

TALLINNA TEHNIKAÜLIKOOL

Majandusteaduskond

Dmitri Zhilyaev

**SENTIMENDI MÕJU ANALÜÜS ERINEVATE
KRÜPTOVARADE HINDADELE**

Bakalaureusetöö

Õppekava Ärindus, peeriala Ärirahandus

Juhendaja: Tõnn Talpsepp, PhD

Tallinn 2024

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

Töö pikkuseks on 6151 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Dmitri Zhilyaev 09.05.2024

(kuupäev)

SISUKORD

LÜHIKOKKUVÕTE	4
SISSEJUHATUS	5
1. KRÜPTOVARAD JA ANALÜÜSI MEETODID	7
1.1. Krüptovarad finantsturgudel ja nende liigitus	7
1.2. Krüptovarade analüüsi meetodid	9
1.3. Varasemad uuringud	10
2. ANDMED JA METOODIKA	12
2.1. Valim ja andmete kirjeldus	12
2.2. Metoodika	15
3. ANALÜÜS JA TULEMUSED	19
3.1. Pikaajaliste mudelite analüüs	19
3.2. Lühiajaliste mudelite analüüs	22
3.3. Sentimendi lisamine mudelisse	24
3.4. Järeldused ja piirangud	26
KOKKUVÕTE	29
SUMMARY	31
KASUTATUD ALLIKATE LOETELU	33
LISAD	36
Lisa 1. Ühikjuure testimine Dickey-Fuller Testiga	36
Lisa 2. Pikaajaliste mudelite testide tulemused	36
Lisa 3. Lühiajaliste mudelite testide tulemused	37
Lisa 4. Sentimendiga mudeli testide tulemused	37
Lisa 5. VIF Test	37
Lisa 6. Lihtlitsents	38

LÜHIKOKKUVÕTE

Käesoleva töö eesmärgiks on analüüsida sentimentide mõju erinevate krüptovarade hindadele. Töö eesmärgi saavutamiseks ning uurimisküsimustele vastamiseks koostab autor pikaajalised ja lühiajalised mudelid. Täiendava analüüsi jaoks püstitatakse mudel ilma sentimentide hinnanguta ning seejärel lisatakse sentiment ja analüüsitakse kahte mudelit sentimentide mõju tuvastamiseks.

Pikaajalised mudelid mis hõlmasid perioodi 18. augustist 2017 kuni 21. jaanuarini 2019, näitavad, et sentimentil ei ole suur mõju krüptovarade hindadele. Lühiajalised mudelid, mis hõlmasid kolme juhuslikult valitud 90-päevast perioodi ajavahemikus 18. augustist 2017 kuni 21. jaanuarini 2019, näitavad, et sentimentil ei ole suurt mõju Bitcoin hindadele. Sentimentide lisamine mudelisse ilma sentimentideta näitas, et sentimentide lisamine võib aidata Bitcoin hinna paremini kirjeldada või prognoosida.

Võtmesõnad: inimeste sentiment, sentimentide mõju krüptovaradele, Bitcoin, Ethereum, Ripple, lühiajaline periood, pikaajaline periood

SISSEJUHATUS

Umbes kõik täisealised ja alaealised teavad krüptovarade olemasolust maailmas. Krüptovarad on populaarsust ajaga kogunud ja on muutunud üheks finantsturgude osaks. Krüptovarade hindade tõus turgudel võib olla kiire, samamoodi võivad suured ja kiired olla ka hindade langused. Erinevad krüptovarad näitavad suurt volatiilsust mis põhjendab ka hindade kõikuvust (Katsiampa, 2018).

Viimastel aastatel on paljud inimesed hakanud rääkima krüptovaradest ning kaaluvad nendesse investeerimist. Sotsiaalmeedias võib leida postitusi Bitcoinist ja Ethereumist ning nende hindade analüüsi. Krüptovarade hinna prognoosimiseks või analüüsimiseks kasutatakse tavaliselt tehnilist analüüsi või fundamentaalset analüüsi (Akgül et al., 2022).

Tehnilise või fundamentaalse analüüsi täiendamiseks võib kasutada ka sentimendi analüüsi. Sentimendi analüüs on loomuliku keele töötlemise uurimisvaldkond, mis tegeleb subjektiivsete elementide meeoleolu või arvamuse tuvastamisega tekstis andmebaasina kasutades kommentaare või postitusi. (Bhadane et al., 2015) Sentimendi roll krüptovarade hindade prognoosimisel tundub loogiline, kuna sentimendi analüüsi kasutatakse ka tava aktsiate turu analüüsimiseks (Xiao & Ihnaini, 2023). Sotsiaalmeedias on palju postitusi ja uudiseid krüptovarade kohta. Sotsiaalmeedia ja uudiste portaalide kaudu levivad uudised, spekulatsioonid ja erinevad seisukohad krüptovarade vastu. Seega on oluline uurida, kuidas sentimendi analüüs võib mõjutada krüptovarade hindu ning kuidas seda mõju saaks kvantitatiivselt kirja panna.

Krüptovarade jaoks sentimendi analüüs ei ole tavaliselt kasutatud. Krüptovarade turg on väga spekulatiivne mis see järel muudab keeruliseks eristada õige informatsiooni ja vale informatsiooni (Peterson, 2020). Isegi sentimendi hinnata on keeruline sest informatsiooni kogumine toimub interneti kaudu ning internetis on palju boteid või roboteid mis võivad mõjutada saadud hinnanguid.

Selle lõputöö eesmärk on analüüsida sentimentide mõju erinevate krüptovarade hindadele. Sentimentide mõju analüüsimiseks autor valis kolm krüptovara Bitcoin, Ethereum ja Ripple ning püstitas järgmised uurimisküsimused:

- 1) Milline on seos inimeste sentimentide (positiivne/neutraalne/negatiivne) ja krüptovara hindade vahel Bitcoin, Ethereum ja Ripple näitel?
- 2) Kas leitud mõju erineb lühiajalisel ja pikaajalisel perioodil?
- 3) Kas sentimentide kasutamine aitab paremini krüptovarade hindu prognoosida/kirjeldada?

Autor rakendab töö koostamisel kvantitatiivsed meetodeid. Andmebaasi saamiseks laeb autor internetis Bitcoin, Ethereum ja Ripple hindu. Sentimentide analüüsiks autor laeb alla internetis inimeste Twitteri kommentaare, mis on liigitatud positiivseteks, negatiivseteks ja neutraalseteks. Pärast kommentaaride liigitamisest antakse hinnang inimeste sentimentidele ja kasutatakse need andmed mudelite püstitamiseks. Andmete töötlemiseks kasutab autor Excelit ja statistiliseks analüüsiks Gretl tarkvara.

Sentimentide mõju leidmiseks kasutab autor erinevaid ajaintervalle, nagu lühiajaline periood (90 päeva) ja pikaajaline periood (üks aasta ja rohkem). Sentimentide mõju tõendamiseks teostab autor Gretlis korrelatsiooni analüüsi. Korrelatsiooni analüüsis testitakse seose Bitcoin, Ethereum ja Ripple hindade ja inimeste sentimentide vahel. Seose leidmiseks püstitab autor OLS mudeleid ning analüüsib neid pikaajaliselt ja lühiajaliselt, kasutades ainult sentimentide. Pärast kasutatakse sentimentide lisamist mudelisse. Detailsema sentimentide mõju analüüsimiseks loob autor statistiliselt olulist mudelit krüptovara hinna kohta ja täiendab seda seejärel sentimentidega. Uurimise käigus analüüsitakse, kas ilma sentimentidega oleva mudeli krüptovara hinna kirjeldamisevõime suurenes pärast sentimentide lisamist.

Bakalaurusetöö koosneb kolmest peatükist. Esimeses peatükis esmaselt räägitakse lühidalt krüptovaradest turgudel ja krüptovarade liigitusest. Pärast räägitakse erinevate krüptovarade analüüsimeetoditest ja seejärel keskendutakse käesolevas töös kasutatud analüüsi meetodile. Teise peatükis on välja toodud uurimise valim ja kvantitatiivsed arvutused koos kasutatud uuringu meetodikaga. Kolmandas peatükis koostatakse ja analüüsitakse mudeleid ning esitatakse saadud tulemused ja järeldused ning pakutakse ettepanekuid edasiste uurimiste läbiviimiseks.

1. KRÜPTOVARAD JA ANALÜÜSI MEETODID

Antud peatükk annab ülevaate krüptovarade kujunemisest maailmaturul. Selles peatükis tuuakse välja krüptovarade põhimõtted, nende hinnad, hinna muutused, krüptovarade liigitused ning erinevad analüüsimeetodid, mida tavaliselt kasutatakse. Lisaks esitakse varasemate sentimendiga seotud uuringute tulemusi.

1.1. Krüptovarad finantsturgudel ja nende liigitus

DeVries (2019) selgitas oma uuringus, et krüptovara on digitaalne valuutavorm. Bitcoin on kasvavas populaarsuses ning näitab tugevat tõusutrendi kui elujõuline valuuta ja varjupaik inflatsiooniga võitlevatele riiklikele valuutadele. Krüptovaradel puudub keskne väljaandev või reguleeriv asutus, selle asemel kasutatakse deentraliseeritud süsteemi tehingute salvestamiseks. Krüptovara omab potentsiaali saavutada aktsepteerimist investorite hulgas, kes otsivad kuhu investeerida ebastabiilsetest globaalsetest turgudest.

Krüptovarad ei sõltu pankadest tehingute kinnitamiseks. See on rahasüsteem, mis võimaldab kellelgi, kus iganes, saata ja vastu võtta makseid. Pärast krüptovarade tehingut salvestatakse seda krüptovara pearaamatusse ehk Blockchain tehnoloogia. Tripathi et al. (2023b) kirjeldas blockchain tehnoloogiat nagu digitaalne hajutatud pearaamat, mis salvestab andmeid plokkide kujul. Plokkid on omavahel seotud krüptograafilise funktsiooni abil.

