

TALLINNA TEHNIKAÜLIKOOL

Majandusteaduskond

Majandusanalüüsi ja rahanduse instituut

Mehis Tambu

**S&P 500 VOLATIILSUSE ENNUSTAMINE
TEHISNÄRVIVÕRKUDE ABIL**

Magistritöö

Ärirahandus

Juhendaja: Kaido Kepp, MA

Tallinn 2020

Deklareerin, et olen koostanud töö iseseisvalt ja olen viidanud kõikidele töö koostamisel kasutatud teiste autorite töödele, olulistele seisukohtadele ja andmetele ning ei ole esitanud sama tööd varasemalt ainepunktide saamiseks.

Töö pikkuseks on 9960 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Mehis Tambu

(allkiri, kuupäev)

Üliõpilase kood: 183228TARM

Üliõpilase e-posti aadress: mehistambu@gmail.com

Juhendaja Kaido Kepp, MA:

Töö vastab kehtivatele nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(nimi, allkiri, kuupäev)

SISUKORD

LÜHIKOKKUVÕTE	1
SISSEJUHATUS	2
1. VOLATIILSUS KUI KAUPLEMISPÄEVA HINNAVAHEMIK	5
1.1. Volatiilsuse ennustamine	5
1.2. Volatiilsusega seotud terminid ja nähtused	7
1.3. Mudelite hindamine	9
2. GARCH JA TEHISNÄRVIVÕRGUD	13
2.1. GARCH	13
2.2. Tehisnärvivõrgud	14
2.2.1. Sissejuhatus tehisnärvivõrkudesse	15
2.2.2. Mitmekihiline pertseptron närvivõrk	17
3. VARASEMATE EMPIIRILISTE UURINGUTE TULEMUSED	23
3.1. GARCH empiirilised tulemused	23
3.2. Tehisnärvivõrkude empiirilised tulemused	25
4. EMPIIRILISED TULEMUSED	28
4.1. Andmed	28
4.2. Tulemused	36
4.2.1. Mudelite seadistus ja tulemused	36
4.2.2. Madala volatiilsuse ennustamine	41
4.2.3. Kõrge volatiilsuse ennustamine	43
KOKKUVÕTE	47
SUMMARY	49
KASUTATUD ALLIKATE LOETELU	51
LIHTLITSENTS	58

LÜHIKOKKUVÕTE

Volatiilsuse ennustamine on oluline riskijuhtimise, portfellide varajaotuste ja optioonide hinnastamise näol. Volatiilsuse ennustamisel on olnud levinumateks tööriistadeks autoregressiivsed mudelid. Tänu arvutite võimekuse arengule ja süvaõppe populaarsuse kasvule on võimalik välja töötada keerulisi tehisnärvivõrkude mudeleid. Käesoleva töö eesmärgiks on ennustada S&P 500 aktsiaindeksi volatiilsust mitmekihilise pertseptron tehisnärvivõrkude mudeliga. Selleks, et võrrelda kuidas erinevad tehisnärvivõrkude ennustused levinud volatiilsuse ennustamise mudelist, kasutatakse etalonina üldistatud autoregressiivse tingliku heteroskedastiivsusega protsessi, lühidalt GARCH (*Generalised AutoRegressive Conditional Heteroskedasticity*) mudelit. Töö käigus ennustatakse S&P 500 aktsiaindeksi volatiilsust, mille raames kasutatakse andmeid vahemikul 03.01.2007-28.02.2020. Mudeli konstrueerimisel, kasutatakse 3223-st vaatlusest 2500 vaatlust, sellest tulenevalt on valimiväliseks volatiilsuse ennustusteks 723 vaatlust vahemikul 05.12.2016-28.02.2020. Lisaks kogu valimivälise ennustustele jagatakse antud 723 vaatlust kvartiilide abil kahte gruppi: madala volatiilsusega vaatlused ja kõrge volatiilsusega vaatlused. Töö tulemusena selgus, et kõigi valimivälise vaatluste ja kõrge volatiilsusega vaatluste puhul on täpsemate ennustustega tehisnärvivõrkude mudel. Seejuures madala volatiilsuse ennustamisel oli GARCH mudel märkimisväärselt täpsemate ennustustega kui tehisnärvivõrkude mudel aga kuna kõrge volatiilsuse ennustamise täpsus on äärmiselt olulisem, siis antud töö järelduseks on, et tehisnärvivõrgud on täpsemad volatiilsuse ennustajad. Lisaks leidis autor kinnitust asümmeetrilisest volatiilsusest ja paksu saba tõenäosusjaotustest, volatiilsuse kuhjumisest ja võimenduse efektist ning keskväärtusele koondumisest.

SISSEJUHATUS

Iga modelleerimise probleemi lahendamise eelduseks on, et ajaloolised andmed sisaldavad või peegeldavad vähemalt mingil määral informatsiooni tuleviku andmete kohta. Finantsvarade volatiilsuse ennustamise puhul eeldame, et ajalugu kordab ennast hinnamustrite näol. Antud hinnamustreid on võimalik modelleerida, kuid teinekord ei piisa ainult vaadeldava vara ajaloolistest hindadest, sest tihtipeale on ennustatavad finantsvarad tugevas seoses teiste finantsvarade, makroökonomiliste näitajate ja erinevate mõõdikutega. Volatiilsuse ennustamine ja volatiilsust ennustavate mudelite kasutamine on viimastel aastakümnetel laialdaselt populaarsust kogunud. Seda eelkõige 1980. aastate aktsiate kriisi, 2000. aastate internetiettevõtete mulli ja 2007-2008 finantskriisi tulemusena. Nimelt sellised sündmused tekitavad suurt volatiilsust. Kui volatiilsus põhjustab investori varade hinnaliikumist ebasoodsas suunas, siis kujuneb see investori jaoks riskiks. Teades, milliseks võib järgmise päeva volatiilsus kujuneda, annab see turuosalistele võimaluse reageerida vastavalt ning minimeerida ebasoodsas suunas liikuvate varadega seotud kaotusi. Investorid ja teised finantsturgudel osalejad kasutavad volatiilsuse ennustamist riskide haldamiseks, portfelli varaklasside jaotamisel ning riski ja kasumi optimeerimiseks tehingute ajastamise näol.

Tänaseks on levinuimaks volatiilsuse ennustamise meetoditeks kujunenud 1980. aastate algul Robert Engle poolt välja arendatud autoregressiivsed mudelid. Antud mudelite sisenditeks on peamiselt hinna ja tehingu andmed ning teiste finantsinstrumentide volatiilsus. Alternatiiviks autoregressiivsetele mudelitele on stohhastilise volatiilsuse mudelid, mis erinevalt autoregressiivsetest mudelitest eeldavad konstantset volatiilsust. Vähemlevinud alternatiiviks volatiilsuse ennustamisel on tehisnärvivõrgud, mis on viimasel ajal masinõppe ja süvaõppe populaarsuse kasvuga muutunud väga levinuks tööriistaks erinevate probleemide lahendamisel.

Sentimenti tõlgendatakse tihtipeale kui tunnetuslikku hinnangut. Finantsturgude kontekstis peegeldab sentiment turuosaliste arvamust või hoiakut teatud vara suhtes. Kui turuosaliste hoiakud või arvamused muutuvad, siis peaks see peegelduma ka finantsinstrumentide hindade muutuses. Hoiakud ja arvamused on psühholoogilised nähtused, mis põhinevad emotsioonidel. Investorite hirm on täiendav emotsioon, mis külvab hoiakuid ja arvamusi suunal, et finantsinstrumentide hinnad võivad langeda. Töö järgnevatest peatükkidest leitakse, et tihtipeale tähendab volatiilsus just hindade langust. Lähtudes eelnevast defineerib autor

sentimendi indikaatoritena mõõdikuid, mis peegeldavad segmentide või kogumi üldist hoiakut või arvamust. Levinuim selline indikaator volatiilsuse ennustamisel on Chicago Board of Trade (CBOT) Volatility Index (VIX), mis peegeldab tuletisinstrumentide volatiilsust. Antud mõõdikut tuntakse ka hirmuindeksi või hirmu mõõdikuna, mis ennustab järgneva 30 päeva volatiilsust. Lisaks eelnevatele sentimendi indikaatoritele on levinud alternatiiviks Google Trends'i märksõna indeks, mis põhineb Google otsingumootori märksõna otsingu populaarsusel.

Käesoleva töö eesmärgiks on ennustada S&P 500 aktsiaindeksi volatiilsust tehisnärvivõrkudega. Selleks, et aru saada, kas tehisnärvivõrkude tulemused on head või halvad, kasutatakse etalonina üldistatud autoregressiivset tingliku heteroskedastiivsusega protsessi, lühidalt GARCH (i.k *Generalised AutoRegressive Condidtional Heteroskedasticity*) mudelit. Selleks, et säilitada mudelite võrreldavust, kasutatakse mõlemat mudelit samades tingimustes ja samade andmetega. Ühtlasi on oluline teada, kuidas mudelite ennustuste täpsused käituvad erinevates turuolukordades. Varasemates töödes on kasutatud volatiilsuse ennustamisel nii GARCH ja tehisnärvivõrkude võrdlust, kui ka mõlema mudeli hübriide. Käesolev töö hakkab erineva varasematest töödelt selle poolest, et kasutab volatiilsuse ennustamisel Google Trends'i ja võrdleb ennustuste tulemusi erinevates turuolukordades. Turuolukorrad eristatakse kvartillide abil kaheks: esimese kvartiili vaatlused vahemikus 0%-25% – madala volatiilsusega ja neljanda kvartiili vaatlused vahemikus 75%-100% – kõrge volatiilsusega.

Modelleerimisel kasutatakse S&P 500 hinna ja –tehinguandmeid, teiste finantsvarade hinnaandmeid, erinevaid USA majandusolukorda peegeldavaid makroökonomilisi mõõdikuid ning varasemalt nimetatud sentimendi indikaatoreid. Järgmise kauplemispäeva volatiilsuse ennustamine on oluline, sest kui suure volatiilsusega turuolukorras on mudelite ennustused täpsed, siis võimaldab see riske paremini hallata ning vajaduse korral kohandada portfellis olevate varade osakaalusid. Lisaks riskide haldamisele võimaldab järgmise päeva volatiilsuse ennustamine optimeerida kasumeid nii õigesti ajastatud positsioonide sulgemise kui ka õigesti ajastatud positsioonide avamise näol.

Töö koosneb neljast peatükist. Esimene peatükk annab ülevaate volatiilsuse, selle ennustamise ja riski käsitlusest ning meetoditest. Teine peatükk koosneb GARCH ja tehisnärvivõrkude ülevaatest. Teise peatüki tehisnärvivõrkude alampeatükk jaguneb kaheks. Esimene

alampeatükk tutvustab ja selgitab tehisnärvivõrkude ajalugu ning teine alampeatükk annab ülevaate käesolevas töös kasutatavast mitmekihilisest pertseptron tehisnärvivõrgust. Kolmanda peatüki moodustavad GARCH ja tehisnärvivõrkude empiiriliste uuringute ülevaated varasematest tulemustest. Neljas ning ühtlasi viimane peatükk koosneb käesoleva töö empiirilisest osast, milles antakse ülevaade töös kasutatavatest andmetest, mudelite seadistustest ning üldistest ennustuste tulemustest ja erinevate turuolukordade ennustuste tulemustest.

1. VOLATIILSUS KUI KAUPLEMISPÄEVA HINNAVAHEMIK

Käesolev peatükk jaguneb kolme alampeatükki. Esimene annab ülevaate volatiilsuse ennustamisest, sealhulgas ka erinevatest volatiilsuse ennustamise meetoditest ning nende ajaloost. Teine alampeatükk annab ülevaate ja selgitab levinuimaid volatiilsusega termineid ja nähtuseid. Ühtlasi esitab autor varasematest töödest märkimisväärsemaid leide ning toob esile töid, kus on esinenud antud nähtuseid. Viimases alampeatükis kirjeldatakse moodsikuid, mille abil saab erinevate mudelite ennustusvõimet võrrelda.

1.1. Volatiilsuse ennustamine

Ennustavad mudelid jagunevad kahte gruppi: regressioonil põhinevad mudelid ja klassifitseerimisel põhinevad mudelid. Kahe grupi peamiseks erinevuseks on see, et regressioonil põhinevad mudelid ennustavad väärtust ning ennustuste erinevused tegeliku väärtusega varieeruvad. Näiteks tegelik väärtus võib olla 10 aga ennustatud väärtus 7 või 12, kus viga on vastavalt 3 või 2. Üheks levinuim regressioonil põhinev mudel on lineaarne regressioon. Klassifitseerimismudelite puhul vea suurus ei varieeru, ehk ennustus on kas õige või vale. Ühtlasi on klassifitseerimismudelite eeluduseks vähemalt kahe grupi olemasolu. Näiteks erinevate sõltumatute muutujate alusel ennustatakse kas tegemist on mehe või naisega, kus mehe grupi väärtuseks on 1 ja naise grupi väärtuseks on 0.

Volatiilsust on võimalik defineerida mitmetel viisidel. Varasemates töodes on käsitletud volatiilsust kui kauplemispäeva sulgemishinna ja eelneva kauplemispäeva sulgemishinna erinevust, ehk vara tootluse muutust käsitleti kui volatiilsust. Antud meetodi puhul on oluline märkida, et tootlus võib olla nii positiivne kui ka negatiivne, sest vara hind võib võrreldes eelmise päevaga kas kasvada või langeda. Mitmetes töodes (Garman, Klass, 1980; Rogers, Satchell, 1991; Yang, Zhang, 2000) on leitud, et kui defineerida volatiilsust kui kauplemispäeva kõrgeima ja madalaima hinna erinevust on võimalik saavutada 5-14 korda efektiivsemad tulemused võrreldes vara tootluse kasutamiselega. Lisaks eelnevale on leitud tõendeid selle kohta, et kui kasutada lisaks kauplemispäevade madalaima ja kõrgeima hinna erinevusele ennustamisel sõltumatute muutujatena kauplemispäevade avamishindu ja sulgumishindu on võimalik saavutada paremad tulemused kui ainult sulgumishindade

ennustamisel (Garman, Klass, 1980; Beckers 1983; Wiggins, 1992; Rogers, Satchell, 1991; Kunitomo, 1992; Yang, Zhang, 2000). Ühtlasi on volatiilsuse ennustamisel levinud libiseva standardhälve kasutamine (Dicle, Levendis 2019). Enamasti kasutatakse 20 või 22 eelneva päeva libiseva volatiilsuse standardhälvet. Lisaks hinnaandmetele on levinud ka tehingumahtude, volatiilsusindeksi, aktsiaindeksite ja muude finantsinstrumentide volatiilsuse kasutamine vaadeldava finantsinstrumendi volatiilsuse ennustamisel.

Olemasolevad volatiilsuse ennustamise mudelid jagunevad kahte peamisesse kategooriasse: sümmeetrilised ja asümmeetrilised mudelid. Sümmeetrilised mudelid sõltuvad tinglikust dispersioonist, mitte volatiilsuse muutumise märgist, seega antud mudelid ei erista seda, kas volatiilsuse muutus on negatiivse või positiivse märgiga. Seevastu võivad asümmeetrilised mudelid tõlgendada sama sisendiga negatiivsest ja positiivsest volatiilsusest erinevad väljundid.

Varasematest uuringutest on leitud, et GARCH mudeli ennustuste täpsus võib sisendandmete mahu kasvades nõrgeneda (Lootus, 2016). Tehisnärvivõrkude varasemates uuringutes (Tang, De Almeida, Fishwick, 1991) on leitud vastupidiseid seoseid. Nimelt leidsid Tang, De Almeida ja Fishwick (1991), et kui suurendada tehisnärvivõrkude sisendandmete mahtu, siis see kajastub positiivselt mudeli ennustuste tulemustes.

Arvestades, et ühe päeva kõrgeima ja madalaima hinna vahe ei anna täielikku ja usaldusväärset informatsiooni vara volatiilsuse kohta ühe päeva juhuslikkuse tõttu, siis sellest tulenevalt kasutatakse tihtipeale ühe aasta või ühe kuu andmeid ning volatiilsust hinnatakse kokkuvõtvalt standardhälvega. Standardhälve näitab, kui kaugel on üks vaatlus andmerea keskmisest. Standardhälve (s) arvutamisel leitakse vaatluste (x_i) ja vaatluste keskmise (\bar{x}) vahe, mis tõstetakse astmele 2 ning seejärel summeeritakse. Antud summa jagatakse vaatluste arvuga (N). Vaatluste arvust on lahutatud 1 kui tegemist on valimiga, mitte kogu populatsiooniga. Eelnevate arvutuste tulemuseks on dispersioon, ning standardhälve on ruutjuur dispersioonist.

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad (1)$$

kus

s – standardhälve

N – vaatluste arv

x_i – vaadeldava vaatluse väärtus

\bar{x} – andmerea keskmine väärtus

Volatiilsust kasutatakse riskide haldamisel, varaklasside jaotamisel ja investeerimisotsuste tegemisel. Lähtudes turu efektiivsuse teooriast (Fama, 1965) on leitud, et mida suurem on vara volatiilsus, seda suurem peaks olema ka vara riskipremia. Kui investor on valiku ees, kus valida on 1€ ja 2€ aastase volatiilsusega vara vahel, siis ideaaltingimustes peaks 2€ volatiilsusega vara tagama suurema riskipremia. Seda muidugi juhul kui vara liigub peale soetamist soovitud suunas, kui 2€ volatiilsusega vara hind hakkab langema, siis langeb see tõenäoliselt kiiremini, kui 1€ volatiilsusega vara.

1.2. Volatiilsusega seotud terminid ja nähtused

Varasemates töodes on leitud, et volatiilsuse ennustamisel esineb mitmeid spetsiifilisi nähtuseid. Käesolev alampeatükk selgitab nelja volatiilsusega seotud nähtust. Sealhulgas tuuakse esile varasemaid töid, kus antud nähtused on esinenud ning millal, kus ja millistes tingimustes on need esinenud.

