Универзитет „Св. Климент Охридски“ – Битола

Факултет за информатички и комуникациски технологии - Битола

ПОДОБРУВАЊЕ НА ПЕРФОРМАНСИТЕ НА ФИНАНСИСКИТЕ СЕРВИСИ СО КОРИСТЕЊЕ НА ТЕХНОЛОГИЈАТА НА БЛОКОВСКИ ВЕРИГИ

ДОКТОРСКА ДИСЕРТАЦИЈА

Изработил

м-р Мимоза Мијоска

Ментор

ред. проф. д-р Благој Ристевски

јуни, 2022 година

ЧЛЕНОВИ НА КОМИСИЈАТА ЗА ОДБРАНА НА ДОКТОРСКАТА ДИСЕРТАЦИЈА

Ред. проф. д-р Пеце Митревски
Факултет за информатички и комуникациски технологии
Универзитет „Св. Климент Охридски“ – Битола, претседател

Ред. проф. д-р Благој Ристевски
Факултет за информатички и комуникациски технологии
Универзитет „Св. Климент Охридски“ – Битола, ментор

Ред. проф. д-р Костандина Вељановска
Факултет за информатички и комуникациски технологии
Универзитет „Св. Климент Охридски“ – Битола, член

Ред. проф. д-р Снежана Савоска
Факултет за информатички и комуникациски технологии
Универзитет „Св. Климент Охридски“ – Битола, член

Ред. проф. д-р Владимир Трајковиќ,
Факултет за информатички науки и компјутерско инженерство,
Универзитет „Св. Кирил и Методиј“ – Скопје, член

До
Наставен совет на трет циклус студии
Наставно-научен совет
на Факултет за информатички и
комуникациски технологии - Битола

ИЗЈАВА

Јас, Мимоза Мијоска, студент на трет циклус студии на Факултетот за информатички и комуникациски технологии, Универзитет „Св. Климент Охридски“ - Битола, изјавувам дека при изработка на докторската дисертација со наслов „**Подобрување на перформансите на финансиските сервиси со користење на технологијата на блоковски вериги**“, ги почитував позитивните законски прописи од областа на заштита на интелектуалната сопственост и авторски права и не користев реченици или делови од трудови на други автори без да ги почитувам методолошките стандарди.

Докторската дисертција е изработена под менторство на ред. проф. д-р Благој Ристевски.

Изјавата ја давам под материјална и кривична одговорност.

Изјавил,

м-р Мимоза Мијоска

Мијоска

БЛАГОДАРНОСТ

Најголема благодарност до моето семејство за нивното разбирање, трпение и поддршка во текот на целокупното студирање и изработката на дисертацијата. Непоколебливо им благодарам за несебичната сила и доверба што ми ја даваат во животот.

Исклучителна благодарност до мојот ментор ред. проф. д-р Благој Ристевски кој несебично ме поддржуваше во текот на студиите, ми пружи професионална соработка и комплетно го споделуваше своето знаење во текот на изработката на дисертацијата.

Содржина

Листа на слики

Апстракт

Технологијата на блоковски вериги има потенцијал да се примени во различни области од секојдневниот живот. Новите решенија во области каде се користат големи податоци што можат да бидат креирани со примена на технологијата на блоковски вериги се: здравството, медицината, децентрализираната заштита на личните податоци, дигиталната сопственост, Интернетот на нештата, дигиталниот идентитет, финансиските услуги и инфраструктурата, е-трговијата, образовниот систем, споделувањето знаење, осигурувањето, прехранбената индустрија, сметководството, ревизијата и електронското гласање. Оваа технологија може да се користи во електронската здравствена евиденција, за формирање и водење матични книги на родени, умрени, венчани, регистрација на деловни активности, но и во организирање на избори. Карактеристиките на оваа технологија можат да се користат за редефинирање на Интернет 3.0 дефиниран како нов тип на децентрализирана инфраструктура или мрежа од мрежи. Како нова технологија, може да се користи за анализа и обработка на податоците преку ефективна интеграција на финансиските ресурси. Се произведуваат нови финансиски формати или модели на услуги за да се надгради финансискиот систем и да се промовира ефикасноста и квалитетот на финансиското работење, врз основа на потребите на клиентите. Технологијата на блоковски вериги може да ѝ помогне на финансиската индустрија автоматски и прецизно да ги идентификува условите за кредитирање на клиентите, да го реструктуира кредитниот систем на финансискиот пазар. Меѓу другите предности што ги нуди оваа технологија е големото забрзување на меѓународниот трансфер на средства, што може да се направи за неколку минути, а во традиционалното банкарство потребни се неколку дена во развиените економии или уште повеќе во земјите во развој.

Од друга страна, машинското учење е една од најзабележливите технологии во последниве години. Двете технологии се водени од податоци, и затоа има брзорастечки интерес за нивно интегрирање за побезбедно и поефикасно споделување и анализа на податоците. Оваа дисертација покажува како овие две технологии, технологијата на блоковски вериги и машинското учење, можат да се комбинираат во предвидувањето на променливоста на биткоинот. За да се анализира и предвиди променливоста на биткоинот, користени се податоци за биткоинот од серии во реално време, а како алгоритам за машинско учење е користен алгоритамот случајни шуми. При предвидување на променливоста на биткоинот, добиени се мали статистички грешки во множеството за обука и множеството за тестирање. Анализирани се вредностите на следните статистички грешки: средната апсолутна грешка (*MAE*), средната квадратна грешка (*RMSE*), просечната апсолутна процентуална грешка (*MAPE*), просечната апсолутна процентуална грешка *(MDAPE)* и симетричната средна апсолутна процентуална грешка *(sMAPE)*. Тие се прикажани визуелно. Од добиените резултати се

извлекуваат соодветни заклучоци и се потврдува дека моделот за предикција е добро дизајниран.

Клучни зборови: технологија на блоковски вериги, децентрализирана заштита на лични податоци, дигитална сопственост, Интернет на нештата, криптовалути, машинско учење, случајни шуми, променливост на биткоин, статистички грешки.

Abstract

Blockchain technology has the potential to be applied in various areas of our daily lives. New solutions in areas where big data is used that can be created by applying blockchain technology are healthcare, decentralized protection of personal data, digital property, Internet of Things, digital identity, financial services and infrastructure, e-commerce, educational records, education system, knowledge sharing, insurance, food industry, accounting, auditing and electronic voting. This technology could be used in electronic health records and the creation and management of registers of births, deaths, marriages and registration of business activities, but also in the organization of elections. The features of this technology can be used to redefine Internet 3.0 defined as a new type of decentralized infrastructure or network of networks. As a new technology, it can be used to analyze and process data through the effective integration of financial resources. New financial formats or service models are produced to upgrade the financial system and promote the efficiency and quality of financial operations, based on customer needs. Blockchain technology can help the financial industry automatically and accurately identify the credit conditions of customers and restructure the credit system of the financial market. Among the other advantages offered by this technology is the great acceleration of the international transfer of funds, which can be done in a few minutes, while in traditional banking it takes a few days in developed economies or even more in developing countries. On the other hand, machine learning is one of the most notable technologies in recent years. Both technologies are data-driven, and thus there is rapidly growing interest in integrating them for more secure and efficient data sharing and analysis. This dissertation shows how these two technologies, blockchain technology and machine learning, can be combined to predict bitcoin volatility. To analyze and predict Bitcoin volatility, real-time series Bitcoin data and random forests as a machine learning algorithm were used. When predicting the volatility of Bitcoin, low statistical errors were obtained in the training set and the test set. The values of the following statistical errors were analyzed: mean absolute error (MAE), mean squared error (RMSE), mean absolute percentage error (MAPE), median absolute percentage error (MDAPE) and symmetric mean absolute percentage error (sMAPE). They are also shown visually. Appropriate conclusions are drawn from the obtained results and it is confirmed that the prediction model is well designed.

Keywords: blockchain technology, decentralized personal data protection, digital property, Internet of Things, cryptocurrencies, machine learning, random forests, bitcoin instability, statistical errors.

1. Вовед

1.1 Предмет и област на истражување

Технологијата на блоковски вериги (blockchain) е воведена во 2008 година со објавување на трудот на Сатоши Накамото - „Биткоин: електронски готовински систем со рамноправен пристап (peer-to-peer)“ [1]. Технологијата на блоковски вериги првпат е употребена кај криптовалутата биткоин. Првите трансакции со биткоини се случиле во јануари 2009 година. Освен нивната употреба во економскиот домен, биткоинот и технологијата на блоковски вериги, решаваат важен проблем во информатиката и компјутерската техника што со години претставувал пречка, за изградба на функционален дигитален монетарен систем. Со оваа технологија решен е проблемот на двојно искористување, т.е. елиминиран е ризикот криптовалутата да може да се користи два или повеќе пати. Развивачите на виртуелна валута мора да ги спречат корисниците да можат да ги трошат своите средства повеќе од еднаш. Интересот на претпријатијата, индустријата и владите ширум светот за технологијата на блоковски вериги е голем, бидејќи примената на оваа технологија е многу поголема од доменот на криптовалутите.

Постојат четири главни типови на примена на технологијата на блоковски вериги:

1. парични средства (валути, плаќања, дознаки, финансии, хартии од вредност и финансиски инструменти како што се: пари, чекови, меници, акции, обврзници, благајнички записи, комерцијални записи и др.),
2. имот (регистри на земјиште, недвижен имот и сопственост на возила),
3. договори (деловни договори, лиценцирање, регистрација, тестаменти, договори за партнерство и регистрација на патенти) и
4. лични документи (пасош, виза, возачка дозвола и матични книги на родени).

Од 2015 година, голем број меѓународни финансиски организации веќе планираат понатаму да ја развиваат технологијата на блоковски вериги. Во 2014 година е основан конзорциум наречен R3, со цел да започне истражување и развој на технологијата на блоковски вериги. Во март 2017 година, оваа група броела околу 75 компании, 200 компании во март 2018 година, за да достигне 400 компании во март 2022 година [14]. Формирањето на толку силна корпорација со многу истражувања и имплементација на технологијата на блоковски вериги, особено во финансискиот сектор, укажува на тоа дека претстои нова ера во развојот на банкарството. Поединци изразија загриженост дека условите за доделување кредити честопати се премногу строги, или дека треба да обезбедат хипотека или дека воопшто немаат пристап до кредит. Каматните стапки се различни од просекот во Европската Унија и се сноси значително поголем ризик, што доведува до нестабилна економија, но исто така и високи оперативни трошоци по кои можат да се применат нови технологии. Со примена на нова технологија, банкарскиот

сектор би создал услови за значително зголемување на ефикасноста на работењето, намалување на изложеноста на голем број ризици што би резултирале со иновации во заемните активности, па дури и пониски каматни стапки на пласманите на хартии од вредност, акции и др.

Имајќи го предвид погоре наведеното, предметот на истражување на оваа дисертација се дефинира како истражување, примена и подобрување на финансиското работење со користење на технологијата на блоковски вериги со интеграција на алгоритмите за машинско учење.

1.2 Структура на докторската дисертација

Оваа докторска дисертација е структурирана од 7 поглавја, во кои се опфатени описот, карактеристиките и примената на технологијата на блоковски вериги. Понатаму, се објаснети алгоритмите за машинско учење, метриките за евалуација, како и имплементацијата на алгоритмите за машинско учење и нивната анализа во насока на подобрување на перформансите при предвидување на променливоста на биткоинот.

Во првото поглавје е даден воведот, додека во второто поглавје е опишан предметот на истражување на оваа дисертација, технологијата на блоковски вериги како основа на криптовалутата биткоин. Опишани се карактеристиките на технологијата на блоковски вериги, и можноста да се избегнат медијаторите (посредниците). Подетално е опишана примената на технологијата на блоковски вериги во различни области.

Во третото поглавје ставен е акцент на алгоритмите за машинско учење и нивната примена. Машинското учење ги „програмира“ компјутерите на оптимизација врз основа на одреден дел од податоци или одредено искуство од минатото. Постои модел во кој се дефинирани одредени параметри, а учењето се извршува за оптимизација на параметрите на моделот преку користење на тренирачко множество или пак минато искуство. Моделот само тогаш може да биде предвидлив односно да дава претпоставки за иднина, или да стане дескриптивен, односно да стекнува знаење од податоците.

Во четвртото поглавје е наведена имплементацијата на ансамбл-алгоритамот случајни шуми во насока на подобрување на перформансите при предвидување на променливоста на биткоинот во програмскиот јазик *R*. Јазикот *R* доаѓа со интегрирана рамка за извршување на напредни анализи што им помага на корисниците да користат повторувачки прашања на различно множество на податоци, при што од голема помош се алгоритмите за машинско учење. Во ова поглавје опишани се и метриките за евалуација на моделот.

Во петтото поглавје прикажани се резултатите што се добиени со примена на моделот што е креиран во програмскиот јазик *R* со примена на технологијата на блоковски вериги во комбинација со машинско учење.

Во наредното поглавје анализирани се добиените резултати при користење на моделот во ситуации со користење на различен број на дрва, различна поделба на множествата за тренирање и тестирање и избор на различен број на хоризонти при предвидувањето на променливоста на биткоинот. Резултатите кои произлегуваат од истражувањата спроведени во оваа докторска дисертација директно можат да влијаат врз секојдневната примена при предикцијата на променливоста на вредноста на биткоинот, а на тој начин би влијаеле на подобрување на финасиските сервиси.

Заклучокот и насоките за идна работа се опишани во последното поглавје.

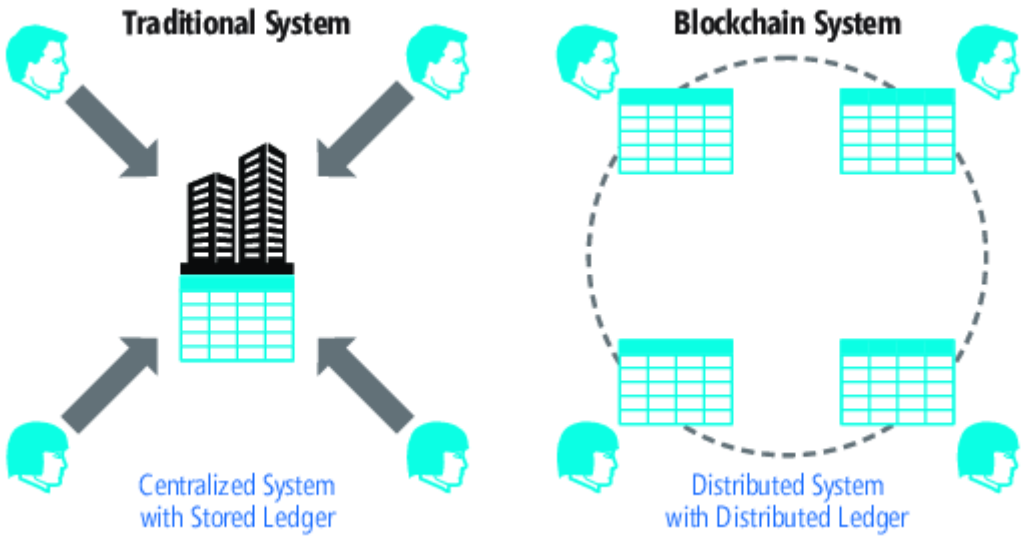
2 Технологија на блоковски вериги

2.1 Опис на технологијата на блоковски вериги

Во 2008 година голем број моќни финансиски институции и осигурителни компании во САД биле на работ на банкрот. Овие околности довеле до непосредна интервенција на федералната влада, со цел да се избегне домашен, а можеби и глобален финансиски колапс [7].

Овие настани ги илустрираат опасностите за живеење во дигитален, поврзан свет, кој зависи од посредниците што генерираат трансакции и ги прават луѓето ранливи на дигитални експлоатирања и криминал. Академски предизвик е да се создаде дигитална инфраструктура за располагање, без посредници, којашто нема корумпирано или централно тело подложно на грешки, и е сигурна и може да ѝ се верува.

Во блоковската верига, регистрите се дистрибуираат низ целата мрежа и нема потреба некој медијатор да биде во средина на трансакцијата. Технологијата одржува повеќекратни копии на податоци, слично како во системот за споделување датотеки со рамноправен пристап (peer-to-peer). Секој јазол добива копија од целата база на податоци, прикажано на слика 1.



Слика 1. Децентрализирана дистрибуирана евиденција [3].

2.2 Медијатори

Секогаш кога има потреба да се испрати нешто на некој кој не е на иста локација, треба да се користат медијатори (посредници). Ако треба да се испрати порака преку мобилен телефон, медијатори можат да бидат мобилните оператори. Исто така посредници може да бидат: Viber, WhatsApp, Telegram, или социјалните мрежи како Facebook, Twittter или LinkedIn. Опции за испраќање порака има многу, но заедничка за сите е потребата од постоење на посредник. Истото важи и кога се испраќа

електронска порака. Тогаш посредник може да биде Gmail, Yahoo, Hotmail или било кој друг давател на услуги преку електронска пошта. Исто така, може да се испрати порака со писмо и во тој случај посредник е поштата. Кога се испраќаат пари во странство, повеќе посредници се вклучени во процесот - повеќе банки, финансиски компании како што се Western Union, PayPal и др. [5].

Посредниците не се совршени. Покрај тоа што се корисни, тие имаат и свои недостатоци и тоа:

• **Цена**. Медијаторските услуги обично не се бесплатни (исклучок е комуникација преку Интернет – Skype, Viber, e-mail, WhatsApp). Цените на овие услуги можат да бидат симболични, но можат да бидат и значително високи.

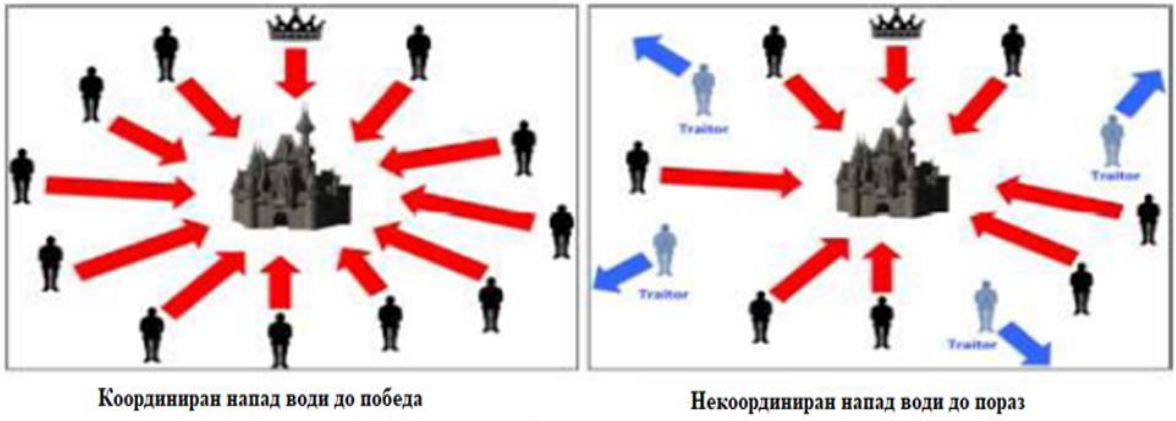
• **Задоцнување**. Генерално, посредниците влегуваат со одредено задоцнување во процесот на испраќање. Исклучок се комуникациите преку Интернет, каде што ова доцнење обично се мери со делови од секунда. Пример за случај во кој ова доцнење може да биде значително е испраќање пари во друга земја преку банкарска сметка. Затоа сервисите како што се Western Union и Moneygram наплаќаат висока провизија што ја оправдуваат со брзината на трансферот.

• **Доверба**. Во повеќето случаи довербата се претпоставува и во повеќето случаи медијаторот ја оправдува оваа доверба. Се случува понекогаш да се загуби пакет или трансакцијата со пари да биде одложена поради неуспехот на медијаторот, но таквите случаи се релативно ретки.

• **Приватност**. Медијаторот често има приватни податоци од своите клиенти. Постои оправдан страв дека посредникот може да ги злоупотреби овие информации. Ако медијаторот не ги чува приватните податоци на безбеден начин, хакери можат да ги нападнат и потоа да ги злоупотребат.

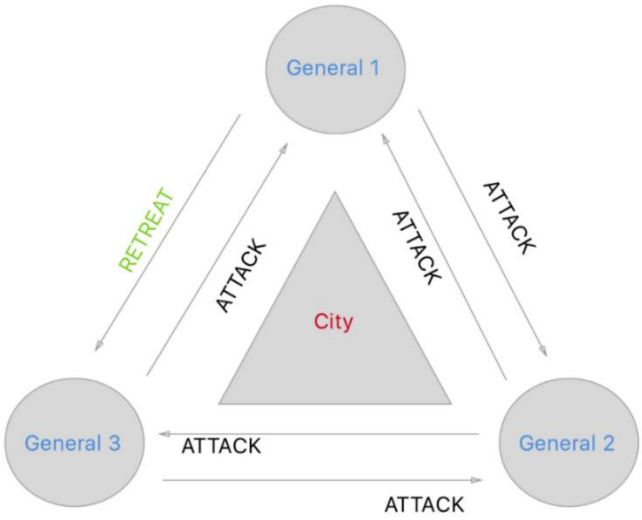
2.2.1 Проблем на византиските генерали

Проблемот на византиските генерали е дефиниран уште во 1982 година и го илустрира проблемот на комуникација преку посредници кои не се доверливи [4]. Византиски генерали организираат опсада на непријателски град и треба да се договорат за заеднички акционен план. За да успее планот потребно е сите да нападнат истовремено. Доколку некој од генералите не го изведе нападот во договореното време, многу е веројатно дека нападот ќе пропадне. Бидејќи генералите се на различни локации, не можат да преговараат во живо, туку комуницираат преку гласници (курири, посредници). Се појавуваат два потенцијални проблеми. Првиот проблем е ако некои од генералите се предавници, тие намерно ќе го саботираат договорот и ќе испраќаат информации преку нивните курири за да доведат до конфузија кај чесните генерали, прикажано на слика 2.



Слика 2. Приказ како византиските генерали организираат опсада на град [4].

Во поедноставно сценарио, сите генерали се искрени, но постојат курири што се предавници. Во уште поедноставна ситуација постојат само 3 генерали, а секој од нив има по 2 гласници. Секој генерал испраќа по еден курир до секој свој колега. Во овој пример, еден од куририте на првиот генерал е предавник и намерно му пренесува погрешна порака на третиот генерал, прикажано на слика 3. Првиот и вториот генерал мислат дека е постигнат договор и тргнуваат во напад. Третиот генерал добива различни пораки од своите колеги и затоа не напаѓа бидејќи смета дека не е постигнат договор. Наместо цела војска, 2/3 одат во напад и со тоа значително се намалува можноста за победа.



Слика 3. Курир - предавник [4].