Esimeseks krüptovaraks maailmas oli Bitcoin (BTC), mis asutati 2009. aastal tehtud anonüümse isiku või rühma poolt, kes kasutas pseudonüümi Satoski Nakamotoga (Wang & Hausken, 2024). See on praegu kõige tunduim ja kõige suurima turu kapitaliseerimisega krüptovara. Ethereum on ka krüptovara oma blockchainiga, mida nimetatakse Etheriks (ETH). See krüptovara oli loodud 2015 aastal ja on kõige populaarsem krüptovara pärast Bitcoin. Kolmas krüptovara mis oli valitud uurimiseks oli loodud 2012 aastal. See krüptovara on Ripple või XRP ja võtab 2024 aastal kaheksas koht turu populaarsuses. Nende krüptovarade tänapäevane turu kapitaliseerimine aastal

2024 on Bitcoinil 1,285 triljoni dollarit, Ethereumil 405 miljardit dollarit ja Ripple 33 miljardit dollarit (Coingecko, 2024).

Krüptovarade volatiilsus on tavaliselt suur (Katsiampa, 2018). Hindade kõikumised võivad olla märkimisväärselt suured lühikese aja jooksul. See pakub investoritele võimalust saada suurt kasumit või tootlust kui ka suurt kahjumit. Krüptovara turud võimaldavad investoritel genereerida ebataavalisi tulusid või tootlusi krüptovara suurenenud ebakindluse perioodidel, nagu COVID-19 perioodil. (Katsiampa et al., 2022)

See hindade volatiilsus tekitab ka hinna anomaaliaid ja püstitab spekulatiivsust uurijatel ja finantsturgude investoritel. Analüüs kinnitas anomaaliaid Bitcoin hindades aastatel 2013, 2017 ja 2019. Võib öelda suure kindlusega, et Bitcoin hind on olnud pettuslikult manipuleeritud mingil hetkel alates selle loomisest 2009. aastal. (Peterson, 2020) Krüptovarad mõjutavad ka tavalaseid ettevõtteid ja finantsturu, sest isegi suured ettevõtet nagu Microsoft, KFC, Wikipedia võtavad Bitcoin makseviisina vastu (Reed, 2024).

Lisaks, krüptovaradel on erinevad liigitused. Krüptovarasid liigitatakse Bitcoiniks ja altcoiniks. Altcoinid viitavad kõikidele krüptovaradele peale originaalse krüptovara Bitcoin. Altcoinid annavad investoritele omadusi ja funktsioone, mis võivad Bitcoin puhul puuduvad olla. Altcoinid tuginevad samale või sarnasele Blockchain tehnoloogiale ning proovivad täiendada või laiendada Bitcoin omadusi. (Ciaian et al., 2018) Üheks altcoini liigituseks on stablecoin, mis on krüptovara mis oli tehtud volatiilsuse leevendamiseks ja hinnakõikumisi kõrvaldamiseks, sidudes selle krüptovara väärtuse metallidega nagu kuld või hõbe ja Fiat valuutaga (Grobys et al., 2021). Tavaliselt on nende stable coinsite hinnad nagu dollaril ehk stable coin USDT maksab tavaliselt üks dollar. Privacy coin on krüptovara mis keskendub tugevalt tehingute turvalisusele ja anonüümsusele. Tavaliselt on need valuutad välja töötatud, et kustutada ja krüpteerida iga tehingu päritolu ja sihtkohta. Selle krüptovara näide on Monero või XMR. (Hilmola, 2021). Selle bakalaurusetöös uuritakse Bitcoin ja kahte altcoini nagu ETH ja XRP.

1.2. Krüptovarade analüüsi meetodid

Krüptovarade hindade analüüsimiseks kasutatakse erinevaid meetodeid. Analüüsi kasutatakse krüptovarade hindade prognoosimiseks ja investeerimisotsuste tegemiseks. Iga meetod keskendub erinevatele aspektidele ja omab oma eeliseid ja riske. Kaks peamist meetodit, mida kasutatakse krüptovarade hindade analüüsiks on (Akgül et al., 2022):

- 1) Tehniline analüüs;
- 2) Fundamentaalne analüüs;

Tehniline analüüs on meetod krüptovara analüüsimiseks, uurides mineviku turuandmeid, peamiselt hinda ja mahtu, et tuvastada signaalid, mida saab kasutada tulevaste hinnaliikumiste ennustamiseks (De Souza et al., 2018). Tehnilise analüüsi puhul kasutatakse sageli graafikuid ja diagramme (Akgül et al., 2022).

Tehniliseks analüüsiks võib kasutada liikuvaid keskmisi, et tuvastada suuna ja trendi tugevust või kasutada oskilaatoreid/indikaatoreid, et tuvastada üle ostetud või üle müüdud tingimusi turul. Tehniline analüüs saab olla kasutatud turul potentsiaalsete sisenemispunktide ja väljumispunktide tuvastamiseks oma strateegiate optimeerimiseks ja kasumi tegemiseks. (De Souza et al., 2018)

Fundamentaalne analüüs hõlmab mitmete aluslike tegurite uurimist ja analüüsimist, nagu projekti kasutuselevõtu määr, arengu progress, kasutusjuhud, konkurents ja regulatiivne keskkond. Fundamentaalne analüüs keskendub rohkem turgude majanduslike ja finantsaspektide kohta. (De Souza et al., 2018)

Sentimendi analüüs on inimeste kommentaaride või teiste andmebaaside kogumine ning sentimendi hinnangu andmine. Sentimendi analüüsi eesmärk on klassifitseerida tekst vastavalt tekstis väljendatud meeleolule või mõtteviisile, mis võib olla positiivne, negatiivne või neutraalne. Analüüsi käigus peamine eesmärk on tuvastada loomuliku keele teksti polaarsust ehk emotsioonide ja omaduste analüüs. (Chalothom & Ellman, 2015) Sentimendi analüüs võib olla kasulik nii fundamentaalse kui ka tehnilise analüüsi kontekstis. Samuti sentimendi kasutamine läks kasuks krüptovarade mudeli püstitamiseks kuna inimesed teevad otsuseid emotsioonide järgi, mitte loogika põhjal. (Frohmanni et al., 2023)

Kuna sotsiaalmeedia populaarsus kasvab ja inimesed teevad palju postitusi ning sisaldavad emotsionaalset aspekti oma otsuste langetamises selle bakalaaursetöös kasutatakse just sentimentide analüüsi, et proovida krüptovarade hindu kirjeldada.

1.3 Varasemad uuringud

Autor annab ülevaate varasematest uuringutest kus analüüsitakse sentimentide mõju krüptovaradele või tehakse krüptovarade hindadele prognoosi kasutades sentimentide. Antakse ülevaate uuringute tulemustest ning metoodikast.

Frohmanni et al. (2023) uurimistöö eesmärgiks oli püüda kinni sentimentide hinnangud, mis olid saadud SA tehnikatest. SA võib kategoriseerida masinõppeks, leksikonipõhisteks meetoditeks ja reeglipõhisteks tehnikateks andmete hankimiseks. Tema uuring oli kombinatsioon aegreade prognoosimisest ja sentimentide analüüsi kasutamisest. Metoodikasse kuulus nii BTC ajalooline hind kui ka ajaliselt joondatud Twitteri kommentaarid, millest sentimentide eraldati. Selle uurimistöö tulemused näitasid, et lineaarne regressioon kasutamine andis parem tulemusi kui keerukamate mudelite kasutamine. Samuti oli saadud, et VADERi ja BERT sentimentide skoorid võivad prognoosimise täpsust suurendada. Teised uuringud näitasid ka potentsiaali VADERi kasutamises nagu Pano & Kashef (2020) uuring VADERi sentimentide analüüsist Bitcoin hinnale. Selles uuringus ka oli nähtav korrelatsioon sentimentide ja Bitcoin vahel.

Critien et al. (2022) uurimustöö eesmärgiks oli ennustada nii Bitcoin hinna muutuste suunda kui ka suurust kasutades Twitteri sentimentide ja andmevoogu koguse. Tema töö metoodika hõlmas Twitteri andmekogumi eeltötluse ülesandeid Twitteri kommentaaridega ja inimeste sentimentide hinnangute määramist kasutades VADER-i. Tema tulemused olid optimistlikumad ja näitasid, et Bitcoin hinnamuutuste suunda ja suurust on võimalik suhteliselt täpselt ennustada, mille raporteeritud täpsus oli 63%.

Ibekwe (2021) uurimustöö keskendub krüptovara turu liikumiste ennustamisele, kasutades Twitteri ja Redditi andmete sentimentide analüüsi. Uuringu eesmärgiks leida sobiv mudel turusuuna ennustamiseks. Antud uurimuses kasutatakse VADER-i analüüsi eeltöödeldud sotsiaalmeedia postituste jaoks. Uuring näitas lineaarse regressioonimudeli kasutamises tulemusi, mis osutasid seose olemasolule sentimentide ja krüptovara hindade vahel.

Uras et al. (2020) uurimustöös ei kasutata sentimentide analüüsi. Metoodika hõlmas masinõppe ja statistiliste tehnikate ja meetodite testimist erinevatel andmebaasidel. Parimad tulemused saavutati nii regressioonimudelite kui ka LSTM ehk pika lühiajalise mälu võrgu kasutamisel. Selgus, et lineaarseid regressioone mudeleid on samuti võimalik kasutada. Uurimistöö vihjas sellele, et Bitcoin ja teised krüptovarad on endiselt varases arengufaasis ning tulevased uurimused võivad anda erinevaid tulemusi.

Sentimendi analüüsi kasutamine muutub populaarsemaks ning seda kasutatakse ka tavalisel turul aktsiate hindade prognoosimiseks. Sentimendi hinnangu saamiseks samuti kasutatakse VADERi ja varasema uurimustöös Xiao & Ihnaini (2023) kasutas sentimentide analüüsi aktsiate turgude trendi prognoosimiseks. Hutto & Gilbert (2014) uuringus oli tuvastanud, et VADER näitas teiste võrdlusnäitajatega paremaid tulemusi sentimentide analüüsi üldistamisel erinevates kontekstides.

