Predikcija budućih vrijednosti kriptovaluta metodama strojnog učenja i analizom blockchain informacija

Frković, Dominik

Sveučilište u Rijeci – Odjel za informatiku

Jednopredmetni preddiplomski studij informatike

Dominik Frković

Predikcija budućih vrijednosti kriptovaluta metodama strojnog učenja i analizom blockchain informacija

Završni rad

Mentor: doc. dr. sc. Sanda Bujačić Babić, v. pred. dr. sc. Vedran Miletić

Rijeka, 23. rujna 2021.

Sažetak

Kripovalute i blockchain tehnologije omogućuju sigurne, transparentne i neizbrisive transakcije, a zbog svojih svojstava imaju i brojne druge primjene pa se vjeruje da će njihova popularnost samo rasti u budućnosti. Golemu popularnost kriptovalute imaju i kod investitora pa je tako mjesečni obujam trgovine kriptovalutama već nekolimo mjeseci iznad bilijun dolara. Tema ovog završnog rada predikcija je vrijednosti kripovaluta metodama strojnog učenja analizom blockchain informacija što može, ujedinivši različite dostupne metode strojnog učenja sa značajkama iz blockchaina, dovesti do boljih odluka pri investiranju. U prvom poglavlju objašnjena je temeljna tehnologija blockchaina, drugo se poglavlje fokusira na same kriptovalute, a treće poglavlje dalje pregled znanstvenih istraživanja na temu predikcije. Na samom kraju predstavljene su primjene metoda strojnog učenja na kriptovalute u stvarnom svijetu.

Ključne riječi

kriptovalute, blockchain, strojno učenje

Tablica sadržaja

1. Uvod

U današnje vrijeme sve smo više okruženi različitim informacijama o kriptovalutama, NFT-ovima i tehnologijama vezanim uz blockchain koje će zahvaljujući svom nevjerojatnom rastu i širenju te raznovrsnim mogućnostima primjene revolucionarizirati i temeljito promijeniti svijet u kojem živimo. Jedna od najvažnijih razlika između kriptovaluta i do sada korištenih valuta je ta da kriptovalute, zahvaljujući blockchainu, nisu vezane za banke niti bilo kakve centralizirane sustave upravljanja. Zbog tog svojstva na vrijednost kriptovaluta ne mogu utjecati, ni bankarski lobiji, ni regulacija od strane države što su dva vrlo bitna faktora kod određivanja vrijednosti valuta.

Kriptovaluta (engl. cryptocurrency) se može definirati kao decentralizirana digitalna valuta zaštićena pomoću raznih kriptografskih metoda što omogućuje osiguranje od duplog dvosturkog trošenja (engl. double spending) i/ili falsificiranja [1]. Ulančani blokovi (engl. blockchain) zapravo su posebna vrsta distribuirane glavne knjige (engl. distributed ledger) i glavna su tehnologija na kojima se baziraju kriptovalute. U prvom će poglavlju detaljnije biti opisane najvažnije kriptovalute i blockchain tehnologija.

Početkom travnja 2021. tržišna kapitalizacija cjelokupnog globalnog tržišta kriptovaluta (zbroj tržišnih kapitalizacija svih kriptovaluta) prvi put je prošao vrijednost 2 bilijuna dolara (10^{12} dolara) , od čega se polovica odnosi na Bitcoin [2].

Tema ovog završnog rada bit će predikcija budućih vrijednosti kriptovaluta metodama strojnog učenja analizom blockchain informacija. Strojno učenje već se, s manjom ili većom uspješnošću, primjenjuje u trgovanju na svjetskim burzama, ali novitet u odnosu na trgovanje dionicama su blockchain informacije koje je moguće analizirati kako bi se što uspješnije predvidjeli trendovi na tržištu kriptovaluta. Postoji nekoliko radova koji su se fokusirali na predikciju vrijednosti samo kriptovalute Bitcoin kao najpoznatije i najkorištenije valute. Za predikciju se koriste razne metode strojnog učenja poput Bayesovih neuralnih mreža, RNN mreža s LTSM ćelijama, itd. Svaka će od osnovnih metoda korištenih u radu biti najprije objašnjena te će se nakon toga komparativno analizirati njihova učinkovitost. Neke su mreže trenirane koristeći samo povijesne podatke o vrijednosti kriptovalute te će i ti rezultati biti uspoređeni s rezultatima dobivenim korištenjem blockchain informacija. Uz to, navodi se i istraživanje koje za predikciju koristi i javno mišljenje utemeljeno na setovima podataka s društvene mreže Twitter.

Primjenom sve važnijih i sve korištenijih metoda strojnog učenja za predikciju vrijednosti kriptovaluta moguće je doći do relativno točnih predikcija, ali treba uzeti u obzir da su one najčešće kratkoročne (iako to nije zapreka ostvarenju velikih profita u svakodnevnom trgovanju kriptovalutama), kao i to da kod trgovanja kriptovalutama preko aplikacija i posrednika postoje troškovi transakcija koji mogu znatno utjecati na isplativost ovakvih strategija.

Kako bi se dao što kvalitetniji uvod u složeni svijet kriptovaluta i njihovog trgovanja prvo je potrebno objasniti nekoliko temeljnih pojmova vezanih uz njih.

2. Blockchain

Osnovni koncept blockchaina, predstavljen je još početkom 90-ih godina u radovima kriptografa Stuarta Habera i W. Scotta Stornetta [3], ali do njegove najvažnije primjene, one u području kriptovaluta, dolazi 2009. godine nastankom Bitcoina.

Blockchain se može definirati kao posebna vrste baze podataka koju čine šifrirani blokovi informacija koji se međusobno povezuju s prethodnim blokom popunjenog kapaciteta tvoreći lanac blokova koji je neraskidiv i kronološki poredan [4].

Blockchain[1] je najvažnija tehnologija koja je omogućila pojavu i korištenje kriptovaluta. Zahvaljujući svom jedinstvenom načinu funkcioniranja kriptovalute mogu ostvariti decentraliziranost, sigurnost i transparentnost, što se navode kao njihove najveće prednosti.

U blockchainu se mogu pohranjivati različite vrste informacija pa se tako sve više tvrtki okreće bilježenju podataka o lancima opskrbe u blockchainu. Na primjer, Walmart koristi ovu tehnologiju za praćenje transporta i dostave svojih pošiljki te za nadzor lanaca opskrbe nekih svojih proizvoda (npr. zelene salate), kako bi se u slučaju kontaminacije mogao lakše pronaći izvor kontaminacije i izbaciti kontaminirane proizvode [5].

Primjena blockchaina su i tzv. pametni ugovori, odnosno „programi pohranjeni u blockchainu koji se izvršavaju kad se ispune određeni unaprijed dogovoreni uvjeti”. Ovakvi ugovori eliminiraju potrebu za trećim stranama (posrednicima), a mogu se koristiti i za automatizaciju radnih procesa [6].

Vjerojatno najveća promjena koju kriptovalute donose u odnosu na današnji način rukovanja novcem jest činjenica da banke više nisu potrebne. U tradicionalnim bankarskim sustavima, banke su središte sustava, čuvaju podatke o transakcijama, ali i klijentima - što nije uvijek dobro i poželjno te, iako jamče određenu razinu sigurnosti, generalno gledano, korisnici ipak gube dio svoje privatnosti. Navedeni bankarski mehanizmi jamče određenu dozu sigurnosti, no ta razina sigurnosti je često manjkava pa se klijentu može dogoditi da uz sve provedene bankarske sigurnosne mjere i regulative ipak podaci budu objavljeni i/ili korišteni u nekontrolirane i nepredviđene svrhe, što se događalo u praksi vrlo često.

Prema MIT Technology Reviewu „poanta korištenja blockchaina je da se ljudima omogući dijeljenje podataka na siguran i naknadno nepromjenjiv način bez da se prethodno morao ostvariti uvjet međusobnog povjerenja” [7]. U blockchainu kao decentraliziranoj tehnologiji navedeni uvjet ne postoji jer ne postoji središnja centralna jedinica (računalo ili grupa računala) koja čuva i posjeduje lanac, nego su transakcije zapisane u blockchainu i distribuirane svim čvorovima (engl. node) u mreži, što blockchain čini oblikom distribuirane glavne knjige (engl. distributed ledger). Svaki čvor, odnosno svaki član mreže, ima vlastitu kopiju svih podataka u blockchainu koja se sinkronizira sa svim ostalim čvorovima u mreži.

Svaka pokrenuta transakcija u banci autorizira se samo od strane banke. Metodom blockchaina, mreža sačinjena od čvorova mora verificirati transakciju konsenzusom i kad se transakcija jednom obavi, ona se upisuje u blok te postaje trajni dio bloka informacija. Zahvaljujući tome što svaki čvor ima vlastitu kopiju blockchaina, ukoliko dođe do greške ili zlonamjerne promjene na lancu na nekom čvoru, takvu je situaciju moguće uočiti i ispraviti uzimajući podatke s nekog drugog čvora iz mreže. Na taj se način održava integritet podataka, odnosno njihova

1 Hrvatski prijevod pojma blockchain je ulančani blok informacija

validnost, konzistentnost i cjelovitost tijekom cijelog životnog procesa [8]. Više o ovom mehanizmu i rudarenju kriptovaluta biti će riječi u nastavku kroz primjer kriptovalute Bitcoin.

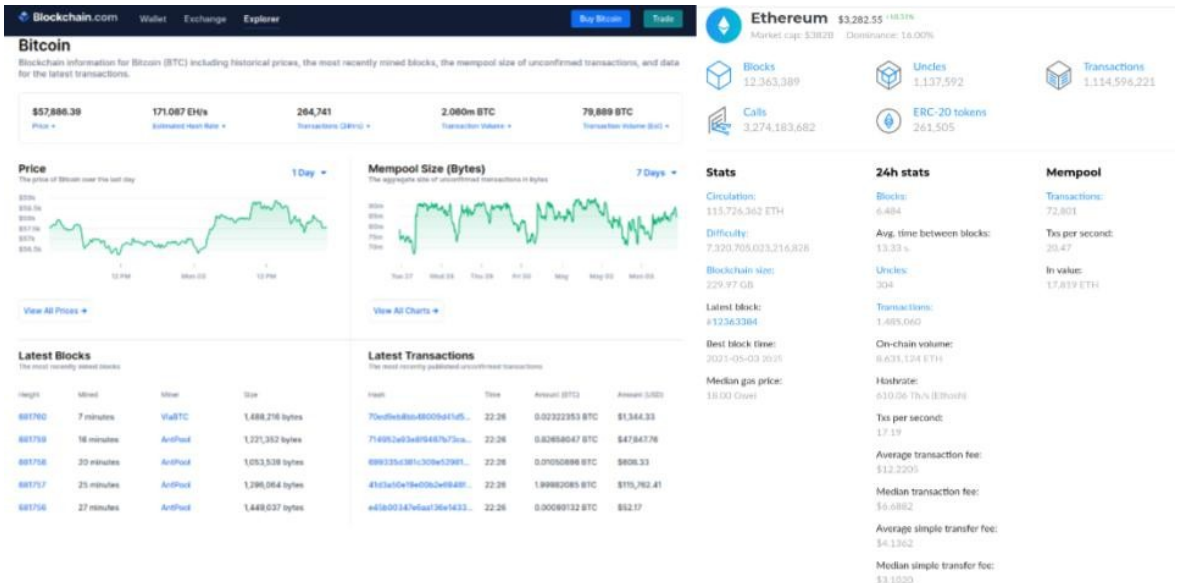
Spomenimo i transparentnost blockchain tehnologije. Uz to što svaki čvor u mreži može vidjeti sve transakcije jer posjeduje kopiju razdijeljenih knjiga, također postoje i online blockchain exploreri, alati kojima bilo tko može uživo pratiti sve trenutne i prošle transakcije koje se odvijaju u blockchainu. Alati te vrste su:

- Blockchair [9],
- Blockchain Explorer [10],
- BTC.com [11].

Većina ovih pretraživača nudi slične informacije poput informacija o blokovima, najnovijim transakcijama (vremenska oznaka, adrese, količina, provizija) i statusu mreže te iz njih izvlače različite statistike. Blockchain nudi najširi izbor kriptovaluta koje se mogu pregledavati, kao i mogućnost preuzimanja statistika u obliku tablica i grafova što je vrlo korisno ako želimo raditi predikciju koristeći metode strojnog učenja koje koriste povijesne podatke.