Kuigi tavaliselt eeldab standardne finantsteooria, et tootlus on normaaljaotusega (st vara hind on lognormaalse jaotusega) ja volatiilsus on konstantne, siis esimest korda mainiti volatiilsuse kuhjumist 1963. aasta Mandelbrot'i uurimustöös, kus autor leidis, et suurtele hinnamuutustele kipuvad järgnema suured hinnamuutused ning väikestele hinnamuutustele kipuvad järgnema väikesed hinnamuutused. Lisaks sellele toodi välja, et kuhjumisele järgnevad muutused võivad olla positiivses või negatiivses suunas. Ding ja Granger (1993) leidsid, et aktsiaindeksi päevased tootlused ei ole korreleerunud, kuid absoluutse tootluse näol esineb vähenev autokorrelatsioon. Volatiilsuse kuhjumisele on viidatud ka teistes töodes, nagu Klaassen (2002), kus modelleeriti valuutakursside volatiilsust ajavahemikul 1978-1997 või Cheung (2000) töös, kus autor modelleeris New Yorgi aktsiaindeksi volatiilsust ajavahemikul 1962-1985. Antud nähtus on levinud ka arengumaade börsidel, näiteks Emekike (2010) leidis

tõendust volatiilsuse kuhjumise kohta Nigeeria börsilt ning Suliman ja Winker (2012) leidsid märke volatiilsuse kuhjumisest Sudaani ja Egiptuse börsilt.

Niinimetatud „paksu sabaga“ jaotus on tõenäosusjaotus, mida iseloomustavad kõrge väärtusega kurtoosus (*kurtosis*) ehk püstakus, mis tähendab seda, et andmekogumi andmepunktid on akumulunud andmekogumi keskmise lähedale. Visualiseerimisel iseloomustab antud jaotust kõrge ja peenike jaotus. Võrreldes normaaljaotusega tähendab see seda, et äärmuslike väärtuste tõenäosuslikkus on suurem. Finantsandmete puhul on tihtipeale tegemist leptokurtoosusega, mis on positiivse jäägiga kurtoosus. Mandelbrot'i (1965) uurimustöö pani suure kahtluse alla seni teada olnud vara hindade kausaalsuse ja normaaljaotuste valiidsuse. Peale põhjalikke testimisi leidis Mandelbrot, et eksponentjaotused pakuvad paremat stabiilsust kui normaaljaotus. Selgituseks toodi välja asjaolu, et jaotuste „sabad“ peaksid olema pikemad, mis tähendab, et jaotuste äärtes peaks olema suurema sagedusega vaatlused – „paksud sabad“.

Asümmeetriline volatiilsus rahanduse kontekstis tähendab ebahürtlaste suurustega hinnaliikumisi, ehk volatiilsus ei ole konstantne. Näiteks 10€ maksev aktsia võib kukkuda 0€ peale, millest enam madalamaks ei saa minna. Seevastu 10€ pealt kõrgemale liikudes on aktsia kasvuvõimalused piiratud. On leitud, et volatiilsus on hindade langedes suurem kui hindade tõustes. Varasemates töodes (Lins, 2009; Dzielinski, Rieger ja Talpsepp, 2011) leiti, et volatiilsuse asümmeetria peamiseks põhjuseks on varaga seotud halvad uudised ning investorite ülereageerimine nendele uudistele. Ühtlasi tõid nad oma töös esile, et Google märksõna populaarsusindeks on hea volatiilsuse ennustaja koos märksõnaga majanduskriis (*recession*). Hea ennustajana pidasid Seega Dzielinski, Rieger ja Talpsepp silmas seda, et kui volatiilsus kasvab, siis kasvab ka Google Trends'i indeksi väärtus. Antud indeks näitab märksõna otsingu populaarsust aegridade indeksina väärtustega 0-100, kus 0 tähistab minimaalset populaarsust antud ajavahemikul ning 100 tähistab maksimaalset populaarsust antud ajahetkel.

Volatiilsuse asümmeetriaga on seotud termin võimenduse efekt. Cont (2001) kirjeldas võimenduse efekti kui nähtust, kus volatiilsus on negatiivses korrelatsioonis vara tootlusega. Seega kui vaadeldava vara volatiilsus kasvab, siis tõenäoliselt on antud vara tootlus negatiivne, ehk vara hinnad langevad. Selle tulemusena kasvab laenukapitali/omakapitali suhte D/E suhtarvu tõus, mis omakorda muudab aktsia omamise riskantsemaks (Indla, 2019). Antud

nähtus on esinenud erinevates volatiilsuse ennustamise töödes, sealhulgas nii arenenud turgudel (Bouchaud, Matacz, Potters, 2001) kui ka arengumaade turgudel (Selcuk, 2005). Viimases leiti volatiilsuse efekti nii Lõuna-Ameerika riikidest kui ka Ida-Aasia arengumaadest. Võrdluses arengumaadega leiti, et volatiilsuse efekt on arengumaades veelgi suurem.

Keskväärtusele koondumine (i.k. *reversion to the mean*) on statistikas laialdaselt levinud nähtus. Finantsinstrumentide volatiilsuse mõistes tähendab see, et kõrge volatiilsusega perioodidele järgneb volatiilsuse taandumine keskmise volatiilsuse juurde ning madala volatiilsusega perioodidele järgneb samuti koondumine keskmise volatiilsuse juurde. Bali ja Demirtas (1997) leidsid, et aastatel 1982-2002 esineb S&P 500 aktsiaindeksi volatiilsuse keskväärtusele koondumist. Antud nähtust on võimalik näha käesoleva töö empiirilise osa joonistelt, kus on selgelt näha, et kõrge volatiilsusega kauplemispäevad taanduvad teatud hetkel keskmise volatiilsuse tasandile.

1.3. Mudelite hindamine

Käesoleva töö eesmärgiks on ennustada S&P 500 aktsiaindeksi volatiilsust tehisnärvivõrkudega. Selleks, et mõista tehisnärvivõrkude ennustusi, võrreldakse ennustuste tulemusi traditsioonilise GARCH mudeli ennustustega. Sellest tulenevalt on vaja teada mõõdikuid, mille alusel mudelite ennustuste täpsust hinnata. Mõõdikutena kasutatakse mudeli jääkliikmete ja tegelike väärtuste erinevuse vea mõõdikuid. Käesoleva töö empiirilises osas kasutatakse järgnevat mõõdikuid:

- Keskmise absoluutviga (*mean absolute error* – MAE)
- Ruutkeskmise viga (*mean squared error* – MSE)
- Keskmise suhteline absoluutviga (*mean absolute percentage error* – MAPE)
- Mediaan absoluuthälve (*mean absolute deviation* – MdAD)
- Mediaan suhteline absoluutviga (*median absolute percentage error* – MdAPE)

Shcherbakov, Brebels, Shcherbakova, Tykov, Janovsky ja Kamaev (2013) leidsid oma töös, et aegridade ennustamisel oleks kõige mõistlikum kasutada MAE, MdAE ja ruutjuur ruutkeskmise viga ehk RMSE (*root mean squared error*) mõõdikuid. Lisaks soovivad autorid kasutada mõõdikuid, mis põhinevad protsentuaalsetel vigadel ainult siis kui andmetes ei ole erindeid ja kui kõik väärtused on kas positiivse või negatiivse märgiga. Käesoleva töö raames

on täidetud mõlemad tingimused: 1) antud töös käsitletakse volatiilsust kui kauplemispäeva kõrgeima ja madalaima hinna vahet – alati positiivse märgiga väärtus ja 2) andmetes ei esine palju erindeid.

Keskmine absoluutviga on üks levinuimatest vea mõõdikutest. Arvutuskäigus jagatakse 1 andmepunktide arvuga ning korrutatakse läbi tegeliku väärtuse ja ennustatud väärtuse erinevuse summaga. Iga jääkliige panustab kogu vea arvestusse proportsionaalselt. Kui MAE = 0 siis on tegemist perfektsete ennustustega, kuid üldjuhul selliseid tulemusi ei esine peaaegu kunagi.

$$MAE = \frac{1}{n} \sum |y - y'| \quad (2)$$

kus

n – andmepunktide arv

y – tegelik väärtus

y' – ennustatud väärtus

Ruutkeskmine viga on väga sarnane keskmisele absoluutveale, kuid peamiseks erinevuseks on tegeliku väärtuse ja ennustatud väärtuse erinevuse ruutu panemine. Sellest tulenevalt ei ole MAE ja MSE omavahel võrreldavad, aga neid saab kasutada erinevate mudelite võrdlemiseks. Täiendavaks erinevuseks on jääkliikmete panus vea mõõdikusse. Kui MAE puhul oli see proportsionaalne, siis MSE puhul kasvab viga palju kiiremini, sest erindid omavad suuremat osakaalu.

$$MSE = \frac{1}{n} \sum (y - y')^2 \quad (3)$$

kus

n – vaatluste arv

y – tegelik väärtus

y' – ennustatud väärtus

Keskmine suhteline absoluutviga on valemi poolest väga sarnane keskmisele absoluutveale. Peamiseks erinevuseks on see, et kõike ei teisendata protsentideks. Ühtlasi mõlemad mõõdikud ei ole mõjutatud erinditest tänu absoluutväärtustele. Kuid MAPE'1 on ka omad puudused – nimelt tegeliku väärtuse ja ennustatud väärtuse jagamistehte tulemusena tekib olukord, kus MAPE'1 on tendents ennustuste kasuks kus ennustatud väärtus on väiksem kui tegelik väärtus.

$$MAPE = \frac{100\%}{n} \sum \left| \frac{y - y'}{y} \right| \quad (4)$$

kus

n – vaatluste arv

y – tegelik väärtus

y' – ennustatud väärtus

Mediaan absoluuthälve (MdAD) erineb ülalolevatest mõõdikutest enim, sest ülalolevad mõõdikud põhinevad erinevatel viisidel keskmisest veast. Seevastu mediaan absoluuthälve näitab vigade jaotuvust ümber keskmise vea. Seega kui mudeli ennustused on suure veaga, siis on MdAD väärtus suur, kui ennustused on täpsed ning väikese veaga, siis on MdAD väärtus väike.

$$MdAD = \text{Mediaan}(|y_t - y'|) \quad (5)$$

kus

y – tegelik väärtus

y' – ennustatud väärtus

Mediaan protsentuaalne absoluutviga (MdAPE) on praktiliselt sama, mis MAPE, kuid keskmise asemel on kasutatud mediaani ning erinevalt MAPE'st ei agregeeri MdAPE kõiki vaatlusi, vaid arvutab iga vaatluse protsentuaalse vea ning arvutab sellest andmerekast mediaani. Mediaani võib defineerida kui kasvavas järjekorras sorteeritud andmerekast keskmist väärtust, ehk 50% vaatlustest on kasvavas järjekorras väiksemad kui mediaan ning teine 50% vaatlustest on kasvavas järjekorras suuremad kui mediaan. Ühtlasi on mediaan robustne erindite ehk ekstreemsete väärtustega vaatluste poolest, mis teeb erinevate andmestike või mudelite võrdluse täpsemaks.

$$MdAPE = \text{Mediaan} \left(\left| \frac{y_t - y'}{y_t} \right| \times 100 \right) \quad (6)$$

kus

y – tegelik väärtus

y' – ennustatud väärtus

Eelnevalt käsitletud mõõdikud annavad mitmekülgse ülevaate ennustuste täpsusest. MAE, MSE ja MAPE kirjeldavad ennustuste ja tegelike väärtuste erinevusi. MdAD ja MdAPE

kirjeldavad samuti ennustuste ja tegelike väärtuste erinevusi, kuid arvutuskäigus kasutatakse keskmise asemel mediaani. Kuna mediaani komponendiga mõõdikud on ekstreemväärtuste ehk erindite osas robustsed, siis võimaldab see keskmise kompoendiga mõõdikute võrdlemisel tuvastada kumb mudel tagastab ekstreemsete väärtustega ennustusi, sest ekstreemsed väärtused kannavad keskmist väärtust ekstreemväärtuse suunda.

2. GARCH JA TEHISNÄRVIVÕRGUD

Käesoleva töö esimene peatükk käsitleb volatiilsust, volatiilsuse ennustamist ja sellega seotud termineid ja nähtusi. Esimene peatükk on oluline, sest see aitab selgitada erinevaid nähtusi ja termineid ning mudeleid, mida järgnevatel peatükkides käsitletakse. Käesolev peatükk annab ülevaate antud töö empiirilises osas kasutatavatest mudelitest ning nende ajaloost. Lisaks esimesele peatükile panustab teine peatükk kolmandasse peatükki, ehk empiiriliste tulemuste tõlgendamisesse. GARCH mudelit käsitletakse esimesena, sest on rohkem tundud ning käesolevas töös kasutatakse seda etalonina tehisnärvivõrkudega võrdlemisel.

2.1. GARCH

Varasemates töödes (Malmsten, Teräsvirta, 2010; Katsiampa, 2017) on välja toodud, et levinumateks tööriistadeks volatiilsuse modelleerimisel on autoregressiivsed mudelid. Tavapärased ökonomeetrilised ja aegridade mudelid eeldavad konstantset dispersiooni, aga Engle (1982) poolt tutvustatud ARCH mudel võimaldab tinglikul dispersioonil aja jooksul muutuda. Sellest tulenevalt arvestab mudel varasemaid vigu, kus tingimusteta dispersioon jääb konstantseks. 1980. aastatel esmalt tutvustatud ARCH mudel oli esimene mudel autoregressiivsete mudelite perekonnas, tänaseks aga on loodud sadu erinevaid autoregressiivseid mudeleid. Ühtlasi on antud mudelid kõige levinumad turu tootluse ja volatiilsuse ennustamisel. ARCH mudelite puhul kasutatakse standardhälvete asemel dispersiooni ning sellest tulenevalt on α^2 defineeritud h tähena allolevas ARCH mudelis:

$$\begin{aligned} r_{t+1} &= \mu + \varepsilon_{t+1}, \\ \varepsilon_{t+1} &= \sqrt{h_{t+1}} z_{t+1}, \\ h_{t+1} &= \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t+1-j}^2, \end{aligned} \tag{7}$$

kus

r_{t+1}	–	tinglik vara tootluse ennustus ajahetkel $t + 1$
μ	–	keskmine vara tootlus
ε_{t+1}	–	jääkväärtused (residuals)
z_{t+1}	\sim	normaaljaotusele alluvad juhuliikmed, ehk valge müra
$\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_q$	–	mudeli parameetrid

Antud mudeli moodustab ARCH(q) protsess, millel sisendparameeter q ehk viitaeg on tavaliselt 1 või 2. Suuremate parameetritega võib kaasneda mudeli efektiivsuse vähenemine (Tsai, Chan, 2006). z_{t+1} skaleerib tinglikku dispersiooni, ehk h_{t+1} , millele järgneb autoregressiivne regressiooni protsess. Parameetrid $\alpha_j \geq 0, j = 1, \dots, q$ tagavad selle, et dispersioon h_{t+1} on positiivne. Ühe sammu võrra tulevikku vaatav ennustus saadakse dispersioonist ruutjuure arvutamisel.

GARCH mudel on ARCH mudeli üldistatud versioon. Erinevus võrreldes ARCH'iga seisneb tinglikus dispersioonis ehk h_{t+1} .

$$\begin{aligned} r_{t+1} &= \mu + \varepsilon_{t+1}, \\ \varepsilon_{t+1} &= \sqrt{h_{t+1}} z_{t+1}, \\ h_{t+1} &= \omega + \sum_{i=1}^p \beta_i h_{t+1-i} + \sum_{j=1}^q \alpha_j \varepsilon_{t+1-j}^2, \end{aligned} \tag{8}$$

kus

r_{t+1} –	tinglik vara tootluse ennustus ajahetkel $t + 1$
μ –	keskmine vara tootlus
ε_{t+1} –	jääkväärtused (<i>residuals</i>)
$z_{t+1} \sim$	normaaljaotusele alluvad juhuliikmed, ehk valge müra
$\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_q$ –	mudeli parameetrid

Nagu ka ARCH mudeli puhul on GARCH mudeli parameetrid $\omega \geq 0, \beta_i \geq 0, \alpha_j \geq 0$ positiivsed. Eelnevate parameetritega kaasnevad ka täiendavad piirangud, mis on prioriteetsemad kui GARCH(1,1) (Tsai, 2006). Ühtlasi tõi Tsai oma töös välja, et antud parameetritega GARCH on kõige leviuim mudel. (1,1) mudeli parameetrites esimene number viitab sellele mitu autoregressiivset viitaega on mudeli võrrandis ning teine number näitab mitu libiseva keskmise viitaega on mudeli võrrandis kasutusel (Engle, 2001).

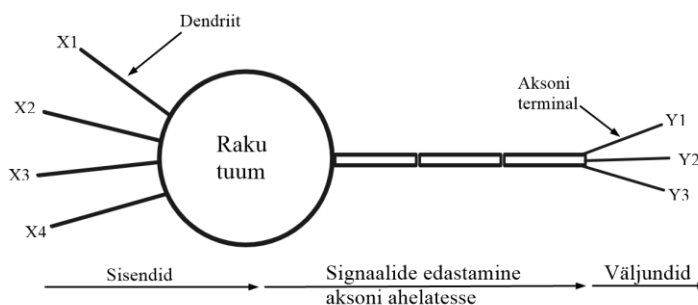
2.2. Tehisnärvivõrgud

Käesolev peatükk jaguneb kahte alampeatükki. Esimeses alampeatükis teeb autor sissejuhatuse bioloogilistesse närvirakkudesse ja tehisnärvivõrkudesse, mille raames tutvustab ja kirjeldab bioloogilise närviraku ja tehisnärvivõrgu erinevusi ja sarnasusi. Lisaks

eelnevale kirjeldab autor tehisnärvivõrkude mudeli algeid ja arenguid. Teises alampeatükis käsitleb autor detailselt käesolevas töös kasutatavat mitmekihilist pertseptron mudelit. Detailse kirjelduse hulka kuulub mudeli struktuur ning levinuimad ja soovituslikud seadistused mudeli konstrueerimiseks.

2.2.1. Sissejuhatus tehisnärvivõrkudesse

Närvirakud on bioloogilised rakud, mis käsitlevad informatsiooni inimajus. Antud rakud koosnevad närviraku kehast, puuoksa sarnastest dendriitidest (vastuvõtjad) ja aksonitest (signaale edastavad ahelad). Kui närvirakk saab signaali (impulsi) teistelt närvirakkudelt läbi dendriitide, siis suunatakse see signaal läbi närviraku tuuma aksoni ahelatesse. Informatsiooni vastuvõtmise ja edastamise protsess on kuvatud Joonisel 1. Kahe närviraku vahelist aksonit ja dendriiti ühendavad sünapsid, mis tekitavad neuronitest tulnud signaalidest keemilisi ühendeid, mida kutsutakse neurotransmitteriteks. Neurotransmitterite eesmärgiks on kas hajutada või rikastada vastuvõtva neuroni võimet edastada signaale. Sünaps suudab ise oma efektiivsust seadistada tänu neuroneid läbivatele signaalidele, mille tulemusena neuronid suudavad õppida tegevusi, millest nad osa võtavad (Jain, Mao, Mohiuddin, 1996).