2. ANDMED JA METOODIKA

Selles peatükiks kirjeldatakse ja selgitatakse kasutatud andmeid, nende töötlemisprotsessi ning tutvustatakse ka metoodikat. Töö empiirilises osas kasutatakse kvantitatiivset analüüsimeetodit, mis põhineb lihtsatel lineaarsetel regressioonimudelitel (OLS). Analüüsiks ja mudelite aktsepteerimiseks on kasutatud olulise nivoo $\alpha=0,05$.

2.1. Valim ja andmete kirjeldus

Autor kasutab andmetena aegridu perioodil 17 august 2017 kuni 21 jaanuar 2019 alla laetud veebi lehest Kaggle.com. Aegrea pikkus on sellest tulenevalt 523 päeva. Valiti see periood kuna teistel perioodidel olid puuduvad andmed. Andmetena olid alla laetud negatiivsed, positiivsed ja neutraalsed Badiola (2019) kogutud 17,7 miljonit Twitteri kommentaari hinnangud. Olid saadud iga tunni kohta andmed, näidis välja toodud (Tabel 1). Selleks, et koguda kõiki kommentaare, andmebaasi autor kasutas programmeerimiskoodi mis liigitas kommentaare, mis olid krüptovaraga seotud ja mis ei olnud. Twitteri kommentaarid, mis ei olnud krüptovaraga seotud olid eemaldatud. Andmebaasi autor kasutas VaderSentiment'i raamatukogu sentimendi analüüsiks, mis näitas head tulemust Hutto & Gilbert (2014), Pano & Kashef (2020), Frohmanni et al. (2023), Critien et al. (2022) ja Ibekwe (2021) uurimistöodes. Ta lisas umbes 30 väljendit ja sõna sõnastikku, mida pärast kasutati Twitteri kommentaaride jagamiseks (Badiola, 2019).

Tabel 1. Sentimendi hinnangud viie tunni kohta perioodil 1 august 2017 kuni 21 jaanuar 2019. Kokku 17,7 miljonit

Date	TotalS	Neg	P	Neu
01.08.2017 00:00	1 027	139	347	397
01.08.2017 01:00	778	111	193	320
01.08.2017 02:00	836	89	273	264
01.08.2017 03:00	984	250	236	314
01.08.2017 04:00	751	114	195	305

Allikas: Badiola, 2019 Kaggle.com (vt elektrooniline lisa)

Badiola (2019) ei lisanud, mis kujul kogutakse kommentaare. Andmete täpsemaks kirjeldamiseks olid võetud teised Twitterist kogutud kommentaarid eeskujuks (Tabel 2). Pärast kogutud kommentaaride juurde lisandub kuupäeva ja sentimendi hinnang mis on tehtud VADER-iga või teise analüüsi tarkvaraga.

Tabel 2. Kogutud Twitterist kommentaaride kuju näide

Id	Text
1	The recent negative news concerning the Bitcoin Lightning Network has brought to light several security issues, including undisclosed vulnerabilities, a significant security flaw, and replacement cycling attacks that exploit payment channel inconsistencies.
2	Bitcoin is the ultimate subject to see someone's character. Especially for the people with uneducated negative opinions. They are easy to spot. Big ego. No proof of work. Exactly that what we need to fix.
3	spot Bitcoin will go negative
4	Bitcoin is going to zero and then it will print negative. Whatever you think, it will be worse.
5	I don't think bitcoin is going to zero or negative bitcoin. It will be taken private instead by the Fed. Wifey of Crypto

Allikas: Alla laetud Kaggle.com (vt elektrooniline lisa)

Iga tunni kohta oli saadav Twitteri sentimendi järgi jagatud positiivne, neutraalne ja negatiivne kommentaaride arv (Tabel 1). Päevase sentimendi hinnangu leidmiseks lõputöö autor kasutas Exceli, kus oli valitud sentimendi positiivseks hindeks hinnang 1, neutraalseks hindeks hinnang 0,5 ja negatiivseks hindeks hinnang 0. Oli leitud keskmine positiivsete, negatiivsete ja neutraalsete kommentaaride arv. Pärast oli sentimendi hinnangu leidmiseks kasutatud järgmine valem:

$$\text{SENT} = (\text{P} + 0.5 \times \text{Neu}) / \text{TotalS} \quad (1)$$

kus,

SENT – päeva sentimendi hinnang

P – positiivsete sentimentidega kommentaaride arv päevas

Neu – neutraalsete sentimentidega kommentaaride arv päevas

TotalS – kommentaaride arv päevas

Järgmine samm oli alla laadida kolme krüptovara turuandmed: Bitcoin, Ethereum ja Ripple dollarites. Krüptovarade turu päevane hinnainfo alates 17. augustist 2017 kuni 21. jaanuarini 2019 alla laaditi kasutades YahooFinance'i. Samuti olid alla laetud BTCE ehk Bitcoin vahetusmaht miljardites dollarites, GOLD ehk kulla hinnad dollarites, MSCI ehk MSCI World Index tuhandetes

dollarites, Wiki ehk mitu korda on Wikipeediast olnud otsitud sõna Bitcoin tuhandetes, OIL ehk nafta hinnad, Tran ehk tehingute maht Bitcoiniga tuhandetes dollarites, BTCS ehk Bitcoin'i pakkumine miljonites, HASH ehk hash rate või kui kiiresti arvuti saab teostada krüptovaraga seotud (hash per second) arvutusi krüptovara teenimiseks miljonites hash per second ning VIX ehk volatiilsusindeks dollarites.

Näitaja VIX võib olla hea valik selle ajaperioodil, kuna see näitas VIX korrelatsiooni teiste krüptovaradega 22 juunist 2017 kuni 23 juuni 2018 (Akyıldırım et al., 2020). HASH võib anda ka tulemusi mudeli püstitamises, sest kõrgem Bitcoin'i hind innustab inimesi seda krüptovara rohkem teenima või mainima (Krištoufek, 2015). Andmetega nagu OIL, GOLD, MSCI ja VIX olid tehtud muudatused seoses sellega, et börs ei tööta nädalavahetustel ning puudusid nädalavahetuse andmed. Nende täitmiseks kasutati eelmise päeva hinna ja järgmise päeva hinna keskmist väärtust. Nende hindade lisamine ei pea uuringu tulemuse oluliselt mõjutada. Krüptovarade hindade andmetega ei tehtud muudatusi.

Järgnevas tabelis (Tabel 3) autor toob välja kasutatud aegriidade kirjendavat statistikat perioodil 17. augustist 2017 kuni 21. jaanuarini 2019. Kirjeldava statistika andmete hulka kuuluvad miinimum, maksimum, keskmine väärtus, standardhälve ja variatsioonikordaja.

Bitcoin'i ja Ethereum'i variatsioonikordajad näitavad mõõdukat varieeruvust, kuid XRP puhul on varieeruvus suurem. SENT, GOLD, MSCI, BTCS, OIL ja Tran näitajatel on varieeruvus väike. Wiki näitajal varieeruvus on suurem kui üks ehk see tähendab, et standardhälve on suurem kui keskmine väärtus. VIX, HASH ja BTCE on varieeruvus suur. Näitajatel nagu BTCS, HASH, MSCI, WIKI, Tran ja BTCE arved on oluliselt suurem kui teistel näitajatel. BTCS oli võetud miljonites krüptovara ühikutes. HASH puhul seda mõõdetakse h/s ehk Hash per second ning oli võetud miljonites. BTCE puhul on kogu summa dollarites mida vahetatakse ning see summa peab olema suur. BTCE oli võetud miljardites dollarites. MSCI ja WIKI oli võetud tuhandes dollarites ja Tran tuhandes ühikutes. Suur kogus näitajaid on valitud kuna suuremad krüptovarad nagu Bitcoin ja Ethereum ilmuvad prognoosi keerukust, volatiilsust ja käitumise ettearvamatus. (Stošić et al., 2019)

Tabel 3. Aegridade kirjeldav statistika

	Min	Max	Keskmine	Mediaan	Standard h�lve	Variatsiooni kordaja
BTC (USD)	3154,9	19497	7465,6	6710,6	3171,6	0,42482
XRP (USD)	0,15531	3,3595	0,57690	0,45377	0,48308	0,83737
ETH (USD)	84,308	1396,4	445,82	370,67	273,21	0,61283
SENT (�hikud)	0,36242	0,56554	0,49508	0,49575	0,025964	0,052444
Gold (USD)	1182,30	1508,00	1276,30	1282,20	49,56	0,03883
MSCI (tuh USD)	1,776	2,25	2,068	2,09	0,0952	0,0452
Wiki (tuh �hikud)	6,521	344,69	32,811	16,792	41,53	1,2657
VIX (USA)	9,01	37,32	15,32	13,93	5,05	0,329760
BTCS (milj �hikud)	16,51	17,49	17,006	17,011	0,287	0,016848
OIL (USD)	50,37	86,07	66,83	67,16	8,90	0,13318
Tran (tuh �hikud)	134,97	490,46	245,1	235,73	59,182	0,24146
HASH (milj h/s)	4,085	61,87	29,59	30,346	15,39	0,52017
BTCE (miljard USD)	0,781	19,61	4,611	3,848	3,063	0,66431

Allikas: Autori koostatud Gretl  konomeetriapaketis

J rgnevas tabelis (Tabel 4) on toodud korrelatsiooni maatriks. Korrelatsiooni maatriksi tulemusena on n ha kr ptovarade vahel k rget korrelatsiooni. See v ib viidata sellele, et  he kr ptovara muutused v ivad p hjustada teiste kr ptovarade hindade muutusi. Kuid selleks tuleb seda edasi uurida. Γκίλλας et al. (2018) m rkis oma uuringus, et kr ptoturul on suur korrelatsioon. Teiselt poolt, korrelatsioon SENT ja BTC vahel on 0,254, samas kui korrelatsioon SENT ja XRP ning ETH vahel on suhteliselt madalam. See v ib n idata, et Twitteri kommentaarid on peamiselt seotud BTC-ga v i et kr ptovara sentiment m jutab BTC-d rohkem kui teisi kr ptovarasid.