Neke od zanimljivijih informacija su mempool (engl. memory pool, memorijski bazen), odnosno mehanizam čvorova gdje se nalaze još nepotvrđene transakcije, tako da možemo reći da je to svojevrsna čekaonica u kojoj transakcije čekaju na verifikaciju i dodavanje u blok. Možemo promatrati i visinu bloka koju predstavlja broj ulančanih blokova (u trenutku pisanja ovog rada visina bloka Bitcoina je 681,762 blokova, a Ethereuma 12,363,488 blokova).

Slika 1 prikazuje dio mogućnosti koje nude blockchain preglednici.



Slika 1: Prikaz web sučelja blockchain explorera

Izvor: https://www.blockchain.com/explorer/?utm_campaign=dcomnav_explorer i https://blockchair.com/ethereum , pristupljeno 2.svibanj 2021.

3. Kriptovalute

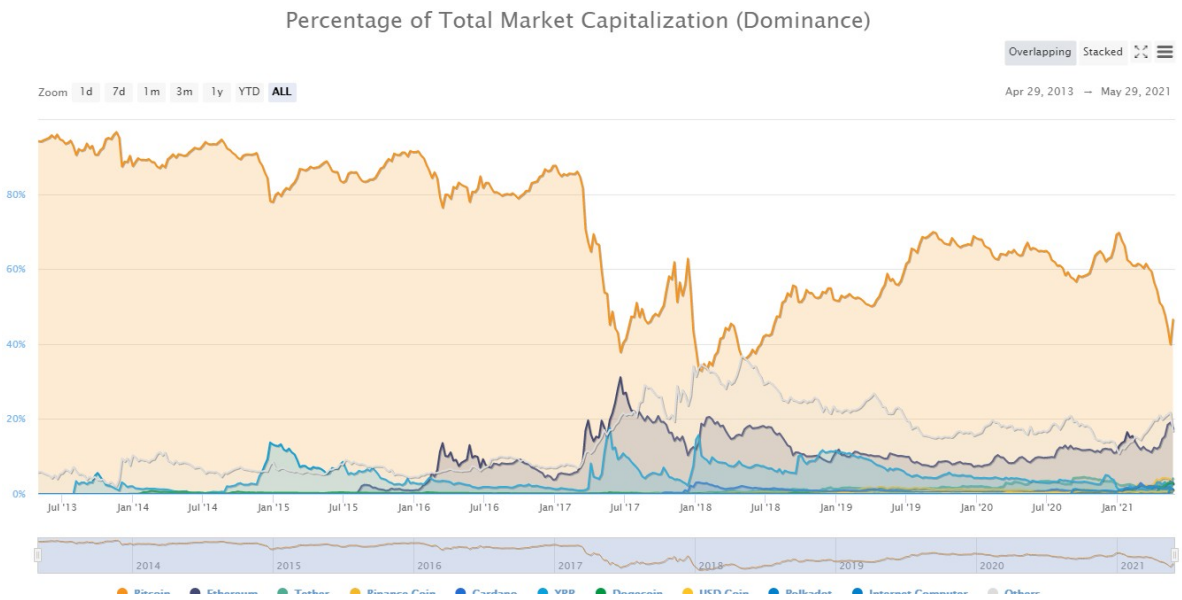
Za razliku od tradicionalnih valuta, kriptovalute su decentralizirane i njihov nastanak nije vezan uz banke i vlasti. Kriptovalute se, analogno kao i zlato, rudare što ih čini otpornima na intervencije (vlade ili centralne banke) koje bi im mogle promijeniti vrijednost, što također čini i predviđanje kretanja vrijednosti kriptovaluta znatno složenijim. Ovakve valute za zaštitu i očuvanje integriteta koriste različite kriptografske tehnike i algoritme šifriranja poput hashiranja i javnih i privatnih ključeva.

Hashiranje se može definirati kao postupak generiranja izlaza fiksne veličine iz ulaza promjenjive veličine korištenjem hash funkcija. Kriptografske hash funkcije su determinističke što znači da će algoritam hashiranja za jednak ulaz uvijek proizvoditi jednak izlaz. Također, ovakvi algoritmi dizajnirani su da budu jednosmjerni, odnosno osmišljeni su tako da su za pronalazak ulaza iz unaprijed poznatog izlaza potrebne velike količine računarskog vremena i resursa [12].

Javni i privatni ključevi su kriptografske metode primjenom kojih se obavljaju transakcije na blockchainu. Javni ključ je kriptografski kod uparen s privatnim ključem koji omogućuje primanje transakcija kriptovaluta. Javni ključevi su javno poznati i svatko može poslati kriptovalutu na adresu, tj. skraćeni oblik javnog ključa. Privatni ključ treba biti poznat samo njegovom vlasniku jer omogućuje dokazivanje vlasništva i trošenje kriptovaluta. Javni ključ se generira pomoću privatnog ključa korištenjem kriptografskih algoritama, a zahvaljujući jednosmjernim funkcijama obratan smjer je gotovo nemoguće realizirati.

Opisat ćemo postupak transakcije na blockchainu. U prvom koraku se transakcija šifrira pomoću javnog ključa te je se nakon toga može dešifrirati samo korištenjem pripadajućeg privatnog ključa. Nakon toga se transakcija potpisuje korištenjem privatnog ključa čime se dokazuje da nije bilo modifikacija, a takav se potpis generira kombinacijom privatnog ključa i podataka koji se šalju. Na kraju se autentičnost transakcije može dokazati samo korištenjem pripadajućeg javnog ključa [13].

U travnju prošle godine prema CoinmarketCapu [14] je postojalo više od 7800 različitih kriptovaluta, od kojih je oko 2000 „mrtvo”, što znači da se više ne koristi. Najveći broj tih valuta je za sad irelevantan i ne zauzima nikakav značajan udio u ukupnom kapitalu tržišta kriptovaluta, ali treba napomenuti da je riječ o 2·1012 dolara vrijednom tržištu pa i samo 0.00005% tog tržišta je vrijedno milijun dolara. Slika 2 prikazuje kretanje udjela najvažnijih kriptovaluta u ukupnom tržištu kapitala. Možemo vidjeti da je na samom početku Bitcoin dominirao tržištem, ali nakon najvećeg pada 2017. godine nije se uspio uzdići iznad 70%. Trenutno, sa snažnim uzletom Ethereuma, ali i drugih kriptovaluta, Bitcoinov udio na tržištu i dalje pada što ne znači da pada njegova vrijednost ili važnost kao kriptovalute, već da se i druge kriptovalute uspijevaju pozicionirati kao relevantne i kvalitetne, neke zahvaljujući dobrim tehnologijama koje stoje iza njih, a neke, poput Dogecoina (kojeg javno podržava Elon Musk), zbog snažnih marketinških kampanja koje se provode na društvenim mrežama pa se manje educiranoj široj publici predstavljaju kao kvalitetan izbor.



Izvor: https://coinmarketcap.com/charts/ , pristupljeno 29.svibanj 2021.

Korištenje kriptovaluta kao općeprihvaćenog sredstva plaćanja kako su to zamislili tvorci kriptovaluta još nije u potpunosti zaživjelo, ali trend širenja mogućnosti plaćanja kriptovalutama svakako postoji. Od ove godine Paypal planira uvesti mogućnost trgovanja kriptovalutama. Odnedavno moguće je kupiti i automobil koristeći kriptovalute, što je usvojena praksa i u Hrvatskoj, gdje je jedna autokuća uvela mogućnost plaćanja kriptovalutom po izboru [15].

Za transakcije kriptovaluta koriste se javni i privatni ključevi pa su te transakcije gotovo anonimne. Iako je koncept kriptovalute najčešće vezan uz vrlo pozitivne karakteristike, postoje i neke negativne okolnosti koje se vezuju uz trgovanje kriptovalutama poput mogućnosti pranja (ilegalno stečenog) novca i raznih načina izbjegavanja plaćanja poreza Zbog anonimnosti pri transakcijama kriptovalute se često koriste kao sredstvo plaćanja u ilegalnim aktivnostima i na tamnom webu [1].

Iako se za kriptovalute može reći da su same po sebi sigurne, zahvaljujući brojnim zaštitnim mehanizmima, isto ne mora vrijediti za trgovanje kriptovalutama. Kako bi se trgovalo kriptovalutama, potrebno je koristiti novčanik (engl. wallet) i alate za razmjenu (engl. exchange) koji mogu biti meta hakerskih napada i krađa.

3.1. Bitcoin

Najpoznatija kriptovaluta s najvećim udjelom na tržištu je Bitcoin. Sam temelj Bitcoina postavio je u bijeloj knjizi (engl. white paper) *„Bitcoin: A Peer-to-Peer Electronic Cash System”* osoba ili grupa ljudi pod pseudonimom Satoshi Nakamoto 2008. godine. Do današnjeg dana nije otkriven Satoshijev identitet, iako postoje brojne teorije o tome tko bi on mogao biti. Najmanji dio Bitcoina zove se Satoshi i iznosi 0.00000001 BTC. Svijet je 2007. zahvatila velika financijska kriza, a jedan od glavnih uzroka bilo je neodgovorno ponašanje bankarskog sektora i velikih financijskih igrača te se može reći da je Bitcoin djelomično nastao kao odgovor na tadašnje aktualne prilike i kao dokaz nepovjerenja u institucije.

Bitcoin, kao i većina drugih kriptovaluta, funkcionira na principu blockchaina, a kako je Bitcoin prva ovakva valuta i u njegovoj bijeloj knjizi je prvi puta stvarno definiran blockchain kakvog danas poznajemo, izrazi Bitcoin

i blockchain su ponekad bili gotovo pa sinonimi. Slijedi opis načina rada Bitcoina kako ga je zamislio Satoshi.

Bitcoin se definira kao peer-to-peer sustav (svatko-sa-svakim), što se zapravo odnosi na komunikaciju između čvorova u mreži. Kao što je opisano kod definiranja blockchaina, u mreži se sve nove transakcije šalju do svih čvorova koji ih potom skupljaju u blok te svaki od čvorova u mreži potom pokušava pronaći dokaz o radu (engl. proof of work) za svoj blok. Čim čvor uspije u tom pronalasku, šalje svim ostalim čvorovima blok, a oni ga prihvaćaju ukoliko su u tom bloku sve transakcije važeće i nema dvostruke potrošnje[2]. Taj blok potom ulazi u lanac ulančanih blokova, a njegov hash se koristi kao prethodni hash za idući blok u lancu. Moguće je da se dogodi da više čvorova emitira različite verzije najnovijeg bloka. Ta situacija će se razriješiti pomoću već navedene činjenice da čvorovi uvijek smatraju da je najdulji lanac ispravan. Kada prime te konfliktne blokove, čvorovi traže dokaz o radu za onaj koji je došao prvi, a druge zadržavaju u slučaju da duljina tog alternativnog lanca pretekne onaj na kojem rade. Kad se pronađe idući dokaz o radu, blok se dodaje u lanac te stoga jedan od lanaca postaje dulji i taj lanac postaje „pravi", odnosno točan. Zahvaljujući razdijeljenosti blockchaina, ukoliko dođe do toga da neki čvor u mreži i ne primi jedan blok, uvijek postoji niz drugih čvorova s ispravnom verzijom lanca od kojih ju može preuzeti i tako se vratiti na pravi put [16].

Transakcije Bitcoina su kao što smo opisali upisane u blokovima i ulančane, a kako bi se transakcija dogodila, vlasnik kovanice (engl. coin) digitalno potpisuje hash prethodne transakcije i javni ključ idućeg vlasnika te ih dodaje na kraj kovanice [16].

Da bi se zaštitilo od dvostrukog trošenja kovanica, uvodi se dokaz o radu i vremenske oznake (engl. timestamp). Kao i sami blokovi, vremenske oznake tvore lanac i to tako da iduća vremenska oznaka sadrži u svome hashu vremensku oznaku prethodne, a navedeno ih čini dodatno otpornima na bilo kakve manipulacije. U slučaju Bitcoina dokaz o radu izvršava se povećavanjem noncea (skraćenica od engl. number used only once, što je u prijevodu „samo jednom korišteni broj") sve do pronalaska hasha sa zadanim (potrebnim) brojem nula za taj blok. Proces zahtjeva određenu količinu utrošene procesorske snage te jednom kad je ta snaga utrošena, bilo kakva izmjena bloka bez trošenja procesorske snage nije moguća. Zbog ulančavanja blokova trebalo bi obraditi i sve blokove koji dolaze nakon tog bloka, što pruža znatnu zaštitu od hakerskih napada. Bitcoin kod hashiranja koristi kriptografski algoritam SHA256 i izlaz je fiksne duljine za sve ulaze, što dodatno povećava sigurnost podataka. SHA256 algoritam hashiranja (Secure Hash Algorithm 256) inačica je SHA-2 algoritma, razvijenog od strane Agencija za nacionalnu sigurnost (NSA). Algoritam generira 256-bitni (64 znaka) slučajni niz slova i brojeva (izlaz) iz bilo kojeg ulaza. Na primjer, za ulaz kripto dobiva se d0d519866210e5dffa7d23fe9fb21ca8dd1d8f642a908ddf166321009b67caa5, a za ulaz kript izlaz je bb1cde507f22c8c10ccdbf6b301e44607b984a5069a03b970c305edf99b3d855.

