussa luotiin yksinkertainen ennustemalli ja sijoitussuunnitelma, joita kehitettiin ensin vaistojen varassa. Manuaalisia vaiheita oli käytössä useita. Ajatus tuottoennusteen ja keskihajonnan hyödyntämisestä oli vähän saman suuntainen kuin portfolioteoriassa, joten näin opinnäytetyöhön löydettiin teoreettinen lähtöpiste (Markowitz 1952). Portfolioteorian pohjalta kehitettyjä menetelmiä, kuten riskinottohalukkuutta ja hyötyfunktiota käytetään mukautetusti valintapäätöskien teossa ja varojen allokoinnin määrittelyssä (Francis & Kim 2013).

Ennustemallin lopputuloksena saadaan kolmella eri valintaperusteella valittu tuottoennuste ja keskihajonta. Sijoitussuunnitelmien simuloinnin lopputulosten avulla saadaan määriteltyä yhtiöihin sijoitettava summa. Lopputulokset ovat mallin kehityksen pääkohteet, jotka vaikuttavat sijoitusvarallisuudelle saatavaan tuottoon. Tuottoennusteeseen ja keskihajontaan vaikuttavat perustietoihin tallennettujen tietojen laatu, tunnuslukujen laskennassa käytetyt kaavat, käytetyt tunnusluvut ja vastaavuuksien tunnistaminen. Simuloinnin lopputuloksiin vaikuttavat tuottoennusteen valintaperuste ja painoarvon määrittelyn tapa.

Vastaavuuksien tunnistaminen tehdään asettamalla raja-arvot, joiden sisällä tunnusluvun pitää olla, koska täsmälleen samanlaisia tilanteita ei ole juuri koskaan. Vastaavuuksia pitää olla riittävän paljon niin, että yksittäiset vastaavuudet eivät vaikuta liikaa tuottoennusteen lopputulokseen. Sijoitussummaan vaikuttaa edellisten lisäksi käytettävä sijoitussuunnitelma. Tällä hetkellä mukana on vain neljä yhtiötä ja kehitystyö on alussa, joten sijoitussuunnitelman osalta otetaan edelleen askeleita realismia kohti. Sijoitettava summa määritellään yhtiökohtaisten painoarvojen avulla. Tietokoneen laskentatehoa hyödynnetään raja-arvojen ja sijoitussuunnitelman painoarvojen määrittelyssä.

### 3.1 Toteutustapa ja lähestyminen

Tutkimus ja kehitys liittyvät usein yhteen, kun ratkaistaan esiin tulleita ongelmia tai luodaan uutta tietoa. Ratkaisuja etsitään yleisesti hyväksytyillä menetelmillä. Tutkimuksellisessa kehittämisessä voidaan käyttää monipuolisesti erilaisia menetelmiä. Soveltavassa tutkimuksessa tulosten avulla voidaan luoda parempia menetelmiä. Kehittämistyössä asioiden kuvailun ja selittämisen lisäksi asioita muutetaan paremmiksi. Raportoinnissa kuvataan lähtökohdat ja tavoitteet, työmuodot ja prosessin eteneminen sekä lopputulos. (Ojasalo ym. 2014, 18-22)

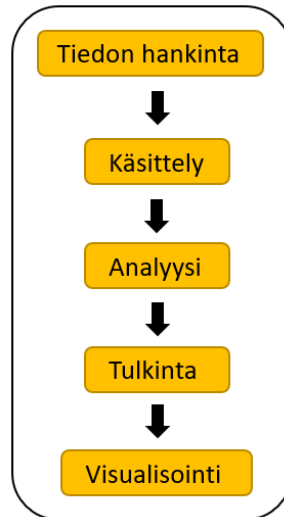
Kehittämistyö voi toisaalta olla myös prosessimaista ja ennakoimatonta, jossa monien ideoiden kehitysvaiheiden jälkeen saavutetaan tavoiteltu lopputulos. Ajureina toimivat käytännölliset tavoitteet, joiden ratkaisemisessa hyödynnetään teoriaa. Ongelmat, joita lähdetään kehittämään pitää havaita ennen kuin niihin voidaan löytää ratkaisu.

Tutkimuksen avulla toimenpiteet ovat paremmin perusteltuja, kun ne otetaan kattavammin ja suunnitelmallisemmin huomioon. Tutkimuksellisen kehittämisen pitää edetä järjestelmällisesti ja tuloksiin pitää suhtautua kriittisesti. (Ojasalo ym. 2014, 18-22)

Aihe on laaja ja hankala kokonaisuus, joten yksittäisiin aiheisiin syvennyttäessä muita asioita ei luonnollisestikaan pystytä kehittämään. Käytännön teknisiä ongelmia on esiintynyt erittäin usein muutoksia tehtäessä. Teoria ei usein vastaa suoraan ratkaistavaan ongelmaan, vaan lupaavia tai kiinnostavia menetelmiä on testattava ennen lopullista käyttöönottoa.

Opinnäytetyössä käsitellään enimmäkseen tämän hetken tilannetta, vaikka useimmat asiat ovat jo ottaneet useampia kehitysaskelaita ja jotkin asiat tulevat muuttumaan kehitystyön jatkuessa. Opinnäytetyön kirjoittaminen on myös osaltaan auttanut paremmin hahmottamaan kokonaisuutta ja löytämään kehityskohteita. Muihin tutkimuksiin ja kirjallisuuteen tutustumalla on usein pyritty löytämään ratkaisu johonkin ongelmaan.

Tekoälyn avulla yritetään toimia järkevällä tavalla tehtävän ja tilanteen mukaisesti. Tekoälyyn kuuluu useita erilaisia osaamisalueita, kuten muun muassa data-analyysi, koneoppiminen, havainnointi ja tilannetietoisuus. Data-analyysiin voidaan katsoa kuuluvan tiedonhankinta, käsittely, analyysi, tulkinta ja visualisointi. Toimintaympäristön havainnoinnin avulla tunnistetaan vallitseva tilanne, jonka avulla luodaan ennuste tulevasta ja suoritetaan tarpeelliset toimenpiteet. Koneoppimisen avulla oppiminen tapahtuu tilastotieteen menetelmiä käyttäen, mikä parantaa suorituskykyä. (Valtioneuvoston www-sivut 2021)



Kuvio 9. Data-analyysi

Tapaustutkimus sopii hyvin menetelmäksi etsittäessä kehitysideoita. Ilmiöstä saadaan syvällistä ja yksityiskohtaista tietoa todellisessa tilanteessa ja toimintaympäristössä. Tapaustutkimus sopii heikosti ymmärrettyjen tilanteiden tai käyttäytymisen tutkimiseen. Sitä käytetään erottamalla fyysinen yksiköiden joukko omaksi tutkittavaksi kohteeksi. Aiheeseen on usein perehdyttävä ennen kuin tiedetään mikä todellinen kehittämistehtävä on, mutta siitä huolimatta kehittämistehtävää voidaan muuttaa prosessin edetessä. (Ojasalo ym. 2014, 52-57)

Toimintatutkimuksessa ratkaistaan käytännön ongelmia ja toteutetaan muutosta. Ongelmat voivat olla teknisiä tai ammatillisia. Ratkaisuilla luodaan samalla uutta tietoa ilmiöstä. Tutkimuksen kohteena on toimintatavat ja toimintatilanne. Toimintayhteisön jäsenten asiantuntemuksen ja tutkijan ulkopuolisen näkökulman sekä teoreettisen osaamisen avulla ongelma on helpompi ratkaista. Prosessissa toistuvat suunnittelu, toiminta ja arviointi, joita toistetaan riittävän monta kertaa. (Ojasalo ym. 2014, 58-62)

Menetelmiä voidaan kehittää monilla erilaisilla tavoilla tai näiden yhdistelmillä. Hyviä kehityskohteita tai kehitysideoita tulee usein eteen kriittisen arvioinnin lopputuloksena. Tärkeintä on muistaa tavoite, jota kohti kehitystyössä pyritään ja priorisoida asioita. Eteen tulee usein ongelmia, jotka olisi ollut mahdoton ennakoida. Osakemarkkinoiden ennustamista koskevat tutkimukset, teorit ja muut julkaisut ovat usein ristiriitaisia, joten pelkästään niiden perusteella on vaikea saavuttaa tyydyttävää ja yhtenäistä

lopputulosta. Tutkimalla kehitysideoita vaikutusta osana käytössä olevia menetelmiä, lopuksi on mahdollista saavuttaa tyydyttävä lopputulos.

### 3.2 Tutkimusmenetelmät

Havainnoinnin avulla mallia voidaan kehittää paremmaksi ja huomioida kaikki tarpeellinen tieto. Kaikkia muutoksien seurauksia on vaikea ottaa huomioon heti ensi yrittämällä, joten usein muutokseen liittyy myös muita toteutettavia muutoksia. Käytettävissä oleva tiedon määrä rajoittaa osaltaan tilanteen määrittelyn tarkkuutta, jolloin sillä on suora vaikutus käytettyihin raja-arvoihin. Finanssikriisin tai koronaviruksen kaltaiset poikkeukselliset ilmiöt ovat harvinaisia ja vastaavia ajanjaksoja esiintyy harvoin, joten vastaavien muutosten toistuessa ennusteiden epävarmuus kasvaa huomattavasti. Normaalimmalta ajalta tietoa on käytettävissä huomattavasti enemmän ja liiketoiminta on huomattavasti ennustettavampaa.

Dokumenttianalyyseissä päätelmiä tehdään yleensä kirjallisesta aineistosta, kuten uutisista, osavuositarkastuksista tai vuosikertomuksesta. Dokumenttien avulla voidaan tunnistaa trendejä ja toistuvia tapahtumia. Sisällön erittelyn ja analyysin tavoitteena on tunnistaa tekstin merkityksiä numeroin. Kriittisen tarkastelun avulla ongelmallinen tieto voidaan tunnistaa ja tarvittaessa korjata. (Ojasalo ym. 2014, 136-138)

Tärkeimmät dokumentit ovat yhtiön itse julkaisemat tiedotteet. Tiedotteiden tiedot kirjataan oikaisuineen yhtiöiden perustietoihin. Tuloslaskelma saattaa sisältää kertaluonteisia eriä tai taseessa olevat osakkeet saattavat olla arvostettu hankintahintaan käyvän arvon sijasta. Vaikutukset voivat ajoittain olla hyvinkin suuria, kuten esimerkiksi UPM-Kymmene pörssitiedotteessa ilmoitettu vuoden 2012 viimeiselle vuosineljännekselle kirjattu 1 770 miljoonan arvonalentuminen (UPM-Kymmene www-sivut 2021). Vuoden 2011 tulos oli esimerkiksi 457 miljoonaa, joten kirjauksella oli merkittävä vaikutus. Tuloslaskelman kertaluonteisten erien määrä on usein ilmoitettu muualla tiedotteessa. Taseessa olevien varojen arvonnäytys on usein hankalaa, eikä varojen käypä arvo useimmiten ilmene tiedotteesta. Lisätietoja omaisuuseristä on saatavilla liitetiedoissa, mutta arvonnäytys voi edelleen olla hankalaa. Laadintaperiaatteet ja yleiset kirjanpitoikäntännöt auttavat ymmärtämään dokumentteja paremmin

ja niiden muutoksilla voi olla myös merkittävä vaikutus raportoituihin lukuihin. IFRS 16-standardin vaikutuksesta esimerkiksi Finnairin taseen loppusumma kasvoi kolmesta miljardista neljään miljardiin (Vähähyppä 2019). Väärän tiedon kirjaaminen vaikuttaa kaikkiin myöhempiin vaiheisiin.

Skenaariotyöskentely on yksi ennakoinnissa käytettävistä tyypillisistä menetelmistä. Skenaarion avulla voidaan kuvata tulevaisuuspolku, joka voi edelleen haarautua erilaisiin lopputuloksiin. Skenaarioita laaditaan useampia vain positiivisen ja negatiivisen skenaarion sijaan. (Ojasalo ym. 2014, 90-94, 146-147)

Mallin avulla yritetään ennakoida toteutuvaa tuottoa, mikä tarkoittaa suurimmaksi osaksi osakekurssin kehityksen ennakoitua. Ennakointiin käytetään menetelmiä, joiden avulla pyritään mahdollisimman hyvään lopputulokseen. Tuottoennuste syntyy historiallisen aineiston pohjalta, josta poimitaan yhtiön tilannetta vastaavaa tietoa. Jokainen vastaavuus luo oman skenaarion osakekurssin tulevasta kehityksestä. Simuloidulla sijoitusvarallisuuden kehitystä erilaisia valintoja tekemällä syntyy myös erilaisia skenaarioita, joiden kaltaisesti tuottojen oletetaan myös tulevaisuudessa toteutuvan.

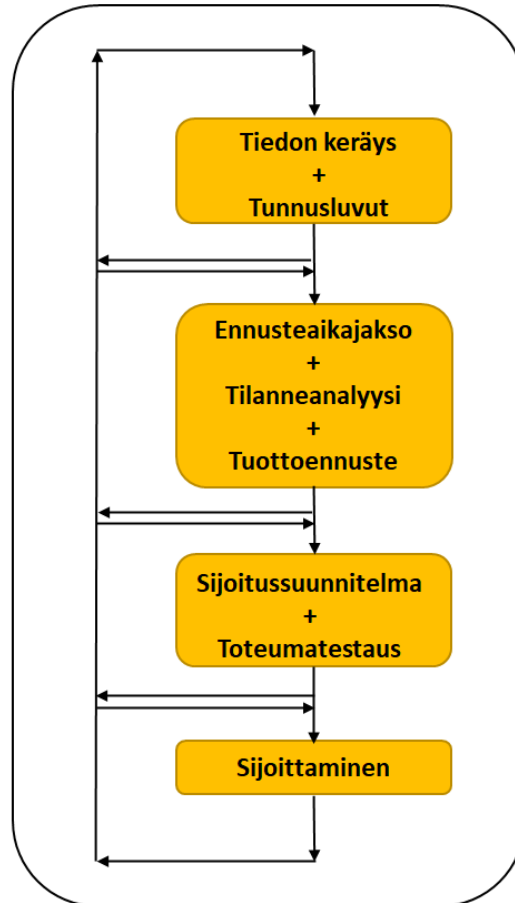
Käytettävät menetelmät ovat kvantitatiivisia, mutta erilaisia vaihtoehtoja ja sijoitusympäristöä tutkitaan kvalitatiivisia menetelmiä hyödyntämällä. Erilaisista sijoittamisen lähestymistavoista löytyy runsaasti tietoa eri lähteistä. Yleisiä lähestymistapoja esimerkiksi arvo- ja kasvusijoittamisesta on otettu tai otetaan tulevaisuudessa osaksi menetelmiä.

### 3.3 Jatkuva parantaminen

Tuottavuuteen vaikuttavat asiat on tunnistettava ja tunnettava niiden nykytila. Päämääränä on tavoite, jota kohti edetään pienin askelin. Nykytilan tunnistamisen jälkeen yritetään tunnistaa erilaisia vaihtoehtoisia lähestymistapoja. Tiedettäessä mitä halutaan saavuttaa ja miten tavoite saavutetaan, on vain toteutettava toimenpiteet. Jälkitoimenpiteenä arvioidaan, millainen lopputulos saavutettiin verrattuna lähtötilanteeseen. (Larikka, Heinilä, Selin & Tuominen 2007, 8-29)

Ongelmana saattaa olla puutteet työkaluissa, tiedoissa, taidoissa tai tahdossa. Puutteita on yritettävä poistaa mahdollisuuksien mukaan. Havainnoimalla kerätään tietoa tapahtumista ja muutosten vaikutusta selvitetään analysoimalla tilannetta. On tarpeellista tuntea havaintojen perussyyt, jolloin tarvittavat muutokset voidaan kohdistaa näihin. Arvoa lisäävää työtä on sellainen, jonka seurauksena kokonaisuus muuttuu paremmaksi. Yksittäiset muutokset ovat yleensä suhteellisen pieniä, mutta niiden yhteismäärä saattaa olla suuri. Ideoita tuotetaan ja jalostetaan, minkä jälkeen niiden toimivuutta kokeillaan ennen lopullista käyttöönottoa. (Larikka, Heinilä, Selin & Tuominen 2007, 113-206)

Mallin kehitysprosessi on jaettu alapuolella olevan kuvion mukaisesti osakokonaisuuksiin, joita voidaan kehittää koskematta muihin osakokonaisuuksiin. Suurin osa ajasta kuluu keskimmäisten osa-alueiden kehittämiseen, mutta uusia yhtiöitä lisättäessä aika kohdistuu tiedon keräämiseen ja tallentamiseen. Uusien tietojen ja taitojen opettelu on työlästä, eikä valmista ja helposti ymmärrettävää ratkaisua useinkaan ole saatavilla. Lisäksi yksi tärkeä kehityskohde on turhan laskennan vähentäminen, jolloin säästyy aikaa. Sijoittaminen viimeisenä osakokonaisuutena jätetään pääasiassa opinäytetyön ulkopuolelle, koska aikaisemmat vaiheet on ensin saatava riittävän hyvälle tasolle. Suurin osa oikealla rahalla sijoittamisen vaiheista on samoja kuin aikaisemmissä vaiheissa, jolloin suuria muutoksia ei ole tarpeen tehdä. Uudet tunnusluvut lasketaan ainoastaan yhdelle päivälle, joten laskenta on suhteellisen nopeaa.



Kuvio 10. Kehitysprosessi

Ensimmäisessä osakokonaisuudessa keskitytään tiedon keräämiseen, laatuun ja tunnuslukujen laskentaan. Mallin toiminta perustuu näiden tietojen varaan, joten on tärkeää, että tieto on riittävän laadukasta. Tietoa saadaan yhtiöiden julkaisemista tiedotteista, Nasdaqomxnordic.com verkkosivulta ja Yahoo Finance palvelusta. Ennustetietoja on saatavilla esimerkiksi Inderes.fi verkkosivuilta tai pankkien analytiikkopalveluista. Tunnuslukujen on oltava laskettuna oikeilla kaavoilla ja tilinpäätöstiетоjen tulee olla kohdistettuna oikeille päiville. Tulevaisuudessa julkaistavia tietoja ennakoidaan alustavasti seuraavan kvartaalin tietoihin saakka. Suurin kehityksen kohde on uusien tunnuslukujen käyttöönotossa ja uusien yhtiöiden lisäämisessä.

Toisessa osakokonaisuudessa lasketaan tuottoennuste ja keskihajonta eri aikajaksoille, joista valitaan yksi teoriaan perustuvalla valintaperusteella. Lopullinen valintaperusteen valinta tehdään sijoitusvarallisuuden kehityksen simulointien perusteella. Tavoitteena on mahdollisimman suuri tuottoennusteen ja toteutuneen tuoton korrelaatio.

Tuottoennusteen keskihajonnan osalta tavoitellaan mahdollisimman pientä arvoa, jolloin myös toteutuneen tuoton pitäisi olla tarkemmin tiedossa. Yhtiön tilanteen mukaan tuottoennustemenetelmä jaetaan voitollisiin ja tappiollisiin tilanteisiin. Suurin kehityskohde on käytettävien tunnuslukujen ja raja-arvojen valinta tietoon perustuen. Haasteena on erilaisten kombinaatioiden määrä. Päivittäisille tunnusluvuille määritellään raja-arvot vastaavuuksien etsimiseen tietokannasta. Raja-arvoille määritellään myös muutosnopeus tilanteisiin, joissa vastaavuuksia ei löydetä riittävästi. Raja-arvoja voidaan kasvattaa jokaisella kierroksella esimerkiksi 20 % tai muu tietty määrä, minkä jälkeen vastaavuuksia etsitään uudestaan. Vastaavuuksien määrä vaikuttaa ennustusten lopputulokseen, jota varten on yritettävä löytää optimaalinen vastaavuuksien määrä.