Tabel 4. Aegridade korrelatsioon

BTC	XRP	ETH	SENT	
1	0,6883	0,7715	0,2540	BTC
	1	0,7862	0,0792	XRP
		1	0,0662	ETH
			1	SENT

Allikas: Autori koostatud Gretl  konomeetriapaketis

2.2. Metoodika

Selles t os regressiooni parameetrite hindamiseks kasutatakse v himruutude meetod lineaarsed mudelit: harilik v himruutude meetod OLS (Ordinary Least Squares). V himruutude meetodi

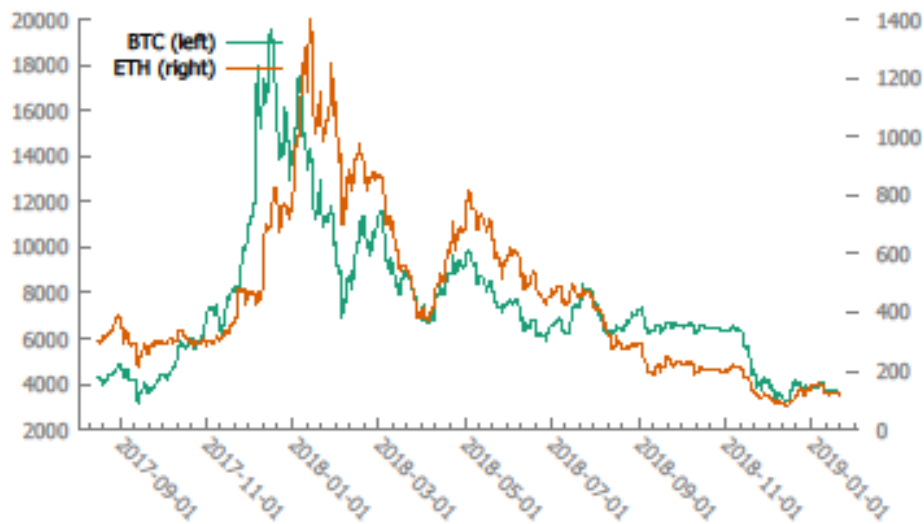
põhjal regressioonimudeli parameetrite hinnangud leitakse nii, et jääkide ruutude summa oleks minimaalne. Lineaarne regressioonimudel on tundlik erinditele ehk üksikutele teistest oluliselt erinevatele väärtustele. (Sauga, 2017, 422-424)

OLS mudelite tulemused võivad oluliselt erineda, sõltuvalt konkreetsetest spetsifikatsioonidest ja sellest, kas mudeli eeldused on täidetud või mitte. Neid eeldusi tuleb hoolikalt kontrollida ja käsitleda täpsema tulemuse saamiseks. Esiteks, aegread peavad olema statsionaarsed, sest mitte statsionaarsete aegrida kasutades võib tulemuseks olla näiv regressioon. Juhuslikud liikmed peavad alluma normaaljaotusele. Normaaljaotuse mitte täitumise korral pole hinnangud mõjusad. Regressisid peavad samuti olema eksogeensed, homoskedastiivsus ning autokorrelatsioon peavad puuduma, et vältida valed standardhälved. Lisaks, mudel peab olema lineaarne parameetrite suhtes, sama sagedusega, regressorite väärtused peavad hajuma ning regressorid ei tohi olla lineaarselt sõltuvad. (Gujarati, 2003, 62-70)

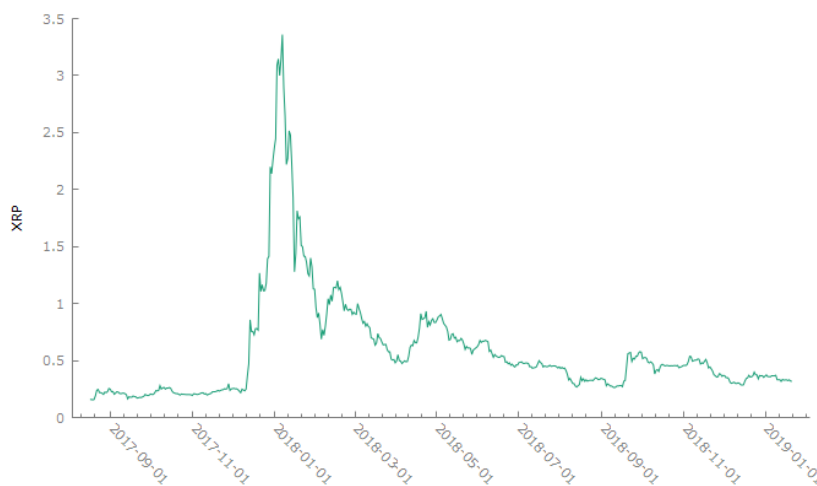
Valitud aegridade baasil koostatakse mitmeid OLS mudeleid ning kontrollitakse OLS mudelite eeldusi. Nende eelduste kontrollimiseks kasutatakse erinevaid teste mis aitavad mudeli paranemisega ning mudeli eelduste täitmisega. Kõigi tarkvarade püstitatud mudelite ja testide tulemuste toorväljavõted on lisatud elektroonilisse lissasse.

Mudeli koostamisel on oluline eeldus, et aegread oleksid statsionaarsed, mis tähendab et nad ei sisalda pikaajalisi trende ja kõikuvad pigem keskmise taseme ümber (Sauga, 2017, 582). Kui mudeli koostamisel kasutatakse mittestatsionaarseid aegridu, ei saa tunnuste seoseid ega mudeli tulemusi korrektselt tõlgendada. Näiteks kui aegread omavad trendi või sesoonsust, ei pruugi nende ajaline liikumine olla omavahel seotud, mis võib põhjustada mudelis multikollineaarsuse tekkimist ja varjutada tunnuste tegelikke seoseid. Seetõttu on oluline tagada, et mudeli ehitamisel arvestatakse aegridade statsionaarsusega, et tagada modelleerimise usaldusväärsus. (Gujarati, 2003, 741)

Vaadates Joonist 1 ja Joonist 2 on näha, et aegridadel BTC, ETH ja XRP puudub statsionaarsus. Statsionaarsuse kontrollimiseks oli teostatud Dickey-Fuller test, mille nullhüpoteesiks on ühikjuure olemasolu ja sisukaks hüpoteesiks on ühikjuure puudumine.



Joonis 1. BTC ja ETH hinnad igapäevaselt
 Allikas: Autori koostatud Gretlis



Joonis 2. XRP hinnad igapäevaselt
 Allikas: Autori koostatud Gretlis

Pärast Dickey-Fuller testi kasutamist kõiki muutujatele, järeldas autor, et aegread BTC, ETH ja XRP on mittestatsionaarsed. Statsionaarsuse saavutamiseks kasutas autor 1. järgu diferentsi ning logaritmi väärtused juhendaja soovitusel (Lisa 1). Tulemuseks olid saadud päevased krüptovara tootlused. Tunnuste diferentseerimise ja logaritmi kasutamise järgi olid saadud uued aegread tähistatud d_1_BTC , d_1_ETH , d_1_XRP . Lisaks olid Dickey-Fuller testiga kontrollitud teised turu näitajad ning olid leitud mittestatsionaarsus. Nende aegreade statsionaarsuse saavutamiseks oli kasutatud 1. järgu diferentsid (Lisa 1). Uued statsionaarsed tunnused olid

d_BTCE, d_Wiki, d_BTCS, d_OIL ja d_Tran. Teised näitajad pärast testimist näitasid statsionaarsust.

Lineaarse regressioonimudeli jaoks identifitseeritakse statistiliselt olulised tunnused. Oluliste tunnuste tuvastamise järel viiakse läbi mudeli testimine nelja standardse testiga, mis põhinevad klassikalise lineaarse regressioonimudeli eeldustel. Need testid hõlmavad heteroskedastiivsuse ja autokorrelatsiooni puudumise hindamist, kasutades vastavalt White'i ja Breusch-Godfrey teste, multikollineaarsuse hindamist VIF testi abil ning jääkide normaaljaotuse alluvuse kontrollimist Doorik-Hanseni testi abil. Lineaarse regressioonimudeli standardkuju korral on (Gujarati, 2003, 192):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (2)$$

kus

Y – sõltuv tunnus

β – mudeli parameeter

X – argumenttunnus

ε - mudeli juhuslik viga

Sõltuvateks muutujateks pikaajaliste ja lühiajaliste mudelite puhul valiti krüptovara hinnad: BTC, ETH ja XRP. Sõltumatu muutujana on valitud sentiment hinnang ehk SENT. Mudelis kuhu lisatakse sentiment hinnangu sentiment mõju kontrollimiseks, on sõltuvaks muutujaks valitud Bitcoin hind (BTC). Mudelis kuhu lisatakse sentiment sõltumatuteks muutujateks on valitud BTCE, GOLD, MSCI, Wiki, OIL, Tran, BTCS, HASH ning VIX. Mudelite hindamisel kasutatakse determinatsioonikordajat ehk R ruut, mis iseloomustab mudeli kirjeldusvõimet ning näitab regressioonihajuvuse osakaalu koguhajuvust (Sauga, 2017, 435).

3. ANALÜÜS JA TULEMUSED

Töö järgnevas osas koostatakse seitse OLS mudelit. Kolm pikaajalist mudelit on koostatud kasutades sõltuva muutujana d_1_BTC , d_1_ETH ja d_1_XRP . Sõltumatu muutujana on võetud SENT ehk inimeste sentimendi hinnang. Aegrea perioodiks on võetud 17. augustist 2017 kuni 21. jaanuarini 2019. Vaatluste arv on 523 päeva.

Pärast pikaajaliste mudelite püstitamist on koostatud kolm lühiajalist OLS mudelit. Kolm mudelit on koostatud kasutades sõltuva muutujana d_1_BTC . Sõltumatu muutujana on võetud SENT ehk inimeste sentimendi hinnang. Selles mudelis proovitakse leida seos juba lühiajalise aegrea perioodi jooksul ning tehakse juhuslik valik keskmiselt 90 päevase kestvusega aegrea kogu perioodil 17. augustist 2017 kuni 21. jaanuarini 2019.

