

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Rain Nõmmik 176415IAAM

**MASINÕPPEL PÕHINEV
KAUPLEMISSTRATEEGIA USA
VÕLAKIRJATURUL**

Magistritöö

Juhendajad: Aleksei Talisainen
MSc

Triin Kriisa
magister

Tallinn 2019

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Rain Nõmmik

13.05.2019

Annotatsioon

Käesoleva töö eesmärk on uurida masinõppe võimalikku rakendamist kasumliku kauplemisstrateegia loomiseks USA võlakirjaturul. Töös koostatakse kaks erinevat tehnilikul närvivõrgul põhinevat LSTM mudelit. Ühes kasutatakse sisendina alusvara enda hinnarida ning teises lisatakse sisenditena makro- ja finantsandmed. Mudelite sisenditele rakendatakse erinevaid viitaegu, ühest kuni 20 päevani. Koostatud mudelite eesmärk on ennustada USA 2- ja 10-aastase võlakirja futuuri hinnaliikumise suunda järgmiseks päevaks. Vastavalt prognoositud suunale ostetakse või müüakse alusvara.

Töö tulemusena koostatud kauplemisstrateegiaid võrreldi teenitud tulu ja Sharpe suhtarvu alusel. Valdav enamus koostatud kauplemisstrateegiatest ületasid klassikalist osta ja hoia strateegiat. Kauplemismudelid olid edukamad, kui sisenditena oli kasutatud makro- ja finantsandmeid.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 53 leheküljel, 4 peatükki, 29 joonist, 3 tabelit.

Abstract

Trading strategy based on machine learning for US bond market

This master's thesis studies the possibility to use machine learning models to build profitable trading strategy on US government bond market. In this paper, two different long short-term memory artificial recurrent neural networks models are created. One of the models uses input values only from the underlying instrument and for the second model different macro economical and financial time series data was added. Different time lags were analysed, starting from one day and up to 20 days lookback was used. The purpose of the models was to predict the direction of price movement for the next day. Underlying instruments were 2- and 10-year treasury futures. Depending on the prediction of the price movement, either buy or sell trading signal was generated.

The results were compared based on the overall return and Sharpe ratio over testing period. The results show that trading strategies based on the LSTM can beat the classical buy and hold strategies when certain parameters are met. Trading models were more successful when macro economical and financial data was added to the input.

The thesis is in Estonian and contains 53 pages of text, 4 chapters, 29 figures, 3 tables.

Lühendite ja mõistete sõnastik

ReLU	Tõkestatud mittenegatiivne lineaarne funktsioon
Tanh	Hüperboolne tangens
MSE	Keskmine ruutviga
RNN	Rekurrentne närvivõrk
LSTM	Pikk lühiajalise mälu närvivõrk

Sisukord

Sissejuhatus	9
1 Investeerimine	11
1.1 Investeeringute haldamine	11
1.2 Võlakiri	12
1.3 Futuur	13
1.4 Riskid	13
1.5 Fundamentaalsed seosed võlakirjaturul	14
2 Masinõpe	16
2.1 Masinõppe algoritmid ja meetodid	16
2.1.1 Tehislikud närvivõrgud	17
2.1.2 Närvivõrgu treenimine	21
2.1.3 Rekurrentne närvivõrk	22
2.1.4 LSTM	23
3 Kirjanduse ülevaade	26
3.1 Erinevate masinõppe tehnikate võrdlus	26
3.2 Masinõpe ja traditsioonilised statistilised meetodid	27
3.3 Erinevate närvivõrkude rakendused võlakirjaturul	28
3.4 Rekurrentsete närvivõrkude rakendused finantsandmetel	29
4 Empiiriline analüüs ja tulemused	30
4.1 Mudeli kirjeldus	30
4.2 Andmed	34
4.3 Tulemused	35
4.3.1 2-aastane USA võlakirja futuur	35
4.3.2 10-aastane USA võlakirja futuur	41
4.4 Edasine töö	47
Kokkuvõte	49
Kasutatud kirjandus	51

Jooniste loetelu

Joonis 1. Pärilevivõrk [21].....	18
Joonis 2. Neuron koos sisendite (x_i), kaalude (w_i) ja väljundiga (y) [22].....	19
Joonis 3. Sigmoid funktsioon [22].....	20
Joonis 4. Tanh funktsioon [22].....	20
Joonis 5. ReLU funktsioon [22].	21
Joonis 6. RNN närvivõrk lahti laotatud kujul [19].	23
Joonis 7. LSTM raku diagramm [19].	24
Joonis 8. Minimaalne treeningu ruutviga erinevate neuronite arvuga mudelis.	31
Joonis 9. LSTM mudel programmikoodis.	33
Joonis 10. USA 2-aastase võlakirjafutuuri hind.	35
Joonis 11. LSTM(AR) 2-aastane, viitaeg 4 päeva, mudeli ruutviga.	36
Joonis 12. LSTM(AR) 2-aastane, viitaeg 4 päeva, mudeli signaalid.	36
Joonis 13. LSTM(AR) 2-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.	37
Joonis 14. LSTM(AR) 2-aastane, viitaeg 1-20 päeva, Sharpe.	38
Joonis 15. LSTM(makro) 2-aastane, viitaeg 6 päeva, mudeli ruutviga.	38
Joonis 16. LSTM(makro) 2-aastane, viitaeg 6 päeva, mudeli signaalid.	39
Joonis 17. LSTM(makro) 2-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.	39
Joonis 18. LSTM(makro) 2-aastane, viitaeg 1-20 päeva, Sharpe.	40
Joonis 19. LSTM(AR) ja LSTM(makro) 2-aastane, kumulatiivsete tootluste võrdlus. .	41
Joonis 20. USA 10-aastase võlakirjafutuuri hind.	42
Joonis 21. LSTM(AR) 10-aastane, viitaeg 10 päeva, mudeli ruutviga.	42
Joonis 22. LSTM(AR) 10-aastane, viitaeg 10 päeva, mudeli signaalid.	43
Joonis 23. LSTM(AR) 10-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.	43
Joonis 24. LSTM(AR) 10-aastane, viitaeg 1-20 päeva, Sharpe.	44
Joonis 25. LSTM(makro) 10-aastane, viitaeg 12 päeva, mudeli ruutviga.	45
Joonis 26. LSTM(makro) 10-aastane, viitaeg 12 päeva, mudeli signaalid.	45
Joonis 27. LSTM(makro) 10-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.	46
Joonis 28. LSTM(makro) 10-aastane, viitaeg 1-20 päeva, Sharpe.	46
Joonis 29. LSTM(AR) ja LSTM(makro) 10-aastane, kumulatiivsete tootluste võrdlus.	47

Tabelite loetelu

Tabel 1. Mudelite parameetrid.	33
Tabel 2. LSTM(AR) ja LSTM(makro) 2-aastane, mudelite võrdlus.....	41
Tabel 3. LSTM(AR) ja LSTM(makro) 10-aastane, mudelite võrdlus.....	47

Sissejuhatus

Võlakirjaturg on riikidele ja ettevõtetele tähtis rahastusallikas. Kogu maailma võlakirjaturu suurus on hinnanguliselt 100 triljonit dollarit. Sellest suurima osa moodustab USA võlakirjaturg, mille suuruseks hinnatakse 40 triljonit dollarit [1]. Keskmise päevane käive USA riiklikul võlakirjaturul on ligikaudu 400 miljardit dollarit [2]. Võrdluseks USA aktsiaturu suurus on 34 triljonit dollarit ning keskmine päevakäive ligikaudu 216 miljardit dollarit [3].

Ajalooliselt on võlakirjaturg pakkunud investorile stabiilset tootlust. Kuigi pikaajaline tootlus on olnud positiivne, pakub igapäevane hinnakõikumine ning aktiivne positsioonide muutmine täiendavaid võimalusi tulu teenimiseks. Võlakirjainvestorite jaoks muutub aktiivne positsioneerumine olulisemaks tõusvate baasintressimäärade keskkonnas, kus võlakirjahinnad langevad. Selleks, et hinnakõikumistest jätkusuutlikult kasu saada, kasutatakse erinevaid kauplemisstrateegiaid.

Viimastel aastatel on lisaks klassikalistele tehnilistele ja fundamentaalsetele kauplemisstrateegiatele uuritud masinõppe rakendamise võimalust investeerimisotsuste tegemiseks. Masinõppe on kiiresti arenev teadusharu. Erinevad masinõppe tehnikad võimaldavad modelleerida seoseid suuremahuliste ja keerukate andmete vahel ning seetõttu rakendatakse masinõppe meetodeid ning lähenemisi väga laias teemades valdkonnas nagu näiteks kõnetuvastus, isesõitvad autod, meditsiin ja finantsturud.

Töö eesmärk on uurida, kas masinõppet on võimalik rakendada USA võlakirjaturul kasumliku kauplemismudeli loomiseks. Eesmärgi saavutamiseks võrreldakse masinõppel põhineva strateegia tulemuslikkust klassikalise osta ja hoida strateegiaga. Selleks, et eesmärki täita, on esmalt oluline välja selgitada, millist masinõppe meetodikat on mõistlik rakendada finantsturgude aegridade analüüsimisel.

Töös käsitletakse kahte strateegiat. Esimeses kasutatakse sisendandmetena võlakirjafuturi enda aegrida ning teises lisatakse sisendandmetele USA makromajanduslikud andmed ning erinevad finantsturgude näitajad. Mõlemat strateegiat

rakendatakse 2- ja 10-aastase lõpptähtajaga USA valitsusvõlakirja futuurile. Töös lähtutakse ratsionaalse investori seisukohast, kelle eesmärk on teenida maksimaalset tulu riskiühiku kohta.

Empiirilises analüüsis kasutatud makromajanduslikud ja finantsturgude andmed koguti USA Föderaalreservi Saint Louis reservpanga majandusandmete andmebaasist ja Bloombergist.

Töö jaguneb neljaks suuremaks peatükiks. Esimeses peatükis keskendutakse investeerimise põhilistele kontseptsioonidele ning kirjeldatakse lühidalt kauplemisega seotud instrumente ning riske võlakirjaturul. Teises peatükis antakse ülevaade masinõppe põhimõtetest ja meetoditest. Täpsemalt kirjeldatakse töös kasutatud mudelit. Kolmandas peatükis antakse ülevaade varasemalt tehtud akadeemilistest uuringutest ning saavutatud tulemustest masinõppe kasutamisel finantsturgudel. Viimases peatüki esitatakse käesoleva uurimustöö tulemused. Töös analüüsitakse erinevate sisendandmetega ja viitaegadega mudeleid erinevatel alusinstrumentidel.

1 Investeerimine

Järgnevas peatükis käsitletakse aktiivse investeerimise olulisemaid printsiipe ja finantsturgude ning võlakirjadega seotud põhimõtteid. Täpsemalt kirjeldatakse Ameerika Ühendriikide valitsuse võlakirju, nendega seotud tuletisinstrumente ning investeerimisega seotud riske.

1.1 Investeeringute haldamine

Investeeringute haldamise saab jagada kahte kategooriasse – aktiivne ja passiivne investeerimine. Passiivne investeerimine on portfelli halduse strateegia, kus investor ostab sobiva alusvara vastavalt mõnele indeksile või ostab endale sobiva instrumendi selleks, et seda hoida pikaajaliselt. Aktiivne investeerimine on portfelli halduse strateegia, mis hõlmab järjepidevat turgude jälgimist ja investeeringute aktiivset muutmist vastavalt turunägemusele [4].

Investeerimisstiilide poolest võib aktiivse varahalduse jagada subjektiivseks, mis põhineb majanduse ja finantsturgude kvalitatiivsel analüüsil, ja süstemaatiliseks, kus turu käitumist ennustatakse kvantitatiivsete seoste abil. Nii subjektiivsete kui süstemaatiliste otsuste tegemisel kasutatakse sisenditena fundamentaalseid või tehnilisi näitajaid ja nende kombineerimist [4].

Süstemaatilise investeerimise puhul kasutatavad kindlad kauplemisreeglid tagavad selle, et investeerimisotsused on objektiivsed, järjepidevad, ei sisalda emotsioone, tuginevad põhjendatud seostel ja võimaldavad teadlikku riskijuhtimist. Lisaks hajutab süstemaatiline investeerimine subjektiivselt tehtud investeerimisotsuseid.

Varahalduse seisukohast on väga oluline portfell erinevate riskide suhtes hajutada. Läbi hajutamise paraneb investeerimisportfelli tulu–riski profiil ja sõltuvalt investeeringutest võib paranemine olla seejuures märkimisväärne. Sama kehtib ka aktiivse halduse kohta – mida hajutatam on aktiivne portfell erinevate vähekorreleeritud strateegiatega vahel, seda parema tulemuse portfell saavutab [5].

Üheks oluliseks võimaluseks portfelliinvesteeringuid hajutada on lisaks erinevatele investeerimisstrateegiatele ja kasutatavatele sisenditele ka hajutamine varaklasside kaupa. Klassikalistes investeerimisportfellides on kombineeritud aktsia- ja võlakirjainstrumente. Vähem konservatiivsed investorid on portfelli lisanud ka alternatiivseid varaklasse, näiteks erakapitali investeeringuid, kinnisvara ja toormeid [4].

Käesolevas uurimuses on keskendunud USA valitsusvõlakirjaturule, kus aktiivseid positsioone võetakse sageli rahalise võlakirjainvesteeringu asemel futuuri kasutades.

1.2 Võlakiri

Võlakiri on võlakohustus, mis seab väljaandjale kohustuse tagasi maksta laenatud summa ning sellega kaasnevad intressid. Valitsused, kohalikud omavalitsused ja ettevõtted annavad välja võlakirju, kui on vaja kaasata kapitali. Investor, kes ostab valitsuse võlakirja, laenab oma raha võlakirja väljaandjale ning saab vastu võlakirja. Tüüpiliselt makstakse võlakirjade eest perioodilisi intressimakseid ning lõpptähtpäeval makstakse tagasi laenusumma [6].

Tavapäraselt jagatakse USA riiklikud võlakirjad lõpptähtaja järgi kolme kategooriasse:

- lühiajalised võlakirjad,
- keskmise pikkusega võlakirjad,
- pikaajalised võlakirjad.

Lühiajalised võlakirjad on kuni aastase tähtajaga võlakirjad. Tavapäraselt on alla aastased võlakirjad nullkupongiga. See tähendab, et võlakirju müüakse alla nimiväärtuse. Keskmise pikkusega võlakirjad on kuni 10-aastase tähtajaga võlakirjad. Pikaajalisteks võlakirjadeks liigitatakse üle 10-aastase tähtajaga võlakirjad [6].

Võlakirjade hinnastamiseks arvutakse välja nende teoreetiline väärtus. Kõik tuleviku rahavood diskonteeritakse käesolevasse päeva. Selliselt saadakse võlakirja hind [7]:

$$P = \sum_{n=1}^N \frac{C}{(1+r)^n} + \frac{M}{(1+r)^N} \quad (1)$$

kus P – võlakirja hind, C – aastane kupongimakse, r – diskontomäär, N – aastate arv võlakirja lõpptähtajani, M – võlakirja nimiväärtus.

Lisaks võlakirja hinnale on võlakirjadel tavapäraselt ka tulumäär. Võlakirja hind ja tulumäär liiguvad vastupidises suunas. Kui võlakirja hind tõuseb, siis tulumäär langeb ning vastupidi. See on otseselt seotud valemiga (1). Tõusev intressimäär suurendab diskontomäära ning seeläbi võlakirjahind langeb. Kõige lihtsam viis intressimäära avaldamiseks on järgnev valem [7]:

$$rc = \frac{c}{p} \times 100 \quad (2)$$

kus rc – intressimäär.

1.3 Futuur

Futuur on üks tuletisväärtpaberite liik. Tuletisväärtpaber annab õiguse või kohustuse, osta või müüa alusvara kindlaks määratud hinnaga kindlaks määratud kuupäeval. Võlakirjafutuuriid on turul kaubeldavad instrumendid, mille alusvaraks on võlakirjad. Järgneva valemiga on võimalik kindlaks teha võlakirjafutuuri hind [8]:

$$F_0 = (S_0 - I)e^{rT} \quad (3)$$

kus F_0 – futuuri hind, S_0 – võlakirja hetkehind, I – võlakirja kupongide nüüdisväärtus futuurilepingu jooksul, T – aeg futuurlepingu lõpptähtajani, r – riskivaba intressimäär.

1.4 Riskid

Investeeringisega on alati lisaks tulule seotud ka riskid. Võlakirjadesse investeerimine ei ole seejuures erand. Nagu kõigi teiste instrumentidega, on ka võlakirjadega seotud hinnarisk. See tähendab, et alusvara hind võib liikuda investori jaoks ebasobivas suunas. Ühel päeval ostetud võlakiri võib olla järgmisel päeval vähem väärt. Võlakirjade puhul seisneb erinevus selles, et kui võlakirja hoitakse lõpptähtajani, siis hinnarisk ei ole investori jaoks probleem, kuna investor saab lõpptähtajal tagasi kogu laenatud summa koos intressiga [6].