Kolmannessa osakokonaisuudessa suunnitellaan ja testataan erilaisia sijoitussuunnitelmia. Sijoituksen painoarvo määräytyy tuottoennusteen ja keskihajonnan perusteella ja se on tärkein tässä vaiheessa määriteltävä asia. Hyväkään ennuste ei johda hyvään lopputulokseen, jos painoarvo sijoitukselle määritellään huonosti. Yksittäisen yhtiön vaikutusta sijoitusvarallisuuteen voidaan rajoittaa asettamalla raja sijoittavalle summalle. Vaikutus riippuu toteutustavasta ja se voi olla joko positiivinen tai negatiivinen. Sijoitusaste kertoo, kuinka suuri osa sijoitusvarallisuudesta on sijoitettuna. Ennusteiden ollessa keskimääräisiä huonompia voidaan sijoittaa pienempi osa sijoitusvarallisuudesta. Tulevaisuuden näyttäessä erityisen hyvältä voisi olla järkevää hyödyntää tilanne ottamalla sijoituslainaa. Yhtiöiden määrän lisääntyessä, mahdollisuudet sijoitusportfolion muodostuksessakin lisääntyvät, kun sijoituskohteiksi voidaan valita vain kaikista houkuttavimmat sijoituskohteet.

Neljännessä osakokonaisuudessa yhdistetään kolme aikaisempaa osakokonaisuutta parhaat menetelmät valitsemalla. Sijoitettavasta varallisuudesta ja omistuksista on pidettävä erikseen kirjaa, koska ne vaikuttavat päätöksentekoon. Ensin haetaan uusimmat tiedot ja lasketaan ajantasaiset tunnusluvut. Toiseksi määritellään ja valitaan parhaalla tiedolla olevalla tavalla tuottoennuste ja keskihajonta. Kolmanneksi lasketaan parhaan sijoitussuunnitelman mukaiset sijoitukset. Sijoitukset voidaan toteuttaa manuaalisesti osakevälittäjän palvelussa tai toimeksiantojen suorittaminen voidaan myös automatisoida. Joihinkin osakevälittäjien palveluihin on mahdollista olla yhteydessä ohjelmointirajapinnan kautta (Interactive Brokers www-sivut 2021).

### 3.4 Automaatio

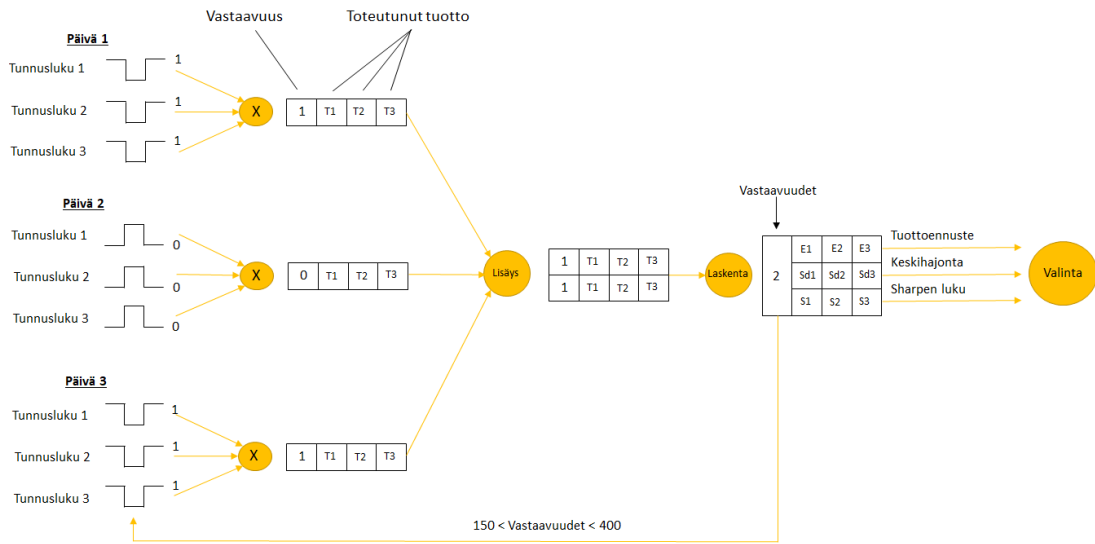
Yksinkertaisia tehtäviä, kuten tiedon hakua verkosta tai tiedon siirtoa tiedostosta toiseen voidaan sanoa ohjelmistorobotiikaksi. Tiedonhaku ja tietojen siirto toteutetaan ohjelmoinnin avulla ilman varsinaista ohjelmistorobotiikan ohjelmistoa. Sen avulla voidaan poistaa manuaalisesti tehtävää työtä ja käyttää aika muuhun hyödyllisempään työhön. Virheitä tietojen kopioinnissa syntyy vähemmän kuin manuaalisesti kirjaimella, mutta harkinta prosessissa on vähäistä. (Kaarlejärvi & Salminen 2018, 54-56)

Älykkäämpää automaatiota on koneoppiminen, jossa laajan aineiston käsittelyn ja luokittelun avulla luodaan ennuste matemaattisen mallin perusteella. Edistynyt ratkaisu pystyy kehittämään itseään paremmiksi. ”Treidausrobotit” suorittavat esimerkiksi osto- ja myyntitoimeksiantoja historiallisen datan perusteella. Tällainen alkeellinen tekoäly vaatii riittävästi dataa toimiakseen luotettavasti. Ohjelman käyttäjällä pitää olla riittävästi ymmärrystä sijoituspäätöksen muodostamisesta. (Kaarlejärvi & Salminen 2018, 60–61)

Koneoppiminen voidaan jakaa oppimisen näkökulmasta ohjattuun ja ohjaamattomaan oppimiseen. Ohjatussa oppimisessa ennustetaan esimerkiksi osakkeen tulevaa tuotto prosenttia, kun ohjaamattomassa oppimisessä lopputuloksena on joko nousu, lasku tai ei muutosta (Trask 2019, 10-13). Oppiminen on toteutettu ohjatusti ja läpinäkyvästi tallentamalla oppimisprosessin tiedot omaan tiedostoon, mikä auttaa paremmin ymmärtämään mallin toimintaa ja mahdollisesti edelleen kehittämään sitä. Tallentamalla myös vastaavuudet omaan tiedostoon voidaan säästää huomattavasti aikaa, kun seuraavalla kerralla vastaavuuksia voidaan etsiä ensin samoilla raja-arvoilla. Viimeisessä vaiheesta voidaan edelleen halutessa palata kaikkiin niihin tietoihin, joihin ennusteet perustuvat.

Neuroverkkojen avulla voidaan toteuttaa ennustaminen, vertailu ja oppiminen. Tieto kulkee alla olevissa kuvioissa vasemmalta oikealle. Jokainen päivä tietokannassa, jonka toteutunut tuotto ja tunnusluvut tunnetaan, voi vaikuttaa tuottoennusteeseen. Jokaisen päivän osalta tunnusluvut menevät aktivointi funktion läpi, jossa tunnusluvun ollessa raja-arvojen sisällä se saa arvon yksi tai muussa tapauksessa arvo on nolla. Kaikkien päivän tunnuslukujen ollessa raja-arvojen sisäpuolella, sen toteutuneita

tuottoja käytetään osana ennustamista. Käytettävät päivät kootaan yhteen, minkä jälkeen lasketaan vastaavuuksien lukumäärä, jonka pitää olla määrättyssä välissä. Kaikkien ehtojen täytyessä lasketaan tuottoennusteet ja keskihajonnat valintaa ja sijoittamista varten. Oppimisessa lasketaan edelleen tuottoennusteiden ja toteutuneiden tuottojen korrelaatio, jonka jälkeen raja-arvoja tai muutosnopeuksia muutetaan.



Kuvio 11. Ennustamisen neuroverkko

Toteutuneita tuottoja on neljän yhtiön osalta käytössä n. 16 000 päivältä. Tunnuslukuja vastaava määrä raja-arvoja ja muutosnopeuksia on määriteltävä olennaisten tuottojen keräämiseksi ennusteita varten. Raja-arvot määrytyvät osittain tunnuslukujen lukuarvon mukaan ja löydettyjen vastaavuuksien perusteella. Keskimäärin vastaavuuksia löytyy yli 200, joiden avulla lasketaan ennustetiedot 21 päivälle. Automaattisesti toistuva prosessi on välttämättömyys, suuren tietomäärän käsittelyyn.

Oppimisprosessissa tuottoennusteiden ja toteutuneiden tuottojen korrelaatio lasketaan, jonka jälkeen raja-arvoja tai niiden muutosnopeuksia muutetaan suuremmiksi tai pienemmiksi. Muutoksen jälkeen tuottoennuste, keskihajonta ja korrelaatio toteutuneen tuoton kanssa lasketaan uudelleen. Korrelaation muuttuessa suuremmaksi uudet raja-arvot tai muutosnopeudet otetaan käyttöön ja ennustaminen kehittyy paremmaksi askel askeleelta. (Trask 2019) Toistamalla automaattista prosessia taustalla voidaan oma aika käyttää muuhun toimintaan.

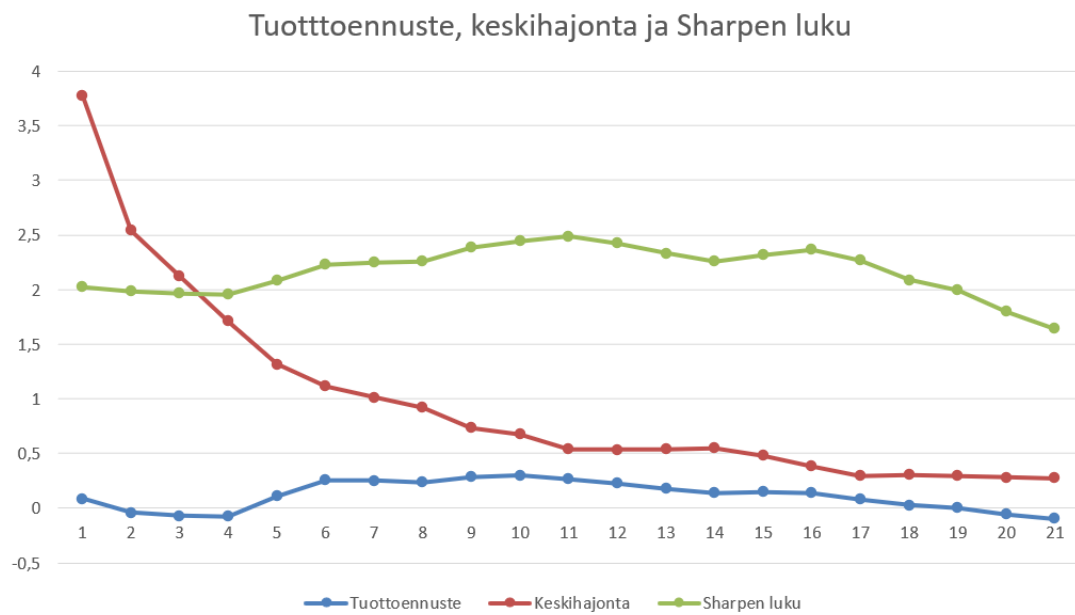
## 4 ENNUSTAMINEN JA SIMULOINTI

### 4.1 Aikajakso

Aluksi ajatuksena oli ennusteajakajakso, joka olisi 1-2 kvartaalin (60-120 kaupankäyntipäivän) pituinen, mutta kuitenkin rationaalisesti perusteltavissa. Ennusteajakajakso oli aluksi n. 90 päivää, mutta paremmin perusteltavissa oleva ennusteajakajakso haluttiin löytää. Ennustevirhe eli tuottoennusteen ja toteutuneen tuoton erotus oli jakautunut epätasaisesti kvartaalin kulumisen mukaan. Ratkaisuna ennustehorisontiksi valittiin sopivan pituinen ajanjakso aina viiden kvartaalipäivän välein, alkaen tilinpäätöstietojen julkaisusta. Ennustehorisontiksi valittiin aikajakso, joka antoi korkeimman korrelaation tuottoennusteen ja toteutuneen tuoton välillä. Keskimääräinen ennustehorisontti oli 34 kaupankäyntipäivän pituinen. Ennusteajakajakson poukkoileminen aiheutti kuitenkin tyytymättömyyttä, koska myös tuottoennuste poukkoili samalla.

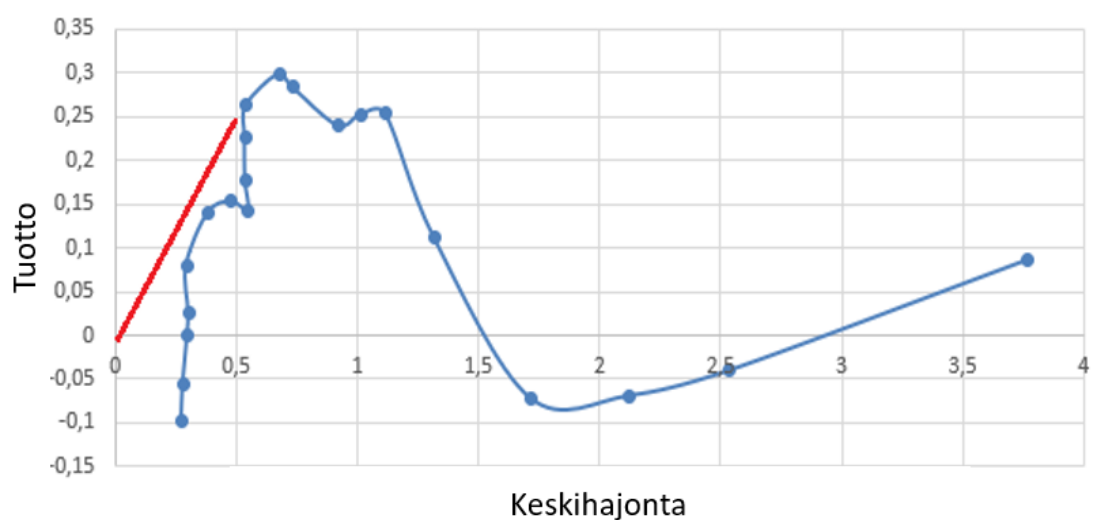
Parempi tutustuminen portfolioteoriasta edelleen kehitettyyn CAPM-malliin ja Sharpen lukuun johti vaihtoehtoisiin menetelmiin, joissa ennustejakso voidaan valita suurimman tuottoennusteen, pienimmän keskihajonnan tai näiden avulla lasketun suurimman Sharpen luvun perusteella. Aivan jokaiselle tulevalle päivälle ennustetta ei tässä vaiheessa tehdä vaan valituille päiville, joita on 21 seuraavan 4,5 kuukauden aikajaksoilla. Ennusteiden määrällä ja aikajakson pituudelle ei ole enimmäismäärää, vaan ennuste voidaan laatia niin pitkälle aikajaksolle kuin halutaan. Skaalaamalla luvut vuoden aikajaksolle, niistä saadaan vertailukelpoisia. Eri aikajaksojen tuottoennusteet muutetaan vuosittaiseksi tuotoksi kertomalla esimerkiksi kuukauden päähän sijoittuva tuottoennuste 12:sta ja vastaavasti kahden kuukauden päähän sijoittuva tuottoennuste 6:lla.

Alla olevassa kuviossa havainnollistetaan yhdelle päivälle laskettuja ennusteita. Sharpen luvun pisteet on nostettu kaaviossa kaksi yksikköä ylemmäs selkeämmän kokonaiskuvan saamiseksi. Suurin tuottoennuste on sinisen viivan ylin piste 10. havaintopisteen kohdalla. Pienin keskihajonta on viimeisen 21. havaintopisteen kohdalla ja suurin Sharpen luku on 11 havaintopisteen kohdalla. Riippuen valitusta ennustejakson valintamenetelmästä, valinta kohdistuu yhteen näistä vaihtoehdoista.



Kuvio 12. Tuottoennuste, keskihajonta ja Sharpen luku viivakaaviossa

Sama tieto voidaan esittää myös pistekaaviossa, jossa pystyakselilla on tuottoennuste ja vaak akselilla keskihajonta. Ylin havaintopiste kuvaa suurinta tuottoennustetta ja vasemmanpuolisin piste kuvaa pienintä keskihajontaa. Punaisen viivan avulla on havainnollistettu suurimman Sharpen luvun antavaa pistettä. Tässä pisteessä tuottoennusteen suhde keskihajontaan on suurin. Edellisen kuvion ensimmäinen havaintopiste kuviossa on oikeanpuoleisin. Toinen havaintopiste löytyy seuraamalla sinistä viivaa seuraavaan pisteeseen.



Kuvio 13. Tuottoennuste, keskihajonta ja Sharpen luku pistekaaviossa

Yksi mahdollinen vaihtoehto olisi voinut olla Triple-Barrier-menetelmä. Menetelmässä valitulla aikajaksolla seurataan saavuttaako tuotto ensin ylärajan, alarajan vai kuluuko valittu aikajakso ennen rajojen saavuttamista. Menetelmä on kiinnostavampi kuin kiinteä aikajakso, koska se huomioi aikajaksolla toteutuneen tuoton, eikä vain aikajakson lopussa toteutunutta tuottoa. (Prado 2019)

#### 4.2 Vastaavuudet ja ennustaminen

Havainnollistetaan ennusteen määräytymistä vielä yksinkertaistetun esimerkin avulla. Tietokannasta löytyvät tiedot, jotka ovat esitetty alla olevassa kuviossa. Tietojen avulla ennustetaan Osakkeiden A ja B tuotto ja määritellään osakkeisiin sijoitettava varallisuus. Ennuste laaditaan kahden tunnusluvun perusteella. Osakkeen A P/B = 1,25 ja P/E = 9.9. Osakkeen B P/B = 2,0 ja P/E 11,0.

Nro.	P/B	P/E	...
1	1,00	10,00	...
2	1,10	10,10	...
3	1,20	10,20	...
4	1,30	10,30	...
5	1,40	10,40	...
6	1,50	10,50	...
7	1,60	10,60	...
8	1,70	10,70	...
9	1,80	10,80	...
10	1,90	10,90	...
11	2,00	11,00	...
12	2,10	11,10	...
13	2,20	11,20	...
14	2,30	11,30	...
15	2,40	11,40	...
16	2,50	11,50	...
17	2,60	11,60	...
18	2,70	11,70	...
19	2,80	11,80	...
20	2,90	11,90	...
21	3,00	12,00	...
...	...	...	...

<b>P/B Toleranssi</b>	25,0 %
<b>P/E Toleranssi</b>	5,5 %

	A	B
<b>P/B</b>	1,25	2,20
<b>P/E</b>	9,90	11,50
<b>P/B alaraja</b>	0,94	1,65
<b>P/B yläraja</b>	1,56	2,75
<b>P/E alaraja</b>	9,36	10,87
<b>P/E yläraja</b>	10,44	12,13

Kuvio 14. Tietokanta ja raja-arvot

Oletetaan, että koneoppimisen perusteella pienin ennustevirhe toteutuneen tuoton ja tuottoennusteen välillä saataisiin, kun P/B-tunnusluvulle annetaan ala- ja ylärajaksi 25 % tunnusluvusta poikkeava luku ja P/E-tunnusluvulle 5,5 %. Molempien tunnuslukujen pitää olla määritellyissä rajoissa, jotka ovat laskettuna kuviossa oikealla. Osakkeen A vastaavuuksia ovat numerot 1-5 ja osakkeen B vastaavuudet ovat numerot 10 - 18. Jokaisella vastaavuudella on yhtä suuri todennäköisyys toteutua ja kaikkien tietokannassa olevien tietojen toteutunut tuotto tunnetaan. Tuottoennuste on vastaavuuksien toteutuneiden tuottojen keskiarvo ja keskihajonta on vastaavuuksien toteutuneiden tuottojen keskihajonta.