Со овој едноставен пример, се заклучува дека е доволно само 1 од 6 курири да е предавник и веројатноста за успех во нападот значително да се намали. Во поголемите и посложени системи, бројот на учесници (и бројот на предавници) е значително повисок, што само дополнително ја усложнува ситуацијата. Овој проблем е посебно изразен кај системи што не се централизирани и каде што бројот на учесници е преголем за да можат директно да комуницираат.

2.2.2 Технологијата на блоковски вериги како решение за проблемот на византиските генерали

Проблемот со византиските генерали, всушност е проблем за доверливоста на посредниците. Технологијата на блоковски вериги го решава овој проблем со зголемување на бројот на посредници. Ако не може да му се верува на еден посредник, се користат илјадници посредници. Поради начинот на кој работи мрежата, ништо не може да се случи и ако некој од учесниците во биткоин мрежата се обиде да ја смени трансакцијата што ја примил пред да биде испратена. Учесниците во мрежата постојано комуницираат и ги споредуваат своите копии од базата на податоци со другите копии. Ако се забележи дека нивната копија е различна од другите, тие ја прилагодуваат својата копија да биде иста со останатите. Сепак, секој учесник е директно поврзан со неколку други. Ако од еден учесник добие една информација, а од сите останати друга информација, тој учесник, едноставно ќе биде игнориран, така што променетите информации не може да се пропагираат преку мрежата.

Во мрежа од неколку илјади учесници, еден „предавник“ не може да направи никаква штета. Се додека бројот на „предавници“ е под 50 % од сите посредници, целата мрежа е безбедна, бидејќи јазлите се програмирани така што позицијата на мнозинството се прифаќа како точна. За успешен напад на мрежата доверливите членови на мрежата треба да се во малцинство, но тоа не е единствениот услов. За нападот да успее, потребно е напаѓачите да бидат во мнозинство и да бидат совршено синхронизирани. Синхронизација на илјадници учесници сама по себе е тежок процес, а нападот го прави уште потежок, бидејќи тие треба да дејствуваат истовремено и во оној момент откако ќе се појави трансакцијата што сакаат да ја изменат. Проблемот е во тоа што е многу тешко однапред да знаат кога точната трансакција ќе се случи и како ќе изгледа. Нападот врз мрежата на биткоин е теоретски можен, но во практиката тоа се покажало невозможно во овие десетина години од постоењето [99].

Со технологијата на блоковски вериги се решава прашањето на доверба во децентрализирана мрежа на учесници што не се познаваат, што е всушност, проблем на византиските генерали.

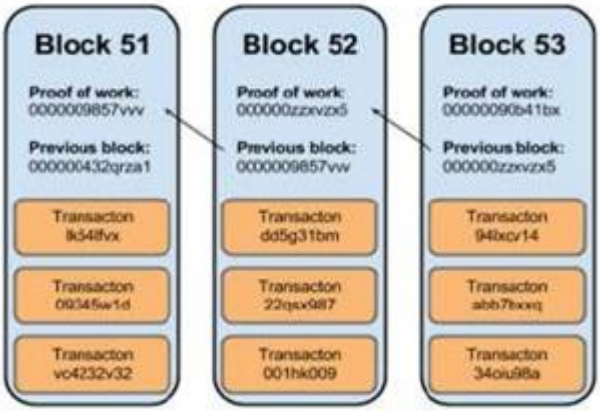
Блоковските вериги се регистар на сите трансакции што некогаш се случиле во блоковската верига на биткоин. Ова само по себе не е ништо необично, бидејќи секој систем во кој се врши исплата води евиденција на овие трансакции и ги чува како база на податоци. Она што е голема иновација во блоковските вериги е начинот на кој информациите за трансакциите се испраќаат и складираат.

Блоковските вериги се постојано растечка листа на дигитални записи во пакети наречени блокови, коишто се поврзани и обезбедени со помош на криптографски алгоритми. Овие дигитално снимени блокови на податоци се складираат во линеарна верига (синџир). Секој блок во веригата содржи податоци на пример трансакција на биткоин за која е запишано и времето кога е извршена трансакцијата. Новиот блок со

податоци се поврзува со блокот којшто се случил пред него во веригата, и на тој начин не постои можност за менување на сите податоци во целата блоковска верига.

Оваа технологија е релативно нов концепт и брзорастечки дел од основната технологија, како што се Интернетот или пресметувањето во облак. Слични структури на податоци постоеле долго пред да се замисли популарниот биткоин, меѓутоа главните теории за архитектури за оваа технологија што се користат денес биле првично наведени и дефинирани во оригиналната статија за биткоинот напишана и објавена од Сатоши Накамото во 2008 година [1].

„Blockchain“ е составен од зборовите „Block“ (блок) и „chain (синџир или верига). Имено, кај биткоинот трансакциите се пакувани во блокови, а блоковите се вградени во еден синџир, прикажано на Слика 4. За да се поврзат блоковите, се користат криптографски алгоритми, поточно хеш-функција, на начин на кој е невозможно да се промени содржината на еден блок, а да не се промени содржината на сите блокови што следуваат по него. Ова е многу важна карактеристика на технологијата на блоковски вериги, бидејќи обезбедува непроменливост на податоците што се внесуваат во неа. Алгоритамот SHA-256 за ист влезен податок секогаш го произведува истиот излез со фиксна должина. Значењето на користењето криптографски хеш функции како што е SHA-256 е дека тие се доволно единствени и можат да служат како дигитален отпечаток од прст додека истовремено дејствуваат како контролна сума. Исто така, еднонасочните хеш-функции не можат да бидат дешифрирани.



Слика 4. Верига на блокови [6].

Табелата 1, покажува како низите од знаци со различни должини секогаш произведуваат 64-цифрена хексадецимална хеш-вредност, и дека мала промена на некоја одредена низа произведува сосема поинаков резултат.

Табела 1. Низи од карактери (стрингови) со различни должини со помош на SHA-256 алгоритам [7].

Влезна низа од знаци	**SHA-256 хеш вредност**
m	62C66A7A5DD70C3146618063C344E531E6D4B59E379808443CE962B3ABD63C5A
M	08F271887CE94707DA822D5263BAE19D5519CB3614E0DAEDC4C7CE5DAB7473F1
M1	2D214CA69B86C255BE416D42CCA977A59B34A7492873105522C35015FAB806F0
M2	0892A10ECE1F933EE98F5D554601B28F8437801D1AA1B77799E4035DDCB6950C
March	9D95A2CF0D7180B5089691163B188A7203B0CDE179346B8CFAA8AB6C2C3E6414
March 1, 2018	767328E7367048FA9DB37354CFA43DBB1691E8330DB54D54F52C1A444CA2E680
March 2, 2018	CCF33BF1C08B74EDE6A7C15C56EEC16269D83967670032ACDA6EE395361B7595

Понекогаш хеш-вредноста е двојно хеширана, што значи дека првиот хеш се враќа повторно со примена на вториот круг од алгоритамот SHA-256. Ако вредностите на табела 1 двојно се хешираат се добива резултатот прикажан во табела 2.

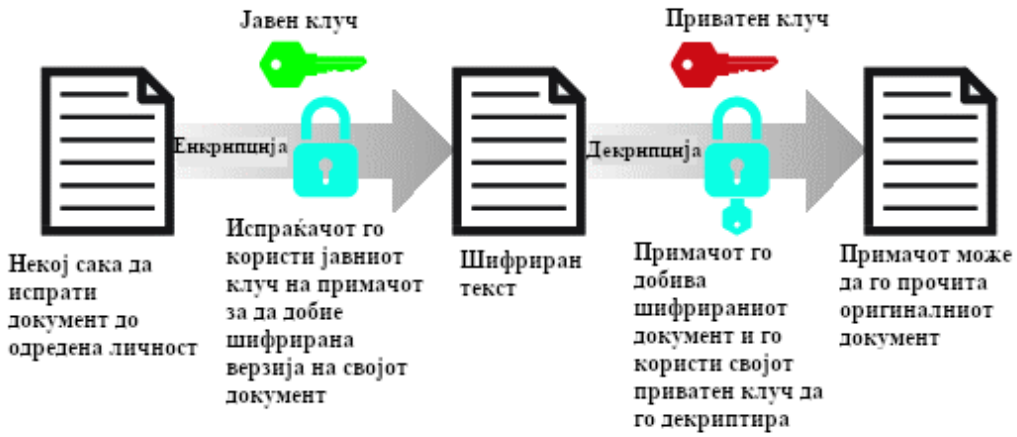
Табела 2. Двојно хеширање на вредностите од табела 1.

Влезна низа од знаци	**Двојни SHA-256 хеш вредности**
m	A5FCE7E78734EC317F83F9019C80FFAF2508689B06EFA02191495A7D21FECE9A
M	6F6DCF58526B0D29EE664A708A939B7CDAC124A6A8569FCACE46FEAD38868E2E
M1	6C5D08BE9FFBBABD24B5F19AFFE6590FD402D347A50B519A59D40E15DCC0A6CB
M2	B2307311CC5877D5A581EDC821F3BFD5F99EB4E3B1D1B4009D9545BCF07E2E1A
March	B5410E155022AE6EB22CA21FADEDE65F0F7296DE14CA1D7A720A4937BD23AA5D
March 1, 2018	345DD725FEE80F8C5953A66C1495605E4ED01C4CE5AEF6C0A6D238999266A1A6
March 2, 2018	3E85B3D910BA77F88ECD5E24D1396457C532C73B89C032DED9AD0CBB4D4D9794

Една од основните функции на технологијата на блоковски вериги е да ја следи сопственоста на дигиталните средства. Дигиталното средство може да биде безвредно или да вреди многу милиони долари, па затоа тестот за сопственост мора да го обезбеди сопственикот дека не може да биде измамен. За да се спроведе таков тест во

дигиталниот свет, технологијата на блоковски вериги се темели на криптографија со јавен клуч (PKC), што му овозможува на сопственикот дигитално да го потпише своето средство за да докаже сопственост и да даде овластување за да може да биде префрлено на друго лице.

Кога и да е потребно, корисниците пристапуваат кон софтверска апликација која користи алгоритам за дигитално потпишување со елиптични криви (ECDSA, Elliptic Curve Digital Signature Algorithm) за генерирање на пар на криптографски клучеви. Корисникот мора да задржи резервна копија на приватниот клуч, бидејќи тој клуч е потребен за пренос или искористување на вредноста што се чува во дигиталното средство, кое е зачувано во блоковските вериги. Ако постои пристап само до приватниот клуч може да се генерира јавниот клуч, бидејќи постои математичка врска меѓу двата клуча. Но, приватниот клуч не може да се генерира од јавниот клуч, што значи дека ако се направи резервна копија само на еден клуч, тоа мора да е приватниот клуч. На слика 5 е прикажано користењето на приватен и јавен клуч кога е потребно некој да испрати шифриран документ/порака што може да ја отвори само тој што го има приватниот клуч.

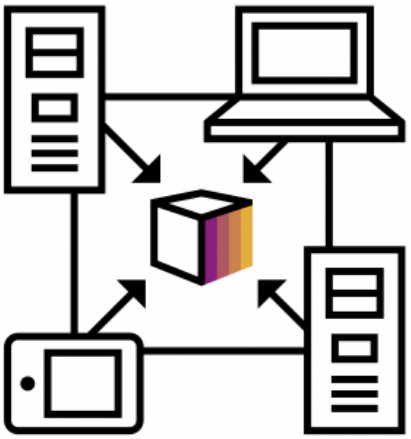


Слика 5. Користење на приватен и јавен клуч кога е потребно некој да испрати шифриран документ/порака што може да ја отвори само тој што го има приватниот клуч [7].

2.3 Карактеристики на технологијата на блоковски вериги

2.3.1 Дистрибуирана евиденција

Дистрибуираната евиденција е консензус на пресликаните, споделените и синхронизираните дигитални податоци географски распространети на повеќе локации, земји и/или институции, прикажано на слика 6.



Слика 6. Дефиниција на дистрибуирана евиденција [8].

Корисниците на технологијата на дистрибуирана евиденција (Distributed Ledger Technology - DLT) имаат значителна корист од ефикасноста и економичноста преку создавањето посилна средина во реално време и безбедно делење на податоците. Спротивно на вообичаеното верување, биткоин-технологијата на блоковски вериги не е единствениот дистрибуиран регистар, всушност, многу други корисници на технологијата на дистрибуирана евиденција користат различни методологии за постигнување на истиот консензус, како на пример Ripple, MultiChain, HyperLedger Project и други.

Технологијата на блоковски вериги е децентрализирана, што значи дека нема централен сервер на којшто корисниците се поврзуваат. Сите корисници се поврзани со мрежата со рамноправен пристап (peer-to-peer) и секој корисник претставува еден јазол во таа мрежа, прикажано на слика 7. Бидејќи мрежата со рамноправен пристап е таква што секој корисник може да биде директно поврзан со само неколку други корисници, тоа значи дека сите други се индиректно поврзани. Информациите преку мрежата со рамноправен пристап патуваат со испраќање на секоја порака само до оние корисници кои се директно поврзани, а потоа секој ја праќа пораката до оние кои се директно поврзани со него, сè додека пораката не стигне до сите учесници во мрежата. Испраќањето на информации преку голем број медијатори е аналогно на примерот со проблемите на византиските генерали.

Слика 7. Мрежа со рамноправен пристап [9].

За да се разбере како технологијата на блоковски вериги го решава проблемот на византиските генерали, треба да се спомене уште едно важно својство. Имено, некои од јазлите, наречени „целосни јазли“, во мрежата имаат комплетен блок на нивниот хард-диск (основата на сите трансакции што се случиле) и бројот на такви јазли е во илјадници [99]. Бидејќи сите тие постојано комуницираат меѓусебно, секогаш проверуваат дали нивната копија на блоковската верига се согласува со останатите. Ако таа не се согласува, тие автоматски ја ажурираат својата верзија за да одговара на останатиот дел од мрежата. Тоа е како база на податоци пресликана на илјадници компјутери, која се ажурира на сите овие компјутери во реално време, што значи дека во таква база на податоци е многу тешко да се сменат некои информации.

Секој корисник на биткоин-технологијата на блоковски вериги има приватен клуч, јавен клуч и биткоин-адреса. Јавниот клуч се создава од приватен клуч, но на начин што е невозможно да се направи обратен процес, односно од јавниот клуч да се креира приватниот клуч. Потоа, од јавниот клуч се креира биткоин адреса на која се добиваат биткоини и нема потреба да се крие. Од друга страна, многу е важно приватниот клуч добро да се чува. Со приватниот клуч се потпишуваат трансакции на биткоин адресата која е поврзана со тој клуч. Аналогија на биткоин адресата е бројот на личната банкарска сметка, а приватен клуч е пинот со кој се потврдуваат плаќањата од таа сметка. Ако се изгуби пинот што ги потврдува трансакциите од банкарската сметка, банката ќе издаде нов по барање на сопственикот, но ако се изгуби приватниот клуч, се губат биткоините засекогаш на адресата која е поврзана со изгубениот приватен клуч. Овие биткоини сè уште постојат, но никој не може да им пристапи без приватен клуч.

Кога треба да се испратат биткоини некому, треба да се внесе биткоин-адресата на примачот, износот што треба да се испрати и да се потпише трансакцијата со

соодветниот приватен клуч. Потоа информациите за оваа трансакција се испраќаат на сите учесници што директно се поврзани со мрежата со рамноправен пристап, така што тие понатаму ќе бидат пренасочени кон примачот. Нема посредници во класичната смисла, меѓутоа, кога се испраќаат биткоини, трансакцијата има посредници, па дури и илјадници, само нивната улога е малку поинаква.

Секој јазол во мрежата на биткоин е посредник што ја проверува трансакцијата и ја пренесува, сè додека не ги достигне сите јазли во мрежата. Кога секој ќе ја провери трансакцијата и таа ќе стигне до сите во мрежата, таа го исполнува условот да се вметне во блок и на тој начин станува дел од блоковска верига. За да се изврши една трансакција се проверуваат следните податоци:

• дали испраќачот има доволно биткоини на својата биткоин адреса за да може да се изврши трансакцијата;

• валидноста на адресата од примачот и

• валидноста на потписот на испраќачот.

Постојат јазли кои имаат комплетна блоковска верига на нивниот хард диск (сите трансакции што се случиле). Затоа тие знаат во секое време точно каде е адресата, колку биткоини има и кои адреси се валидни, а кои не се. Ако трансакцијата ги помине сите три проверки, таа се проследува понатаму преку мрежата со рамноправен пристап сè додека не пристигне до сите јазли од блоковската верига. Иако ова делува како многу сложен процес, во практиката, трансакцијата на биткоинот трае помалку од секунда за да ги достигне сите јазли во мрежата на биткоин.

Проблемот со посредниците коишто се нелојални во примерот со византиските генерали, биткоинот го решава со зголемување на бројот на посредници. Ако не може да му се верува на еден посредник, тогаш се користат илјадници посредници. Ако некој од учесниците во биткоин-мрежата се обиде да ја смени трансакцијата што ја примил пред да биде испратена, нема ништо да се случи, поради начинот на кој работи мрежата. Учесниците во мрежата постојано комуницираат и ги споредуваат своите копии од базата на податоци со другите копии. Ако се забележи дека нивната копија е различна од другите, тие ја прилагодуваат својата копија да биде иста со останатите. Сепак, секој учесник е директно поврзан со неколку други. Ако од еден учесник добие една информација, а од сите останати друга информација, тој еден учесник едноставно ќе биде игнориран, така што променетите информации не може да се пропагираат преку мрежата.

Технологијата на блоковски вериги е само еден вид на дистрибуиран регистар, и сите дистрибуирани регистри не мора секогаш да користат блокови или верижни трансакции. Иако во дискусиите терминот „blockchain“ се користи почесто од „distributed ledger“, технологијата на блоковски вериги е само еден од многуте типови на структури на податоци кои обезбедуваат сигурно и валидно постигнување на

дистрибуиран консензус. Биткоин-технологијата на блоковски вериги, користи „доказ за работа“ (PoW, Proof-of-Work Mining). Тоа е најстариот јавно докажан метод што се користи за постигнување на дистрибуиран консензус [51].

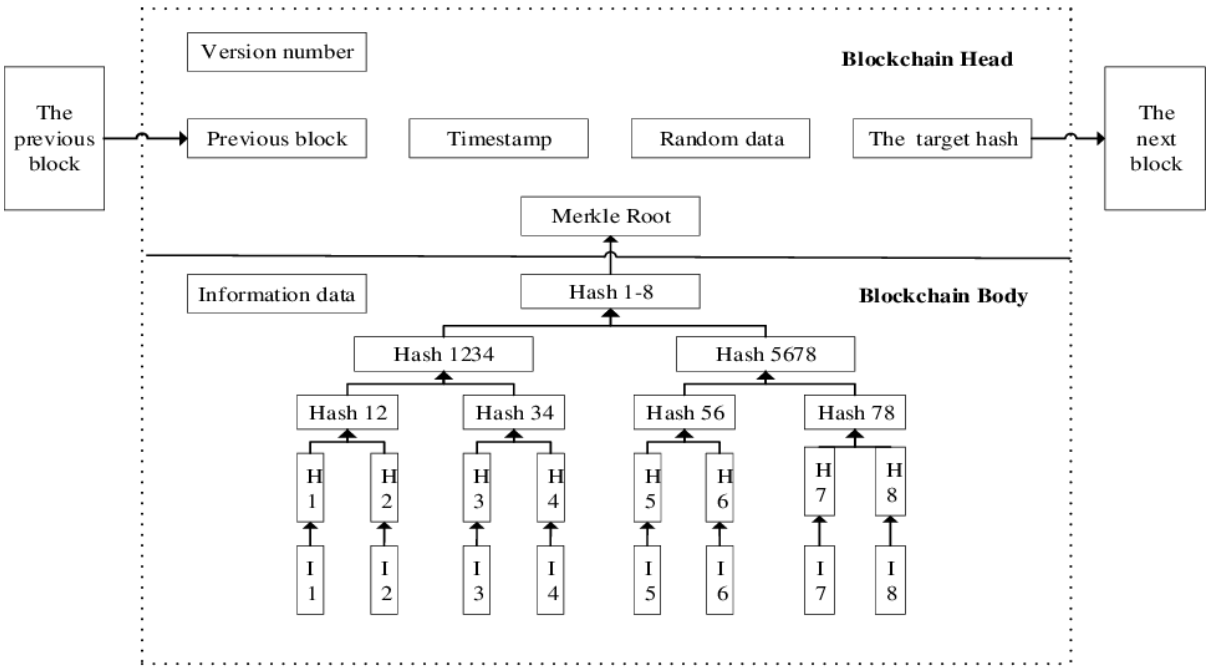
Блок е збир на податоци што се собираат и обработуваат за да се вклопат во него преку процесот на рударење. Секој блок се идентификува преку криптографски хеш и временски печат (time stamp). Кога ќе се формира нов блок, тој ќе содржи хеш од претходниот блок, така што блоковите можат да формираат хронолошки нареден синџир од првиот блок што некогаш бил генериран во целата блоковска верига (исто така наречен „genesis block“) до новоформираниот блок. Овој процес се повторува одново и одново за да се развива и одржува мрежата.

Многу често се дискутира за придобивките од технологијата на блоковски вериги, или како таа ќе револуционизира разни застарени индустрии. Спротивно на тоа, малку се дискутира за тоа како таа, всушност, функционира, првенствено од аспект на вистинската софтверска архитектура.

Овој концепт се објаснува со разбивање на концептот на блоковски вериги, во две одделни компоненти - блок и верига.

Блокот може да се смета за контејнер за податоци. Во случај на блоковската верига од биткоинот, секој блок содржи податоци (како што се биткоин-трансакции), блок-заглавија, блок-идентификатори и т.н. Мерклови дрва (Merkle trees).

Структурата на податоци во блоковските вериги е прикажана на слика 8.



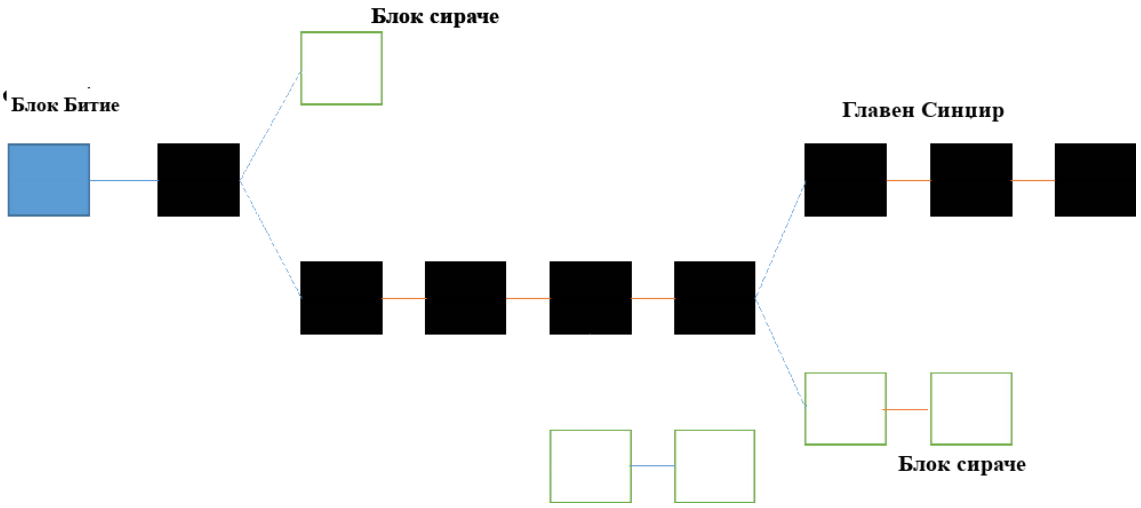
Слика 8. Структурата на податоци во блоковски вериги [86].