Teine oluline risk, mida võlakirjadesse investeerimisel tuleb arvestada, on krediidirisk. See tähendab, et laenuvõtja sattub makseraskustesse ning laenuandja ei pruugi oma raha tagasi saada. Krediidiriski hindavad tavaliselt reitinguagentuurid, andes võlakirja väljaandjale vastava reitingu. Arenenud riikide valitsusvõlakirjasid peetakse väga madala krediidiriskiga võlakirjadeks [6].

Kolmas risk, mida võlakirjade puhul tavapäraselt välja tuuakse, on likviidsusrisk. See tähendab, et investoril ei ole võimalik võlakirja maha müüa või ei ole võimalik mõislikel tingimustel müüa. Samuti nagu krediidirisk on USA valitsusvõlakirjadel madal likviidsusrisk [6].

Ratsionaalse investori jaoks on oluline tulude kõrval kõikide riskidega arvestada. Enimkasutatud suhtarv tulu ja hinnariski mõõtmiseks on Sharpe suhe [4]:

$$SR = \frac{r}{\sigma} \quad (4)$$

kus r – perioodi tootlus, σ – tootluse standardhälve.

1.5 Fundamentaalsed seosed võlakirjaturul

Intressid ning võlakirjaturud on sõltuvad majanduse käekäigust. Intressid tõusevad ja langevad vastavalt majanduse tsüklitele. Makromajanduslikud näitajad võiksid seetõttu anda informatsiooni, kuhu suunas intressid tulevikus liikuda võiks.

Akadeemilistes artiklites on laialdaselt uuritud makromajanduslike näitajate ja finantsturu andmete mõju võlakirjade hinnadünaamikale ja tootlusele. Uuringud näitavad, et majanduse fundamentaalsete näitajate abil on võimalik edukalt ennustada USA valitsusvõlakirjade tootlust ning valimivälisel testimisel on mitmetest makronäitajatest kombineeritud faktoritel tugev ennustusvõimekus, mis on ajas stabiilne ja statistiliselt oluline [9]. Mitmed akadeemilised uurimused [10], [11], [12] kinnitavad, et USA valitsusvõlakirjade tootlus on ennustatav finantsturu andmete pealt.

Majanduse ja võlakirjaintresside vahelisi seoseid on uuritud palju, kasutades erinevaid meetodeid. Traditsiooniliste ökonomeetriliste mudelite (nt. lihtne lineaarne regressioonimudel, ARMA mudelid) puhul kasutatavad matemaatilised ja statistilised meetodid eeldavad üldjuhul üsna rangelt mudeli lineaarsust tema parameetrite suhtes [13]. See tähendab eeldust, et uuritava tunnuse hetkeväärtus avaldub lineaarse kombinatsioonina kas tema enda minevikuväärtustest või muude kirjeldavate tunnuste hetke- või minevikuväärtustest.

Kuigi matemaatiliselt on lineaarse sõltuvusega protsesse kõige lihtsam kirjeldada ja mudeli lihtsus on omakorda oluline nii rakendatavuse kui ka tõlgendatavuse aspektist, on finantsvaradel traditsiooniliste meetodite rakendamine raskendatud. See tuleneb

asjaolust, et finantsvarade hinnaread on mittelineaarsed [14]. Veelgi enam, lisaks mittelineaarsusele on finantsturud ja finantsvarade hinnad oma põhiolemuselt keerukad, dünaamilised, mitteparameetrilised ja sisaldavad palju müra [15]. Finantsandmete mitteparameetrilisus tähendab, et andmete üldkogumi jaotus ei ole teada. Selliste keeruliste ja mittelineaarsete aegridade puhul klassikaliste statistiliste lineaarsete meetodite prognoositäpsus langeb. Mitmed masinõppe meetodid on aga võimelised keerukaid mittelineaarseid seoseid kirjeldama [15].

2 Masinõpe

Järgnevas peatükis käsitletakse peamiseid masinõppe kontseptsioone ning sellega kaasnevaid põhimõisteid. Antakse detailsem ülevaade töös kasutatud masinõppe meetodist.

Masinõpe tähendab arvuti võimet õppida talle ette antud andmete põhjal sooritama mingit ülesannet. Arvutile antakse ette andmed, mille vahel leitakse seosed ning neid seoseid kasutatakse hilisemate ülesannete lahendamiseks. Masinõpe on kiirelt kasvav tarkavarateaduste haru. See ei tulene ainuüksi asjaolust, et andmete mahud kasvavad, vaid ka valdkonnad, kus masinõpet rakendatakse, on oluliselt laienenud. Masinõpet kasutatakse nii bioloogias kui astroloogias, samuti leiab see kasutust meie igapäevaelus. Interneti otsingumootorid proovivad arvata meie jaoks kõige sobilikumaid tulemusi ning rämpsposti filtreerivad algoritmid [16].

2.1 Masinõppe algoritmid ja meetodid

Masinõppe jagatakse peamiselt kolme kategooriasse [17]:

- juhendatud õpe,
- juhendamata õpe,
- stiimulõpe.

Juhendatud õppe korral õpitakse ette antud andmete ja sõltuva muutuja vahel selgeks võimalik seos. Ajalooliste andmete peal õpitakse selgeks andmehulga ja ennustatava väärtuse vahelised seosed ning seda rakendatakse hiljem sõltuva muutuja prognoosimisel. Juhendamata õppe korral ei anta ette seost ning algoritmid püüavad luua ette antud andmetest struktuuri. Seda lahendust kasutatakse suurandmete analüüsil ja andmetevaheliste seoste tuvastamisel. Juhendamata õpe on abiks andmete kirjeldamisel. Stiimulõpe ehk kinnitusega õpe on sarnane juhendamata õppele. Tegemist on nn käitumusliku õppemudeliga, kus süsteemi mitte ei õpetata treeningandmetega, vaid õpe toimub katse-eksituse meetodil. Erinevus seisneb selles, et stiimulõppele seatakse teatud

reeglid ja piirangud, millised tulemused on soovitud ja millised tulemusi peaks vältima [17].

Juhendatud õppe ülesanded jagunevad kaheks – klassifikatsioon ja regressioon. Klassifikatsiooni ülesande korral püütakse algoritmi abil tuvastada, millisesse klassi objekt kuulub. Seega ennustatav väärtus on diskreetne. Näiteks punane või sinine, haige või mitte. Regressiooni ülesande korral proovitakse ennustada elemendi reaalselt väärtust, näiteks mis on väärtus dollarites [18].

Enim kasutatavad meetodid juhendatud õppe korral on logistiline regressioon (ingl *logistic regression*), naiivne Bayesi klassifitseerija (ingl *naive Bayes*), tugivektorklassifitseerija (ingl *support vector machines*), otsustusmets (ingl *random forests*) ja tehislikud närvivõrgud (ingl *artificial neural networks*). Tehislikud närvivõrgud omakorda kuuluvad masinõppe ühte alamklassi, milleks on süvaõpe.

Süvaõpe on üks masinõppe tehnikaid, mille eesmärk on tundma õppida järjest keerulisemaid mustreid hierarhiliste kihtide kaupa [19].

Käesolevas töös on kasutatud juhendatud õpet, täpsemalt masinõppe ühte alaliiki, milleks on süvaõpe. Eesmärk on ajalooliste andmete pealt ennustada järgneva perioodi hinnadünaamikat. See tähendab, et ajalooliste andmete pealt õpitakse selgeks seoseid ning neid rakendatakse sõltuva muutuja prognoosimisel.

2.1.1 Tehislikud närvivõrgud

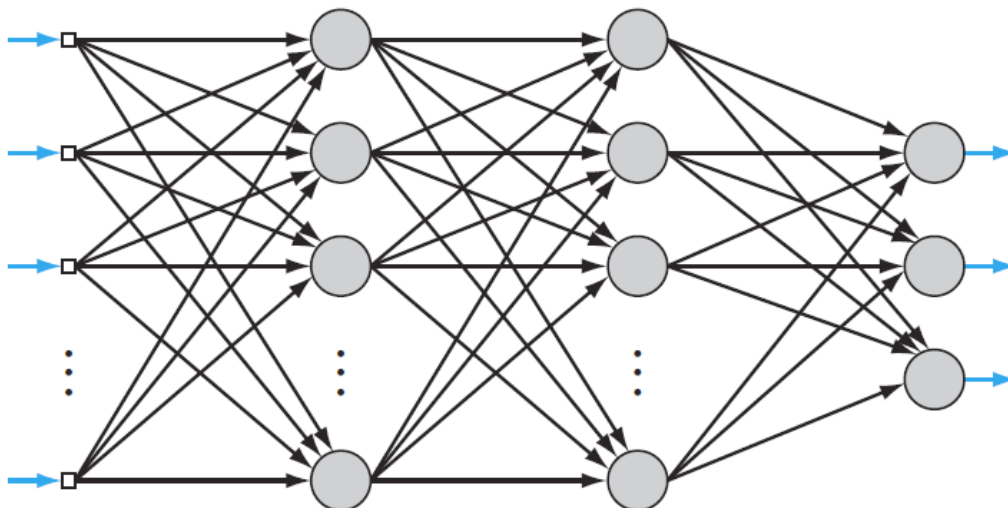
Närvivõrgud on masinõppe üks tehnikaid. Oma esialgse idee ning ülesehituse on närvivõrgud saanud inimese närvivõrkude kujutamisest. Sellega prooviti jäljendada inimese õppimisvõimet. Reaalsuses ollakse aga inimese ajutegevuse jäljendamisest veel kaugel ning otseselt selle tegevust jäljendada ei proovita. Pigem keskendutakse närvivõrkude tehnika arendamisel kindla ülesande lahendamiseks vajaminevate konfiguratsioonide täpsustamisele [19].

Närvivõrkude metoodikat kasutatakse muude rakendusvaldkondade hulgas ka statistilises analüüsis ja andmete modelleerimises. Lahendatakse nii klassifitseerimis- kui ka regressiooniprobleeme. Enimlevinud kasutusala on kõnetuvastus, tekstiliste sümbolite tundmine, meditsiin ning finantsandmete prognoos [20].

Närvivõrk on oma olemuselt matemaatiline mudel, mille ülesanne on informatsiooni töötlemine. Informatsiooni töödeldakse läbi elementide, mida nimetatakse neuroniteks. Närvivõrke iseloomustakse kahe peamise omaduse järgi – närvivõrgu arhitektuur ja õppimise meetod [20].

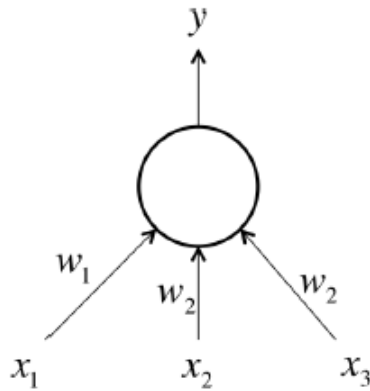
Närvivõrkude arhitektuur kirjeldab, millised on neuronite omavahelised sidemed. Peamisteks närvivõrkude arhitektuuri tüüpideks on pärilevivõrk (ingl *feedforward network*) ja rekurrentne närvivõrk. Lisaks kirjeldab arhitektuur mitmest kihist närvivõrk koosneb ning mitu neuronit asub igas närvivõrgu kihis [19].

Järgnevalt on kujutatud kahe varjatud kihiga kolmekihiline pärilevivõrk (Joonis 1). Esimene kiht on sisendite kiht ning seda võrgu kihtide loendamisel arvesse ei võeta. Sellise närvivõrgu arhitektuuri korral liigub informatsioon vasakult paremale, ehk sisenditest väljundite suunas.



Joonis 1. Pärilevivõrk [21].

Närvivõrgu iga sisend on ühendatud neuroniga (Joonis 2) ning iga neuron on omakorda seotud järgmise kihi neuroniga. Viimase kihi neuronid on seotud väljunditega [21].



Joonis 2. Neuron koos sisendite (x_i), kaalude (w_i) ja väljundiga (y) [22].

Seega, igasse neuronisse tulevad sisendid x_1, \dots, x_n , mis korrutatakse kaaludega w_1, \dots, w_n ja summeeritakse omavahel [22]:

$$z = \sum_{i=0}^n x_i w_i \quad (5)$$

Sageli lisatakse sellele summale nihe b [22].

Neuronid on erinevate kihtide vahel omavahel seotud ning nad suhtlevad omavahel signaale edastades. Omavahelised sidemed võivad olla tugevamad või nõrgemad ning sellega määratakse kindlaks, kuidas informatsiooni edastatakse. Ühenduste vahelised kaalud määravad, kui suur osa infost liigub järgmisesse kihti, ehk sisuliselt, kui suur mõju on igal sisendil. Igal neuronil on oma seisund, mis määratakse teiste neuronitega loodud sidemete kaudu [19].

Neuroni väljundi genereerimiseks rakendatakse neuroni seisundile mittelineaarset aktiveerimisfunktsiooni σ [22]:

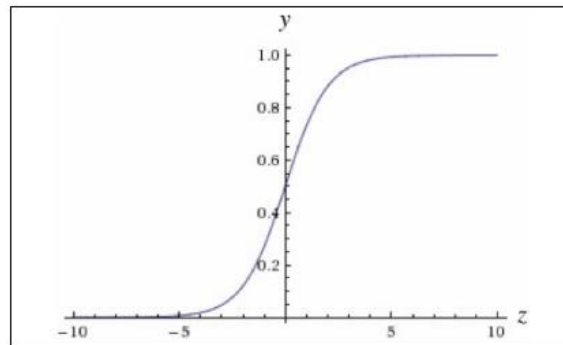
$$y = \sigma(z) = \sigma\left(\sum_{i=0}^n x_i w_i\right) \quad (6)$$

Mittelineaarse aktiveerimisfunktsiooni rakendamise tulemusena on iga neuroni väljundiks standardiseeritud väärtus. See väärtus on omakorda järgmises kihis asuva neuroni üks sisendväärtustest. Aktiveerimisfunktsioonidel on tähtis osa närvivõrkude struktuuris, kuna nende abil õpitakse looma mittelineaarseid seoseid. Enimkasutatavad aktiveerimisfunktsioonid on sigmoid, hüperboolne tangens (Tanh), tõkestatud mittenegatiivne lineaarne funktsioon (ingl *rectified linear unit*, e ReLU) [22].

Sigmoid aktiveerimisfunktsioon jagab väljundid monotoonselt kasvavalt 0 ja 1 vahele. See tähendab, et väga suured väärtused lähenevad ühele ja väga väikesed väärtused lähenevad nullile. Sigmoidfunktsiooni rakendamiseks kasutatakse järgmist valemit:

$$f(z) = \frac{1}{1+e^{-z}} \quad (7)$$

Üldiselt võib kirjelda funktsiooni tulemit S tähe kujuga (Joonis 3) [22].

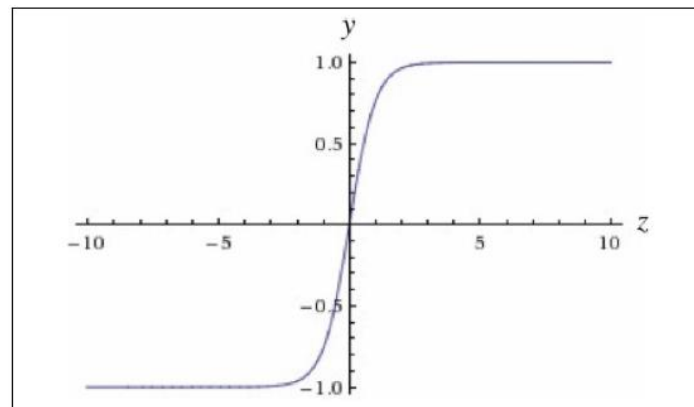


Joonis 3. Sigmoid funktsioon [22].

Tanh funktsioon kasutab väljundina sarnast S tähe kuju, kuid väljund jagatakse -1 ja 1 vahele. Sellisel juhul koonduvad keskväärtused nulli ümber. See on ka üks põhjuseid, miks teatud juhtudel seda eelistatakse Sigmoidi funktsioonile. Tanh funktsiooni arvutamiseks kasutatakse järgmist valemit [22]:

$$f(z) = \tanh(z) \quad (8)$$

Tanh funktsiooni kuju on toodud järgneval joonisel (Joonis 4).