Joonis 1. Bioloogilise närviraku mudel

Allikas: autor

Tehisnärvivõrkude alged ulatuvad tagasi 1943. aastasse, kus McCulloch ja Pitts tutvustasid esmakordselt arvutuslikku mudelit neuronist, mida nimetatakse lineaarseks lävendi ühikuks. Nende teooria lähtus järgmistest eeldustest:

- 1) Neuroni aktiivsus sõltub „kõik või mitte midagi“ protsessist, mis matemaatiliselt tähendab binaarset väljundit, 1 kui neuron aktiveerub ja 0 kui neuron ei aktiveeru.
- 2) Fikseeritud arv sünapse peab olema aktiveeritud teatud perioodi viitaja lisandumisega selleks, et aktiveerida neuronit igal ajahetkel ning antud number on neuroni suhtes sõltumatu eelnevast neuroni aktiivsusest ja positsioonist.
- 3) Närvisüsteemi ainukene märkimisväärne viivitus on sünapside vaheline viivitus.
- 4) Pärssiva sünapsise aktiivsus hoiab igal ajahetkel ära neuroni aktiveerumise.
- 5) Närvivõrgu struktuur on ajas muutumatu.

Kui eelnevaid eeldusi lahti seletada, siis esimese punkti all peetakse silmas neuroni aktiveerimise binaarset väljundit – ehk kas neuron aktiveerub või mitte. Kui neuron ei aktiveeru, siis sellest neuronist signaale edasi ei edastata. Teise punkti all peetakse silmas, et fikseeritud arv sünapse peab olema aktiveeritud teatud viitaja lisandumise ajavahemikul selleks, et aktiveerida neuronit ning fikseeritud sünapside arv on sõltumatu neuroni eelnevast aktiivsusest ja positsioonist. Neljanda punkti all peetakse silmas, et pärssiv sünaps hoiab ära neuroni aktiveerumise ehk mingisugune impulss jõuab sünapsini, kuid sünaps ei kannu seda edasi neuronini.

$$f(x) = \begin{cases} 1 & \text{kui } w \cdot x + b > 0, \\ 0 & \text{kui } w \cdot x + b \leq 0 \end{cases} \quad (9)$$

kus

w – kaalud
 x – sisend
 b – vabaliige

$$w \cdot x = \sum_{i=1}^m w_i x_i \quad (10)$$

kus

m – pertseptroni sisendite arv

Neuroplastilisus on aju võime ajas muutuda, ehk teatud sünapsid võivad aja jooksul kas tugevneda või nõrgeneda. Donald Hebb (1949) tutvustas õppimise hüpoteesi, mis põhines neuroplastilisusel – antud hüpotees hiljem sai tuntuks kui Hebb'i õpe (*Hebbian learning*).

Hebb'i töö viis selleni, et 1954. aastal tutvustasid Farley ja Clark (1954) arvutuslike masinaid, mis põhinesid Hebb'i võrgustikel. Mõned aastad hiljem tutvustas Rosenblatt (1958) pertseptroni (ühekihiline tehisnärvivõrk) algoritmi. Antud mudel on lävendifunktsiooniga binaarne klassifitseerimise mudel, mis võtab sisendi x ja annab väljundiks 0 või 1. Ülaloleval valemil (vt Valem 10.) on kuvatud pertseptroni valem.

2.2.2. Mitmekihiline pertseptron närvivõrk

Tehisnärvivõrkude arhitektuure on erinevaid, neist levinuimad on mitmekihiline pertseptron, LSTM (*Long Short Term Memory*) ja konvolutsioonilised tehisnärvivõrgud. Käesolevas töös kasutatakse mitmekihilist pertseptron tehisnärvivõrkude mudelit, mis kogusid populaarsust 1980. aastatel kõne- ja pildituvastuse ning masintõlke aladel (Wasseman ja Schwartz, 1988). LSTM mudel erineb pertseptronist selle poolest, et suudab üksikute andmepunktide asemel töödelda andmete jadasid. Antud mudel on levinud kõnetuvastuse, käekirjatuvastuse ja muusika modelleerimise aladel (Greff, Srivastava, Koutnik, Steunebrink ja Schmidhuber, 2016). Konvolutsioonilised tehisnärvivõrgud on enim levinud pildi- ja videotuvastamise aladel, sest antud mudel suudab visuaalset informatsiooni töödelda pikslite haaval (Long, Shelhamer ja Darrell, 2015). Eelneva põhjal võib järeldada, et erinevad tehisnärvivõrkude arhitektuurid on head mudelid klassifitseerimise probleemide lahendamisel.

Mitmekihilist pertseptron tehisnärvivõrku tutvustas esmakordselt Ivakhnenko (1970). Võrreldes ühekihilise närvivõrguga oli Ivakhnenko idee selles, et kui ühekihilises närvivõrgus tehakse viimane otsus või ennustus esimeses andmete töötamise kihis, siis mitmekihilises närvivõrgus kantakse signaalid läbi mitmete kihtide. Kõikidest kihtidest läbi kantud signaalid on seotud mõõdiku muutustega, summaatorite ja lävendipõhiste ühikutega. Mittelõplike otsuste tegemine läbi erinevate kihtide võimaldab erinevatel heuristikutel reageerida informatsioonivoole mitu korda ning see tagab märkimisväärselt kõrgema heuristiliselt iseorganiseeruvate süsteemide täpsuse.

Sarnaselt bioloogilistele närvivõrkudele vajavad tehisnärvivõrgud sisendeid. Dendriitide asemel kasutatakse algandmeid ning sünaptsid esindavad neuronite vahelisi kaalusid ehk parameetrite hinnanguid. Kaalud on alguses juhuslikud, kuid neid uuendatakse tagasilevi

protsessiga, mis leiab iga sisendi jaoks optimaalse kaalu, minimeerides ennustava väärtuse ja tegeliku väärtuse erinevust. Kaalude ja sisendite abil arvutatakse kaalutud sisendsignaalide summa, mis peegeldab närviraku aktiivsust. Sisendsignaalide summale rakendatakse aktiveerimisfunktsioon, mille tulemusena saadakse mudeli väljund, sellele omakorda rakendatakse tagasilevi protsess, mis minimeerib väljundi eksimust tegeliku väärtuse suhtes.

Lähtudes eelnevast koosneb mitmekihiline pertseptron närvivõrk sisendandmete kihist, peidetud kihtidest ja nendes kihtides olevates neuronitest, kihtidevahelistest aktiveerimisfunktsioonidest ning tagasilevi protsessist, kus on kaalusid uuendav algoritm. Modelleerijal on võimalik valida peidetud kihtide arvu ja igas kihis olevate neuronite arvu. Antud teemadel on läbi viidud mitmeid uurimustöid, kuid siiani ei ole jõutud täpsete või universaalsete lahendusteni, mis vastaksid küsimusele kui palju peaks ühes tehisnärvivõrgus olema peidetud kihte ja kui palju neuroneid peaks neis kihtides olema.

Mitmekihilise pertseptron mudeli sisendandmete kiht koosneb sõltumatutest muutujatest, kus iga muutuja esindab ühte neuronit (Karsoliya, 2012). Üldistatuna tagavad sisendkihti edastatud muutujad väljundi, kuid sisendandmete kihi ja väljundi kihi vahele on võimalik konstrueerida detailsem närvivõrgu arhitektuur. Detailsem närvivõrgu arhitektuur ja võimalikud seadistuse võimalused on kirjeldatud järgnevates lõikudes.

Peidetud kihid on kogumid neuronitest, mis asuvad sisendandmete kihi ja väljundi kihi vahel. Antud kihtidesse tulevad signaalid teistest peidetud kihtidest või ühe peidetud kihi olemasolul sisendandmete kihist. Kõik neuronid on ühendatud eelnevates ja järgnevates kihtides olevate neuronitega. Iga kihi neuronile arvutatakse eelnevast kihist signaalide ja kaalude kaalutud summa ning antud summale rakendatakse aktiveerimisfunktsioon. Aktiveerimisfunktsioonist tekkinud uus signaal suunatakse järgmise peidetud kihi neuronitesse või väljundi kihti. Kaalud on olulised tagasilevi protsessis, sest siis tagasilevi protsess leiab optimaalse kaalu iga neuroni jaoks, mis tagab kõige täpsema või lähedasema väljundi tegelikule väärtusele.

Aktiveerimisfunktsioone on erinevaid – neist levinuimad on sigmoidfunktsioon, hüperboolne tangensfunktsioon ning mittenegatiivne lineaarfunktsioon. Käesolevas töös kasutatakse sigmoidfunktsiooni, sest Sibi, Jones ja Siddarth (2013) leidsid, et sigmoid annab parimaid tulemusi kui sisend on vahemikus 0-1. Arvestades, et käesolevas töös on volatiilsust käsitletud kui kauplemispäeva kõrgeima ja madalaima hinna protsentuaalset erinevust, on 0-1 vahemiku

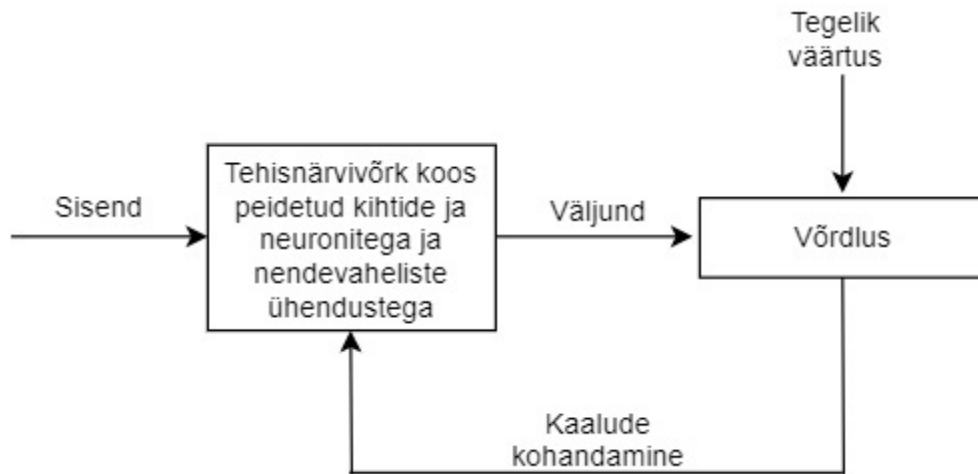
tingimus täidetud, sest volatiilsus jaguneb enamasti 0.00 – 0.15 vahemikku, kus enamus vaatlusest on 0.00-0.05 (vt Joonis 6. Volatiilsuse jaotuvus). Sigmoidfunktsioon on kuvatud valemil 11.

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (11)$$

kus

x – sisendi ja kaalu kaalutud summa

Tagasilevi protsessi tutvustasid esmakordselt Rumelhart, Hinton ja Williams (1986). Antud protsess uuendab neuronite vahelisi kaalusid, et saavutada minimaalne erinevus tegeliku väärtuse ja ennustatava väärtuse vahel. Mudeli esialgsed kaalud on juhuslikud ning neid kaalusid optimeerib tagasilevi algoritm iga iteratsiooniga. Üks tagasilevi protsessi iteratsioon tähistab ühe vaatluse arvutamise protsessi ja tagasilevi protsessi ehk sama vaatluse arvutamist uuendatud kaaludega. Üks iteratsioon mudeli treenimisel tähistab kogu andmestiku läbi töötamist ühe korrana. Mudeli treenimise all peetakse silmas mudeli konstrueerimist valimiga, eesmärgil teha valimiväliseid ennustusi, kus treenimise käigus leitakse optimaalne mudel. Levinud iteratsioonide arvudeks on S. Hosein ja P. Hosein (2017) töö põhjal 400 iteratsiooni, kuid see ei ole reegel, mis ühildub üheselt kõikide andmete ja mudeli seadistustega. Samas 400 iteratsiooni on hea koht, millest lähtuda modelleerimise alustamisel. Tagasilevi protsess on lihtsustatult kuvatud Joonisel 2., millelt on näha, et protsess algab vaatluse sisendist. Seejärel peale mudeli neuronite, kaalude ja kihtide vahelisi arvutusi jõutakse väljundini. Peale väljundi ja tegeliku väärtuse võrdlust liigutakse tagasi väljundi arvutamise protsessi, milles on kasutusel uued kaalud ning seda korratakse niikaua kuni modelleerija poolt määratud iteratsioonide arv on saavutatud.



Joonis 2. Tagasilevi protsess

Allikas: autori koostatud

Stohhastiline gradientlaskumine (*stochastic gradient descent*) ja ruutjuur ruutkeskmise leviku (*root mean square propagation*) algoritmid on tagasilevi protsessis laialdaselt levinud kaalusid uuendavad algoritmid. Stohhastiline gradientlaskumise funktsioon sai alguse Kiefer ja Wolfowitz (1952) tööst. Ruutjuur ruutkeskmise leviku meetodit tutvustasid Hinton ja Tielman (2012) Coursera kursusel. Antud meetodi eesmärgiks on leevendada haihtuvate gradientide probleemi, mis esineb kaalude uuendamisel. Haihtuvate gradientide probleem kujutab endas kaalude uuendamise protsessi olukorda, kus kaalusid uuendatakse nii väiksel määral võrreldes eelneva kaalude uuendamisega, et ennustused uute kaaludega on täpselt samad või erinevad väga väikesel määral võrreldes eelneva ennustusega (Mosca, Magoulas, 2017).

Masinõppes ja statistikas on õppimise määr gradientlaskumise parameeter, mis optimeerib iga iteratsiooniga sammu suurust (kaalude muutust) minimaalse vea suunas. Gradientlaskumise arvutuskäigus korrutatakse õppimise määr gradiendiga, mille tulemusena leitakse järgmises iteratsioonis kasutatav sammu suurus. Õppimise määr võib võrrelda tavapärase sammude astumisena, kus õppimise määr on sammu suurus või pikkus, mida vea minimeerimisel astuda. Kujutledes hüperboolset joonist, kus minimaalne viga on hüperbooli väikseim väärtus, siis väikeste sammudega kulub rohkem aega, et jõuda minimaalse veani, kuid seevastu suurte sammude astumisel võib minimaalsest veast liiga kaugemale või üle astuda. Enamasti jääb antud õppimise määr vahemikku 0.001-0.1. Optimaalse kaalu leidmisel korrutatakse õppimise määr

muutuse kiirusega ehk gradiendiga. Valemil 12 on kuvatud ruutjuur ruutkeskmise leviku algoritmi kaalude uuendamise valem.

$$w_{ij}(t) = w_{ij}(t-1) + (1-\alpha) \left(\frac{\partial E(t)}{\partial w_{ij}} \right)^2 \quad (12)$$

kus

w_{ij} – kaal i -nda sisendandmete kihi neuroni ja j -nda peidetud kihi neuroni vahel

α – õppimise määr

E – ruutviga

t – vaatlus

$\left(\frac{\partial E(t)}{\partial w_{ij}} \right)^2$ – gradient

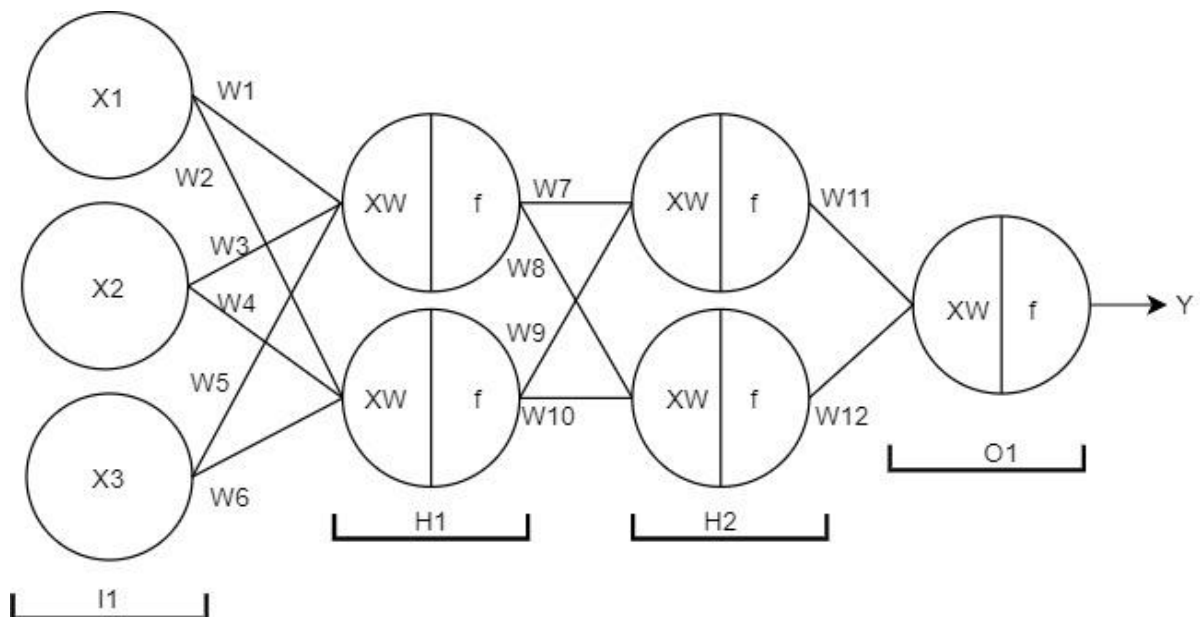
Ruutvea arvutus:

$$E = \sum [O_k(n) - O'_k(n)]^2 \quad (13)$$

kus

$O_k(n)$ – närvivõrgu n -da vaatluse ennustus

$O'_k(n)$ – tegeliku n -da vaatluse väärtus



Joonis 3. Mitmekihiline pertseptron tehisnärvivõrk

Allikas: autori koostatud

Joonisel 3. On kuvatud mitmekihiline pertseptron närvivõrgu mudeli arhitektuur, kus:

$X1, X2, X3$ – sisendandmed

$I1$ – sisendandmete kiht

$W1, W2, W3 \dots WN$ – neuronite vahelised kaalud

XW – neuronite vahelised kaalutud summad

f – aktiveerimisfunktsioon

$H1, H2$ – peidetud kihid

$O1$ – väljundi kiht

Y – ennustatud väärtus

Neuronite vaheliste kaalude kaalutud summa arvutatakse järgnevalt:

$$XV = \sum x_n w_n \quad (14)$$

kus

x_n – sisendandmed

w_n – kaalud

Kui Joonis 3. abil antud alampeatükk kokku võtta, siis on näha, et mitmekihilised tehisnärvivõrgud koosnevad vähemalt kolmest kihist: 1) sisendandmete kiht – I_n , 2) peidetud kiht või kihid – H_n , 3) väljundi kiht – O_n . Üldjuhul on igas kihis erinev neuronite arv. Sisendandmete kihis on alati iga muutuja kohta üks neuron – X_n . Peidetud kihi neuronite arv moodustab 70% – 90% kõikidest sisendandmete muutujate arvust. Väljundi kiht koosneb alati ühest neuronist – väljundist. Iga peidetud kihi neuron on seotud kõikide eelneva ja järgneva kihi neuronitega. Iga peidetud kihi neuroni väärtus arvutatakse eelneva kihi neuroni ja kaalu kaalutud summana – XW_n , kus kaal W_n on mudeli poolt sätestatud juhuslik väärtus. Iga kihi igale neuronile rakendatakse aktiveerimisfunktsioon – f , mis sõltuvalt sisendist hajutab või rikastab signaali. Viimasena optimeerib tagasilevi protsess neuronite vahelisi juhuslikke kaalusid, eesmärgil minimeerida ennustuse ja tegeliku väärtuse erinevust.