Seitsmes mudel on erinevate sõltumatu tunnuste mudel. Mudel on püstitatud sõltuva muutujaga d_1_BTC ja sõltumatu muutujatega BTCE, GOLD, MSCI, Wiki, OIL, Tran, BTCS, HASH ja VIX. Mudel on analüüsitud ning pärast on mudelisse lisatud lõputöö autori SENT ja mudel on veel kord hinnatud ja analüüsitud sentimendi mõju tuvastamiseks.

Mudelite koostamise aluseks on statistiliselt olulised sõltumatud tunnuste leidmine, OLS mudeli eelduste täitmine ning mudeli parandamine. Esitatakse esialgsed mudelid ning pärast lõplikud mudelid. Kõik läbiviidud testide tulemused on esitatud käesoleva töö lisades ja tarkvara toorväljavõtted elektroonilises lisas. Lõplikud järeldused on tehtud lühiajaliste, pikaajaliste ning sentimendi lisatud mudelisse järgi. Järelduste osas on toodud uuringu tulemused, piirangud ja võimalikud parandused.

3.1. Pikaajaliste mudelite analüüs

Esimeses mudelis testitakse seost BTC ja SENT vahel. Proovitakse kirjeldada Bitcoinini hinna ainult kasutades sentimendi. Vaadatakse, kas ainult sentimendi kasutades on võimalik saada kirjeldava mudeli ning mis krüptovara valida teiste mudelite tegemiseks. Sõltuvaks muutujaks on võetud

BTC ehk Bitcoin hinnad ja sõltumatu muutujaks on võetud SENT ehk inimeste sentimendi hinnang. Kuid OLS mudeli eelduseks on muutujate statsionaarsus siis võetakse mudelise sõltuvaks muutujaks d_1_BTC ehk Bitcoin tootlus ning SENT ehk inimeste sentimendi hinnang. Lisas 1 on kirjeldatud Dickey-Fuller testid.

Pärast Bitcoin logaritmeerimist ja diferentseerimist vaatluste arv vähenes 1 võrra ning nüüd on vaatluste arv $t = 522$ ja perioodiks on võetud 18. augustist 2017 kuni 21. jaanuarini 2019. Sama asi kehtib nii ETH kui XRP mudelite kohta. Bitcoin mudeli kuju on järgmine:

$$d_1_BTC = -0,166 + 0,335SENT + \epsilon \quad (3)$$

(0,039) (0,078)

Determinatsioonikordaja $R^2 = 0,0324$

Valimi maht $t = 522$

Esimene mudel ei vaja korrigeerimist, sest ei esine mitteolulisi tunnuseid, mis mõjutavad mudeli väärtusi. Mudeli märgid on autori jaoks loogilised, sest SENT kasvades peab BTC tootlus ka kasvama. Järgmise sammuna kontrollib autor mudeli heteroskedastiivsuse White testi abil, autokorrelatsiooni puudumist Breusch-Goldfrey testiga ja jääkliikmete alluvust normaaljaotusele Doornik-Hanseni testi abil (Lisa 2). Multikollineaarsuse testimine pole vajalik, sest mudelis on ainult üks sõltumatu muutuja.

White testi abil saadi tulemusena Test-statistiku olulisuse tõenäosuse 0,962. Nullhüpoteesiks on heteroskedastiivsuse puudumine ning lähtudes olulisuse nivoole, on võimalik öelda, et mudelis ei esine heteroskedastiivsust ehk võetakse vastu nullhüpotees. Breusch-Goldfrey testi abil saadi tulemusena Test-statistiku LMF olulise tõenäosuse 0,294. Nullhüpoteesiks on autokorrelatsiooni puudumine ning lähtudes olulisuse nivoole on võimalik öelda, et mudel ei esine autokorrelatsiooni ehk võetakse vastu nullhüpotees. Doornik-Hanseni testi abil saadi tulemusena 1.593×10^{-20} . Nullhüpoteesiks on, et jääkliikmed alluvad normaaljaotusele ning lähtudes olulisuse nivoole on võimalik öelda, et jäägid ei allu normaaljaotusele ehk võetakse vastu sisukas hüpotees.

Lähtudes testide tulemustest mudelis puudub autokorrelatsioon ja heteroskedastiivsus, kuid jäägid ei allu normaaljaotusele. Suure valimi korral, ehk rohkem kui sada, ei tekita jääkide jaotuse kõrvalekaldumine normaaljaotusest suurt probleemi (Sauga, 2017).

Teises mudelis testitakse seos ETH ja SENT vahel. Kuid OLS mudeli eelduseks on muutujate statsioonarsus siis võetakse mudelisse sõltuvaks muutujaks d_1_ETH ehk Ethereumi tootlus ning sõltumatu muutujaks SENT ehk inimeste sentimendi hinnang. Ethereumi mudeli kuju on järgmine:

$$d_1_ETH = -0,135 + 0,268SENT + \varepsilon \quad (4)$$

(0,048) (0,096)

Determinatsioonikordaja $R^2 = 0,0128$

Valimi maht $t = 522$

Mudelite sarnasuste tõttu ehk nii BTC kui ka ETH mudelis on üks sõltuv ja üks sõltumatu muutuja, teostatakse samasugused eelduste kontrollimise testid nagu BTC mudeliga (Lisa 2).

Teises mudelis puuduvad mitteolulised tunnused, mis võivad mudelit mõjutada. Mudeli märgid on autori jaoks loogilised, sest SENT kasvades peab ETH tootlus ka kasvama. White testi abil saadi tulemusena Test-statistiku olulisuse tõenäosuseks 0,604 ehk võetakse vastu nullhüpotees, heteroskedastiivsust ei esine. Breusch-Goldfrey testi abil saadi tulemusena Test-statistiku LMF olulise tõenäosuse 0,055 ehk võetakse vastu nullhüpotees, autokorrelatsiooni ei esine. Doorik-Hanseni testi abil saadi tulemusena 5.678×10^{-14} ehk võetakse vastu sisukas hüpotees, jäägid ei allu normaaljaotusele. Lähtudes testide tulemustest mudelis puudub autokorrelatsioon ja heteroskedastiivsus kuid jäägid ei allu normaaljaotusele. Suure valimi korral, ehk rohkem kui sada, ei tekita jääkide jaotuse kõrvalekaldumine normaaljaotusest probleeme (Sauga, 2017).

Kolmas mudelis testitakse seost XRP ja SENT vahel. Kuna OLS mudeli eelduseks on muutujate statsioonarsus, võetakse mudelisse sõltuvaks muutujaks d_1_XRP ehk Ripple'i tootlus ning SENT ehk inimeste sentimendi hinnang. Esialgse Ripple mudeli kuju on järgmine:

$$d_1_XRP = -0,160 + 0,326SENT + \varepsilon \quad (5)$$

(0,065) (0,131)

Determinatsioonikordaja $R^2 = 0,0099$

Valimi maht $t = 522$

Mudelite sarnasuste tõttu ehk nii BTC kui ka XRP mudelis on üks sõltuv ja üks sõltumatu muutuja, teostatakse samasugused eelduste kontrollimise testid nagu BTC mudeliga (Lisa 2).

Teises mudelis puuduvad mitteolulised tunnused mis võivad mudelit mõjutada. Mudeli märgid on autori jaoks loogilised, sest SENT kasvades peab XRP tootlus ka kasvama. White testi abil saadi tulemusena Test-statistiku olulisuse tõenäosuse 0,364 ehk võetakse vastu nullhüpootees, heteroskedastiivsust ei esine. Breusch-Goldfrey testi abil saadi tulemusena Test-statistiku LMF olulise tõenäosuse 0,006 ehk võetakse vastu sisukas hüpootees, autokorrelatsioon esineb. Doorik-Hanseni testi abil saadi tulemusena $1,807 \times 10^{-50}$ ehk võetakse vastu sisukas hüpootees, jäägid ei allu normaaljaotusele.

Pärast eelduste kontrollimist oli heteroskedastiivsuse puudumine, kuid jäägid ei allunud normaaljaotusele ning esines autokorrelatsioon. Mitte alluvuse normaaljaotusele ei tekita suurt probleemi aga autokorrelatsiooni olemasolu võib anda mudeli valesid tulemusi ehk valesid standardvead. Kuid pole soovitatav kasutada kohandatud standardvigu kui valimimaht on suurem kui 50, aga proovitakse neid kasutada autokorrelatsiooni arvestamisega. (Sauga, 2017)

Lõpliku mudeli kuju on järgmine:

$$d_1_XRP = -0,160 + 0,326SENT + \varepsilon \quad (6)$$

(0,069) (0,142)

Determinatsioonikordaja $R^2 = 0,0099$

Valimi maht $t = 522$

Pärast OLS mudeli eelduste testimist olid saadud sama tulemused (Lisa 2). Autokorrelatsiooniga arvestada ei õnnestunud.

3.2. Lühiajaliste mudelite analüüs

Pärast seose testimist krüptovarade tootluste ja sentimendi vahel pikaajaliselt, proovitakse leida seost lühiajaliselt. Lühiajalise ajaperioodi uurimine on tehtud determinatsiooni kordaja varieeruvuse leidmiseks ning sentimendi mõju hindamiseks. Pikkajaliseks perioodiks oli võetud 522 päeva kuid nüüd lühiajalise seose testimiseks valiti kolm kuud ehk 89-90 päeva. BTC

kasutatakse lühiajalise seose leidmiseks. Valikusse läks BTC kuna pikaajalistel mudelites BTC-ga seotud mudelil oli kõike suurem R2 ehk 0,0324. Igas allolevas mudelis võetakse sõltuvaks muutujaks krüptovara tootluse ehk d_1_BTC ja sõltumatuks muutjaks SENT ehk sentimendi hinnangu.

Võimalike valede tulemuste vältimiseks testitakse juhuslikult 90 päevase ajaperioodi erinevatel aegrida hetkedel. Kahe mudeli 90 päeva periood on võetud 2018 aastal ja kolmanda mudeli periood 2017 aastal juhuslikult valitud ajal.