Во телото на блоковската верига е сместен дел од Мерклово хеш-дрво кое може да биде бинарно дрво или мултидрво во структурата на податоци. Поточно, податоците или информациите се евидентираат како хеш-вредност зачувана во телото на блокот, а генерираниот корен преку процесот на хеш на Меркел-дрвото ќе се запише во заглавјето на блокот.

Блок-заглавијата содржат метаподатоци за тој посебен блок, како што се:

1. криптографскиот хеш од блокот хронолошки пред него;
2. податоци за рударење;
3. структура на податоци за да ги сумираат трансакциите во блокот - исто така наречени корен на Меркловото дрво.

Криптографскиот хеш во заглавието на блокот на секој блок, е она што хронолошки ги поврзува блоковите заедно во синџирот (веригата), прикажано на слика 9.



Слика 9. Хронолошко поврзување на блоковите во синџирот [10].

Како што е забележано претходно, блоковските вериги не се биткоин, туку форма на база на податоци зачувана во децентрализиран систем. Затоа, блоковските вериги можат да се прилагодат за употреба во различни области. Технологијата на блоковски вериги е приспособлива технологија која може да се модифицира за да одговара на специфични цели.

2.3.2 Консензус

Најважната карактеристика на технологијата на блоковски вериги е консензусот. Консензусот се однесува на способноста на сите анонимни учесници во мрежата. Тие се согласуваат да се следат правилата на мрежата и има само една вистина запишана во блоковите. Консензусот може да се постигне на многу различни начини, со алгоритам за доказ за работа (PoW) или со алгоритам за докажување на влогот (PoS, Proof of stake).

Дистрибуирано пресметување

Еластичноста на мрежата на блоковски вериги во голема мера ѝ се припишува на нејзината дистрибуирана архитектура. Продолжувајќи со примерот на блоковската верига на биткоин, секој корисник на биткоин-блоковската верига, којшто работи на целосен јазол на својот компјутер, ќе преземе целосна копија од целата блоковска верига. Секоја целосна копија ќе вклучува податоци за сите трансакции снимени на биткоин-блоковската верига. Откако ќе се преземе копијата, тогаш јазолот може да работи независно за да ги обработува трансакциите и да ги пропагира понатаму низ целата мрежа. Јазлите, исто така, можат да придонесат за мрежен консензус преку рударството преку вклучување на трансакциски податоци во блок и потоа наоѓање доказ за работа за блокот. Важен концепт за дистрибуираната мрежа на блоковски вериги е дека не постои обработка на централни јазли и дистрибуција на податоците, но секој јазол може да работи независно и да ја емитува која било работа што е докажана.

Складирање на информации

Во случај на блоковската верига на биткоин, информациите што се чуваат во рамките на блоковите се трансакциски податоци. Сепак, оваа функција се протега понатаму подалеку од трансакциите и може да се прошири во паметни договори, како што се користат во блоковската верига на етериум.

Потекло

Блоковските вериги обезбедуваат следење на податоците, на претходно програмиран начин. Во традиционалното банкарство се знае дека вложените пари се наоѓаат во банката, бидејќи банката тоа го гарантира. Во трансакција што користи блоковска верига, секоја активност се следи, евидентира и целосно може да се следи без потреба трета страна да ја потврди конкретната акција.

Непроменливост

Ниту еден учесник во мрежа што користи блоковска верига не може да ја менува трансакцијата откако ќе биде снимена, и притоа нема исклучоци. Ако е направена грешка, не може да се уредува или да се врати назад. Погрешен запис не може да се избрише и секогаш ќе биде видлив откако ќе се сними. За да се поправи грешката, мора да се генерира нова трансакција, која ќе се повикува на погрешен запис.

Контрола на пристап

Во споделена отворена јавна база на податоци, како што е блоковската верига на биткоин, секој има пристап за гледање и додавање на блокови. Спротивно на тоа, блоковската верига може да биде повеќе приватизирана и да има построг пристап во кој има дозволи за прегледување и уредување. Овие типови на приватизирани вериги

обично се наоѓаат во приватни претпријатија, каде што податоците имаат тенденција да бидат почувствителни.

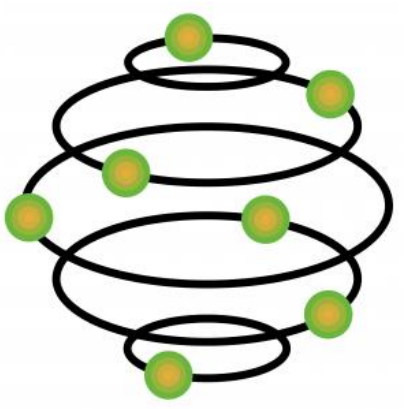
2.3.3 Видови на технологии на блоковски вериги

Технологијата на блоковски вериги постојано се развива. Поради своите основни технолошки карактеристики, постојано се развиваат нови апликации користејќи ја нејзината рамка. Во зависност од дозволата за пристап блоковските вериги се категоризираат во следните категории [41]:

- јавни блоковски вериги,
- приватни блоковски вериги и
- хибридни блоковски вериги.

Јавна блоковска верига

Јавна блоковска верига се нарекува онаа којашто не содржи апсолутно никакви ограничувања. Јавните блоковски вериги дозволуваат секој да достави податоци до главната книга со сите учесници кои поседуваат идентична копија од главната книга. Бидејќи нема единствен сопственик на главната книга, оваа методологија е посоодветна за апликациите отпорни на цензура (на пример, биткоин). Јавните блоковски вериги често нудат економски стимулации за оние кои ја обезбедуваат мрежата, прикажано на слика 10.



Слика 10. Јавна блоковска верига [10].

Јавната блоковска верига ги има следните карактеристики:

• Секој може да пристапи до неа, односно да ги види сите трансакции што се појавуваат на блокот. Постојат голем број сервиси што овозможуваат да се прегледа јавната блоковска верига познати како Block Explorer, а најпознат е blockchain.info, на кој може да се следи блоковската верига на биткоин.

• Секој може да врши трансакции. Доволно е да се преземе мобилен или десктоп паричник (wallet) или да се користи еден од онлајн паричниците и слободно да се извршуваат трансакции.

• Секој може да учествува во создавањето на блокови и во поделбата на наградата што следува за додавање на блокови. Со други зборови, секој може да биде „рудар“.

• Секој може да има удел во одлучувањето за промените и дополнувањата на протоколот кој управува со криповалутата. Кај некои криптовалути, рударите донесуваат одлуки, но постојат криптовалути во кои и други учесници имаат контролен удел.

• Протоколот што го контролира системот е во форма на отворен изворен код. Секој може да го прегледа овој код и секој може да предложи промени во тој код. Доколку мнозинството ги прифати предложените измени, тие промени стануваат составен дел од протоколот. Исто така, многу чест случај е да се земе протоколот од една криптовалута, малку да се измени и потоа да се објави како нова криптовалута.

Од целосната отвореност на јавните блоковски вериги се појавуваат некои од нивните предности и недостатоци. Главните предности на јавните блоковски вериги се:

• Блоковската верига е отпорна на потенцијални напади. Поради фактот што секој може да биде јазол во мрежа со рамноправен пристап, бројот на овие јазли е многу голем и затоа е потешко да има повеќе од 50 % недоверливи учесници, што е многу скапо.

• Базата на податоци е непроменлива. Ова повторно произлегува од фактот дека целата база на податоци со сите трансакции се пресликува на илјадници компјутери. Исто така, блоковската верига одржува мрежа на рудари со многу голема процесорска моќ. Историјата може да се промени доброволно од страна на мнозинството учесници кои се согласуваат да ја направат промената. Ова е малку веројатно во практиката, бидејќи повеќето ќе треба да се договорат за нешто што го загрозува интегритетот на мрежата и со тоа ја намалува цената на криптовалутата поврзана со блоковската верига. Историјата, исто така, може да се промени насилно, но напаѓачот треба да контролира повеќе од 50 % од процесорската моќ на целиот систем во текот на подолг временски период. Ова би било исклучително скап потфат и нема начин таква инвестиција да биде финансиски исплатлива за напаѓачот.

Главните недостатоци на јавните блоковски вериги се:

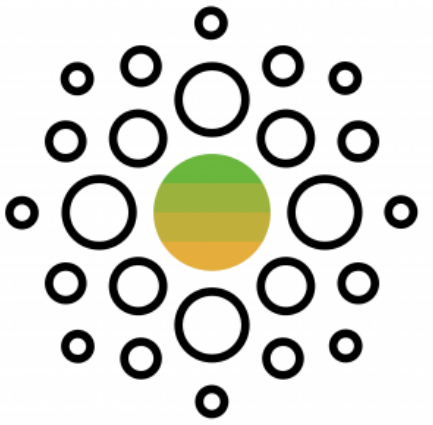
• Капацитетот на блоковските вериги е многу ограничен, и според бројот на трансакции кои можат да бидат обработени во временска единица и од количеството податоци што може да се зачуваат во блокот. За да може повеќе луѓе да учествуваат во мрежното одржување, барањата мора да бидат релативно скромни, и во количеството на простор на хард дискот кој блоковската верига го зазема и во брзината на Интернет-

врската. Мрежата на биткоин во моментов може да процесира само неколку трансакции во секунда, а за споредба, Visa може да процесира 1700 трансакции во секунда [81].

• Режимот за управување е неефикасен. Со цел да се спроведе која било, дури и најмала промена на системот, неопходно е поголемиот дел од членовите во мрежата да се согласат за оваа промена. За споредба, тоа би било приближно како да се организира референдум за секоја индивидуална одлука во земјата. Јасно е дека овој принцип го отежнува управувањето и ја намалува флексибилноста на системот. Дури и за добри предлози, не е лесно да се обезбеди мнозинство, и често не е лесно да се утврди кој предлог е добар. Одличен пример за овој проблем што се случува во биткоин мрежата последниве неколку години е дека од 2015 година наваму, постојат различни предлози за тоа како да се подобри протоколот со цел да се зголеми капацитетот на блоковската верига на биткоин, но мрежата не може да се согласи за еден од овие предлози.

Приватна блоковска верига

Понекогаш ја нарекуваат регистар со дозвола, бидејќи дозволено им е само на поканети учесници да се приклучат на мрежата. Овие мрежи се контролирани од страна на еден или серија назначени мрежни администратори. Приватните блоковски вериги дозволуваат дистрибуирани идентични копии од регистарот, но само до ограничен број доверливи учесници. Бидејќи мрежата може да има сопственик, оваа методологија е подобра за апликации кои бараат едноставност, брзина и поголема транспарентност, прикажано на слика 11.



Слика 11. Приватна блоковска верига [11].

Во последно време сè повеќе институции и компании почнуваат да покажуваат интерес за технологијата на блоковски вериги. Таа привлекува со својот многу иновативен начин на испраќање и складирање на податоци, како и со начинот на решавање на проблемот на доверба во систем со повеќе корисници. Она што не им се допаѓа е транспарентноста и тоа дека системот е достапен за секого. Затоа се појавила идејата за приватни блоковски вериги што ги задржува повеќето придобивки од

јавните блоковски вериги, но исто така ќе ги елиминира недостатоците што не им одговараат на компаниите.

Разлики меѓу приватни и јавни блоковски вериги:

• Приватната блоковската верига не е видлива за секого. Таа е видлива само за оние што имаат дозвола. Лиценцата најчесто се издава од страна на создавачот или сопственикот на блоковската верига, со можност за издавање на различни категории на лиценци во кои количината на видливи податоци е различна.

• Не може секој да прави трансакции. Трансакциите се овозможени само за оние со дозволи. Лиценцата може да биде издадена од сопственикот на блоковската верига, но исто така и од некои од организациите што учествуваат во системот.

• Не може секој да креира блокови, туку само оние кои се овластени од страна на сопствениците на блоковските вериги. За разлика од повеќето јавни блоковски вериги, тука создавањето на блокови обично не вклучува ангажирање на моќен и скап хардвер за рударење.

• Во одлуките не учествува секој. За сите промени во протоколот одлучуваат само оние што се делегирани од сопствениците на блоковската верига.

• Кодот не е јавно достапен. Понекогаш тоа не е достапно за сите учесници во системот. Промена на кодот не може да предложи кој било, но постои тим кој се занимава со развојот на протоколот.

• Приватната блоковска верига не мора да има криптовалута. Во јавните блоковски вериги, криптовалутата служи за мотивирање на рударите да помогнат во одржувањето на системот. Овде, системот најчесто се одржува од организации што се дел од системот. Вредноста на блоковските вериги расте со бројот на корисници, како јавните, така и приватните. Затоа постојат многу чести примери на група на компании што работат заедно за развој на проект за блоковски вериги кој подоцна ќе го користат заедно. Најпознати компании коишто учествуваат во вакви проекти се Hyperledger, Enterprise Ethereum Alliance и R3 [5].

Главните предности на приватните блоковски вериги се:

• Капацитетот на таква блоковска верига може да биде екстремно голем. Приватните блоковски вериги се дел од компанија, којашто може да обезбеди голем простор за складирање каде што се чува блоковската верига, како и брза Интернет-врска конекција што би можела да поддржи голем број трансакции. Како резултат на тоа, таквата блоковска верига може да биде значително поголема и побрза од јавната.

• Управувањето е поефикасно. За разлика од јавните блоковски вериги каде што илјадници луѓе се вклучени во донесување одлука, приватните блоковски вериги најчесто се занимаваат со десетици компании, а понекогаш и само неколку. Затоа

многу полесно може да се постигне консензус за секоја одлука, така што промената на протоколот може да се спроведе полесно и побрзо. Сепак, фактот дека бројот на учесници е помал не значи дека секогаш ќе биде лесно да се договорат за сè. На пример, неколку компании го напуштиле конзорциумот R3 бидејќи не се согласувале со насоката на развојот на блоковски вериги што им се наметнува.

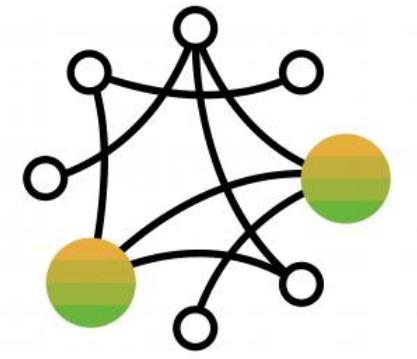
Главните недостатоци на приватните блоковски вериги се:

• Премногу моќ во рацете на мал број компании. Дури и со приватни блоковски вериги, одлуките обично се донесуваат со консензус. Сепак, тука е потребно значително помал број учесници за да се обезбеди мнозинство кое може да ја преземе контролата врз системот. Често, бројот на оние кои го контролираат системот е уште помал отколку на прв поглед. Имено, бидејќи на системот му е дозволен пристап, некој мора да издаде дозвола. Бројот на издавачи често е помал од вкупниот број на корисници. Оние кои ја издаваат лиценцата, всушност, го контролираат системот, бидејќи можат да ја одземат дозволата на оние што се непослушни, и наместо тоа, ги стават во послушни и со тоа го обезбедуваат мнозинството.

• Нов тип на напад на системот. Податоците што се зачувани во блоковските вериги се достапни само за учесниците во системот. Овие податоци се од голема вредност за нив, и тие стануваат многу интересни цели за хакери кои, со пристап до овие податоци, би можеле да ги загрозат учесниците во системот или да ги уценуваат.

Хибридна блоковска верига

Хибридните блоковски вериги, наречени блоковски вериги на конзорциум, се сметаат за полу-децентрализирани и користат карактеристики на јавни и приватни блоковски вериги. Хибридните блоковски вериги содржат групи на дозволи, слични на приватните блоковски вериги, меѓутоа, наместо една организација, контролираат група на договорени организации. Администраторите на секоја организација може да ги ограничат правата на читање на корисниците како што сакаат и дозволуваат само ограничено множество на доверливи јазли за извршување на консензус протокол, прикажано на слика 12.



Слика 12. Хибридна блоковска верига [11].

Иднината на технологијата на блоковски вериги, најверојатно нема да резултира во една блоковска верига штоа ги отфрла другите [11]. Наместо тоа, најверојатно ќе има многу примени на технологијата на блоковски вериги и би било од корист за сите, ако тие се меѓусебно поврзани. Како резултат на ова се појавуваат различни примени со отворен код, за неутрални протоколи и стандарди. Најпознатите конзорциуми и колаборативни проекти кои ја користат оваа технологија се Hyperledger, R3 и Interledger.

2.3.4 Рударење

Креирање блокови

При креирање на блок, јазлите ги потврдуваат трансакциите, а работата на рударите е да ги спакуваат трансакциите коишто претходно биле потврдени од јазлите во блокови. Постојат одредени правила за создавање блокови [100]:

• сите рудари во светот се натпреваруваат во создавањето блокови.

• блокот не смее да биде поголем од 1 мегабајт (MB), така што бројот на трансакции што може да се сместат во блокот е ограничен.

• кога некој ќе предложи победнички блок, тој ќе им го испрати на другите во peer-to-peer мрежата со рамноправен пристап за проверка. Ако јазлите утврдат дека блокот е валиден, тој се додава во синџирот од блокови.

• победникот добива 12,5 биткоини + провизија за сите трансакции коишто се внесени во тој блок (провизии се плаќаат, бидејќи бројот на трансакции што се кандидати за да влезат во блокот е поголем од капацитетот на блокот, а рударите ги избираат оние трансакции со поголеми провизии со цел да заработат повеќе приходи).

• штом ќе се додаде нов блок, трката започнува повторно и сите се обидуваат први да го креираат следниот валиден блок.

• победник е оној кој прв создава блок чиј хеш е помал од даден број. Дадениот број е директно поврзан со тежината на рударењето. Колку е поголем бројот на рударите и нивната збирна процесорска моќ толку тежината е поголема, а бројот е помал.

• овој зададен број автоматски се прилагодува, така што блоковите се создаваат во просек на секои 10 минути. Прилагодувањето се врши на секои 2016 блокови, што во просек е 2 недели [101]. Ако во претходните 2016 блокови се зголемила процесорската моќ на мрежата, тој број ќе се намали. Ако процесорската моќ се намали, овој број ќе се зголеми.

• на секои 4 години, наградата на рударите се преполовува. Во првите 4 години, наградата била 50 биткоини по блок, а потоа паднала на 25 биткоини по блок, а во моментов изнесува 12,5 биткоини по блок [101].

Алгоритам - доказ за работа (Proof-of-Work, PoW)

Еден блок ги содржи следните делови:

• хеш на претходниот блок, така што блоковите се поврзани во синџир.

• *coinbase* трансакција - ова е трансакција со која рударите самите назначуваат нови 12,5 биткоини, како награда за тоа што први создавале валиден блок. Тие 12,5 биткоини не ги земаат од никаде, туку ги создаваат, и само на тој начин се внесуваат во оптек нови биткоини.

• трансакциите кои рударите избрале да ги спакуваат во тој блок (обично околу 2500, зашто не може да се стават повеќе поради ограничувањето на големината на блокот до 1 MB).

• Nonce - број во опсег од 0 до 2^{32}, во обид да се најде оној број за кој хешот на блокот ќе има соодветна вредност (пониска од стандардната).

Поради особината на хеш-функцијата, рударите немаат начин да ги најдат потребните влезови за да го добијат резултатот што го сакаат. Затоа, тие треба да пробаат различни вредности на „nonce“ број еден по еден за да го добијат посакуваниот резултат. Нема скратен начин за да се добие резултат, туку мора да се пробуваат сите броеви еден по еден додека не се погоди вистинскиот. Овој метод се нарекува метод на груба сила (brute force). Поради ова, важно е рударите да имаат што помоќен хардвер, за да направи што повеќе обиди за пократко време, зголемувајќи ја шансата да биде прв што ќе создаде валиден блок. За да се пронајдат неопходните влезни податоци што даваат вистински резултат, потребен е многу голем број на обиди и за да се направи голем број обиди за кратко време, потребна е голема процесорска моќ и големо количество работа. Бидејќи рударот мора да вложи многу работа за да добие резултат, овој алгоритам за рударење се нарекува доказ за работа (PoW). Кај овој алгоритам, добро е тоа што е потребно многу малку време и работа за да се провери валидноста на резултатот. Кога еден рудар создава блок што ги исполнува бараните услови, тој го испраќа до остатокот од мрежата за да се потврди. Многу лесно и брзо се проверува дали за дадените влезни податоци се добива потребниот хеш-резултат и ако се потврди дека е во право, тој блок официјално станува составен дел од блоковската верига и добива награда од 12,5 биткоини + провизија. Ако блокот не е валиден, мрежата ќе го отфрли тој нов блок и потрагата по „победникот“ што прв ќе создаде нов блок продолжува [16].

Блокови сирачиња

Понекогаш се случува двајца рудари речиси во исто време да најдат нов блок што ги исполнува бараните услови. Ако и двата блока се валидни, и двата ќе бидат прифатени од мрежата, но не од целосната мрежа. Во мрежата со рамноправен пристап (peep-to-peer) информацијата патува од јазол до јазол сè додека не стигне до сите. Кога некој рудар ќе најде блок, тој прво ја испраќа оваа информација до јазлите со кои е директно поврзан. Ако друг рудар најде нов блок во речиси исто време, тој сè уште не добил информација дека некој друг создал блок и тој, исто така, испраќа информации за својот нов блок до јазлите со кои е директно поврзан. Од две различни страни на мрежата, во исто време се лансираат два различни блока во мрежата и секој е валиден. Секој од јазлите ќе го прифати блокот што ќе стигне прв, а другиот ќе го отфрли. Кога препраќањето на новите блокови е завршено, ситуацијата ќе биде таква што на дел од мрежата има еден блок како последен, но дел од мрежата ќе има друг блок како последен (сите други блокови се исти). Во моментов постојат две верзии на блоковската верига, дел од мрежата има една верзија, а дел од мрежата друга верзија, се разликуваат само во последниот блок. Првиот дел не ги препознава трансакциите во последниот блок од вториот дел од мрежата и обратно (иако некои трансакции веројатно се совпаѓаат). Рударите во првиот дел од мрежата се обидуваат да го надградат новиот блок на нивната верзија на блоковската верига, а рударите во другиот дел на мрежата се обидуваат да го надградат новиот блок на нивната верзија на блоковската верига. Која група ќе успее прва да го направи тоа, ќе ја наметне својата верзија на блоковската верига на целата мрежа.