3. VARASEMATE EMPIIRILISTE UURINGUTE TULEMUSED

Kui esimene peatükk selgitas ja kirjeldas volatiilsust ning sellega seotud termineid, nähtuseid ja mudeleid ning teine peatükk tutvustas käesoleva töö empiirilises osas kasutatavaid mudeleid, siis käesolev peatükk tutvustab empiirilises osas kasutatavate mudelite varasemaid tulemusi. Peatüki raames käsitletakse mõlema mudeli varaseimaid empiirilisi uuringuid, mille käigus on modelleeritud erinevaid varasid erinevatel ajavahemikel.

3.1. GARCH empiirilised tulemused

Emenike (2010) kasutas GARCH(1,1) mudelit Nigeeria börsi aktsia tootluste volatiilsuse ennustamisel. Uurimuse tulemusena leiti, et aastate vahemikul 1985 kuni 2008 esineb volatiilsuse kuhjumine. Antud töö eesmärgiks ei olnud mõõta GARCH mudeli valimivälist täpsust, kuid Aliyu (2009) tuvastas töö raames, et Nigeeria börsil esineb volatiilsuse kuhjumist, asümmeetrilist volatiilsust ja leptokurtooset jaotust. Sama avastuse tegid Suliman ja Abdalla (2012), uurides Saudi Araabia börsi aastatel 2007 – 2011.

Tihti peale on GARCH mudelid volatiilsetel perioodidel ebatäpsed, andes ennustatavast väärtusest märkimisväärselt suuremaid või väiksemaid ennustusi. Klaassen (2002) pakkus välja, et see on seotud asjaoluga, et tavapärased GARCH mudelid on üherežiimilised (*single regime*) ja seejärel esitas Klaassen režiimi muutvad GARCH mudeleid, kus lähtus GARCH mudeli kahe režiimiliseks tegemisest. Klassen lähtus Cai (1994), Hamilton, Susmel (1994) ja Fong (1998) uurimustööst, mis käsitlesid mitme režiimiga ARCH mudeleid. Kasutades 1978.–1997. aasta USA dollari, Briti naela ja Saksa marga andmeid, saavutas Klassen valimiväliste tulemuste alusel ühepäevastel ennustustel 22% parema tulemuse kui ühe režiimiline GARCH. Kümnapäevastel ennustustel saavutati 58% parem tulemus kui ühe režiimilisel GARCH mudelil.

Viimase paarikümne aasta jooksul on tutvustatud sadu erinevaid autoregressiivseid mudeleid. Hansen ja Lunde (2005), viisid läbi uurimuse, kus võrdlesid 330 autoregressiivse mudeli tulemusi GARCH(1,1) mudeli valimiväliste tulemustega. Antud töö käigus kasutati järgmise päeva volatiilsuse ennustamiseks IBM'i aktsia, USA dollari ja Saksa marga andmeid

vahemikul 1987-1993. Uurimuse tulemusena ei leitud ühtegi mudelit, mille valimivälaliste ennustuste täpsus oleks valuutakursside ennustamisel parem kui GARCH(1,1), kuid märkimisväärne erinevus oli IBM'i aktsia volatiilsuse ennustamisel, kus mudelid, mis suudavad arvestada võimenduse efekti olid GARH mudelist selgelt paremad. Samas võrdluses osutus Katsiampa (2017) töös parimaks mudeliks CGARCH, kuid tavapärase GARCH'i mudeli tulemused ei olnud märkimisväärselt kehvemad. Antud töö raames kasutati 2010-2016 aasta Bitcoin hinnaandmeid, mille põhjal ennustati järgneva kauplemispäeva volatiilsust.

Chan ja Grant (2016) viisid läbi uuringu, mille raames võrreldi GARH(1,1) ja stohhastilise volatiilsuse mudelit, millele rakendati AR(1) logaritmitud volatiilsuse protsess. Antud töö käigus kasutati kokku 9 erineva toornafta, petrooleumi ja maagaasi varade hinnaandmeid vahemikul 1997-2015. Töö tulemusena jõuti järgnevate järeldusteni: 1) t jaotus parandab märkimisväärselt GARCH mudeli tulemusi, 2) libiseva keskmise komponendi lisamine mudelitesse parandab mõlema mudeli tulemusi, 3) võimenduse efekt on oluline ainult toornafta modelleerimisel, 4) stohhastilise volatiilsuse mudel koos libiseva keskmise komponendiga saavutas kõikide aegridade puhul parimad tulemused.

Tabel 1. GARCH empiirilised tulemused

Autor (aasta)	Mudel	Andmed	Järeldused
Klassen (2002)	Ühe ja mitmerežiimilised GARCH mudelid	USA dollar, Briti nael ja Saksa mark (1979-1997)	Mitmerežiimiline on täpsem kui üherežiimiline GARCH
Hansen ja Lunde (2005)	330 autoregressiivset mudelit	IBM'i aktsia, USA dollar ja Saksa mark (1987-1992)	GARCH mudel oli valuutakursside ennustamisel parim, aktsiate puhul olid võimenduse efekti arvestavad mudelid paremad
Katsiampa (2017)	CGARCH, GARCH veel	Bitcoin (2010-2016)	CGARCH oli kõige väiksema veaga, GARCH ei olnud oluliselt halvem
Chan ja Grant (2016)	GARCH ja AR protsessiga stohhastilise volatiilsuse mudel	Erinevad toornafta, petrooleumi ja maagaasi varad (1997-2015)	Libiseva keskmise komponendiga stohhastilise volatiilsuse mudel on kõige täpsem

Allikas: autori koostatud

Tabelist 1. on näha, et enamasti kõigi varasemate tööde puhul on GARCH võrreldes teiste mudelitega olnud hea või kõige täpsemate ennustustega. Ühtlasi on näha, et antud tulemused kajastuvad ka erinevate varaklasside lõikes, nagu valuutad, aktsiad ja maavarad.

3.2. Tehisnärvivõrkude empiirilised tulemused

Arvestades, et valdav enamus tehisnärvivõrkude modelleerimistest viiakse läbi programmeerimiskeeltes, siis annab see võimaluse luua vähese vaevaga levinud mudelite ja tehisnärvivõrkude hübriidmudeleid. Hübriidmudelite ajendiks on peamiselt lineaarsete ja mitte-lineaarsete mudelite kombineerimine, kus tehisnärvivõrgud on mitte-lineaarsed mudelid. Ühe sellise hübriidmudeli ennustuste tulemusi demonstreeris Zhang (2003). Autor viis läbi aegridade analüüsi kasutades ARIMA mudelit, tehisnärvivõrke ja mõlema mudeli hübriidi. Hübriidmudeli aluseks oli ARIMA mudel, mis modelleeris lineaarseid suhteid ning ARIMA mudeli jääkliikmetele ehk mittelineaarsetele andmepunktidele rakendati tehisnärvivõrgud. Modelleerimisel kasutati kolme erinevat andmekogumit: Wolf'i päikesepunktide andmed (1700–1987), Mackenzie jõkke lõksu jäänud Kanada ilveste andmed (1821-1934) ning Briti naela ja USA dollari valuutakursi andmeid 1980-1993). Päikesepunktide andmete tulemustest oli näha lühikese etteennustuse perioodil närvivõrkude tugevust väikese vea näol. Pikemal etteennustuse perioodil oli hübriidmudel väikseima veaga. Ilveste lõksu jäämise andmete puhul oli hübriidmudelil valimiväliseks veaks 18%, mis võrreldes tehisnärvivõrkude ja ARIMA'ga oli märkimisväärselt kehvem, kus mõlema mudeli viga oli 7% ringis. Valuutakursi ennustati kolme perioodi lõikes: 1 kuu, 6 kuu ja 12 kuu ennustus. Lühemal perioodil olid väikseima veaga hübriidmudel ning tehisnärvivõrgud, keskmisel perioodil oli kõikide mudelite viga samas suurusjärgus, seejuures keskmise perioodi viga oli kõikidel mudelitel suurem kui pika ja lühikese perioodi viga. Pikal ennustusperioodil oli väikseima veaga hübriidmudel.

Guresen, Kayakutlu ja Daim'i (2011) läbiviidud uurimustöös uuriti mitmekihilise närvivõrgu, GARCH mudeli ja tehisnärvivõrkude hübriidi. Kumar, Ravi (2007) uurimustöös kasutati Nasdaq'i aktsiaindeksi üheksa kuu hinna andmeid. Töö tulemusena leiti, et mitmekihiline tehisnärvivõrgu mudel saavutas veidi paremad tulemused kui tehisnärvivõrkude ja GARCH'i hübriid. Tavapärase GARCH mudel saavutas selgelt halvimal tulemused.

Nagu juba varasemalt on käesolevas töös nimetatud, on autoregressiivsed mudelid ühed laialdasemalt levinud volatiilsuse ennustamise mudelid. Dondaldson ja Kamstra (1996) viisid läbi uuringu, mille raames võrreldi autoregressiivsete mudelite ja tehisnärvivõrkude valimi ja valimiväliste ennustuste täpsust. Autoregressiivseteks mudeliteks olid GARCH, EGARCH ja GJR mudelid. Mudelite sisendandmeteks olid Londoni, Tokyo, New Yorgi ja Toronto aktsiaindeksite hindade volatiilsus aastavahemikul 1980-1990. Kõigi nelja aktsiaindeksi puhul saavutas väikseima veaga tulemused tehisnärvivõrkude mudel. Sarnase uurimuse viisid läbi ka Hu ja Tsoukalas (1999), kus autorid ennustasid 11 valuutakursi volatiilsust GARCH, EGARCH, IGARCH, MAV ja tehisnärvivõrkudega. Sisendandmetena kasutati 1979.-1994. aasta valuutakursside hinnaandmeid. Kõige väiksemad vead saavutas valimivälistel ennustustel EGARCH mudel, millele järgnesid MAV ja tehisnärvivõrgud. Ühtlasi näitasid tehisnärvivõrgud parimaid tulemusi 1993. aasta valuutakriisi perioodil.

Hu, Zhang, Jiang ja Patuwo (1999) viisid läbi aegridade ennustuse analüüsi, mille raames ennustati Briti naela ja USA dollari valuutakursi hindu tehisnärvivõrkude ja ARIMA mudeliga. Töös kasutatavad andmed pärinesid aastatest 1976 – 1992. Tehisnärvivõrkude puhul oli tegemist kolmekihilise pertseptron mudeliga, mille aktiveerimisfunktsioonideks olid logistilised funktsioonid. Mudeli tagasilevi protsessi vigade optimeerimise funktsioonina kasutati lineaarset funktsiooni. Autorid kasutasid oma töös mitmeid mõõdikuid, mille abil mudelite tulemusi võrrelda, sealhulgas: MAE, MAPE ja erindite osas robustne MdAPE. Töö käigus kasutati kolme erinevat etteennustuse perioodi ning tulemuste analüüsi käigus leiti, et tehisnärvivõrgud on täpsemate ja stabiilsemate tulemustega kui ARIMA mudel kõigi kolme ennustusperioodi lõikes.

Tabel 2 Tehisnärvivõrkude tulemused

Autor (aasta)	Mudel	Andmed	Järeldused
Zhang (2001)	ARIMA, tehisnärvivõrgud, ARIMA ja ARIMA ning tehisnärvivõrkude hübriid	Wolf'i päikesepunktid (1700–1987), ilveste jõkke lõksu jäämised (1821-1934), GBP/USD (1980-1993)	Lühikesel ennustusperioodil on tehisnärvivõrgud parimad, pikal perioodil hübriidmudel
Guresen, Kayakutlu ja Daim (2011)	GARCH, tehisnärvivõrgud ja GARCH-i ja tehisnärvivõrkude hübriid	Nasdaq'i aktsiaindeksi 9 kuu hinnaandmed	Tehisnärvivõrgud olid veidi paremad kui hübriidmudel, GARCH saavutas suurimad vead
Dondaldson ja Kamstra (1996)	GARCH, EGARCH, GJR ja tehisnärvivõrgud	Londoni, Tokyo, New Yorgi ja Toronto aktsiaindeksite hinnad (1980-1990)	Tehisnärvivõrgud saavutasid väikseimad vead
Hu ja Tsoukalas (1999)	GARCH, EGARCH, IGARCH, MAV ja tehisnärvivõrgud	11 valuutakurssi (1979-1994)	Väiksemad vead saavutasid EGARCH ja MAV, tehisnärvivõrgud olid parimad 1993. valuutakriisi perioodil
Hu, Zhang, Jiang ja Patuwo (1999)	Kolmekihiline pertseptron logistilise aktiveerimisfunktsiooniga, ülejäänud parameetrid on leitud katseeksitus meetodiga, ARIMA	Briti naela ja USA dollari vahetuskurss (1976-1982)	Tehisnärvivõrgud tagavad täpsemad ning stabiilsemad tulemused kõikide vea mõõtmise moodsikute puhul.

Allikas: autori koostatud

Tabelis 2. on kuvatud varasemad empiirilised uuringud tehisnärvivõrkudega volatiilsuse ennustamise teemal. Sarnaselt Tabel 1.-le on näha, et ka tehisnärvivõrkude varasemate uuringute puhul on tehisnärvivõrgud kujunenud heaks või parimaks volatiilsuse ennustajaks. Ühtlasi võrreldes GARCH mudeli empiiriliste tulemustega on ka tehisnärvivõrkude empiiriliste uuringute puhul näha, et ennustatavad varaklassid ei mõjuta mudeli ennustuste täpsust.

4. EMPIIRILISED TULEMUSED

Käesoleva töö eesmärgiks on ennustada S&P 500 aktsiaindeksi volatiilsust tehisnärvivõrkude abil. Selleks, et mõista tehisnärvivõrkude ennustusi, kasutatakse võrdlusena samade andmete ja samades tingimustes koostatud GARCH mudelit. Siiani on antud töös käsitletud järgnevaid teemasid. Esimeses peatükis selgitati volatiilsuse olemust ning sellega seotud termineid ja nähtuseid. Lisaks sellele tutvustati levinuimaid volatiilsuse modelleerimise mudeleid ning mõõdikuid, mille abil erinevate mudelite tulemusi võrrelda. Sellele järgnes teine peatükk, mille raames käsitleti käesolevas töös kasutatavaid mudeleid ning nende mudelite tausta ja ajalugu. Viimasena, enne empiirilist analüüsi, uuris autor varaseimad tulemusi volatiilsuse ennustamisel GARCH ja tehisnärvivõrkude mudelitega.

Antud peatükk on jaguneb kahte alampeatükki. Esimene alampeatükk on fokuseeritud töös kasutatavatele andmetele. Andmete peatükis selgitab autor, milliseid andmeid kasutatakse ning kirjeldab varasemate tööde põhjal, mis peitub antud valikute taga ning millised on antud andmetega seotud eeldused. Töö teine alampeatükk jaguneb omakorda kolme alampeatükki. Esimeses viiakse läbi üldine empiiriline analüüs kõikide valimiväliste vaatluste põhjal. Teises ja kolmandas alampeatükis viiakse läbi madala ja kõrge volatiilsuse ennustamise empiiriline analüüs.

4.1. Andmed

Valitud andmed on autori poolt grupeeritud nelja kategooriasse. Hinna ja tehingumahtude kategooriasse kuulub sõltuv muutuja S&P 500 ja sõltumatutest muutujatest S&P 500 tehingumahud ja volatiilsuse libisev standardhälve. Töös kasutatakse andmeid ajavahemikust 03.01.2007 – 28.02.2020, kus on kokku 3223 vaatlust. Antud ajavahemikku jääb 2007. aasta turu tipp, sellele järgnenud majanduslanguse periood ning majanduskasvu periood kuni 2020 aasta Covid-19 viirusest tingitud börsilanguse alguseni (vt Joonis 4.).

Wu ja Hou (2019) soovivad modelleerimisel kasutada volatiilsuse ennustamisel teisi aktsiaindeksid sõltumatute muutujatena. Sellest tulenevalt kuuluvad hinnaandmete hulka Dow Jones ja Nasdaq aktsiaindeksite andmed. Lisaks sellele autor usub, et kulla, toornafta ja USA dollari hinnaandmed on olulised S&P 500 volatiilsust ennustamisel, sest teatud

ettevõtteid mõjutavad kulla ja toornafta hinnad ning valuutakursside määrad. Sellest tulenevalt on ka viimased hinnaandmete kategoorias. Lisaks eelnevale leidsid Tully ja Lucey (2007), et USA dollar on kulla volatiilsuse ennustamisel oluline muutuja.

Liu ja Pan (2019) saavutasid VS-GARCH (*spillover GARCH*, Liu ja Pan, (2019)) mudeliga makroökonomiliste andmete kasutamisel GARCH (1,1) mudelist paremad ennustuste tulemused. Samale järeldusele jõudsid Tully ja Lucey (2007), kes uurisid kuidas makroökonomilised indikaatorid mõjutavad kulla hinda. Töö tulemusena leiti, et kõige suuremat mõju kulla hindadele omab USA dollar. Tuginedes eelnevale kasutab autor USA majanduse makroökonomilisi indikaatoreid.