Ukoliko većinu procesorske snage kontroliraju pošteni čvorovi, neće biti nekonzistentnosti u mreži i ispravan će lanac biti najdulji jer se odluke donose konsenzusom, odnosno dogovorom većine čvorova u mreži. Satoshi je u obzir uzeo i činjenicu da brzine rada računala rastu, kao i da se ne može predviđati interes za uključenjem čvorova u mrežu pa se tako težina dokaza o radu (engl. proof of work difficulty) prilagođava s obzirom na prosječni broj blokova po satu.

Promotrimo i analizirajmo scenarij u kojem se jedna osoba odlučuje promijeniti dio lanca u svoju korist, odnosno zaraditi. Ranije je opisano da bi, ukoliko se promijeni samo jedna kopija lanca, glasovanjem većine ta kopija bila

2 Mogućnost paralelnog trošenja iste jedinice kriptovalute dva ili više puta, tehnički problem koji se pojavljuje samo kod digitalnih valuta.

odbačena kao pogrešna ili prepravljena. Ovakve postavke impliciraju da napadač mora promijeniti više od polovice kopija lanca da bi njegova verzija lanca postala izglasana, što bi zahtijevalo golemu količinu novca i resursa, s obzirom na količinu već uključenih čvorova u mrežu, a što bi bilo skuplje nego potencijalna moguća zarada. Također, blokovi u lancu su kronološki poredani i povezani hash funkcijama te sadrže vremenske oznake. Kad bi potencijalni napadač želio promijeniti neki podatak u bloku, neizbježno bi promijenio hash i vremensku oznaku tog bloka, a zato što su blokovi ulančani (svaki blok sadrži vlastiti hash, hash prethodnog bloka i vremensku oznaku) morao bi zatim izmijeniti i sve blokove nakon njega što prepravljanje podataka zapisanih u blockchainu čini gotovo nemogućom misijom [17].

3.2. Rudarenje

Rudarenje Bitcoina je, prema službenoj Bitcoin stranici, proces u kojem računalno sklopovlje izvodi matematičke izračune u Bitcoin mreži koji služe za potvrdu transakcija i povećanje sigurnosti [18]. U zaglavlju bloka nalazi se broj verzije bloka, vremenska oznaka, hash prethodnog bloka, hash Merkelovog korijena, nonce i ciljani hash. U praksi, računa se hash funkcija trenutnog bloka algoritmom SHA256. Merkleovo stablo ili binarno hash stablo, struktura je podataka koja se koristi za učinkovito sažimanje i provjeru integriteta velikih skupova podataka u što se može ubrojiti i blockchain.

U slučaju Bitcoina, Merkleovo stablo sažima sve transakcije u bloku stvarajući cjelokupni digitalni otisak svih transakcija čime se omogućuje učinkovita provjera je li transakcija uključena u blok. Merkleovo stablo konstruira se rekurzivnim dvostrukim SHA256 hashiranjem parova čvorova sve dok se ne dobije samo jedan hash nazvan Merkleov korijen [19].

Uzmimo kao primjer blok visine 676767. Njegov hash je 00000000000000000000041296aa8726457c499f865e4e5c41b8c82c5bb320740e, vremenska oznaka 2021-03-29 01:53, broj verzije bloka 0x20000000, nonce 638,161,599, a Merkleov korijen 2e6aeb657c456f28925f3dd8fe320071ac941ef1c919b264f4aaa1cfe0c516a6. U njemu je zapisano ukupno 2,021 transakcija. Svi su ovi podaci javno dostupni u različitim blockchain explorerima, ali i sigurni zahvaljujući kriptografskim tehnikama koje ih štite od manipulacija.

Rudari prvo moraju potvrditi blok veličine 1MB, a zatim kako bi dodali novi blok u lanac moraju prvo pogoditi (pronaći pogađanjem) 64-znamenkasti heksadekadski broj koji je manji ili jednak ciljnom hashu mijenjajući nonce (SHA256 algoritam sam po sebi za isti ulaz, tj. podatke u bloku daje uvijek isti izlaz pa se nadodaje nonce kako bi se „pogodio” ciljani hash). Za ilustraciju postoji 18,446,744,073,709,551,615 različitih 64-znamenkastih brojeva pa je jasno da rudarenje nije brz ni jednostavan proces. Težina se prilagođava pomoću ciljanog hasha svakih 2016 blokova (u praksi otprilike svaka dva tjedna)[3]. Novi se blok trenutno izrudari svakih 10-ak minuta [20]. Još jedan pojam vezan za rudarenje je brzina hashiranja, odnosno mjerna jedinica ukupne procesorske snage mreže koja ukazuje na to koliko se u vremenskom periodu obavi operacija hashiranja, a trenutno iznosi $161 \cdot 10^{18}$ izračuna po sekundi.

Nagrada koja se dobiva za uspješno izrudaren blok na samom je početku iznosila 50 Bitcoina, ali se taj iznos prepolavlja (engl. halving) nakon svakih 210,000 blokova (otprilike svake 4 godine) pa od svibnja 2020. iznosi

3 U svibnju 2021. težina iznosi $23 \cdot 10^{12}$ što znači da je vjerojatnost da se pogodi ciljani hash $1:23 \cdot 10^{12}$, a vjerojatnost dobitka Eurojackpota u iznos od preko 37 milijuna eura je 1:95,344,200

6.25 Bitcoina odnosno 339.966 dolara (na dan 4. svibanj 2021.) [21].

3.3. Ether (Ethereum)

Nakon nastanka Bitcoina i njegovog širenja nastaju i druge kriptovalute koje zajedničkim imenom zovemo altcoinovi ili alternativni novčići. Većina ih je kao i Bitcoin utemeljena na tehnologiji blockchaina, ali nemaju sve jednake ciljeve te koriste blockchain tehnologiju na različite načine tražeći tako svoje mjesto na tržištu.

Trenutno druga najproširenija kriptovaluta je Ether platforme Ethereum. Dok Bitcoin teži rasprostranjenju kao platežno sredstvo, načini upotrebe Ethereuma prema njegovim tvorcima, su mnogo širi. Ether je podloga „svjetskog računala“, kako se naziva Ethereumova mreža, a koristi se za pokretanje i monetizaciju pametnih ugovora i decentraliziranih aplikacija nastalih na Ethereumovoj platformi, za razliku od Bitcoina koji se koristi za trgovanje i plaćanje. Ethereum je nastao kao zamisao Vitalika Buterina, 21-godišnjeg zaljubljenika u blockchain. Mreža Ethereuma puštena je u rad u 2015. godine i u samo nekoliko godina njegova je cijena porasla s 1.25 dolara u kolovozu 2015. na današnjih 3393 dolara. Glavni pokretač takvog skoka je Ethereumova odlična integracija s rastućim svijetom DeFi-ja (decentraliziranih financija) i Nezamjenjivim tokenima (NFT – non-fungeable tokeni) koji su dio Ethereumovog blockchaina kojima se predstavlja vlasništvo nad jedinstvenim predmetom, npr. umjetničkim djelom ili kolekcionarskom akvizicijom [22].

Ideja Ethereuma je da postane „protokol za izgradnju decentraliziranih aplikacija“ s naglaskom na iznimno brzo vrijeme njihova razvitka, kao i sigurnost za rijetko korištene i manje aplikacije uz mogućnost vrlo učinkovite interakcije između aplikacija [23]. Kao i kod Bitcoina, želja je da takve aplikacije, popularno nazvane dapps (engl. decentralized applications) rade bez zastoja, različitih prevara i posebno bez kontrole i uplitanja trećih strana.

Ethereum se može rudariti kao i Bitcoin, a za razliku od Bitcoina koristi se nešto drugačiji algoritam nazvan Ethash [23].

Stvaranje pametnih ugovora, koji se aktiviraju, tj. izvode tek nakon ispunjenja određenih uvjeta jedan je od popularnih načina upotreba ove platforme.

Pametni ugovori se spajaju u blockchain pomoću DAppova. DAppovi se dijele u tri kategorije:

- financijske aplikacije koje korisnicima omogućuju upravljanje i sklapanje ugovora te upravljanje novcem,
- polufinancijske aplikacije u kojima novac nije najvažniji čimbenik, kao što su samoaktivirajuće naknade (financijske) za izvršavanje računalnih zadataka ili isplata putnog osiguranja u slučaju otkazivanja leta,
- aplikacije koje koriste blockchain tehnologiju, npr. online glasanje, decentralizirano upravljanje, aplikacije za praćenje distribucije energije [24].

U idućih nekoliko godina u planu je i Ethereum 2.0 kao nadogradnja na trenutni Ethereumov blockchain koji bi trebao povećati skalabilnost, brzinu i učinkovitost mreže, te istovremeno eliminirati uska grla i povećati broj transakcija.

3.4. Link (ChainLink)

Link je kriptovaluta Chainlinka, nastala 2017. na bazi Ethereuma, a razlikuje se od drugih po svojoj funkcionalosti koja pokušava povezati podatke u blockchainu s podacima i događanjima iz vanjskog svijeta.

ChainLink je decentralizirana mreža sigurnih oracleova koja donosi poboljšanje kod pametnih ugovora. Oracle je upravljački sustav baza podataka koje spajaju vanjske podatke s blockchain mrežom [25]. Problem koji se pojavljuje pri korištenju pametnih ugovora je njihov rad s eksternim podacima (potrebno ih je slati prema van) te verificiranjem i zaštitom tih podataka. Pomoću ChainLinka je moguće razmjenjivati točne i vjerodostojne podatke između različitih blockchain mreža ili ih izravno slati pametnom ugovoru.

Glavna načela dizajna i rada ChainLinka su ranije opisana decentraliziranost i modularnost koja pojednostavljuje i omogućuje fleksibilan dizajn sustava što se postiže načelom izgradnje više manjih alata koji kvalitetno odrađuju svoje funkcije. Nadalje, dizajn se vodi politikom otvorenog koda te se potiče suradnja unutar zajednice kako bi se platforma kontinuirano poboljšavala i rasla [26].

Uzmimo kao primjer upotrebe ChainLinka pametne ugovore s vrijednosnim papirima (engl. securites smart contracts), npr. obveznice. Njima će sigurno trebati pristup API-jima koji izvještavaju o cijenama i referentnim podacima na tržištu, kamatne stope i sl. što će se izvršavati preko oracleova. Osim toka podataka prema unutra (prema blockchainu) moguć je i obrnut smjer, primjerice slanje poruke o uspješnom izvršenom plaćanju prema nekoj financijskoj instituciji kao što je banka [25].

3.5. Volatilnost kriptovaluta

Volatilnost tržišta može se definirati kao osnovna mjera rizika, a njome se mjeri opseg promjena u cijeni financijske imovine. Ako je ona visoka, to znači da je ulaganje rizično, tj. da postoji veliki rizik da se dogodi gubitak, ali, istovremeno, velika je i kratkoročna mogućnost zarade. Niska volatilnost znači da je neko ulaganje sigurno, odnosno da je rizik od gubitka nizak [27].

Vrijednost kriptovaluta vrlo je volatilna što donosi rizike i nesigurnost u investiranju, iako novija istraživanja [28] pokazuju da čak 29% dionica S&P 500 indeksa pokazuje veću volatilnost vrijednosti tokom perioda od godine dana nego vrijednost Bitcoina. Unatoč tome tome, zahvaljujući brojnim prednostima koje donosi korištenje kriptovaluta, danas je njihova primjena sve češća i sve više ljudi se odlučuje na investiranje u kriptovalute.

Zanimljivo je istaknuti i činjenicu da je u veljači 2011. godine jedan Bitcoin imao vrijednost 1 dolar, a danas jedan Bitcoin vrijedi oko 57 tisuća dolara, što znači da bi s ulaganjem 2011. od $17 danas postao milijunaš, što je također jedan od faktora zašto se danas ljudima čini jednostavno obogatiti na kriptovalutama te se sve više odlučuju na investiranje.