Јазлите комуницираат едни со други и ја споредуваат нивната блоковска верига со „соседите“. Ако се забележи дека некоја блоковска верига е со поголем број на блокови, се претпоставува дека таа е поддржана од страна на мнозинството на мрежата и се презема оваа верзија на блоковската верига. Во ваков случај, тие практично ќе го отфрлат нивниот последен блок како невалиден, заедно со сите трансакции во него. Затоа овој блок се нарекува „сираче“. Теоретски е можно два пати по ред двајца рудари во исто време да создадат нов блок, но шансите за тоа се многу мали. Веројатноста тоа да се случи три пати по ред е занемарливо мала, и затоа повеќето сметаат дека трансакцијата со биткоини е конечна ако од моментот на испраќање на овие трансакции се додадат три нови блока (конечната трансакција мора да биде во првиот од трите блока) [100].

Рударски здруженија

Бројот на рудари во светот е висок и постојано расте. Можноста за некој поединечен рудар прв да го креира следниот блок е многу мала. За среќа, блоковите се појавуваат во просек на секои 10 минути, но честопати тоа не е доволно брзо. Да се земе пример еден рудар чија процесорска моќ е 0,01 % од вкупната процесорска моќ за рударство во светот (тоа претставува многу сериозен рудар). Статистички гледано, на

10.000 додадени блокови еден е негов, што значи дека, во просек, на секои 70 дена тој ќе создаде еден блок и ќе заработи 12.5 биткоини + провизија [101]. Но, тоа е само статистика и во никој случај не е гарантирано. Во практиката, таков рудар може да создаде два блока за еден ден, но може да не создаде ниту еден блок во период од две години. Бидејќи рударите имаат фиксни трошоци, важно е да имаат предвидливи приходи. Затоа, постојат здруженија за рударење, попознати како базени (pools). Таквото здружение обединува голем број рудари кои имаат договор дека наградата што ќе ја добие тој што ќе додаде нов блок во блоковската верига во рамките на здружението, ја делат сите, пропорционално на процесорската моќ што ја даваат на располагање во здружението. Сопственикот на базенот ја зема оваа посебна такса, која обично е неколку проценти. Ако здружението за рударење е поголемо, тоа ќе создава блокови секој ден, така што сите рудари во базенот ќе имаат дневни исплати, колку и да се мали. Ваквиот начин на работа обезбедува редовни приходи за рударите.

Рударење во облак

Рударството бара соодветен физички простор, добри електрични инсталации, добра вентилација/ладење и звучна изолација. Некои се откажуваат од рударството поради овие компликации. Како решение за овие проблеми создадено е „рударење во облак“. Тука човекот наместо да купува опрема за рударење, практично плаќа закупнина за некоја постоечка рударска опрема. На овој начин учествува во рударството, иако физички не е сопственик на хардвер, и затоа не треба да се грижи за струја, ладење, бучава, сигурност и слични работи што им прават проблеми на рударите со сопствена опрема.

Иако рударењето во облак звучи одлично во теорија, во практиката не е така. Проблемот е во тоа што не може со сигурност да се знае што всушност се изнајмува и дали изнајменото работи за тој што ја плаќа киријата. Многу висок процент на компании кои нудат услуги како што е рударење во облак се всушност измами. Тие немаат опрема или имаат многу мала и не рударат. Измамничката шема е таква што, тие со парите од новите инвеститори ги исплаќаат старите. Од гледна точка на инвеститорите, функционираат сè додека има редовни плаќања. Овој систем функционира сè додека редовните исплати не го надминуваат износот на парите што се добиени од нови инвеститори. Кога тоа ќе се случи, тие престануваат да плаќаат. Некои го прават тоа под изговор дека се нападнати од хакери, некои се правдаат со привремени технички проблеми, а некои, едноставно, ја деактивираат веб-страницата и исчезнуваат. Компании коишто вистински се занимаваат со рударењето во облак, за жал, се многу ретки.

Кога во јануари 2009 година започнал да се рудари биткоинот, тоа го вршеле процесори. Биткоинот тогаш вредел многу малку, па никој не гледал интерес за оптимизирање на рударството и за поефикасно работење [101].

Во средината на 2010 година, полека почнало да се тргува со биткоинот, па затоа рударството полека почнало да станува профитабилно и луѓето почнале да бараат нови начини за рударство. Во јули 2010 година се дошло до заклучок дека со соодветен софтвер, графичките картички можат значително поефикасно да рударат од процесорите. Графичките картички трошат многу електрична енергија, генерираат голем шум и генерираат голема количина на топлина, па затоа морало да се бара алтернатива. Во мај 2011 година се појавиле алтернативите FPGA (Field-Programmable Gate Array) плочи. Тие не биле многу силни, но трошеле неспоредливо помалку електрична енергија отколку графичките картички. Никогаш не успеале целосно да ја заменат графичката картичка, бидејќи биле релативно скапи. На крајот на јануари 2013 година се појавиле ASIC (Application-Specific Integrated Circuit) машини со кои се отишло еден чекор понатаму [88]. Ова се уреди кои се дизајнирани и веќе изработени во фабрика за само една специфична функција, а тоа е рударството.

Други пристапи кон рударството

Високата потрошувачка на електрична енергија се споменува како сериозен проблем на биткоин рударството. Алгоритамот за доказ за работа (PoW) е таков што растот на мрежата неизбежно води до зголемување на потрошувачката на електрична енергија. Биткоинот е најочигледен пример, но, генерално, сите криптовалути кои го користат алгоритамот доказ за работа (PoW) го имаат истиот проблем. Други позначајни криптовалути кои го користат алгоритамот доказ за работа (PoW) се етериум, зикеш и лајткоин (Ethereum, Zcash, Litecoin).

Најпопуларниот алтернативен алгоритам за рударство е доказ за вложување (Proof-of-Stake, PoS). Во овој модел, рударите примаат награда која не е пропорционална на моќта на нивниот хардвер, туку во зависност од количеството на криптовалути што веќе се поседуваат. Најчесто, потребна е одредена минимална сума што некој мора да ја има за да може да рудари. Поради многуте различни пристапи, многу луѓе не го нарекуваат овој процес „рударство“, туку „ковање“. Некои од примерите каде се користи алгоритамот доказ за вложување (PoS) се криптовалутите стратис и нео (Stratis, NEO). Правото на одлучување во овој алгоритам е пропорционално со количеството на криптовалутите што некој ги поседува и овој алгоритам работи со следната логика: еден човек да има поголема количина од некоја криптовалута нема да донесува одлуки што се на штета на таа криптовалута (во спротивно би си штетел самиот на себе). Ова е слично на модел за управување со компанија каде што „тежината“" на нечиј глас е директно пропорционална на бројот на акции од компанијата што тој човек ги поседува.

Варијација на овој модел претставува алгоритамот делегиран доказ за влог (Delegated Proof-of-Stake, DPoS) каде на одреден број на „делегати“ им е доверено одржување на мрежата - како што е BitShares [89]. Постојат криптовалути кои комбинираат доказ за работа (PoW) и доказ за влог (PoS) и, на некој начин, се

хибридни системи. Тие се обидуваат да ги искористат предностите на секој од алгоритмите. Во хибридниот систем, дел од придобивките од создавање на нови блокови оди на оние кои придонесуваат за одржување на хардверската мрежа, а дел оди на оние кои имаат одреден број на криптовалути. Примери за таков рударски алгоритам се деш и пиркоин (Dash, Peercoin).

Доказ за простор (Proof-of-Space) или доказ за капацитет (Proof-of-Capacity**)** е алгоритам во кој место процесорската моќ, системот на располагање го дава просторот за складирање на хард-дискот. Колку повеќе простор е достапен во мрежата, толку поголема награда добива во криптовалутата што се рудари на тој начин. Таков пример е бурсткоин (Burstcoin). Во приватни блоковски вериги се користат и следниве алгоритми: доказ за авторитет и доказ за изминато време (Proof-of-Authority, Proof-of-Elapsed-Time) [90].

2.4 Преглед на примена на технологијата на блоковски вериги

Криптовалута. Првата апликација на технологијата на блоковски вериги е криптовалутата биткоин. Можноста да се воспостави доверба во децентрализирана мрежа им овозможува на трансакциите да се извршат на брз, евтин и безбеден начин, без никакви ограничувања.

Складирање на податоци. Кога датотеката е многу важна се прави резервна копија. Блоковската верига може да има илјадници ажурирани копии од комплетната база на податоци во секое време. Јавните блоковски вериги имаат ограничен капацитет во однос на количеството податоци што можат да го зачуваат, додека приватните не можат да бидат доволно сигурни за таа намена.

Интегритет на податоци. Речиси невозможно е да се променат податоците во блоковските вериги.

Контрола на потрошувачката. Сите трансакции кои некогаш се случиле се видливи во блоковските вериги и тоа може да овозможи да се провери дали одредени средства се трошат на оној начин на кој е планирано.

Определување на време. За секоја трансакција снимена во блоковската верига е запишано и времето кога таа се случила, така што секогаш може прецизно да се утврди текот на настаните. Ефективно, биткоинот користи блоковска верига за децентрализирање на плаќањата.

Би било корисно сите да имаат пристап до децентрализиран извор на евиденција (катастар) во кој е запишана сопственоста за секоја парцела на земјиште. Со оглед на тоа што државните удари и војни честопати прераспределуваат земја неправедно и/или погрешно, тоа не само што може да се покаже како корисно, туку може и да има и хуманитарни импликации. Кога е договорена распределба на земјиште, таа може да

биде запишана во дистрибуирана книга. Постојат компании што работат на овој принцип и нивниот број веројатно ќе расте.

Во иста насока, блоковската верига може да се искористи за да се евидентира сопственост над кое било физичко средство - автомобил, уметничко дело, музички инструмент итн. Евиденцијата за сопственост на хартија подлежи на фалсификување и/или физичка деградација. Централизирани бази на податоци се мета на хакерски напади и човечки грешки (случајни или намерни). А блоковска верига значи дека нема единствен ентитет што ја контролира главната книга. Затоа, евидентирањето на физичките средства за следење на сопственоста на блоковски вериги е одличен пример за тоа каде технологијата би можела да биде корисна.

Технологијата на блоковски вериги може да се покаже како применлива во виртуелната реалност. Ако се создаде виртуелен свет - за игри или за други цели - оваа технологија може да им овозможи на корисниците да купуваат и да поседуваат делови од тој виртуелен свет, исто како што би можеле да купат физичка парцела. Децентраленд (Decentraland) е проект што веќе работи на тоа со својот токен MANA, и ветува дека ќе ја изгради „првата виртуелна платформа во сопственост на своите корисници“ [82].

Идентитетот, исто така, може да биде пример каде што би можела да се примени оваа технологија. Наместо државата или владата која го издава, идентитетот може да биде верификуван на отворена, глобална блоковска верига, што не е контролирана од никој, а е доверлива за сите. Така, корисниците можат да го контролираат сопствениот идентитет. Голем број на компании работат на оваа проблематика, вклучувајќи ја и глобалната алијанса посветена на подобрување на животот преку дигиталниот идентитет (ID2020, Digital Identity Alliance).

Слично на тоа, компанијата за „јавно добро“, Blockstack се надева дека ќе изгради нов децентрализиран Интернет, каде што корисниците ги поседуваат своите податоци и апликации на локално ниво [91]. Ако тоа проработи, Blockstack може да им го наруши работењето на големите Интернет гиганти кои дејствуваат како посредници денес Google и Facebook.

2.4.1 Алткоини

Името алткоини се користи за да се опише категорија на криптовалути коишто се алтернатива на биткоинот. Кога биткоинот станал првата успешна дигитална валута, многу други се обиделе да го копираат. Алткоините се обидуваат да надминат некои од ограничувањата што ги има биткоинот, т.е. да се замени или надгради барем една компонента што ја има биткоинот. На пример, новиот и посовремен алткоин влегува на пазарот и ги поместува границите на брзината на трансакција, капацитетот и приватноста.

Првите алтернативи додале само малку поголема вредност од биткоинот и, првично се обиделе да го копирааат неговиот успех. Покрај тоа, повеќето од денешните криптовалути се само копии од биткоинот со мали, непогрешливи промени. Тие имаат помала рударска моќ зад нив, имаат помалку програмери кои ги подобруваат и се помалку корисни поради малиот ефект што мрежата го создава (мал број на учесници) [14].

Раздвојување (fork)

Раздвојување е настан што се јавува кога блоковската верига се дели на два дела. Ова може да се случи кога се создава ново правило што опишува која трансакција е валидна. Корисниците на блоковската верига треба да ги поддржат предложените модификации на протоколот или да ги отфрлат. Ако корисниците не се согласат за поддршка за предложените промени, мрежата се дели на два дела и две различни блоковски вериги се создаваат преку процес познат како „раздвојување“ (fork).

Биткоинов кеш

Биткоинот (BTC) доживеал тешка поделба (fork) и на тој начин се формирал биткоинов кеш (BCH). Оваа поделба се случила кога бројот на трансакции на биткоинот се зголемил толку многу што корисниците морале да чекаат долго време за извршување на нивните трансакции. Ова создало низа трансакции кои се акумулираат, чија цена е значително повисока од редовната и за обичните корисници премногу голема за извршување. За да се надмине ова, неопходно било да се бара ново решение. Протоколот за зајакнување на биткоинот (BIP) 148 започнал измени со предлагање на нов протокол SegWit што создава повеќе простор за повеќе трансакции во блоковите [109].

За оние што не сакале да ја прифатат оваа промена, Bitmain Network претставил HardFork, UAHF решение со големина на блокот од 8 MB, без SegWit, и уште три други начини за подобрување на оригиналниот биткоин. Бидејќи биткоиновиот кеш (BCH) е резултат на тешка поделба (HardFork), секој што поседувал биткоини, добил иста вредност во токени на биткоинов кеш [109]. Биткоиновиот кеш варира од неговото создавање до денешен момент со вредноста на своите токени, така што во моментов е четврта криптовалута, по биткоин, етериум и рипла [16].

Лајткоин

Лајткоин (Litecoin LTC) се базира на отворениот изворен код на биткоин, но со неколку технолошки разлики: алгоритамот за рударство е „scrypt“ наместо „SHA-256“, рударството бара повеќе меморија, а не повеќе процесорска моќ. „Scrypt“ е исто така криптографска функција, но многу поедноставна од SHA-256. Затоа, рударството базирано на скрипти е побрзо и поедноставно, но, од друга страна, таков систем е поранлив од безбедносен аспект. Во системот на лајткоин, додавање на нов блок во блоковската верига се одвива во просек на секои 2,5 минути, така што произведува

повеќе парични единици, а вкупниот износ на пари што ќе биде издаден е 84 милиони [16].

Зикеш

Зикеш (Zcash, ZEC) е криптовалута што потекнува од проектот Zerocoin. Протоколот Zerocoin, кој првично требаше да ја подобри анонимноста на корисниците на биткоинот, стана ZeroCash, што резултираше со појава на нова криптовалута, зикеш, во 2016 година. Трансакциите на зикеш јавно се објавуваат во блоковска верига, но корисниците имаат можност да ја користат опцијата за приватност и да го сокријат испраќачот, примачот и износот на токени во трансакцијата. Ова е една од клучните разлики во однос на Monero, каде што сите трансакции се чисто приватни [112].

Даш

Даш (DASH) е криптовалута што има софтер со отворен код и рамноправен пристап, чија цел е да биде што е можно попријателски ориентирана кон корисниците. Во прилог на сите карактеристики што ги поседува биткоинот, даш нуди инстант трансакции (InstantSend) и приватни трансакции (PrivateSend). Можноста за управување и буџетскиот систем ја направи даш децентрализирана автономна организација (Decentralized Autonomous Organization, DAO). Во јануари 2014 година, првично била објавена како екскоин (XCoin, XCO), во февруари името било променето во даркkоин (Darkcoin). Сегашното име го добила во март 2015 година [113].

Рипла

Рипла (Ripple, XRP) се базира на протокол кој ги регулира трансакциите кои се користат за размена меѓу банки и фирми со официјален токен XRP. Мрежата е целосно децентрализирана и може да функционира без Рипла како организација. Во суштина, риплата се базира на јавно споделена база на податоци, која користи консензус за одобрување на плаќања и размена [114].

Initial Coin Offering (ICO)

Initial Coin Offering (ICO) е нов начин на зголемување на капиталот за финансирање на проекти, каде што стартап или онлајн проекти добиваат инвестиции преку создавање и продажба на нивните криптовалути без да продаваат акции на своите компании или да пристапат до големи капитални фондови. Кога стартап претпријатието сака да ги зголеми средствата преку ICO, треба преку низа чекори да ја приближи идејата до инвеститорите. Најчесто, тоа се прави преку изработка на бела книга (whitepaper), во која се специфицираат следните информации:

- за што е проектот и што е потребно за негово успешно спроведување?
- колку пари се потребни и од кој тип?
- колку пари основачите планираат да задржат за себе?

- колку трае кампањата?

Првата кампања од овој тип е направена од страна на Mastercoin во јули 2013 година [16]. Овој релативно нов начин на финансирање за развој на нови апликации предизвика многу контроверзии околу нивоата на ризици што произлегуваат од овој начин на инвестирање. Стравот најмногу произлегува од тоа што овој екосистем функционира целосно надвор од моменталниот финансиски систем, регулаторните стандарди и практики. Притоа, многу проекти се обидуваат да ги измамат инвеститорите со тоа што по подигнувањето на средствата ќе заврши проектот. Привлечна работа околу ICO е тоа што инвеститорите можат да инвестираат колку што сакаат, што е многу слично на кампањите за групно финансирање (crowdfunding) на кикстартер (kickstarter) [16].

Иако официјално луѓето купуваат токени за да можат да ги користат услугите на апликациите, во реалноста најголема причина за купување е очекувањето дека цената на токенот ќе се зголеми. Овие очекувања во суштина не се неосновани. На пример, кога токенот стратис излегол во продажба во јули 2016 година, бил вреден помалку од цент, а веќе во октомври 2017 година изнесувал 3,41 долари, што е 487 пати повеќе [5]. Вредноста на средствата што се подигнати преку ICO во 2017 година е преку 2 милијарди долари, што е повеќе од 10 пати поголемо отколку во 2016 година, што јасно покажува каков тренд има оваа појава.

Пиркоин (Peercoin)

Пиркоин е создаден во 2012 година и е првата криптовалута за која се користи алгоритам за докажување на влогот (proof-of-stake, PoS) за да се постигне консензус. Алгоритамот за доказ за работа (PoW) се користи комбиниран со алгоритамот за доказ за влогот (PoS) [110]. Пиркоин нема дефинирана горна граница за максималниот износ на издадена валута, а другите карактеристики се слични на биткоин. Бидејќи алгоритамот на биткоин е објавен како отворен код, на овој начин се создадени многу нови криптовалути. За да се разликуваат од биткоинот и да се подобрат нивните својства, новите валути вообичаено менуваат некои од функциите на биткоин, како што се: максималниот износ на валута, наградата по блок, времето за потврда на трансакцијата, трансакциските такси и сл.

Прајмкоин (Primecoin)

Прајмкоин е создаден во 2013 година и припаѓа на новата генерација на криптовалути чии креатори го сметаат за бескорисен алгоритамот за доказ за работа (PoW) за да постигнат консензус. Таквите системи се обидуваат да ја искористат моќта на компјутерот што се троши кога се создава нов блок за решавање на друг корисен проблем. Со потрагата по задоволителен хеш, рударите партнери во системот на прајмкоин градат синџир на прости броеви. На овој начин рударите откриваат и пресметуваат сè поголеми и поголеми прости броеви коишто можат да се користат во

други научни дисциплини. Секој блок во блоковската верига на прајмкоин има снимен и откриен прост број, така што дневниците на трансакциите и евиденцијата на научните податоци на истата блоковска верига се одржуваат истовремено. Генерирање на нов блок во случај на оваа енкрипција трае околу 1 минута и нема горна граница за бројот на единици од криптовалутата кои ќе бидат издадени во иднина [111].

Бајткоин (Bytecoin)

Се појавил во 2012 година и припаѓа на групата анонимни криптовалути. Најпопуларната криптовалута, биткоин, е критикувана поради тоа што не може да остане целосно анонимна при употребата, зашто со анализа на големи податоци може да го предвиди однесувањето на корисниците. Партнерите во системот на биткоин ја содржат адресата на корисникот во форма на јавен и таен клуч. Кога корисникот ќе изврши трансакција, таа го содржи неговиот јавен клуч, и велиме дека корисникот ја потпишал трансакцијата. Бајткоин воведува иновација наречена прстенско потпишување [115]. За секоја трансакција во мрежата бајткоин, дефинирана е корисничка група, секоја со свои јавни и тајни клучеви. Трансакцијата го содржи јавниот клуч на некои од членовите на групата, но остатокот од системот е невозможно да утврди кое е тоа лице. Нов блок во системот се генерира на секои 2 минути, а вкупниот број на единици на оваа криптовалута е ограничен на 184 трилиони [115].

Фрајкоин (Freicoin)

Криптовалутата фрајкоин била воведена во 2012 година. Таа користи ист алгоритам за докажување на работа (PoW) како биткоионот и вклучува нов блок на секои 10 минути. Ова е пример за криптовалута кај којашто за прв пат е воведна негативната каматна стапка на вредноста на складирање. 4.5 % камата се одзема од вредноста на фрајкоин, со цел да се стимулира потрошувачката и да се намалат заштедите или да се акумулираат пари. Фрајкоин како валута не заживеала, но интересен пример е колку разновидни монетарни политики може да се спроведат преку криптовалута [116].

Нејмкоин (Namecoin)

Нејмкоин е, исто така, криптовалута, но ова, всушност, е систем што користи технологија на блоковски вериги во која криптовалутата има секундарна улога. Нејмкоин е изведен од биткоин и ги има истите карактеристики со биткоинот како криптовалута. Всушност, овие две валути го користат истиот програмски код во позадина. Нејмкоин, од друга страна, претставува дистрибуирана платформа за регистрација која ги поддржува записите во форма клуч-вредност. Корисниците на таква блоковска верига можат да снимаат информации како електронски адреси, криптографски клучеви, SSL сертификати, дигитални потписи (такви потписи овозможуваат идентификација и верификација на содржината на компјутерските датотеки) и многу повеќе. Најважната улога на оваа блоковска верига е да се одржи

евиденција на регистрирани имиња на домени, исто како дистрибуиран систем за регистрација на домен кој го користи Интернетот [117].