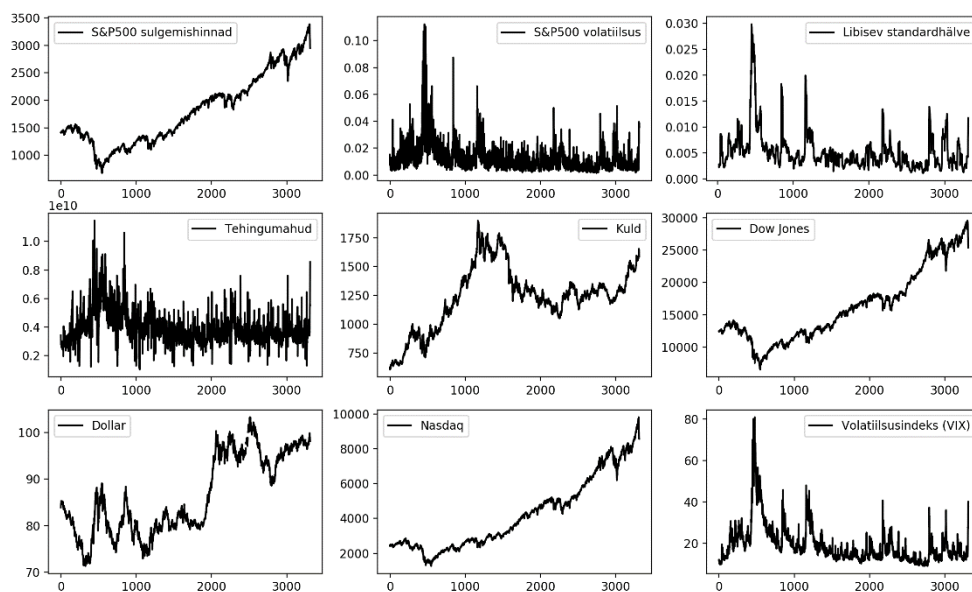
Volatiilsuse modelleerimisel laialdaselt levinud volatiilsusindeks VIX (Dicle, Levendis, (2019); Bardgett, Gourier, Leippold, (2019)) on autori poolt defineeritud sentimentide indikaatoriks. VIX'i on kasutatud mitmetes volatiilsuse ennustamise töödes, sealhulgas ka Kambourdis ja McMillan (2016) töös, kus autorid uurisid, kas VIX volatiilsusindeks ja ennustatava vara tehingumahud parandavad GARCH mudeli ennustusi. Autorid leidsid, et nii VIX kui ka sõltuva muutuja tehingumahud parandavad mudeli ennustamisvõimet, kuid täiendav ennustusvõime on marginaalne. Lisaks volatiilsusindeksile on volatiilsuse ennustamisel kasutatud ka Google Trends'i (Hamid ja Heiden (2015); Dzelinski, Rieger ja Talpsepp (2011)). Dzelinski, Rieger ja Talpsepp (2011) leidsid, et Google Trends'i märksõnaotsingu indeks on majandlanguse (*recession*) märksõna kasutamisel heaks kõrge volatiilsuse ennustajaks. Tabelis 3. on kuvatud töös kasutatavad andmed, nende andmete intervallid ja allikad.

Tabel 3. Töös kasutatavad andmed

Muutuja	Muutuja kategooria	Ajavahemik	Allikas
S&P 500 volatiilsus	Hinna ja tehinguandmed	Päev	Yahoo Finance
VIX	Sentimendi indikaator	Päev	Yahoo Finance
S&P 500 tehingumahud	Hinna ja tehinguandmed	Päev	Yahoo Finance
S&P 500 libisev standardhälve	Hinna ja tehinguandmed	20 päeva	Autori arvutused
Toornafta	Hinnaandmed	Päev	U.S. EIA
Kuld	Hinnaandmed	Päev	Investing.com
Dow Jones 30	Hinnaandmed	Päev	Yahoo Finance
Nasdaq 100	Hinnaandmed	Päev	Yahoo Finance
USA dollar	Hinnaandmed	Päev	Yahoo Finance
USA töötuse määr	Makroökonomiline indikaator	Kuu	U.S. Bureau of Labour statistics
USA loodud töökohtade arv	Makroökonomiline indikaator	Kuu	U.S. Bureau of Labour statistics
USA tarbija hinnaindeks	Makroökonomiline indikaator	Kuu	U.S. Bureau of Labour statistics
USA tootja hinnaindeks	Makroökonomiline indikaator	Kuu	U.S. Bureau of Labour statistics
USA kinnisvarahinnaindeks	Makroökonomiline indikaator	Kuu	Federal Housing Finance Agency
USA keskpanga intressimäär	Makroökonomiline indikaator	Kuu	Federal Reserve Bank
Google Trends	Sentimendi indikaator	Nädal	Google Trends

Allikas: autori koostatud

Joonisel 4. on näha esimene osa käesolevas töös kasutatavatest andmetest. Täpsemalt kuvab antud joonis S&P500 sulgemishindu, volatiilsust, libisevat standardhälvet, S&P 500 tehingumahtusid ning teiste aktsiaindeksite ja toormete hindu, sealhulgas ka volatiilsuseindeksit VIX'i.



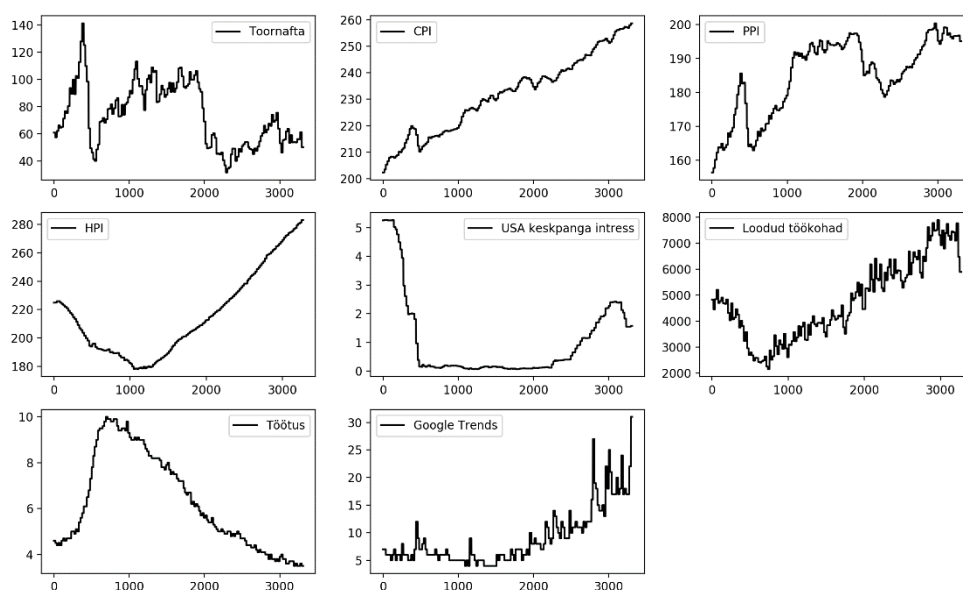
Joonis 4. Andmed 1

Allikas: autori koostatud

Ülaloleval joonisel on kaks mõõdikut autori poolt arvatud. Esmalt on autor defineerinud volatiilsust kui kauplemispäeva kõrgeima ja madalaima hinna protsentuaalset vahet. Autor valis protsentuaalse erinevuse, sest alates 2008-2009 majanduskriisist on S&P 500 väärtus kasvanud ja kui keskmine volatiilsus on 1%, siis see 1% võib reaalarvuna olla väga erinev, sest 1% 1000-st (vt Joonis 4. S&P 500 sulgemishindade esimesi vaatluseid) ja 1% 3000-st (vt Joonis 4. S&P 500 sulgemishindade viimaseid vaatluseid) on reaalarvuna kolmekordse erinevusega. Seetõttu on mõistlikum kasutada protsentuaalset volatiilsust. Antud mõõdik on ühtlasi käesoleva töö ennustatav ehk sõltuv muutuja. Teiseks autori poolt arvatud mõõdikuks on libisev standardhälve, mille arvutamisel kasutatakse eelneva 20 vaatluse (kauplemispäeva) volatiilsust.

Joonise visuaalsel uurimisel on näha, et volatiilsus on tugevas korrelatsioonis libiseva standardhälbe, tehingumahtude ja volatiilsusindeksiga. Ühtlasi on tugevas korrelatsioonis S&P500 aktsiaindeksi sulgemishinnad, kulla, Dow Jones'i, ja Nasdaq'i aktsiaindeksiga. Korrelatsioon on nõrgem dollariga. Antud seoseid on näha ka korrelatsioonimaatriksist (vt Joonis 6.). Lisaks sellele on näha, et 2008-2009 aastatel esineb võimenduse efekt – hindade langedes kasvab volatiilsus.

Joonisel 5. on kuvatud järgmised andmed: Toornafta – West Texas Index’i toornafta barreli hind, CPI – tarbijahinnaindeks, PPI – tootjahinnaindeks, HPI – kinnisvarahinnaindeks, USA keskpanga intressimäär, loodud töökohtade arv, töötuse määr, Google Trends – Google otsingumootori märksõna otsingu põhinev indeks, mille märksõnaks on S&P 500. Sarnaselt hinna andmetele on näha, et makroökonomilised andmed on nii negatiivselt kui ka positiivselt seotud S&P 500 aktsiaindeksiga, väiksem seos on näha toornafta puhul.



Joonis 5. Andmed 2

Allikas: autori koostatud

Eelneva kahe lõigu põhjal oli nii tehingumõõdikute kui ka makroökonomiliste andmete puhul näha tugevaid seoseid S&P 500’ga. Antud seosed ning nende seoste tugevused leiavad kinnitust Jooniselt 6., millel on kuvatud töös kasutatavate andmete korrelatsioonimaatriks. Korrelatsioon on mõõdik, mis kirjeldab kahe mõõdiku vahelist seost. Korrelatsiooni koefitsient on mõõdik, mis kirjeldab kahe mõõdiku vahelise korrelatsiooni tugevust ja suunda. Korrelatsiooni koefitsiendid saavad olla positiivsed või negatiivsed vahemikus -1 kuni 1, kus 1 on väga tugev positiivne korrelatsioon ning -1 on väga tugev negatiivne korrelatsioon ja 0 tähistab kahe mõõdiku vahelise seose puudumist. Korrelatsioonimaatriks on tabel, mis koosneb andmestiku vahelistest korrelatsiooni koefitsientidest.

Nagu hinnaandmete lõigus sai mainitud tugevat korrelatsiooni libiseva standardhälve, tehingumahtude ja VIX'i ja volatiilsuse vahel on antud seos saanud kinnitust ka allolevas korrelatsioonimaatriksis, kus vasakul üleval nurgas asuvad eelnevalt nimetatud muutujad on tugevas positiivses korrelatsioonis. See omakorda viitab sellele, et kui ühe eelnevalt nimetatud mõõdiku väärtus kasvab, siis kasvab ka teiste eelnevalt nimetatud mõõdikute väärtus.

Teine huvitav nähtus asetseb korrelatsioonimaatriksi ülemisel real, mis kirjeldab volatiilsuse suhteid aktsiaindeksite ja makroökonomilised näitajatega. Nimelt on näha, et kui volatiilsus kasvab, siis aktsiaindeksid langevad. Nagu Jooniselt 1. oli näha on ka korrelatsioonimaatriksilt näha, et andmetes esineb esimeses peatükis mainitud volatiilsuse võimenduse efekti. Seega kui volatiilsuse kasv tähendab hindade liikumist, siis võimenduse efekti puhul on eeldatavasti tegemist hindade langusega. Ühtlasi on näha, et volatiilsuse kasv omab negatiivset mõju makroökonomilistele andmetele. Ainukesena makroökonomilistest andmetest on töötuse määr volatiilsusega positiivses suhtes.

Kolmandana on allolevalt korrelatsioonimaatriksilt selgelt näha, et aktsiaindeksid ja dollar on väga tugevas positiivses korrelatsioonis makroökonomiliste andmetega, välja arvatud töötus, mis on negatiivses seoses. Seega kui aktsiaindeksite või dollari väärtus kasvab, siis tõenäoliselt kasvavad makroökonomiliste andmete väärtused ja sama kehtib vastupidiselt makroökonomiliste andmete väärtuste kasvades.

	Volatiilsus	Libisev STD	Tehingumahud	VIX	Dow Jones 30	Dollar	Nasdaq	Töökohad	Kuld	CPI	PPI	HPI	Intressimäär	Töötus	Google	Toornafta
Volatiilsus	1,00	0,74	0,66	0,81	-0,39	-0,21	-0,37	-0,31	-0,24	-0,35	-0,31	-0,24	-0,01	0,20	-0,02	0,07
Libisev STD	0,74	1,00	0,51	0,84	-0,36	-0,18	-0,35	-0,28	-0,25	-0,34	-0,33	-0,20	0,01	0,17	0,05	0,01
Tehingumahud	0,66	0,51	1,00	0,68	-0,42	-0,19	-0,38	-0,40	-0,23	-0,35	-0,37	-0,28	-0,17	0,36	-0,07	-0,01
VIX	0,81	0,84	0,68	1,00	-0,55	-0,30	-0,52	-0,48	-0,28	-0,50	-0,46	-0,37	-0,08	0,38	-0,12	0,05
Dow Jones 30	-0,39	-0,36	-0,42	-0,55	1,00	0,76	0,99	0,93	0,23	0,91	0,63	0,91	0,21	-0,80	0,82	-0,43
Dollar	-0,21	-0,18	-0,19	-0,30	0,76	1,00	0,80	0,77	0,05	0,75	0,30	0,77	0,01	-0,68	0,71	-0,80
Nasdaq	-0,37	-0,35	-0,38	-0,52	0,99	0,80	1,00	0,92	0,25	0,94	0,64	0,90	0,13	-0,77	0,83	-0,48
Töökohad	-0,31	-0,28	-0,40	-0,48	0,93	0,77	0,92	1,00	0,09	0,82	0,51	0,91	0,33	-0,89	0,79	-0,48
Kuld	-0,24	-0,25	-0,23	-0,28	0,23	0,05	0,25	0,09	1,00	0,50	0,75	-0,11	-0,56	0,22	0,05	0,19
CPI	-0,35	-0,34	-0,35	-0,50	0,91	0,75	0,94	0,82	0,50	1,00	0,82	0,73	-0,16	-0,59	0,73	-0,36
PPI	-0,31	-0,33	-0,37	-0,46	0,63	0,30	0,64	0,51	0,75	0,82	1,00	0,31	-0,34	-0,26	0,37	0,19
HPI	-0,24	-0,20	-0,28	-0,37	0,91	0,77	0,90	0,91	-0,11	0,73	0,31	1,00	0,46	-0,89	0,87	-0,57
Intressimäär	-0,01	0,01	-0,17	-0,08	0,21	0,01	0,13	0,33	-0,56	-0,16	-0,34	0,46	1,00	-0,56	0,24	-0,07
Töötus	0,20	0,17	0,36	0,38	-0,80	-0,68	-0,77	-0,89	0,22	-0,59	-0,26	-0,89	-0,56	1,00	-0,68	0,45
Google Trends	-0,02	0,05	-0,07	-0,12	0,82	0,71	0,83	0,79	0,05	0,73	0,37	0,87	0,24	-0,68	1,00	-0,53
Toornafta	0,07	0,01	-0,01	0,05	-0,43	-0,80	-0,48	-0,48	0,19	-0,36	0,19	-0,57	-0,07	0,45	-0,53	1,00

Joonis 5. Korrelatsioonimaatriks

Allikas: autori koostatud

Tabelis 4. on kuvatud töös kasutatavate andmete kirjeldavat statistikat. Antud tabelis on iga andmerea kohta 8 kirjeldava statistika mõõdikut, mis on defineeritud järgnevalt: 1) Vaatluste arv – näitab vaatluste arvu iga muutuja kohta, 2) Keskmine – andmerea keskmine väärtus, 3) Standardhälve – andmerea standardhälve, 4) Min – andmerea väikeseim väärtus, 5) 25% - andmerea esimene kvartiil, 6) 50% - andmerea teine kvartiil, mida tihtipeale nimetatakse ka mediaaniks, 7) 75% andmerea kolmas kvartiil, 8) Max – andmerea maksimaalne väärtus.