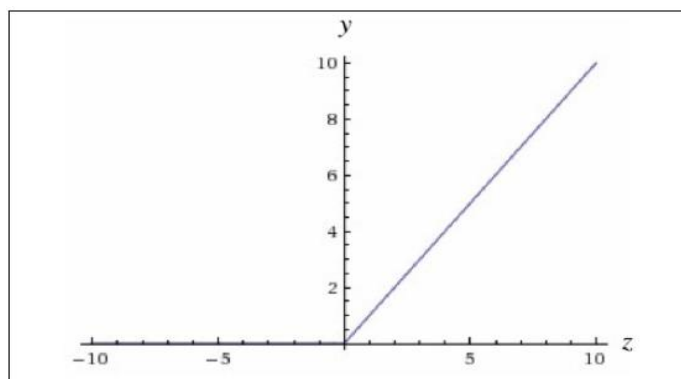


Joonis 4. Tanh funktsioon [22].

Neist kolmest kõige rohkem erineb ReLU funktsioon, mille arvutamiseks kasutatakse järgmist valemit [22]:

$$f(z) = \max(0, z) \quad (9)$$

ReLU funktsiooni kuju on toodud järgneval joonisel (Joonis 5).



Joonis 5. ReLU funktsioon [22].

Antud töös kasutatav närvivõrk rakendab arvutamisel Sigmoidi ning Tanh aktiveerimisfunktsioone.

Tulenevalt eelnevast avaldub närvivõrgu väljundi matemaatiline funktsioon kujul [19]

$$y = \sigma(z) = \sigma\left(\sum_{i=0}^n x_i w_i + b\right) \quad (10)$$

kus y – väljundkihi vektor, σ – aktiveerimisfunktsioon, x_i – sisendite vektor, w_i – kaaluvektor ja b – nihe.

2.1.2 Närvivõrgu treenimine

Närvivõrgu treenimine tähendab sisuliselt võrgu parameetrite iteratiivset muutmist närvivõrgu prognoosivea vähendamiseks.

Esialgne närvivõrk on justkui oletuslik funktsioon, kus kaalud on seatud juhuslikult ning mille tulemus ei pruugi võrduda soovitud tulemusega. See tähendab, et närvivõrgu prognoosi ja tegeliku tulemuse vahel on viga. Tüüpiliselt mõõdetakse mudeli ennustamise täpsust keskmise ruutveaga (ingl *mean squared error*, e MSE), mis arvutatakse järgmiselt [19]:

$$J = MSE = \frac{1}{n} \sum_{i=0}^n (\hat{y}^i - y^i)^2 \quad (11)$$

kus J – veafunktsioon, \hat{y} – ennustatav väärtus, y – soovitud väärtus.

Närvivõrgu treenimise eesmärk on vähendada prognoosi ruutviga ning seeläbi saavutada tegelikule tulemusele võimalikult lähedane tulemus. Kuna ruutvea funktsioon J koosneb kaaludest, sisenditest ja soovitud väljundist, siis treenimine toimub läbi kaalude muutmise. Eesmärk on kaalusid muuta selliselt, et prognoosi ruutviga väheneks võimaliku miinimumini ehk nullini [19].

Selleks, et saada teada, kuidas ning kummale poole peab kaalusid muutma, arvutatakse vea funktsioonist gradient kaalude järgi [19]:

$$\Delta J(w) = \frac{\partial J(w)}{\partial w_j} = \frac{\frac{\partial}{\partial w_j} \sum_{i=0}^n (\hat{y}^i - y^i)^2}{\partial w_j} \quad (12)$$

Vea minimiseerimiseks kasutatakse optimeerimise algoritmi, mida nimetatakse gradientlaskumiseks (ingl *gradient descent*). Selle algoritmiga tehakse kindlaks, millises suunas ning kui palju on vaja kaaluvektoreid muuta, et vähendada mudeli ruutviga [19].

Eelpool kirjeldatud viisil on võimalik leida kaalud, mis minimeerivad närvivõrgu ruutviga. Mitmekihilise närvivõrgu korral rakendatakse lisameetodeid, kuna närvivõrk koosneb paljudest kihtidest ning nende vahel olevatest kaaludest, mis kõik mõjutavad lõplikku tulemust. Selleks, et optimeerida tervet närvivõrku, rakendatakse gradiendi arvutamiseks ahelreeglit (ingl *chain rule*) ning lisaks tagasilevi (ingl *backpropagation*) meetodit. Sellisel juhul arvutatakse gradient mööda närvivõrku tagurpidi ehk liigutakse mööda närvivõrku paremalt vasakule ning muudetakse igas kihis neuronitele rakendatud kaale kuni sisendkihini välja [19].

Optimeerimise tulemusena kohandatakse väiksema veafunktsiooni saamiseks kaalusid järknevalt [23]:

$$w_{t+1} = w_t - \eta_t \Delta J(w_t) \quad (13)$$

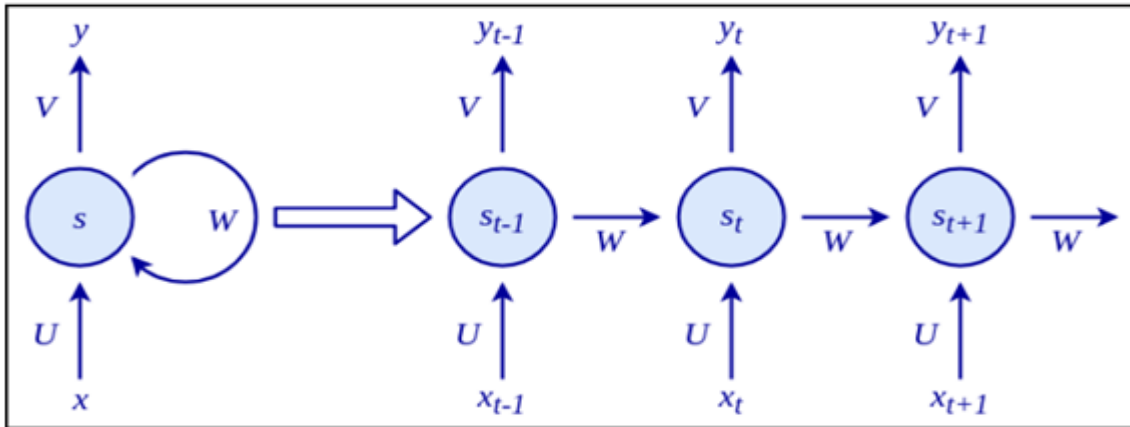
kus, w_i – kaalud, η – õppimise kiirus, $\Delta J(w)$ – veafunktsiooni gradient.

Õppimise kiirus on sammu pikkus, millega proovitakse iteratiivselt leida funktsiooni miinimumi.

2.1.3 Rekurrentne närvivõrk

Närvivõrke on väga palju erinevaid tüüpe ning kõik need püüavad lahendada erinevat ülesannet. Mõned on arendatud pildi tuvastamiseks, mõned kõne tuvastamiseks jne. Rekurrentsete närvivõrkude (RNN) eesmärk on töödelda jadas olevat järjestatud infot nagu näiteks lauses olevad sõnad või aktsia hind ajas [19].

Rekurrentse närvivõrkude erinevus tavaliselt ühesuunalisest närvivõrgust seisneb selles, et kasutatakse sisendite ajalist järjestust. See tähendab, et iga uus seisund on sõltuv eelnevast seisundist ehk jada liikmete järjestus kannab endas lisainformatsiooni [24].



Joonis 6. RNN närvivõrk lahti laotatud kujul [19].

Joonis 6 esitab rekurrentse närvivõrgu ühe neuroni laiili laotatud kuju, kus x on sisend, U on sisendi kaal, s on seisund ning V on väljundkaal ja y on väljund. Erinevalt ühesuunalisele närvivõrgule on lisandunud kolmas seisundi kaalumõõde W . See tähendab, et iga eelnev seisund võetakse arvesse järgneva seisundi arvutamisel [19].

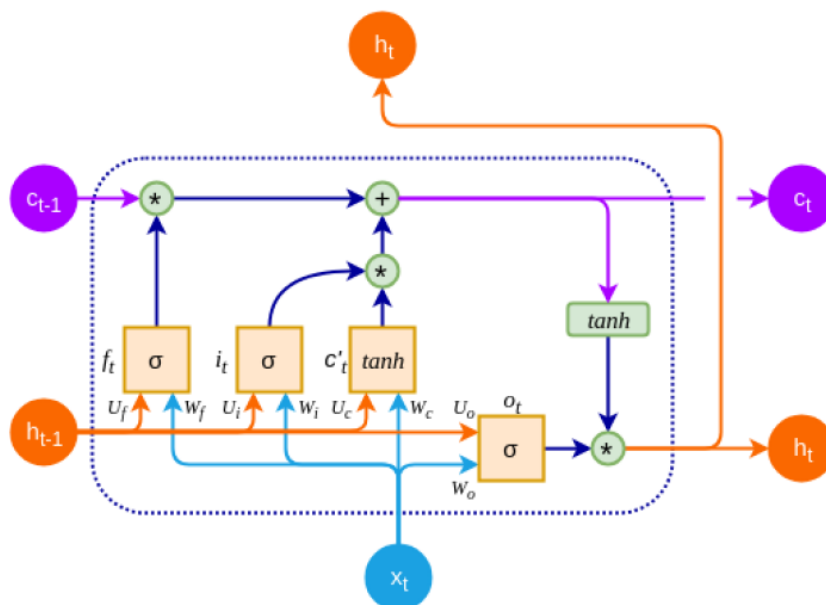
Kuigi rekurrentne närvivõrk on hea meetod ajalooliste andmete analüüsimiseks, esineb sellel ka üks nõrkus. Pidevalt ajalugu kaaludes tekib efekt, kus ajaloos kaugemal olevate kaalude muutmisel on väljundile oluliselt väiksem (või suurem) mõju kui ajaloos lähemal olevate kaalude muutmisel. Seda nimetatakse kahaneva (või plahvatava) gradiendi probleemiks (ingl *vanishing gradient problem*). Selle probleemi ühe võimaliku lahendina on välja töötatud täiendatud versioon rekurrentsest närvivõrgust – pikk lühiajalise mälu (ingl *long short-term memory* e LSTM) närvivõrk [19].

2.1.4 LSTM

LSTM ehk pikk lühiajalise mälu närvivõrk on rekurrentse närvivõrgu edasiarendus, mille avaldasid Hochreiter ja Schmidhuber 1997. aastal [25] ning Schmidhuber, Gers ja Cummins täiendasid seda 1999. aastal [26]. Nagu eelmises peatükis mainitud, on rekurrentsel närvivõrgul üks nõrkus. Nimelt mida rohkem ajas tagasi liikuda, seda vähem olulisemaks muutuvad ajas kaugemal olevad väärtused ning see muudab võrgu treenimise raskeks ning prognoosi ebatäpseks.

LSTM võrk on sisuliselt rekurrentne närvivõrk pikaajalise mälukomponendiga. Selleks, et oluline info jääks ajaloost kättesaadavaks, täiendati rekurrentse närvivõrgu rakku mitme funktsiooniga. LSTM rakud (ingl *LSTM cells*) on sisemise rekursiooniga (ingl *self-loop*). Igal LSTM rakul on samad sisendid ja väljundid nagu tavalisel rekurrentsel närvivõrgul, kuid sellele lisandub rohkem parameetreid ning nn väravate süsteem (ingl

system of gate units), mis kontrollib infovoogu [27]. See võimaldab närvivõrgul õppida, millal eelmiste peidetud kihtide olek unustada ning millal peitekihi olek uue infoga uuendada. Peidetud kiht on sisuliselt raku väljund. LSTM-i rakk koosneb kolmest väravast – sisendvärav (ingl *input gate*), unustamise värav (ingl *forget gate*) ja väljundvärav (ingl *output gate*). Need väravad aitavad otsustada, millised väärtused on vaja meeles pidada ja millised võib unustada. Lisaks väravatele on LSTM rakul ka olek ehk seisund. Seisund on sisuliselt võrgu mälu, millele väravad lisavad või millelt väravad võtavad andmed [28].



Joonis 7. LSTM raku diagramm [19].

Joonis 7 esitab LSTM raku diagrammi, millel on esitatud info liikumine raku. Joonisel esitatud väärtused tähendavad järgmist [19]:

- x_t on LSTM sisend ajahetkel t , c_t on LSTM raku seisund aja hetkel t , h_t on LSTM väljund (peidetud väärtus) ajahetkel t ;
- W ja U on vastavalt sisend- ja väljundvektorite kaalud;
- f_t, i_t, o_t on vastavalt unustamise värav, sisendvärav ja väljundvärav;
- σ ja \tanh on aktiveerimisfunktsioonid;
- $*$ on elemendiviisiline korrutis ja $+$ elemendiviisiline liitmine.

Selleks, et info jõuaks sisendist väljundini, tehakse läbi mitmed sammud. Alustatakse unustamise väravast. Selles otsustatakse sisuliselt, millist osa eelmise LSTM raku seisundist säilitatakse. Otsus tehakse sisendi x_t ja eelmise LSTM raku väljundi h_{t-1} põhjal, kasutades aktiveerimisfunktsioonina sigmoidi [25], [26]:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (14)$$

Selle funktsiooni tulemus on surutud 0 ja 1 vahele. See tähendab, et unustusvärava 0 tulemuse juures unustatakse kogu raku seisundi andmed ja 1 juures lubatakse kõik andmed läbi. Selle väravaga hinnatakse, millised väärtused on olulised ja millised mitte.

Sisendväravas otsustatakse, millist osa LSTM olekust tuleks uuendada. Seda tehakse kahes osas. Esimeses osas otsustatakse, milliseid väärtusi uuendatakse, kasutades sigmoidi [25], [26]:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (15)$$

Järgmise sammuna arvutatakse nn kandidaatväärtused, mida uuendatud olekusse võiks lisada. Otsus tehakse samuti sisendi x_t ja peidetud kihi väljundi h_{t-1} põhjal, kuid sellele rakendatakse \tanh funktsiooni [25], [26]:

$$c'_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (16)$$

Kolmanda sammuna uuendatakse raku seisundit [25], [26]:

$$c_t = f_t * c_{t-1} \oplus i_t * c'_t \quad (17)$$

Väljundväravas otsustatakse, milline saab olema raku väljund, millest omakorda saab järgmise raku sisend peidetud väärtuse kujul [25], [26]:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (18)$$

$$h_t = o_t * \tanh(c_t) \quad (19)$$

Kuna kõikidest nendest valemiteist on võimalik võtta tuletis, siis saab rakendada optimeerimisel tagasilevi meetodit ning võrku on võimalik treenida täpselt nagu rekurrentset närvivõrku. Eelis rekurrentse närvivõrgu ees seisneb selles, et olenemata võrgu sügavusest, ei unusta LSTM kauges ajas olevaid väärtuseid. Need väärtused säilitatakse või vajadusel kustutatakse unustamise värava kaudu [19].

Käesolevas töös kasutatakse prognoosi tegemiseks LSTM mudelit, kuna finantsaegridade puhul võivad ajaloolised andmed prognoosi tegemisel osutada olulisteks.

3 Kirjanduse ülevaade

Alljärgnevas peatüks antakse ülevaade akadeemilistest uuringutest ning saavutatud tulemustest masinõppe rakendustest finantsturgudel. Varaseimad uuringud finantsturgude modelleerimisest tehisnärvivõrkude abil viidi läbi juba 1980ndate lõpus ja 1990ndate alguses. Tehisnärvivõrke rakendati peamiselt aktsiahindade, valuutakursside ja toormehindade prognoosimisel. Seevastu võlakirjaturge on tehisnärvivõrkude abil modelleeritud oluliselt vähem. Seniste uurimuste tulemused näitavad, et tehisnärvivõrkude finantsvarade hindade ennustusvõime on väga kõrge ning ületab seejuures märgatavalt ja pidevalt lihtsamate lineaarsete mudelite oma.

3.1 Erinevate masinõppe tehnikate võrdlus

Salem jt [29] uurisid 2017.a masinõppe võimekust investeerimisotsuste tegemiseks USA võlakirjaturul. Nad kasutasid erinevaid juhendatud masinõppe tehnikaid, et finantsturgude toorandmetest luua efektiivsed kauplemismudelid. Õppimine limiteeriti klassifitseerimisülesandena, seega ennustatav väärtus oli binaarne (osta või müü). Masinõppe klassikalistest tehnikatest kasutati k -lähima naabri (ingl *k-nearest neighbours*), otsustuspuu (ingl *decision trees*) ja tugivektorklassifitseerija algoritme, ansambelõppe tehnikatest otsustusmetsa ja sügavõppe tehnikatest tehisnärvivõrke. Analüüsis kasutatud tehisnärvivõrgu täpset arhitektuuri ei täpsustatud. Tulemused näitasid, et kõikide kasutatud tehnikate puhul saavutati positiivne kauplemisstrateegia mediaan infohuhe testperioodi pikkuse varieerimisel. Seejuures annavad osad tehnikad (otsustusmets, tugivektorklassifitseerija ja tehisnärvivõrgud) 95% usaldusnivool statistiliselt olulise tulemuse positsioonide ajastamise ja suuruse valikute tegemisel võlakirjaturul.