Esimeses lühiajalises mudelis testitakse seost BTC ja SENT vahel. Kuid OLS mudeli eelduseks on muutujate statsionaarsus, siis võetakse mudelise sõltuvaks muutujaks d_1_BTC ehk BTC tootlus ning SENT ehk inimeste sentimendi hinnang. Esimese lühiajalise mudeli aegrea perioodiks on võetud 1. jaanuar 2018 kuni 31. märts 2018 ehk 90 päeva. Esimese mudeli kuju on järgmine:

$$d_1_BTC = -0,465 + 0,920SENT + \epsilon \quad (7)$$

(0,171) (0,344)

Determinatsioonikordaja $R^2 = 0,065$

Valimi maht $t = 90$

Esimese lühiajalises mudelis puuduvad mitteolulised tunnused, mis võivad mudelit mõjutada. Mudeli märgid on autori jaoks loogilised, sest SENT kasvades peab BTC tootlus ka kasvama. White testi abil saadi tulemusena Test-statistiku olulisuse tõenäosuse 0,517 ehk võetakse vastu nullhüpotees, heteroskedastiivsust ei esine. Breusch-Goldfrey testi abil saadi tulemusena Test-statistiku LMF olulise tõenäosuse 0,539 ehk võetakse vastu nullhüpotees, autokorrelatsiooni ei esine. Doorik-Hanseni testi abil saadi tulemusena 0,265 ehk võetakse vastu nullhüpotees, jäägid alluvad normaaljaotusele (Lisa 3). OLS mudeli eeldused on täidetud.

Teise lühiajalise mudeli aegrea perioodiks on võetud 18. august 2017 kuni 14. november 2017 ehk 89 päeva. Testitakse mudeli eeldused (Lisa 3) ning heteroskedastiivsust ei esine, tõenäosus 0,145. Autokorrelatsiooni ei esine, tõenäosus oli 0,642. Doorik-Hanseni testi abil saadi tulemusena $7,352 \times 10^{-8}$ ehk võetakse vastu sisukas hüpotees, et jäägid ei allu normaaljaotusele. Teise lühiajalise mudeli kuju on järgmine:

$$d_1_BTC = -0,245 + 0,512SENT + \varepsilon \quad (8)$$

(0,096) (0,196)

Determinatsioonikordaja $R^2 = 0,062$

Valimi maht $t = 89$

Kolmanda lühiajalise mudeli aegrea perioodiks on võetud 1. juuni 2018 kuni 29. august 2018 ehk 90 päeva. Mudelis ei esine statistiliselt olulisi näitajaid. Mudelit ei saa kasutada (vt elektrooniline lisa).

3.3. Sentimendi lisamine mudelisse

Viimase mudeli järgi proovitakse välja selgitada, kas sentimendi kasutamine aitab paremini krüptovarade hindu prognoosida. Khamisa (2019) juba proovis teha OLS mudelit oma uuringus. Tema lõpliku mudeli R^2 oli 0,335, kuid tema mudelis ei esinenud inimeste sentimendi ning olid võetud mudelisse teised krüptovarad, mis võivad oma vahel korreleeruda (Γκίλλας et al., 2018). Statistiliselt ebaolulisi näitajaid oli palju. Tema mudel ei ole sobilik kasutamiseks, kuid olid võetud tema mudeli näitajad ja lisatakse autori poolt näitajad HASH mudelisse. Viimase mudeli püstitamiseks analüüsitakse mudelit ning pärast statistiliselt olulise mudeli saamist lisab autor inimeste sentimendi mudelisse. Pärast SENT mudelisse lisamist võrreldakse kaks mudelit.

Perioodiks on võetud 18. augustist 2017 kuni 21. jaanuarini 2019 ning vaatluste arv on 522. Sõltuvaks muutujaks on d_1_BTC ja sõltumatuteks muutujateks on võetud d_BTCE , $GOLD$, $MSCI$, d_Wiki , d_OIL , d_Tran , d_BTCS , VIX , $HASH$ ja $SENT$. Esialgses mudeli tulemuseks tuli determinatsioonikordaja 0,0207 ning olid palju statistilisi ebaolulisi muutujaid (vt elektrooniline lisa). Proovitake statistilised ebaolulised muutused eemaldada suurema p-väärtuse järgi. Pärast ebaoluliste tunnuste eemaldamist oli saadud mudel järgmisel kujul:

$$d_1_BTC = 0,0914 - 0,0444MSCI + 0,00049d_Wiki + 0,00173d_OIL + \varepsilon \quad (9)$$

(0,0442) (0,0213) (0,000153) (0,000912)

Determinatsioonikordaja $R^2 = 0,0268$

Valimi maht $t = 522$

Selles mudelis on statistiliselt mitteoluline tunnus d_OIL . Selle tunnuse p-value on 0,059 aga otsustati lisada sellesse mudelisse SENT ja hinnata uut mudelit ning vaadata, kas näitaja muutub statistiliselt oluliseks. Pärast SENT lisamist oli saadud lõplik mudel järgmisel kujul:

$$d_1_BTC = -0,077MSCI + 0,000458d_Wiki + 0,321SENT + \varepsilon \quad (10)$$

(0,0143) (0,00015) (0,0598)

Determinatsioonikordaja $R^2 = 0,07$

Valimi maht $t = 522$

Mudel ei vaja korrigeerimist sest ei esine mitteolulisi tunnuseid, mis mõjutavad mudeli väärtusi. Mudelist oli eemaldatud d_OLD kuna see ei muutunud statistiliselt oluliseks. Uuringust İçellioglu & Öner (2019) oli saadud, et kuld, SPX 500 indeks ja nafta ehk OIL hinnad tõstavad krüptovarade hindu, kuid selles mudelis seda tõestada ei saadud.

Kontrollitakse mudeli heteroskedastiivsuse White testi abil, autokorrelatsiooni puudumist Breusch-Goldfrey testi abil, jääkliikmete alluvust normaaljaotusele Doornik-Hanseni testi abil, kas mudelil on õige kuju RESET testi abil (Lisa 4). Mudelis esinevad kolm muutujaid, millepärast kontrollitakse ka multikollineaarsuse esinemist VIF testiga (Lisa 5).

White testi abil oli saadud tulemus Test-statistiku olulisuse tõenäosus 0,822. Nullhüpoteesiks on heteroskedastiivsuse puudumine ning lähtudes olulisuse nivoole on võimalik öelda, et mudelis ei esine heteroskedastiivsust ehk võetakse vastu nullhüpotees. Breusch-Goldfrey testi abil saadi tulemusena Test-statistiku LMF olulise tõenäosuse 0,196. Nullhüpoteesiks on autokorrelatsiooni puudumine ning lähtudes olulisuse nivoole on võimalik öelda, et autokorrelatsiooni ei esine ehk võetakse vastu nullhüpotees. Doornik-Hanseni testi abil saadi tulemusena $1,496 \times 10^{-18}$. Nullhüpoteesiks on, et jääkliikmed alluvad normaaljaotusele ning lähtudes olulisuse nivoole on järeldatud, et jäägid ei allu normaaljaotusele ehk võetakse vastu sisukas hüpotees. RESET testi tulemusena oli saadud olulisuse tõenäosus 0,165. Nullhüpoteesiks on mudeli õige kuju ehk võetakse vastu nullhüpotees. VIF varieeruvus tulemusena oli saadud, et VIF indeksid on alla väärtuse 10 ning ei olnud identifitseeritud seda, et mudelis võiks esineda multikollineaarsus.

Lähtudes testide tulemustest mudeli kuju on õige, puudub autokorrelatsioon, heteroskedastiivsus, multikollineaarused, kuid jäägid ei allu normaaljaotusele.

3.4. Järeldused ja piirangud

Pärast kõigi mudelite tulemuste kontrollimist, esimene asi, mida tuvastati on, et mudelitel esineb madal R² või madal kirjeldav võimekus. Krüptovara hinna prognoosimine ja kirjeldamine on väga keeruline, nagu oli leitud Stošić et al. (2019) uuringus, seega tulemuseks on võimalik saada väikest R ruutu.

Esiteks näitas korrelatsioonimaatriks (Tabel 4) ligikaudu 0,3 suurust korrelatsiooni BTC ja SENT vahel. Kuid ainult korrelatsioonile tugineda ei saa. Olid püstitatud pikaajalised ning lühiajalised mudelid.

Pärast esimeste pikaajaliste mudelite püstitamist autor jõudis järeldusele, et inimeste sentiment ja krüptovarade hindade vahel suurt seost ei ole. Pärast pikaajalist Bitcoin mudeli püstitamist oli saadud determinatsioonikordaja 0,0324 (vt Tabel 5). Mudeli tunnused olid statsionaarsed, puudus heteroskedastiivsus, autokorrelatsioon ja jäägid ei allunud normaaljaotusele. Samal ajal oli mudel statistiliselt oluline. Pikaajalise Ethereumiga mudeli püstitamisega saadi determinatsioonikordaja 0,0128. Mudelis puudus autokorrelatsioon ning heteroskedastiivsus, kuid jäägid ei allunud normaaljaotusele. Mudel oli statistiliselt oluline. Pärast pikaajalist Ripple-i mudeli püstitamist saadi determinatsioonikordaja 0,0099. Puudus heteroskedastiivsus, kuid jäägid ei allunud normaaljaotusele ja esines autokorrelatsioon, mida eemaldada ei õnnestunud.

Mudelite järjestikku determinatsioonikorda vähenemised on loogilised sest Bitcoin näitab seoseid teiste krüptovaradega ja traditsiooniliste investeerimisvahenditega (Kartal, 2022). Sellepärast inimeste sentiment rohkem hõlmab Bitcoin hindu ja tootlusi võrreldes teiste krüptovaradega nagu Ethereum ja Ripple. Pikaajaliste mudelite loomine viitab sellele, et ainult sentiment ei saa kasutada krüptovarade hindade kirjeldamiseks või prognoosimiseks.