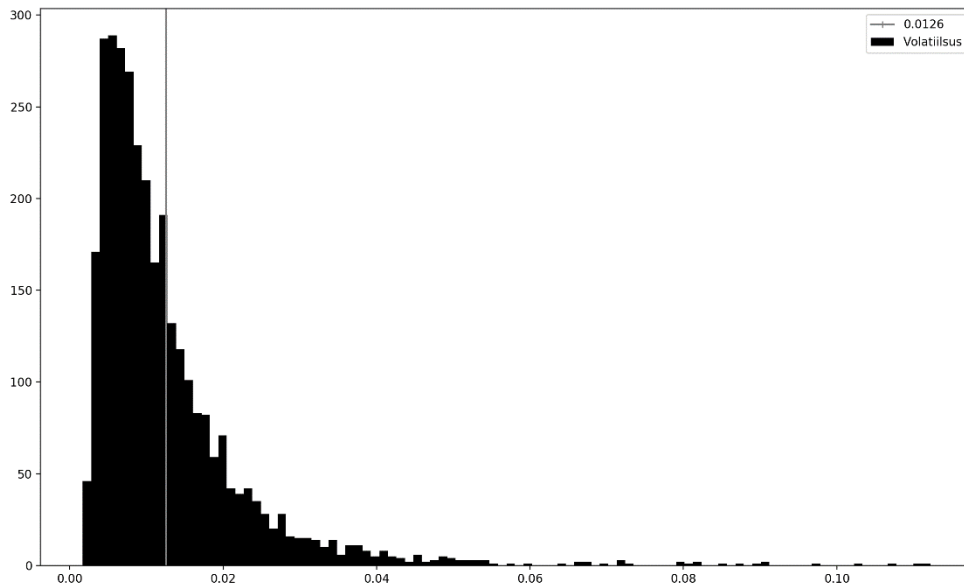
Tabel 4. Töös kasutatavate andmete kirjeldav statistika

	Vaatlu ste arv	Kesk- mine	Standard- hälve	Min	25%	50%	75%	Max
S&P 500 volatiilsus	3223	1,26%	1,06%	0,17 %	0,62 %	0,97 %	1,52 %	11,22 %
Libisev standardhälve	3223	0,54%	0,38%	0,12 %	0,31 %	0,45 %	0,61 %	2,98%
Tehingumahud milj	3223	3946,84	1093,51	1025, 00	3268, 35	3692, 41	4320, 46	11456, 23
Volatiilsusindeks Dow Jones	3223	19,3	9,1	9,1 6547,	13,5 12152	16,6 15176	22,1 18486	80,9 28645,
USA dollar Nasdaq	3223	16120,5	5378,5	0 71,3	,5 79,7	,1 82,9	,7 95,3	3 103,3
Loodud töökohad	3223	4163,8	1975,5	1268, 6	2523, 3	3587, 6	5244, 6	9022,4
Kuld	3223	4772,1	1517,5	2157, 0	3575, 0	4606, 0	5867, 0	7884,0
Toornafta	3223	1235,8	267,2	635,8 1	1109, 1	1255, 0	1356, 2	1898,1
Tarbijaaindeks	3223	74,1	22,8	31,6 202,4	53,6 218,3	72,7 233,1	93,5 241,4	141,1 257,3
Tootjaaindeks	3223	231,6	14,2	156,4 178,1	176,8 191,2	187,6 209,3	194,2 235,7	200,4 283,0
Kinnisvarahinnain- deks	3223	185,0	11,1	178,1	191,2	209,3	235,7	283,0
Keskpanga intressimäär	3223	216,1	29,8	178,1	191,2	209,3	235,7	283,0
Töötusemäär	3223	1,0	1,4	0,1	0,1	0,2	1,7	5,3
Google trends	3223	6,3	2,1	3,5	4,6	5,7	8,2	10,0
	3223	8,8	4,8	4	5	7	11	27

Allikas: autori koostatud

Kui selgitada Tabeli 4. sisu võttes aluseks S&P 500 volatiilsuse rea, siis on sellelt näha, et antud muutuja kohta on 3223 vaatlust, nagu ka kõigi teiste andmeridade kohta. S&P 500 keskmine volatiilsus on 1,26%, mis tähendab, et keskmine kõrgeima ja madalaima hinna protsentuaalne vahe antud ajavahemiku jooksul on 1,26%. Kui keskmine volatiilsus on 1,26%, siis volatiilsuse standardhällbeks on 1,06%, mis tähendab, et umbkaudselt 68% vaatlustest jäävad vahemikku

1,26% +- 1,06%. Kvartiilide puhul on näha, et 25% vaatlustest jääb vahemikku 0,17% - 0,62%. 50% vaatlustest jääb vahemikku 0,17% – 0,97%. 75% vaatlustest jääb vahemikku 0,17% - 1,52%. Volatiilsuse andmerea ekstreemumiteks on 0,17% ja 11,22%, mis tähendab, et kõige madalama volatiilsusega kauplemispäevaks oli 0,17% ning kõige kõrgema volatiilsusega kauplemispäevaks oli 11,22% .



Joonis 6. Volatiilsuse jaotus

Allikas: autori koostatud

Joonisel 6. on kuvatud S&P 500 kauplemispäevade kõrgeima ja madalaima päeva protsentuaalset vahet, ehk volatiilsust graafilise jaotusena. Lähemal vaatlusel on leiate kinnitust volatiilsuse asümmeetria olemasolule, kus on näha, et kõige rohkem vaatlusi on esinenud 0% lähedal ning ekstreemsematel juhtudel on volatiilsus tõusnud üle 10%. Kui normaaljaotuse puhul paiknevad vaatlused tihedalt ümber keskmise, siis antud juhul esineb volatiilsuse asümmeetria selles, et suur osa vaatlustest asuvad keskmisest volatiilsusest vasakul pool. Joonisel 6. on kauplemispäevade keskmine volatiilsus tähistatud vertikaalse halli joonega, mille väärtuseks on 1,26%.

4.2. Tulemused

Käesolev peatükk annab ülevaate GARCH mudeli ja tehisnärvivõrkude valimiväliste ennustuste tulemustest. Ülevaade jaguneb kolme alampeatükki, millest esimeses selgitatakse mudelite seadistust ning üldiseid tulemusi. Kuna antud töö raames on oluline teada, kuidas käitub mudelite ennustusvõime erinevates turuolukordades, on kaks teist alampeatükki mõeldud mudelite tulemuste analüüsimiseks erinevates turuolukordades. Autor on grupeerinud turuolukorrad kaheks – madala volatiilsusega kauplemispäevad ja suure volatiilsusega kauplemispäevad. Sellest tulenevalt analüüsitakse teises alampeatükis madala volatiilsusega vaatluste ennustusi ning kolmandas alampeatükis analüüsitakse kõrge volatiilsusega vaatluste tulemusi. Varasemalt on volatiilsuse eristamiseks kasutatud mitmeid viise. Granger ja Hyung (1999) kasutasid erinevate volatiilsusega perioodide eristamiseks ruutkumulatiivsete iteratsioonide algoritmi ehk ICSS'i (*iterative cumulative sum of squares*). Conrad ja Kleen (2019) kasutasid oma töös volatiilsuse eristamiseks kvartiile. Nimelt defineerisid nad madalaks volatiilsuseks vaatlused mis olid vahemikus 0%-25%, keskmiseks volatiilsuseks vaatlused vahemikus 25%-75% ja kõrgeks volatiilsuseks vaatlused vahemikus 75%-100%. Käesolevas töös kasutatakse Conradi ja Kleeni (2019) meetodit ehk volatiilsust eristatakse kvartiilide abil.

4.2.1. Mudelite seadistus ja tulemused

Antud peatükk annab ülevaate mudelite seadistustest ning nende tulemustest. Modelleerimisel kasutatavaid andmeid on iga sõltumatu muutuja kohta 3223 vaatlust vahemikul 03.01.2007-28.02.2020. Varasemates töödes ei ole üheselt välja toodud parimat vaatluste arvu või osakaalu andmetest, mida peaks kasutama valimiväliste ennustuste puhul. Küll aga erinevates töödes on kasutatud valimiväliste ennustuste jaoks erinevad osakaale, mis enamasti jäävad vahemikku 10-30% kõikidest vaatlustest. Käesolevas töös kasutatavaid andmeid on iga mõõdiku kohta 3223 vaatlust, neist 0-2500 vaatlust perioodil 03.01.2007-05.12.2016 on mõeldud mudelite konstrueerimiseks ning vaatlused 2501-3223 perioodil 05.12.2016-28.02.2020 kasutatakse valimiväliste ennustuste andmetena. Sellest tulenevalt on mudeli konstrueerimise andmete ja valimiväliste andmete suhe ümardatult 75%/25%. Seega 75% andmetest on kasutatud 25%

ennustuste sooritamise jaoks. Mudelite tulemusi võrreldakse peatükis 1.3 tutvustatud MSE, MAPE, MAE, MdAPE ja MdAD mõõdikutega.

GARCH mudeli seadistusel lähtus autor Tsai (2006) tööst, milles leiti, et kõige levinum ja efektiivsem mudel on GARCH(1,1) parameetritega mudel. Sellest tulenevalt kasutab autor käesolevas töös ühe autoregressiivse viitajaga ja ühe libiseva keskmise viitajaga mudelit. Autor koostas mudelid nii standardiseeritud andmetega, kui ka standardiseerimata andmetega. Andmete standardiseerimiseks kasutas autor Z-skoori, mis näitab mitme standardhälbe kaugusel on üks vaatlus andmerea keskmisest.

$$Z = \frac{x_i - \bar{x}}{s} \quad (15)$$

kus

Z – Z-skoor

x_i – vaadeldava vaatluse väärtus

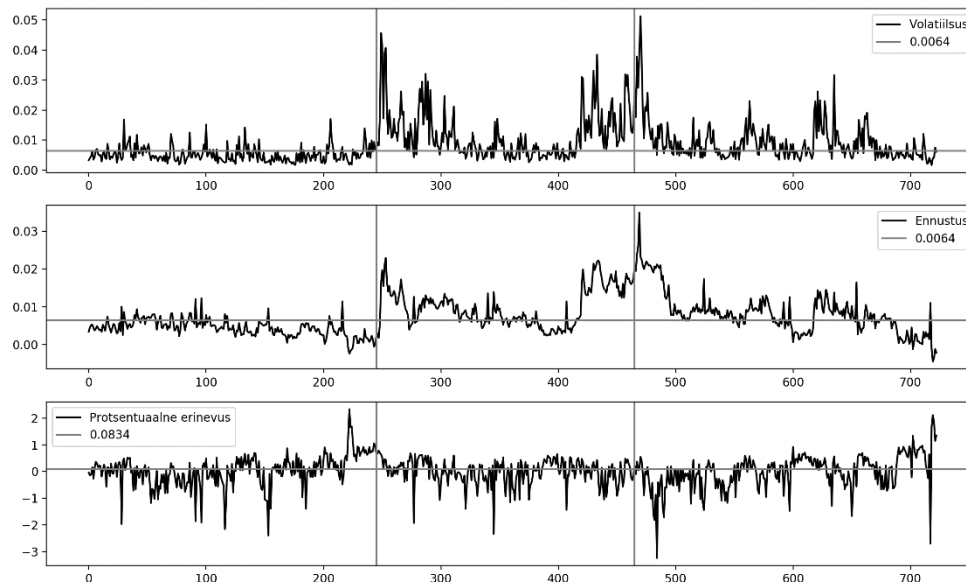
\bar{x} – andmerea keskmine väärtus

s – andmerea standardhälve

GARCH mudel andis märkimisväärselt halvemad tulemused standardiseeritud andmetega. Märkimisväärse erinevuse all peab autor silmas standardiseeritud andmetega mudeli MAPE't, mis oli ligikaudu 190%, seevastu, standardiseerimata andmetega mudeli oli MAPE 42,95%. Sellest tulenevalt otsustas autor GARCH mudeli koostamisel kasutada standardiseerimata andmeid.

Joonis 7. kuvab joonise ülemises plokis ennustusperioodi tegelikku volatiilsust, keskmises plokis on kuvatud GARCH mudeli ennustused ning alumises plokis kuvatakse tegeliku volatiilsuse ja GARCH mudeli ennustuse protsentuaalset erinevust. Horisontaalsed jooned tähistavad vaadeldava andmerea mediaani. Antud juhul on näha, et tegeliku volatiilsuse mediaaniks on 0,64%. See on täpselt sama, mis GARCH ennustuste mediaan. Alumise graafiku mediaan tähistab protsentuaalset viga, milleks on 8,3%. Graafiku alumine plokk on mõeldud täiendavaks selgituseks, sest kuigi tegeliku volatiilsuse ja ennustatud volatiilsuse mediaan on sama, esineb tegelikult protsentuaalseid vigu ning kohati on need vead suured. Alumise ploki lähemal vaatlusel on näha, et tihti esineb ennustusi, mille viga ületab 100% vea taset. Vertikaalsed jooned tähistavad vaatlust vahetult enne kõrge volatiilsusega vaatlust. Nende abil saab võrrelda kas ja kuidas erinevad kõrgeima volatiilsusega vaatluse vead. Lähtudes alumisest

plokist on hea tõdeda, et hallide vertikaalsete joonte läheduses on ennustuste vead väikesed. Seega võib järeldada, et kõrge volatiilsusega vaatlused ei tähenda proportsionaalselt suuremaid tegeliku väärtuse ja ennustatud väärtuse erinevusi.

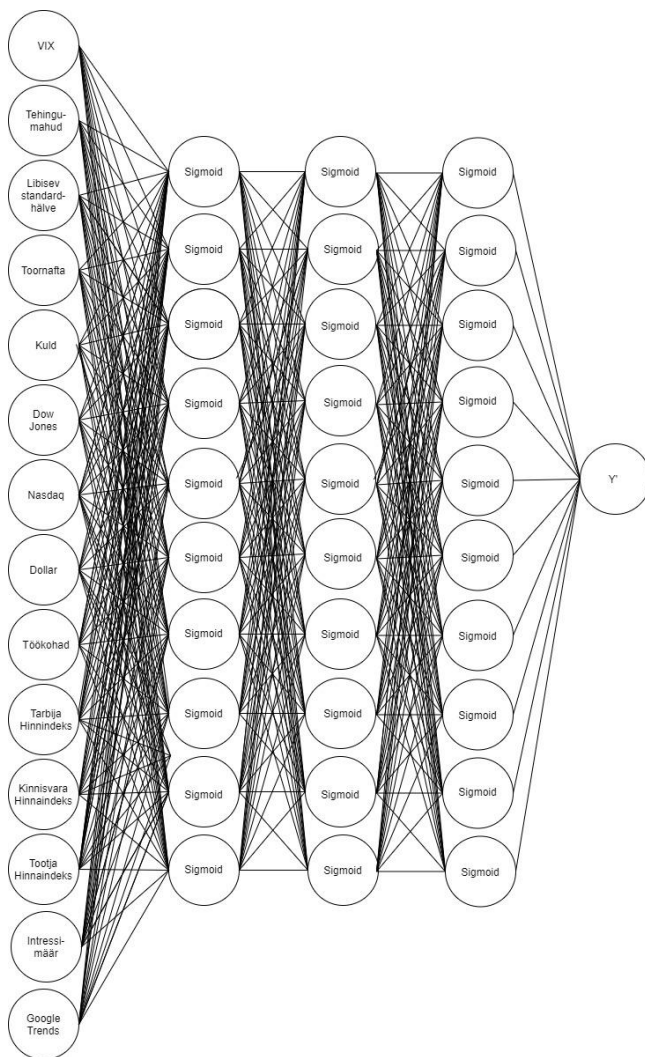


Joonis 7. Volatiilsus, GARCH mudeli ennustus ning protsentuaalne erinevus

Allikas: autori koostatud

Võrreldes GARCH mudeliga on optimaalse tehisnärvivõrkude mudeli saavutamiseks vaja sooritada märkimisväärselt rohkem katsetusi. Osaliselt on see seotud sellega, et tehisnärvivõrgud on väga dünaamilised mudelid – nendega saab lahendada nii regressiooniprobleeme kui ka klassifikatsiooniprobleeme. See ühtlasi tähendab seda, et modelleerijal on väga palju erinevaid võimalusi närvivõrgu mudeli seadistamisel. Käesolevas töös kasutatakse mitmekihilist pertseptron närvivõrku, mis koosneb sisendandmete kihist, peidetud kihtidest, kihtidevahelistest aktiveerimisfunktsioonidest ning tagasilevi protsessist. Antud mudelite seadistamine on keeruline ning nõuab palju katsetamist, sest kuigi tehisnärvivõrkude alal on läbi viidud palju erinevaid töid, ei ole loodud reeglit, mis töötaks üheselt kõikide närvivõrkude või andmete puhul. Siinkohal on oluline lähtuda varasemast kirjandusest ning antud töö raames lähtus autor Karsoliya (2012) tööst, kes pakkus välja, et mitmekihilise pertseptron mudeli optimaalne peidetud kihtide arv on 3 ning igas kihis peaks olema neuronite arvuks 70%-90% sisendandmete andmeridade arvust. Sellest tulenevalt on käesoleva töö peidetud kihtide arvuks 3 ja iga kihis on 10 neuronit. Kuna sisendandmeid ehk

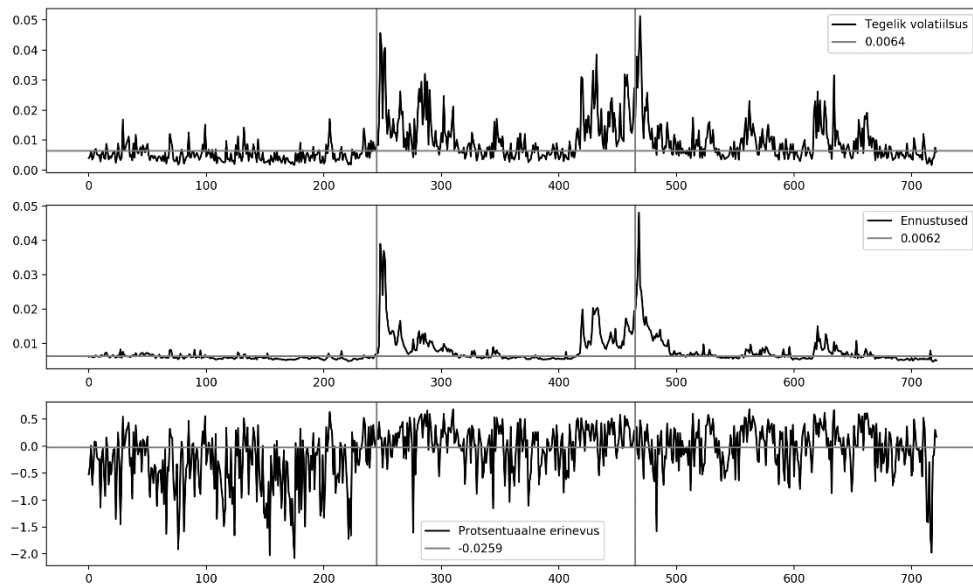
sõltumatuid muutujaid on 14, siis 10 neuronit, moodustab 71,4% sõltumatutest muutujatest ühe peidetud kihi kohta. Arvestades, et ennustatav volatiilsus on vahemikus 0.00 – 0.15, lähtus autor Sibi, Jones ja Siddarth (2013) tööst, milles viidati, et sigmoidfunktsioon on parim aktiveerimisfunktsioon kui ennustatav väärtus jääb vahemikku 0-1. Õppimise määrana kasutas autor Python programmeerimiskeele tehisnärivõrkude modelleerimise paketi Keras poolt vaikimisi sätestatud määra – 0,001. Tagasilevi protsessi vigade minimeerimise algoritmina kasutas autor ruutjuur ruutvea levikut. Tehisnärivõrgud kasutavad mudeli treenimisel iteratsioone ja S. Hosein ja P. Hosein (2017) töö põhjal otsustas autor kasutada 400 iteratsiooni. Seejuures 1 iteratsioon tähendab seda, et mudel on ühe korra andmestiku läbi töödeldud. Seega enne valimiväliste ennustusteni jõudmist on mudeli kontrueerimise käigus 2500 vaatluse suuruse andmestik läbi töödeldud 400 korda.