U radu Domagoja Sajtera *„Financijska analiza kriptovaluta u odnosu na standardne financijske instrumente“* izvršena je financijska analiza javno dostupnih podataka o trgovanju 10 kriptovaluta (Bitcoin, Bitcoin Cash, Cardano, Ether, IOTA, Litecoin, NEM, NEO, Ripple i Stellar) u razdoblju od 1. siječnja 2016. do 17. siječnja 2018. te su na kraju uspoređene s ponašanjem dva najveća burzovna indeksa, tečajem EUR/USD i cijenom nafte i zlata. Na kraju se dolazi do zaključka da je cijena kriptovaluta puno promjenjivija u usporedbi s klasičnom financijskom imovinom. Također je uočeno da cijene kriptovaluta nisu značajnije povezane sa standardnim financijskim tržištima. Zahvaljujući tome moguće je gledati kriptovalute kao zasebnu investicijsku klasu. Analizom se također dolazi do zaključka kojeg je moguće uočiti i u često senzacionalističkim vijestima o brzoj i velikoj zaradi na kriptovalutama, ali i pričama o tome kako se u jednoj noći mogu izgubiti golemi iznosi, što čini tržište kriptovaluta je iznimno volatilnim: s potencijalno visokim zaradama, ali i gubitcima [29].

Predikcija budućih vrijednosti kriptovaluta iznimno je složena i česte su pogreške čak i kod iskusnih analitičara. Jedan od razloga je činjenica da je na tržištu sve više novih kompanija, različitih valuta i tehnologija što postaje jako teško pratiti i analizirati kvalitetu i potencijal svakog projekta. Kriptovalute su vrlo mlado tržište i ne postoji još dovoljno istraženih modela i strategija koji su se pokazali kao učinkoviti za korištenje u predikciji [30].

Prema Forbesu jedan od uzroka za volatilnost je činjenica da zapravo nitko ne zna koja bi trebala biti cijena određene kriptovalute te se ne zna pretplaćuje ili se ili potplaćuje određena kriptovaluta, već je sve na špekulativnoj razini [31].

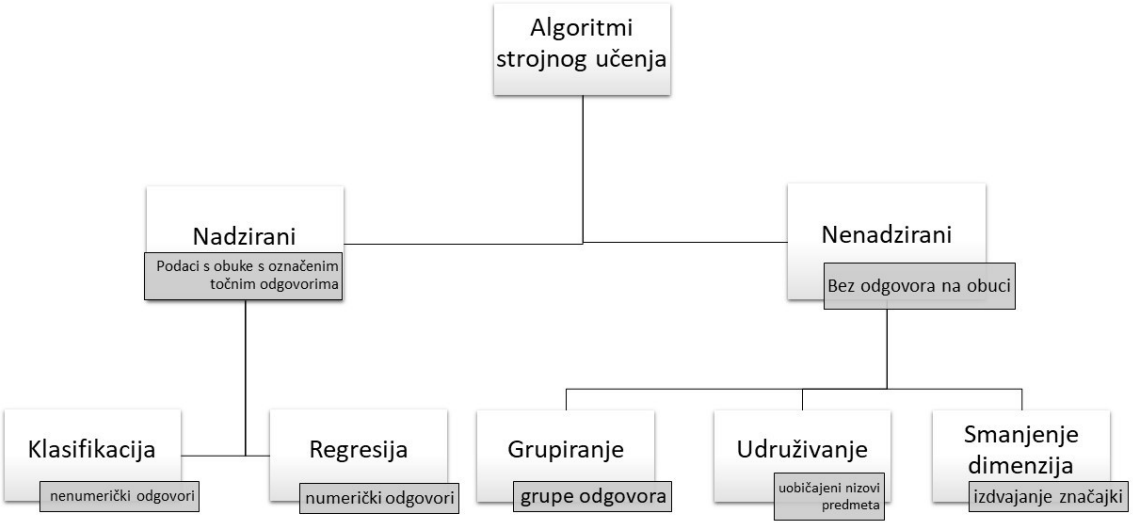
Na volatilnost nepovoljno utječe brzina širenja vijesti i, kako se čini njihov velik i trenutak utjecaj na vrijednost kriptovalute [32]. Kad se proširi neka negativna vijest o kriptovaluti (hakerski napad, korištenje u ilegalnim aktivnostima, potencijalna naplata poreza), tada kriptovalute u par sati mogu izgubiti na vrijednosti i po nekoliko desetaka postotnih poena što se ne događa s tradicionalnim valutama. Ipak, treba naglasiti da je u budućnosti moguće smanjivanje visoke volatilnosti s ulaskom velikih institucionaliziranih investitora na tržište, kao i sa stvaranjem sve šireg ekosustava kriptovaluta i tehnologija koje se temelje na blockchainu [30].

4. Predikcija budućih vrijednosti kriptovaluta metodama strojnog učenja

4.1. Strojno učenje

Kombinacija strojnog učenja i kriptovaluta postaje sve popularnija tema i sve se više istraživača uključuje u nju. Poznata je činjenica da iskoristivost modela strojnog učenja iznimno ovisi o kvaliteti i točnosti podataka koji se koriste u fazi uvježbavanja, a kod kriptovaluta je zahvaljujući blockchainu baš taj dio uvijek osiguran i podaci o kriptovalutama skupljaju se od prvoga dana njihovog postojanja, što za npr. Bitcoin znači da postoje podaci unutar zadnjih 12 godina. Postoje i brojne kvalitetne stranice od kojih su neke predstavljene u prvom poglavlju gdje je moguće pronaći različite javno dostupne blockchain informacije.

Razvoj umjetne inteligencije počeo je davno, već nakon Drugog svjetskog rata [33]. Graf (Slika 3) ispod prikazuje podjelu algoritama strojnog učenja.



Izvor: vlastita izrada

Poseban oblik strojnog učenja je duboko učenje (engl. deep learning) koje podrazumijeva korištenje neuronskih mreža koje se modeliraju i mogu naučiti prepoznavati uzorke [34].

Godinama unazad ove se metode koriste kod predikcija financijskih događanja (banke, burze, osiguranja) pa je prirodno postaviti pitanje može li se isto postići i za tržište kriptovaluta i s kojom točnošću.

4.2. Predikcija temeljena na analizi blockchaina

Nakon što je predstavljen uvod u svijet kriptovaluta i njihovog trgovanja, u nastavku će biti predstavljeni konkretni rezultati istraživanja na području predikcije budućih vrijednosti kriptovaluta metodama strojnog učenja

analizom blockchain informacija. Ovo je područje, kao i same kriptovalute, još relativno novo i stoga nedovoljno istraženo. Zbog određenih sličnosti s područjem trgovanja dionicama na burzi postoje radovi koji istražuju predikciju vrijednosti kriptovaluta metodama strojnog učenja koje se koriste na burzi. Metode iz tih radova koriste blockchain informacije koje nisu toliko strogo vezane uz samu blockchain tehnologiju zbog čega se gubi pregršt dodatnih informacija potencijalno korisnih za predikciju. Najčešće se koriste samo metrike poput: cijene, dinamike ponude i potražnje, tržišnog kapitala, udjela na tržištu (engl. market share), volumena trgovanja i sl. Slične se metrike koriste i za predikciju cijena dionica, ali se kao velika prednost kod korištenja blockchaina uzima činjenica da su te informacije nepromjenjive i samim time uvijek vjerodostojne, a najčešće su i javno dostupne.

U *„Anticipating Cryptocurrency Prices Using Machine Learning“* [35] dano je jedno takvo istraživanje gdje su rangirane performanse tri modela za predikciju dnevnih vrijednosti kriptovaluta. Istraživanje je napravljeno 2018. godine kada je kriptovaluta na tržištu bilo znatno manje nego danas, a uključivalo je 1681 kriptovaluta. Kako bi se umanjio efekt tržišnih promjena, jer kroz većinu vremena tržište kriptovaluta raste, vrijednosti kriptovaluta prikazane su u Bitcoinu umjesto u tradicionalnoj valuti. Informacije, odnosno podaci, koji su preuzeti o njima su: vrijednost u USD, tržišni kapital određen kao produkt vrijednosti i cirkulirajuće ponude (engl. circulating supply) i volumen trgovanja (engl. trading volume) predstavljen kao broj razmijenjenih novčića u danu, za razdoblje od 11.studenog 2015. do 24. travnja 2018.g. Kako bi prikazali rezultate, odnosno pokazali korisnost svakog od prikazanih algoritama, koristi se povrat ulaganja (engl. ROI, Return of investment) na konkretan dan s obzirom na cijenu, a kako je fokus istraživanja bio na kratkoročnoj predikciji, računa se povrat ulaganja nakon jednog (svakog) dana. ROI za neku kriptovalutu se računa prema formuli:

$$\frac{vrijednostkriptovalutenadan - vrijednostkriptovalutenaprethodnidan}{vrijednostkriptovalutenaprethodnidan}$$

Uspoređene su tri metode koje se i inače koriste za kratkoročnu predikciju vrijednosti, a kao osnovna metoda s kojom će se usporediti sva tri algoritma uzima se Strategija jednostavnog pomičnog prosjeka (SMA, engl. simple moving average strategy). SMA predikciju za dan d radi tako što računa prosječnu vrijednost kriptovalute između *d-prozor* i *d-1*. SMA je prosjek vrijednosti kriptovalute tijekom određenog vremenskog razdoblja i tehnički je indikator koji pomaže kod određivanja hoće li se vrijednosti nastaviti ponašati po trendu ili će doći do preokreta trenda (iz rastućeg u padajući ili obratno).

Prve dvije metode koriste XGBoost[4], skalabilni sustav strojnog učenja temeljen na stablu odlučivanja, a treća metoda se bazira na algoritmu dugog kratkoročnog pamćenja (LSTM, engl. Long short-term memory) za ponavljajuće neuronske mreže. Kod prve i druge metode model je skup regresijskih stabala izgrađen korištenjem XGBoost algoritma, a njegova svojstva su karakteristike valute u vremenu između *t-w* i *t-1*, gdje t označava vrijeme, a *w* cilj povrat ulaganja u vremenu *t*, gdje je *w* parametar koji se treba odrediti. Karakteristike valute su: vrijednost, udio na tržištu, tržišni kapital, povrat ulaganja, tržišni volumen i rang, a njihova svojstva u istraživanju su prosjek, standardna devijacija, medijan, zadnja vrijednost te raspon između prve i zadnje vrijednosti. Razlika između ove dvije metode je što prva koristi jedan zajednički model tj. cijelo tržište kriptovaluta za opisivanje

4 XGBoost je optimizirana distribuirana programska biblioteka za pojačavanje gradijenta dizajnirana da bude učinkovita, fleksibilna i prenosiva. Primjenjuje algoritme strojnog učenja u softverskom okviru Gradient Boostinga. XGBoost izvodi paralelno pojačavanje stabla koje na brz i točan način rješava mnoge probleme iz znanosti o podacima.

promjena u vrijednosti svih kriptovaluta, dok druga ima različit model za svaku kriptovalutu, odnosno algoritam uči predviđati vrijednost kriptovalute na osnovu karakteristika svih valuta u sustavu. Treća metoda koristi LTSM mreže te je zbog toga dobar izbor za učenje dugoročnijih međuovisnosti, a kao i kod druge metode za svaku se valutu radi zaseban model koji predviđa povrat ulaganja valute na dan t prema vrijednostima povrata ulaganja te valute između dana *t-w* i uključivo *t-1* [35]. Parametri koji se mogu prilagođavati kod LTSM-a su broj točnih prolaza kroz set podataka tijekom uvježbanja (broj epoha, gdje epohu definiramo kao jedno iteriranje kroz čitav skup podataka za učenje), broj neurona u mreži te veličina prozora, a prilagodba se vrši optimizacijom predikcije budućih vrijednosti Bitcoina, Ripplea i Ethereuma. Ipak, nije uočen poseban utjecaj izbora broja neurona ili epoha, a za sve tri metode uočeno je da dulji period uvježbavanja ne vodi uvijek k točnijem predviđanju što se pripisuje činjenici da se tržište s vremenom mijenja i zbog nekih vanjskih faktora i sl. [35].

Pomoću svake metode se gradi portfelj preko kojeg će se procjenjivati učinkovitost i to na način da se početni kapital jednoliko dijeli između prvih *n* valuta za koje se predviđa pozitivan povrat ulaganja.