Osake A			Tuotto, pv					
Nro.	P/B	P/E	2	5	10	15	20	...
1	1,00	10,00	-3,325	1,088	0,333	0,141	0,317	...
2	1,10	10,10	1,915	-0,104	-1,561	-1,505	-0,896	...
3	1,20	10,20	2,646	-1,432	-0,691	0,154	0,841	...
4	1,30	10,30	0,000	-0,298	-0,119	-0,159	-0,238	...
5	1,40	10,40	-6,484	-2,051	-1,357	-1,669	-2,262	...
Keskiarvo			-1,050	-0,559	-0,679	-0,607	-0,448	...
Kulut			0,251	0,100	0,050	0,033	0,025	...
Tuottoennuste			-1,301	-0,660	-0,729	-0,641	-0,473	...
Keskihajonta			3,084	0,946	0,629	0,783	0,905	...
Sharpen luku			-0,422	-0,698	-1,160	-0,818	-0,522	...

Osake B			Tuotto, pv					
Nro.	P/B	P/E	2	5	10	15	20	...
10	1,90	10,90	-4,593	-2,012	-0,219	-0,146	0,044	...
11	2,00	11,00	3,808	-0,087	0,174	-0,290	-0,315	...
12	2,10	11,10	-3,162	5,827	3,930	2,288	1,965	...
13	2,20	11,20	-2,694	-1,149	0,198	0,886	1,347	...
14	2,30	11,30	1,965	1,352	0,755	1,195	0,810	...
15	2,40	11,40	2,061	2,473	1,379	0,502	0,995	...
16	2,50	11,50	-1,667	-0,480	-3,787	-3,004	-1,313	...
17	2,60	11,60	-2,993	-1,719	-1,397	0,184	-0,437	...
18	2,70	11,70	3,345	2,458	2,054	0,270	-1,416	...
Keskiarvo			-0,437	0,740	0,343	0,209	0,186	...
Kulut			0,251	0,100	0,050	0,033	0,025	...
Tuottoennuste			-0,688	0,640	0,293	0,176	0,161	...
Keskihajonta			2,872	2,033	1,499	0,910	0,971	...
Sharpen luku			-0,239	0,315	0,195	0,193	0,166	...

## Kuvio 15. Vastaavuuksien perusteella laskettu tuottoennusteen ja keskihajonta

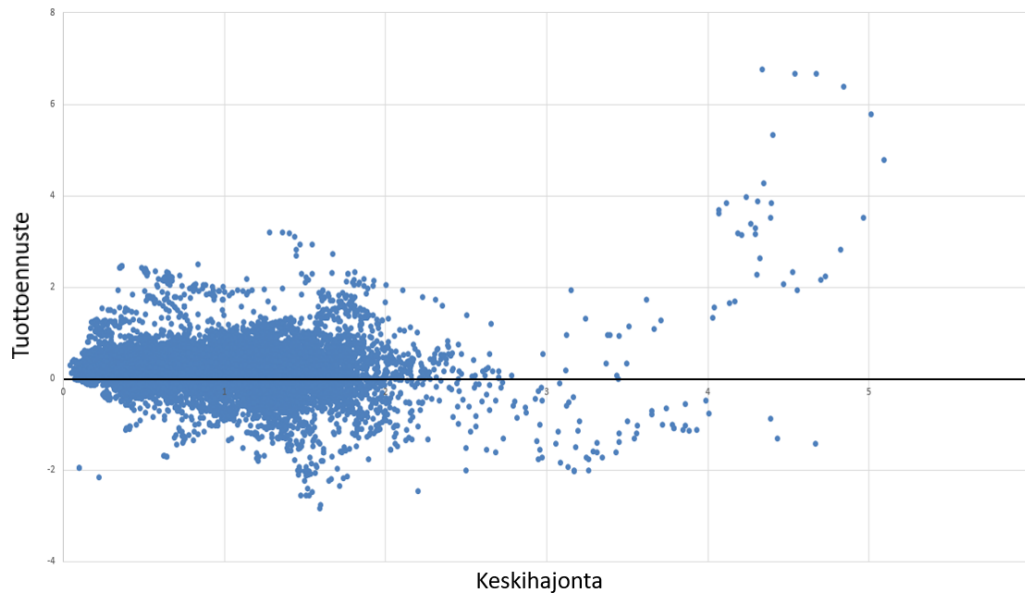
Tuottoennuste ja keskihajonta valitaan tuottoennusteen, keskihajonnan tai Sharpen luvun perusteella. Oletetaan, että Sharpen luvun avulla valittu ennuste johtaa haluttuun lopputulokseen. Osakkeen tuottoennuste ja keskihajonta valitaan suurimman Sharpen luvun perusteella, jolloin osakkeen A tuottoennusteeksi tulee -1,050 ja keskihajonnaksi 3,084 ja osakkeen B tuottoennusteeksi tulee 0,740 ja keskihajonnaksi 2,033. Luvut ovat skaalattu ja ne saadaan muutettua prosenteiksi kertomalla ne 100 %. Ennustettavaksi aikajaksoksi valikoitui osakkeelle A 2 päivää ja osakkeelle B 5 päivää.

Osakkeen B tuottoennuste 0,740 tarkoittaa  $100 \% * 0,740 * (251/5) = 1,47 \%$  potentiaalista tuottoa viidessä päivässä. Tuottoennusteessa on huomioitava tuottoennustetta pienentävät kaupankäynnin kulut, jolloin 0,2 % arvioitujen kulujen jälkeen tuottoennusteeksi saadaan 0,640. Käytännössä osakkeita saattaa jo olla omistuksessa, jolloin kaupankäynnin kulujen osuus saattaa olla eri suuruinen.

### 4.3 Simulointimenetelmät

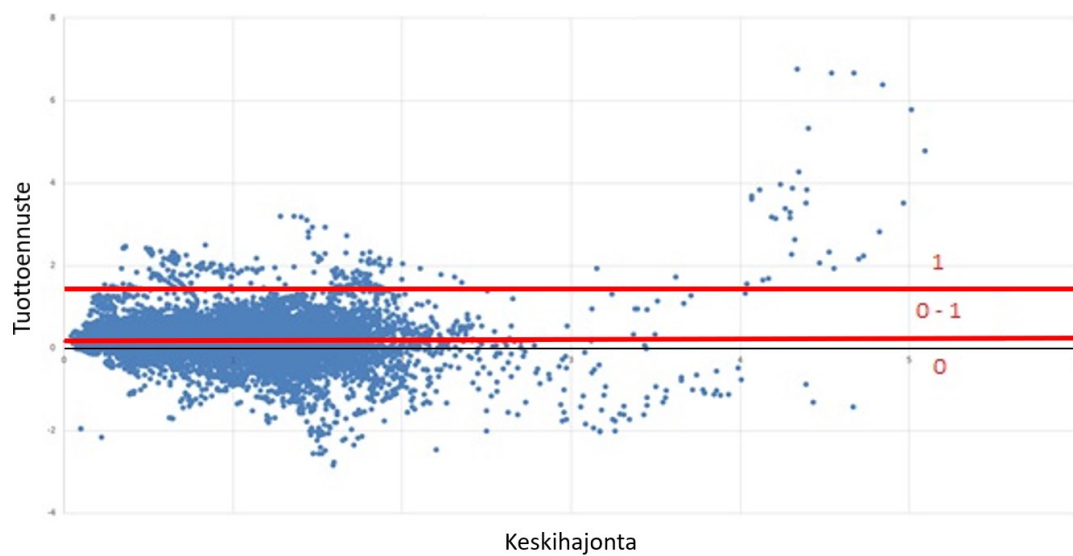
Simuloinnin alkutiedot ovat valittu tuottoennuste ja keskihajonta, joiden perusteella painoarvo määritellään yhtiökohtaisesti. Painoarvon lukuarvo on nollan ja yhden välillä. Tasahajautuksella yhtiö saa painoarvon 1, jos tarkastelupäivälle on määritelty tuottoennuste. Mikäli tuottoennuste on laadittu tasahajautuksella neljälle yhtiölle, sijoitetaan kuhunkin yhtiöön neljäsosa sijoitusvarallisuudesta.

Painoarvojen laskentaa varten on määriteltävä ala- ja yläraja, joiden mukaan painoarvot lasketaan. Ala- ja ylärajaan verrattavia tietoja voivat olla tuottoennuste, Sharpen luku tai koordinaatistoon sijoitettavan tuottoennuste–keskihajonta -pisteen etäisyys toiseen määriteltyyn pisteeseen. Tuottoennusteen ja keskihajonnan arvojen suhteellisia etäisyyksiä voidaan manipuloida tai vaihtoehtoisesti toimenpide voidaan kohdistaa suoraan painoarvoon. Alla olevassa kuviossa on yhdellä menetelmällä valitut noin 16 000 tuottoennustetta ja näiden keskihajonnat.



Kuvio 16. Yhden valintaperusteen mukaiset tuottoennusteet ja keskihajonnat

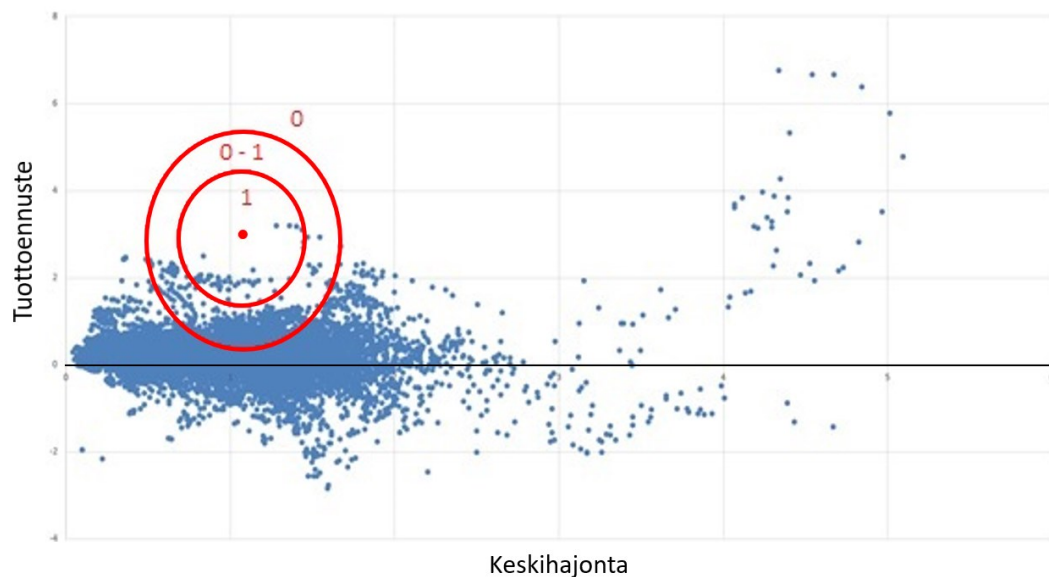
Painoarvon määrittelyyn voidaan käyttää erilaisia menetelmiä, joita käydään läpi seuraavaksi. Tuottoennustemenetelmällä painoarvoja varten määritellään alaraja, jota pienemmät tuottoennusteet saavat painoarvon 0, ja yläraja, jota suuremmat tuottoennusteet saavat painoarvon 1. Ala- ja ylärajan puolella välissä annetaan painoarvoksi 0,5. Päiväkohtaiset painoarvot lasketaan yhteen ja sijoitusvarallisuudesta sijoitetaan painoarvon suuruinen osuus painoarvojen summasta. Ala- ja ylärajaa on havainnollistettu alla olevassa kuviossa.



Kuvio 17. Tuottoennustemenetelmä

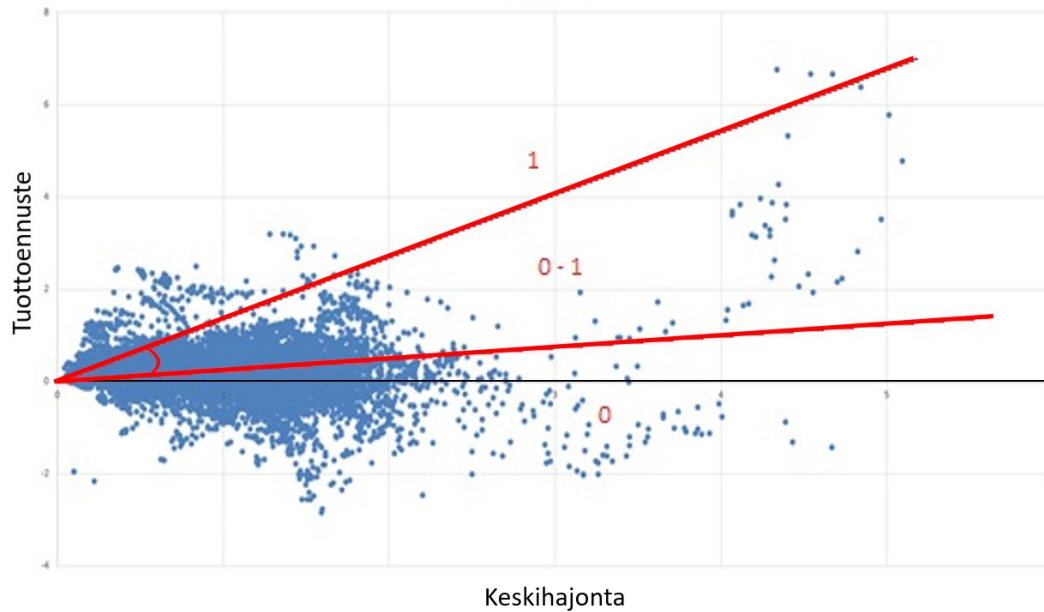
Pelkän keskihajonnan mukaan ei ole mielekäästä sijoittaa, koska tuottoennuste voi olla joko positiivinen tai negatiivinen. Ottamalla tuottoennuste mukaan, voidaan tuottoennusteen ja keskihajonnan mukaan määräytyvä ennustepiste sijoittaa koordinaatistoon, jossa ennustepisteen etäisyys punaisen ympyrän keskipisteeseen lasketaan. Punaisia ympyröitä voidaan määritellä myös useampia, jolloin lasketaan ennustepisteen etäisyys lähimmän punaisen ympyrän keskipisteeseen.

Ala- ja yläraja painoarvon määrittämiseen valitaan kokeilemalla eri arvoja. Rajoja on havainnollistettu alla olevassa kuviossa, jossa sisempi ympyrä on yläraja ja ulompi ympyrä alaraja. Menetelmä on varsinkin useampia pisteitä käytettäessä optimoinnin näkökulmasta monimutkaisempi kuin muut lähestymistavat. Etäisyys lasketaan hyödyntämällä Pythagoraan lausetta. Ennustepisteet, jotka ovat pienemmän ympyrän sisällä saavat painoarvon 1. Painoarvo nollan ja yhden välillä annetaan ennustepisteille, jotka ovat sisemmän ja ulomman ympyrän välissä. Ympyröiden ulkopuolella olevat pisteet saavat painoarvon 0.



Kuvio 18. Etäisyysmenetelmä

Sharpen luku -menetelmällä painoarvon määrittely on samankaltainen kuin tuottoennustemenetelmällä. Erona on, että tuottoennusteen sijasta käytetään Sharpen lukua. Alarajan alittavat ennusteet saavat painoarvon 0 ja ylärajan ylittävät saavat painoarvon 1. Raja-arvojen välillä painoarvo on nollan ja yhden välillä.



Kuvio 19. Sharpen luku menetelmä

Riskitön korko voidaan ottaa mukaan teorioiden mukaisesti, jolloin riskitön korko vähennetään tuottoennusteesta ja Sharpen luku lasketaan uudestaan. Kovinkaan suurta vaikutusta riskittömän koron huomioimisella ei kuitenkaan pitäisi olla. Riskitöntä korkoa ei ole määritelty, vaan se on valittava saatavilla olevista vaihtoehdoista. Yksi vaihtoehto on käyttää Suomen pankin ilmoittamaa euriborkorkoa, joka on saatavilla päivittäin koko tarkasteluajanjaksolle (Suomen pankin www-sivut 2021).

Erilaisia menetelmiä yhdistelemällä voisi myös päästä parempaan lopputulokseen. Koordinaatiston nollakohtaa voidaan siirtää johonkin muuhun kohtaan, jossa simuloinnissa saatava tuotto sijoitusvarallisuudelle on parempi. Yhtä tai useampaa menetelmää voisi käyttää joidenkin sijoitusten rajaamiseen sijoituspäätöksen ulkopuolelle. Rajauksen jälkeen toisella menetelmällä painoarvon määrittelyssä voisi päästä parempaan lopputulokseen kuin käyttämällä vain yhtä menetelmää.

#### 4.4 Painoarvo

Painoarvojen suhteellisia etäisyyksiä voidaan muuttaa korottamalla painoarvo esimerkiksi toiseen potenssiin. Pienet painoarvot ovat muutoksen jälkeen suhteellisesti pienempiä verrattuna suurempiin arvoihin. Normaalisti puolessa välissä oleva painoarvo

0,5 olisi muutoksen jälkeen 0,25. Muutos korostaisi suurempien painoarvojen merkitystä ja saattaisi johtaa suurempaan tuottoon. Ainakaan toistaiseksi tulokset eivät ole olleet kovin lupaavia, mutta tilanne saattaa muuttua sijoituskohteiden lisääntyessä. Ehdoiksi voidaan lisäksi asettaa enintään tietyn prosentuaalisen osuuden sijoittaminen yhteen yhtiöön. Sijoitettavaa varallisuutta voidaan myös rajoittaa painoarvojen summan jäädessä pieneksi, jolloin pienen tuottoennusteen yhtiöön ei sijoiteta suurta osuutta sijoitusvarallisuudesta.

Tuottoennustemenetelmää käytettäessä tarvitaan sijoittamista varten vain osakkeiden tuottoennusteet. Aikaisemmin selvitettiin, että osakkeen A tuottoennuste on -1,301 ja osakkeen B 0,640. Oletetaan että, simulointien perusteella on valittu menetelmä, jossa tuottoennusteen alaraja on 0,05 ja yläraja 0,70. Osakkeen A tuottoennuste on pienempi kuin alaraja, joten sen painoarvo on 0, eikä siihen sijoiteta. Osakkeen B tuottoennuste on raja-arvojen välissä, joten sille lasketaan painoarvo ja sijoitettava summa. Vähennetään ensin tuottoennusteesta alaraja  $0,64 - 0,05 = 0,59$ . Toiseksi ylärajasta vähennetään alaraja  $0,70 - 0,05 = 0,65$ . Kolmanneksi lasketaan painoarvo, joka on  $0,59 / 0,65 = 0,91$ .

Simuloinnissa on voinut olla kolmantena muuttujana painoarvon potenssi. Valitun potenssin ollessa 2, painoarvo korotetaan tähän potenssiin, jolloin painoarvoksi saadaan  $0,91^2 = 0,82$ . Sijoitussuunnitelmassa on voitu laatia yhteen yhtiöön sijoitettavan enintään 10 % sijoitusvarallisuudesta painoarvolla 1. Koko sijoitusvarallisuuden sijoittamista varten painoarvojen summan pitää tällöin olla vähintään 10. Sijoitusvarallisuudesta sijoitetaan kaupankäynnin kulut sisältäen osakkeeseen B  $100 \% * 0,82 / 10 = 8,2 \%$ . Sharpen luku -menetelmää käytettäessä tuottoennusteen sijasta käytetään Sharpen lukua, mutta muuten prosessi on samanlainen.

Etäisyysmenetelmää varten on määriteltävä vähintään yksi piste etäisyyden laskentaa varten. Esimerkiksi osakkeen B tuottoennuste on 0,640 ja keskihajonta 2,033. Valitun pisteen tuottoennuste voisi olla 3,000 ja keskihajonta 1,200. Apuna käytetään Pythagoraan lausetta, laskemalla ensin tuottoennusteiden etäisyys  $0,640 - 3,000 = -2,360$  ja keskihajontojen etäisyys  $2,033 - 1,200 = 0,833$ . Lasketut luvut korotetaan toiseen potenssiin, jonka jälkeen summaksi saadaan  $(-2,360)^2 + 0,833^2 = 6,263$ . Lopuksi

summasta otetaan neliöjuuri. Etäisyydeksi saadaan 2,502, jota voidaan käyttää painoarvon määrittelyyn kuten tuottoennustemenetelmässä.