Ganguli ja Dunmon [30] viisid 2017.a läbi analoogse uurimuse võlakirjaturgudel, kasutades masinõppe erinevaid tehnikaid ja algoritme võlakirjahinna määramiseks. Võrreldi üldistatud lineaarset mudelit (ingl *generalized linear models*), regressioonipuud (ingl *regression tree*), peakomponentide regressiooni (ingl *principal component regression*), tugivektorklassifitseerijat, otsustusmetsa, vähimruutude võimendamist (ingl

LS-Boost) ja tehisnärvivõrke. Uurimuse lõppjärelendusena tõdetakse, et tehisnärvivõrgud annavad täpseima võlakirjade hinna prognoosi. Närvivõrgus kasutati üht peitkihti ning õpialgoritmina Levenberg-Marquadt tagasilevi optimeerimist. Lisajärelendusena mainitakse, et närvivõrgu keerukuse tõstmine (neuronite arvu suurendamine) parandab tulemust.

3.2 Masinõpe ja traditsioonilised statistilised meetodid

Nunes jt [31] tõdesid 2018.a tehtud uurimuses, et võlakirjaturge on tehisnärvivõrkude abil modelleeritud oluliselt vähem kui aktsia- ja valuutaturgusid. Oma uurimuses kasutasid nad mitmekihilisi erinevate sisenditega närvivõrke euroala tulukõvera prognoosimiseks ja võrdlesid saadud tulemusi lineaarsete meetoditega saadud prognoosidega. Analüüsis selgus, et mitmekihilised närvivõrgud saavutasid lineaarse regressiooniga võrreldes kõrgema täpsuse. Olulise nüansina selgus, et erinevate tulukõvera punktide ja erineva prognoosiperioodi puhul omab sisendite valik olulist rolli ja see määrab ära prognoosi täpsuse.

Dunis ja Morrisson [32] prognoosisid oma 2007.a avaldatud töös tehisnärvivõrkude abil 10-aastaste USA, UK ja Saksamaa valitsusvõlakirjade intressimäärasid nendel turgudel kauplemissignaali genereerimiseks. Lisaks võrdlesid nad keerulisema närvivõrkude meetodil saadud kauplemissignaali edukust traditsiooniliste ökonomeetriste mudelitega nagu autoregressiivse libiseva keskmise (ARMA) mudel ja logistiline regressioon (LOGIT). Neid võrreldi omakorda tehnilise analüüsi libiseva keskmise konvergens divergens (MACD) strateegiaga. Sisenditena kasutati erinevaid finantsturgude andmeid (võlakirjaintressid, intressivahed, pankadevahelised intressimäärad, aktsia- ja toormehinnad, vahetuskursid). Kasutati ühe peitkihi ja viie neuroniga tehisnärvivõrku. Peitkihi aktiveerimisfunktsioonina kasutati hüperboolset tangensit. Analüüsi tulemus näitas närvivõrkude paremat võimekust võlakirjaintresside prognoosimisel, võrreldes traditsiooniliste prognoosimudelitega ning samuti saavutati parem tulemus kui tehnilise analüüsi strateegiaga. Tehisnärvivõrgul põhineva kauplemisstrateegiaga saavutati kõrgeim kumulatiivne tulemus ja riskiga kohandatud tootlus.

Bianchi jt [33] võrdlesid 2018.a erinevaid masinõppe meetodeid traditsioonilise peakomponentide regressiooniga võlakirjade riskipremia prognoosimisel. Nende

tulemused näitavad, et mittelineaarsete masinõppemeetodite, täpsemalt närvivõrkude täpsus on võlakirjatootluste valimiväliste prognooside ennustamisel kõrgem, võrreldes lineaarse peakomponentide regressiooniga. Lisaks näitab nende uurimus, et makronäitajate lisamine tõstab oluliselt valimiväliste prognooside täpsust. Tulemustele omas olulist mõju närvivõrgu kihtide arv. Võrreldes 2- ja 4-kihilisi närvivõrke, leiti, et kahe kihi lisamine suurendas täpsust 10% võrra. Parimad tulemused saavutati sügavamate 4-kihiliste närvivõrkudega, kus sisenditena kasutati nii makronäitajaid kui ka forvardintressimäärasid. Analüüsis kasutati masinõppe meetoditest regressioonipuud, otsustusmetsa ja närvivõrkudest traditsioonilist pärilevivõrku. Närvivõrgu aktiveerimisfunktsioonina kasutati mittenegatiivset lineaarfunktsiooni ReLU.

3.3 Erinevate närvivõrkude rakendused võlakirjaturul

Cheng jt [34] kasutasid 1996.a tehtud uurimuses tehisnärvivõrke 30-aastase USA valitsusvõlakirja nädalase liikumissuuna prognoosimiseks. Uuriti pärilevi närvivõrke, kaht peitkihti kahe erineva aktiveerimisfunktsiooniga (siinus ja sigmoid). Sisenditena kasutati tulukõvera andmeid, aktsia- ja toormehindu ning vahetuskurse. Tulemused kinnitasid tehisnärvivõrkude väga head ennustusvõimet. Seejuures oli hinnaliikumise suuna tabavus oluliselt kõrgem võrreldes ulatusega. Leiti, et närvivõrgu prognoositäpsuse määravad nii võrgu struktuur, kui ka treeningperioodi pikkus, õpetamisperioodide arv ning sisendandmete eeltöötlus. Huvitava aspektina näitas analüüs, et närvivõrgu treeningu lühendamine andmete kogumahult viimase 2-4 aastale parandas prognoosi täpsust.

Din [35] kasutas 2003.a avaldatud uurimuses nii lühi- kui pikemaajaliste intressimäärade prognoosimiseks ühe peitkihi ja 5-10 neuroniga pärilevi närvivõrku. Intressimäärade kuise dünaamika prognoosimiseks kasutati 10-aastast perioodi. Sisenditena intressimäärade enda ajalugu ja lisaks sellele makro- (inflatsioon, SKP) ja turunäitajaid (pankadevahelised intressimäärad, valuutakursid). Uuringu tulemus näitab, et närvivõrkude kasutamisel saadud prognoosi tabavusmäär on väga kõrge (83% 12-kuulise prognoosiperioodi korral).

3.4 Rekurrentsete närvivõrkude rakendused finantsandmetel

Kui pärilevi närvivõrkude rakendusi võlakirjaturul on akadeemilistes uurimustes mõnevõrra käsitletud, siis rekurrentseid närvivõrke on võlakirjaturul rakendatud seni väga vähe. Küll aga on rekurrentseid närvivõrke järjest enam hakatud rakendama aktsiaturul ja seal on senised tulemused paljulubavad.

Naeini jt [36] võrdlesid 2010.a kaht erinevat närvivõrkude arhitektuuri – mitmekihilist pärilevivõrku ja Elmani rekurrentset närvivõrku, rakendades neid aktsiahindade prognoosimiseks. Uuringu tulemusena järeldati, et rekurrentne närvivõrk ennustab aktsiahindade liikumise suunda täpsemalt võrreldes mitmekihilise närvivõrguga. Samal ajal on rekurrentse närvivõrgu aktsia hinna prognoosi keskmine ruutviga suurem kui mitmekihilisel närvivõrgul.

Nakagawa jt [37] rakendasid 2019.a masinõppe meetodeid, sh rekurrentset närvivõrku Jaapani aktsiaturul. Kasutati pika lühiajalise mälu (LSTM) närvivõrku ja lisaks selle edasiarendust, võttes võrgu optimeerimisel arvesse ka sisendite olulisust igas kihis. Analüüsi tulemus näitas, et LSTM ületab prognoosi täpsuse poolest nii mitmekihilise pärilevi närvivõrgu, lineaarse regressioonimudeli, tugivektorklassifitseerija kui ka otsustusmetsa.

Nelson jt [38] kasutasid 2017.a aktsiahindade prognoosimiseks LSTM mudelit ja võrdlesid seda muude masinõppe meetoditega (mitmekihiline pärilevi närvivõrk ja otsusmets). LSTM mudeli sisenditena kasutati väga suurt hulka tehnilise analüüsi näitajaid (libisevad keskmised, momentum, MACD, RSI jms), mida kasutatakse laialdaselt nii subjektiivsete investeerimisotsuste sisendina kui ka automaatsete investeerimisstrateegiate väljatöötamisel. Koos aktsiahinna andmetega (avamis-, sulgemis-, kõrgeim, madalaim hind ja kauplemismaht) oli sisendeid kokku 180. Kuigi vaadeldud aktsiate arv antud analüüsis oli tagasihoidlik (vaid viis), selgus, et mälu komponenti sisaldav närvivõrk on aktsiahindade prognoosimisel statistiliselt oluliselt edukam traditsioonilisest pärilevi närvivõrgust enamuse aktsiate korral.

4 Empiiriline analüüs ja tulemused

Varasemate akadeemiliste uurimustööde tulemuste põhjal [32] otsustati USA võlakirjafutuuride hinnadünaamika prognoosimiseks kasutada masinõppe meetoditest tehisnärvivõrke.

Mudelid koostatakse erinevate tulukõvera punktide, täpsemalt 2- ja 10-aastase valitsusvõlakirjade futuuride jaoks. Mudeli sisenditena kasutatakse nii futuuride enda hinnaridasid kui ka finantsturgude andmeid ja makronäitajaid, mis varasemates uuringutes on prognoosi täpsust tõstnud [33], [35].

Närvivõrgu arhitektuuri valikut ja vastavat konfiguratsiooni (sisendite arv, peidetud kihtide arv, neuronite arv, neuronite aktiveerimisfunktsioon igal kihil, kaalukoefitsientide ja nihete algväärtused, optimeerimismeetod, õpisamm, õpetamisperioodide arv jms) akadeemilises kirjanduses võlakirjaturu jaoks palju käsitletud pole. Lisaks keskendub varasem kirjandus võlakirjaturu puhul pigem pärilevi närvivõrkudele ning rekurrentseid võrke võlakirjaturul seni palju rakendatud pole. Kuna rekurrentsed närvivõrgud on finantsaegridade analüüsiks sobivamad ning aktsiaturul häid tulemusi näidanud [36]-[38], otsustati käesolevas töös samuti rekurrentse, täpsemalt pika lühiajalise mälu (LSTM) närvivõrgu kasuks. LSTM närvivõrgu võime modelleerida järjestusi ja ajas muutuvaid sõltuvusi on finantsandmete puhul oluline omadus. Närvivõrgu parameetrite valik toimus võimalusel finantsturgudel (nt aktsiaturul) läbiviidud empiiriliste uurimustööde alusel ja muul juhul eksperimentaalselt.

4.1 Mudeli kirjeldus

Uuringu läbiviimiseks oli vajalik erinevate rakenduste installimine. Uuringu teostamiseks kasutati üht enim tuntud masinõppe rakendust Keras. Keras on kõrgetasemeline rakendusliides, mis sisaldab endas erinevaid närvivõrkude algoritme. Keras kasutab taustsüsteemina Tensorflow-d, CNTK-d või Theano-t [39]. Käesolevas töös kasutati taustsüsteemina Tensorflow-d. Tensorflow on avatud lähtekoodiga masinõppe platvorm. See on kõikehõlmav ja paindlik tööriistade süsteem, mis võimaldab kasutada erinevaid

masinõppe meetodeid [40]. Andmete analüüs ning mudelite rakendamine on teostatud programmeerimiskeeles Python.

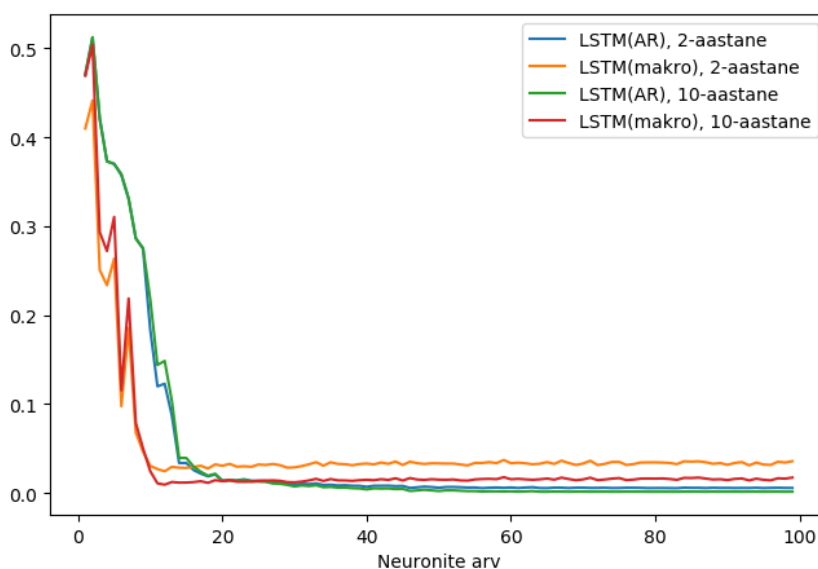
Mudeli eesmärk on ennustada 2- ja 10-aastase valitsusvõlakirja futuuri hinna järgmise päeva liikumissuunda. Vastavalt ennustatud suunale genereeritakse ostu- või müügi-signaal ning võetakse positsioon järgmiseks päevaks. Positsiooni suund arvutatakse uuesti iga päev.

Töö käigus koostati kaks erinevat mudelit:

- 1) autoregressiivne LSTM(AR) mudel, kus sisendina kasutati futuuri enda hinnareala ajalugu erinevate viitaegadega;
- 2) makro- ja finantsandmetel põhinev LSTM(makro) mudel, kus sisendina kasutati lisaks futuuri enda viitajale suuremat hulka finantsturgude ja makromajanduslikke näitajaid.

Mõlemal juhul kasutati analüüsiks kolme peidetud kihiga LSTM närvivõrku. Varasemad uuringud [33] on näidanud, et mitmekihiliste närvivõrkude prognoositäpsus on kõrgem, võrreldes ühekihiliste närvivõrkudega.

LSTM mudeli neuronite arvaks valiti 50 kõigis kihis. Erandiks on väljundkiht, kus jääb alles ainult üks väärtus. See tuleneb asjaolust, et mudeliga ennustatakse ühte ainsat muutujat.



Joonis 8. Minimaalne treeningu ruutviga erinevate neuronite arvuga mudelis.

Neuronite arv valiti autori katsetuste tulemusel. Sobiva neuronite arvu leidmiseks kasvati seda 1st 100ni (Joonis 8). Võrreldavuse tekitamiseks võeti kõigi mudelite viitajaks 1 päev. Kuni 50 neuronini olid muutused suuremad ning parandasid mudeli tulemust. Edasine neuronite arvu suurendamine enam tulemusele statistiliselt olulist mõju ei avaldanud ning mudeli täpsus ei kasvanud.

Juhuslik kadu (*dropout*) on mudelis seatud 20%-le. Juhuslik kadu on reguleerimistehnika, mis ette antud tõenäosusega eemaldab juhuslikult närvivõrgust neuronid ning temaga seotud kaalud. Juhuslik kadu muudab mudeli robustsemaks ning aitab vältida olukordi, kus õpitakse valesid seoseid. 20% on kirjanduses levinud juhusliku kao suurus [19].

Mudeli headuse hindamiseks jagati analüüsis andmed kaheks. Mudel treeniti 70% andmete pealt ja saadud parameetreid kasutati mudeli hindamiseks ülejäänud 30% andmete peal. Treeningute arvuks valiti 100, kuna teostatud arvutuste käigus leiti, et see on piisav leidmaks võimalikud optimaalsed parameetrite väärtused. See tähendab, et kasutades rohkem kui 100 treeningkordust, paranes mudeli ruutviga vähem kui 0,0005 võrra.

Optimeerimismeetod ning õpisamm jäid muutmata kujule. Optimeerimismeetod Adam ning õpisamm 0,001 on Keras poolt algseadistatud LSTM mudeli parameetrid. Treeningplokki otsustati mitte osadeks jagada, kuna analüüsitavate andmete maht võimaldas kasutada igas õpisammus kogu sisendväärtuste andmemahtu.

Viitaeg oli üks parameeter, mida töö käigus uuriti. Viitaeg on andmeploki suurus, mille pealt ennustatakse sõltuv väärtus. See tähendab, et uue väärtuse ennustamiseks vaadatakse ajaloos tagasi vastav päevade arv. Käesolevas analüüsis muudeti viitaega 1 päevast kuni 20 päevani.

Mudeli optimeerimise vea funktsioonina kasutati Keras poolt algseadistatud keskmist ruutviga.

Kokkuvõtvalt on mudelite parameetrid toodud järgnevas tabelis (Tabel 1). Töö käigus varieeriti kaht parameetrit – sisendite arvu ja viitaega.

Tabel 1. Mudelite parameetrid.