Tabel 5. Pikaajalised mudelid

Muutujad	d_1_BTC	d_1_ETH	d_1_XRP
Konstant	-0,166*** (0,039)	-0,135*** (0,048)	-0,160** (0,065)
SENT	0,335*** (0,078)	0,268*** (0,096)	0,326** (0,131)
Valimi maht	522	522	522
R2	0,0324	0,0128	0,0099

Allikas: Autori koostatud, vt elektrooniline lisa

Märkus: Statistiline olulisus *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Vastuseks krüptovarade ja inimeste sentimendi vahel seotud küsimustele näitavad pikaajalise ja lühiajalise 90-päevase analüüsi tulemused erinevaid tulemusi. R-ruut on suurenenud, kui hakati lühiajalist perioodi analüüsima (vt Tabel 6). Siiski võeti kolm juhuslikku perioodi, millest kaks näitasid kõrgemat R-ruutu võrreldes pikaajalise mudeliga ja üks näitas, et mudel pole statistiliselt oluline. Lühiajaliste mudelite jaoks kasutati ainult Bitcoinit. Lühiajaliste mudelite analüüs näitas, et ainult sentimendi kasutades ei saa lühiajalised Bitcoinit hinda kirjeldada või prognoosida. Kolmas püsitatud mudel oli statistiliselt mitteoluline (Tabel 6).

Tabel 6. Lühiajalised mudelid.

Muutujad	d_1_BTC	d_1_BTC	d_1_BTC
Konstant	-0,465*** (0,171)	-0,245** (0,096)	-0,08 (0,067)
SENT	0,92*** (0,344)	0,512** (0,196)	0,16 (0,134)
Valimi maht	90	89	90
R2	0,065	0,062	0,0049

Allikas: Autori koostatud, vt elektrooniline lisa

Märkus: Statistiline olulisus *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Viimaseks analüüsiks kasutati OLS mudelit ilma SENT ning pärast oli SENT mudelisse lisatud. Sentimendi mõju leidmiseks lõputöö autor lisas statistiliselt olulisse mudeli inimeste sentimendi. Oli saadud, et mudelis ilma SENT oli determinatsioonikordaja 0,0268 ning pärast sentimendi lisamist suurenes determinatsioonikordaja ja see sai 0,07 (vt Tabel 7). See näitab, et sentimendi lisamine tegi mudeli paremaks, kuid mitte suures ulatuses. Sentimendi lisamine näitas, et juba keeruliste mudelitega on võimalik sentimendi lisada, kuid ainult sentimendile lootmine pole mõistlik. Ennustusmeetodid ainult sentimendi kasutades ei toimi üldiselt nii hästi aga neid saab kasutada hinnaennustuste täiustamiseks, eriti keerukamate mudelite puhul (Frohmann et al., 2023).

Siiski madala R-ruudu tõttu neid mudeleid ei saa kasutada krüptovara hindade ennustamiseks ega kirjeldamiseks. Madala R-ruudu mudel vahemikus 0 kuni 0,09 pole sotsiaalteaduste uurimises aktsepteeritavad ning tuleks tagasi lükata. Mudel, mille R-ruut on vahemikus 0,10 kuni 0,50 on aktsepteeritav, kui mõned või enamik seletavatest muutujatest on statistiliselt olulised. (Ozili, 2022) Selles uuringus puuduvad mudelid R-ruuduga üle 0,1.

Tabel 7. Statistiliselt oluline mudel ilma ja koos sentimendiga

Muutujad	d_1_BTC	d_1_BTC
Konstant	0,0914** (0,0442)	
MSCI	-0,0444** (0,0213)	-0,077*** (0,0143)
d_WIKI	0,00049*** (0,000153)	0,00046*** (0,00015)
d_OIL	0,00173* (0,000913)	
SENT		0,32*** (0,0598)
Valimi maht	522	522
R2	0,0268	0,07

Allikas: Autori koostatud, vt elektrooniline lisa

Märkus: Statistiline olulisus *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Sellel uuringul on ka piirangud seotud kogutud andmete mitmetähenduslikkusega. On raske teha korralikku sentimentide analüüsi, kuna Internetis on palju boteid või roboteid, mis võivad samuti uuringu tulemusi mõjutada. Himelein-Wachowiak et al. (2021) näitas, et desinformatsiooni saab levitada nii inimene, kui ka botid. Twiiteris nad saavad teha liigset postitamist, mis võib mõjutada sellega sentimentide hinnangu. Krüptovarad võivad ka olla pettuslikult manipuleeritud mis võib ka tulemusi mõjutada (Peterson, 2020). Krüptovarade hindu võivad mõjutada mitmed tegurid või hinnangud, mis ei pruugi alati kajastada sotsiaalmeedia sentimentides (Ibekwe, 2021). Lisaks võeti analüüs ja andmed aastatest 2017 kuni 2019, mistõttu andmed võivad olla vananenud ja mitte tulevaste uuringute jaoks olulised. Mõned muutujad, mis olid ka leitud teistes uuringutes mõjukateks, polnud võetud mudelisse. Nendest muutujatest on rahapakkumine ja krüptovara teenimise raskus, nagu leiti Krištoufek (2015) uuringus.

KOKKUVÕTE

Käesoleva töö eesmärgis oli analüüsida sentimendi mõju erinevate krüptovarade hindadele. Sentimendi mõju analüüs oli tehtud pikaajalisel, lühiajalisel perioodil ning oli analüüsitud sentimendi lisamine mudelisse ilma krüptovara hinna sentimendi. Kolm regressioonimudelid olid tehtud pikaajaliseks analüüsiks Bitcoin, Ethereum ja Ripple-iga. Lühiajaliseks analüüsiks oli tehtud kolm mudelit Bitcoiniga. Sentimendi lisamisega oli tehtud üks mudel.

Vale tulemuse vältimiseks olid kontrollitud OLS regressioonimudeli eeldused nagu statsionaarsus, autokorrelatsioon, heteroskedastiivsus, multikollineaarsus ja jääkide alluvus normaaljaotusele. Samuti mudelis, kuhu lisati sentimendi, oli kontrollitud mudeli õige kuju RESET testiga. Statsionaarsuse saavutamiseks kasutati diferentseerimist.

Töö alguses püstitati kolm uurimisküsimust:

- 1) Milline on seos inimeste sentimendi (positiivne/neutraalne/negatiivne) ja krüptovara hindade vahel Bitcoin, Ethereum ja Ripple näitel?
- 2) Kas leitud mõju erineb lühiajalisel ja pikaajalisel perioodil?
- 3) Kas sentimendi kasutamine aitab paremini krüptovarade hindu prognoosida/kirjeldada?

Koostatud mudelite alusel on võimalik vastata uurimisküsimustele. Pikaajalised mudelid, mida kasutati Bitcoin, Ripple-i ja Ethereum hindade kirjeldamiseks on statistiliselt olulised ja täidavad suure osa OLS mudeli eeldustest, aga jäägid ei allu normaaljaotusele. Mudelitel oli väike determinatsioonikordaja mis näitas, et ainult sentimendi kasutamine ei anna tulemusi ning suurt seost ei esine.

Lühiajalistest mudelitest saadi kaks statistiliselt olulist mudelit. Determinatsioonikordaja oli suurem, kui pikaajalistel mudelitel aga kolmas lühiajaline mudel oli statistiliselt ebaoluline. Lühiajalise perioodi analüüs näitas, et ainult sentimendi kasutamine ei anna tulemusi ning suurt seost ei esine.

Sentimendi lisamiseks mudelisse oli esialgselt tehtud statistiliselt oluline mudel ilma sentimentita teiste näitajatega. Oli võetud sentiment, mida lisati statistiliselt olulise mudelisse. Tulemuseks saadi, et sentimentide lisamine mudelisse võib aidata paremini Bitcoin'i hinda kirjeldada või prognoosida. Determinatsioonikordaja pärast sentimentide lisamist suurenes ja sai 0,07.

Ettepanekuna edasisteks analüüsideks soovib autor analüüsida sentimentide mõju, kasutades sentimentide andmeteks käesoleva aasta päevaseid hinnanguid ja mitte võtta vanu sentimentide hinnanguid. Täiendavalt on võimalik leida suurem mõju näitajatel nagu hash rate ja bitcoin supply käesoleval aastal.

SUMMARY

SENTIMENT EFFECT ANALYSIS ON DIFFERENT CRYPTOCURRENCY PRICES

Dmitri Zhilyaev

The aim of this study was to investigate the impact of sentiment effect on different cryptocurrency prices. Sentiment effect analysis was conducted over the long-term and short-term periods. Additionally, sentiment inclusion was done into the statistically significant model without sentiment to try to analyse the effect of sentiment on the model without the sentiment. Three regression models were constructed for the long-term analysis with Bitcoin, Ethereum, and Ripple. Three models were constructed for short-term analysis with Bitcoin. One model was constructed with sentiment inclusion.

To avoid obtaining errors or wrong results, the assumptions of the OLS regression model were verified, including stationarity, autocorrelation, heteroscedasticity, multicollinearity, and the adherence of residuals to a normal distribution. Additionally, the model with sentiment inclusion underwent verification of the correct model shape using the RESET test. Differentiation was added to the variables to achieve stationarity.

In this study three research questions were posed:

- 1) What is the relationship between people's sentiment scores (positive/neutral/negative) and cryptocurrency prices, using Bitcoin, Ethereum, and Ripple as examples?
- 2) Does the identified effect differ between short-term and long-term periods?
- 3) Does the use of sentiment assist in better forecasting/describing cryptocurrency prices?

Based on the constructed models, it is possible to address the research questions. The long-term models used to describe Bitcoin and Ethereum prices were statistically significant and met the assumptions of the OLS model, however residuals did not adhere to a normal distribution. The models exhibited a small coefficient of determination. This indicated that only the use of sentiment

to describe cryptocurrencies does not give results and those models can not be used to describe cryptocurrencies' prices.

From the short-term models, two statistically significant models were obtained. The coefficient of determination was higher than in the long-term models, although the third short-term model was statistically insignificant. Sentiment had no effect on short-term models.

For the inclusion of sentiment in the model, an initially statistically significant model without sentiment but with other indicators was constructed. Sentiment was then added to this statistically significant model. The results indicated that the use of sentiment can help in constructing a model for Bitcoin. The coefficient of determination increased to 0.07 after sentiment inclusion.