Нејмкоин се користи како алтернатива на постоечката централизирана апликација DNS (Domain Name System) доменски именски систем, со главен домен наречен .bit. Централизираниот доменски именски систем е пример за апликација изградена на клиент : сервер модел. Единствена функција на DNS-серверот е да преведува симболични имиња на компјутери во IP адреси. Дури и со намалување на пониските хиерархиски нивоа на серверот, властите не можат да воведат правила и да влијаат на доменот на највисоко ниво.

Криптовалутата нејмкоин (ознака за криптовалутна единица е NMC) е дел од системот на Namecoin. Се користи за плаќање на таксата за регистрација и пренесување на името. Namecoin поддржува неколку видови трансакции чија цел е да се зачуваат податоците во блоковски вериги. Преглед на овие трансакции е даден во табела 3, а надомест од 0,005 NMC се додава на трансакциската цена. Ограничување на вкупниот број на издадени NMC е 21 милион, а за создавање на секој нов блок на рударот му се доделуваат 50 новоформирани единици од криптовалутата нејмкоин [12].

Табела 3. Преглед на трансакции во системот на Namecoin.

Име	**Цена**	**Опис**
name_new	0.01 NMC	Дозволува корисникот да се претплати на домен
name_firstupdate	0 NMC	Трансакцијата го регистрира името и го објавува
name_update	0 NMC	Ви овозможува да го промените или обновите доменот

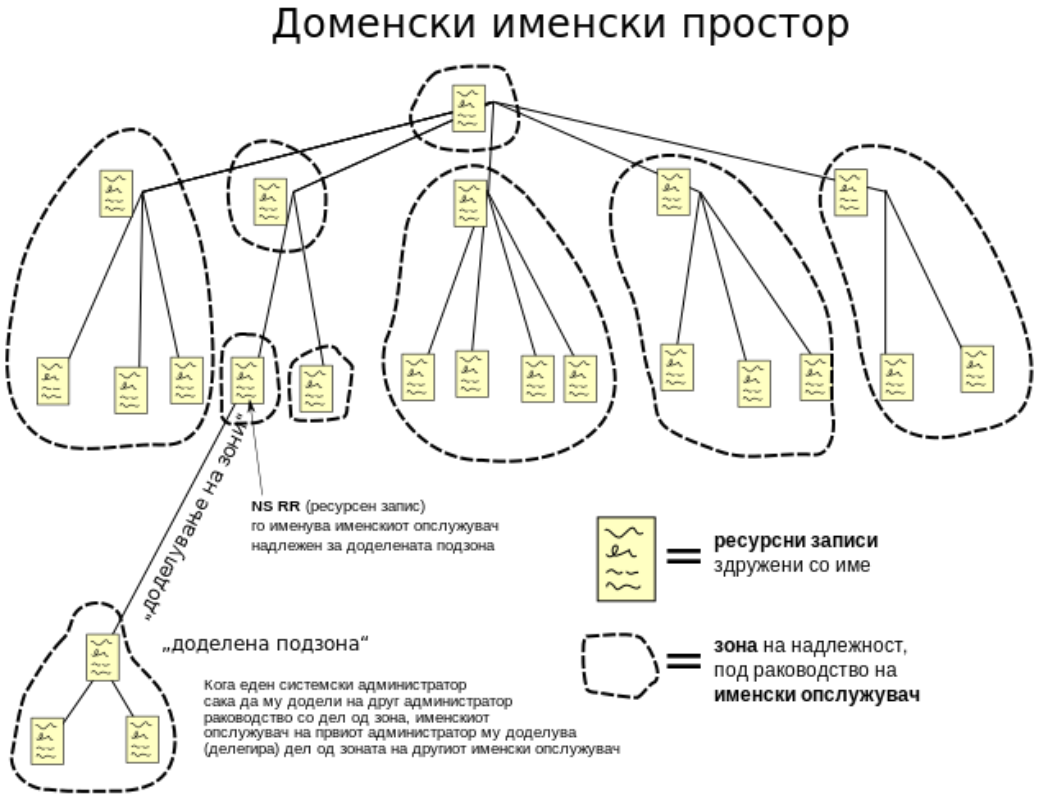
2.4.2 Споредба на централизиран и децентрализиран доменски именски систем (DNS)

Централизиран доменски именски систем (DNS)

Централизираниот доменски именски систем (DNS) има дистрибуирана база на податоци која е индексирана според имињата на домените. Секое име на доменот претставува само патека во една голема структура податоци на инвертирано дрво, наречено простор за имиња на доменот.

На слика 13, е прикажана структурата на Unix систем на датотеки. Дрвото има единствен корен на врвот. Во Unix систем за датотеки тој е наречен „root“ директориум и е претставен со симбол (/). Доменскиот именски систем овој директориум го

нарекува „корен“. Како и датотечните системи, DNS-дрвото може да се разгранува на поголем број на гранки после јазлите. Висината на дрвото е ограничена на 127 нивоа.



Слика 13. Доменски именски простор [92].

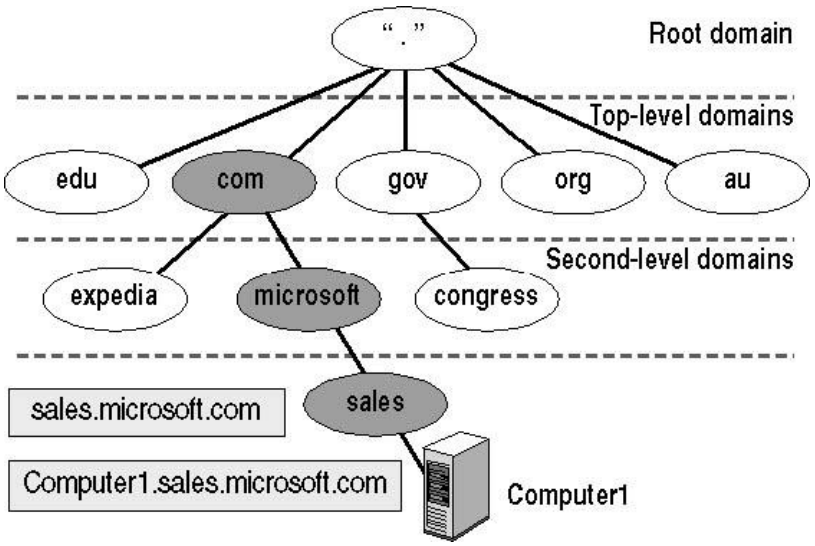
Секој јазол во дрвото има поле со текст кое може да биде со должина од 63 карактери. За коренот е резервирано поле за текст со нулта должина (null). Целосното име на доменот за кој било јазол во дрвото е низа од полиња со текст од тој јазол до коренот. Имињата на доменот секогаш се читаат од јазолот кон коренот, имињата во патеката се одделуваат со точка.

Доколку полето со името на коренот се појавува во името на доменот на јазолот, името ќе се прикаже како да завршува со точка, што, всушност, е точката која го одделува полето за текст со нулта должина на коренот. Кога единствено се појавува името на коренот се запишува со една точка.

Некои софтвери точката која е на крајот на името ја интерпретираат како индикатор дека се работи за апсолутно име на доменот. Апсолутното име се пишува релативно на коренот и недвосмислено ја означува локацијата на јазолот во хиерархијата. Апсолутното име на доменот уште се нарекува целосно квалификувано име на доменот (FQDN, Fully Qualified Domain Name). Наспроти ова, имињата кои не завршуваат со точка понекогаш се интерпретираат како релативни за одредени имиња

на домени не само за коренот – како што се имињата на директориумите без коса линија (/), што се интерпретираат како релативни за тековниот директориум.

Доменот претставува поддрво на доменот на просторот со имиња. Името на доменот е исто како и името на доменот на јазолот на врвот од доменот. На пример, врв на microsoft.com е јазолот со име microsoft.com, што е прикажано на слика 14.



Слика 14. Домени и поддомени [92].

Доменот може да има неколку свои поддрва наречени поддомени. Термините домен и поддомен се користат наизменично. Терминот поддомен е релативен, бидејќи доменот претставува поддомен на друг домен доколку коренот на поддоменот се наоѓа во друг домен.

Наједноставен начин за да се утврди дали доменот е поддомен на друг домен е преку споредување на нивните имиња на домени. На пример доменот sales.microsoft.com мора да е поддомен на microsoft.com, бидејќи sales.microsoft.com завршува со microsoft.com, понатаму и microsoft.com е поддомен на com. Покрај тоа што кон домените може да се обраќа со релативни термини – како поддомени на други домени, кон домените може да се обраќа и според нивото. Нивото ја определува позицијата на доменот во доменот на просторот со имиња. Домените од прво ниво претставуваат деца на коренот. Домени од второ ниво претставуваат деца на домените од прво ниво итн.

Поради забрзаната интернационализација на Интернетот, наместо да постојат само основните домени од прво ниво кои ќе ја опишуваат целокупната организациска структура се одлучило да се внесат и географски одредници. На тој начин за сите земји биле резервирани одредници што ќе одговараат на земјите.

Во 2000 година, организацијата која го одржува доменскиот именски систем (DNS), креирала дополнителни 7 домени во прво ниво за посоодветно да се сместат организациите во доменот со имиња.

DNS-серверите се управувани од големи организации и влади на држави. Тие можат да ја злоупотребат својата моќ преку цензурирање и контролирање на нашата употреба на Интернетот. Ова се случува низ целиот свет низ кратката историја на Интернетот, а се случува и денес. За да се надминат постоечките проблеми, се развива децентализиран DNS систем базиран на технологијата на блоковски вериги.

Децентрализиран доменски именски систем

Предности на децентрализираниот DNS систем во однос на претходно опишаниот централизиран:

• Регистрација на домен е значително поевтино со користење на децентрализиран систем базиран на технологија на блоковски вериги.

• Децентрализираниот систем ја почитува приватноста на корисниците. Системот не бара од корисниците да обезбедат податоци за идентификација кои подоцна можат да се користат против нив. На пример, во централизиран DNS-систем, владата има способност да контролира независни медиуми.

• Владите или организациите можат да го блокираат или одземаат името на редовен домен. Технологијата на блоковски вериги тоа го оневозможува.

• Парите потрошени за регистрирање на кориснички домен помагаат во одржување и понатамошно развивање на други децентрализирани јавни услуги.

• Децентрализиран DNS-систем овозможува постоење на домен со највисоко ниво кој не е во ничија сопственост (.bit домен). Сè додека постојат партнери или јазли во мрежата што ја вршат улогата на DNS-серверите, сите други корисници имаат пристап до кој било домен во хиерархијата пониска од .bit доменот. Таквите партнери се всушност партнери во системот од блоковски вериги од слика 15.

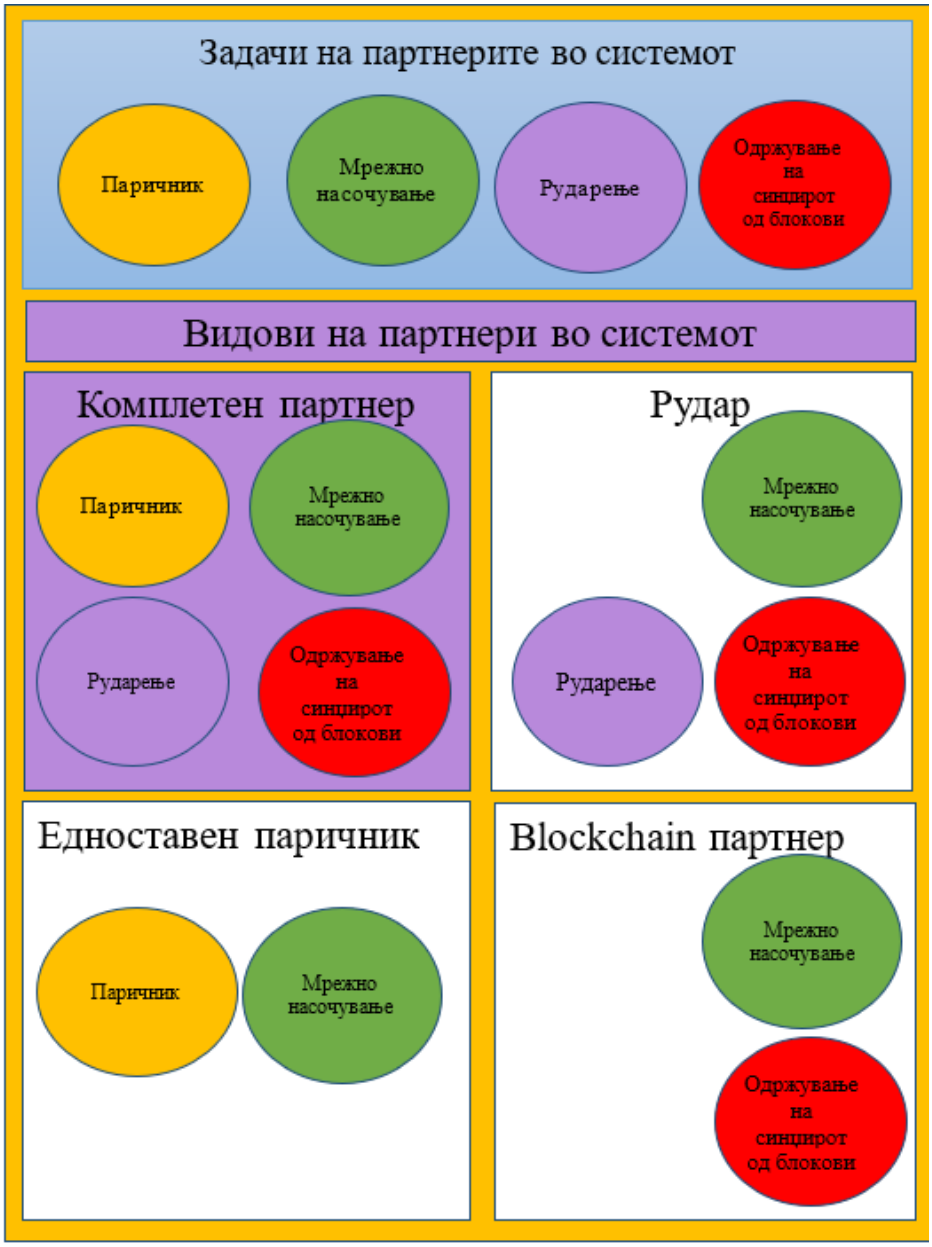
Како дел од технологијата на блоковски вериги, еден партнер во основа работи на компјутерот на еден корисник. Постојат четири задачи што може да ги изврши партнерот. Тоа се: паричник, мрежно насочување, рударство, одржување на целата блоковска верига. Во приватна блоковска верига, по правило, секој партнер ги извршува сите четири задачи, додека кај јавните блоковски вериги ги разликуваме партнерите според задачите што ги извршуваат. Постојат следниве типови на партнери: целосен партнер, рудар, едноставен паричник, партнер со блоковска верига. Сите типови на партнери ја извршуваат задачата за мрежно насочување. Причината за тоа е потребата на секој партнер да воспостави и одржува врски со некои од другите партнери во моделот на рамноправни партнери. Слично на тоа, секој од партнерите кои учествуваат во системот е одговорен за валидација и дифузија (емитување) на нови записи и нови блокови.

Партнерот со блоковска верига ја одржува веригата со сите записи, почнувајќи од првиот блок, кој се нарекува генерички блок, на кој се додаваат сите други блокови до последниот што е создаден.

Во јавните блоковски вериги, поради големата количина на податоци, не секој корисник има можност да ја складира целата верига. Таквиот корисник се нарекува едноставен паричник. Главната задача што ја извршува едноставниот паричник е создавање на нови записи во согласност со протокол пропишан од системот.

Партнерите рудари ги преземаат новите записи создадени од паричниците, ги формираат во блокови и ги додаваат во блоковската верига.

Комплетниот партнер може да ги извршува сите улоги на другите партнери.



Слика 15. Партнери во системот од блоковски вериги [12].

2.4.3 Паметни договори

Етериум (ETH, Ethereum) е платформа која овозможува развој на децентрализирани апликации на технологијата на блоковски вериги. Блоковската верига на етериум се користи како база на податоци за складирање на информации. Целата мрежа на етериум е група на голем број на јазли (компјутери) поврзани меѓусебно кои можат да се сфатат како единствен ентитет наречен етериум виртуелна машина (EMV, Ethereum Virtual Machine). Сите трансакции се ажурираат автоматски и се внесуваат во дистрибуирана база на податоци. Етериум е најстариот познат пример за децентрализиран дистрибуиран систем што овозможува да се извршат паметни договори користејќи ја технологијата на блоковски вериги. Системот етериум ја користи криптовалутата етер. Етер има двојна улога: се користи за плаќање на поставување на договорот на мрежата, но, исто така, се користи и како средство за исплата на бонуси на партнерите рудари во создавањето на нова единица, исто како и во системот биткоин.

Етериумот има своја валута наречена етер (ETH), рударите ја валидираат секоја трансакција и за тоа добиваат одреден дел од него или „*Wei*“ – како најмала единица на оваа криптовалута [118]. Разновидноста на можни апликации кои можат да бидат изградени врз оваа платформа го направија етериумот многу популарен.

Како работи етериум, користејќи ја терминологијата од слика 15: Партнерите паричници или целосните партнери создаваат записи со паметни договори. Договорите се извршуваат кога се исполнети одредени услови. Партнерите рудари одржуваат низа нерешени записи, што вклучува постоечки договори кои треба да се извршат и договори кои допрва треба да бидат напишани во блоковски вериги. Кога се креира нов блок, тие избираат некои од овие записи и работат на нивните компјутери со постоечките паметни договори. Извршување на кодот ја менува состојбата на паричникот на партнерот во начинот на кој тој автоматски ја извршува трансакцијата со која еден корисник му плаќа на друг одреден број на единици од криптовалутата етер, која е дефинирана во паметниот договор. При извршувањето на трансакцијата, исто така, се креира и документ што се поврзува со паметен договор, и овие податоци треба да се додадат во новиот блок. Откако рударите ги запишуваат добиените податоци заедно со податоците за новите договори во нов блок, следи извршувањето на алгоритамот за докажување на работа (PoW) за тој блок. Ако го најде бројот nonce, кој обезбедува задоволителна хеш-порака, со порака се испраќа новиот блок низ целата мрежа.

Секој друг партнер што ја одржува блоковската верига по добивањето на нов блок повторно ќе ги изврши паметните договори по редоследот што е запишан во блокот што го добил. Покрај тоа, ги евидентира и ги потврдува резултатите од извршувањето на паметните договори. Исто така, партнерот проверува дали бројот

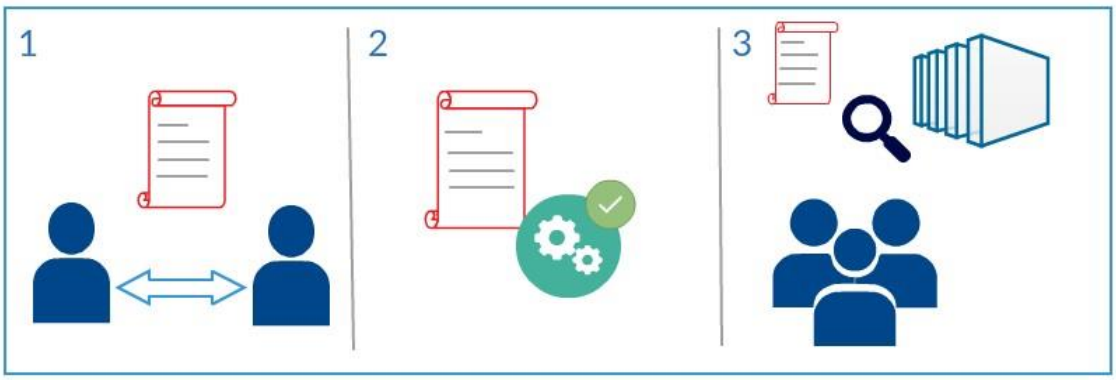
nonce што го пронашол рударот дава задоволувачки хеш и, ако е така, го прифаќа тој блок и го запишува во својата локална копија на блоковската верига.

Сите партнери во системот треба да се согласат со резултатите од извршувањето на паметните договори со тоа што ќе го стават својот код на своите компјутери. Затоа страните коишто ја напишале својата соработка во паметните договори може да му веруваат на системот.

Највисоката вредност што ја нуди етериумот, што ја прави повеќе платформа за апликации отколку валута, е дека нарачките се контролираат од самиот код – паметни договори (smart contracts), а не од корисниците. Паметните договори се збир на инструкции напишани во програмскиот јазик Solidity. Едноставноста на логиката е она што ги прави привлечни и речиси универзално применливи. Секоја трансакција извршена преку паметниот договор се евидентира и снима на мрежата. Може да се каже дека паметните договори се користат за регулирање на деловните односи меѓу страните што немаат взаемна доверба. Овој код може да биде запишан на блоковски вериги и извршен на кој било компјутер во дистрибуирана мрежа. Паметниот договор автоматски се извршува кога се исполнети специфични услови. Бидејќи кодот на паметниот договор е запишан на блоковски вериги, извршувањето се одвива без никаква можност за цензура, прекини, измама или упад од страна на трети страни. Може да се рече дека блоковската верига на која се чуваат паметните зделки е дистрибуиран оперативен систем [13].

Во контекст на блоковските вериги, паметни договори постоеле и пред етериум. Основната идеја на паметните договори е дека многу видови на договорни клаузули можат да бидат инкорпорирани во хардверот и софтверот со кој се комуницира, на таков начин што прекршувањето на договорот е скапо за оној што сака да не го почитува [13]. Средствата или валутите се пренесуваат во програмата. При извршување на програмата, таа автоматски ја препознава состојбата во која се наоѓа системот. Исто така, кога повеќе луѓе бараат трајна или привремена сопственост на овие вредности, програмата одредува на кого треба да ги додели. Постои можност кога никој не ги исполнува условите, тогаш вредностите се враќаат на почетниот сопственик. Во меѓувреме, еден вид на документ во кој е запишана одлуката на програмата се чува на блоковска верига, што му дава одредена сигурност и непроменливост.

Слика 16 го покажува процесот на креирање на паметни договори со следните чекори:



Слика 16. Процес на креирање на паметни договори [12].

1. Можна размена на стоки меѓу двајца (или повеќе) партнери се евидентира како програмски код и се чува на блоковска верига. Партнерите остануваат анонимни, но содржината на паметниот договор е јавно достапна за сите партнери во системот.

2. Променливите како датум или одреден износ на пари го поттикнуваат извршувањето на договорот според правилата дефинирани во кодот.

3. Другите корисници на системот можат да ја пребараат блоковската верига за да ги разберат активностите утврдени со договорот или да го проверат исходот од извршувањето на договорот.