Parameetrid	LSTM(AR)	LSTM(makro)
Sisendite arv	1	22
Peidetud kihtide arv	3	3
Neuronite arv	50	50
Juhuslik kadu	20%	20%
Treeningperiood	70%	70%
Testperiood	30%	30%
Treeningkorduste arv	100	100
Optimeerimismeetod	Adam	Adam
Õpisamm	0.001	0.001
Treeningplokk	maksimaalne	maksimaalne
Viitaeg	1-20	1-20
Veafunktsioon	MSE	MSE

Mudeli tulemust hinnati kahe parameetri järgi. Esiteks hinnati mudeli suutlikkust suuna ennustamisel. Selleks võrreldi mudeli kumulatiivset tulemust passiivse osta ja hoia strateegia vastu. Teise parameetrina hinnati mudeli tulu-riski suhet (Sharpe suhtarvu).

```

from keras import Sequential
from keras.layers import Dense, LSTM, Dropout

model = Sequential()
model.add(LSTM(units=50, return_sequences=True, input_shape=(X.shape[1], X.shape[2])))
model.add(Dropout(0.2))
model.add(LSTM(units=50, return_sequences=True))
model.add(Dropout(0.2))
model.add(LSTM(units=50))
model.add(Dropout(0.2))

model.add(Dense(units=1))
model.summary()
model.compile(optimizer='adam', loss='mean_squared_error')

model.fit(X_train, Y_train, epochs=100, batch_size=len(X_train),
validation_data=(X_test, Y_test), verbose=2, shuffle=False)

predicted_value = model.predict(X_test)

```

Joonis 9. LSTM mudel programmikoodis.

Mudel on rakendatud programmeerimiskeeles Python (Joonis 9). Mudeli sisendandmed on standardiseeritud, ehk teisendatud kujule, kus kõik sisendväärtused on seatud 0 ja 1 vahele. Viitaegade muutmisel kohandatakse vastavalt treening ja test andmestikku.

4.2 Andmed

LSTM(makro) mudelis sisenditena kasutatud andmed pärinevad USA Föderaalreservi majandusuuringute andmebaasist (*Federal Reserve Economic Data*, e FRED) [41]. Tegemist on makromajanduslike teadustööde tarbeks koostatud väga põhjaliku USA makro- ja finantsnäitajate andmebaasiga. Andmeid uuendatakse reaalajas ning andmebaas on veebis vabalt kättesaadav. Võimalik on pärida nii kuise kui kvartaalse sagedusega majandusnäitajaid. Käesoleva töö sisenditeks valiti antud andmebaasist kuise sagedusega järgmised võlakirjaturule olulised makronäitajad:

- reaalne sissetulek,
- reaaltarbimine,
- jaemüük,
- tööstustoodang,
- uute töötüabiraha saajate arv,
- hõive kasv,
- keskmine töötundide arv,
- uued eluasemed,
- laovarud,
- laovarude käibekordaja,
- tarbijahinnaindeks (THI),
- eratarbimise deflaator,
- keskmine nädalapalk.

Finantsturu andmetest kasutati mudeli sisenditena järgmisi näitajaid:

- USA Föderaalreservi baasintress,
- 3-kuulise rahatähe intress,
- 10-aastane krediidipreemia,
- S&P 500 aktsiaindeks,
- dollari vahetuskursi indeks,
- USDJPY vahetuskurss,
- nafta hind,
- kulla hind.

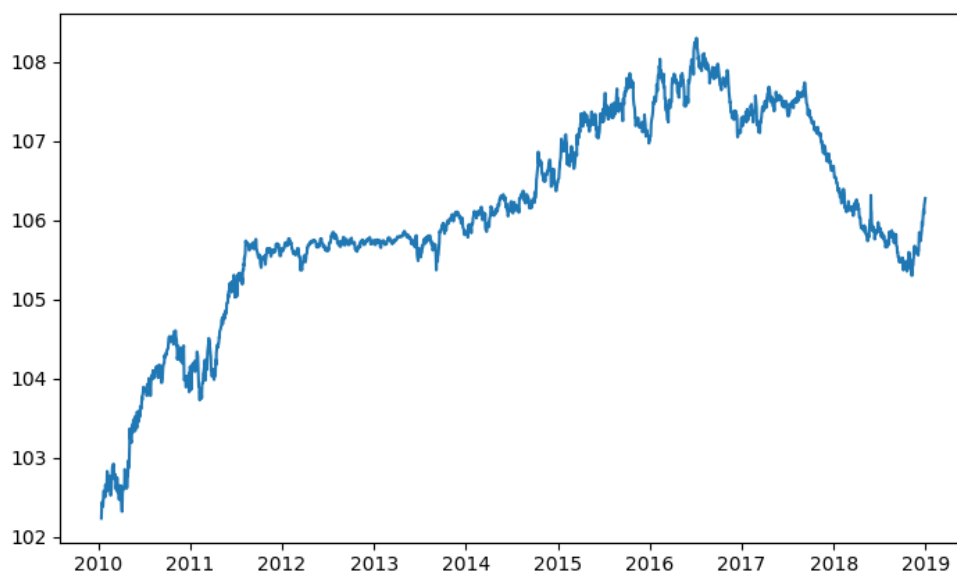
Sisendandmete valimisel lähtuti sellest, et makromajanduslikud näitajad kataksid võimalikud laia ulatust. Finantsandmete valikul võeti arvesse turul enim jälgitavad näitajad. Sisendandmete täpne selekteerimine ning parimate välja selgitamine ei olnud käesoleva töö eesmärk.

4.3 Tulemused

Järgnevas peatükis antakse ülevaade analüüsi tulemustest. Peatükk on jagatud kaheks suuremaks alapeatükiks, kus igas käsitletakse erinevat alusinstrumenti.

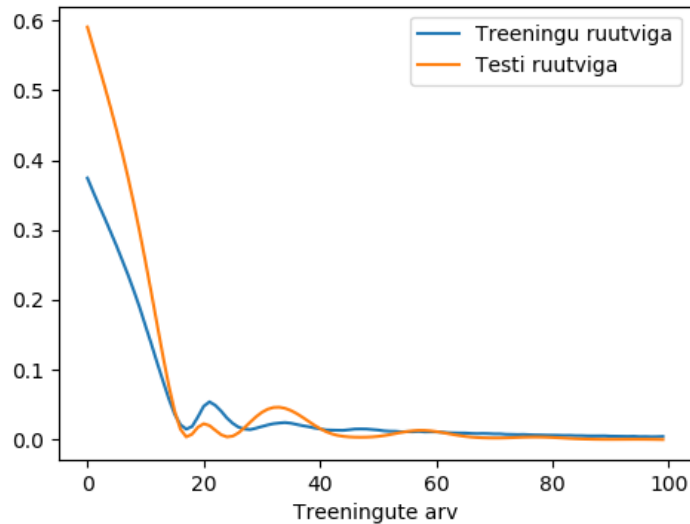
4.3.1 2-aastane USA võlakirja futuur

USA 2-aastase võlakirjafutuuri hind on vaadeldaval perioodil kasvanud (Joonis 10). Esimestel aastatel kasvas hind kiiremini ning vahepealsetel aastatel on hind püsinud stabiilsena. Viimastel aastatel on hinnavolatiilsus taas kasvanud. Kui investor oleks kogu perioodil kasutanud osta ja hoia strateegiat, oleks tema portfell kasvanud 3,95%.



Joonis 10. USA 2-aastase võlakirjafutuuri hind.

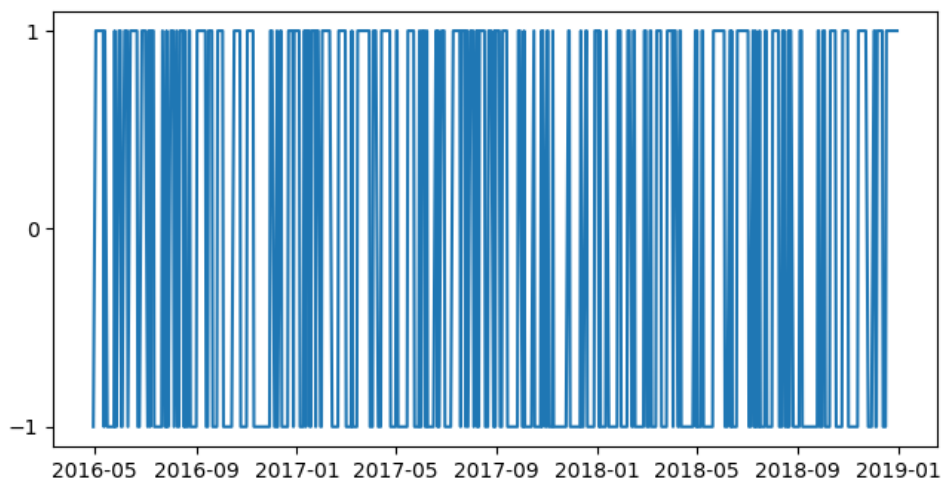
Käesoleva töö empiirilises osas töötati esmalt välja LSTM(AR) mudel, mille sisendiks oli ainult aegrea enda ajalugu. Mudelit treeniti vastavalt etteantud parameetritele (Tabel 1) treeningperioodil ning hinnati testperioodil, varieerides mudeli viitaegade arvu. Parim kumulatiivne tulemus ja tulu-riski suhe saavutati 4-päevase viitajaga.



Joonis 11. LSTM(AR) 2-aastane, viitaeg 4 päeva, mudeli ruutviga.

Mudeli treeninguruutviga kahanes esimeste treeningkorduste juures kiiresti (Joonis 11) ning saavutas oma madalaima taseme ligi 20 treeningkorduse juures. Stabiilselt madal ruutviga saavutatakse 40-60 korduse vahel. Testi ruutvea dünaamika sarnanes treening ruutveale. Sellest võib järeldada, et mudeli parameetrid on optimaalsed ning mudel on saavutanud maksimaalse võimekuse.

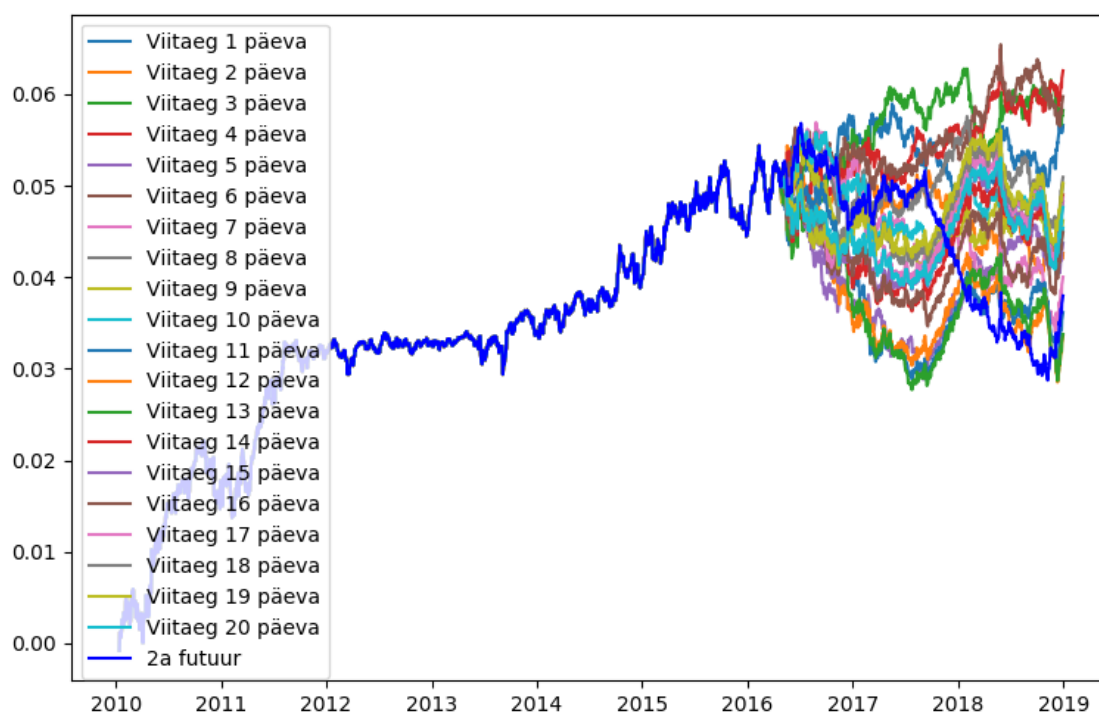
Saadud mudeli abil genereeriti vastavalt ostu- ja müügisignaale (Joonis 12). Ostusignaal on tähistatud numbriga 1 ning müügisignaal numbriga -1. Mudel on investeeritud kogu perioodi vältel ning neutraalset positsiooni ette ei nähtud. Signaalid genereeriti ainult testperioodi jaoks. Treeningperioodil positsioone ei võetud. Seetõttu on võrreldavaks perioodiks testperiood. Mudel muudab testperioodi jooksul signaale korduvalt ja tihti. Kohati muudetakse signaale igapäevaselt. Tuvastatavad on mõned üksikud pikemad samasuunalise signaaliga perioodid.



Joonis 12. LSTM(AR) 2-aastane, viitaeg 4 päeva, mudeli signaalid.

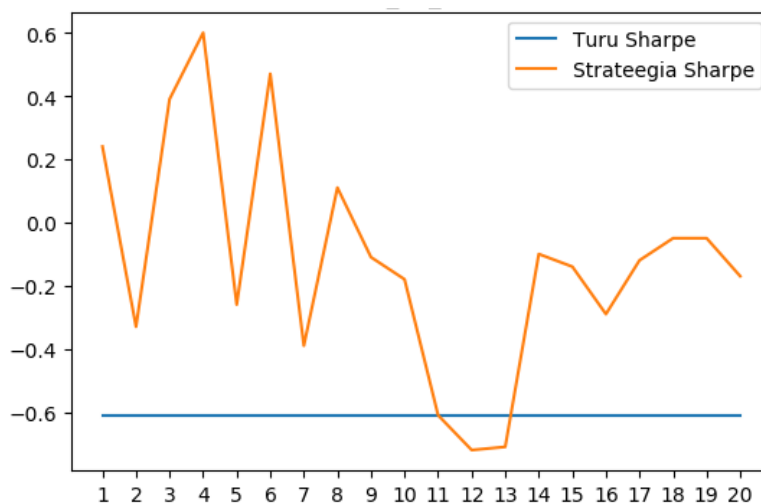
Vastavalt genereeritud signaalidele võeti positsioonid ning seeläbi saavutati treeningperioodil osta ja hoia strateegiaga võrreldes erinev tulemus.

Kasutades mudelis erinevaid viitaegu, olid tulemused testperioodi lõpuks valdavalt osta ja hoia strateegiast paremad (Joonis 13). Turu kumulatiivne tulemus testperioodil oli -1,33%. LSTM(AR) strateegiatest andis parima tulemuse viitaeg 4 päeva kumulatiivse tulemusega 1,32%. Samuti saavutati turust parem tulemus teiste viitaegadega. Üldistades võib öelda, et suurimat tulu suutsid vaadeldaval perioodil teenida lühema viitajaga mudelid. Suurem osa strateegiatest ületasid küll turgu, kuid lõpetasid perioodi negatiivse tootlusega.



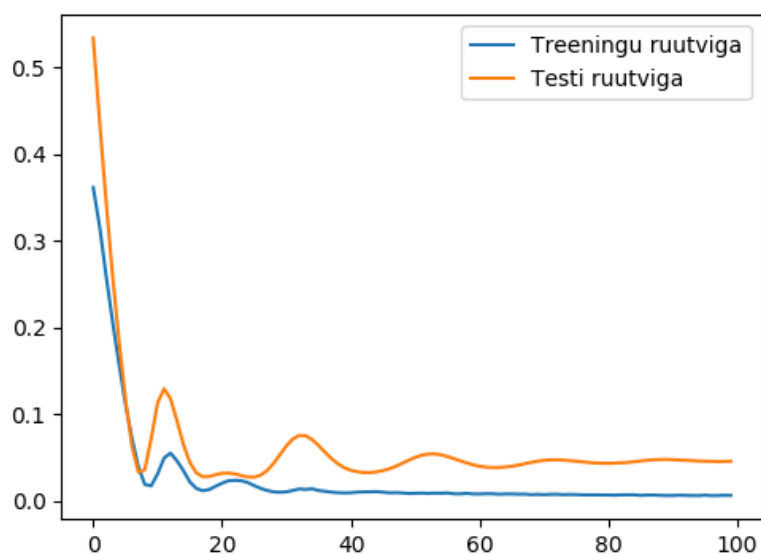
Joonis 13. LSTM(AR) 2-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.

Võrdlemaks erinevate mudelis kasutatud viitaegade tulu-riski suhet kasutati Sharpe suhtarvu (Joonis 14). Sharpe võrdlusest on näha, et parim tulemus saavutati pigem lühemate viitaegadega. Enamus valitud viitaegadest on parema tulemusega kui vaadeldava perioodi turu Sharpe suhe. Parim tulemus saavutatakse 4-päevase viitajaga ning see annab Sharpe suhteks 0,6. Turu Sharpe suhtest madalam tulemus saavutati viitaegadega 12 ja 13 päeva. Keskmine Sharpe suhe üle kõikide viitaegade oli -0,12 ning turu Sharpe vaadeldaval perioodil oli -0,61. See tähendab, et keskmiselt ületas LSTM(AR) strateegia turu Sharpe suhet 0,49 võrra.