#### 4.5 Simulointi

Simuloinnissa kokeillaan sijoitussuunnitelmien toimintaa 100 000 €:n sijoitusvarallisuudella ensimmäisestä tuottoennusteesta alkaen. Uusien yhtiöiden lisääminen tai optimoinnista johtuvat muutokset vaikuttavat simuloinnin lopputulokseen. Simuloinnilla etsitään optimaalinen alaraja, yläraja ja potenssi painoarvon määrittämiseen, minkä perusteella yhtiöihin sijoitettava summa määräytyy.

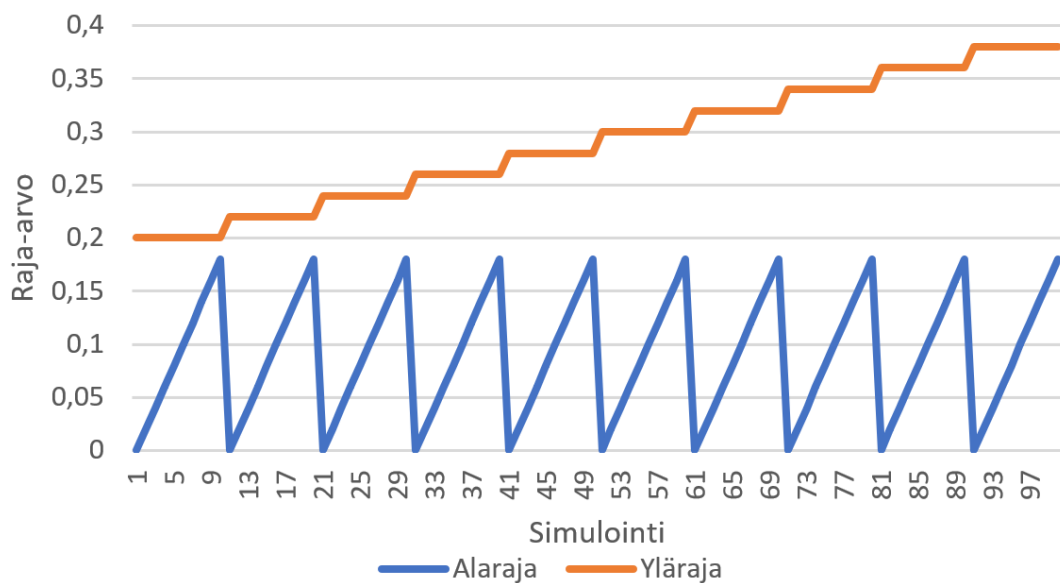
Simulointi aloitetaan asettamalla ala- ja yläraja ensimmäisellä kierroksella nollassi, minkä jälkeen 10 kierroksen ajan alarajan arvoa kasvatetaan kierros kierroksella. Kolmen muuttujan simuloinnissa erilaisia lopputuloksia syntyy  $10 \times 10 \times 10 = 1000$ . Lopputuloksena saadaan jokaisen simuloinnin keskimääräinen vuosituotto ja tuottojen keskijakoa. Simulointi on suoritettava erikseen jokaisella tuottoennusteen valintaperusteella ja jokaisella painoarvon määrittämisen menetelmällä. Kolmella valintaperusteella ja kolmella painoarvon määrittämisen menetelmällä syntyy yhteensä 9000 erilaista lopputulosta.

Pienet muutokset sijoituksissa jätetään tekemättä, kun sijoittamisen kulut pienentävät liikaa tuotto-odotusta. Kaupankäynninkuluissa voi olla huomattavia eroja, esimerkiksi Nordnet:ssä vähän kauppaa käyvän sijoittajan kaupankäyntikulut ovat Helsingin pörsissä 0,20 % tai minimissään 9 €. Yli 100 kauppaa kuukaudessa tekevän sijoittajan kaupankäyntikulut voivat olla alle 0,055 % tai minimissään 3 €. (Nordnet [www-sivut](http://www.nordnet.fi) 2021)

Sijoitukset tasapainotetaan säännöllisesti ja kaupankäyntikulut vaikuttavat siihen, kuinka usein tasapainotus kannattaa tehdä. Verotus on rajattu työn ulkopuolelle, koska verotuksessa on runsaasti erilaisia vaihtoehtoja, joihin ei tässä vaiheessa syvennyttä. Verot on mahdollista maksaa ennen kalenterivuoden päättymistä tai sen jälkeen. Sijoitukset voidaan tehdä yksityishenkilönä tai osakeyhtiön kautta, jolloin verotus on erilaista. Lisäksi pitäisi huomioida ainakin osakesäästötilin mahdollisuudet,

osinkoverotus ja muut pääomatulot. Pitkällä aikavälillä tavoite on kuitenkin mahdollisimman hyvin ottaa huomioon kaikki sijoituspäätökseen vaikuttavat asiat.

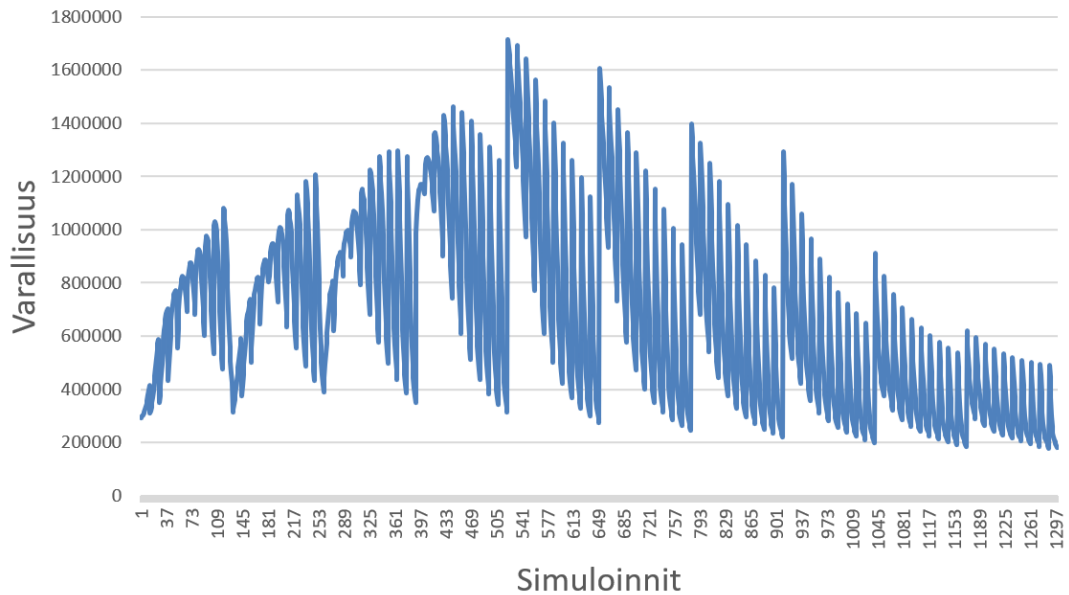
Simuloinnin avulla pyritään hyödyntämään tuottoennuste ja keskihajonta mahdollisimman tuottavasti. Yhtenä vertailukohtana voidaan pitää tasahajautusta, jossa sijoitetaan tasasumma jokaiseen yhtiöön. Sijoitukset tehdään samoina päivinä, kun tuottoennuste on laskettavissa. Vertailukohtana voi olla myös muut sijoitussuunnitelmat. Sijoitussalkun arvoksi simuloinnin alussa on valittu 100 000 €, koska summalla voidaan tehdä riittävän monta toimeksiantoa ja muuttaa sijoitusten kokoa. Jos pienin sijoitussumma olisi 4 000 €, niin varoilla voidaan toteuttaa 25 toimeksiantoa. Näin pystytään jo helposti muuttamaan sijoitusten kokoa. Mukana on vain neljä yhtiötä, joten sijoitussuunnitelmat pidetään yksinkertaisena ja pyritään kohti todenmukaisia lopputuloksia. Alla on havainnollistettu simulointien raja-arvoja muuttamalla ala- ja ylärajaa yhteensä 100 kertaa. Raja-arvojen nostoväliksi on määritelty 0,02.



Kuvio 20. Raja-arvot

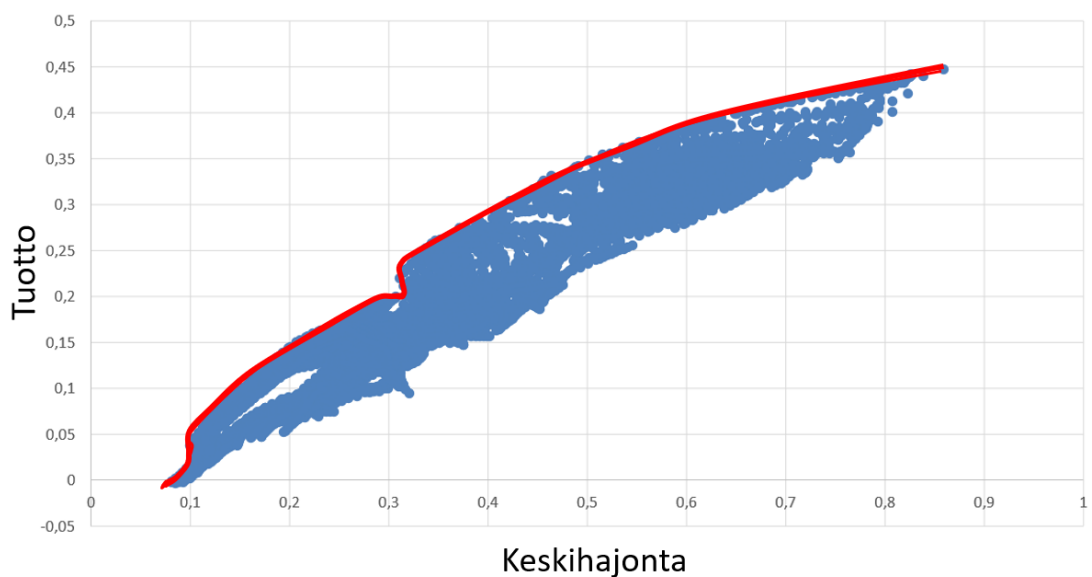
Luomalla viivakuvio sijoitusvarallisuuden kehityksestä nähdään missä vaiheessa simulointia saavutetaan suurin sijoitusvarallisuus. Alla olevassa kuviossa kolmella muuttujalla simulointeja on suoritettu 1300. Raja-arvojen alueet voidaan määrittellä paremmin, jos sijoitusvarallisuuden huippu olisi kuvion reunassa. Pelkkä varallisuuden määrä simuloinnin lopussa ei ole paras vertailutieto, vaan arvioinnissa voidaan

käyttää apuna tutuksi tullutta tuottoa ja keskihajontaa. Kuviosta on nähtävissä, että monilla erilaisilla arvoilla voidaan saavuttaa sama tuotto euroissa.



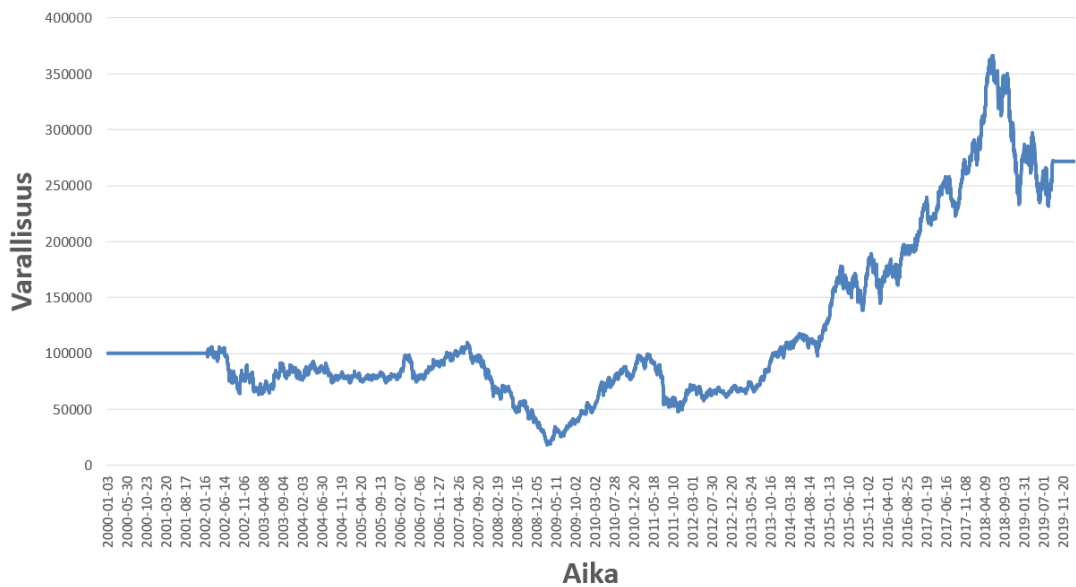
Kuvio 21. Varallisuus simulointien lopussa

Simulointien tiedoista lasketaan keskimääräinen vuosituotto ja tuottojen keskihajonta. Näistä tuloksista valitaan sopivin tuotto ja keskihajonta tulevaan sijoitustoimintaan. Valinta kohdistuu simulointiin, jonka tulokset sijaitsevat punaisen viivan alueella, koska näissä pisteissä keskimääräinen tuotto on korkein samalla tuottojen keskihajonnalla. Punaisen viivan vasemmassa reunassa sijaitsee pienin keskihajonta ja oikeassa reunassa sijaitsee suurin tuotto.



Kuvio 22. Simulointien tuotto ja keskihajonta

Valinnan jälkeen sijoitussuunnitelman toimintaa voidaan tarkastella yksityiskohtaisemmin. Kiinnostavaa on ainakin nähdä, miten sijoitusvarallisuus on kehittynyt. Päivakohtaisiin ennusteiden valintoihin, painoarvoihin, sijoituksiin ja kaupankäyntiin voidaan lisäksi syventyä. Tärkeintä on saada tieto käytettävästä ennusteiden valintaperusteesta, painoarvon määräytymisperusteesta ja raja-arvoista, joita käytetään oikeiden sijoituspäätösten tekemisessä.



Kuvio 23. Sijoitusvarallisuuden kehitys

Yllä olevassa kuviossa on havainnollistettu tasahajautetun sijoitusportfolion kehitystä. Sijoitukset on tasapainotettu päivittäin eikä kaupankäyntikuluja ole otettu huomioon. Ensimmäisen vuosikymmenen aikana kehitys on ollut heikkoa. Jälkimmäisellä vuosikymmenellä kehitys on ollut erittäin hyvää.

## 5 HALLINTA

Kokonaisuudesta on hyvä olla visio siitä, millaisia menetelmiä ja tietoja käytetään tulevaisuudessa. Missio on näkemys siitä, miten visio saavutetaan. Jatkuvassa parantamisessa nähdään mahdollisuus tehdä jokin asia paremmin, tehokkaammin tai

helpommin. Prosessiajattelun avulla käsitellään asioita osana kokonaisuutta. Helpoin tapa johtaa tosiasioihin perustuen, on käyttää tilastollisia menetelmiä, jossa merkittävät asiat voidaan erotella sivuseikoista. Esiin tulevat asiat on käsiteltävä tärkeysjärjestyksessä. Syy-seuraussuhteet prosessissa pyritään tunnistamaan mahdollisimman hyvin. Kehitystyötä auttaa, kun pystytään määrittelemään, kuinka paljon prosessia on jo pystytty kehittämään. Tarkkaan määritelty prosessi luo parempia kehitysmahdollisuuksia. Ennustamisessa korrelaatio tuottoennustuksien ja toteutuneiden tuottojen osalta pyritään saamaan mahdollisimman suureksi. Simuloinnissa kehitystä arvioidaan keskimääräisen tuoton ja keskihajonnan muutoksen avulla. (Salomäki 1999, 9-29)

Päätöksenteon ongelma on usein, ettei valmista ratkaisumallia ole, vaan ratkaisu ongelmaan on itse löydettävä. Menetelmiä valvomalla on löydettävissä vinkkejä jo olemassa olevista tai tulevista ongelmista ennen kuin ne aiheuttavat suurempia ongelmia. Kehityskohteiden puuttuminen ei yleensä ole ongelma, vaan ennemminkin päättäminen siitä, mikä ongelma ensimmäisenä lähdetään ratkaisemaan. Tuloksellisuuden kannalta saattaa olla parasta aloittaa selkeimmistä ongelmista. Lähtötilanteen tunnistaminen on suurin ja tärkein askel kehitystoiminnan aloittamisessa. (Salomäki 1999, 37-73)

Tuottoennusteen osalta visiona on määritellä tunnuslukujen raja-arvot erikseen jokaiselle ennusteajaksolle. Yksittäiset tunnusluvut vaikuttavat tuottoon lyhyellä aikajaksolla luultavasti eri tavalla kuin pitkällä aikajaksolla, jolloin tunnuslukujen raja-arvot ja muutosnopeudet voisi myös määritellä erikseen. Sijoitussuunnitelman osalta visiona on mahdollisimman hyvin oman varallisuuden määrän huomioiminen ja uusien lähestymistapojen etsiminen. Suunnitelman laadinnassa voi olla vähän luovuutta mukana, jolloin kokeilemalla erilaisia lähestymistapoja voidaan saavuttaa parempi lopputulos.

## 5.1 Tunnusluvut ja toteutunut tuotto

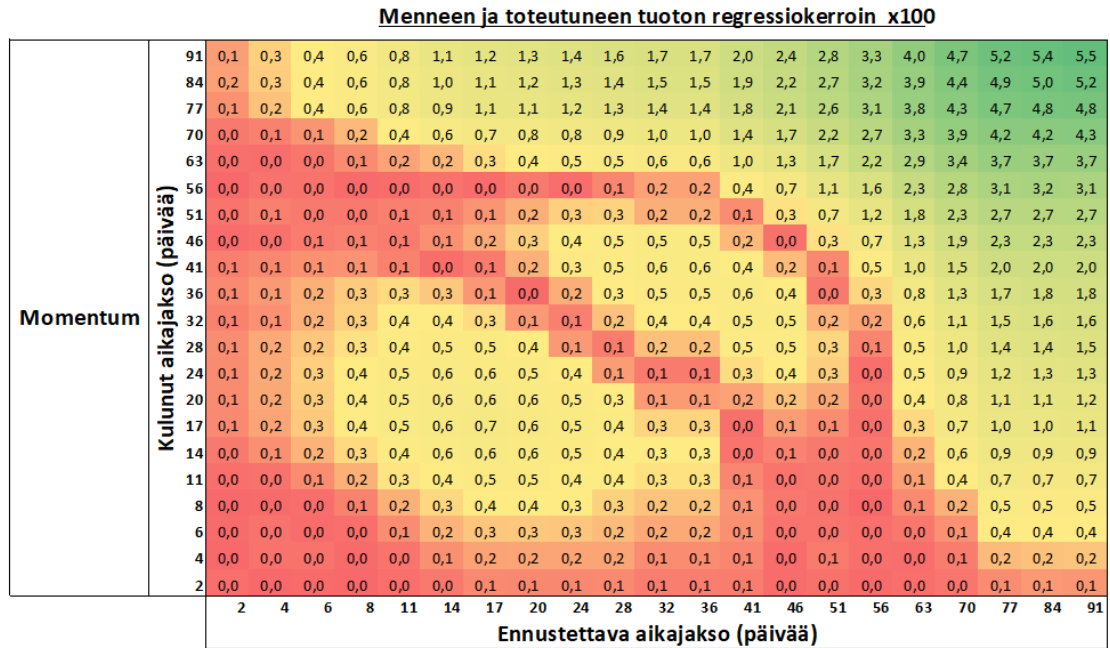
Ennustamisessa voidaan käyttää apuna suurta määrää erilaisia tunnuslukuja. Kaikki erilaiset raja-arvojen ja muutosnopeuksien yhdistelmät on mahdotonta käydä läpi. Koneoppimista hyödyntämällä tunnuslukujen vaikutusta tuoton selittäjänä voidaan yrittää

selvittää tunnusluku kerrallaan. Tuloslaskelma, tase, rahoituslaskelma, pörssin arvostuskertoimet ja osakekurssin käyttäytymistä mittaavat tunnusluvut muuttuvat erisuuruisesti suhteessa toisiinsa, jolloin niiden tiedot täydentävät toisiaan. Alma Talent käyttää esimerkiksi yli 50 tunnuslukua tilinpäätösanalyysissä ja tiivistää ne yhdeksi arvosanaksi (Alma Talentin www-sivut 2021).