Да се разгледа следната ситуација: лицето А сака да го изнајми својот стан на лицето Б. Лицето Б може да го плати износот потребен за изнајмување на стан со потпишување трансакција меморирана на блоковска верига. Значи тој добива дигитална сметка која е содржана во виртуелен договор помеѓу двете страни. Лицето А потоа му испраќа на лицето В дигитален клуч што ќе му биде достапен на лицето Б од договорениот датум. Ако лицето Б не го добие клучот навреме, кодот што го содржи паметниот договор автоматски ќе ги врати платените средства. Ако А го испрати клучот предвреме, функцијата во рамките на програмата го задржува до датумот кога договорот за изнајмување е договорен. Заедно со клучот, функцијата, исто така го задржува криптовалутниот надоместок што го исплаќа на лицето А кога лицето Б го прима клучот. Кодот е запишан во блоковска верига која одржува илјадници партнери во системот и Б не мора да се грижи за какви било грешки или измами. Исто така, лицето А може да биде сигурно дека неговите изнајмени услуги ќе бидат платени ако го испрати клучот. Договорот автоматски се исклучува по договореното време, кодот не може да се промени од страна на кој било учесник без знаење на другиот, и сите учесници ќе бидат истовремено известувани за промените.

Примерот за изнајмување станови е само една од можните употреби на паметните договори. Други примери што се користат или постои можност да се користат во иднина се:

- автоматизација на гласачкиот систем, каде што технологијата на блоковски вериги може да помогне во кредибилитетот на целиот систем,
- клиничко истражување спроведено од неколку институции со заштита на личните податоци на испитаниците,
- следење на сопственоста и инвестирањето од страна на компанијата која работи на проект во кој инвестираат странски инвеститори,
- автоматизација на плаќањето и отплата на кредити, како и следење на каматните стапки, итн.

Во комбинација со други технологии, се добива уште поширока употреба на паметни договори. На пример, може да се автоматизира осигурувањето на возила и да се дозволи веднаш да се плати осигурувањето на оштетениот. Во случај на несреќа, исплатата ќе се базира на податоците собрани од сензорите што ги следат параметрите на состојбата на возилото во паметни автомобили.

2.4.4 Паричник

Вовед во паричници

Луѓето се навикнати да ги чуваат своите пари во паричник или на сметка и може да се направи слична аналогија со криптовалутите. Криптовалутите се наоѓаат на адреси. Секоја криптовалута има свој систем на адреси и тие се разликуваат по број на знаци, почетни знаци и слично. Кога се испраќаат средства, тој што испраќа средства мора да докаже дека е и сопственик на адресата од која се обидува да испрати одредено количество на криптовалути. Ова се случува на тој начин што со својот приватен клуч ја потпишува трансакцијата и на тој начин им докажува на сите дека е сопственик на таа адреса и она што е на неа. Процесот на дигитално потпишување на трансакција е низа криптографски и математички функции што ја докажуваат сопственоста на некои средства.

За најголем број на корисници на криптовалути, информацијата дека нивните криптовалути се наоѓаат на некоја адреса и дека е потребен приватен клуч за да потрошат средства не значи премногу. За да се олесни користењето на криптовалутите и да се приближи до сите, потребен е поедноставен начин за складирање и испраќање на криптовалутите. Така се појавуваат дигитални паричници за чување и праќање на криптовалути. Дигиталниот паричник (wallet) претставува софтверско решение за складирање на шифрирани адреси и приватни клучеви за да се отклучат овие адреси. Еден паричник може да содржи неограничено многу адреси.

Десктоп паричници

При користење на овие паричници, зависно од типот на десктоп паричникот, корисникот на својот компјутер може да инсталира и „целосен“ клиент, кој доаѓа со

целокупната блоковска верига. Зависи од тоа која криптовалута се користи, оваа акција може да бара многу мемориски простор на хард-дискот. Приватните клучеви заедно со соодветните адреси, се чуваат на хард дискот на компјутерот на кој се инсталирани. Од аспект на безбедноста ова е доста добар начин за зачувување на податоците. Криптовалутата е на компјутерот од корисникот и ако се успее да се одржи безбедноста на компјутерите на доволно високо ниво, тогаш и средствата од корисникот се безбедни.

Често, новите валути немаат доволно квалитетни и доволно добро тестирани десктоп паричници, па затоа се применуваат и некои други опции. Десктоп паричниците се, главно, едноставни за користење, но недостаток е тоа што не можат лесно да се пренесуваат. По инсталацијата, ако е во прашање и целосен клиент, потребно е да помине одредено време десктоп паричникот да се синхронизира со остатокот од мрежата. Кога се работи за биткоин и етериум, чии блоковски веригии се најголеми, синхронизацијата може да трае и неколку дена, зависно од брзината на Интернет врската.

Мобилни паричници

Мобилните паричници претставуваат апликации што се наоѓаат на мобилниот телефон од корисникот. Инсталацијата на овие паричници трае многу кратко и не бара многу меморија. Во овој случај, приватните клучеви се зачувани на соодветниот мобилен телефон. Ова внесува потенцијални ризици во случај на губење или крадење на мобилниот телефон. Постојат неколку механизми за заштита. Првиот е поставување на пин код што го спречува оној кој доаѓа во сопственост на телефонот неовластено да ги потроши средствата. Сепак, останува проблемот со загубените средства.

Без оглед на тоа што некој кој неовластено пристапил до соодветниот телефон не може да ги потроши средствата, важно е да има и начин со којшто може да се вратат изгубените средства. Детерминистичкиот паричник ни овозможува со помош на „*seed words*“, едноставно да се направи резервна копија на паричникот. Сè што треба да се направи е да се запишат најчесто 12 зборови и безбедно да се зачуваат. Ова е доволно за да се вратат податоците назад во секој момент.

Мобилните паричници се погодни за мали и чести плаќања. При користење на мобилни паричници не е потребно рачно да се внесе адресата на примачот, доволно е да се скенира QR (quick response) кодот што ја претставува таа адреса. Ова ја олеснува примената на криптовалутите во секојдневниот живот. Со помош на телефонот можно е да се плати за некоја услуга, да се купат стоки или да се подигнат пари од некој од специјализираните автомати.

Онлајн паричници

Онлајн паричниците го претставуваат најнебезбедниот тип на дигитални паричници, бидејќи приватните клучеви не се наоѓаат кај корисникот. Нема потреба од преземање или инсталирање на кој било софтвер. Паричникот се добива со тоа што ќе се направи кориснички налог на некој онлајн сервис што ја нуди оваа услуга. Сè се сведува на тоа да се направи кориснички налог на некоја од веб-страниците што овозможуваат чување на криптовалутата. До средствата подоцна може да се пристапи со помош на кој било уред што има пристап до Интернет. Ова дава дополнителна слобода и мобилност, бидејќи не постои опасност да се изгубат криптовалутите ако се изгуби мобилен телефон или да мора да се носи компјутер со себе за да се изврши некоја трансакција со криптовалути.

Недостатокот на овие паричници се гледа во тоа што во овој случај, приватните клучеви се наоѓаат во сопственост на давателот на онлајн паричници. Затоа не е препорачливо да се чуваат поголеми износи на пари во ваков тип на паричници.

Парични банкноти

Паричните банкноти најчесто се користат во ситуации во кои некој сака да ги чува своите банкноти долгорочно или да ги подари некому. Со користење на специјализирани веб-страници, се прават парични банкноти кои поседуваат криптоадреса и приватен клуч за пристап до средства од оваа адреса. Потоа ваквите парични банкноти се печатат и се чуваат на некое безбедно место. За да се извлечат средства од дадените адреси, едноставно со скенирање или со внесување на приватен клуч се префрлаат потребните средства во некој паричник. За дополнителна сигурност се практикува компјутерот и печатачот да бидат исклучени од мрежата за да се оневозможи некој малициозен корисник да пристапи до чувствителните податоци. Оној кој одлучува за овој вид на чување на своите биткоини мора да има предвид дека секој што доаѓа во контакт со паричните банкноти лесно може да пристапи и до средствата што се запишани на хартијата, освен ако при креирањето на банкнотите не е вметната дополнителна шифра. Покрај тоа, ако хартиените банкноти се изгубат, скинат или се оштетат на некој друг начин, средствата се изгубени долгорочно. Затоа, многу е важно на правилен начин да се заштитат хартиените банкноти. Се препорачува да се чуваат во сеф или слично место заштитено од оштетување и неовластен пристап.

Хардверски паричник

Најсигурен начин за чување на криптовалутите претставуваат хардверските паричници, но тие не се бесплатни. Овие паричници се хардверски уреди како USB флеш уреди. Тие имаат сложен систем за складирање и чување на парови од приватни клучеви и соодветни адреси. При првото користење на овие паричници, корисникот од уредот препишува 12 или 24 зборови што подоцна може да се користат за враќање на средствата во случај на губење или оштетување на уредот [12]. Хардверските

паричници доаѓаат со систем за внесување шифра. Во случај шифрата да се внесе погрешно одреден број пати, паричникот ќе се заклучи и потоа ќе треба да внесат горенаведените 12 (24) зборови за да се отклучи паричникот [102].

При испраќањето на биткоините, најчесто се користат онлајн апликации или десктоп паричници, кои се поддржани од страна на производителот на хардверските паричници. Хардверскиот паричник најпрво се поврзува на уесбе-порта со компјутерот, потоа на хардверскиот паричник му се испраќаат непотпишаните трансакции, следно корисникот го гледа на екранот на својот уред износот кој го испраќа и адресата на примачот и по проверка на податоците и внесувањето на шифрата на хардверскиот паричник, тој ја враќа назад потпишаната трансакција и со тоа на сите им докажува дека е сопственик на овие средства и навистина сака да ја изврши таа трансакција. Овој вид чување на криптовалута, иако поскап од сите други, претставува најсигурен начин за чување на криптовалута и обезбедува одличен степен на сигурност за сите што планираат долгорочно да чуваат криптовалути. Хартиените и хардверските паричници се посигурни од останатите, затоа што кај тие паричници приватните клучеви се наоѓаат офлајн.

Едноставна имплементација на паричник

Со воведувањето на технологијата на блоковски вериги, развиени се повеќе проекти со отворен код, кои им овозможуваат на програмерите да развиваат апликации како што се паметни договори или примена на партнерски паричници. Овие проекти со отворен код (open-source projects), исто така, овозможуваат да се комуницира со некоја јавна блоковска верига или да се создаде приватна блоковска верига. Паричник е компјутерска програма која се користи за примање и испраќање биткоини. Улогата на паричникот е да ги чува приватните клучеви на корисниците, да ги прикажува трансакциите што се користат за испраќање или примање биткоини со користење на адресата на корисникот, и да ја прикажува количината на биткоините што ги поседува корисникот. Биткоините се префрлаат од еден паричник во друг со биткоински трансакции. Паричник може да се инсталира на компјутер, паметен телефон или таблет. Исто така, тие се достапни како веб-апликации, до кои може да се пристапи од кој било уред поврзан на Интернет [12].

Најважните поими за работа со паричникот се: приватен клуч, јавен клуч, адреса и трансакција. Приватен клуч е бројот што може да се генерира на различни начини. За да се дефинира врската меѓу приватниот и јавниот клуч, прво треба да се дефинира елиптичната крива, типот на крива што често се користи во криптографијата.

2.4.5 Дефиниција на елиптична крива

Нека K е алгебарска структура - поле. Карактеристика на полето K е најмалиот природен број n, така што

$$1 + 1 + \cdots + 1 = n \cdot 1 = 0,$$

каде што 0 и 1 се неутрални елементи за собирањето, односно множењето во *K*, соодветно [17]. Ако $n \cdot 1 \neq 0$ за секој природен број *n*, тогаш се вели дека полето *K* има карактеристика 0. Концептот на елиптична крива може да се дефинира преку произволно поле *K*, но најважните случаи се кога *K* е полето на рационални броеви *Q*, полето на реални броеви *R*, полето на комплексни броеви *C*, и, конечното поле Fq од q елементи. Полињата *Q*, *R* и *C* се со карактеристика 0, додека карактеристиката на Fq е еднаква на *p*, каде што *p* е прост број и $q = p^m$ за некој природен број *m* [17].

Дефиниција: Нека *K* е поле со карактеристики различни од 2 и 3, и нека

$$f(x) = x^3 + ax + b$$

каде што $a, b \in K$ е кубен полином без повеќекратни корени. Елиптичната крива *E* над *K* е множество на сите точки (*x*, *y*) во $K \times K$, што го задоволуваат равенството:

$$y^2 = x^3 + ax + b$$

заедно со друг елемент што го означуваме со *O* и се нарекува „точка во бесконечност“ [17].

Елиптичните криви се користат за генерирање на јавен клуч од приватен клуч. Најчесто, генерираниот приватен клуч се сведува на избор на случаен број помеѓу 1 и 2^{256} за кој се применува хеш-функцијата SHA-256. Откако ќе се генерира приватниот клуч, со помош на елиптична крива се генерира јавен клуч од приватниот клуч.

Биткоин-системот ја користи елиптичната крива дадена со равенката:

$$y^2 = (x^3 + 7) \; mod \; p$$

кадешто, $p = 2^{256} - 2^{32} - 2^9 - 2^8 - 2^7 - 2^6 - 2^4 - 1$.

Со стандардната крива се поврзува точката *G* наречена генератор на елиптична крива. Нека со *pk* е означен приватниот клуч, тогаш соодветниот јавен клуч *jk* се пресметува со формулата:

$$jk = pk * G$$

Значи, јавен клуч *jk* е точка на кривата добиена со математичка операција множење на точката *G* со скалар *pk*, што, всушност, претставува собирање на точката *G* самата со себе *pk* пати. Потоа, со користење на хеш-функции, вклучувајќи ја и хеш-функцијата SHA-256 и на јавниот клуч може да се пресмета биткоин-адресата A. Ако составот на хеш-функциите го означиме со H_{adr}, адресата *A* може да се пресмета според следнава формула:

$$A = H_{adr}(jk).$$

При креирање трансакција се користи приватниот клуч. Приватните клучеви може да се чуваат во паричник во различни формати. WIF (Wallet Import Format) е еден од најчестите формати на датотеки со приватни клучеви [12]. Користејќи приватен клуч, корисникот потврдува дека тој е сопственик на одредена адреса. Значи, може да се каже дека биткоините на одредена адреса му припаѓаат на лице коешто има приватен клуч за таа адреса. Важно е да се напомене дека елиптичната крива и хеш-функциите се еднонасочни или неинверзни. Тоа значи дека во секој момент лесно може да се генерира јавен клуч од приватен и од јавниот клуч да се генерира адресата. Но, не постои начин да се генерира приватен или јавен клуч користејќи ја адресата.

Трансакцијата е запис во мрежата на биткоин, и со неа се пренесува одреден износ на биткоини од една адреса (или повеќе) во друга адреса (или повеќе). Трансакциите се јавни и можат слободно да се прелистуваат, на пример, користејќи прегледувачи на блоковски вериги достапни на Интернет. Тие недвосмислено потврдуваат дека одредено количество биткоини биле пренесени од една адреса во друга. Паричникот му овозможува на корисникот да креира нова трансакција од адресата што му припаѓа на тој паричник. За да се изврши трансакција, потребно е корисникот да ја наведе адресата на примачот и износот што сака да го плати. Со трансакцијата го поврзуваме записот, што се состои од претходни трансакции што се плаќаат на адресата на корисникот. Со приватниот клуч, корисникот потврдува дека одреден износ на пари му припаѓа и дека овие трансакции може да се користат како запис за нова трансакција. Секоја трансакција има свој излез, со кој се дефинира кој дел од вкупниот износ на парите се префрлени на адресата на примачот, и кој дел останува на адреса на испраќачот.

2.4.6 Вовед во криптоберза

Криптоберзите се специјализирани платформи за тргување со криптовалути. Целта на криптоберзата е да им овозможи на корисниците едноставно и лесно да разменуваат криптовалути за пари или да разменат еден вид криптовалута за друг. Првата криптоберза што овозможила тргување со криптовалути била јапонската берза MtGox, која е создадена во 2006 година, а првично била онлајн платформа за играчи на социјални игри. Оваа платформа им овозможувала да тргуваат и разменуваат карти за играње на социјални игри. Во јули 2010 година, само два месеци откако биткоин првпат се користел за плаќање, MtGox лансирала платформа за трговија со биткоини [94]. Во моментот кога се појавила, берзата понудила нешто сосема ново за луѓето, способност да се купат или продадат биткоини секому, од дома, со неколку кликнувања на нивниот компјутер.

Оваа берза имала огромно влијание и врз цената на биткоините и врз трговијата со криптовалути. Се проценува дека во текот на 2013 и почетокот на 2014 година, дури 70 % од глобалната светска трговија со биткоини се одвивала на оваа берза [94]. Во 2014 година се случува голем хакерски напад, по што многу корисници на MtGox

остануваат без голем дел од своите биткоини. Ова значително влијаело на цената на биткоините, кои брзо паднале од над 1.200 американски долари на нешто повеќе од 200 долари.

Безбедносни проблеми

Во меѓувреме, се развиле и други берзи како што се Bitstamp.net, Poloniex.com итн. Нападот на берзата MtGox најдобро покажува колку е важно средствата да се чуваат безбедно и сигурно. Кога биткоините или други криптовалути се чуваат на берза, приватните клучеви за адресите се во сопственост на некој друг. Ова значи дека ако некој може да успее да ја загрози безбедноста на берзата и продира во нивниот систем за складирање на приватните клучеви, многу веројатно трајно ќе отуѓи голем дел од средствата, што директно значи дека се загрозени сите што ги чуваат своите криптовалути на берза. Во такви ситуации берзите често не ја преземаат одговорноста, па затоа корисниците на берзата се тие кои губат.

Многу луѓе практикуваат своите криптовалути да ги чуваат на берза. Причините за ова се различни. Некои сакаат нивните криптовалути да бидат лесно достапни во случај да одлучат да ги продадат. Некои од нив се, едноставно, мрзливи да си отворат свој паричник и сами да се грижат за својот имот, а постојат и такви кои не се свесни за ризиците за чување на криптовалутите на берза. Не постои помалку сигурно место за складирање на криптовалутите од берзата. Поради големата сума на пари, берзите се под постојани напади од хакери. Не смее да се размислува за користење на берзата без двоен фактор за проверка, но со тоа може да се спречи само да не биде пробиен тековниот налог, а ако се пробие самата берзата, нема помош.

По падот на MtGox, берзите ја сфатиле опасноста од потенцијалните закани, па почнале значително повеќе да инвестираат во безбедноста и креирањето на механизми за безбедно чување на доверливи податоци. Сепак, и покрај сите мерки на заштита, успешните напади на големите берзи продолжија да се случуваат дури и по инцидентот на MtGox. Во 2015 година е пробиен Bitstamp, а во 2016 година Bitfinex [95].

Тргување на берза

Најчеста причина за користењето на берзите е трговијата со криптовалути. Принципот на тргување е ист како и на секоја друга берза. Лицето кое сака да тргува со криптовалути најпрво мора да обезбеди одредени средства на берзата. Најчесто, криптовалутите се депонираат на берзата. На некои берзи може да се депонираат средства во национални валути, како што се еврото или доларот. По тоа, се купува одредена криптовалута, која подоцна се продава по поповолна цена.

Разликата меѓу берзата на криптовалути и класичната берза е првенствено попустливата регулаторна рамка, ако постои, што предизвикува поголем ризик, но, исто така, значително поголемо осцилирање на цената и можности за остварување на огромни профити. Кога се тргува, многу е важно да се земат предвид условите на

користење, како и законските прописи во земјата на лицето кое инвестира пари на берзата, како и во земјата која е седиште на одредена берза. Ова ги спречува правните проблеми што можат да се појават при подигнување на пари од берзата. Ситуацијата тука малку е компликувана, бидејќи трговијата со криптовалути во многу земји не е правно регулирана, но тоа полека се менува.

Трговијата со криптовалути е со растечки тренд. Дневниот обем често надминува 10 милијарди долари, а бројот на нови корисници кои се регистрираат секој ден во берзите се мери во десетици, па дури и стотици илјади луѓе [12]. Значителен број на трговци од други платформи или од други системи, како што се трговијата со благородни метали, индекси и обврзници, на крајот се префрлуваат на тргување со криптовалути бидејќи постои потенцијал за повисок профит. Изработени се специјализирани алатки за да им се олесни на корисниците да ги следат разните берзи и трендовите во движењето на цените. Се очекува дека со напредокот на дефинирање на правна регулатива во целата област, ќе се зголеми бројот на големите инвеститори што ќе влезат на пазарот на криптовалути, а сега го избегнуваат поради правната несигурност.

Видови на корисници на берза

Берзата претставува институција за организирано тргување со производи. Главние видови активности на берзата се:

- инвестиции,
- еднократно купување,
- шпекулации,
- арбитража.

Инвестицијата и еднократното купување се сосема различни работи, но во двата случаи берзата е многу ограничена. Инвеститорите ја купуваат посакуваната криптовалута што ја гледаат како долгорочна инвестиција. По купувањето, криптовалутите треба да се повлечат од берзата, бидејќи немаат повеќе потреба од користење на берзите. Повторно ќе се појават на берзата кога ќе одлучат да го продадат тоа што го имаат или, евентуално, да купат уште криптовалути.

Еднократно купување и продавање, всушност, е ситуација кога некој треба да купи одредена криптовалута за нешто, па ја користи берзата за таа цел или има некој износ на криптовалути што сака да го продаде и да дојде до пари (често рударите спаѓаат во оваа група). Дури и тука, берзата се користи само за одредена цел и откако оваа цел ќе се исполни, најчесто, употребата на берзата завршува многу брзо.

Оние што се занимаваат со шпекулации и арбитража, за разлика од претходните две групи, трошат многу време на берзата и, ако се квалификувани, добро и заработуваат. Шпекулантите се оние што се обидуваат да профитираат од својата вештина за добро да го проценат движењето на цените. Се разбира, има повеќе и

помалку успешни шпекуланти. Јасно е дека не постои шпекулант кој секогаш добива, но тоа не е ни потребно, доволно е да се биде во право во повеќе од 50 % од случаите.

Арбитражата е активност која е генерално многу помалку ризична од шпекулациите. Овде, трговецот не се обидува да го предвиди движењето на цените, но профитира на ценовните разлики помеѓу берзите. Ако, на пример, на една берза цената на биткоинот е 7500 долари, а на друга 7600 долари, тој може да купи биткоини на првата берза по 7500 долари, а на другата берза истото количество да го продаде по 7600 долари и така заработува 100 долари на биткоин. Трговецот е целосно отпорен на трендот на движење на цените на криптовалутите. Добрите можности за арбитража не се појавуваат толку често, особено кога пазарот е стабилен. Трговецот мора да има вложено значителен износ на средства во неколку берзи за да може арбитражата ефикасно да работи. А секогаш е ризично да се чуваат многу пари на берза. Арбитража е токму причината зошто цените на главните берзи се генерално изедначени, бидејќи секогаш кога ќе се појави малку поголема разлика, трговците го користат моментот за арбитража.