Joonis 14. LSTM(AR) 2-aastane, viitaeg 1-20 päeva, Sharpe.

Teise sammuna viidi läbi uuring 2-aastase võlakirja futuuri peal, kus sisendandmetena kasutati lisaks iseenda aegreale ka makromajanduslike ning teisi finantsandmete aegridu. LSTM(makro) mudelit treeniti vastavalt etteantud parameetritele (Tabel 1), varieerides mudeli viitaegu. Parim tulemus saavutati 6-päevase viitajaga.

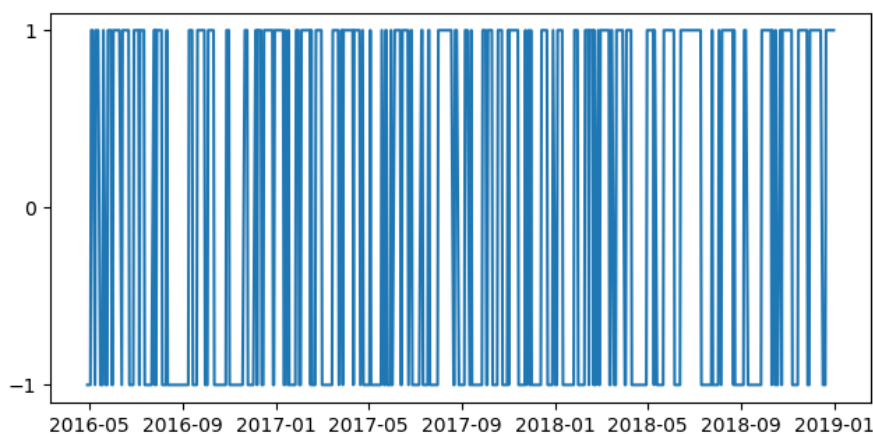


Joonis 15. LSTM(makro) 2-aastane, viitaeg 6 päeva, mudeli ruutviga.

Sarnaselt LSTM(AR) mudelile kahanes treeningu ruutviga esimeste treeningkorduste juures kiiresti (Joonis 15) ning saavutas oma madalaima taseme ligi 25 treeningkorduse juures. Stabiilselt madal ruutviga saavutatakse 40-60 korduse vahel. Testi ruutvea dünaamika sarnanes treeningruutveale. Sellest võib järeldada, et mudel oli suuteline seoseid tuvastama. Erinevalt LSTM(AR) mudelist jääb LSTM(makro) mudeli puhul treeningu ja testi ruutvea vahe suuremaks. Sellest võib järeldada, et makro- ja finantsandmete lisamisega hinna ennustamise täpsus väheneb. Antud töö kontekstis ei ole

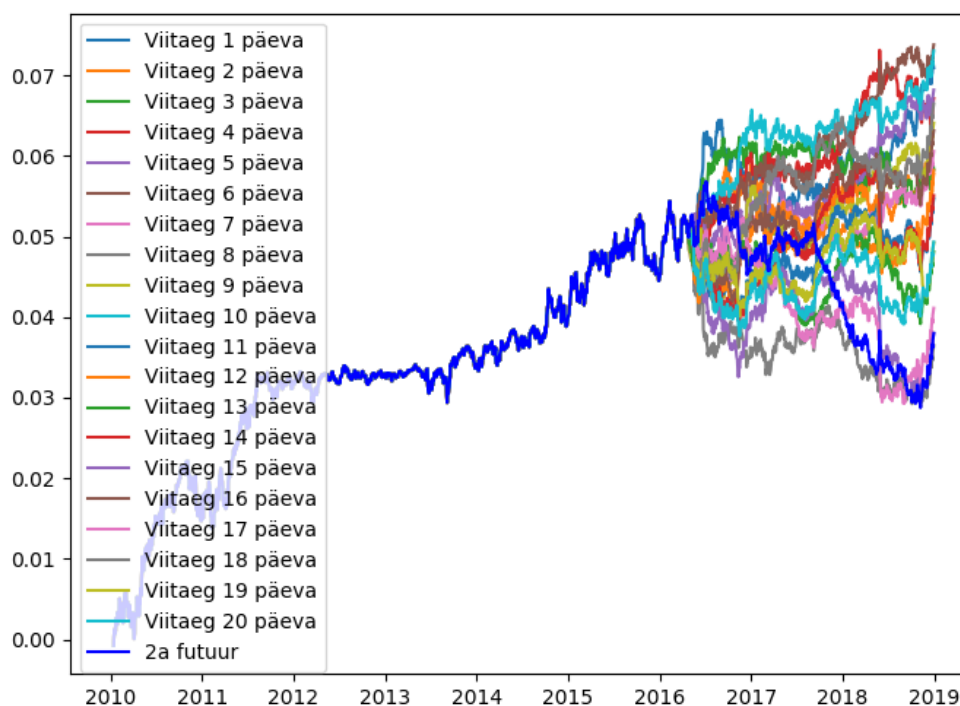
hinna prognoosi täpsus kõige prioriteetsem, sellest olulisem on hinna liikumise suuna ennustamise võimekus.

LSTM(makro) mudeli signaalid (Joonis 16) on võrreldava pikkusega LSTM(AR) mudeli signaalidega (Joonis 12). Sarnaselt muutub signaali suund tihti ning esineb üksikuid samasuunalise signaaliga perioode.



Joonis 16. LSTM(makro) 2-aastane, viitaeg 6 päeva, mudeli signaalid.

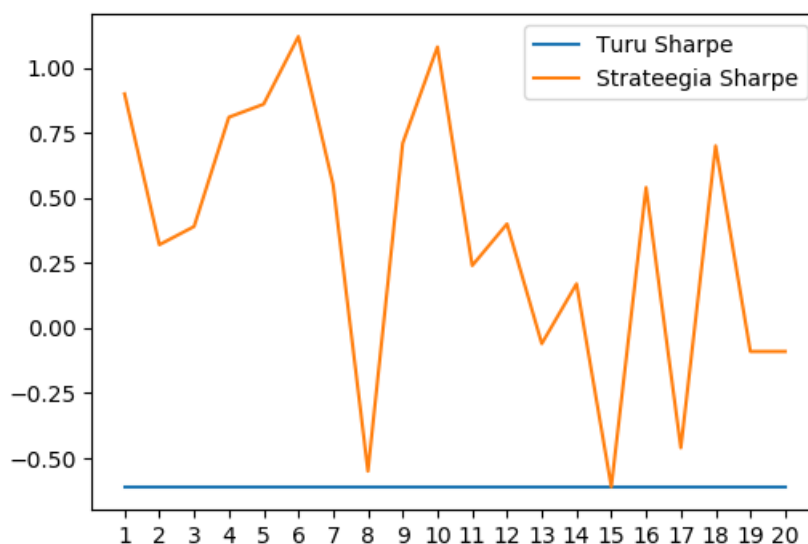
Kasutades mudelis erinevaid viitaegu, olid tulemused (Joonis 17) testperioodi lõpuks valdavalt osta ja hoia ja LSTM(AR) strateegiast (Joonis 13) paremad. Turu kumulatiivne tulemus testperioodil oli -1,33%. LSTM(makro) strateegiatest andis parima tulemuse viitaeg 6 päeva kumulatiivse tulemusega 2,44%.



Joonis 17. LSTM(makro) 2-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.

Üldistades võib taas öelda, et suurimat tulu suutsid vaadeldaval perioodil teenida lühema viitajaga mudelid. Olulise muutusena võrreldes LSTM(AR) mudeliga suutsid suurem osa valitud strategiatest näidata testperioodil positiivset tootlust.

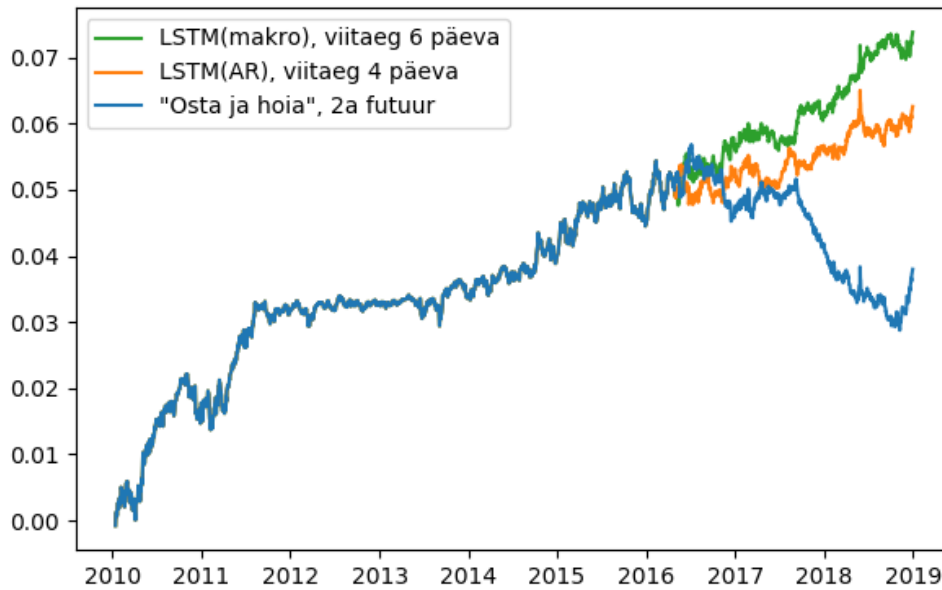
Sarnaselt LSTM(AR) mudelile ei peaks tähelepanu pöörama ainult positiivsele tulemusele vaid ka tulu-riski suhtele.



Joonis 18. LSTM(makro) 2-aastane, viitaeag 1-20 päeva, Sharpe.

Sharpe võrdlusest on näha, et parem tulemus saavutati pigem lühemate viitaeagadega. Enamus valitud viitaeagadest on parema tulemusega kui vaadeldava perioodi turu Sharpe. Parim tulemus saavutatakse 6-päevase viitajaga ning see annab Sharpe suhteks 1,12. Turu Sharpe suhtega sama tulemus saavutati 15-päevase viitajaga. Keskmise LSTM(makro) strateegia Sharpe suhe üle kõikide viitaeagade oli 0,35 ning turu Sharpe vaadeldaval perioodil oli -0,61. See tähendab, et keskmiselt ületas LSTM(makro) strateegia turu Sharpe suhet 0,96 võrra ja LSTM(AR) mudeli Sharpe suhet 0,47 võrra. Makro mudeli keskmiselt suurem Sharpe on tingitud positiivsematest tulemustest.

Valides parimad viitajaga tulemused mõlemast 2-aastase futuuri strateegiast, näeme, et LSMT(makro) strateegia on vaadeldaval perioodil kõige edukam (Joonis 19).



Joonis 19. LSTM(AR) ja LSTM(makro) 2-aastane, kumulatiivsete tootluste võrdlus.

Võrreldes mudelite ruutvigasid, leiti, et LSTM(makro) mudeli hinnaprognooosi täpsus on madalam kui LSTM(AR) mudelil, samas LSTM(makro) mudeli kumulatiivne tootlus on parem kui LSTM(AR) mudelil. Kahe mudeli võrdlusest võime järeldada, et makroandmete kasutamisel suureneb futuuri hinna liikumissuuna ennustamise võimekus ja kauplemisstrateegia signaali täpsus tõuseb (Tabel 2).

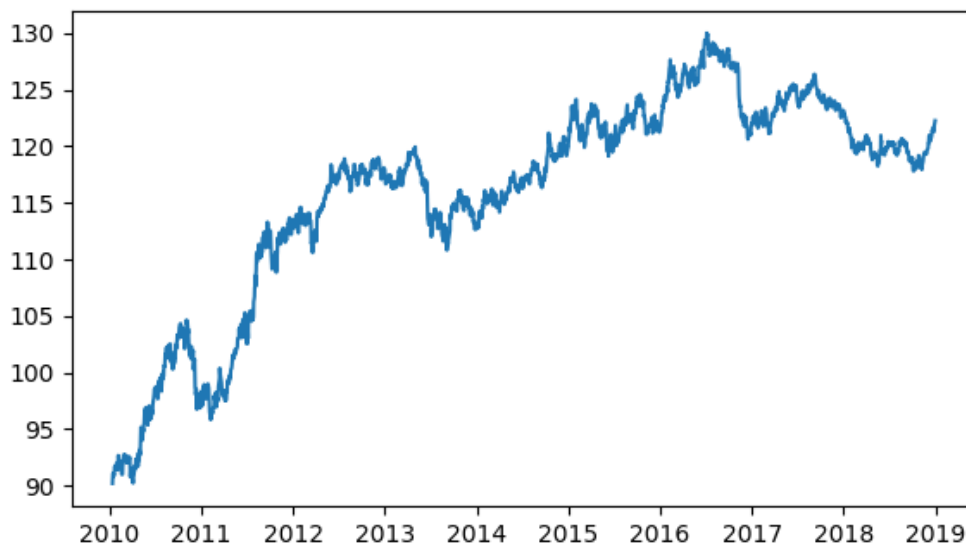
Tabel 2. LSTM(AR) ja LSTM(makro) 2-aastane, mudelite võrdlus.

	Turg	LSTM AR	LSTM makro
Testperioodi kumulatiivne tootlus	-1.33	1.32	2.44
Sharpe suhe	-0.61	0.60	1.12

Mõlemad LSTM mudelid ületasid vaadeldaval perioodil turu tootlust ning olid seejuures kasumlikud. Kolmest võimalikust strateegiast saavutas parima tulemuse LSTM(makro) mudel, kogudes testperioodi kogutootluseks 2,44% ning Sharpe suhtarvuks 1,12. See on oluliselt parem tulemus kui turu tootlus. Eriti kaalukas on asjaolu, et samal perioodil oli turu tootlus negatiivne.

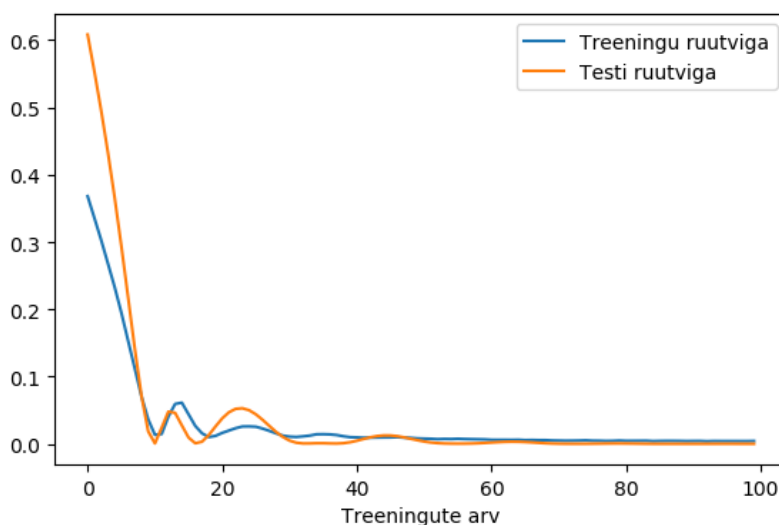
4.3.2 10-aastane USA võlakirja futuur

USA 10-aastase võlakirjafutuuri hind on vaadeldaval perioodil kasvanud (Joonis 20). Hinnadünaamika on sarnane 2-aastasele futuurile. Esimestel aastatel kasvas hind kiiremini ning vahepealsetel aastatel on hinna kasvutempo aeglustunud. Viimastel aastatel on hinnavolatiilsus püsinud sama, kuid hind on langenud. Kui investor oleks vaadeldaval perioodil kasutanud osta ja hoia strateegiat, oleks tema portfell kasvanud 35,5%.



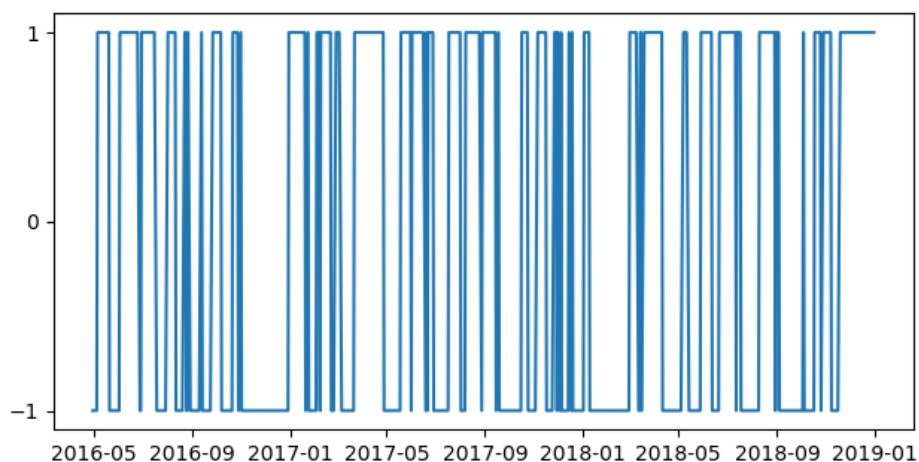
Joonis 20. USA 10-aastase võlakirjafutuuri hind.