Tunnuslukujen avulla muodostetaan ja kuvaillaan tarinaa yhtiön kehityksestä. Yrityksen arvoa määritetään diskonttaamalla ennustettuja tulevaisuuden tuotot nykyhetkeen, jolloin kaikki tuottoihin vaikuttavat asiat on otettava huomioon (Damodaran 2017). Tilikauden tuloksen osalta tarinaa luovat tilikauden tuloksen ja hinnoittelun suhdetta kuvaava P/E-luku yhdessä historiallisen tuloksen muutoksen ja ennustetun muutoksen kanssa.

Tarinan muodostamista voidaan hyödyntää seuraavien tunnuslukujen käyttöönotossa. Tarinan kerronta aloitetaan kartoittamalla, mitä perustietoja ja tunnuslukuja on jo hyödynnetty ja mitä seuraavaksi voitaisiin hyödyntää. Tunnusluvut jaetaan niihin liittyville perustiedoille ja selvitetään karkeasti yhteys toteutuneeseen tuottoon regressio-kertoimen avulla. Selvitysten jälkeen eteneminen on suunnitelmallisempaa ja prosessi on järjestelmällisempi.

Kvantitatiivisen tutkimuksen avulla tutkitaan yleensä lukumääriä ja prosentuaalisia osuuksia. Otoksoon pitää olla tarpeeksi suuri edustamaan tutkittavaa kohdetta. Tuloksia yritetään yleistää tilastollisen päättelyn keinoin. Tietojen yhteyksiä voidaan esittää korrelaatiokertoimen avulla, joista useimmiten käytetty on Pearsonin korrelaatiokerroin. Korrelaatiota voidaan käyttää apuna jatkotutkimuksia varten. Tietoa voidaan esimerkiksi jakaa ryhmiin yhden muuttujan perusteella ja tutkia ryhmien välisiä eroja. (Heikkilä 2014, 15, 192-195, 210, 222-223.)



Kuvio 24. Momentum-lämpökartta

Yllä olevan kuvion perusteella aikaisemmin toteutunut tuotto selittää osittain tulevaa tuottoa ja sen tiedot ovat jatkoa kuvion 26 tiedoille. Viimeisen 91 päivän tuotto selittää regressiokerroimen perusteella tulevaa tuottoa parhaiten kaikkina ennustettavina aikajaksoina, minkä perusteella se on valittu käytettäväksi ennustamisessa. Kahden tai useamman aikajakson käytöllä kurssikehityksestä saadaan enemmän tietoa, jolloin lopputuloskin saattaisi olla parempi. Lopputuloksen toteaminen vaatii kuitenkin toisen aikajakson mukaan ottamista osaksi ennustamista tai jonkin muun menetelmän.

	Liikevaihto	Liikevoitto	Tulos ennen veroja	Tulos
Liikevaihto	-	0,41	0,23	0,24
Liikevoitto	0,41	-	0,82	0,80
Tulos ennen veroja	0,23	0,82	-	0,95
Tulos	0,24	0,80	0,95	-

Kuvio 25. Tuloslaskelman perustietojen korrelaatiot

Tuloslaskelman tietojen välisistä korrelaatioista voidaan nähdä, että korrelaatio tulos ennen veroja ja tuloksen välillä on 0,95. Voimakkaan korrelaation takia molempia ei

ainakaan aluksi kannata ottaa käyttöön, koska ne muuttuvat lähes samankaltaisesti. Liikevoiton ja tulos ennen veroja sekä tuloksen kanssa on myös korkea, joten näiden yhteiskäyttöä kannattaa harkita. Liikevaihto sen sijaan saattaisi tuottaa enemmän lisäarvoa hyödynnettäessä yhdessä tuloksen kanssa.

Perustieto	Tunnusluku	Regressiokerroin toteutuneeseen tuottoon																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Liikevaihto	P/S-luku																					
	Hist.muutos-%																					
	Tul.muutos-%																					
Liikevoitto	Liiketulos-%																					
	EV/EBIT-luku																					
	Hist.muutos-%																					
Voitto ennen veroja	Tul.muutos-%																					
	Lainojen hoitokate																					
	Hist.muutos-%																					
Tilikauden voitto	Tul.muutos-%																					
	P/E-luku																					
	Nettotulos-%																					
	Tulostuotto-%																					

Kuvio 26. Tuloslaskelman tunnuslukujen ja perustietojen kohdistusta

Tunnusluvut kohdistetaan yhdelle tai useammalle perustiedolle. Laskemalla tämän jälkeen regressiokerroin toteutuneen tuoton kanssa eri aikajaksoilla saadaan karkea kuva perustiedon ja siihen liittyvien tunnuslukujen tärkeydestä. Mitä enemmän regressiokerroin poikkeaa nolasta positiiviseen tai negatiiviseen suuntaan, sitä hyödyllisempää tunnusluvun käyttäminen saattaa olla.

Yhdessä perustietojen korrelaatioiden kanssa regressiokertoimet antavat paremman kuvan. Näin voidaan arvioida miten tiedot liittyvät toisiinsa ja miten ne liittyvät saadaan tuottoon. Karkea ennakkotyö on nopea toteuttaa ja auttaa ymmärtämään tunnuslukujen laajaa joukkoa. Tunnuslukujen yhteisvaikutus selviää myöhemmin otettaessa tunnusluku osaksi ennustamista.

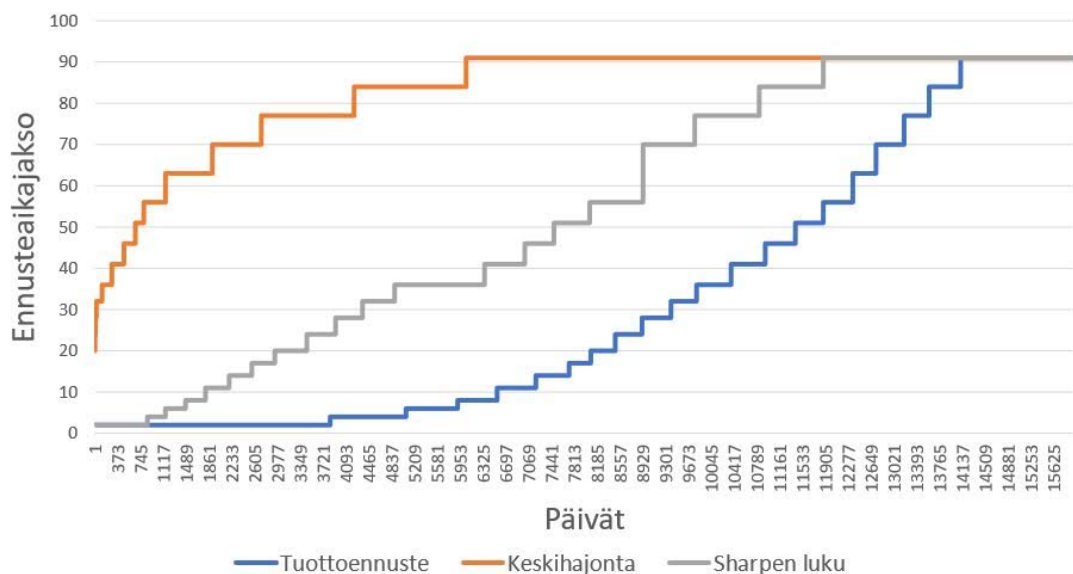
Kvantitatiivisia menetelmiä käytetään tuottoennusteen määrittelyssä. Olennaisten tunnuslukujen ja niiden välisten suhteiden tunteminen auttaa kehittämään menetelmistä parempia. Tuottoennusteen ja toteutuneen tuoton korrelaatio on tärkein ennustamisen mittari, jonka ehtona on, että vastaavuuksia on löydyttävä riittävä määrä jokaiselle päivälle. Optimoinnissa korkeamman korrelaation antava vaihtoehto on aina parempi ja

se valitaan käytettäväksi. Valinta voi liittyä esimerkiksi käytettävien tunnuslukujen raja-arvoihin tai niiden muutosnopeuksiin.

## 5.2 Ennusteajakajako

Sijoitettaessa odotuksena on saada tuottoa sijoitukselle jonkin ajan kuluessa. Lyhyen aikajakson tuottoennusteella on yleensä selvästi suurempi keskihajonta suhteessa tuottoennusteeseen kuin pidemmillä aikajaksoilla. Lyhyellä aikajaksolla potentiaalisten tuottojen määrä suhteessa aikaan on suuri, mutta niin on myös potentiaalisten tappioiden. Lisäksi lyhyen aikajakson kaupankäynnissä kulujen merkitys on suurempi. Optimaalinen aikajänne riippuu käytössä olevista tuottoennusteista ja keskihajonnoista.

Ennusteen valintaperusteella on selkeä vaikutus ennustevalintojen aikajaksoihin. Pisin ennustettava aikajakso on 91 kaupankäyntipäivää. Se voisi olla vielä jonkin verran pidempi, koska pisimmän aikajakson osuus kaikista aikajaksoista vaikuttaa suurelta. Paras ennusteen valintaperuste sijoittamisen kannalta selviää simulointien perusteella myöhemmin, kun aineistoa valinnan tekemiseen on riittävästi. Alla olevasta kuviosta voidaan nähdä, että eri valintaperusteilla on merkittävä vaikutus valittuihin ennusteajakajaksoihin.



Kuvio 27. Valintaperusteen vaikutus ennusteajakajaksoihin

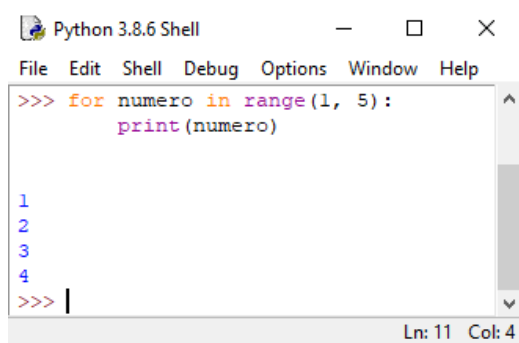
Keskihajonnan perusteella valittaessa ennusteajakajakso on keskimäärin selvästi pidempi kuin tuottoennusteen perusteella valittu ennusteajakajakso. Sharpen luvun perusteella valittu ennusteajakajakso on keskimäärin keskihajonnan ja tuottoennusteen perusteella valittujen ennustejaksojen välissä. Aikajaksojen erojen takia ennustemenetelmää kehitetään yleisemmin kaikkien aikajaksojen tuottoennusteiden ja toteutuneiden tuottojen korrelaatioiden keskiarvon muutoksen perusteella.

### 5.3 Ryhmittely ja iterointi

Tuottoennusteen määrittelyssä aineisto käsitellään jakamalla se ryhmiin, joissa havaintojen väliset etäisyydet tunnetaan. Tunnuslukujen raja-arvojen ja muutosnopeuksien muutokset ovat pääasiallinen tapa vaikuttaa vastaavuuksiin ja lopputulokseen. Muutos tunnuslukujen hyödyntämisessä tuottaa aina omanlaisensa lopputuloksen. Valmista menetelmää ei ole olemassa, vaan ryhmittely tehdään tapauskohtaisesti. (Holopainen T. 2012, 2-6.)

Iteroinnissa likiratkaisun avulla voidaan ohjata seuraavan likiratkaisun laskentaa, jolloin syntyy tavoiteltavaa lopputulosta kohti tarkentuva lopputulos. Puolitusmenetelmää käytettäessä alkuarvoina annetaan alku ja päätepiste, joiden avulla ratkaisua aleataan lähestymään. Menetelmää voidaan jatkaa niin kauan, kunnes haluttu tarkkuus on saavutettu. (Arvio V. 2017, 27-34.)

Iteroinnin avulla käsitellään toistuvia prosesseja, kuten tunnuslukujen, tuottoennusteiden, painoarvojen ja kaupankäynnin toimenpiteiden laskenta. Tuottoennusteet lasketaan esimerkiksi jokaiselle aineiston päivälle aina samalla toistuvalla menetelmällä. Aineiston määrä tunnistetaan ja haluttu prosessi toistetaan jokaiselle päivälle ensimmäisestä päivästä lähtien. Jokaisen päivän kohdalla iteroidaan aineiston läpi vähintään kerran vastaavuuksien tunnistamiseksi. Iterointia on havainnollistettu alla olevalla ohjelmalla, jossa lähdetään liikkeelle numerosta 1 ja tulostetaan numero (sinisellä), niin kauan kuin se on alle viiden. (Towards data science www-sivut 2018)



```
Python 3.8.6 Shell
File Edit Shell Debug Options Window Help
>>> for numero in range(1, 5):
        print(numero)
1
2
3
4
>>> |
Ln: 11 Col: 4
```

Kuvio 28. Iterointi

Tuottoennusteen määrittelyyn on valittu aika, joka on kulunut viimeisimmästä tulosjulkistuksesta. Vuosineljänneksen alkupuolella vastaavuuksia etsitään myös muiden vuosineljänneksien alkupuolelta, jolloin tilanteet vastaavat paremmin toisiaan. Raja-arvoksi voidaan asettaa esimerkiksi 10 päivää ennen ja jälkeen kyseisen päivän.

Tuloslaskelman tiedoista käytetään liiketuloksen ja tuloksen muutoksen vaikutusta. Viimeisen 12 kuukauden juoksevaa tulosta verrataan seuraavan vuosineljänneksen mukaan laskettuun ennustettuun tuloksen muutokseen. Liiketuloksen osalta tarkastellaan vastaavasti historiallista kehitystä. Kannattavuuden muutos menneisydessä ja tulevaisuudessa lasketaan aina vuoden ajanjaksolla. Raja-arvona voidaan käyttää 8 %-yksikköä.

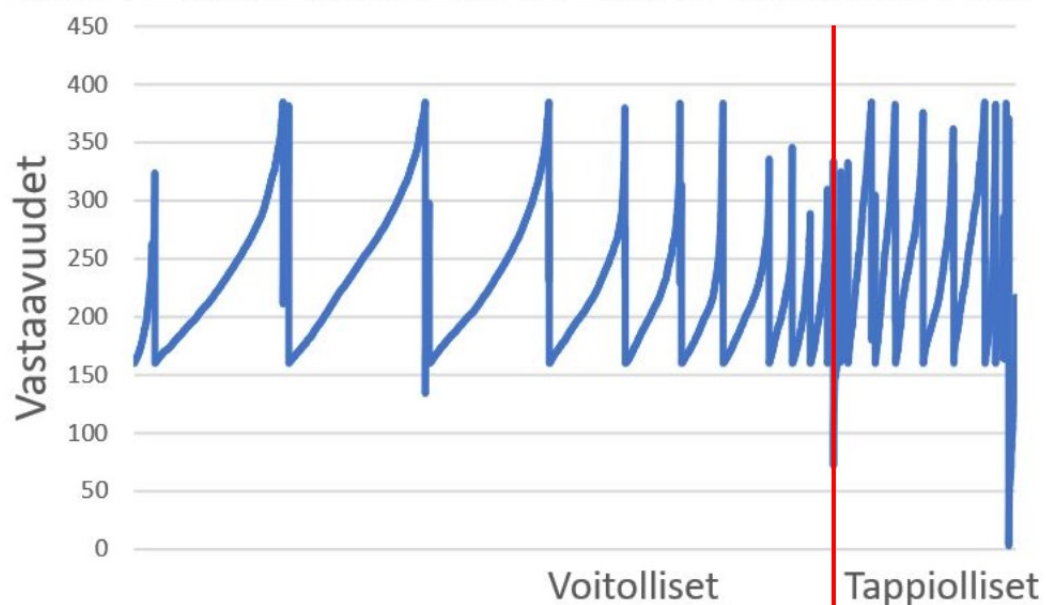
Nykytilannetta tarkastellaan pörssin arvostuksen kautta P/E-tunnusluvulla, jossa suhteutetaan yhtiön tulos markkina-arvoon. Käytännössä tarkastelu on tehty tunnusluvun käänteisluvulla eli tulostuotolla. Kummallakin tavalla päästään samaan lopputulokseen. Raja-arvoksi voidaan määrittellä 5 %-yksikköä. Viimeisenä tarkastellaan osakkeen tuottoa viimeisen 91 kaupankäyntipäivän ajalta. Raja-arvona käytetään 10 %-yksikköä.

Tietokanta iteroidaan läpi tarkastamalla kaikki tunnusluvut raja-arvojen osalta. Yhdellä rivillä on aina yhden yhtiön päiväkohtaiset tiedot. Valittujen tietojen ollessa raja-arvojen sisäpuolella tulee tarkasteltavasta päivästä vastaavuus, jota hyödynnetään tuottoennusteen laskemiseen. Vastaavuuksille on asetettu tavoiteltavaksi määräksi vähintään 150 ja enintään 400 vastaavuutta. Mikäli vastaavuuksia ei löytynyt tavoiteltua määrää, muutetaan raja-arvoja seuraavaa tarkistuskierrosta varten. Seuraavalle

kierrokselle raja-arvoja kasvatetaan kullekin tunnusluvulle yksilöllisesti valitulla muutosnopeudella.

Puolitusmenetelmää käytetään, kun vastaavuuksien määrä muuttuu raja-arvojen kasvatuksen jälkeen liikaa. Tällöin vastaavuuksia löytyy ensin vähemmän kuin 150 ja seuraavalla kierroksella vastaavuuksia löytyy yli 400. Toimenpiteessä edellinen muutosnopeus puolitetaan, jolloin päästään lähemmäs tavoiteltua vastaavuuksien määrää.

Lopputulosta arvioidaan karkeasti alla olevan kuvion avulla, jossa tiedot on lajiteltu kierroksien perusteella. Tappiollisten päivien määrä on selvästi pienempi kuin voitollisten ja niiden ryhmittely alkaa punaisesta viivasta. Ensimmäinen kierros alkaa vasemmasta reunasta ja päättyy ensimmäiseen huippuun. Kun vastaavuudet eivät osu tavoiteltavaan haarukkaan, niin raja-arvoja muutetaan puolitusmenetelmällä. Puolitusmenetelmän hyödyntäminen näkyy sahalaitana huippujen yhteydessä. Vastaavuuksia etsitään enintään kaksitoista kierrosta, minkä jälkeen siirrytään eteenpäin. Tietokannan kasvaessa raja-arvoja ja muutosnopeuksia joudutaan muuttamaan, kun tavoitteena on tietty määrä vastaavuuksia.



Kuvio 29. Vastaavuudet kierroksittain

Tunnusluvun ja toteutuneen tuoton välinen yhteys vaikuttaa siihen, kuinka paljon tunnusluvun avulla rajataan vastaavuuksien määrää. Voimakkaan yhteyden kohdalla

vastaavuuksia rajataan enemmän asettamalla tiukat raja-arvot kuin heikon yhteyden kohdalla. Raja-arvot määräytyvät koneoppimisen tulosten perusteella.

#### 5.4 Luokittelu

Tietokannassa olevien päivien tuottoennuste tai toteutunut tuotto voidaan luokitella tuottoennusteen määrittelyssä käytettyjen tunnuslukujen avulla. Määrittelyssä on käytetty viittä tunnuslukua, joiden avulla tuottoennuste jaetaan 144 erilaiseen luokkaan. Tuottoennusteen aikajaksoksi on valittu 91 kaupankäyntipäivää. Taulukko muodostaa yleiskatsauksen tuottoennusteista tällä aikajaksolla.