Тргувањето со берзите е многу примамливо и навистина може да биде многу профитабилно. Сепак, треба да се биде многу внимателен. Освен фактот дека берзите се инхерентно ризични, проблемот е што тие се нарекуваат „игра со нулта сума“. Ова значи дека збирот на добивките на берзата на оние што добиваат е еднаков на збирот на загуби на губитниците. Не е тешко да се погоди во која група ќе заврши неискусниот трговец. Потребно е многу знаење, дисциплина, време и малку среќа за да се добие добар приход од берзите на долг рок и да се живее од берзата.

Алтернативи на берзите

Иако берзите се главниот начин за купување и продавање на криптовалути, постојат и други начини. Овие други начини честопати се многу поскапи (просечната провизија на берзите е околу 0,2 %), но најчесто побрзи и честопати посигурни.

Сервисите за купопродажба на криптовалути работат слично како менувачница. Кога станува збор за традиционални валути, треба да се депонираат пари на берзите, а потоа да се тргува таму со други учесници. Од друга страна, во менувачница се носат пари и се разменуваат во друга валута во самата менувачница. На ист начин функционира и сервисот менувачница, криптовалутите се продаваат на компанијата која го нуди тој сервис или од неа се купуваат. Таквите сервиси често се нарекуваат менувачница, но ова не е сосема прецизно, бидејќи во повеќето земји криптовалутите не се сметаат за пари.

Криптовалутите може да се купат и продадат на специјализирани машини. Криптовалутите може да се разменуваат за пари и обратно. Често, ова е најскапиот начин да се купи или да се продаде криптовалута (провизијата за продажба во просек е блиску до 7 %, а за купување блиску до 10 % [5], но некои од нив се практични,

затоа што за тоа не е неопходно да се има банкарска сметка, некои се достапни 24/7, и може да се користат од странци чиишто локални услуги обично не се достапни.

Постојат служби кои за одреден надомест им помагаат на купувачите и продавачите да се поврзат, водејќи сметка за безбедноста на двете страни, и најчеста е употребата на системот „escrow“. Овде, биткоинот од продавачот е заклучен во внатрешноста на услугата и останува заклучен додека продавачот не потврди дека ги добил парите од купувачот. Кога ќе го потврди тоа, биткоинот автоматски се испраќа до купувачот. Најпознат таков сервис е LocalBitcoins. Исто така, постои можност за купување и продавање преку оглас. Овој метод е многу ризичен, бидејќи никогаш не се знае со кого имаш работа, иако понекогаш тој може да биде најдобра опција, бидејќи не постои посредник.

2.4.7 Децентрализирана заштита на лични податоци

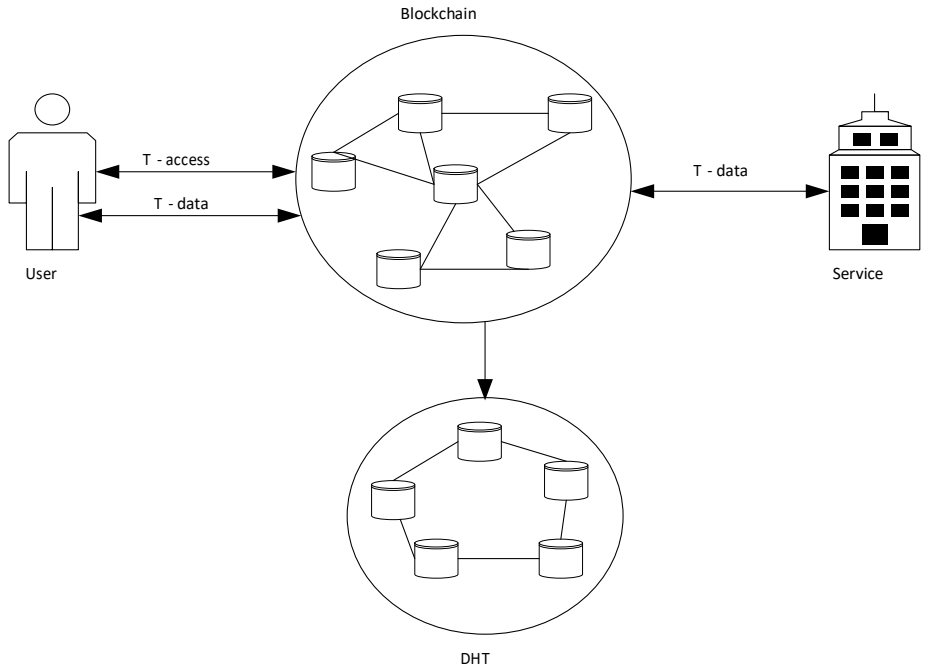
Од социјалните мрежи постојано се собираат лични податоци, активности и навики на корисниците, а со тоа се губи приватноста на корисниците. Корисниците сè уште немаат јасен преглед за тоа кои податоци точно се собираат и за која цел. Тие ја губат целосната контрола на она што се случува со податоците потоа и не можат да ги повлечат дозволите. Обично, има страница за поставување приватност на повеќето од страниците на социјалните медиуми, каде што корисниците можат да го ограничат она што другите луѓе го гледаат за нив. Она што не можат да го контролираат и конфигурираат е тоа што го гледаат другите во социјалните медиуми. Луѓето се навикнати на договори за приватност, дадени на начин што не е лесен за корисниците, објаснувајќи ги површните аспекти за прибирање лични податоци. Но, собирањето лични податоци не запира на страницата. Социјалните медиуми ги следат интересите за прелистување на веб-страници на корисниците и интеракциите што ги прават со други мобилни или веб-апликации. Чувствителните податоци, како што се списоци со контакти и локацијата на корисникот, се собираат при инсталирање на апликациите за мобилни телефони. Од корисниците се бара само да го прифатат пристапот на трети страни при инсталирање мобилна апликација без детални информации или опција за делумно прифаќање.

Загриженоста за приватноста на податоците расте кога се соочуваме со последиците од она што другите го виделе или научиле за нас. Во неодамнешната студија за состојбата со приватноста направена во поглед на бројни нарушувања на податоци од висок профил, 74 % од испитаниците велат дека за нив е „многу важно“ тие да можат да контролираат кој може да добие информации за нив, а 65 % велат дека за нив е „многу важно“ да контролираат какви информации се собираат за нив. А 91 % од испитаниците се согласуваат дека потрошувачите изгубиле контрола врз тоа како се собираат личните информации и се користат од страна на компаниите [18]. Предложеното решение од Зискинд [19] се однесува на вообичаени прашања поврзани со приватноста, како што се сопственоста на податоците, транспарентноста на

податоците и контролата на пристапот. Ова решение е систем за управување со контрола на пристап кој главно се фокусира на мобилните платформи и неможноста на корисникот да го отповика одобрениот пристап до приватни податоци. Со инсталирање на мобилна апликација, дозволите се даваат на неодредено време и корисникот треба да ја деинсталира апликацијата и да престане да ги користи услугите ако сака да го отповика пристапот. Целта на новото решение е корисникот да може да контролира и ревидира кои податоци се зачувани и како се користат. Идејата е да се чуваат полисите за пристап до личните податоци на блоковска верига и потоа, јазлите во блоковската верига да го контролираат пристапот до дистрибуирана хеш-табела **(**DHT, Distributed Hash Table).

Решението е составено од три субјекти: корисникот, компанијата што ја обезбедува услугата и блоковската верига. Кога корисникот сака да дозволи или да го отповика пристапот до неговите лични податоци, блоковската верига доаѓа до активност како посредник. Тука, блоковската верига поддржува два вида на трансакции: трансакција за пристап и трансакција за податоци. Овие типови на трансакции овозможуваат управување со контрола на пристап, складирање на податоци и земање податоци. Кога корисникот инсталира нова апликација, се создава заеднички идентитет и се испраќа до блоковската верига заедно со конфигурираните дозволи по желба на корисникот. Сите доделени дозволи се наведени во т.н. полиса. Споделените клучеви (јавниот клуч на корисникот и јавниот клуч на услугата) и полисата се испраќаат преку трансакција за пристап во блоковската верига. Во предложениот систем се воведува нов комплексен идентитет. Сложениот идентитет е заеднички идентитет меѓу корисникот и услугата. Корисникот е сопственик на клучот, а услугата е гостин. Комплексниот клуч е составен од парови за потпишување на двете страни, така што податоците ќе бидат заштитени од сите други страни во системот, освен од сопственикот и сите негови гости.

Чувствителните податоци за корисникот се шифрираат со споделениот клуч за шифрирање и се испраќаат со податоците за трансакција за складирање. Блоковската верига ги испраќа податоците во дистрибуирана хеш-табела (DHT) и ја задржува само вредноста на хешот како покажувач до податоците. Вредноста поставена во DHT е шифрирана со комплексниот клуч. Покажувачот на вредноста им е познат и на корисникот и на услугата. DHT ги исполнува само веќе одобрените функции за читање и пишување. И корисникот и услугата можат да ги пребаруваат податоците користејќи го покажувачот на податоците. Секој пат кога услугата ќе пристапи до податоците, нејзините дозволи се проверуваат споредени со последната трансакција за пристап. Корисникот може да ги отповика дозволите во секое време или да ги модификува, со иницирање нова трансакција за пристап. За да се следи ова, лесно може да се развие веб-табела што ги покажува тековните дозволи на корисникот, прикажано на слика 17.



Слика 17. Преглед на системот за децентрализирани дозволи [20].

Друго истражување, засновано врз социјалните медиуми контролирани од корисници на блоковската верига, е прикажано во [21]. Според ова истражување, корисниците треба да го контролираат своето присуство преку Интернет со набљудување на објавите што ги споделуваат и да ја контролираат можноста за повторно споделување. Користејќи ги можностите за P2P, создадена е децентрализирана мрежа за дистрибуција на содржини [21]. Алатката со која управува блоковската верига во овој случај се податоците што ги објавуваат корисниците. Предложеното решение се состои од: табела со шифриран хеш, споделена од корисник, систем на контрола на максималниот број на дејства што ги извршуваат корисничките кругови, локален орган за сертификати (PCA, Personal Certificate Authority), кој управува со круговите на корисниците, и блоковска верига.

Кога корисникот споделува објава со својот круг, неговиот PCA ги шифрира податоците со јавниот клуч на кругот. Шифрираните податоци се чуваат во дистрибуирана хеш-табела DHT. Оваа дистрибуирана хеш-табела DHT има три колони што им овозможува на корисниците да ги споделат веќе објавените податоци што ги гледаат. Секој пат кога корисникот споделува објава, во првата колона се зачувуваат хеш шифрирани податоци за објавата што се гледа и се споделува. Втората колона го евидентира хешот на шифрираните податоци со јавниот клуч на неговиот круг. Во третата колона се чува податочната ставка што се шифрира. Причината за користење DHT во ова второ решение е иста – податоците со големи димензии, како документи, слики и видеа, треба да се чуваат на децентрализиран начин. Во блоковската верига се чуваат само трансакции за споделување на објави на корисници. Не може да се чуваат вистинските податоци затоа што преземањето на целиот синџир на сите јазли ќе

создаде ограничувања во компјутерските пресметки и временски ограничувања. Кога корисникот создава објава, испраќа нова трансакција до блоковската верига со својот идентитет, хеш-клуч на шифрираните податоци и токен во кој го наведува дозволениот број на акции. Следно, корисникот испраќа посебна трансакција до секој член на својот круг со клучот за шифрирање на податоци. Ако друг корисник што ја добил објавата, сака да ја сподели, тој испраќа нова трансакција со својот идентитет и клучот на податоците шифрирани со клучот на кругот на овој нов корисник. Повторно, повеќе нови трансакции се испраќаат до следни корисници што можат да ја разгледаат повторно споделената трансакција. Бројот на токени се намалува со секоја акција.

Сите напори да се создадат овие две решенија на блоковска верига се затоа што личните и чувствителни податоци не треба да се доверуваат на трети страни. Бидејќи корисниците создаваат и објавуваат податоци, исто така, тие треба да останат главни сопственици на податоците. Во однос на надзорот направен со следење на процедурите и интересите на корисниците, корисниците треба, барем, да го знаат тоа. Блоковската верига може да биде филтер за дозволи за пристап до приватни податоци или може да спроведе целосна децентрализирана социјална мрежа, како што е прикажано во второто решение.

2.4.8 Дигитална сопственост

Многу специфична имплементација на блоковската верига во областа на големи податоци е поврзана со интелектуалната сопственост на дигиталната уметност. Уметниците, дизајнерите и креативните работници можат лесно да споделуваат свои дела на Интернет, но чувањето на правото на сопственост или добивање на соодветна компензација за авторство е многу тешко во дигиталниот свет. Не постои транспарентен начин да поседувате нешто што може лесно да се копира и целосно да се преслика без знаци на оригиналот. Според проектот наречен Ascribe, Интернетот е изграден со клучен недостаток во однос на прашањето за сопственоста [22]. Штом делото се стави на Интернет, дури и ако се продава преку Интернет, авторот ја губи контролата. Тоа е причината зошто визијата на Ascribe е да се изгради сопственички слој за дигитални содржини на Интернет. Тие создаваат алатка за авторство со што ќе може да се следи каде се шири уметничкото дело.

Авторите, каналите и потрошувачите на Интернет имаат голем проблем. Заедничко решение за споделување видеа, филмови, музика, слики и фотографии, дигитални графики во 2Д и 3Д сè уште не е пронајдено.

World Wide Web ја започна својата работа со едноставни линкови и ја постави основата на тој начин што може да биде видливо што е оригинал, а што се копира. По тоа, луѓето размислуваа за начин како да се зачува авторството со тоа што авторот ќе го споменат во референци. Но, овој систем е далеку од идеален. Луѓето секогаш можат да најдат начин да копираат работи, а авторот нема да знае и нема да биде известен за тоа. Или, луѓето можат да доделуваат право на авторот, но тоа е во една насока, така

што авторот сè уште нема да знае дека некој се осврнал на неговата работа. Или уште полошо, луѓето можат да наведат некој друг, кој не е оригинален автор. Се заклучува дека главната алатка за дигитална содржина и дигитално споделување (www) ја занемарува потребата за дигитална сопственост. Но, на почетокот од развојот на сервисот www било поинаку. Проектот Xanadu бил првиот хипертекстуален проект, основан во 1960 година од Тед Нелсон [23]. Уште тогаш на проблемот за дигиталната сопственост му се пристапило со воведување на шема за објавување врз основа на авторските права, во систем што ќе обезбеди услуги за складирање и објавување. Доделувањето на правото на авторство е изградено во овој систем. Двонасочните врски требало автоматски да се поставуваат секој пат кога некој ги користел податоците на друг корисник. Се дошло до заклучок дека станува збор за комплицирана, неизводлива технологија и проектот бил затворен.

Сега, со технологијата на блоковски вериги, Ascribe се обидува да ги постигне целите на Xanadu со изнаоѓање решение за регистарот на дигитална сопственост и видливоста на копиите. Во однос на видливоста, се обидуваат да ги пронајдат сите копии на заштитената содржина што постојат на Интернет. Ова може да се направи со пребарување на целиот Интернет и да се изврши споредба за сличност со содржината на креаторот. Овој проблем би се решил со машинско учење со пребарување по сличност. Кога ќе се најдат копиите, системот извршува автоматски двонасочни врски. Тогаш, авторот треба да одлучи дали ќе побара плаќање за лиценца или, можеби, барање за одземање.

Кога станува збор за продажба на интелектуална сопственост на дигиталната уметност, тоа не е само продажба на копија, туку се продава сопственост и право на користење, изменување или препродажба на содржината. За да се направи овој вид продажба на сопственост е потребно склучување правен акт, договор, ангажирање адвокат и сл. Идејата за користење на блоковска верига за чување и продажба на сопственост на дигитални податоци ќе биде едноставна како испраќање е-пошта со потпис дека корисникот ја пренесува сопственоста на неговата содржина. Условите за услуги што ги обезбедува Ascribe се прават во консултација со специјализирани адвокати. Сложеноста на процесите на легално лиценцирање и сопственост се постигнува со прифаќање на условите за услуга. Блоковската верига е јавно достапна книга со доверба и ќе ги обезбеди авторските права на сите корисници. Временскиот печат (timestamp) на трансакциите може да се користи како доказ на суд, во случај на спор за сопственост. Во врска со имплементацијата на блоковска верига, Ascribe направи свој протокол наречен SPOOL (Secure Public Online Ownership Ledger) [24]. Овој протокол е направен специјално за документирање на трансакции поврзани со дигиталната сопственост. Ascribe му дозволува на уметникот да постави фиксен број изданија на дело што потоа може да се пренесе, гарантирајќи дека секое издание е автентично и од уметникот. Значи, кога вршите трансакции со трансфери, корисникот може да пренесе сопственост за едно или повеќе изданија. Изданијата се под едно дело што е зачувано во BigchainDB и се користи за употреба во блоковска верига [25].

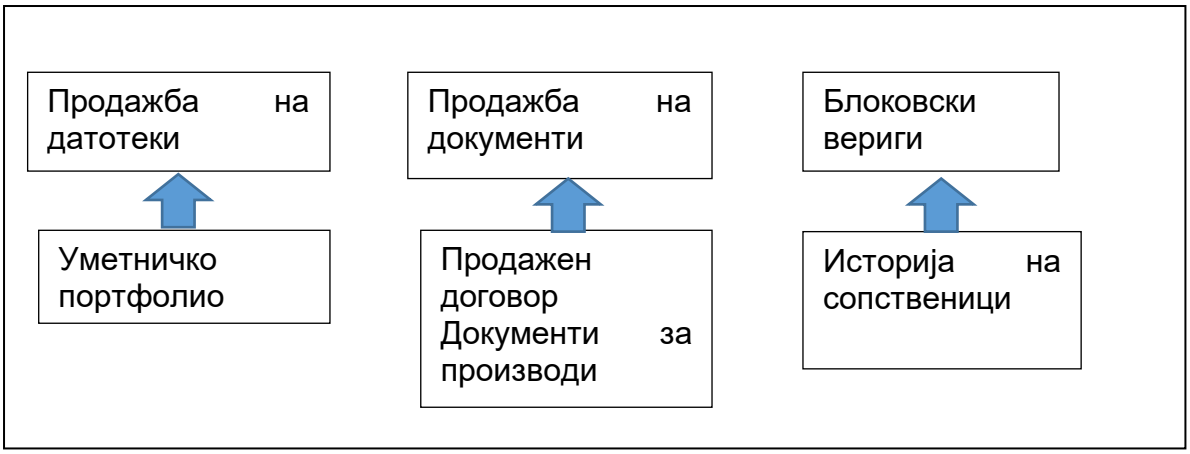
Значи, при пренесување на сопственоста, корисникот создава трансакција за едно од своите дела, ја вклучува вредноста на хешот, изданието и новиот сопственик. Корисникот ја потпишува трансакцијата и ја испраќа. Бидејќи Ascribe ја користи блоковската верига на биткоин, јавен истражувач што го знае хешот на делото може да ја следи сопственоста на делото и да ги пронајде сите адреси што ги поседува секое издание.

Друга имплементација на блоковска верига е Monegraph. Monegraph е доказ дека може да се изгради нов, модерен и дигитален пазар што е лесен за корисниците. Името Monegraph потекнува од името „Monetized Graphics“ бидејќи им помага на уметниците и на сопствениците да ги бараат правата и комерцијалната вредност на нивните дигитални медиуми. Од гледна точка на корисниците, со Monegraph е лесно да се купуваат и продаваат целосно лиценцирани дигитални медиуми директно со услови, права и цени што ги контролираат авторите. Monegraph го олеснува процесот на лиценцирање и примање приходи за уметност во многу дигитални форми создадени од фотографи, дизајнери, илустратори и други креатори на медиуми [27].

Monegraph им овозможува на авторите сами да создадат и прилагодат договор за лиценца со кој се утврдуваат параметрите за употреба за нивните медиуми. Постојат четири типа на лиценци:

- лиценца за уметнички дела, за некомерцијална употреба;
- лиценца за фотовести, за уредување;
- лиценца за слика на производ, комерцијална лиценца, и
- слика на состојбата (snapshot) е лиценца која дава целосно право.

Авторите имаат јавен каталог како портфолио на нивната работа што е јавно достапен и се продава. Блоковската верига води сметка за историјата на сопственоста. Но, чувањето информации само за сопствениците може да не биде доволно на пазарот. Постојат многу податоци за договорот за продажба и метаподатоци за производот, кои се подеднакво важни како и производот. Поради големината на тие податоци, не може да се вклучат во блоковската верига. Затоа, Monegraph ја гледа потребата од систем на блоковска верига што може да се имплементира и за други дигитални пазари. Она што е потребно за решението е интеграција со други услуги. Значи, податоците за сопственоста се чуваат во блоковската верига за да останат доверливи, следливи и неповратни. Но, другите документи поврзани со производот можат да се чуваат во документ – ориентирани бази на податоци, како што се MongoDB или CouchDB. Документите можат да бидат јавни, шифрирани или не. Самата дигитална уметност може да се чува во складиште за документи до кое може да се пристапи со HTTP или P2P. На пример, датотеката може да се чува во Amazon Simple Storage Service (S3). Екосистемот треба да најде начин да ги поврзе блоковската верига, складиштето со документи и складиштето на дигитална уметност, прикажано на слика 18.



Слика 18. Екосистем на блоковски вериги [20].

Системот за трансфер на недвижнини Bitmark овозможува пренесување и на дигитални и на физички објекти [28]. При зачувувањето информации за физички предмети, како автомобил, компјутер или куќа, како апстрактна сопственост, се појавува проблем. Лесно може да се земе отпечаток од прсти за дигитални податоци со примена на хеш-алгоритам, но кога станува збор за физички објекти, постои ново решение наречено „ObjectMinutiae“ [29]. Ова е рамка за идентификација на физички средства врз основа на уникатни обрасци на текстура на површинско ниво. Секое средство во регистарот Bitmark прво се запишува со отпечаток, метаподатоци и потпис на регистраторот. Нов Bitmark се создава кога средството се пренесува на нов сопственик и сопственоста може да се промени со трансакција за пренос на сопственост.

Може да се заклучи дека досега Интернетот остави многу материјали со изгубено авторство. Но, гледајќи во иднината, технологијата на блоковски вериги може да обезбеди сигурен доказ за сопственоста со зачувување хеш-вредност на дигиталната уметност во временска рамка за трансакција. Со претходните примери се обезбедуваат начини за потврдување на автентичноста на уметничките дела преку Интернет и во реално време. Кај сите нив концептот е ист: сопственикот ќе го поседува приватниот клуч и оригиналната копија на хешираната уметност. Никој не може да докаже поинаку и ниту една институција не може да ги промени податоците. Авторот може да го продаде дигиталното дело, а новиот сопственик сега ќе го има приватниот клуч за дигиталното дело. Тоа е сè што е потребно за дигиталната сопственост, и сега тоа е можно.