Dobiveni su sljedeći zaključci:

- Prva i druga metoda pokazale su se najučinkovitije kod kratkoročnih vremenskih okvira (5 ili 10 dana), a LSTM kod dužih (oko 50 dana), iz čega se može zaključiti da su one bolje u iskorištavanju kratkoročnih međuovisnosti, dok LSTM bolje iskorištava dugoročnije međuovisnosti i stabilniji je kad je u pitanju volatilnost.
- Sve tri metode bile su bolje od osnovne strategije SMA, a sve su strategije uspjele stvoriti profit.
- Kao najvažniji faktori za predikciju ponašanja vrijednosti kriptovalute prepoznati su njena vrijednost i povrat vrijednosti u posljednjih nekoliko dana uoči predikcije.
- Paralelno predviđanje trendova na cjelokupnom tržištu i trendova vezanih uz određenu kriptovalutu mnogo je složenije nego samo predviđanje za jednu određenu kriptovalutu.

Treba naglasiti određena ograničenja ove studije, kao što su zanemarivanje dnevnih fluktuacija vrijednosti kriptovaluta, netočna pretpostavka da je dostupnost Bitcoina neograničena i zanemarivanje utjecaja samog trgovanja. Također, ovo istraživanje nije uzelo u obzir činjenicu da postoji više burzi za trgovanje kriptovalutama koje imaju različite vrijednosti po kojima trguju kriptovalutama te one naplaćuju određene transakcijske troškove (0.1% - 0.5% iznosa). Ako su ti troškovi do 0.2%, sve strategije i dalje ostvaruju profit. Nije loše ovdje istaknuti da je značajna prednost LSTM-a je što on ostvaruje profit do čak 1% transakcijskih troškova.

U prvom izdanju knjige *„Essentials of Blockchain Technology“* objavljen je rad s naslovom *„Prediction of Cryptocurrency Market Price Using Deep Learning and Blockchain Information“* [36] koji je nastao kao rezultat istraživanja na Sveučilištu u Melbourneu. Ovo istraživanje posebno je zanimljivo jer kod predikcije u obzir uzima do sad nerazmatrane informacije blockchaina. Nažalost, analizirane su vrijednosti samo dviju najvažnijih kriptovaluta, Bitcoina i Ethereuma. Korištene su informacije u razdoblju od 2015. do 2018. godine, otkad su na tržištu obje kriptovalute.

Blockchain informacije korištene u predikciji su:

- tržišna vrijednost,
- težina, odnosno vrijednost koja opisuje koliko je teško izrudariti novi blok,

- brzina hashiranja,
- prosječna veličina bloka (MB),
- broj potvrđenih transakcija,
- zarada rudara, zbroj ukupne nagrade za izrudaren osnovni blok i naknade za transakciju dodijeljene rudaru.

Nad tim je informacijama provedena i statistička analiza za dobivanje njihovog geometrijskog prosjeka, geometrijske standardne devijacije, minimuma, maksimuma i intervala pouzdanosti.

80% podataka iskorišteno je u fazi uvježbavanja, a ostatak za testiranje.

Da bi se dodatno naglasila važnost blockchain informacija kod predikcije budućih vrijednosti, korištena su dva različita seta podataka, jedan sa iznad navedenim blockchain informacijama i podacima o kriptovaluti poput njene vrijednosti i trenda (skup podataka A), a drugi samo sa podacima o kriptovaluti (skup podataka B).

Cilj istraživanja je utvrditi može li kombinacija dubokog učenja i blockchain informacija biti učinkovito rješenje za predikciju vrijednosti kriptovaluta. Duboko učenje, da bi bilo učinkovito, zahtjeva puno označenih podataka koji se mogu ekstrahirati iz blockchaina pa bi zato bilo još bolje kada bi se u istraživanju koristili svi dostupni podaci o Bitcoinu (od 2009.), što se zbog usporedbe s Ethereumom odbacilo. Primijenjen je poseban model dubokog učenja, model dugog kratkoročnog pamćenja (LSTM), koji se pokazao najučinkovitiji baš kod ovakvih serija vremenski povezanih podataka i, kao što je spomenuto ranije, LSTM mreže mogu prepoznati dugoročne međuovisnosti. Ćelije u ovakvoj mreži sadrže vrata za pamćenje i zaboravljanje što im omogućava biranje koje će informacije biti proslijeđene dalje uzevši u obzir njihovu važnost i kvalitetu pa slabiji signali mogu biti blokirani čime se izbjegava problem nestajućeg gradijenta. U neuralnim mrežama koje se sastoje od više slojeva stanja iz prethodnog koraka utječu na sljedeće stanje tako da se gradijenti prethodnog koraka djelovanjem funkcija na intervalu [-1,1] i matričnim množenjem kombiniraju s ulazom. Na taj se način ubrzo dođe do malih vrijednosti ili nule, pa se može reći da gradijent nestaje. Kod ovakvih mreža gradijent određuje koliko mreža nauči tijekom uvježbavanja, a ukoliko uvježbavanje nije dovoljno predikcije će biti neprecizne [37].

Broj epoha promjenjiv je parametar LSTM-a. Za broj epoha iznad 30 postiže se visoka preciznost predviđanja kod skupa podatka A za Bitcoin (85.9785%), a preciznost iznad 90% već za 110 epoha, dok kod skupa podataka B tek na 130 epoha. U slučaju Ethereuma prosječna razlika u preciznosti između ova dva skupa podataka iznosila je oko 2% (s manje epoha čak i 5%, a iznad 100 epoha između 1% i 2%, što je svejedno značajan iznos). Time se dokazuje da su informacije koje se nalaze u blockchainu vrlo značajne u predikciji vrijednosti kriptovaluta i trebalo bi dodatno poraditi na tome da se što veća količina navedenih informacija iskoristi kod predikcije. Uspoređene su i točnosti predikcije na Bitcoin i Ethereum s jednakim brojem epoha (100) i pokazano je da je predviđena vrijednost za Ethereum bila vrlo blizu stvarne vrijednosti, dok za Bitcoin to nije slučaj (bilo je određenih odstupanja). Za Bitcoin prosječna stopa predviđanja iznosi 88.84%, a za Ethereum 91.70%. Također, uočeno je da Ethereum ima manju stopu pogreške kod predviđanja (Bitcoin: 0.0054, Ethereum: 0.0047) [36]. Time je pokazana i veća dugoročna, ali i kratkoročna reaktivnost Ethereuma (njegove vrijednosti) na stanje na tržištu, odnosno njegove reakcije na stanje na tržištu veće su u usporedbi onih Bitcoina. Na kraju istraživanja donesen je zaključak da blockchain informacije daju visoku sigurnost u digitalnom trgovanju i pokazuju veliku preciznost kada se upotrebljavaju za predikciju.

U moru informacija koje je moguće izvući iz blockchaina neke su bitnije od drugih. U *„Combining Blockchain and Machine Learning to Forecast Cryptocurrency Prices“* [38] prvo su uzete sve informacije, a zatim je testirana njihova korisnost za predikciju te su izbačene one redundantne. Na početku skup podataka ima 17 značajki. Postoji nekoliko redundantnih značajki pa je tako primijećeno da je, ako postoji vrijednost u dolarima, vrijednost u Bitcoinima redundantna. Također, postoje neke značajke koje ne donose nove informacije o samim transakcijama (npr. URL adresa kriptovalute) pa su i takva svojstva eliminirana. Možemo spomenuti da je ipak ponekad korisno analizirati popis URL adresa kriptovaluta, odnosno novčanika u koje se šalju. Pomoću popisa može se prepoznati koje adrese drže najviše valuta, a kakve transakcije vrše te adrese može nam biti indikator o trendovima na tržištu što u ovom istraživanju, kao ni u drugima javno dostupnim istraživanjima još nije analizirano. Na kraju filtriranja podataka ostaju samo četiri značajke: volumen, tržišna kapitalizacija, promjena kapitalizacije u danu i maksimalna ponuda kriptovalute u cirkulaciji. Volumen možemo definirati kao količinu kriptovalute kojom se trguje na određeni dan, tržišna kapitalizacija umnožak je dnevnog volumena i dnevne vrijednosti kriptovalute, a promjena kapitalizacije pokazuje razliku u tržišnoj kapitalizaciji u odnosu na prethnodni dan. Maksimalna ponuda je ukupna količina kriptovalute dostupne u opticaju. Sama po sebi ponuda nije ovisna o vrijednosti kriptovalute i kod Bitcoina je ona ograničena na 21 milijardu Bitcoina, ali postoje i određene kriptovalute koje nemaju gornju granicu. Volumen kriptovaluta također nije ovisan o vrijednosti kriptovalute, ali tržišna kapitalizacija kao umnožak njenog volumena i vrijednosti jest te se može reći da ona predstavlja ukupnu vrijednost kriptovalute.

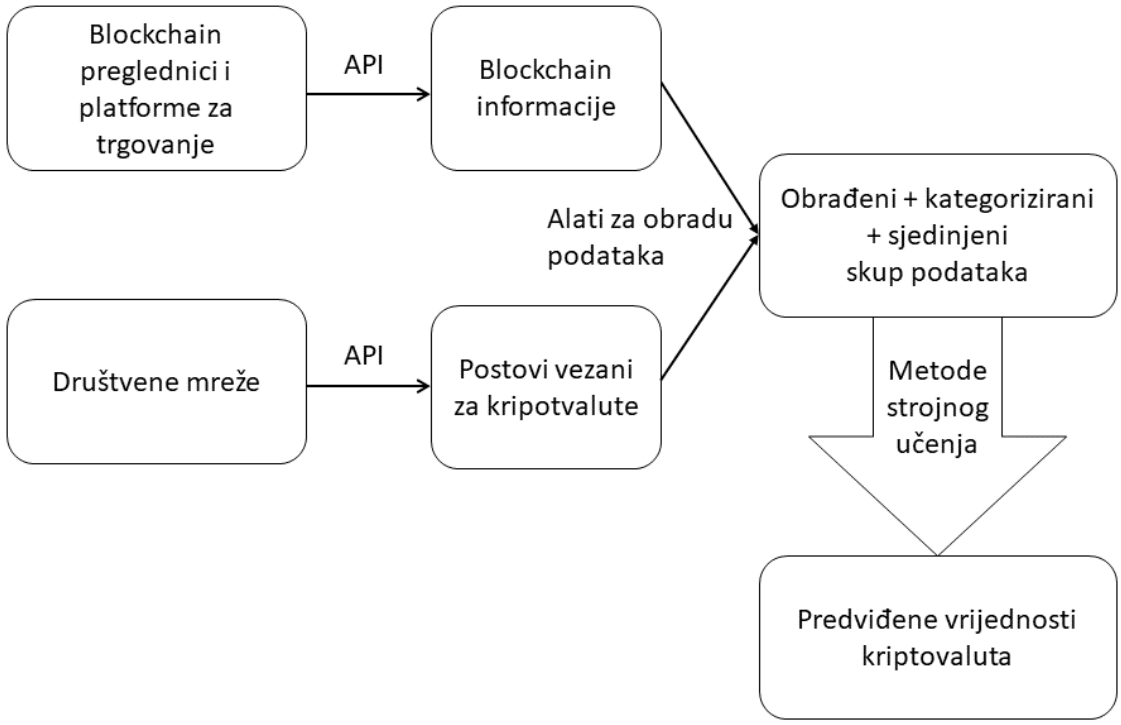
Provodi se analiza glavnih komponenti (PCA, engl. Principal component analysis) koja se koristi za redukciju dimenzije podataka i interpretaciju podataka [39]. Zatim se mjeri i F-Score, mjera točnosti koja uravnotežuje preciznost i opoziv. Na kraju su određena tri najvažnija svojstva navedena prema padajućoj važnosti: tržišna kapitalizacija, volumen i dnevna promjena kapitalizacije.

U ovom su istraživanju korištena četiri pristupa predikciji vrijednosti. Prvi je pristup fokusiran samo na izabrana svojstva i koristi se regresijski pristup zbog činjenice da je vrijednost kriptovalute kontinuirana varijabla. Izabran je regresijski model pojačavanja gradijenta. Pojačavanje gradijenta je tehnika strojnog učenja za regresiju, klasifikaciju i druge zadatke, koja stvara model predviđanja u obliku skupa slabih modela predviđanja, obično stabala odlučivanja. Kada je stablo odluke slabo u učenju, rezultirajući algoritam naziva se stablom pojačanih gradijenta i obično daje bolje rezultate u odnosu na nasumičnu šumu. Izrađuje se model u fazama i generalizira ih se dopuštajući optimizaciju proizvoljne diferencijabilne funkcije gubitka [40].