Sarnaselt 2-aastasele futuurile, töötati kõigepealt välja LSTM(AR) mudel, mille sisendiks oli aegrea enda ajalugu. Mudelit trenniti vastavalt etteantud parameetritele (Tabel 1), varieerides mudeli viitaegu. Parim tulemus saavutati 10-päevase viitajaga.



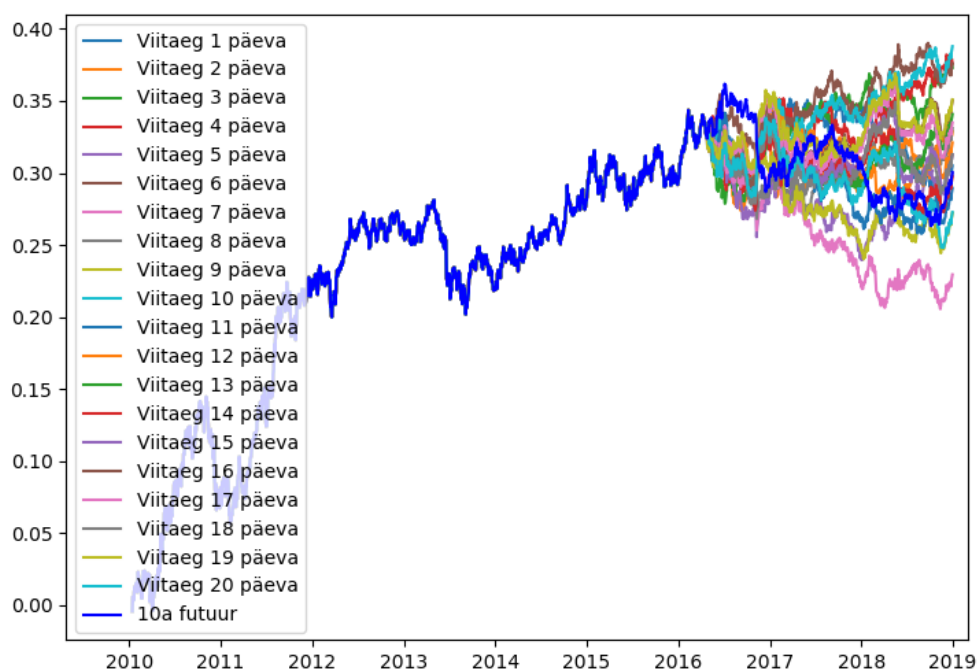
Joonis 21. LSTM(AR) 10-aastane, viitaeg 10 päeva, mudeli ruutviga.

Mudeli treeningu ruutviga kahanes esimeste treeningkorduste juures kiiresti (Joonis 21) ning saavutas oma madalaima taseme ligi 20 treeningkorduse juures. Stabiilselt madal ruutviga saavutatakse 40-60 korduse vahel. Testi ruutvea dünaamika sarnanes treening ruutveale. Sellest võib järeldada, et mudeli parameetrid on optimaalsed ning mudel on saavutanud maksimaalse võimekuse.



Joonis 22. LSTM(AR) 10-aastane, viitaeg 10 päeva, mudeli signaalid.

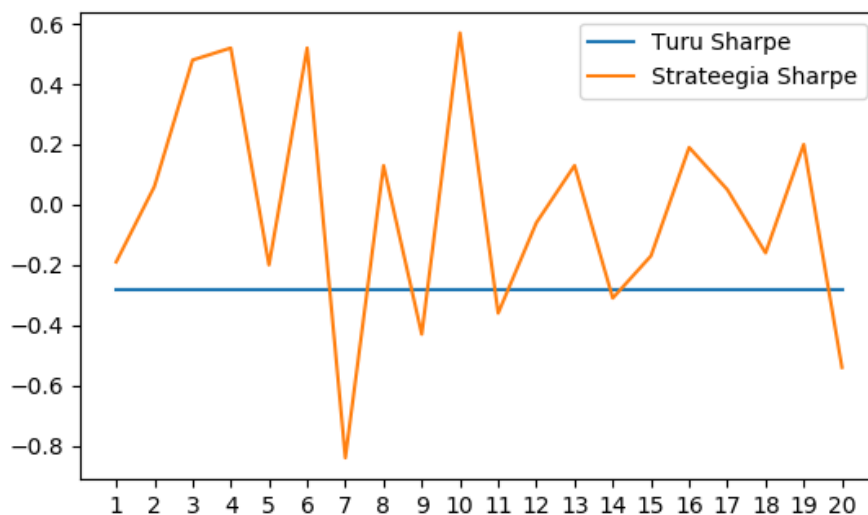
Sarnaselt 2-aastase futuuri LSTM(AR) mudeli signaalidele esineb ka 10-aastase futuuri signaalide (Joonis 22) hulgas nii ostu- kui müügisignaale. Ostusignaalidega võrreldes esineb antud strateegial müügisignaale rohkem, mis langevate futuurihindade puhul annab eelise osta ja hoia strateegia suhtes. Võrreldes 2-aastase LSTM(AR) mudeliga, on signaalid pikemad. Selle üks põhjus on pikem viitaeg.



Joonis 23. LSTM(AR) 10-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.

Kasutades mudelis erinevaid viitaegu, olid tulemused testperioodi lõpuks valdavalt osta ja hoia strateegiast paremad (Joonis 23). Turu kumulatiivne tulemus testperioodil oli -2,98%. LSTM(AR) strateegiatest andis parima tulemuse viitaeg 10 päeva kumulatiivse tulemusega 6,04%. Samuti saavutati turust parem tulemus teiste viitaegadega. Võrreldes 2-aastase futuuri LSTM(AR) mudeliga ei saa väita, et suurimat

tulu suutsid vaadeldaval perioodil teenida lühema viitajaga mudelid. Väga selge seos viitaegade ja tulemuslikkuse vahel puudub. Suurem osa strateegiatest ületasid küll turgu, kuid lõpetasid perioodi negatiivse tootlusega.

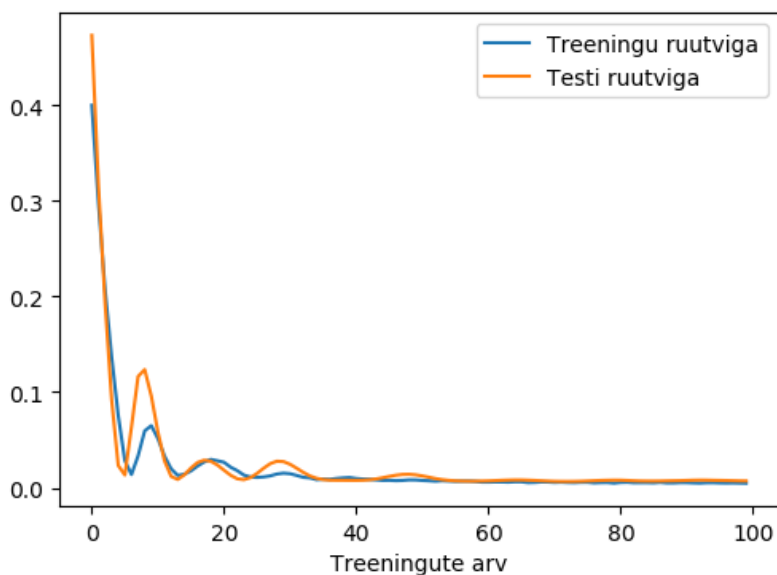


Joonis 24. LSTM(AR) 10-aastane, viitaeg 1-20 päeva, Sharpe.

Sharpe suhte võrdlusest on näha, et selgelt eristuvat parimat viitaega ei ole (Joonis 24). Samuti ei ole Sharpe suhtarvud stabiilselt üle turu. Parim tulemus saavutatakse 10-päevase viitajaga ning see annab Sharpe suhteks 0,57. Samas 7-päevane viitaeg annab turust oluliselt kehvema tulemuse. Keskmine Sharpe suhe üle kõikide viitaegade oli -0,02 ning turu Sharpe vaadeldaval perioodil oli -0,28. See tähendab, et keskmiselt ületas LSTM(AR) strateegia Sharpe turgu 0,26 võrra.

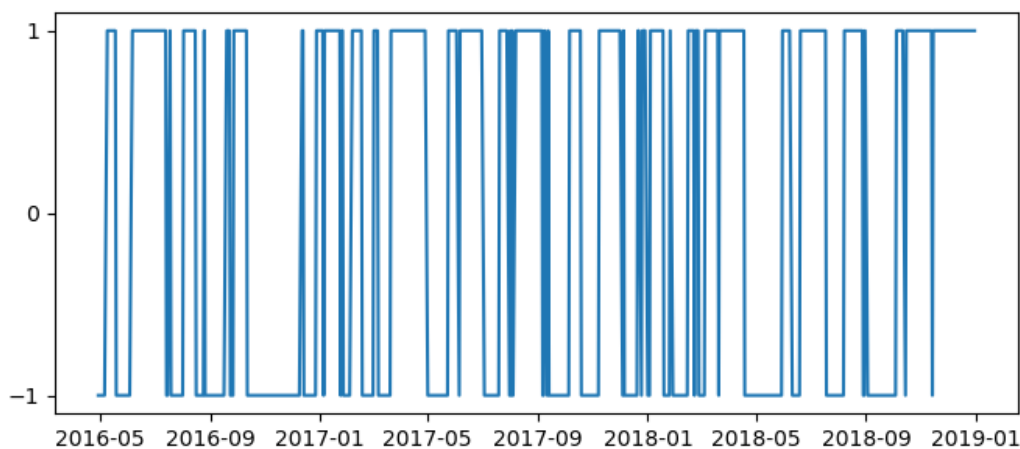
Sarnaselt 2-aastasele futuurile testiti makronäitajatega mudelit 10-aastase võlakirja futuuri peal. LSTM(makro) mudelit treeniti vastavalt etteantud parameetritele (Tabel 1) erinevate viitaegadega. Parim tulemus saavutati 12-päevase viitajaga.

Sarnaselt LSTM(AR) mudelile (Joonis 21) kahanes treeninguruutviga esimeste treeningkorduste juures kiiresti ning saavutas oma madalaima taseme ligi 25 treeningkorduse juures (Joonis 25). Stabiilselt madal ruutviga saavutatakse 40-60 korduse vahel. Testi ruutvea dünaamika sarnanes treening ruutveale. Sellest võib järeldada, et mudel oli suuteline seoseid tuvastama. Erinevalt 2-aastase futuuri makro mudelist ei suurene 10-aastase futuuri mudeli ruutviga kui lisada makroandmed. Sellest võib järeldada, et olenemata suurematest sisendandmetest, suudab mudel leida optimaalsed parameetrid.



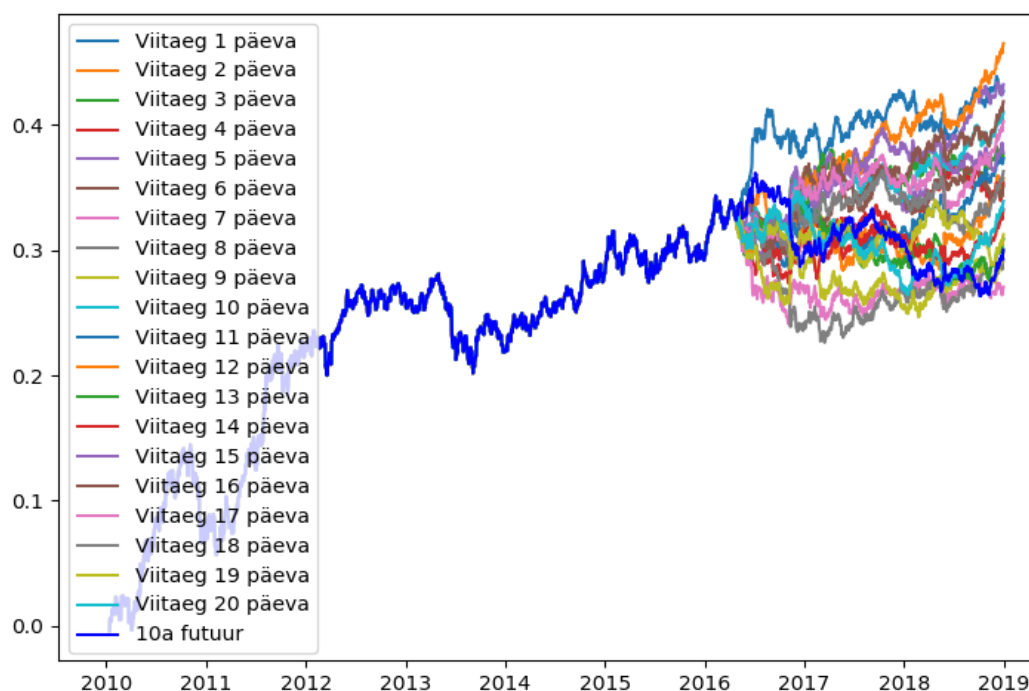
Joonis 25. LSTM(makro) 10-aastane, viitaeg 12 päeva, mudeli ruutviga.

LSTM(makro) mudel muudab signaali suunda harvemini, st signaalid on pikemad (Joonis 26) võrreldes LSTM(AR) mudeli signaalidega (Joonis 22). Selle põhjuseks võib olla mõnevõrra pikem viitaeg ning samuti muutuvad makronäitajad harvemini ning on stabiilsemad kui futuuri enda hinnarida.



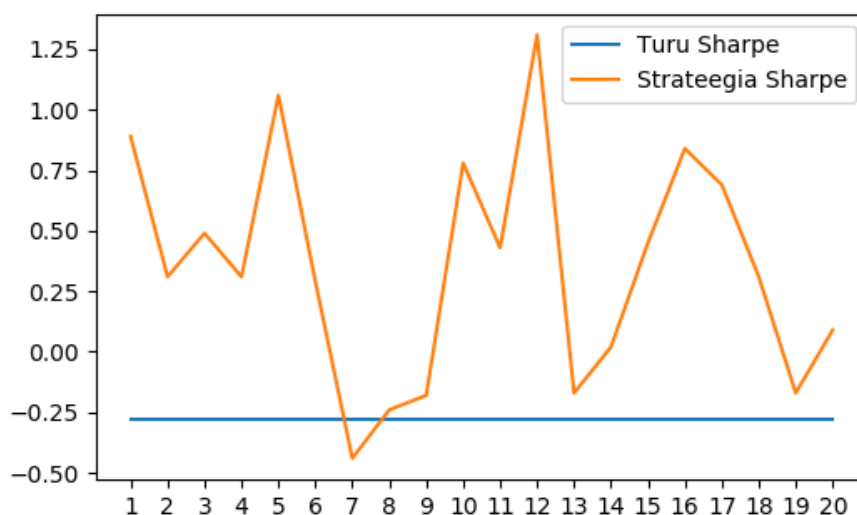
Joonis 26. LSTM(makro) 10-aastane, viitaeg 12 päeva, mudeli signaalid.

Kasutades makro mudelis erinevaid viitaegu, olid tulemused testperioodi lõpuks valdavalt osta ja hoia strateegiast paremad (Joonis 27). Turu tulemus oli testperioodil -2,98%. Parima tulemuse andis viitaeg 12 päeva kumulatiivse tulemusega 13,79%. Võrreldes 10-aastase futuuri LSTM(AR) mudeliga, esineb LSTM(makro) mudelis selge seos viitaegade ja tulemuslikkuse vahel. Enamus viitaegasid annab turuga võrreldes parema tulemuse, kuid sarnaselt 10-aastasele LSTM(AR) mudelile annab negatiivse tulemuse 7-päevane viitaeg. Suurem osa strateegiatest ületasid turu ja lõpetasid perioodi positiivse tootlusega. Seevastu turu tootlus oli testperioodil negatiivne.



Joonis 27. LSTM(makro) 10-aastane, viitaeg 1-20 päeva, kumulatiivne tulemus.