Luokkien määrittelyksiä:

- Momentum 1: Tuotto alle -10 % viimeisen 91 päivän aikana
- Momentum 2: Tuotto yli -10 % ja alle + 10 % viimeisen 91 päivän aikana
- Momentum 3: Tuotto yli 10 % viimeisen 91 päivän aikana
- Alkukvartaali: kvartaalin ensimmäinen puolisko
- Loppukvartaali: kvartaalin jälkimmäinen puolisko
- LV-: Liikevoitto laskee seuraavalla kvartaalilla
- LV+: Liikevoitto nousee seuraavalla kvartaalilla
- Tulos--: Tulos laskenut yli -15 %
- Tulos-: Tulos laskenut 0 – (-15 %)
- Tulos+: Tulos noussut 0 – 15 %
- Tulos++: Tulos noussut yli +15 %
- Kallis: P/E-luku on suurempi kuin 15,4
- Halpa: P/E-luku on pienempi kuin 15,4
- Tappiollinen: Tulos on negatiivinen

Aineiston määrällinen jakautuminen luokkiin on hyvä tarkistaa ja tarvittaessa muuttaa luokkien rajoja. Joihinkin luokkiin voi silti kertyä vain vähän tai ei lainkaan vastaavuuksia, mikä on otettava huomioon lukujen tulkinnassa. Luokittelemalla tiedot saadaan karkea kuvaus tuottoennusteen käyttäytymisestä. Tunnuslukujen määrä ja raja-arvot muuttuvat tulevaisuudessa, joten myös luokittelu tulee muuttumaan.

## Tappiolliset

Keskimääräinen tuottoennuste	
LV-	4 %
LV+	7 %
Alkukvartaali	5 %
Loppukvartaali	6 %
Tulos--	5 %
Tulos-	4 %
Tulos+	4 %
Tulos++	9 %
Momentum 1	3 %
Momentum 2	4 %
Momentum 3	10 %

Tuottoennuste		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
<b>Momentum 1</b>	Tulos--	3 %	2 %	12 %	7 %
	Tulos-	-2 %	0 %	-5 %	14 %
	Tulos+	0 %	2 %	4 %	8 %
	Tulos++	-2 %	7 %	-7 %	4 %
<b>Momentum 2</b>	Tulos--	-2 %	1 %	9 %	6 %
	Tulos-	2 %	10 %	3 %	10 %
	Tulos+	1 %	0 %	4 %	0 %
	Tulos++	0 %	11 %	-3 %	8 %
<b>Momentum 3</b>	Tulos--	6 %	5 %	9 %	6 %
	Tulos-	1 %	9 %	0 %	9 %
	Tulos+	4 %	8 %	6 %	12 %
	Tulos++	38 %	16 %	19 %	11 %

Määrä		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
<b>Momentum 1</b>	Tulos--	222	113	189	95
	Tulos-	33	-	67	12
	Tulos+	33	32	10	31
	Tulos++	65	33	52	38
<b>Momentum 2</b>	Tulos--	151	73	142	78
	Tulos-	15	21	3	49
	Tulos+	23	-	17	-
	Tulos++	31	123	45	74
<b>Momentum 3</b>	Tulos--	139	102	169	94
	Tulos-	16	75	-	52
	Tulos+	40	32	48	16
	Tulos++	32	196	45	236

Kuvio 30. Tappiolliset

Tappiollisuus on johtanut pääasiassa positiivisiin tuottoennusteisiin ja tuotoissa esiintyy huomattavaa hajontaa. Tappiollisuus ei ole johtanut negatiivisiin tuottoennusteisiin, vaan negatiiviset tuottoennusteet ovat ajoittuneet aikaan ennen tappiollisuutta. Suuri osa negatiivisista tuottoennusteista sijoittuu finanssikriisin aikaan ja tulosten muututtua negatiiviseksi tulevaisuus on jo alkanut näyttää huomattavasti paremmalta. 91 kaupankäyntipäivän keskimääräinen tuottoennuste on kolmen ja kymmenen prosentin välillä. Liikevoiton parannuttua viimeisimmässä osavuositarkastuksessa, tuloksen parantuessa seuraavalla kvartaalilla ja yli 10 % tuotto viimeisen 91 kaupankäyntipäivän aikana on johtanut hyviin tuottoennusteisiin seuraavan 91 kaupankäyntipäivän aikana.

## Halpa

Keskimääräinen tuottoennuste	
LV-	2 %
LV+	5 %
Alkukvartaali	3 %
Loppukvartaali	3 %
Tulos--	2 %
Tulos-	5 %
Tulos+	4 %
Tulos++	2 %
Momentum 1	2 %
Momentum 2	3 %
Momentum 3	4 %

Tuottoennuste		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
<b>Momentum 1</b>	Tulos--	0 %	0 %	2 %	1 %
	Tulos-	2 %	6 %	4 %	7 %
	Tulos+	2 %	3 %	3 %	6 %
	Tulos++	-4 %	7 %	-5 %	4 %
<b>Momentum 2</b>	Tulos--	-1 %	1 %	2 %	1 %
	Tulos-	4 %	5 %	4 %	6 %
	Tulos+	3 %	6 %	2 %	8 %
	Tulos++	0 %	6 %	-2 %	5 %
<b>Momentum 3</b>	Tulos--	7 %	2 %	2 %	1 %
	Tulos-	3 %	4 %	3 %	10 %
	Tulos+	3 %	5 %	2 %	5 %
	Tulos++	2 %	6 %	5 %	3 %

Määrä		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
<b>Momentum 1</b>	Tulos--	117	-	110	3
	Tulos-	182	60	171	91
	Tulos+	158	148	71	235
	Tulos++	81	24	113	79
<b>Momentum 2</b>	Tulos--	123	35	138	37
	Tulos-	264	40	240	36
	Tulos+	135	354	162	281
	Tulos++	248	292	228	228
<b>Momentum 3</b>	Tulos--	27	29	27	27
	Tulos-	189	9	134	3
	Tulos+	134	189	135	158
	Tulos++	58	257	21	233

Kuvio 31. Halpa

Voitollisten päivien osalta halvemmassa puoliskossa voidaan havaita pienempiä eroja tuottoennusteissa tunnuslukujen välillä. Tuotto on vähän suurempi silloin, kun seuraavan kvartaalin liikevoitto nousee vertailukaudesta verrattuna tilanteisiin, joissa seuraavan kvartaalin liikevoitto laskee vertailukaudesta. Positiivinen tuotto viimeisen 91 kaupankäyntipäivän aikana on hyvä merkki ainakin positiivisesta tuottoennusteesta. Vain neljässä ryhmässä tuottoennuste on negatiivinen. Keskimääräiset tuottoennusteet ovat kahden ja viiden prosentin välissä.

Kalliimmassa puoliskossa tuottoennuste on negatiivinen 8 ryhmässä ja osakkeiden ennustetaan laskevan enemmän kuin halvemmassa ryhmässä. Seuraavan kvartaalin liikevoiton noustessa tuottoennuste on suurempi kuin liikevoiton laskiessa. Keskimääräiset tuottoennusteet ovat nollan ja kuuden prosentin välissä.

## Kallis

Keskimääräinen tuottoennuste	
LV-	0 %
LV+	6 %
Alkukvartaali	3 %
Loppukvartaali	3 %
Tulos--	4 %
Tulos-	5 %
Tulos+	3 %
Tulos++	0 %
Momentum 1	3 %
Momentum 2	3 %
Momentum 3	3 %

Tuottoennuste		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
Momentum 1	Tulos--	2 %	5 %	4 %	6 %
	Tulos-	8 %	5 %	4 %	4 %
	Tulos+	3 %	5 %	0 %	6 %
	Tulos++	-7 %	10 %	-12 %	11 %
Momentum 2	Tulos--	0 %	3 %	6 %	5 %
	Tulos-	5 %	4 %	4 %	7 %
	Tulos+	-1 %	7 %	-4 %	7 %
	Tulos++	-4 %	9 %	-10 %	5 %
Momentum 3	Tulos--	3 %	4 %	8 %	9 %
	Tulos-	3 %	5 %	3 %	5 %
	Tulos+	1 %	6 %	2 %	5 %
	Tulos++	-5 %	7 %	-12 %	6 %

Määrä		Alkukvartaali		Loppukvartaali	
		LV-	LV+	LV-	LV+
Momentum 1	Tulos--	213	67	182	55
	Tulos-	48	6	38	34
	Tulos+	17	70	-	61
	Tulos++	12	56	14	17
Momentum 2	Tulos--	230	189	187	212
	Tulos-	53	126	33	157
	Tulos+	104	281	82	252
	Tulos++	147	317	107	395
Momentum 3	Tulos--	91	64	88	28
	Tulos-	62	335	108	258
	Tulos+	124	323	128	357
	Tulos++	30	462	58	410

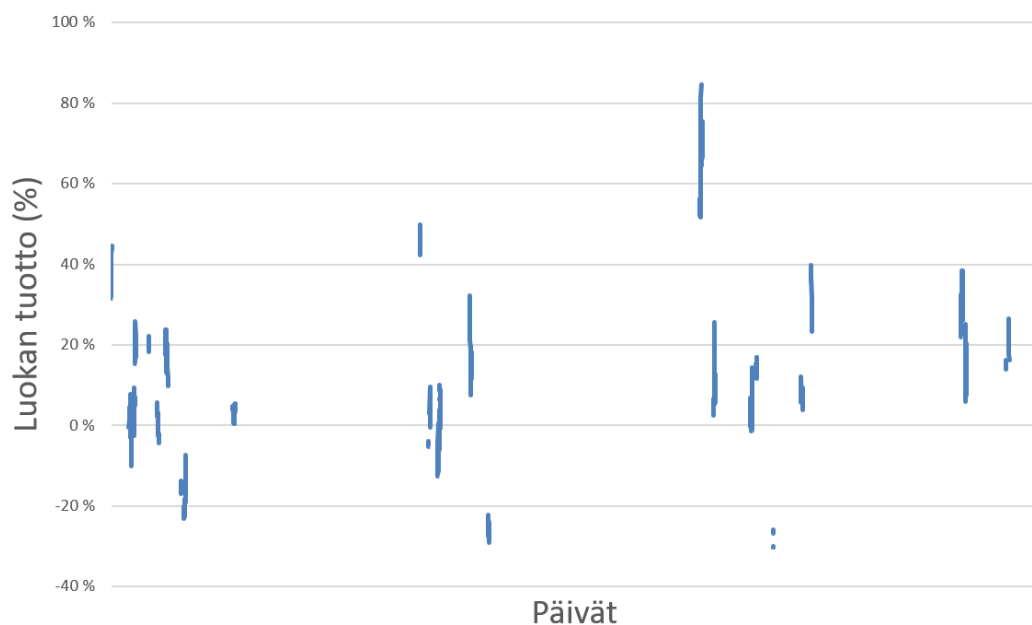
Kuvio 32. Kallis-ryhmä

Tietoja voidaan vielä koota suurempiin luokkiin parempaa yleiskuvaa varten. Yhtiöiden tuloksien ollessa tappiollinen on ollut parempi aika sijoittaa yhtiöihin kuin tuloksen ollessa voitollinen. Liikevoiton noustessa seuraavalla kvartaalilla tuotto on parempi kuin liikevoiton laskiessa. Kvartaalin ajankohdalle ei ole keskimäärin ollut suurtaakaan vaikutusta tuottoennustuksiin. Tuloksen muutoksella viimeisen 12kk aikana ei ole ollut juurikaan vaikutusta tuottoennusteisiin. Momentumin osalta tuottoennuste on suurempi silloin, kun tuotto on ollut positiivista viimeisen 91 päivän aikana.

Päälluokat	Keskituotto	Määrä
Kallis	3,0 %	6688
Halpa	3,2 %	6114
Tappiollinen	5,6 %	3162
LV-	2,0 %	7009
LV+	5,9 %	8955
Alkukvartaali	3,8 %	8180
Loppukvartaali	4,1 %	7784
Tulos--	3,8 %	3846
Tulos-	4,5 %	3020
Tulos+	3,6 %	4241
Tulos++	3,5 %	4857
Momentum 1	2,9 %	3558
Momentum 2	3,3 %	6558
Momentum 3	5,6 %	5848

Kuvio 33. Keskimääräinen tuottoennuste luokittain

Yksityiskohtaisen tarkastelun avulla tuotoista saadaan enemmän tietoa. Yksittäisen luokan toteutuneita tuottoja tarkastellessa alla olevassa kuviossa nähdään, että tuotot on ryhmän osalta ollut enimmäkseen saman suuntaisia. Kuviossa on esitetty luokkaan kuuluvien päivien toteutuneet tuotot, jossa tuottojen hajonnan voidaan nähdä olevan suurta. Vastavuoksuja löytyy jokaisesta yhtiöstä ja tuottoennuste vaikuttaa visuaalisen tarkastelun perusteella houkuttelevalta.



Kuvio 34. Yksittäisen luokan toteutuneet tuotot seuraavan 91 päivän aikana

## 5.5 Raportti

Kokonaisuudesta laaditaan automaattinen raportti, johon kootaan olennaiset asiat perustiedoista, tuottoennusteen määrittelystä ja simuloinnista. Raportin avulla voidaan helpommin tunnistaa ja arvioida suorituskyvyn muutoksen kannalta olennaisia asioita. Tunnusluvut lasketaan kokonaisuudesta tai yhtiökohtaisesti. Yksittäisten ennusteiden tai muun kohinan takia ei ole tarkoitus tehdä muutoksia, vaan keskittyä kokonaisuuteen.

Korrelaatiota käytetään mittarina tuottoennusteen ja toteutuneen tuoton välillä, koska se on tavallisin mittari kahden muuttujan väliselle yhteydelle (Heikkilä 2014, 90). Korrelaatiota voidaan vertailla muiden yhtiöiden lukuihin, jolloin voidaan havaita missä yhtiöissä malli suoriutuu parhaiten tai minkä yhtiöiden ennusteisiin kannattaisi kiinnittää enemmän huomiota. Tietoja voidaan vertailla aikaisempien raporttien tietoihin.

<b>Korrelaatio</b>		
	<b>Vastaavuudet</b>	<b>Korrelaatio</b>
UPM		
Stora Enso		
Metsä Board		
Ponsse		
Voitolliset		
Tappiolliset		
Kaikki		

<b>Vastaavuudet</b>		
	<b>Min.</b>	<b>Maks.</b>
Tavoitevastaavuudet		

<b>Voitollinen</b>		
	<b>Raja-arvo</b>	<b>Muutosnopeus</b>
Kvartaalipäivä		
Tulostuotto (P/E)		
Momentum		
Liikevoiton hist. muutos		
Tuloksen tul. muutos		

<b>Tappiollinen</b>		
	<b>Raja-arvo</b>	<b>Muutosnopeus</b>
Kvartaalipäivä		
Momentum		
Liikevoiton hist. muutos		
Tuloksen tul. muutos		

Kuvio 35. Tuottoennusteraportti

Sijoitussuunnitelman tärkein asia on sijoitusvarallisuuden allokointi tuottoennusteen ja keskihajonnan perusteella. Tiedot simulointiin valitaan kolmella eri valintaperusteella. Allokointi suoritetaan kolmella erilaisella menetelmällä, jolloin syntyy yhdeksän erilaista yhdistelmää simulointia varten. Simuloinnin tulosten perusteella tehdään edelleen valinta suurimman keskimääräisen tuoton, tuottojen keskihajonnan tai jonkin näiden välissä olevan pisteen kuten esimerkiksi suurimman Sharpen luvun perusteella.

Erilaisia lopputuloksia poimitaan yhteensä 27 kappaletta, joista on valittava yksi. Väärää valintaa ei ole, kun valitaan saman keskihajontaluokan korkein tuottoennuste. Valinta riippuu sijoittajan mieltymyksestä. Valinnan mukaiset tiedot ja menetelmät otetaan käyttöön mallin ylätasolla tehtäviä sijoituspäätöksiä varten.

Perustiedot							
Tasapainoitusväli							
Kaupankäynnin kulut							

Maksimi tuotto							
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku	Alaraja	Yläraja	Potenssi
Tuottoennuste- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Sharpen luku- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Etäisyysmenetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						

Minimi keskihajonta							
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku	Alaraja	Yläraja	Potenssi
Tuottoennuste- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Sharpen luku- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Etäisyysmenetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						

Maksimi Sharpen luku							
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku	Alaraja	Yläraja	Potenssi
Tuottoennuste- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Sharpen luku- menetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						
Etäisyysmenetelmä	Tuottoennuste						
	Keskihajonta						
	Sharpen-luku						

Kuvio 36. Simuloinnin raporttipohja

## 6 TEKNINEN TOTEUTUS

Ensimmäinen versio mallista toteutettiin Excelillä, mutta ongelmana oli, että tuottoennuste pystyttiin laskemaan ainoastaan yhdelle päivälle. Toteutus sisälsi paljon manuaalisia välivaiheita, mutta kaikki olennaiset toiminnot kuitenkin pystyttiin toteuttamaan. Tuottoennusteongelma ratkaistiin laskemalla ennusteet Python ohjelmoinnin avulla, mikä mahdollisti myös paljon muunlaisia ratkaisuja. Käytännössä saman Excel taulukon laskenta toistettiin ohjelmallisesti 10 000 kertaa, josta ennusteet saatiin kahdelle yhtiölle tietyllä aikajaksolla. Ratkaisun jälkeen oli selvää, ettei suurempia teknisiä ongelmia pitäisi tulla vastaan ja automaation määrää voidaan lisätä aikaisemmin manuaalisesti suoritettuihin vaiheisiin. Suuntana on edelleen hyödyntää enemmän ohjelmointia ja käyttää vähemmän Exceliä.

### 6.1 Excel

Jokaiselle yhtiölle luodaan oma Excel-tiedosto valmiin pohjan avulla, jolloin kaikki kerätyt tiedot löytyvät juuri oikeista soluista. Tietojen syöttäminen on suoraviivaista toimintaa. Tiedostossa on kolme välilehteä, jotka ovat kvartaali, kurssitieto ja analyysi. Kvartaalivälilehdelle kirjataan perustiedot tuloslaskelmasta, taseesta ja rahavirtalaskelmasta. Tuloslaskelman eriin voidaan tehdä oikaisuja kertaluonteisten erien osalta. Kirjanpitoikäntöjen muutoksia ja merkittäviä yrityskauppoja varten on vuosineljänneksen vertailutiedot tallennettava erikseen aikaisemmin raportoitujen tietojen lisäksi. Oikaisujen ja vertailutietojen avulla liiketoiminnan kehityksestä saadaan riittävän oikea kuva. Kurssitieto-välilehdelle haetaan tiedot historiallisesta kurssikehityksestä. Perustietojen avulla lasketaan liiketoimintaa kuvaavat tunnusluvut analyysivälilehdelle. Laskenta suoritetaan Pythonilla, koska soluissa olevia kaavoja ei pystytä hyödyntämään.

	Osakkeiden lukumäärä	Liikevaihto	Poistot	Olkaisu	Liikevoitto	Olkaisu	Tulos Ennen veroja	Olkaisu	Tilikauden tulos	Pitkäaikaiset varat	Liikearvo	Lyhytaikaiset varat	Vaihto-omaisuus	Myyntisaamiset	Taseen loppusumma	Oma pääoma	Pääomalainat	Pitkäaikaiset velat	Lyhytaikaiset velat	Ostovelat	Korolliset velat	Liiketoiminnan rahavirta	Investointien rahavirta	Rahoituksen rahavirta	Rahavarat
Q1 2001																									
Q2 2001																									
Q3 2001																									
Q4 2001																									

Kuvio 37. Esimerkki kvartaalikohtaisista tiedoista

Kurssitieto-välilehdelle haetaan päiväkohtaiset tiedot osakekurssista, joko manuaalisesti osoitteesta <http://www.nasdaqomxnordic.com/> tai ohjelmallisesti Yahoo Finance palvelusta. Historialliset tiedot on yhtenäistettävä satunnaisten erojen osalta kaupankäyntipäivissä. Kurssitiedon avulla lasketaan kurssikehitystä kuvaavat tunnusluvut ja pörssin arvostustasoa kuvaavat tunnusluvut. Haettujen tietojen mukana tulee myös muita kaupankäynnin tietoja, joita on mahdollista hyödyntää.