Ovaj model je očekivano dao vrlo neprecizna predviđanja, ali je koristan za usporedbu s drugim modelima. Jasno je da za preciznu predikciju nije dovoljno gledati samo svojstva jer na vrijednost može utjecati stav javnosti prema tržištu ili zamah koji se gradi na tržištu ako kriptovaluta raste u očima ulagača. Za usporedbu modela koriste se srednja kvadratna pogreška (MSE, engl. Mean squared error) i srednja prosječna pogreška (MAE, engl. Mean average error) jer se kod njihovog izračuna gleda na cjelokupni period izračuna pa su dobar pokazatelj za procjene vrijednosti bilo koje kontinuirane varijable. MSE mjeri prosječnu kvadratnu razliku između stvarne vrijednosti i procijenjenih vrijednosti, a MAE prosječnu pogrešku između tih istih vrijednosti. Drugi testirani model je algoritam hrbatne regresije [41], (engl. ridge regression) koji se koristi kad postoji gotovo linearan odnos između nezavisnih varijabli. Hrbatna regresija je tehnika za analizu kolinearnih podataka. Kolinearnost označava postojanje gotovo linearnih odnosa među nezavisnim varijablama. U ovakvim situacijama, procjene najmanjeg kvadrata su nepristrane, ali njihove su varijance velike pa mogu biti daleko od prave vrijednosti. Dodavanjem

stupnja pristranosti procjenama regresije, smanjuje se standardna pogreška [42]. Kod hrbatne regresije prvo se oduzima prosjek svake varijable i on se podijeli s vlastitom standardnom devijacijom. Niti ovaj model nije bio posebno učinkovit u predikciji iako se, najvjerojatnije zbog standardizacije samih podataka, pokazao nešto učinkovitiji od prvog modela. Uočeno je ipak malo poboljšanje preciznosti u slučaju volatilnih kretanja. Treći je model baziran na korištenju regresije vektora linearne potpore (engl. Linear support vector regression SVR). Zadnja metoda koristi sekvencijalnu neuralnu mrežu (SNN, engl. Sequential neural network i dobra je za predikciju kontinuiranih varijabli poput cijene, a prednost joj je i mogućnost istovremenog fokusa na srednju kvadratnu i srednju prosječnu pogrešku. Prema pretpostavkama zbog kontinuirane varijable ova je metoda trebala dati najbolje rezultate, ali na kraju je dala lošije rezultate nego druge, znatno jednostavnije metode.

Nakon četiri isprobana modela uočene su prednosti i mane svakog od njih te, kako bi se iskoristila činjenica da je svaki u nečemu precizniji, na kraju je isproban i grupni pristup koji uspoređuje predikcije sva četiri modela i iz njih donosi zajedničku predikciju. Prvi grupni pristup, glasovanje (engl. hard voting) više se koristi kod postavki klasifikacija, a u ovom slučaju implementirano je računanjem modaliteta predviđene vrijednosti sva četiri modela. Drugi je pristup koristio težine (engl. weights) koje se množe s predviđenom vrijednošću za kriptovalutu, a rezultati su bili najbolji kad je četvrti model koji je bio najbolji u predikciji imao najveću težinu. Ovaj se pristup pokazao najučinkovitiji, odnosno dao je najpreciznija predviđanja jer je u slučaju da je jedan algoritam iz nekog razloga jako promašio vrijednost njegov učinak na ukupnu predikciju bio umanjen. Kao najbolji model pokazao se model hrbatne regresije zbog standardizacije svojstava koja pomažu očuvati točnost predikcije i tokom volatilnih kretanja. Najlošijim su se pristupom i modelom pokazali glasovanje i SNN što je i očekivano za glasovanje jer je ono ipak namijenjeno problemima vezanim uz klasifikaciju, a ne uz predikciju. Neočekivano su loši rezultati SNN-a koji bi trebao biti optimiziran za predviđanje temeljeno na slijednim vremenskim podacima (engl. time series data), a na kraju se pokazao gori i od prva dva modela.



Izbor: vlastita izrada

Nakon ovih nekoliko istraživanja evidentno je da bi korištenje metoda strojnog učenja moglo biti, uz izbor prigodnih metoda i relevantnih blockchain informacija, pravi put za kvalitetnu i preciznu predikciju vrijednosti kriptovaluta što bi omogućilo porast povjerenja investitora, ali i običnih ljudi koji zbog volatilnosti nisu sigurni isplati li se uložiti novac u ovu relativno mladu tehnologiju. Autor ovog rada osobno smatra blockchain tehnologije jednom od najvećih ideja unazad nekoliko godina što će ostaviti dubok trag u svijetu, još više ukoliko priljev novca bude kontinuiran i dobro usmjeren.

U današnje vrijeme nemoguće je zanemariti i utjecaj društvenih mreža na vrijednost kriptovaluta pa smo tako mogli vidjeti neke, ne nužno kvalitetne projekte da su za samo par mjeseci ili čak dana doživjeli eksponencijalni rast vrijednosti, a nakon toga često jednako brz ili čak brži pad. Zbog toga bi bilo korisno u predikciju uključiti i stav javnosti (sentiment). Upravo će *„Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis“* [43] spojiti blockchain informacije sa stavom javnosti iznesenim na društvenoj mreži Twitter kako bi postigao još točnije rezultate predikcije. Ovo se istraživanje fokusira samo na Bitcoin što ga na neki način ograničava, ali ostavlja prostor za daljnje istraživanje i promišljanje.

Analiza sentimenta (engl. Sentiment analysis) koristit će se za određivanje stava javnosti na društvenim mrežama Twitter i Reddit, a to je zapravo sustav koji kategorizira mišljenja napisana u tekstualnom obliku u kategorije (npr. pozitivno, negativno, neutralno). Prije provedbe SA-a prvo je potrebno prikupiti podatke što je odrađeno putem Twitterovog API-ja i alata Tweepy koji omogućava filtriranje tvitova po hashtagu ili riječi, a za ovo istraživanje, kako bi količina podataka bila dostatna i ne preopširna, zadržani su samo tvitovi sa riječima Bitcoin i BTC. Nakon toga, kako bi se dodatno očistili podaci provedeni su tokenizacija, odnosno razdvajanje tvitova na karakteristične riječi te uklanjanje emotikona i sličnih „viškova“, eliminacija riječi koje ne nose nikakvo mišljenje, odnosno značaj poput pomoćnih glagola ili čestica, i na kraju uklanjanje posebnih i ponavljajućih znakova korištenjem regularnih izraza. Na postovima pribavljenim s Reddita učinjeno je isto. Nadalje, provedena je analiza sentimenta, odnosno svaki je tvit ili post svrstan u kategoriju pozitivan/negativan/neutralan pomoću alata Textblob [44] i API servisa Haven On Demand [45].

Idući korak bilo je prikupljanje blockchain informacija iz četiri različita izvora od kojih je svaki pridonio dodatnim informacijama. Iz Coinmarketcapa i Blockchain Info-a prikupljeno je 11 ključnih svojstava koja su slična kao i u prethodnim istraživanjima (vrijednost, volumen, dostupna ponuda Bitcoina na tržištu, ukupna količina Bitcoina, tržišna kapitalizacija, itd.). Iz Bitstampove baze prikupljeno je još 10 svojstava vezanih za transakcije (najveća vrijednost u minuti, zadnja vrijednost u minuti,...), a iz Coinbasea su prikupljeni podaci iz blockchaina u stvarnom vremenu. Na samom kraju obrade podataka ova su dva seta spojena korištenjem softverske biblioteke Pandas za Python, a sjedinjeni skup podataka sadržavao je vrijednost, sentiment i vremenska svojstva.

Idući korak je predikcija koja je u ovom istraživanja uključivala dvije metode. Prva metoda je već opisani LSTM koji svojim djelovanjem omogućuje mreži da sazna više o vremenskim koracima održavajući pogrešku stabilnijom. Odnedavno su ponavljajuće neuralne mreže (RNN, engl. Recurrent Neural Network) zbog činjenice da svaki neuron može pristupiti vlastitoj unutarnjoj memoriji kako bi spremio podatke o prethodnom unosu, postale sve popularnije za korištenje sa sekvencijalnim podacima kao što je npr. vremenski slijed blockchain informacija. Ipak, treba znati da su kod RNN-a slojevi i vremenski koraci isprepleteni pa je moguća pojava eksplodirajućeg ili nestajućeg gradijenta.

Drugi promatrani model je ARIMA ili autoregresijski integrirani pokretni prosjek (engl. Auto regression integrated moving average) kojim je moguće procijeniti trend (srednju vrijednost tijekom vremenskoj perioda koji može biti uzlazni/silazni), sezonalnost, cikluse, odstupanja i slično. Takvi se podaci različitim metodama uklanjaju pa je lakše primijetiti dinamična vremenska ponašanja u vremenskom slijedu.

Kako bi se ove dvije metode usporedile, računa se standardna devijacija predikcijskih grešaka (engl. Root mean-square error, RMSE). Za LSTM on iznosi 198.448 kod jednog svojstva, a 197.515 kod više njih, a za ARIMA-u 209.263. Valja uočiti da se javila razlika od gotovo 6% što je vrlo značajna razlika. Kod LSTM-a gubici su minimalni. Može se zaključiti da model LSTM bolje odgovara predikciji vrijednosti kriptovalute, što se moglo i pretpostaviti znajući činjenicu da je LSTM dobar za predviđanje većih fluktuacija u vremenskim slijedovima podataka, a to svakako vrijedi za volatilnu vrijednost kriptovaluta.

4.3. Prijedlozi za proširenje

Nakon što su u radu kratko predstavljena znanstvena istraživanja na temu predikcije, dobro je predstaviti i neke buduće smjerove u kojima se može napredovati, tj. razvijati kvalitetu predikcije pomoću metoda strojnog učenja jer, s obzirom na brzinu razvoja i sve veći fokus na ove dvije discipline, njihovo ispreplitanje će biti sve češće i korisnije.

Ranije istražene metode u strojnom učenju i predviđanju se oslanjaju na skupove podataka organizirane u tablice. Međutim, to nije jedini oblik čuvanja podataka i nije uvijek najpraktičniji. Na primjer, moguće je podatke prikazati u obliku grafa što postaje sve popularnije zbog njihovog integriteta, skalabilnosti i boljih performansi i učinkovitosti te činjenice da rastuća kompleksnost i količina podataka ne utječu negativno na ovakve baze podataka [46]. Prirodno bi bilo prikazati transakcije u blockchainu kao graf gdje čvorovi (eng. node) predstavljaju adrese, a bridovi (engl. edge) predstavljaju transakcije. Ranije spomenute metode nisu predviđene za rad s takvom vrstom podataka pa je potrebno naći druge. U velikim tvrtkama koje u svom radu produciraju mnogo podataka (Uber, Google) sve se češće koriste grafičke neuralne mreže (GNN, engl. Graph Neural Network) koje se fokusiraju na modele koji učinkovito djeluju na podacima strukturiranim u grafove. Moglo bi se dakle iskoristiti GNN i ranije opisani graf s transakcijama za uočavanje utjecaja transakcija (pritok novca i sl.) na vrijednost kriptovaluta.

Istaknimo da još jedna situacija može izazvati problem kod predikcije. U primjerima istraživanja najčešće su uzimani setovi podataka u trajanju od nekoliko godina i to za godine kada postoje podaci o svim kriptovalutama koje su predmet predikcije. Kada bismo radili predikciju za Bitcoin i neku novu kriptovalutu nastalu 2020. godine, u tom bi slučaju svi podaci o Bitcoinu od 2009. do 2020. ostali neiskorišteni, a zasigurno bi pridonijeli kvaliteti predikcije. Također, postoji mogućnost da je za skup podataka za period od godine dana nedovoljan da bi neuralne mreže mogle generalizirati bilo kakvo znanje, odnosno donijeti predikciju [47].

Ovaj problem moguće je riješiti korištenjem generativnih modela. Generativni modeli su modeli koji generiraju umjetne podatke koji odgovaraju (rađeni su prema) distribuciji skupa podataka za uvježbavanje (engl. training dataset). Na taj bismo način stvorili nove transakcije koje bi se ponašale u skladu sa prijašnjima te bi spajanjem stvarnih i generiranih podataka stvorili skup podataka dostatan za uvježbavanje neuralne mreže. Ovakvi modeli već se uspješno koriste za vremenski povezane skupove financijskih podataka.