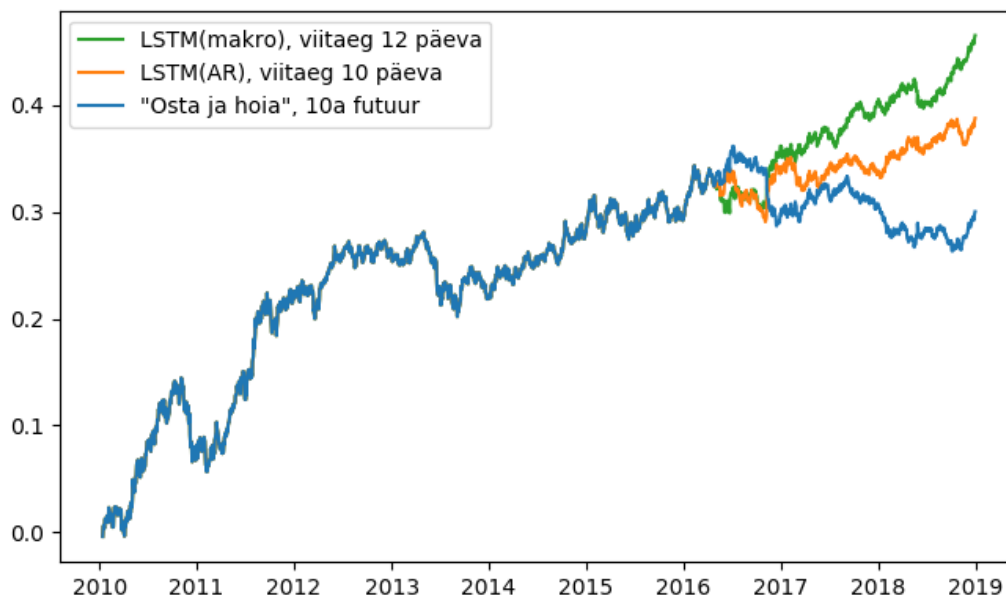
Sharpe võrdlusest (Joonis 28) on näha, et turust parem tulemus saavutati pea kõikide vaadeldud viitaegadega. Turu Sharpe suhtest madalam tulemus saavutati viitajaga 7 päeva. Parim tulemus saavutatakse 12-päevase viitajaga ning see annab Sharpe suhteks 1,31.



Joonis 28. LSTM(makro) 10-aastane, viitaeg 1-20 päeva, Sharpe.

Keskmine Sharpe suhe üle kõikide viitaegade oli 0,35. See tähendab, et keskmiselt ületas LSTM(makro) strateegia Sharpe turgu 0,63 võrra ja LSTM(AR) strateegia keskmist 0,37 võrra. Seega, võrreldes LSTM(AR) mudeliga on LSTM(makro) mudeli erinevate viitaegade keskmine Sharpe oluliselt kõrgem. Makro mudeli keskmiselt suurem Sharpe on tingitud positiivsematest tulemustest võrreldes LSTM(AR) mudeliga.

Valides parimad tulemused mõlemast 10-aastase futuuri strateegiast, näeme, et LSTM(makro) strateegia 12-päevase viitajaga on vaadeldaval perioodil kõige edukam (Joonis 29).



Joonis 29. LSTM(AR) ja LSTM(makro) 10-aastane, kumulatiivsete tootluste võrdlus.

Võrreldes 10-aastase futuuri mudelid leiti, et LSTM(makro) mudeli kumulatiivne tootlus on oluliselt suurem kui LSTM(AR) mudelil. Kahe mudeli võrdlusest võime ka 10-aastase futuuri puhul järeldada, et makroandmete kasutamisel suureneb futuuri hinna liikumissuuna ennustamise võimekus ja signaali täpsus tõuseb (Tabel 3).

Tabel 3. LSTM(AR) ja LSTM(makro) 10-aastane, mudelite võrdlus.

	Turg	LSTM AR	LSTM makro
Testperioodi kumulatiivne tootlus	-2.98	6.04	13.79
Sharpe suhe	-0.28	0.57	1.31

Kolmest võimalikust strateegiast saavutas parima tulemuse LSTM(makro) mudel, kogudes testperioodi kogutootluseks 13,79% ning Sharpe suhtarvuks 1,31. See on märgatavalt parem tulemus kui turu tootlus. Eriti kaalukas on asjaolu, et samal perioodil oli turu tootlus negatiivne.

4.4 Edasine töö

Käesoleva töö tulemused näitavad, et LSTM närvivõrgu abil on võimalik konstrueerida edukas kauplemisstrateegia USA valitsusvõlakirjade futuuridele. Erinevad alusvarad seevastu võivad anda erinevaid tulemusi ning erinevate sisendandmete kasutamine

muudab samuti tulemust. Need muudatused vajaksid täiendavaid uuringuid ning ei kuulu antud töö koosseisu.

Töös kasutatud analüüsiperioodi muutmine võib täiendavalt muuta strateegiate tulemust. Huvitav oleks uurida, kas ning kui palju erinevate ajaakende nihutamine kogutulemust mõjutab. Näiteks võiks eraldi uurida ainult väga volatiilseid või väga stabiilseid perioode ning mudeli edukust intresside tõusu- ja langusperioodidel. Käesolevas töös analüüsiperioodi ei muudetud ning vaadeldi viimast 9 aastat.

LSTM mudel sisaldab väga paljusid parameetreid. Kõikide parameetrite muutmine omab mudeli tulemuslikkusele mõju. Täiendavalt oleks vajalik läbi viia hüperparameetrite testimine. Kindlasti tuleb arvestada, et ajaloolistel andmetel on mudelit võimalik üle optimeerida. Käesolevas töös uuriti viitaja muutuse ja sisendandmete mõju mudeli tulemustele. Teised parameetrid jäeti võrdluse tekitamiseks konstantseks.

Tulevikus oleks huvitav uurida investeerimishorisoni muutuse mõju tulemusele. Uurida tuleks nädalast ning kuist horisonti, kus positsioon võetakse vastavalt kogu nädalaks või kuuks. Käesolevas töös oli investeerimishorisoni pikkuseks üks päev ning signaalid arvutati igapäevaselt.

Kokkuvõte

Masinõpe on kasutusel väga paljudes valdkondades. Üks võimalik viis masinõppe rakendamiseks on finantsandmete prognoosimine. Uuringud on näidanud, et masinõppe mudelid suudavad saavutada paremaid tulemusi võrreldes klassikaliste ökonomeetriliste mudelitega.

Käesolevas töös uuriti erinevaid masinõppe meetodeid ning varasema kirjanduse põhjal leiti, et võlakirjaturul rakendamiseks on kõige sobivam pikk-lühiajalise mälu närvivõrk. Töös hinnati LSTM-i suutlikust prognoosida USA võlakirjaturu liikumise suunda. Koostatud mudeli eesmärk oli prognoosida, kas alusvara hind järgmisel päeval tõuseb või langeb. Vastavalt mudeli prognoosile võeti järgmiseks päevaks kas ostu- või müügiotsus. Mudeli hindamiseks võrreldi selle tulemust klassikalise osta ja hoia strateegiaga.

Uuringus rakendati kahte lähenemist. Esimesel juhul oli mudeli sisendiks ainult alusvara enda eelnev hinnarida. Teisel juhul lisati sisenditena makro- ja finantsandmed. Lisaks kahele erinevale lähenemisele uuriti erinevate viitaegade mõju prognoositäpsusele. Viitaegade vahemik ulatus ühest päevast kuni 20 päevani.

USA 2-aastase valitusvõlakirja futuurile rakendatud LSTM mudelitest suutis parima tulemuse anda LSTM(makro) mudel viitajaga 6 päeva, mille testperioodi tootluseks oli 2,44% ning Sharpe suhtarvuks 1,12. Parim LSTM(AR) mudel oli viitajaga 4 päeva, mille testperioodi tootlus oli 1,32% ning Sharpe suhtarv 0,60. Turupõhine osta ja hoia strateegia testperioodi tootlikkus oli -1,33% ning Sharpe suhe -0,61. Mõlemad LSTM mudelid suutsid valdavalt näidata paremaid tulemusi kui passiivne lähenemine. Makro- ja finantsandmete lisamine LSTM mudelile tõstis märgatavalt mudeli tulemuslikkust.

USA 10-aastase valitusvõlakirja futuurile rakendatud LSTM mudelitest suutis parima tulemuse anda LSTM(makro) mudel viitajaga 12 päeva, mille testperioodi tootluseks oli 13,79% ning Sharpe suhtarvuks 1,31. Parim LSTM(AR) mudel oli viitajaga 10 päeva, mille testperioodi tootlus oli 6,04% ning Sharpe suhtarv 0,57. Turupõhine osta ja hoia strateegia testperioodi tootlikkus oli -2,98% ning Sharpe suhe -0,28. Mõlemad LSTM

modelid suutsid valdavalt näidata paremaid tulemusi kui passiivne lähenemine. Makro- ja finantsandmete lisamine LSTM mudelile tõstis märgatavalt mudeli tulemuslikkust.

Oluline on märkida, et testperioodil oli passiivse strateegia tulemus negatiivne, kuid aktiivsed lähenemised suutsid vaatamata langevale turuhinnale pakkuda teatud parameetrite juures positiivset tootlust.

Üle kõikide töös testitud viitaegade olid keskmised Sharpe suhted LSTM(AR) mudelite puhul vastavalt 2-aastase futuuri korral -0,12 ning 10-aastase futuuri korral -0,02. LSTM(makro) mudeli keskmised Sharpe suhted üle kõikide viitaegade olid 2-aastase futuuri korral 0,35 ning 10-aastase futuuri korral 0,35. Turu Sharpe vaadeldaval perioodil oli 2-aastaselt futuuril -0,61 ja 10-aastaselt futuuril -0,28. Võrreldud 2- ja 10-aastase futuuri hinnaliikumised on erinevad, kuid tulemused näitavad, et makro- ja finantsandmete lisamine muutis LSTM-i tulemused paremaks nii LSTM(AR) kui passiivse strateegia suhtes. Seda valitud viitaegade hulgas kui ka keskmiselt üle kõigi viitaegade.

Kokkuvõtteks võib järeldada, et LSTM mudel on võimeline töötlemas laiemat hulka andmeid ja selle põhjal hinna liikumissuunda ennustama ning seeläbi tootlust parandama. Lisades LSTM mudelile täiendavaid makro- ja finantsandmed on võimalik strateegia edukust tõsta.

Edasine analüüs oleks vajalik täiendavate mudeli parameetrite testimiseks. Huvitavaid tulemusi võivad anda erinevate sisendandmete testimine ning erineva volatiilsusega perioodide analüüs.

Kasutatud kirjandus

- [1] „International banking and financial market developments,“ Bank for International Settlements, 2019.
- [2] CME Group, „Interest Rate Futures Liquidity Update – CME Group,“ CME Group, 07 01 2019. [WWW] <https://www.cmegroup.com/education/featured-reports/interest-rate-futures-liquidity-update.html> (13.02.2019)
- [3] „2018 Full Year Market Highlights report,“ The World Federation of Exchanges, 2019.
- [4] J. L. Maginn, D. L. Tuttle, D. W. McLeavey ja J. E. Pinto, „Managing investment portfolios : a dynamic process,“ John Wiley & Sons, Inc., New Jersey, 2007.
- [5] R. C. Grinold, „The fundamental law of active management,“ *The Journal of Portfolio Management*, kd. 15, nr 3, pp. 30-37, Spring 1989.
- [6] F. J. Fabozzi, Fixed Income Analysis for the Chartered Financial Analyst Program, New Hope: Frank. J Fabozzi Associates, 2000.
- [7] C. Moorad, An Introduction to Bond Markets, Cornwall: John Wiley & Sons, Incorporated, 2010.
- [8] J. C. Hull, Options, Futures and other derivatives 8th ed., Harlow: Pearson Education Limited, 2011.
- [9] S. C. Ludvigson ja N. Serena, „Macro Factors in Bond Risk Premia,“ *Review of Financial Studies*, kd. 22, nr 12, pp. 5027-5067, 2009.
- [10] J. H. Cochrane ja M. Piazzesi, „Bond Risk Premia,“ *American Economic Review*, kd. 95, 2005.
- [11] E. F. Fama ja R. H. Bliss, „The Information in Long-Maturity Forward Rates,“ *American Economic Review*, kd. 77, 1987.
- [12] J. Y. Campbell ja R. J. Shiller, „Yield Spreads and Interest Rates: A Bird’s Eye View,“ *Review of Economic Studies*, kd. 58, p. 495–514, 1991.
- [13] T. Paas, Sissejuhatus ökonomeetriasse, Tartu: Tartu Ülikooli Kirjastus, 1995.
- [14] D. A. Hsieh, „Nonlinear Dynamics in Financial Markets: Evidence and Implications,“ *Financial Analysts Journal*, kd. 51, nr 4, pp. 55-62, 1995.
- [15] R. Cavalcante, R. Brasileiro, V. Souza, J. Nobrega ja A. Oliveira, „Computational Intelligence and Financial Markets: A Survey and Future Directions,“ *Expert Systems with Applications*, kd. 55, pp. 194-211, 2016.
- [16] S. Shalev-Shwartz ja S. Ben-David, Understanding Machine Learning From Theory to Algorithms, New York: Cambridge University Press, 2014.
- [17] J. P. Mueller ja L. Massaron, Machine Learning For Dummies, New Jersey: John Wiley & Sons, 2016.
- [18] J. Brownlee, Master Machine Learning Algorithms, 2016 [WWW] <http://index-of.es/Varios-2/Master%20Machine%20Learning%20Algorithms.pdf> (04.04.2019)
- [19] I. Vasilev, D. Slater, G. Spacagna, P. Roelants ja V. Zocca, Python Deep Learning Second Edition, Birmingham: Packt Publishing Ltd, 2019.

- [20] K. Gurney, *An introduction to neural networks*, London: UCL Press, 1997.
- [21] S. Haykin, *Neural Networks and Learning Machines* 3rd, New Jersey: Pearson Education, 2009.
- [22] N. Buduma, *Fundamentals of Deep Learning*, Sebastopol: O'Reilly Media, 2017.
- [23] A. Smola ja S. Vishwanathan, *Introduction to Machine Learning*, Cambridge: Cambridge University Press, 2008.
- [24] J. Patterson ja A. Gibson, *Deep Learning A Practitioner's Approach*, Sebastopol: O'Reilly Media, 2017.
- [25] S. Hochreiter ja J. Schmidhuber, „Long Short-Term Memory,“ *Neural Computation*, p. 1735–1780, 1997.
- [26] J. Schmidhuber, F. Cummins ja F. A. Gers, „Learning to forget: Continual prediction with LSTM,“ *Neural computation*, kd. 12, pp. 2451-2471, 2000.
- [27] I. Goodfellow, Y. Bengio ja A. Courville, *Deep Learning*, MIT Press, 2016.
- [28] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Toronto: Springer, 2012.
- [29] M. Y. Salem, Z. Zhao, J. Barry, J. Hunter, D. Sarkar, P. White, A. Tepper ja L. Y. Chang, „Do androids dream of electric bonds? Machine learning in interest rate markets,“ J. P. Morgan, 2017.
- [30] S. Ganguli ja J. Dunnmon, „Machine Learning for Better Models for Predicting Bond Prices“, 2017 [WWW] <https://arxiv.org/abs/1705.01142> (06.03.2019)
- [31] M. Nunes, E. Gerding, F. McGroarty ja M. Niranjana, „Artificial Neural Networks in Fixed Income Markets for Yield Curve Forecasting“, 2018 [WWW] <https://ssrn.com/abstract=3144622> (02.02.2019)
- [32] C. L. Dunis ja V. Morrison, „The economic value of advanced time series methods for modelling and trading 10-year government bonds,“ *European Journal of Finance*, pp. 333-352, 2007 .
- [33] D. Bianchi, M. Büchner ja A. Tamoni, „Bond risk premia with machine learning“, 2018 [WWW] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3232721 (01.03.2019)
- [34] W. Cheng, L. Wagner ja C. H. Lin, „Forecasting the 30-year U.S. Treasury Bond with a System of Neural Network,“ *NeuroWest Journal* 1.2, 1996.
- [35] A. Din, „Forecasting of interest rates using neural network models. Geneva Research Collaboration“, Geneva, 2003 [WWW] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.96.324&rep=rep1&type=pdf> (25.02.2019)
- [36] M. P. Naeini, H. Taremian ja H. B. Hashemi, „Stock market value prediction using neural networks,“ *International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*, 2010.
- [37] K. Nakagawa, T. Ito, M. Abe ja K. Izumi, „Deep Recurrent Factor Model: Interpretable Non-Linear and Time-Varying Multi-Factor Model“, 2019 [WWW] <https://arxiv.org/abs/1901.11493v1> (20.02.2019)
- [38] D. Q. Nelson, A. M. Pereira ja R. A. Oliviera, „Stock Market's Price Movement Prediction With LSTM Neural Networks,“ *International Joint Conference on Neural Networks (IJCNN)*, pp. 1419-1426, 2017.
- [39] Keras: The Python Deep Learning library [WWW] <https://keras.io/> (15.03.2019)

[40] Tensorflow [WWW] <https://www.tensorflow.org/> (20.03.2019)

[41] F. Federal Reserve Economic Data [WWW]
<https://research.stlouisfed.org/econ/mccracken/fred-databases/> (25.02.2019)