Date	Bid	Ask	Opening price	High price	Low price	Closing price	Average price	Total volume	Turnover	Trades
25.1.2021	30,73	30,76	31,32	31,39	30,66	30,75	30,895	969061	29946529,86	4763
26.1.2021	30,81	30,84	30,6	30,96	30,51	30,88	30,778	1661921	51149168,37	5735
27.1.2021	30,01	30,04	30,72	30,73	29,75	30,1	30,13	2115020	63719534,35	9095
28.1.2021	29,87	29,9	30,1	30,32	29,56	29,9	29,923	1468998	43957649,55	7734
29.1.2021	29,57	29,6	29,8	30,06	29,11	29,49	29,605	1880457	55675405,51	7342

Kuvio 38. Kurssitieto esimerkki

Kvartaalikohtaiset liiketoimintaa kuvaavat tiedot on kohdistettava vastaamaan päiväkohtaista kurssitietoa. Kohdistuspäivä saadaan yhtiön tiedotehistoriasta yleensä kerran vuodessa julkaistavasta taloudellisten raporttien aikataulusta. Kurssitiedot ja tilinpäätöstiedot kohdistetaan analyysi-välilehdelle oikeille päiville. Analyysi-välilehdelle lisätään myös tiedot osingoista ja osakeanneista.

Tietojen manuaalisen kirjauksen lisäksi Exceliä voidaan käyttää tietojen lukemiseen ja tarkistamiseen. Kehitystyössä Excel on hyvä apuväline, koska kaavioiden tekeminen on helppoa ja nopeaa. Väliaikaiset datan visualisoinnit ja uusien kehitysideoiden ensimmäiset prototyypit on usein suhteellisen helppoa tehdä ensin Excelillä. Myöhemmissä vaiheissa samat ominaisuudet toteutetaan ohjelmallisesti. Tiedot tallennetaan Python moduulien tukemaan \*.xls -tiedostomuotoon, kun tietoja tarkastellaan itse. Nopeaa lukemista ja tallennusta varten käytetään \*.csv -tiedostomuotoa.

## 6.2 Ohjelmointi

Opinnäytetyössä ohjelmointia käsitellään vain pintapuolisesti käymällä läpi tärkeimmät asiat opinnäytetyön kannalta. Ohjelmointi suoritetaan Python ohjelmointikielellä ja ohjelmat kirjoitetaan IDLE-tekstieditorilla. Apuna käytetään valmiita alla lueteltuja moduuleita, jotka nopeuttavat ja yksinkertaistavat ohjelmointia. Ohjelmoinnissa lähettiin liikkeelle käyttämällä perustasoisia toimintoja, jonka jälkeen on siirrytty kohti kehittyneempiä menetelmiä.

Moduulit:

- Pandas
- Math
- Statistics
- Time
- Concurrent futures
- Matplotlib
- Tkinter

Ohjelman perustoimintoja on laskea tunnusluvut oikeilla kaavoilla. Oikea etumerkki tunnistetaan if-else-valintarakenteella, jonka avulla myös vertaillaan raja-arvojen ja tunnuslukujen suuruutta. Luvuista muodostetaan listoja, jolloin niiden käsitteleminen ja hallinnointi on yksinkertaisempaa. For-silmukkaa käytetään toistuvien toimenpiteiden toteuttamiseen. Toistuvia toimintoja varten ohjelmoidaan omia funktioita, joita kutsutaan tarvittaessa. Funktioihin viedään tietoa ja lopputuloksena niistä palautetaan yleensä lista tai useampia listoja seuraavaa vaihetta varten. (Python www-sivut 2021)

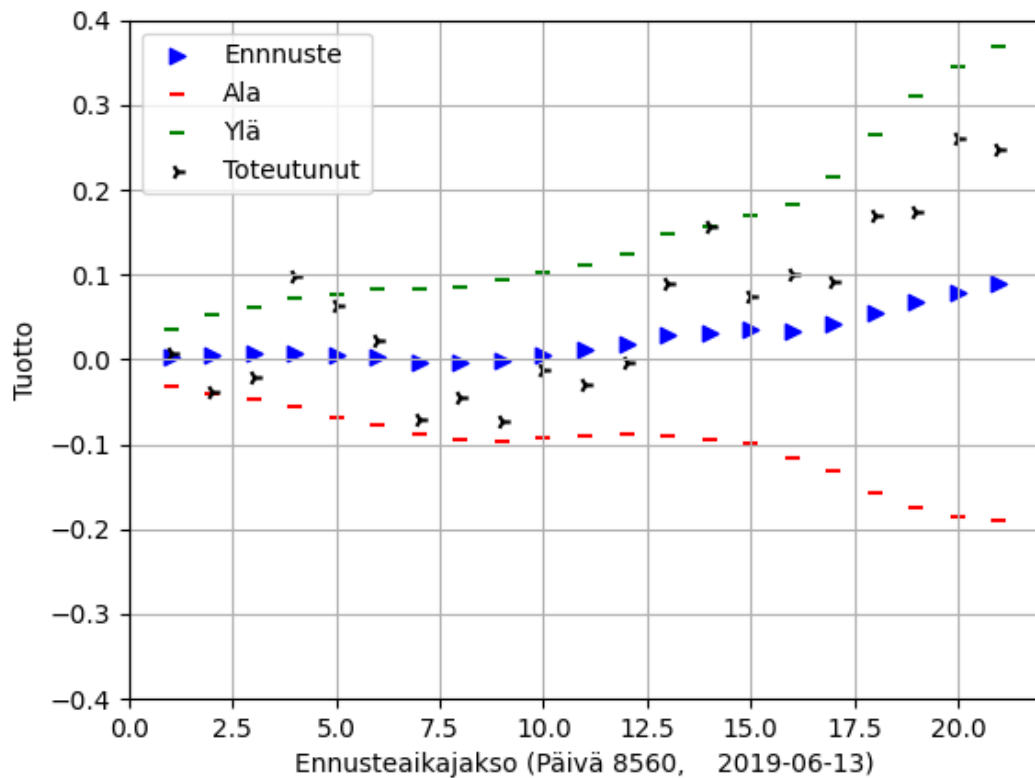
Ohjelmallisesti käytettävät tiedostot luetaan ja tallennetaan \*.xls tai \*.csv -tiedostoina käyttämällä Pandas-moduulin read\_csv, to\_csv, read\_excel ja to\_excel -funktioita. Puutteellisten tietojen osalta rivit poistetaan dropna-funktiolla ennen tietojen yhdistämistä uuteen tiedostoon concat-funktiolla. Yksittäisen sarakkeen keskiarvon laskemiseen käytetään mean-funktiota. Arvon lukeminen tai kirjoitus yksittäisiin soluihin suoritetaan joko at- tai iat-funktiolla. (Pandas www-sivut 2021)

Math-moduulin avulla suoritetaan matemaattisia toimenpiteitä kuten neliöjuuren laskenta, potenssiin korottaminen ja lukujen pyöristäminen. Statistics-moduulin avulla lasketaan keskiarvo mean-funktiolla ja keskihajonta stdev-funktiolla silloin, kun tiedot ovat kerätty listaksi. (Math [www-sivut 2021](#), Numpy [www-sivut 2021](#))

Time-moduulin avulla määritellään suorituksen alku- ja loppuajankohta time-funktiolla. Ajankohtien perusteella lasketaan toimenpiteen kesto vähentämällä loppuajankohdasta alkuaajankohta. Päivämäärän osalta käytetään datetime-moduulin fromtimestamp-funktiota. Päivämäärää käytetään lisäinformaationa optimointitietojen tallennuksen yhteydessä sekä kurssitiedon hakemisen ja tallennuksen yhteydessä. (Time [www-sivut 2021](#), Datetime [www-sivut 2021](#))

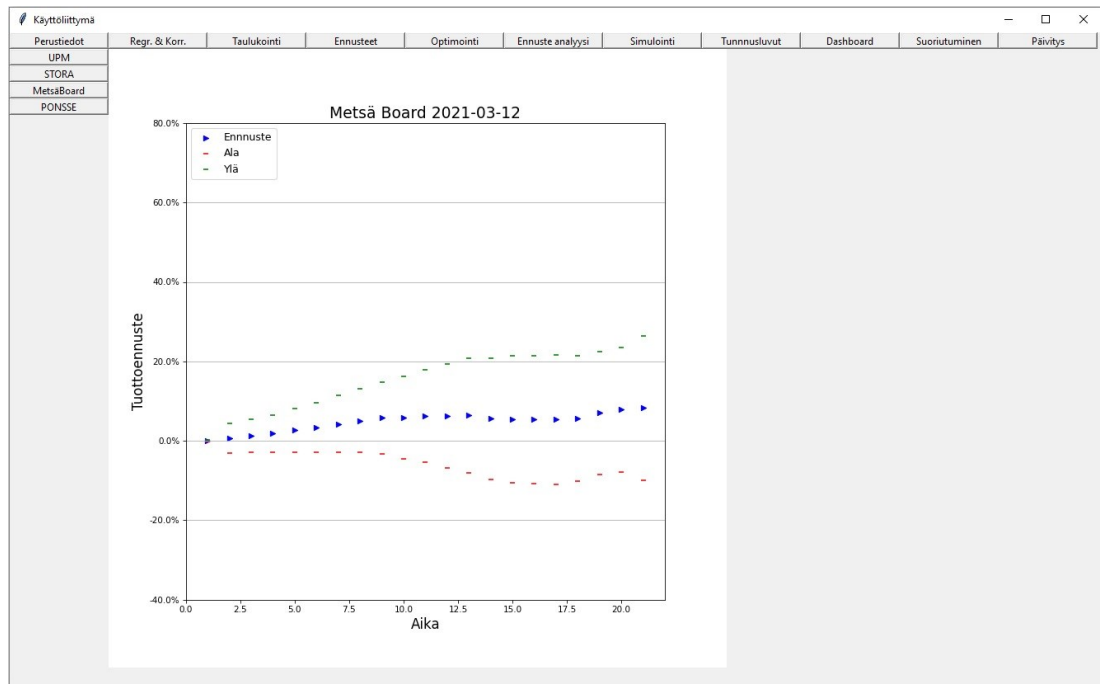
Concurrent futures-moduulilla on merkittävä vaikutus koneoppimisen nopeuteen. Moduulin avulla suoritetaan samanaikaisesti useita prosesseja. Kahdeksaa samanaikaista prosessia suorittamalla keskimääräinen tuottoennusteen laskemiseen käytetty aika tehostui 85 %. (Concurrent futures [www-sivut 2021](#))

Matplotlib-moduulin avulla tietoja voidaan esittää visuaalisesti. Lopputuloksena voi olla yksittäinen kuvio tai animaatio. Alla olevassa kuviossa havainnollistetaan yhden päivän ennustetta ja toteutunutta tuottoa. Ala-tiedossa tuottoennusteesta on vähennetty keskihajonta ja ylä-tiedossa tuottoennusteeseen on lisätty keskihajonta. (Matplotlib [www-sivut 2021](#))



Kuvio 39. Matplotlib visualisointi

Tkinter-moduulin avulla kaikki aikaisempi yhdistetään graafiseksi käyttöliittymäksi, jolloin kokonaisuus on helpommin käsiteltävissä. Käyttöliittymässä on mallin alataso, keskitaso ja ylätaso. Kuvioiden ja taulukoiden avulla voidaan esittää halutut tiedot. Yläreunassa olevien painikkeiden avulla voidaan siirtyä perustietojen lukemisesta myöhempisiin vaiheisiin. Viimeisenä vaiheena sijoituspäätösten tekemisen jälkeen on suoriutumisen seuraaminen ja analysointi. Alla olevassa kuviossa oikealla on havainnollistettu varallisuuden kehitystä ja neljän yhtiön painoarvojen muodostumista tuottoennustemenetelmällä. (TkDocs www-sivut 2021)



Kuvio 40. Käyttöliittymän ennustenäkymä

## 7 LOPPUTULOS

Kokonaisuus koostuu useista osa-alueista, joten arvioinnissakin huomioidaan tärkeimmät osa-alueet. Ennustamisen osalta oppimisessa hyödynnetään korrelaatiokerrointa ja sitä voidaan käyttää myös arvioinnin mittarina, kuten on tehty myös joissakin muissa tutkimuksissa (Juvonen 2020). Simuloinnin arvioinnissa tuottojen vertailutietona käytetään tasahajautuksen keskimääräistä tuottoa, tuottojen keskihajontaa ja Sharpen lukua. Teknisten muutosten avulla on jo saatu aikaan huomattavia parannuksia laskentaan kuluvaan aikaan.

### 7.1 Validiteetti

Validiteetilla tarkoitetaan sitä, miten käytetty mittaus- tai tutkimusmenetelmä mittaa tutkittavaa ilmiötä. Tutkimuksen kohderyhmän ja tutkimuskysymysten on oltava oikeat. Tutkimusotteen on sovittava ilmiön olemukselle ja kysymyksen asettelulle. Käsitteiden pitäisi olla tutkijayhteisön sopimia, jolloin käsitteitä mitataan tietyillä mitta-reilla. Huonossa tapauksessa tutkitaan eri asiaa kuin oli tarkoituksena. (Hiltunen 2009)

Tutkittavien yhtiöiden osakkeet ovat listattuina Helsingin pörssiin ja ne julkaisevat taloudelliset tiedot vuosineljänneksittäin, joten opinnäytetyön kohderyhmä on oikea. Yhtiöt ovat valittu suunnilleen samalta toimialalta ja tiedot ovat kerätty aikaisintaan vuodesta 2000 lähtien. Ensimmäisten tuottoennusteiden aloitusvuodet ovat 2002, 2003, 2004 ja 2006. Tulokset saattavat rajausten takia näyttää turhan hyviltä, eivätkä välttämättä ole riittävän yleistettävissä. Perustietojen hyödyntämisessä on otettu vasta ensimmäiset askeleet. Tarkoituksena oli kehittää yhtenäinen ja looginen kokonaisuus sijoituspäätösten tekemiseen. Perustiedoista siirrytään loogisesti vaiheittain kohti sijoituspäätöksiä. Yksityiskohdista on kuitenkin löydettävissä myöhemmin kehitettäviä parannuskohteita. Tunnuslukujen sijaan tutkimuksissa puhutaan yleisemmin faktoista, mutta kyse on kuitenkin samoista asioista. Keskimääräisen tuoton ja keskihajan hyödyntämisessä on mukailtu portfolioteorian ja CAPM-mallin menetelmiä.

Opinnäytetyössä käytetyt menetelmät sopivat saumattomasti yhteen. Eteneminen yhdestä vaiheesta seuraavan on suoraviivaista eikä vaadi käyttäjän tekemiä toimenpiteitä tai arviointia. Tekoälyn, koneoppimisen ja neuroverkkojen avulla ratkaistaan monimutkaisia ongelmia, joiden osaamisen ja ymmärryksen tasoa on edelleen lisättävä. Niiden avulla etsitään pääasiassa parasta lähestymistapaa tuottoennusteen määrittelyyn. Tekoäly on syntynyt yli 70 vuotta sitten, koneoppiminen on syntynyt 40 vuotta sitten ja neuroverkot ovat syntyneet vasta 10 vuotta sitten, joten kokonaisuuden ymmärtäminen on aikaa vievää (Nvidia [www-sivut](http://www.nvidia.com) 2021). Simuloinnin menetelmät ovat yleisesti käytettyjä toteutuneiden tuottojen mittaamiseen.

Teorioita ei suoraan seurata, mutta tehtävät valinnat ovat kuitenkin hyvin saman tyyppisiä, kuten aikaisemmin opinnäytetyössä on esitetty. Osakokonaisuuksien yhteensovittamisen kannalta arviointi on hankalaa, joten paras mittari on vertailla simuloinnin tuloksia. Pienen aineiston perusteella on nähtävissä selviä eroja eri menetelmien välillä.

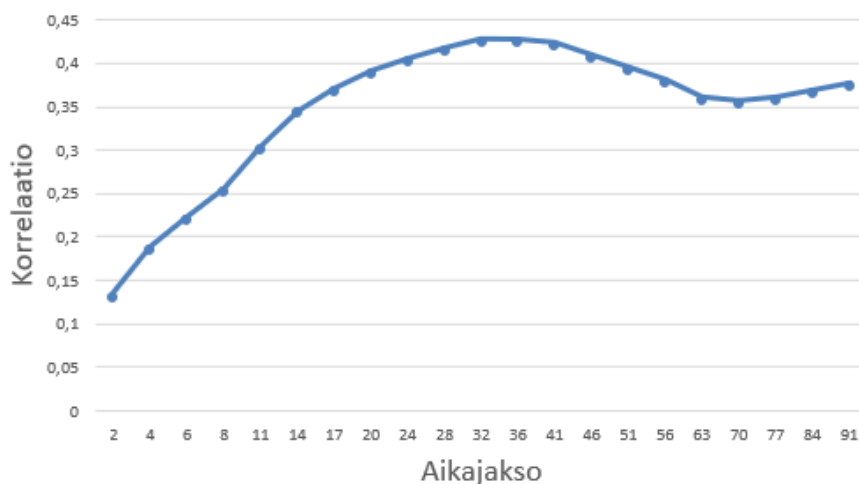
Portfolioteorian mukaisia lähestymistapoja ei useinkaan käytetä, koska käytetyt menetelmät vaativat paljon laskentaa ja menetelmät ovat herkkiä pienille muutoksille. Täysin valmis lopputulos ei opinnäytetyötä aloitettaessa ollut tavoitteena, vaan saada menetelmät jollekin ennalta määrittelemättömälle tasolle, joka on nähdäkseni selvästi

saavutettu. Jatkuvan parantamisen myötä moni asia saattaa jossakin vaiheessa muuttua. Tulevaisuudessa menetelmiä tullaan kehittämään edelleen ja aineiston määrää kasvatetaan, koska lähestymistapa näyttää edelleen hyvin lupaavalta. Kokonaisuutena opinnäytetyön validiteettia pidän hyvänä.

## 7.2 Reliabiliteetti

Reliabiliteetilla tarkoitetaan tutkimusmenetelmän luotettavuutta ja toistettavuutta. Virheen määrä ilmoitetaan usein reliabiliteetin yhteydessä. Stabiliateetti ja konsistenssi voidaan erotella reliabiliteetista. Stabiliateetilla tarkoitetaan mittarin pysyvyyttä ajan muuttuessa. Konsistenssilla tarkoitetaan useista väittämistä koostuvan mittarin jakamista kahteen, jolloin niiden pitäisi mitata samaa asiaa. (Hiltunen 2009)

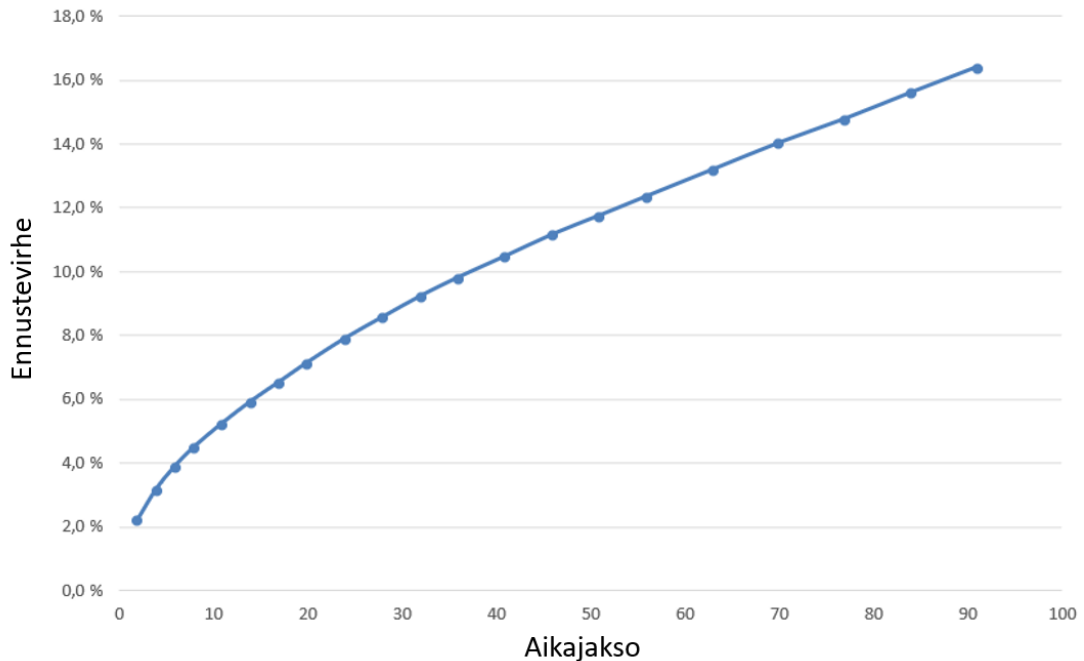
Tuottoennusteen ja toteutuneen tuoton korrelaatio vaihtelee eri aikajaksoilla. Lyhyellä aikajaksolla ennustaminen on selvästi epävarmempaa kuin, pidemmällä aikajaksolla. Ennustemenetelmä vaikuttaa vähän liiankin hyvältä, mikä saattaa johtua aineiston pienestä määrästä. Tavoitteena on ollut kehittää ja sovittaa yhteen menetelmiä, mikä näyttää korrelaation perusteella onnistuneen hyvin.