Joonis 8. Töös kasutatav tehisnärivõrgu arhitektuur

Allikas: autori koostatud

Joonisel 8. on kuvatud käesolevas töös kasutatav närvivõrgu arhitektuur. Sellelt on näha, et sisendandmete kihis on 14 sõltumatut muutujat. Antud muutujad kantakse edasi kolme peidetud kihti, kus igas kihis on 10 neuronit ning iga neuroni kaalutud sisendsignaalide summat aktiveerib sigmoidfunktsioon. Viimasena saadakse ennustus väljundandmete kihist, milles on üks neuron ehk ennustatav väärtus.



Joonis 9. Tegelik volatiilsus, tehisnärvivõrkude ennustused ning protsentuaalne erinevus

Allikas: autori koostatud

Nagu ka GARCH mudeli puhul, viis autor läbi mudeli konstrueerimise nii standardiseeritud kui ka standardiseerimata andmetega. Erinevalt GARCH mudelist, kus standardiseerimata andmed andsid parema tulemuse, olid standardiseeritud andmed tehisnärvivõrkude puhul täpsemad ennustajad kui standardiseerimata andmed. Sarnaselt GARCH mudeli tulemuste joonisele (vt Joonis 7.) kuvab Joonis 9. tehisnärvivõrkude tulemusi, kus ülemisel graafikul on tegelik volatiilsus, keskmisel graafikul tehisnärvivõrkude ennustus ning alumisel graafikul tegeliku väärtuse ja ennustuse protsentuaalne erinevus. Kui võrrelda tegeliku volatiilsuse ja ennustatud volatiilsuse andmeridade mediaane, siis on näha, et tehisnärvivõrkude mediaan volatiilsus 0,62% on veidi väiksem kui tegeliku volatiilsuse mediaan 0,64%. Sarnaselt GARCH mudelile on näha, et suurema volatiilsusega vaatlused ei mõjuta ennustuste täpsust. Seega kui volatiilsus kasvab, siis vaatluse tegeliku ja ennustuse protsentuaalne erinevus jääb mediaani

lähedale. Lisaks on näha, et tehismärkivõrkude tegeliku ja ennustuse protsentuaalne erinevuse ekstreemumid jäävad enamasti 100% vea piiridesse, seevastu GARCH mudelil esines kuni 300% vigu.

Tabel 5. Mudelite võrdlus kõikide valimiväliste vaatluste põhjal

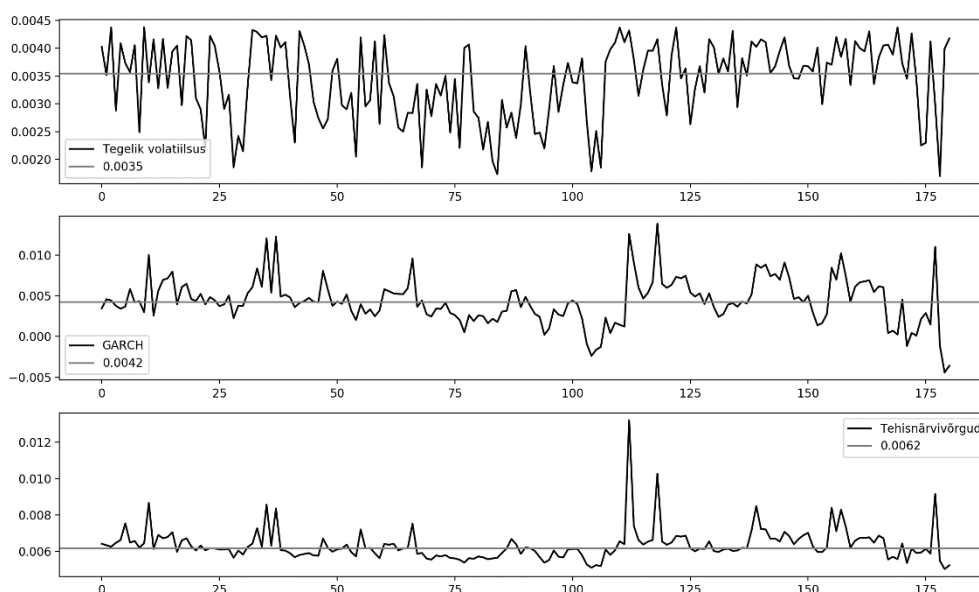
	GARCH	Tehis- märkivõrgud
MAE	0,0033	0,0029
MSE	$2,408^{-05}$	$1,889^{-05}$
MAPE	42,95%	37,86%
MdAPE	8,3%	-2,6%
MdAD	0,0034	0,0029

Allikas: autori koostatud

Tabelis 5. on kuvatud mudelite vaheline võrdlus kasutades mõõdikuid, mida käsitleti esimese peatüki alampeatükis 1.3.. Kui lähtuda asjaolust, et mida väiksem on veamõõdiku väärtus, seda täpsemad on ennustused, siis saab Tabel 5. põhjal järeldada, et kõikide mõõdikute puhul on tehismärkivõrkude mudel täpsemate ennustustega. Detailsemalt vaadates on näha, et märkimisväärsed erinevused esinevad MdAPE ja MAPE mõõdikute puhul. MAPE puhul on erinevus pigem vähem märkimisväärne, kuid MdAPE puhul on näha, et tehismärkivõrkude ennustused on märkimisväärselt täpsemad. Arvestades, et MdAPE on mediaani komponendiga, siis viitab see viitab sellele, et GARCH mudelil esines ekstreemsete vigadega ennustusi.

4.2.2. Madala volatiilsuse ennustamine

Autor eristas kõrge ja madala volatiilsusega kauplemispäevi kvartiilide abil. Sellest tulenevalt jagunesid kauplemispäevad kaheks: esimene kvartiil (0%-25% vaatlustest) – madal volatiilsus, neljas kvartiil (75%-100% vaatlustest) – kõrge volatiilsus. Käesolevas alampeatükis võrreldakse GARCH mudeli ja tehismärkivõrkude ennustusi tegeliku volatiilsusega madalal volatiilsuse perioodil.



Joonis 10. Madala volatiilsuse ennustused

Allikas: autori koostatud

Joonis 10. jaguneb kolme ploki: 1) ülemises ploki on kuvatud tegelikud madala volatiilsusega vaatlused, 2) keskmises ploki on GARCH mudeli ennustused, 3) alumises ploki on tehisnärvivõrkude tulemused. Igas ploki on kuvatud horisontaalne joon, mis tähistab iga andmerea (ploki) mediaani. Joonise lähemalt vaatlusel selgub, et madala volatiilsusega kauplemspäevad on väga stohhastilised võrreldes GARCH ja tehisnärvivõrkude ennustustega. Ühtlasi on vaatlustelt 110 – 120 näha mõlema mudeli puhul tavapärasest ennustust kõrgemaid ennustusi, kuid tegeliku volatiilsuse graafikul ei ole antud vaatluste puhul näha tavapärasest kõrgemat volatiilsust.

Tegelike kauplemspäevade volatiilsuse mediaaniks on 0.35%, GARCH mudeli ennustused on sellele lähedal omades ennustuste mediaani väärtusega 0.42%. Seevastu tehisnärvivõrkude mediaan ennustus on ligemale kaks korda suurem – 0.62%, mis on võrreldes tegeliku mediaan volatiilsusega on märkimisväärne erinevus.

Tabel 6. Mudelite võrdlus madala volatiilsusega vaatluste põhjal

	GARCH	Tehis- närvivõrgud
MAE	0,0021	0,0029
MSE	$8,091^{-06}$	$9,351^{-06}$
MAPE	61,85%	92,45%
MdAPE	-24,71%	-83,42%
MdAD	0,0021	0,0007

Allikas: autori koostatud

Jooniselt 10. jääb mulje nagu GARCH mudel oleks parem ennustaja, kuid ainuüksi visuaalne hinnang ja ennustuste mediaani arvutamine ei anna terviklikku ülevaadet mudelite ennustustest. Tabel 6. kuvab mõlema mudeli madala volatiilsusega vaatluste mõõdikuid. Võrreldes Tabel 6. mõõdikuid Tabel 5. mõõdikutega on näha, et GARCH puhul on MAE vähenenud, tehisnärvivõrkude MAE on mõlema tabeli puhul sama. Sama kalduvust on näha ka MdAD puhul, kus tehisnärvivõrkude ennustused on väga väikese hälbumisega. See viitab sellele, et madala volatiilsusega turuolukorras jaotuvad tehisnärvivõrkude ennustused väga kitsasse vahemikku. See leiab kinnitust ka Joonis 10. ülemise ja alumise ploki võrdlusel, millelt on näha, et tegelik volatiilsus on väga stohhastiline aga tehisnärvivõrkude ennustused ei ole nii stohhastilised. Stohhastilisuse puudumine viib tehisnärvivõrkude MAPE ja MdAPE mõõdikud võrreldes Tabel 5.-ga kordades suuremaks.

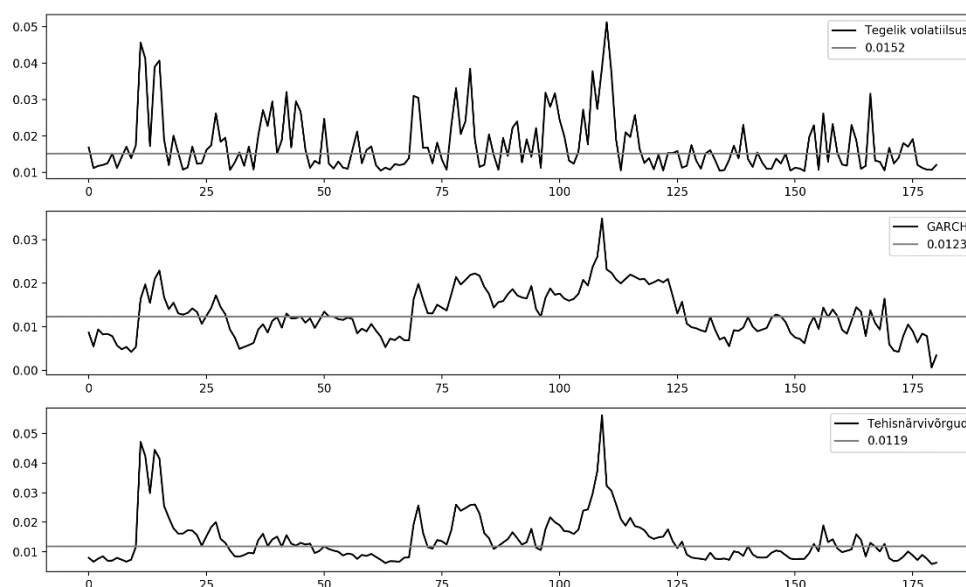
Kokkuvõttes on madala volatiilsusega turuolukorras mõlema mudeli absoluutsed vead väiksemad, kuid vastupidiselt on mõlema mudeli protsentuaalsed vead suuremad. Võrreldes eelneva alampeatüki tulemustega on madala volatiilsusega turuolukorras tehisnärvivõrkude mudeli ennustused halvenenud märkimisväärselt. Sellest tulenevalt võib järeldada, et kumbki mudel ei ole hea madala volatiilsuse ennustaja, kuid GARCH mudel ennustas madalat volatiilsust tehisnärvivõrkudest märkimisväärselt paremini.

4.2.3. Kõrge volatiilsuse ennustamine

Eelnevas alampeatükis kirjeldati madala volatiilsusega kauplemispäevade ennustuste tulemusi. Käesolevas alampeatükis analüüsib autor kõrge volatiilsusega kauplemispäevade ennustuste

tulemusi. Lähtudes eelnevas alampeatükis esile toodud volatiilsuse eristamisest kvartiilide abil analüüsitakse käesolevas alampeatükis neljanda kvartiili (75%-100% vaatlused) ennustusi.

Sarnaselt eelneva alampeatüki Joonis 10.-le on Joonisel 11. kasutatud sama loogikat mudelite ennustuste võrdlemisel. Sellest tulenevalt on ülemises plokis tegelik volatiilsus, keskmises GARCH mudeli poolt ennustatud volatiilsus ning alumises plokis tehiskäivõrkude ennustused. Ühtlasi kuvavad horisontaalsed sirgjooned iga andmerea (plokki) mediaani. Kui madala volatiilsuse puhul oli tegeliku volatiilsuse ja GARCH ennustuse mediaanid väikese erinevusega, siis kõrge volatiilsusega vaatluste puhul on erinevus suurem. Nimelt tegelik kauplemispäevade volatiilsuse mediaan on 1,52% ning GARCH ennustuste mediaan on 1,23%. Võrreldes madala volatiilsusega ennustuste erinevusega on vahe kasvanud 0,07% pealt 0,29% peale. Ühtlasi on tehiskäivõrkude ennustuste mediaan tegelike ennustuste mediaanist suurema erinevusega kui GARCH ennustused. Nimelt on tehiskäivõrkude mediaan ennustus 1,19%, mis võrreldes tegeliku volatiilsuse mediaaniga on 0,33% võrra väiksem. Nagu GARCH mudeli ennustused, on kõrge volatiilsusega vaatluste puhul kasvanud ka tehiskäivõrkude ennustuste mediaani erinevus. Kui madala volatiilsusega vaatluste puhul oli tegeliku volatiilsuse mediaani ja tehiskäivõrkude ennustuste mediaani erinevuseks 0,27%, siis kõrge volatiilsusega vaatluste puhul on erinevuseks 0,33%. Võrreldes GARCH mediaani muutusega võib järeldada, et tehiskäivõrkude mediaani muutus on püsinud muutumatuna.



Joonis 11. Kõrge volatiilsuse ennustused

Allikas: autori koostatud

Kõrge volatiilsuse ennustuste tulemused on esitatud Tabelis 7.. Võrreldes kõikide valimiväliste ennustustega (vt Tabel 5.) ja madala volatiilsuse ennustustega (vt Tabel 6.) on Tabel 7.-st näha, et mõlema mudeli ennustused on kõige täpsemad kõrge volatiilsusega turuolukordades. Nagu ka Tabel 5. puhul oli näha, et tehisnärvivõrgud on veidi täpsemate ennustustega kui GARCH, siis võib sama järelduse teha ka kõrge volatiilsuse ennustamisel. Kokkuvõttes on mõlema mudeli ennustused kõige täpsemad kõrge volatiilsusega olukorras. Seejuures on tehisnärvivõrkude ennustused veidi paremad, kuid erinevus GARCH mudeli täpsusega ei ole märkimisväärne.

Tabel 7. Mudelite võrdlus kõrge volatiilsusega vaatluste põhjal

	GARCH	Tehis- närvivõrgud
MAE	0,0062	0,0053
MSE	$6,501^{-05}$	$4,453^{-05}$
MAPE	33,77%	30,37%
MdAPE	26,22%	25,01%
MdAD	0,0054	0,0048

Allikas: autori koostatud

Tabelis 8. on kokku koondatud mõlema mudeli tulemused erinevates turuolukordades. Tabelist joonistub kõikide moodsikute puhul selgelt välja kaks asja: 1) kõrge volatiilsusega vaatluste ja kõikide vaatluste puhul on täpsemate ennustustega tehisnärvivõrkude mudel, 2) madala volatiilsusega vaatluste puhul ei ole kumbki mudel väga hea ennustaja. Tehisnärvivõrgud on küll kõrge volatiilsusega vaatluste ja kõikide vaatluste põhjal täpsemate ennustustega, kuid erinevus GARCH mudeliga ei ole märkimisväärne. Ühtlasi ei ole kumbki mudel väga hea madala volatiilsuse ennustaja aga GARCH mudeli ennustused on märkimisväärselt täpsemad kui tehisnärvivõrkude ennustused.

Tabel 8. Mudelite võrdlus

	Kõrge volatiilsus		Madal volatiilsus		Kõik vaatlused	
	GARCH	Tehis- närvivõrgud	GARCH	Tehis- närvivõrgud	GARCH	Tehis- närvivõrgud
MAE	0.0062	0.0053	0.0021	0.0029	0,0033	0,0029
MSE	$6,501^{-05}$	$4,453^{-05}$	$8,091^{-06}$	$9,351^{-06}$	$2,408^{-05}$	$1,889^{-05}$
MAPE	33.77%	30.37%	61.85%	92.45%	42,95%	37,86%
MdAPE	26.22%	25,01%	-24,71%	-83,42%	8,3%	-2,6%
MdAD	0,0054	0,0048	0,0021	0,0007	0,0034	0,0029

Allikas: autori koostatud

Kui vastata küsimusele kumb mudel on parem, siis kahel põhjusel selleks on pigem tehisnärvivõrgud. Esiteks tehisnärvivõrgud on täpsemad kõrge volatiilsusega vaatluste ja kõikide vaatluste puhul. Teiseks, kuigi GARCH oli madala volatiilsuse ennustamisel märkimisväärselt parem ei oma madala volatiilsuse ennustamine ärilist väärtust, sest madal volatiilsus ei mõjuta oluliselt riskide haldamist ja varaklasside osakaalusid. Kuna kõrge volatiilsuse ennustamine on äriselt oluline, siis käesoleva töö põhjal peaks volatiilsuse esimeseks valikuks olema tehisnärvivõrgud. Sellele lisab ka kaalu asjaolu, et tehisnärvivõrgud olid täpsemad ennustajad kõikide vaatluste puhul.

KOKKUVÕTE

Riskide haldamine ning optimaalsete vara osakaalude leidmine on olulised teemad eelkõige kõrge volatiilsusega perioodidel. Lähtudes masinõppe ja tehisnärvivõrkude pidevast populaarsuse kasvust ning arvutite võimekuse arengust on võimalik luua kõrgetasemelisi lahendusi antud probleemide lahendamisel. Autoregressiivsed mudelid on kujunenud levinuimaks volatiilsuse ennustamise mudeliteks alates esmasest tutvustamisest 1980. aastatel. Seevastu on tehisnärvivõrgud tänu mitmekülgsele võimelised lahendama keerukaid probleeme. Käesolev töö uuris eelnevalt nimetatud mudelite tulemusi S&P 500 volatiilsuse ennustamisel nii kõrge kui ka madala volatiilsusega turuolukordades.

Käesolevas töös selgitati seoses volatiilsuse ennustamise ja modelleerimisega järgnevaid teemasid. Esimene peatükk andis ülevaate volatiilsuse ennustamisest, sellega seotud nähtustest ja terminitest ning mudelite tulemuste hindamisest. Teises peatükis tutvustati mudeleid, mida käesolevas töös kasutati. Selle raames andis autor ülevaate GARCH ja tehisnärvivõrkude mudelite taustast ja ajaloost. Kolmandas peatükis käsitleti töös kasutatavate mudelite varasemaid empiirilisi tulemusi. Neljandas peatükis tutvustati töös kasutatavaid andmeid ning viidi läbi empiiriline analüüs S&P 500 aktsiaindeksi volatiilsuse ennustamisel.