U istraživanju [38] posebna je pažnja posvećena izboru relevantnih svojstava koja i jesu jedna od najvažnijih komponenti uspjeha metoda strojnog učenja pa ih je za uspješnost predikcije potrebno izvući iz blockchaina i što bolje izabrati najrelevantnija svojstva. Dobar izbor je pogotovo bitan jer još ne postoji dovoljno saznanja kako koje pojedinačno svojstvo utječe na vrijednost, a ne postoji ni konsenzus o najvažnijim svojstvima. Na primjer, bilo bi mnogo lakše kada bismo znali da kod predikcije budućih vrijednosti najveći značaj ima kapitalizacija tržišta, a ne neko drugo od par desetaka potencijalnih svojstava koja se mogu ekstrahirati iz blockchaina.

Učenje reprezentacijom (engl. representation learning) fokusira se na automatizaciju učenja reprezentacija ili svojstava kojima bi se povećala efektivnost modela. U tom slučaju ne oslanja se na manualni izbor svojstava pri modeliranju, nego se svojstva direktno izvlače iz neoznačenih setova podataka što bi u konkretnom slučaju značilo da se analizira blockchain i identificiraju se potencijalna svojstva koja bi mogla poslužiti u analizi.

4.4. Potencijalni problemi

Uz sve prednosti strojnog učenja čini se da je nemoguće pogriješiti u predikciji, ali opet, iz samih dosadašnjih analiza, vidljivo je da nije tako. Za sad ne postoji tehnologija koja bi bila 100% točna u predikciji, a i kad bi se takva situacija dogodila, stav je autora da se to ne bi dobro odrazilo na tržište te bi potencijalno imalo negativne posljedice na cijeli financijski sustav. Postoji nekoliko razloga zbog kojih strojno učenje nije sasvim precizno u svojim procjenama.

Strojno učenje, kao što smo mogli vidjeti, može naći modele koji točno predviđaju velik dio kretanja vrijednosti neke valute. Ipak, moguće su situacije u kojima se u malim dijelovima koji nisu točno predviđeni dogode velike promjene. Na primjer, predikcija se radi za svaki sat tokom dana, postoji prepoznatljiv uzorak za 15 sati tokom kojeg vrijednost poraste za 5%, međutim u 16. satu se, npr. zbog nekog vanjskog uzorka, dogodi nagla promjena i vrijednost u nekoliko minuta padne, na primjer, više od 20%, a što model nikako nije mogao predvidjeti prepoznavši prethodni uzorak koji se čak može opet pojaviti nakon ovog skoka. Ovaj problem naziva se pristranost prepoznavanja uzoraka (engl. pattern matching bias) [48].

Slično se događa i kod overfittinga, gdje modeli vrlo učinkovito rade sa povijesnim skupovima podataka, ali ne i sa novim podacima. Zbog toga što je model uvježban da savršeno odgovara zadanim podacima, ali ukoliko dođe do promjene u ponašanju, unatoč tome što je za tu promjenu ponovno moguće uočiti uzorak, predikcija će nadalje biti sve nepreciznija. Kako bi se ovaj problem izbjegao potrebno je pronaći generalizirane uzorke na što više tržišta sa različitim uvjetima [48].

Na kraju, ipak je glavna svrha predikcije ostvarenje zarade. Kako bi se to ostvarilo potrebno je predvidjeti i kada će određena kriptovaluta prestati s rastom što je zapravo najveći izazov jer kada postoji trend povećanja vrijednosti (engl. bull run), osobito ako je on velik i brz kao što je često slučaj kod kriptovaluta, najčešće slijedi i puno brži i veći pad. Ako se kriptovalutu proda prerano, gubi se dio zarade, a ako se ona proda prekasno, mogući su gubici. Čovjeku je vrlo teško predvidjeti kada je pravo vrijeme za prodaju (treba biti što bliže samom vrhu vrijednosti) jer se često vodi pohlepom koja smanjuje objektivnost. Kod strojnog učenja ovaj problem ne postoji, ali zbog iznad opisanih problema nije vjerojatno da će model biti savršeno uspješan u tempiranju prodaje.

5. Primjene

Na kraju rada ponudit ćemo neke primjere korištenja strojnog učenja u predikciji iz stvarnog života.

5.1. Alati

Nomics je alat koji nudi analitiku blockchain informacija poput blockchain explorera, ali i izgradnju strategija upravljanja portfeljem baziranih na strojnom učenju jer je moguće vidjeti predikcije budućih vrijednosti kriptovalute za sedam dana unaprijed i po tome izabrati u koje kriptovalute želimo uložiti. Prikaz za Bitcoin vidljiv je na Slika 5, gdje crvena linija predstavlja stvarnu, a isprekidana linija procijenjenu vrijednost. Za predikciju se koristi LSTM algoritam pomoću podataka iz Open-high-low-close-volume (Početna vrijednost-najviša vrijednost-najniža vrijednost-zadnja-volumen, OHLCV) dijagrama[5] u obliku svijeća[6] [49]. Nakon testiranja alata primijećeno je da točnost predikcije degradira što je moguće pripisati overfittingu, ali općenito su rezultati upravljanja cijelim portfeljem bili zadovoljavajući jer unatoč odstupanjima pojedinačnih vrijednosti predikcije, prosječni rezultat cijelog portfelja bio je u plusu [50].

Ovaj alat, kao i druge slične, trebalo bi koristiti s oprezom jer ne mogu garantirati točnost, ali u svakom slučaju mogu biti dobar indikator trendova na tržištu i u kombinaciji s drugim faktorima pomoći u ostvarenju što većeg profita.

Izvor: https://nomics.com/assets/btc-bitcoin#markets , pristupljeno 18. svibanj 2021.

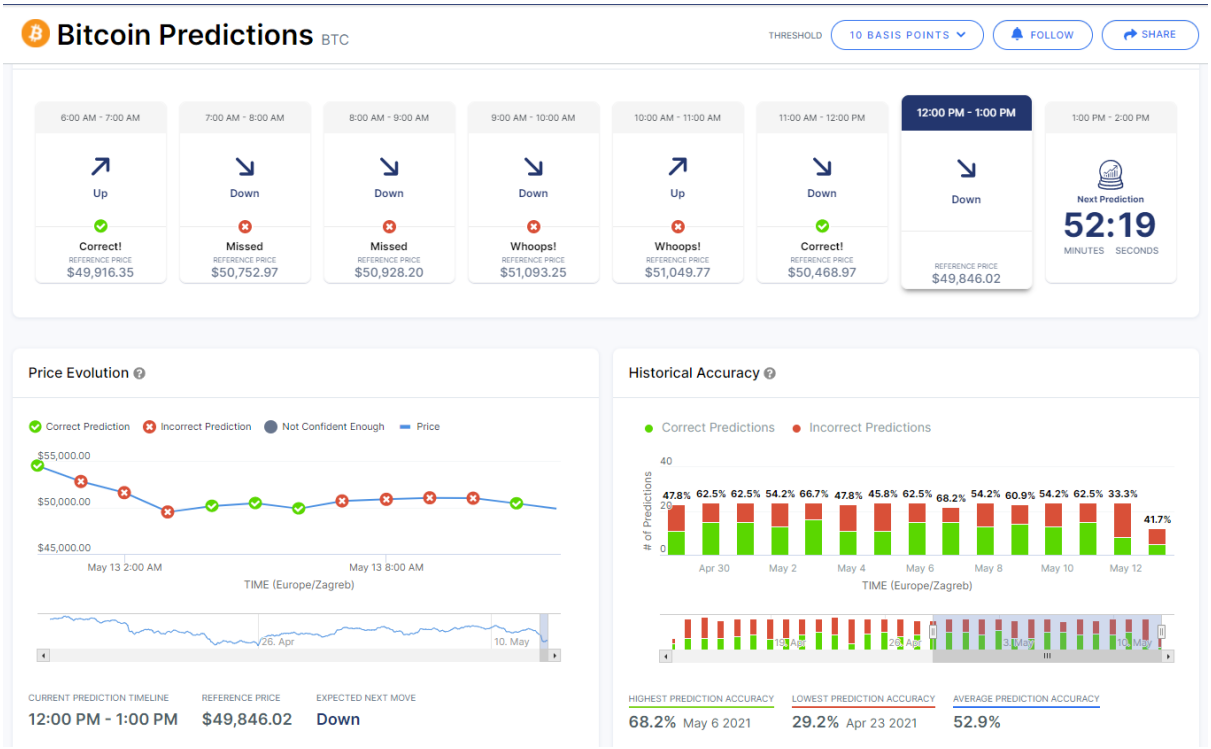
Drugi potencijalno koristan alat je IntoTheBlock koji, za razliku od Nomicsa, generira predviđanja za svaki sat. Nudi i brojne dodatne sadržaje poput analize profitabilnosti, indikatora adresa i transakcija te analize profila investitora koje je moguće izvući iz blockchaina. Zbog kvalitete grafova i analiza ostvarena su i partnerstva sa najvećim burzama kriptovaluta i pretraživačima poput Binancea, CoinMarketCapa, CoinGecka, eTora i drugih. Postoji besplatna verzija koja nudi samo neke mogućnosti i predikciju ograničenu na Bitcoin i verzija koja se

5 Dijagram koji se koristi za prikaz kretanja vrijednosti dionice (ili neke druge imovine kojom se trguje) na tržištu kroz vrijeme. Iz njega je za niz vremenskih intervala moguće isčitati vrijednosti na otvaranju i zatvaranju intervala, najviše i najniže vrijednosti te volumen trgovanja

6 Vrsta grafikona kojeg često koriste trgovci i financijski profesionalci koji ukazuje na cijeli raspon vremena trgovanja (od početne cijene do zaključne cijene), uz pomoć osjenčanih pravokutnih oblika

plaća sa rastućom listom kriptovaluta [50].

Za pronalazak složenih nelinearnih veza između varijabli koje bi bilo teško prepoznati drugim metodama koriste se duboke neuralne mreže za koje se tvrdi da imaju veću sposobnost donošenja generaliziranih odluka o tržištu u usporedbi s drugim metodama strojnog učenja, ali duboke su mreže skuplje za skaliranje, njihova je izgradnja složena i teške su za interpretaciju.



Izvor: https://app.intotheblock.com/predictions/BTC/btcOb60mAvg60m, pristupljeno 2.lipanj 2021.

Na Slika 6 vidljiva je predikcija budućih vrijednosti za Bitcoin koja pokazuje da je prosječna točnost 52.9%.

5.2. Botovi

Internet je vrlo širok i pun informacija pa je teško pronaći vjerodostojne podatke o kvaliteti botova, većina ih se plaća unaprijed i na njihovim se stranicama navode rješenja za velikom i brzom zaradom.

Posljednjih godina proces automatizacije postaje sve bitniji i zahvaljujući razvoju alata za automatizaciju sve je više procesa moguće automatizirati. Sve se više koriste botovi, automatizirani programi koji obavljaju ponavljajuće zadatke učinkovitije od ljudi i time im olakšavaju poslove. Isto vrijedi i za botove za trgovanje kriptovalutama koji pomoću umjetne inteligencije utemeljene na unaprijed određenim parametrima trguju na tržištu kriptovaluta. Prednost ovakvog trgovanja je što rade učinkovito i objektivno te u bilo koje doba dana (24/7) obavljaju trgovanje [51].

Izdvaja se projekt B-cube.ai [52] koji nudi sveobuhvatnu platformu namijenjenu trgovcima kriptovalutama baziranu na umjetnoj inteligenciji. Platforma je nastala u suradnji sa pariškim sveučilištem i postoji opsežna dokumentacija i plan rada, kao i vlastita kriptovaluta koja će služiti za rad platforme. Pomoću B-cube.ai-ja trebalo bi biti moguće koristiti već gotove botove, izrađivati svoje botove koristeći već gotove modele ili uvesti vlastite algoritme (npr. implementaciju GNN-a) koji bi se koristili u botovima ili prodavali na vlastitoj tržnici (engl.

marketplace). Postojeći botovi na platformi koriste u ovom radu opisane tehnike analize sentimenta na temelju podataka iz vijesti i društvenih mreža i strojno učenje kod kojeg svojstva mogu biti povezana sa analizom sentimenta, podacima o trgovanju ili blockchain informacijama [53]. Ova je platforma nastala 2019. godine, ponuda početne valute (engl. ICO, Initial coin offering) bila je u travnju 2021. pa se njezina šira primjena tek očekuje u budućnosti.