Kuvio 41. Korrelaatio

Keskimääräinen ennustevirhe riippuu aikajakson pituudesta. Lyhyimmällä aikajaksolle se on 2,2 % ja pisimmällä aikajaksolla 16,4 %. Kovinkaan tarkoiksi ennustuksia ei voida sanoa. Keskimääräinen ennustevirhe nousee alla olevassa kuviossa ensin

nopeammin kuin aikajakso, mutta aikajakson pidentyessä ennustevirheen kasvunopeus hidastuu aikajaksoa nopeammin.



Kuvio 42. Ennustevirhe

Vaikka ennustevirhettä esiintyykin runsaasti, niin sijoitukselle voidaan silti saada hyvä tuotto. Sijoitukset toteutetaan todennäköisyyksien osoittaessa selkeästi positiivista tuottoennustetta, jolloin on todennäköisempää, että myös toteutunut tuotto on positiivista (Marks 2018).

Positiivista tuottoa ennustettaessa tuoton suunta on ennustettu oikein 8375 kertaa ja väärin 1152 kertaa 91 kaupankäyntipäivän ajanjaksolla. Negatiivisen tuoton ennustaminen ei ole yhtä onnistunutta. Tuoton suunta on ennustettu oikein vain 1698 kertaa ja väärin 4692 kertaa. Suunta on ennustettu oikein 63 % kaikista päivistä, mitä voidaan kuitenkin pitää sijoittamisessa hyvänä tarkkuutena. Suunnan ennustamisen tarkkuus on hieman huonompaa lyhyemmällä aikajaksolla.

	Oikein	Väärin
Positiivinen	8376	1152
Negatiivinen	1698	4692

Kuvio 43. Sijoituksen suunnan ennustaminen

Käytännössä ongelmallisia ovat päivät, jolloin sijoitus oikeasti tehdään ja tuoton suunta on ennustettu väärin. Tuottoennusteen ollessa lähellä nollaa on paljon todennäköisempää, että tuoton suunta ennustetaan väärin. Vaikutus sijoituksen lopputulokseen riippuu lisäksi määritellystä painoarvosta, jonka avulla osittain vähennetään negatiivisen tuoton mahdollista toteutumista.

Toteutunut tuotto on jakautunut suunnilleen normaalijakauman mukaisesti tuottoennusteen ympärille. Toteutuneista tuotoista 68,2 % on alle yhden keskihajonnan päässä tuottoennusteesta. Yhden keskihajonnan etäisyys on esitetty esimerkiksi kuviossa 39. Kahden keskihajonnan päässä tuottoennusteista on 94,5 % tuottoennusteista ja kolmen keskihajonnan päässä tuottoennusteesta on 99,4 % tuottoennusteista. Toisinaan ennustevirhe on siis hyvinkin suuri.

Kuviossa 44 on esitetty tähän mennessä saatuja tuloksia, jossa tasahajautuksen tuloksia käytetään vertailutietona. Tasahajautuksella, päivittäin tasapainottamalla ja ilman kaupankäynnin kuluja keskimääräinen vuosituotto on 10,9 %, keskihajonta 31,7 % ja Sharpen luku 0,34. Simuloinnin tuloksista on poimittu tiedot: maksimi tuottoennusteen, minimi keskihajonnan ja maksimi Sharpen luvun perusteella.

Tasahajautuksen keskimääräinen tuotto on ylittynyt 12 menetelmällä 18 menetelmästä. Tasahajautuksen tuottojen keskihajonta on alittunut 14 menetelmällä 18 menetelmästä. Tasahajautuksen Sharpen luku on ylittynyt 17 menetelmällä 18 menetelmästä. Näiden perusteella kokonaisuus on erittäin onnistunut, koska voidaan valita menetelmät, joilla olisi saavutettu parempi lopputulos kuin tasahajautuksella. Kahdeksalla menetelmällä saavutettiin korkeampi keskimääräinen tuotto ja pienempi tuottojen keskihajonta kuin tasahajautuksella, jolloin myös Sharpen luku on korkeampi.

Tasahajautus			
Menetelmä	Keskituotto	Keskihajonta	Sharpen luku
Tasahajautus	10,9 %	31,7 %	0,34

Maksimi tuotto				
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku
Tuottoennuste- menetelmä	Tuottoennuste	35,8 %	43,3 %	0,83
	Keskihajonta	22,1 %	34,4 %	0,64
	Sharpen-luku	25,4 %	38,8 %	0,66
Sharpen luku- menetelmä	Tuottoennuste	12,9 %	26,4 %	0,49
	Keskihajonta	16,0 %	24,1 %	0,66
	Sharpen-luku	19,3 %	32,8 %	0,59

Minimi keskihajonta				
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku
Tuottoennuste- menetelmä	Tuottoennuste	29,1 %	28,3 %	1,03
	Keskihajonta	7,6 %	18,9 %	0,40
	Sharpen-luku	8,9 %	20,1 %	0,44
Sharpen luku- menetelmä	Tuottoennuste	1,4 %	4,4 %	0,31
	Keskihajonta	5,8 %	12,2 %	0,47
	Sharpen-luku	2,8 %	7,3 %	0,39

Maksimi Sharpen luku				
Menetelmä	Ennustevalinta	Keskituotto	Keskihajonta	Sharpen luku
Tuottoennuste- menetelmä	Tuottoennuste	30,2 %	28,4 %	1,06
	Keskihajonta	19,8 %	26,0 %	0,76
	Sharpen-luku	21,3 %	26,4 %	0,81
Sharpen luku- menetelmä	Tuottoennuste	7,2 %	11,9 %	0,61
	Keskihajonta	14,1 %	19,3 %	0,73
	Sharpen-luku	14,8 %	20,0 %	0,74

Kuvio 44. Simuloinnin tuloksia

Simulointien perusteella ainakin ilman kulujen huomioimista on tilastollisen mallin hyödyntämisellä mahdollista saavuttaa suurempaa tuottoa kuin tasahajautuksella. Kulujen huomioiminen saattaa kuitenkin muuttaa tilannetta huomattavasti. Sijoittajan ollessa kiinnostunut ainoastaan suurimmasta sijoitusten keskituotosta, valittaisiin ennusteajakaksista suurin tuottoennuste ja sitä vastaava keskihajonta, minkä jälkeen painoarvo määriteltäisiin tuottoennuste-menetelmällä. Samat valinnat tehtäisiin myös suurimman Sharpen luvun perusteella. Eroa olisi kuitenkin ala- ja ylärajassa painoarvon määrittelyssä.

Simuloinnin tulokset ovat yhdenmukaisia ja ylittävät tasahajautuksella saadut tulokset. Niiden perusteella menetelmien kokonaisuutta voidaan pitää luotettavana lähestymistapana. Tulokset tulevat muuttumaan aineiston määrän lisääntyessä ja kulujen huomioimisen jälkeen, joten tulokset eivät ole toistettavia. Menetelmät ovat kuitenkin

toistettavissa, mikä on opinnäytetyön kannalta olennaisempi asia. Reliabiliteetti on kokonaisuudessa hyvällä tasolla.

### 7.3 Loppusanat

Opinnäytetyön tekeminen on opettanut paljon sijoittamiseen liittyvän teorian soveltamisesta, ohjelmoinnista, ohjelmistorobotiikasta, koneoppimisesta ja tekoälystä. Monet opituista asioista eivät lopulta päätyneet osaksi opinnäytetyötä. Ohjelmoinnissa voidaan toteuttaa muutamilla riveillä valmiiden moduulien avulla monimutkaisia toimenpiteitä, mikä helpottaa huomattavasti toteuttamista. Ohjelmistorobotiikka, koneoppiminen ja tekoäly tulivat paremmin tutuiksi näin käytännön tekemisen kautta. Youtuben ja Googlen avulla on löydettävissä ratkaisu lähes jokaisen ominaisuuden tai menetelmän hyödyntämisestä, kunhan vain tietää mitä etsiä.

Sijoittamisen osalta yksittäisten tunnuslukujen tärkeyden arviointi sijoittajan näkökulmasta on lähtenyt hyvin käyntiin. Tunnuslukujen merkityksen ymmärtäminen olikin yksi opinnäytetyön alkuun panevista asioista. On erittäin vaikeaa tulkita sijoituskohdetta kokonaisuutena, kun laskettuna voi olla yli 100 tunnuslukua. Lisäksi sijoitusvarallisuus pitäisi pystyä allokoimaan tunnuslukujen perusteella. Allokointiin käytettyjä menetelmiä on edelleen tutkittava ja kehitettävä parempien menetelmien löytämiseksi. Ennustetuilla tilinpäätös tiedoilla näyttäisi olevan huomattavasti suurempi merkitys tulevaan tuottoon kuin historiallisilla tiedoilla.

Tuottoennuste eri aikajaksoille lasketaan tällä hetkellä samojen vastaavuuksien perusteella. Aikajaksot olisi kuitenkin mahdollista jakaa kahteen tai useampaan ryhmään, joille vastaavuudet haettaisiin erikseen. Kolmella ryhmällä ensimmäiseen voisi kuulua alle 20 kaupankäyntipäivän aikajaksot, toiseen seuraavat 30 kaupankäyntipäivää ja kolmanteen viimeiset 41 kaupankäyntipäivää. Ryhmäkohtaiset tunnuslukujen raja-arvot ja muutosnopeudet opeteltaisiin tämän jälkeen koneoppimisen avulla erikseen kullekin ryhmälle. Tällä tavalla tunnuslukujen raja-arvot voisivat paremmin kuvata kyseisten aikajaksojen tuottoa. Laskennan määrä tosin ainakin kolminkertaistuu jakamalla aikajaksot kolmeen ryhmään.

Kokonaisuus on sen verran onnistunut ja opinnäytetyölle asetetut tavoitteet saavutettu, että kehitystyötä jatketaan myös opinnäytetyön valmistumisen jälkeen. Uusi tietokone ohjelmointia ja varsinkin laskentaa varten on jo hankittuna. Tuore kurssitieto haetaan jo automaattisesti verkosta, jonka mukaisesti tunnusluvut lasketaan. Myös uudet en-nusteet ja toimenpide ehdotukset oikean varallisuuden allokointia varten on toteutettu. Uusien tietojen käsittelyyn kuluu aikaa noin 4 minuuttia, josta arviolta 75 % kuluu tiedostojen tallentamiseen. Tallentamista voidaan kuitenkin nopeuttaa huomattavasti käyttämällä \*.CSV-tiedostomuotoa. Osa-alueet yhdistävään käyttöliittymään on li-sätty lähes kaikki opinnäytetyössä esitetyt asia. Aluksi oikean rahan sijaan kokonai-suuden toimintaa testataan kuitenkin virtuaalisalkulla tehtävällä kaupankäynnillä, jol-loin muutoksien toteuttaminen on helpompaa.

## LÄHTEET

Advian www-sivut 2021. Viitattu 10.1.2021.  
<https://www.advian.fi/mita-on-tiedolla-johtaminen>

Alma Talentin www-sivut 2021. Viitattu 1.1.2021.  
<https://www.almatalent.fi/tietopalvelut/palvelut/yritystietopalvelut/tilinpaatosanalyysi>

Arvio, V. 2017. Numeeriset yhtälönratkaisumenetelmät. Pro Gradu -tutkielma. Jyväskylän yliopisto. Viitattu 18.4.2020  
<http://urn.fi/URN:NBN:fi:jyu-201801231313>

Concurrent futures www-sivut 2021. Viitattu 6.1.2021  
<https://docs.python.org/3/library/concurrent.futures.html>

Damodaran, A. 2017. Numbers and Narratives. Viitattu 3.1.2021. <https://www.youtube.com/watch?v=MN1RU9i3ngg>

Datetime www-sivut 2021. Viitattu 6.1.2021  
<https://docs.python.org/3/library/datetime.html>

Dung, K. 2013. The Profitability of Technical Trading System in Vietnamese Stock Market. AMK-opinnäytetyö. Kajaanin ammattikorkeakoulu. Viitattu 6.1.2021  
<http://urn.fi/URN:NBN:fi:amk-2013121721534>

Fama, E & French K. 1993. Common Risk Factors in the Returns On Stocks And Bonds. Viitattu 10.1.2021.  
<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.139.5892>

Fama, E & French K. 2013. A Five-Factor Asset Pricing Model. Viitattu 10.1.2021.  
<https://ssrn.com/abstract=2287202>

Figuroa, A. 2019. Data Demystified – DIKW model. Viitattu 12.1.2021.

<https://towardsdatascience.com/rootstrap-dikw-model-32cef9ae6dfb>

Francis, J & Kim, D. 2013. Modern Portfolio Theory: Foundation, Analysis and New Developments + website. Wiley. New Jersey: Wiley

Heikinmatti, K., Jahkonen, E., Kanervisto, M., Kekki, S., Marjomaa, J., Ruusulaakso, J. & Toivio, A. 2017. Yritystutkimus ry: Yritystutkimuksen tilinpäätösanalyysi. Helsinki: Gaudeamus

Heikkilä, Tarja. 2014. Tilastollinen tutkimus. Edita

Hiltunen, L. 2009. Validiteetti ja reliabiliteetti. Viitattu 16.1.2021.

<https://docplayer.fi/20584479-Validiteetti-ja-reliabiliteetti-leena-hiltunen-gradu-ryhma-18-2-2009.html>

Holopainen, T. 2012. Klusterointi kombinatorisilla ja hierarkisilla meneelmillä. Pro Gradu -tutkielma. Jyväskylän yliopisto. Viitattu 18.4.2020  
<http://urn.fi/URN:NBN:fi:jyu-201203301498>

Hovi, A., Hervonen, H. & Koistinen, H. 2009. Tietovarastot ja business intelligence. Jyväskylä: WSOYpro, Docendo.

Härmä, A. 2010. Tilastollisten menetelmien soveltaminen tuotekehitysympäristössä: Powerwave Finland Oy. AMK-opinnäytetyö. Oulun seudun ammattikorkeakoulu.

Interactive Brokers www-sivut 202. Viitattu 10.1.2021.

<https://www1.interactivebrokers.com/en/index.php?f=5041>

Juvonen, J. 2020. Neuroverkot taloudellisen aikasarjan ennustamisessa: Empiirinen tutkimus S&P 500 -indeksillä. Pro gradu -tutkielma. Turun yliopisto. Viitattu 2.2.2021. <http://urn.fi/URN:NBN:fi-fe2020040812002>

Kaarlejärvi, S. & Salminen, T. 2018. Älykäs taloushallinto: automaation aika. Helsinki: Alma Talent.

Larikka, M., Heinilä, P., Selin, K. & Tuominen, J. 2007. Tuottavuuden Jatkuva Parantaminen. Helsinki: Teknova.

Markowitz, H. 1952. Portfolio Selection. Viitattu 10.1.2021.  
<https://www.jstor.org/stable/2975974>

Marks, H. 2018. Mastering the market cycle: Getting the odds on your side. New York: Houghton Mifflin Harcourt.

Math www-sivut 2021. Viitattu 4.1.2021.  
<https://docs.python.org/3/library/math.html>

Matplotlib www-sivut 2021. Viitattu 17.1.2021.  
<https://matplotlib.org/3.3.3/contents.html>

Niemonen L. 2017. Käyttäytymispsykologia osana piensijoittajan päätöksentekoa. Opinnäytetyö. Haaga-Helia ammattikorkeakoulu. Viitattu 9.4.2020

Niskanen, M. 2016. Yritysrahoitus. Helsinki. Edita.

Nordnet www-sivut 2021. Viitattu 9.1.2021.  
<https://www.nordnet.fi/fi/palvelut/hinnasto>

Numpy www-sivut 2021. Viitattu 4.1.2021.  
<https://numpy.org/doc/>

Nvidia www-sivut 2021. Viitattu 2.2.2021.  
<https://developer.nvidia.com/deep-learning>

Ojasalo, K., Moilanen, T. & Ritalahti, J. 2014. Kehittämistyön menetelmät. Helsinki. Sanoma Pro.

Oppia www-sivut 2020. Viitattu 27.12.2020. <https://www.oppia.fi/courses/tietojen-mallintaminen--data-modeling/6ea333a889b62dbc7c929312531c70b4>

Pandas www-sivut 2021. Viitattu 4.1.2021.  
[https://pandas.pydata.org/docs/user\\_guide/index.html#user-guide](https://pandas.pydata.org/docs/user_guide/index.html#user-guide)

Prado, M. 2019. The 7 Reasons Most Machine Learning Funds Fail Marcos Lopez de Prado from QuantCon 2018. Viitattu 6.1.2021.  
<https://www.youtube.com/watch?v=BRUISm4gdQ4&t=2107s>

Puttonen, V. 2018. Moderni rahoitus. Alma Talent.

Python www-sivut 2021. Viitattu 6.1.2021.  
<https://docs.python.org/3.9/tutorial/index.html>

Salomäki, Rauno.1999. Suorituskykyiset prosessit- hyödynnä SPC. Metalliteollisuuden kustannus

Sandholm, J. 2020. Organisaation datalukutaidon arvioiminen. AMK-opinnäytetyö. Haaga-Helia. Viitattu 12.1.2021  
<http://urn.fi/URN:NBN:fi:amk-2020090220045>

Sharpe, W. 1966. Mutual Fund Performance. Viitattu 10.1.2021.  
<https://www.jstor.org/stable/2351741?seq=1>

Suomen Pankin www-sivut 2021. Viitattu 1.1.2021.  
<https://www.suomenpankki.fi/fi/Tilastot/korot/taulukot2/>

Time www-sivut 2021. Viitattu 6.1.2021.  
<https://docs.python.org/3/library/time.html>

TkDocs www-sivut 2021. Viitattu 31.1.2021.  
<https://tkdocs.com/>

Towards data science www-sivut 2018. Viitattu 3.1.2021.

<https://towardsdatascience.com/python-basics-iteration-and-looping-6ca63b30835c>

Trask, A. 2019. Grokking deep learning. Manning

Tyrylahti, J. 2011. Momentum-anomalia ja siihen perustuvat sijoitus strategiat Yhdysvaltojen osakemarkkinoilla. Kandidaatintyö. Lappeenrannan yliopisto. Viitattu 7.4.2020 <http://urn.fi/URN:NBN:fi-fe201201251185>

UPM-Kymmene www-sivut 2021. Viitattu 24.1.2021.

<https://www.upm.com>

Valtioneuvoston www-sivut 2021. Viitattu 12.1.2021.

<http://urn.fi/URN:ISBN:978-952-287-549-5>

Vähähyppä, P. 2019. IFRS 16-standardin käyttöönoton vaikutukset Finnairin taloudellisiin tietoihin. Viitattu 24.1.2021.

[https://www.youtube.com/watch?v=BBKZvFjMBYk&feature=emb\\_title](https://www.youtube.com/watch?v=BBKZvFjMBYk&feature=emb_title)