Antud töö eesmärgiks oli ennustada järgmise päeva S&P 500 aktsiaindeksi volatiilsust tehisnärvivõrkudega. Selleks, et tehisnärvivõrkude ennustuste täpsusest paremini aru saada, kasutati etalonina GARCH mudelit. Ühtlasi oli töö teiseks eesmärgiks leida, et kas ja kuidas erinevad mõlema mudeli ennustused erinevates turuolukordades. Erinevaid turuolukordasid defineeriti kvartiilide abil, kus esimeses kvartiilis olevad vaatlused on madala volatiilsusega vaatlused ja neljandas kvartiilis olevad vaatlused on kõrge volatiilsusega vaatlused.

Selleks, et leida antud uurimusküsimustele vastused, kasutas autor GARCH (1,1) mudelit standardiseerimata andmetega. Erinevalt GARCH mudelist kasutati tehisnärvivõrkude mudelit standardiseeritud andmetega. Mitmekihilise pertseptron mudeli koostamisel kasutati kolme peidetud kihiga närvivõrku. Mudeli igas peidetud kihis oli neuronite arvuks $\frac{3}{4}$ sisendandmete arvust ja iga kihi aktiveerimisfunktsioonina kasutati sigmoidfunktsiooni. Tagasilevi protsessi vigade optimeerijaks oli ruutjuur ruutvea leviku algoritm ning mudeli treenimiseks kasutati 400 iteratsiooni.

Töö empiirilises osas leidis autor, et S&P 500 andmete põhjal esineb asümmeetrilisest volatiilsust ja paksu saba tõenäosusjaotust, volatiilsuse kuhjumist ja võimenduse efekti ning keskväärtusele koondumist. Mudelite hindamisel kasutati MAPE, MSE, MdAPE ja MdAD mõõdikuid. Kõrge volatiilsusega ja kõikide vaatluste puhul oli kõikide mõõdikute alusel täpsemate ennustustega tehisnärvivõrkude mudel, kuid erinevus GARCH mudeliga ei olnud märkimisväärne. Madala volatiilsusega vaatluste puhul oli kõikide mõõdikute alusel märkimisväärselt täpsemate ennustustega GARCH mudel. Kuna kõrge volatiilsuse ennustamine on äriliselt oluline, siis võib järeldada, et mitmekihiline pertseptron mudel on antud analüüsi põhjal parem volatiilsuse ennustaja. Antud järelduse täiendavaks kinnituseks on asjaolu, et tehisnärvivõrgud olid täpsemad ennustajad kõikide vaatluste puhul.

Riskijuhtimise kontekstis ilmselt ei ole oluline, kui mudel ennustab volatiilsuseks 1,03% olukorras kus tegelik volatiilsus on 1%, sest nii väike erinevus ei oma suurt rolli varade väärtuse kujunemisel. Sellest tulenevalt oleks üheks alternatiiviks edasiste uuringute mõistes lahendada volatiilsuse ennustamine klassifitseerimise probleemina. Selle abil oleks ilmselt võimalik saavutada veelgi parem ennustuste täpsus. Antud mõttel on alust kui riskijuhte huvitavad suurusjärgud, mitte konkreetne arvuline väärtus. Modelleerimise näol võiks volatiilsuse grupeerida nelja kvartiili. Hüpoteesilise volatiilsusena võiksid grupid olla järgnevad: 1.) 0%-0,249%, 2) 0,250% - 0,499%, 3) 0,500%-0,749%, 4) üle 0,75%.

Varasematest töödest ja käesolevast tööst on näha, et volatiilsus jaotub tihedalt keskmise volatiilsuse ümber. Kui ei ole oluline, et grupid oleksid vaatluste arvult ühtsed, siis oleks mõistlik kasutada standardhälvet. Standardhälbe abil saaks grupid luua näiteks nii, et kõik vaatlused mis on esimesest standardhälvest kõrgemal (suuremad kui 1) on kõrge volatiilsusega vaatlused ja kõik vaatlused mis on esimesest standardhälvest madalamad (väiksemad kui -1) on madala volatiilsusega vaatlused.

Varasemates töödes on kasutatud ka teisi närvivõrgu arhitektuure nagu LSTM (*Long Short Term Memory*) ja ja konvolutsioonilised tehisnärvivõrgud. Autor ei ole leidnud, et neis töödes oleks ennustusi eristatud kõrge ja madala volatiilsusega gruppidesse. Sellest tulenevalt oleks edasiste uuringutena hea teada kui täpseid ennustus eelnevad närvivõrkude mudelid annavad erinevates turuolukordades.

SUMMARY

FORECASTING S&P 500 VOLATILITY WITH ARTIFICIAL NEURAL NETWORKS

Mehis Tambu

Volatility forecasting is important in the context of risk management and asset allocation. Models that forecast high volatility situations with high precision are especially important. From business perspective high volatility means high risk and if risk managers and investors have good understanding of how the volatility might develop on the next day, then they might have a chance to manage their risks and allocate their assets accordingly.

This thesis gives an overview of volatility forecasting and volatility forecasting models as follows. First chapter explains different ways to define volatility, and presents stylized facts in the context of volatility and gives an overview of measures that can be used for comparing different models. Second chapter introduces models that are used in the thesis – GARCH (*Generalised AutoRegressive Conditional Heteroskedasticity*) and multilayer perceptron artificial neural network. Third chapter investigates previous literature in the context of both models and their forecasting performance. Lastly, fourth chapter an overview of the data used in the empirical analysis and second subchapter contained the results of empirical analysis. Empirical analysis was divided into three sub chapters. First chapter analyzes all the samples for out of sample forecasts. Second and third subchapter analyzes low volatility and high volatility forecasts. Volatility was divided into groups with quartiles – first quartile 0-25% was grouped as low volatility, fourth quartile 75%-100% was grouped as high volatility.

Aim of the thesis was to forecast the volatility of S&P 500 stock index with neural networks. In order to understand the neural network performance, the author decided to use a common benchmark model for volatility forecasting – GARCH model. Therefore, a research question was raised – how neural network volatility forecasts differ from GARCH model forecasts. Additional research question was to find out how both models forecasts differ from each other in different market conditions.

To answer these questions, a GARCH model was used with one autoregressive lag and one moving average lag. Multilayer perceptron artificial network was used with three hidden layers, $\frac{3}{4}$ of input layer neurons formed 10 neurons in all hidden layers and each neuron in each layer was activated with sigmoid activation function. Backpropagation was used with 400 iterations with root mean square propagation as optimizing algorithm.

In the empirical analysis chapter, it was found that there are signs of volatility clustering, reversion to the mean, leverage effect, asymmetric volatility and fat tail distributions. Forecasts revealed that in the general out of sample forecasts artificial neural networks provide slightly more accurate results than GARCH. Same conclusion can be drawn from high volatility forecast observations where both models performed better when compared to general out of sample results. In low volatility forecast samples both models did not perform well, but GARCH model performed much better when compared to artificial neural networks.

In the context of risk management, it is probably not important if model forecasts volatility of 1.03% where actual volatility turns out to be 1.00%. Since high volatility forecast precision is important, it might be reasonable to forecasts certain groups instead of certain values. For example, 3% daily volatility is considered quite high therefore when forecasting certain value, we would probably get an output of +- something to the actual value, but if we group that 3% as high volatility, we would get a binary output which reflects whether next day will be low to medium or high volatility. Therefore, suggestion for further research would be to group volatility and train models on grouped volatility.

There are many ways to differentiate volatility or market conditions. Previous studies show that quartiles and ICSS (iterative cumulative sum of squares) algorithm have been used to differentiate volatility. Alternative way to eliminate observations around the mean could be using Z-score or standard deviations and see how they perform.

Previous studies show that many different artificial neural network architectures have been used, but there are no studies with these architectures that analyze forecasts in different market conditions. Therefore, it might be a good idea to use LSTM (Long Short-Term Memory) or convolutional networks to forecast volatility and differentiate volatility in order to see how these artificial neural networks perform in different market conditions.

KASUTATUD ALLIKATE LOETELU

- Abdalla, S.Z.S., Suliman, Z., 2012. Modelling stock returns volatility: Empirical evidence from Saudi Stock Exchange. *International Research Journal of Finance and Economics*, 85, 166-179.
- Abdalla, S.Z.S., Winker, P., 2012. Modelling stock market volatility using univariate GARCH models: Evidence from Sudan and Egypt. *International Journal of Economics and Finance*, 4(8), 161-176.
- Aliyu, S.U.R., 2009. Impact of oil price shock and exchange rate volatility on economic growth in Nigeria: An empirical investigation.
- Bal, C., Demir, S., Aladag, C.H., 2016. A comparison of different model selection criteria forecasting EURO/USD exchange rates by feed forward neural network. *International Journal of Computing, Communication and Instrumental Engineering*, 3, pp.271-275.
- Bardgett, C., Gourier, E., Leippold, M., 2019. Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics*, 131(3), 593-618.
- Becker, R., Clements, A.E., Doolan, M.B. and Hurn, A.S., 2015. Selecting volatility forecasting models for portfolio allocation purposes. *International Journal of Forecasting*, 31(3), 849-861.
- Beckers, S., 1983. Variances of security price returns based on high, low, and closing prices. *Journal of Business*, 97-112.
- Bouchaud, J.P., Matacz, A., Potters, M., 2001. Leverage effect in financial markets: The retarded volatility model. *Physical review letters*, 87(22), p.228701.
- Cai, J., 1994. A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, 12(3), 309-316.
- Chan, J.C., Grant, A.L., 2016. Modeling energy price dynamics: GARCH versus stochastic volatility. *Energy Economics*, 54, 182-189.
- Cheung, D.W.W., 2000. The impulse of stock market volatility and the market crash of October 1987. *Journal of Business Finance & Accounting*, 27(5-6), 761-776.

- Chkili, W., Hammoudeh, S., Nguyen, D.K., 2014. Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Economics*, 41, 1-18.
- Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and volatility. *Journal of financial economics*, 50(2), 125-150.
- Conrad, C., Kleen, O., 2019. Two are better than one: volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics*.
- Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues.
- Dicle, M.F., Levendis, J., 2019. Historic risk and implied volatility. *Global Finance Journal*, p.100475.
- Ding, Z., Granger, C.W., Engle, R.F., 1993. A long memory property of stock market new model. *Journal of empirical finance*, 1(1), 83-106.
- Donaldson, R.G., Kamstra, M., 1996. Forecast combining with neural networks. *Journal of Forecasting*, 15(1), 49-61.
- Donaldson, R.G., Kamstra, M., 1997. An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance*, 4(1), 17-46.
- Dzielinski, M., Rieger, M.O., Talpsepp, T., 2011. Volatility asymmetry, news, and private investors. *The handbook of news analytics in finance*, 255-270.
- Emenike, K.O., 2010. Modelling stock returns volatility in Nigeria using GARCH models.
- Engle, R., 2001. GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of economic perspectives*, 15(4), 157-168.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
- Fama, E.F., 1965. The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.
- Farley, B.W.A.C., Clark, W., 1954. Simulation of self-organizing systems by digital Chapman & Hall.

- Fong, W.M., 1998. The dynamics of DM/£ exchange rate volatility: A SWARCH analysis. *International Journal of Finance & Economics*, 3(1), 59-71.
- Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical 1 stock data. *Journal of business*, 67-78.
- Granger, C.W., Hyung, N., 1999. Occasional structural breaks and long memory.
- Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R. and Schmidhuber, J., 2016. LSTM: space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10), 2222-2232.
- Guresen, E., Kayakutlu, G., Daim, T.U., 2011. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8), 10389-10397.
- Guresen, E., Kayakutlu, G., Daim, T.U., 2011. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8), 10389-10397.
- Hamid, A., Heiden, M., 2015. Forecasting volatility with empirical similarity and Google Trends. *Journal of Economic Behavior & Organization*, 117, 62-81.
- Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics*, 64(1-2), 307-333.
- Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. *Journal of applied econometrics*, 20(7), 873-889.
- Hebb, D.O., 1949. *The organization of behavior: a neuropsychological theory*. J. Wiley; Chapman & Hall.
- H.Indla, (2019). *Volatiilsuse Modelleerimisprotsessi Lohtsustamine Stata Statistikatarkvara Programmeerimise Tulemusena*, (Magistritöö), Eesti Ettevõtkluskõrgkool Mainor ärijuhtimise teaduskond, Tallinn.
- Hosein, S., Hosein, P., 2017, April. Load forecasting using deep neural networks. In *2017 IEEE -propagating errors. nature*, 323(6088), 533-536.
- Hu, M.Y., Tsoukalas, C., 1999. Combining conditional volatility forecasts using neural networks: an application to the EMS exchange rates. *Journal of International Financial Markets, Institutions and Money*, 9(4), 407-422.

- Hu, M.Y., Zhang, G., Jiang, C.X., Patuwo, B.E., 1999. A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting. *Decision Sciences*, 30(1), 197-216.
- Ivakhnenko, A.G., 1970. Heuristic self-organization in problems of engineering and cybernetics. *Automatica*, 6(2), 207-219.
- Jain, A.K., Mao, J., Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44.
- Kambouroudis, D.S., McMillan, D.G., 2016. Does VIX or volume improve GARCH volatility forecasts?. *Applied Economics*, 48(13), 1210-1228.
- Karsoliya, S., 2012. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International Journal of Engineering Trends and Technology*, 3(6), 714-717.
- Katsiampa, P., 2017. Volatility estimation for Bitcoin: A comparison of GARCH models of models. *Economics Letters*, 158, 3-6
- Katsiampa, P., 2017. Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158, 3-6.
- Kiefer, J., Wolfowitz, J., 1952. Stochastic estimation of the maximum of a regression function. *The IEEE -propagating errors. nature*, 323(6088), 533-536.
- Klaassen, F., 2002. Improving GARCH volatility forecasts with regime-switching GARCH. In *Advances in Markov-switching models* (223-254). Physica, Heidelberg.
- Kumar, P.R., Ravi, V., 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European journal of operational research*, 180(1), 1-28.
- Kunitomo, N., 1992. Improving the Parkinson method of estimating security price volatilities. *Journal of Business*, 295-302.
- Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- Malkiel, B.G., Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.

- Malmsten, H., Teräsvirta, T., 2010. Stylized facts of financial time series and three popular models of volatility. *European Journal of pure and applied mathematics*, 3(3), 443-477.
- Mandelbrot, B., 1963. New methods in statistical economics. *Journal of political economy*, 71(5), 421-440.
- Mandelbrot, B., 1965. Self-similar error clusters in communication systems and the concept of conditional stationarity. *IEEE Transactions on Communication Technology*, 13(1), 71-90.
- McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous *bulletin of mathematical biophysics*, 5(4), 115-133.
- Mosca, A., Magoulas, G.D., 2017, April. Training convolutional networks with weight-wise adaptive learning rates. In *ESANN*.
- Pan, Z., Wang, Y., Liu, L., Wang, Q., 2019. Improving volatility prediction and option valuation using VIX information: A volatility spillover GARCH model. *Journal of Futures Markets*, 39(6), 744-776.
- Prokopczuk, M., Simen, C.W., 2014. The importance of the volatility risk premium for volatility forecasting. *Journal of Banking & Finance*, 40, 303-320.
- R.Lootus, (2016). *Forecasting of value at risk using nonlinear autoregressive volatility models methodology*. (Magistritöö) TalTechi majandusteaduskond, Tallinn.
- Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 504-512.
- Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 504-512.
- Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), p.386.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Selçuk, F., 2005. Asymmetric stochastic volatility in emerging stock markets. *Applied retarded Financial Economics*, 15(12), 867-874.

- Shcherbakov, M.V., Brebels, A., Shcherbakova, N.L., Tyukov, A.P., Janovsky, T.A. and Kamaev, V.A.E., 2013. A survey of forecast error measures. *World Applied Sciences Journal*, 24(24), 171-176.
- Sibi, P., Jones, S.A., Siddarth, P., 2013. Analysis of different activation functions using backpropagation neural networks. *Journal of Theoretical and Applied Information Technology*, 47(3), 1264-1268.
- Silvennoinen, A., Teräsvirta, T., 2009. Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model. *Journal of Financial Econometrics*, 7(4), 373-411.
- Siven, J.V., Lins, J.T., 2009. Gain/loss asymmetry in time series of individual stock prices and its relationship to the leverage effect. *arXiv preprint arXiv:0911.4679*.
- Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
- Tang, Z., De Almeida, C., Fishwick, P.A., 1991. Time series forecasting using neural Box-Jenkins methodology. *Simulation*, 57(5), 303-310.
- Tieleman, T., Hinton, G., 2012. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2), 26-31.
- Theodossiou, P., Kahya, E., Koutmos, G. and Christofi, A., 1997. Volatility reversion and structure of returns in major international stock markets. *Financial Review*, 32(2), 205-224.
- Tsai, H., Chan, K.S., 2008. A note on inequality constraints in the GARCH model. *Econometric Theory*, 24(3), 823-828.
- Tully, E., Lucey, B.M., 2007. A power GARCH examination of the gold market. *Research in International Business and Finance*, 21(2), 316-325.
- Wasserman, P D., and Tom Schwartz. Neural networks. II. What are they and why is o interested in them now?. *IEEE Expert* 3, no. 1 (1988): 10-15.
- Wiggins, J.B., 1992. Estimating the volatility of S&P 500 futures prices using the extreme-value of *Business*, 97-112.
- Wu, X., Hou, X., 2019. Forecasting volatility with component conditional autoregressive range model. *The North American Journal of Economics and Finance*, p.101078.

Yang, D., Zhang, Q., 2000. Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business*, 73(3), 477-492.

LIHTLITSENTS

Mina Mehis Tambu

1. Annan Tallinna Tehnikaülikoolila tasuta loa (lihtlitsentsi) enda loodud teose S&P 500 volatiilsuse ennustamine tehisnärvivõrkudega,

mille juhendaja on: Kaido Kepp

1.1 reprodutseerimiseks säilitamise ja elektroonilise avaldamise eesmärgil, sealhulgas TTÜ raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas TTÜ raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta kolmandate isikute intellektuaalomandi ega isikuandmete kaitse seadusest ja teistest õigusaktidest tulenevaid õigusi.

allkiri

kuupäev