6. Zaključak

Prilikom pisanja ovog završnog rada autor je pokušao ponuditi kratki uvod u svijet kriptovaluta predstavivši najvažnije mehanizme i valute. Stav je autora da će blockchain tehnologije u idućih nekoliko desetljeća oblikovati svijet u kojem živimo i da će njihova primjena biti proširena na sve više aspekata naših života: od poslovnog svijeta do svakodnevnih kupnji u lokalnom dućanu. Uz ove primjene, ipak je za sad najbitnija primjena blockchaina u svijetu kriptovaluta. Kako je tema ovog rada predikcija budućih vrijednosti kriptovaluta prvo je bilo važno opisati tržište i njegovu volatilnost. Tijekom pisanja ovog rada, između travnja i srpnja 2021. godine, vrijednost Bitcoina pala je sa 64 tisuće dolara na 29 tisuća dolara uz najveći dnevni pad od 35%. Ostale relevantne kriptovalute ponašale su se gotovo identično ili barem slično. Kako je autor već tada bio upoznat s alatima i metodama strojnog učenja koje se koriste za predikciju vrijednosti mogu se donijeti neki novi zaključci. Gledajući graf koji prikazuje predviđanja i stvarne vrijednosti kriptovaluta u ranije opisanom alatu Nomics koji koristi LSTM algoritam, možemo zaključiti da je na dane iznenadnih velikih promjena (o proučavanom periodu te promjene su padovi) alat zakazao, ali da se njegova učinkovitost popravlja čim se pojavi neka vrsta trenda. Odnosno, ako dođe do pojave negativnog trenda od nekoliko dana algoritam se relativno brzo adaptira (nauči) te dalje prilično točno predviđa vrijednosti do iduće velike promjene. Dio ovih situacija može se pripisati i overfittingu i pristranosti prepoznavanja uzoraka. Korisno je, ipak, naglasiti da velike promjene najčešće događaju zbog vanjskih i za računalo nepredvidljivih vijesti (npr. zabrana kriptovaluta u Kini, vijesti o hakerskim napadima i sl.) te je jako teško očekivati da će strojno učenje koje radi sa čvrstim činjeničnim podacima ekstrahiranim iz blockchaina moći predvidjeti ovakve situacije. Da bi se poboljšala učinkovitost kod ovakvih vijesti svakako je dobro u modele uklopiti i analizu sentimenta javnosti koja bi mogla, ako se pravilno postavi, barem malo kompenzirati vanjske utjecaje na vrijednost. U većini radova proučavanih za potrebe pisanja ovog završnog rada donesen je zaključak da je LSTM algoritam najučinkovitiji za predviđanje vrijednosti što je bilo moguće i pretpostaviti jer je njegova učinkovitost već dokazana u trgovanju na burzama.

Autor ovog rada vjeruje u strelovit razvoj strojnog učenja i tehnologija poput big date koje će posljedično svojim razvitkom omogućiti i sve točnije predikcije. Kroz sve širu primjenu i rasprostranjenost kriptovaluta stvarat će se mnogo više blockchain informacija koje će se moći koristiti u predikciji, a sve naprednije neuronske mreže i algoritmi moći će ih procesuirati što bi moglo dovesti do povećanja preciznosti predikcije.

Mogućnost predikcije budućih vrijednosti kriptovaluta korištenjem blockchain informacija i metodama strojnog učenja je sa znanstvene strane i kao analitički alat (kod procjene kriptovaluta) već danas prilično dobro razvijena te daje prilično točne rezultate u odnosu na neke starije strategije trgovanja i predviđanja. No, zbog vanjskih faktora nerealno je očekivati apsolutnu točnost i naivno bi bilo misliti da je moguće stvoriti takav algoritam, a predikciju kod trgovanja treba gledati kao pomoć, a ne jedino rješenje.

Literatura

[1] J. Frankenfield, „Investopedia“, *Investopedia: Cryptocurrency*, ožu. 07, 2021. https://www.investopedia.com/terms/c/cryptocurrency.asp

[2] „Global crypto market crosses $2 trillion, bitcoin accounts for over 50% of the market cap“, *Business Insider*, tra. 07, 2021. https://www.businessinsider.in/business/news/global-crypto-market-crosses-2-trillion-bitcoin-accounts-for-over-50-of-the-market-cap/articleshow/81945257.cms.

[3] A. Narayanan, J. Bonneau, E. Felten, i A. M. S. Goldfeder, *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*. Princeton University Press, 2016.

[4] M. Crosby, P. P. Nachiappan, i S. V. V. Kalyanaraman, „BlockChain Technology: Beyond Bitcoin“, *Appl. Innov. Rev. Br*, sv. 2, 2016.

[5] „How Walmart brought unprecedented transparency to the food supply chain with Hyperledger Fabric“. https://www.hyperledger.org/learn/publications/walmart-case-study.

[6] „What are smart contracts on blockchain?“, *IBM*. https://www.ibm.com/topics/smart-contracts.

[7] „MIT Technology Review: Blockchain“, u *Pokušaj pristupa*, 2021. [Na internetu]. Dostupno na: https://www.technologyreview.com/topic/blockchain/.

[8] I. Hasib, „Blockchain & Data Integrity“, *Medium*, 2019, [Na internetu]. Dostupno na: https://medium.com/@hmishfer17/blockchain-data-integrity-e70e17cac086.

[9] „Blockchair“, *Blockchair.com*. https://blockchair.com/

[10] „Blockchain Explorer“. https://www.blockchain.com/explorer

[11] „btc.com“, *BTC.com*. https://btc.com/

[12] « »Hash, „Binance Academy“. 2021. [Na internetu]. Dostupno na: https://academy.binance.com/en/glossary/hash.

[13] „Public and Private Keys“, *Gemini.com*. https://www.gemini.com/cryptopedia/public-private-keys-cryptography#section-what-does-it-mean-to-digitally-sign-a-transaction.

[14] «, „CoinmarketCap“, *CoinmarketCap*. https://coinmarketcap.com/all/views/all/.

[15] S. Vrbanus, „U Hrvatskoj prvi puta automobil plaćen bitcoinom“, pros. 02, 2020. https://www.bug.hr/kriptovalute/u-hrvatskoj-prvi-puta-automobil-placen-bitcoinom-17545.

[16] S. Nakamoto, „Bitcoin: A Peer-to-Peer Electronic Cash System“. 2009.

[17] L. Conway, „Blockchain Explained“. stu. 17, 2020. [Na internetu]. Dostupno na: https://www.investopedia.com/terms/b/blockchain.asp.

[18] „Bitcoin Vocabulary“. https://bitcoin.org/en/vocabulary#mining.

[19] A. M. Antonopoulos, *Mastering Bitcoin*. O’Reilly Media, Inc, 2017.

[20] E. Hong, „How Does Bitcoin Mining Work?“, *Investopedia*, svi. 04, 2021. https://www.investopedia.com/tech/how-does-bitcoin-mining-work/.

[21] N. Borate, „What is bitcoin halving and will it affect the rate?“, sij. 11, 2021. https://www.livemint.com/money/personal-finance/what-is-bitcoin-halving-and-will-it-affect-the-rate-11610295621496.html.

[22] „NFT“. https://ethereum.org/en/nft/.

[23] V. Buterin, „Ethereum white paper“. 2013.

[24] A. Hetman, „DApps blockchain applications“. https://espeoblockchain.com/blog/dapps-blockchain-applications.

[25] „Chainlink – veza vanjskog svijeta s blockchainom“. lip. 12, 2020. [Na internetu]. Dostupno na: https://mojkripto.com/chainlink-veza-vanjskog-svijeta-s-blockchainom/.

[26] S. Ellis i A. J. S. Nazarov, „ChainLink: A Decentralized Oracle Network“. 2017.

[27] „Što je volatilnost cijene?“, *Admiral markets*, ruj. 04, 2020. https://admiralmarkets.com/hr/education/articles/forex-basics/sto-je-volatilnost.

[28] G. Gurbacs, „Bitcoin: Less Volatile Than Many S&P 500 Stocks?“ stu. 20, 2020. [Na internetu]. Dostupno na: https://www.vaneck.com/us/en/blogs/digital-assets/bitcoin-less-volatile-than-many-sp-500-stocks/?country=us&audience=retail.

[29] D. Sajter, *Financijska analiza kriptovaluta u odnosu na standardne financijske instrumente,« u Financije - teorija i suvremena pitanja*. Osijek: Ekonomski fakultet u Osijeku, 2018.

[30] N. Reiff, „Why Bitcoin Price Predictions Are Unreliable“, *Investopedia*, tra. 02, 2021. https://www.investopedia.com/tech/why-bitcoin-price-predictions-are-unreliable/.
[31] J. Adkisson, „Why Bitcoin Is So Volatile“, *Forbes*, velj. 09, 2018. https://www.forbes.com/sites/jayadkisson/2018/02/09/why-bitcoin-is-so-volatile/?sh=7ccb342739fb.
[32] N. Reiff, „Why Bitcoin Has a Volatile Value“, *Investopedia*, lip. 16, 2020. https://www.investopedia.com/articles/investing/052014/why-bitcoins-value-so-volatile.asp.
[33] «, „Elements of AI“, *Elements of AI*, 2020. https://course.elementsofai.com/hr/4.
[34] „Strojno učenje vs. duboko učenje“, *PC Chip*. https://pcchip.hr/ostalo/tech/strojno-ucenje-vs-duboko-ucenje/.
[35] L. Alessandretti, A. E. Bahrawy, i L. M. A. A. Baronchelli, „Anticipating Cryptocurrency Prices Using Machine Learning“, *Complex. Svez*, str. 16, 2018.
[36] G. S. M. Halgamuge, *Prediction of Cryptocurrency Market Price Using Deep Learning and Blockchain Information: Bitcoin and Ethereum*. Taylor and Francis press, 2019.
[37] „Vanishing gradient problem“, *DeepAI ML glossary and terms*. https://deepai.org/machine-learning-glossary-and-terms/vanishing-gradient-problem
[38] K. Martin, I. Alsmadi, i M. R. M. Ayyash, „Combining Blockchain and Machine Learning to Forecast Cryptocurrency Prices“, 2020.
[39] N. B. B. Dalbelo-Bašić, *Otkrivanje znanja u skupovima podataka: bilješke sa predavanja*. Zagreb: FER, 2003.
[40] „Gradient boosting“, *Wikipedia: Gradient Boosting*. https://en.wikipedia.org/wiki/Gradient_boosting.
[41] J. Šnajder, *Linearan model regresije*. u Strojno učenje, 2017.
[42] L. NCSS, „Ridge Regression, Chapter 335“, u *u NCSS 2021*, 2021.
[43] S. M. R. A. M. Tarif, *Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis*. Gombak: International Islamic University Malaysia, 2020.
[44] „Textblob“, *TextBlob*. https://textblob.readthedocs.io/en/dev/.
[45] „Haven on Demand github repository“. https://github.com/HPE-Haven-OnDemand.
[46] „Graph databases explained“, stu. 07, 2019. https://www.ionos.co.uk/digitalguide/hosting/technical-matters/graph-database/.
[47] J. Rodriguez, „Five Machine Learning Methods Crypto Traders Should Know About“, *CoinDesk*, lis. 16, 2020. https://www.coindesk.com/five-machine-learning-methods-crypto-traders-should-know-about.
[48] „Crypto Trading Bot Service“, *Crypto ML: Trading Bot Service*. https://crypto-ml.com/crypto-trading-bot/.
[49] „Candlestick Chart / Grafikon oblika svijeće“. [Na internetu]. Dostupno na: https://www.fortrade.com/hr/glossary/candlestick-chart.
[50] „Top Machine Learning Products for Cryptocurrency Price Predictions“, *Top Machine Learning Products for Cryptocurrency Price Predictions*, ruj. 02, 2020. https://blog.shrimpy.io/blog/top-machine-learning-products.
[51] D. Igoe, „Crypto trading bots: The ultimate beginner’s guide“, stu. 17, 2020. https://www.trality.com/blog/crypto-trading-bots.
[52] „B-cube“, *B-cube.ai*. https://www.b-cube.ai/index.html.
[53] „They didn’t believe AI can make profits in Crypto trading, but when they saw the results of the first 8 months“, *Becoming Human AI*, stu. 18, 2020. https://becominghuman.ai/they-didnt-believe-ai-can-make-profits-in-crypto-trading-but-when-they-saw-the-results-of-the-4347f5f37fb0.