As a recommendation for further analysis, the author suggests examining the impact of sentiment using daily sentiment ratings for the current year and avoiding the utilization of outdated sentiment ratings. Additionally, it is possible to explore greater impacts from indicators such as hash rate and bitcoin supply for the current year.

KASUTATUD ALLIKATE LOETELU

- Akgül, A., Şahin, E. E., & Şenol, F. Y. (2022). Blockchain-based Cryptocurrency Price Prediction with Chaos Theory, Onchain Analysis, Sentiment Analysis and Fundamental-Technical Analysis. *Chaos Theory and Applications* :, 4(3), 157–168. <https://doi.org/10.51537/chaos.1199241>
- Akyıldırım, E., Corbet, S., Lucey, B. M., Şensoy, A., & Yarovaya, L. (2020). The relationship between implied volatility and cryptocurrency returns. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3758513>
- Badiola, J. (2019, June 15). Bitcoin 17.7 million Tweets and price. Kaggle. https://www.kaggle.com/datasets/jaimebadiola/bitcoin-tweets-and-price?select=df_Final.csv
- Bhadane, C., Dalal, H., & Doshi, H. (2015). Sentiment Analysis: Measuring opinions. *Procedia Computer Science*, 45, 808–814. <https://doi.org/10.1016/j.procs.2015.03.159>
- Chalothom, T., & Ellman, J. (2015). Simple approaches of sentiment analysis via ensemble learning. In *Lecture notes in electrical engineering* (pp. 631–639). https://doi.org/10.1007/978-3-662-46578-3_74
- Ciaian, P., Rajčániová, M., & Kanacs, D. (2018). Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions & Money*, 52, 173–195. <https://doi.org/10.1016/j.intfin.2017.11.001>
- CoinGecko. (2024). *Cryptocurrency prices, charts, and crypto market cap* | CoinGecko. <https://www.coingecko.com/>
- Critien, J. V., Gatt, A., & Ellul, J. (2022). Bitcoin price change and trend prediction through twitter sentiment and data volume. *Financial Innovation*, 8(1). <https://doi.org/10.1186/s40854-022-00352-7>
- De Souza, M. J. S., Ramos, D. G. F., Pena, M. G., Sobreiro, V. A., & Kimura, H. (2018). Examination of the profitability of technical analysis based on moving average strategies in BRICS. *Financial Innovation*, 4(1). <https://doi.org/10.1186/s40854-018-0087-z>
- DeVries, P. D. (2016). An analysis of cryptocurrency, Bitcoin, and the future. *ResearchGate*. https://www.researchgate.net/publication/316656878_An_Analysis_of_Cryptocurrency_Bitcoin_and_the_Future

- Frohmann, M., Karner, M., Khudoyan, S., Wagner, R. F., & Schedl, M. (2023). Predicting the price of Bitcoin using Sentiment-Enriched Time Series forecasting. *Big Data and Cognitive Computing*, 7(3), 137. <https://doi.org/10.3390/bdcc7030137>
- Gujarati, N. D., Porter, C. D. (2003). *Basic Econometrics: Fifth Edition*. New York: McGraw Hill/Irwin.
- Grobys, K., Junttila, J., Kolari, J. W., & Sapkota, N. (2021). On the stability of stablecoins. *Journal of Empirical Finance*, 64, 207–223. <https://doi.org/10.1016/j.jempfin.2021.09.002>
- Hutto, C. J., & Gilbert, É. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the . . . International AAAI Conference on Weblogs and Social Media/Proceedings of the International AAAI Conference on Weblogs and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Hilmola, O. (2021). On prices of privacy coins and bitcoin. *Journal of Risk and Financial Management*, 14(8), 361. <https://doi.org/10.3390/jrfm14080361>
- Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M. M., Ungar, L., Schwartz, H. A., Epstein, D. H., Leggio, L., & Curtis, B. (2021). Bots and misinformation spread on social media: Implications for COVID-19. *JMIR. Journal of Medical Internet Research/Journal of Medical Internet Research*, 23(5), e26933. <https://doi.org/10.2196/26933>
- Ibekwe, A. (2021). Using sentiment analysis of social media posts in cryptocurrency price prediction. ResearchGate. <https://doi.org/10.13140/RG.2.2.16515.43047>
- İçellioğlu, C. Ş., & Öner, S. (2019). An Investigation on the Volatility of Cryptocurrencies by means of Heterogeneous Panel Data Analysis. *Procedia Computer Science*, 158, 913–920. <https://doi.org/10.1016/j.procs.2019.09.131>
- Kartal, C. (2022). ANALYSIS OF THE CORRELATION BETWEEN CRYPTO CURRENCIES, S&P500 AND US 10-YEAR TREASURY BOND INDEX WITH GRANGER CAUSALITY TEST. *Yönetim Ve Ekonomi Araştırmaları Dergisi*, 20(2), 274–291. <https://doi.org/10.11611/yead.1080595>
- Katsiampa, P. (2018). An empirical investigation of volatility dynamics in the cryptocurrency market. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3202317>
- Katsiampa, P., Yarovaya, L., & Zięba, D. (2022). High-frequency connectedness between Bitcoin and other top-traded crypto assets during the COVID-19 crisis. *Journal of International Financial Markets, Institutions and Money*, 79, 101578. <https://doi.org/10.1016/j.intfin.2022.101578>
- Khamisa, Z. (2019). An analysis of the factors driving performance in the cryptocurrency market. Do these factors vary significantly between cryptocurrencies? *ResearchGate*. https://www.researchgate.net/publication/333967521_An_analysis_of_the_factors_driving_performance_in_the_cryptocurrency_market_Do_these_factors_vary_significantly_between_cryptocurrencies

- Krištofek, L. (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. *PLOS ONE*, 10(4), e0123923. <https://doi.org/10.1371/journal.pone.0123923>
- Ozili, P. K. (2022). The acceptable R-Square in empirical modelling for social science research. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4128165>
- Pano, T., & Kashef, R. (2020). A Complete VADER-Based Sentiment Analysis of Bitcoin (BTC) Tweets during the Era of COVID-19. *Big Data and Cognitive Computing*, 4(4), 33. <https://doi.org/10.3390/bdcc4040033>
- Peterson, T. (2020). To the Moon: A History of Bitcoin Price Manipulation. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3639431>
- Reed, A. (2024). Who accepts Bitcoin in 2024? List of 20+ major companies. 99 Bitcoins. <https://99bitcoins.com/bitcoin/who-accepts/>
- Sauga, A. (2017). *Statistika õpik majanduseriala üliõpilastele*. Tallinn: TTÜ Kirjastus
- Stošić, D., Stošić, D., Ludermir, T. B., & Stošić, T. (2019). Exploring disorder and complexity in the cryptocurrency space. *Physica. A*, 525, 548–556. <https://doi.org/10.1016/j.physa.2019.03.091>
- Tripathi, G., Ahad, M. A., & Casalino, G. (2023). A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges. *Decision Analytics Journal*, 9, 100344. <https://doi.org/10.1016/j.dajour.2023.100344>
- Uras, N., Marchesi, L., Marchesi, M., & Tonelli, R. (2020). Forecasting Bitcoin closing price series using linear regression and neural networks models. *PeerJ*, 6, e279. <https://doi.org/10.7717/peerj-cs.279>
- Wang, G., & Hausken, K. (2024). Unravelling the global landscape of Bitcoin research: insights from bibliometric analysis. *Technology Analysis & Strategic Management*, 1–18. <https://doi.org/10.1080/09537325.2024.2306931>
- Xiao, Q., & Ihnaini, B. (2023). Stock trend prediction using sentiment analysis. *PeerJ. Computer science*, 9, e1293. <https://doi.org/10.7717/peerj-cs.1293>
- Γκίλλας, K., Bekiros, S., & Siriopoulos, C. (2018). Extreme correlation in cryptocurrency markets. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3180934>
- Zhilyaev, D. (2024). *Elektroonile lisa*. Google Docs. https://docs.google.com/document/d/1cnD3lNyRTfZzuWt8HwuFLgyKO_zHta8K2SqYnWu1KE0/edit?usp=sharing

LISAD

Lisa 1. Ühikjuure testimine Dickey-Fuller Testiga

	Dickey-Fuller Test
	p-value
d_1_BTC	5.274e-51
d_1_ETH	2.806e-12
d_1_XRP	7.149e-05
SENT	1.343e-05
d_BTCE	2.345e-07
d_BTCS	0.0005108
GOLD	4.994e-12
d_Trans	1.369e-07
MSCI	1.203e-05
HASH	0.0001
VIX	0.0297
d_OIL	5.424e-14
d_Wiki	7.61e-06

Lisa 2. Pikaajaliste mudelite testide tulemused

	Mudel 1	Mudel 2	Mudel 3	Mudel 4
	d_1_BTC	d_1_ETH	d_1_XRP	d_1_XRP
White's test for heteroskedasticity	0,962114	0,603653	0,363889	0,363889
LM test for autocorrelation	0,294254	0,0553185	0,00621055	0,0062106
Test for normality of residual	1,59E-20	5,68E-14	1,81E-50	1,81E-50

Lisa 3. Lühiajaliste mudelite testide tulemused

	Mudel 1	Mudel 2
	d_1_BTC	d_1_BTC
White's test for heteroskedasticity	0,517377	0,145466
LM test for autocorrelation	0,80239	0,641843
Test for normality of residual	0,264945	7,35E-08

Lisa 4. Sentimendiga mudeli testide tulemused

	Mudel 3
	d_1_BTC
White's test for heteroskedasticity	0,821891
LM test for autocorrelation	0,195595
Test for normality of residual	1,50E-18
RESET test for specification	0,165215

Lisa 5. VIF test

	VIF TEST
MSCI	1,006
d_Wiki	1,001
SENT	1,007

Lisa 6. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina Dmitri Zhilyaev

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Sentimendi mõju analüüs erinevate krüptovarade hindadele“,

mille juhendaja on Tõnn Talpsepp,

1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

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

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

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

09.05.2024

¹ Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. jq 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.