2.4.9 Интернет на нештата

Интернетот на нештата (IoT, Internet of Things) е напреден развоен концепт. Досегашната позната технологија за градење IoT системи резултира со разновидни протоколи кои се комплексни и со спротивставени конфигурации. Тековните IoT системи се потпираат на централизирана клиент: сервер архитектура. Сите уреди се идентификуваат, автентицираат и се поврзуваат преку серверите за облак (cloud).

Врската меѓу уредите оди преку Интернет. Дури и ако ова решение функционира добро засега, можеби нема да може да одговори на потребите на поголемите системи на IoT во иднина [30]. Продолжувањето на развојот на IoT на децентрализиран начин се смета како вистинска насока, но поголемиот дел од технологијата недостасува, на пр. приватност и безбедност во огромни мрежи на IoT. Технологијата на блоковски вериги може да стане идеална компонента и основен елемент за следење на милијарди поврзани уреди, обработка на трансакции и координација на уредите. Таа ќе овозможи пораки со рамноправен пристап (peer-to-peer), дистрибуција на датотеки и автономна координација меѓу уредите без потреба за централизиран облак. Блоковската верига е рамка што ги олеснува трансакциите и координацијата меѓу уредите. Секој уред ќе ја има својата улога и ќе управува со своето однесување во новиот Интернет на децентрализирани и автономни работи [31].

Уредите IoT ги контролира корисникот од централна точка. Централна точка може да биде мобилниот уред на корисникот. Сите активности, команди и правила ги поставува корисникот. Иако ова е добро за лична контрола, таа не е автоматизирана на многу начини. Вистинската револуција може да се случи ако сите уреди се контролираат од блоковска верига наместо со директната контрола на корисниците. Ова е можно со користење на паметни договори. Паметниот договор е збир на услови и деловни правила што мора да бидат исполнети пред трансакцијата да биде вклучена во блоковската верига. Трансакцијата што е напишана во блоковската верига може да биде покомплексна отколку само пренесување на сопственост. Паметните договори имаат интегриран механизам за спроведување на различни типови договори меѓу јазлите. Паметниот договор е, исто така, автономен и, технички гледано, тоа е компјутерски код кој може да се самоодржува и да се самоизвршува. Откако ќе стапи на сила, не е потребен човечки фактор за да се контролира [32]. Извршувањето на паметни договори е овозможено од етериум, платформа за создавање на системи на блоковски вериги. Етериум има своја мрежа, јазли и рудари, исто како и биткоин. Slock.it е прва имплементација на IoT и системите на блоковски вериги со помош на платформата етериум [33]. Таканаречените слокс (Slocks) се реални физички објекти кои можат да бидат контролирани со блоковска верига. Тие користат компјутер етериум, кој е парче електроника што ја внесува технологијата на блоковска верига во целиот дом, што овозможува да се изнајми пристап до кој било компатибилен паметен предмет и да прифати плаќања без посредници.

Slock.it овозможува секој да изнајмува, продава или споделува нешто – без посредници. Функционира на следниов начин: сопственикот на паметен предмет (Slock) создава паметен договор за негова употреба поставувајќи ја цената и депозитот. Корисниците можат да го најдат слокот и потоа да извршат исплата на блоковска верига на етериум, со што ќе добијат дозвола да го отворат или затворат тој слок, што значи да го користат според договореното. Паметниот договор автоматски се спроведува, при што депозитот се враќа на корисникот минус цената на изнајмувањето. Практично, ова значи дека со инсталирање паметно заклучување на

вратите од станот, корисниците можат да изнајмат стан користејќи блоковска верига. Паметниот договор ќе го отклучи и ќе го направи достапен како што се утврдува со договорот. Покрај паметните врати, овој систем овозможува изнајмување, продажба или споделување на кој било паметен предмет што има вградена технологија Slock.it. Велосипеди, автомобили и кој било предмет што може да се обезбеди со физичко заклучување е случај на потенцијална употреба. Поради недостатоците на досегашните архитектури на IoT се предлага нова безбедна, приватна и лесна архитектура на IoT за паметен дом, заснована на технологијата на блоковска верига [34]. Овде, рударството на блоковите се смета за прв проблем затоа што IoT-уредите се уреди со ограничен ресурс и не можат да вршат таква операција. Исто така, IoT-уредите треба да дејствуваат во истата секунда кога е откриена или зададена наредба. Потребното време за рударство во блокови, во повеќето случаи не може да биде прифатливо. Со технологијата IoT ќе се поврзат милијарди уреди, а тоа се многу повеќе јазли од досегашното искуство со блоковски вериги. Затоа, предложеното решение [34] содржи три нивоа:

- локална мрежа,
- мрежа за преклопување и
- складирање на облак.

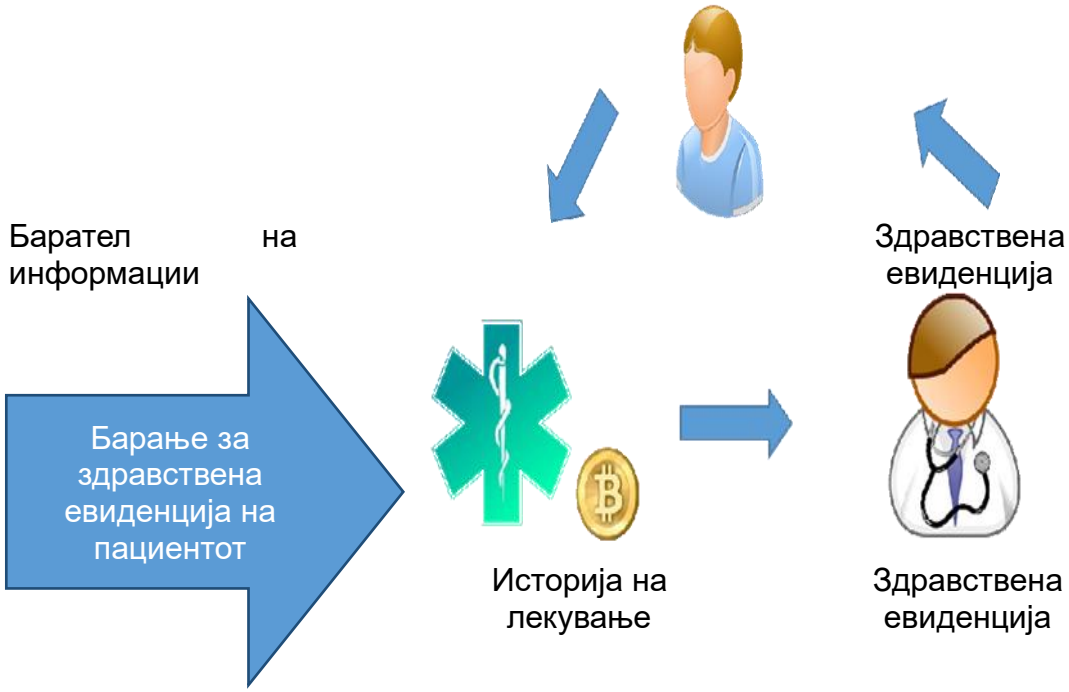
Локалната мрежа ги содржи сите паметни домашни објекти и локален компјутер што дејствува како локална блоковска верига која постојано е на Интернет. Оваа локална блоковска верига е централно управувана од нејзиниот сопственик. Кога има нов паметен уред во домот, корисникот го додава во блоковската верига. Сите трансакции поврзани со одреден уред се поврзани во синџир. Не постои стандардно рударство, па кога е примена трансакција, автоматски се става во блок и се смета за валидна. Мрежата за преклопување е мрежа на рамноправен пристап, која поврзува повеќе паметни домови и корисници. Оваа мрежа управува со јавните клучеви, на корисниците им е дозволен пристап до податоците за паметните домови и јавните клучеви на паметните домови кои обезбедуваат достапни податоци. Складирањето на облакот е вклучено како решение за уредите кои можеби сакаат да складираат податоци во облакот, така што трета страна може да пристапува до податоците и да обезбеди одредени паметни услуги. Користејќи блоковски вериги, IoT може да премине кон мрежа на уреди што можат да комуницираат едни со други и со околината без човечка интервенција. Уредите, исто така, ќе донесуваат паметни одлуки, така што многу текови на работа ќе се автоматизираат на нови начини, постигнувајќи значителни заштеди на време и трошоци.

2.4.10 Здравство

Големите количества податоци во здравството доаѓаат од најразлични извори, како што се клинички испитувања, електронски записи, бази на податоци на пациенти, медицински мерења и слики. Сите овие податоци доаѓаат во широк спектар на формати

и од различни текови на податоци. Податоците треба да се проценат и интерпретираат навремено за да им бидат од корист на пациентите. Но, на лекарите им требаат нови алатки за да ги следат, проследуваат и да обезбедат брза повратна информација за индивидуалните пациенти. Правилното управување со податоците, исто така, ќе помогне во стратегиите за предвидување, интервенциите, здравствените услуги и здравствените политики. Бидејќи медицината е секогаш во чекор со технологијата, придвижувајќи многу иновации, можеме да заклучиме дека податоците за здравствената заштита имаат потреба од круцијална трансформација, исто како и големите податоци.

Една интересна имплементација на технологијата на блоковски вериги е во здравствениот систем, каде што се вклучени сите засегнати страни, како што се болници, здравство, здравствени власти, преку задоволување на потребите на потрошувачите и заштита на приватноста на пациентите со користење на блоковска верига за плаќање на услугата со биткоин [10]. Во постоечкиот систем, каде целосната евиденција се води на хартија, доколку барателите на информации треба да видат здравствена евиденција на пациентот, тие мораат да пополнат формулар за барање и да го испратат до канцеларијата за регистрација за одобрување. По добивањето на одобрение, барателот на информации ќе плати за копија на благајната и ќе добие сметка за потврда на плаќањето. Барателот на информации потоа ја покажува сметката во канцеларијата за регистрација за да добие копија од здравствената евиденција на пациентот. Сепак, здравствената евиденција на пациентот може да се изгуби или да се направат копии за нелегални цели. Концептот на електронски систем за здравствени записи со употреба на технологијата на блоковска верига е прикажан на слика 19.



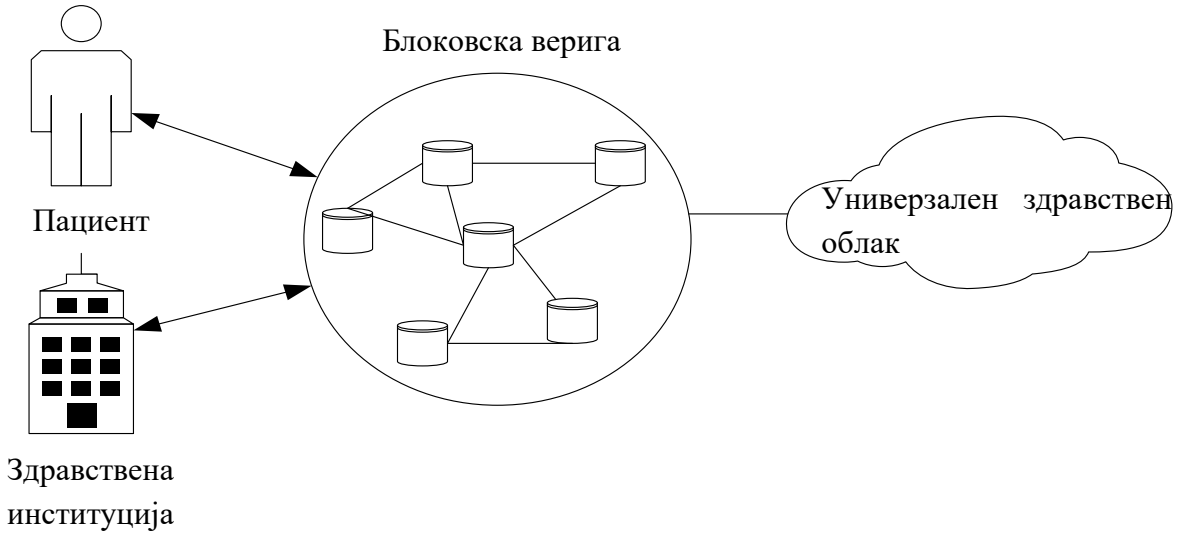
Слика 19. Е-здравствен систем со употреба на блоковска верига.

Кога барателот на информации ќе испрати барање за здравствена евиденција на пациентот до издавачот (болница или установа за здравствена заштита) и издавачот се согласува со барателот на информации, ќе биде поставен биткоин. Пред да се испратат здравствените записи на пациентот до барателот на информации, потребно е одобрување од матичниот лекар и од пациентот, така што, потребно е за пациентот да се испраќаат само конкретни записи, на пример, записи за ментално здравје.

Тал Рапки во својот пристап кон здравството [35] дискутира дека можеби промената треба да дојде од точка каде што луѓето поседуваат и можат да пристапат до податоци за своето здравје. Технологијата на блоковска верига со својот концепт за работа има начин да го донесе овој пристап насочен кон потрошувачите во здравствениот сектор. Податоците од резултатите и процедурите можат да се чуваат на блоковска верига која нема да се базира на еден централен објект за складирање. Ова ќе им помогне на владите и другите претпријатија да бидат ослободени од одговорноста за тие податоци. Во исто време, податоците ќе престојуваат на најновата безбедна технологија и со употреба на криптографија. Како сопственици на податоците, потрошувачите ќе имаат овластување да одлучат со кого ги споделуваат своите податоци. Здравјето ќе биде ориентирано повеќе кон потрошувачот, но сепак е во рамнотежа со останатите важни играчи во здравствениот систем.

Начинот како технологијата на блоковски вериги може да овозможи интероперативна и безбедна размена на електронски здравствени записи во која здравствените потрошувачи се крајни сопственици е објаснет во [36]. Предложеното сценарио е да се зачуваат само метаподатоците за здравствени и медицински настани на блоковска верига. Во спротивно, инфраструктурата на блоковска верига ќе мора масовно да се обедини за да поддржат целосни здравствени записи. Значи, метаподатоците како што се идентитет на пациентот, лична карта за посета, лична карта на давателот, лична карта на исплаќачот итн., може да се чуваат на блоковска верига, но вистинските записи треба да се чуваат во посебен универзален здравствен облак. На пример, ако пациентот посети две болници денес, тие ќе ги чуваат податоците за него во две бази на податоци што пациентот не ги поседува. Ако болниците треба да комуницираат, тие ќе користат медијатор за стандардизирана комуникација како веб-услуги, е-пошта или складиште за споделени датотеки. Во сценарио каде што се применува блоковска верига, првата болница создава запис на универзалниот здравствен облак. Потоа, болницата создава трансакција во блоковска верига со метаподатоците за посети и URL-то за записот во облакот. Пациентот ја потпишува оваа трансакција со својот клуч. Кога пациентот сега ја посетува втората болница, тој мора да го обезбеди својот клуч за да ги прочита трансакциите на блоковска верига. Само засегнатите страни со клучот на пациентот можат да ги дешифрираат трансакциите. Значи, ова е пример за тоа како луѓето можат да ги поседуваат податоците и да го одобрат потребниот пристап, прикажано на слика 20. Дури и паметните договори можат да се кодираат во блокови за да носат упатства за

осигурување, итни контакти, тестаменти, итн. Овие паметни договори ќе се активираат со настани што блоковската верига може да ги чита од друга веб-услуга [96].



Слика 20. Преглед на системот за здравствена заштита на блоковската верига [20].

Според друго истражување, нема смисла да се пресликуваат податоците на сите пациенти во јазли на блоковската верига [20]. Трансакциите во блоковите треба да содржат единствен идентификатор на корисникот, шифрирана врска до здравствениот запис и временска ознака за моментот кога е направена трансакцијата. Трансакцијата може да содржи и тип на податоци што се зачувани. Во зависност од имплементацијата, ова може да помогне во пребарувањето и обработката на пристапните податоци. Оваа блоковска верига треба да содржи историја на сите медицински податоци, вклучително и формални медицински досиеја, како и податоци од мобилни апликации и сензори што се носат. Одржувањето на податоците подалеку од блоковската верига, во базен на облак, може да претставува добра основа за пребарувања, рударство, аналитика и машинско учење. Овој вид на анализа не смее да влијае врз приватноста на секој пациент. Овие податоци треба да бидат шифрирани и дигитално потпишани за да се обезбеди приватност и автентичност на информациите. Корисникот би имал опција да додели збир на дозволи за пристап и да назначи кој може да ги побара и да напише податоци на неговата блоковска верига. Понатамошниот развој на кориснички интерфејс за пациентот да ги прегледува своите здравствени податоци и да управува со привилегиите за пристап, се разбира, е сосема можен и потребен.

Друга студија на случај за блоковска верига во здравството ги користи паметните договори на етериум за да создаде репрезентации на постојни медицински досиеја [38]. Овие договори се чуваат директно во одделни јазли на мрежата. Предложеното решение наречено MedRec, ги структурира големите количества податоци во три вида договори. Првиот е договор за регистратор. Во него се чува идентитетот на учесниците со сите потребни детали и, секако, со јавните клучеви. Овој вид на регистрација на

идентитет може да се ограничи само на овластени институции. Втор договор е договорот за односи меѓу давателот на грижа и пациентите. Главната употреба ќе биде кога постои паметен договор меѓу давателот на грижата и пациентот. Последниот е склучен договор што му помага на пациентот да ја лоцира својата медицинска историја. Како резултат на овој договор, наведени се сите претходни и тековни ангажмани со други јазли во системот. MedRec, исто така, предлага и рударски модел кој ја вклучува целата здравствена заедница во рударството. Медицински истражувачи и засегнати страни од здравствената заштита можат да работат во мрежата. Наградата за рударство, всушност, може да биде одобрен пристап до збирни анонимизирани медицински податоци.

Технологијата на блоковски вериги нуди иднина што ветува дека пациентот ќе има централна улога во здравството и дека ќе им помогне на пациентите да ги откријат и да управуваат со сопствените медицински досиеја. Но, потребен е глобален стандард за чување, пристап и споделување на шифрирани податоци на облакот. Гледајќи кон иднината, сите предложени решенија имаат потенцијал да ангажираат милиони лица, даватели на здравствени услуги и медицински истражувачи. Решенијата може да влијаат на огромниот напредок во медицинските истражувања.

2.4.11 Дигитален идентитет

Поради безбедносниот механизам што штити од нарушување, блоковските вериги можат да играат клучна улога во обезбедувањето дигитални идентитети [39]. Блоковските вериги можат да го заштитат идентитетот со шифрирање. Исто така, блоковските вериги можат да се користат за изградба на многу силен, безбеден и непробоен систем за идентификација, што може да спречи неовластена активност. Технологијата на блоковски вериги има потенцијал да ги замени сите постоечки физички идентитети и да ги премести на дигитална платформа. Опсегот на идентитет, како што се пасоши, возачки дозволи, лични карти, па дури и избори за гласови, може дигитално да се генерира со помош на технологијата на блоковски вериги [39]. Исто така, можно е сите идентитети да се чуваат заедно и да се обезбедат со блоковски вериги. Употребата на блоковски вериги не може да изврши неовластена промена на разни сертификати, како што се диплома за завршено образование, извод од матична книга на венчани, изводи од матична книга на родени или умрени, со што се спречува неовластена и злонамерна модификација.

2.4.12 Финансиски услуги и инфраструктура

Технологијата на блоковски вериги може да обезбеди платформа за подобри финансиски услуги и можности за плаќање. Употребата на криптовалути, како што е биткоин, може да ги замени постоечките системи за плаќање и други финансиски услуги [41]. На пример, ако едно лице испраќа пари на други лица во друга земја, можни средства за трансфер се банки, апликации за плаќање (како што е PayPal) или други посреднички организации. Но, нивните трошоци за услуга се високи, дури и за

мали трансакции. Сите овие посредници може да се елиминираат и парите да се префрлат директно од испраќачот до примачот користејќи криптовалути, без вклучување на посредник. Следење на трансакциите и правата на сопственост, исто така, може да се имплементираат во финансиските сектори користејќи блоковски вериги. Употребата на технологија за блоковски вериги во финансискиот сектор не само што ќе обезбеди бесплатен систем за плаќање, туку и ќе обезбеди сигурен начин за спроведување трансакции преку Интернет.

2.4.13 Е-трговија

Ако се имплементира правилно, технологијата на блоковски вериги може масовно да поддржува е-трговија и малопродажба во однос на растот, продажбата и маркетингот. Трговијата на мало веќе започна да забележува раст и профит од продажба на стоки и услуги за широка потрошувачка користејќи технологија на блоковски вериги. Интернетот делува како одлична платформа за промовирање локални бизниси и друга содржина, но секогаш постои ризик содржината да се користи без соодветна дозвола. Севкупниот плагијат може да биде ограничен со техники за тампонирање на блоковски вериги и со тоа да се зачува оригиналноста на секоја содржина. Спроведувањето на технологија на блоковски вериги во малопродажната индустрија, исто така, ќе обезбеди јасен и транспарентен систем за управување со синџирот на снабдување, што ќе им овозможи на корисниците да добијат увид во потеклото на нивната храна и други производи. Веб-страниците за е-трговија и други компании, како што се OpenBazaar, Provenance, Everledger, Ascribe, BlockVerify, се некои од деловните активности поддржани од блоковските вериги вклучени во малопродажната индустрија [39]. Ова ги елиминира брокерите и провизиите и може да воспостави директен канал за трансакции помеѓу купувачите и продавачите. На овој начин, продажбата преку Интернет ќе ги промовира деловните трансакции на мало и економијата.

2.4.14 Образовни записи

Корпорацијата Sony, во партнерство со „Sony Global Education“, аплицирала за патент за складиште базирано на блоковски вериги, што ќе вклучува евиденција на студенти, вклучувајќи завршени курсеви, резултати од тестови, дипломи и друго, во форма на дигитален запис [42]. Таквиот систем може да им овозможи на наставниците и учениците да пристапат до релевантните податоци додека ја одржуваат приватноста. Исто така, може да им обезбеди на потенцијалните образовни институции и потенцијалните работодавачи транспарентно, сигурно место за стекнување на акредитивите на апликантите.

2.4.15 Образовен систем

Во времето на пандемија со коронавирусот Ковид-19, кога наставата во училиштата се одвиваше преку Интернет, се појавија многу проблеми. Часовите беа

скратени и наставникот требаше да посвети драгоцено време постојано да проверува кои ученици се присутни на часот. Исто така, директорот на секое училиште сака да знае дали наставниците од персоналот редовно и навремено ги одржуваат часовите. Овие проблеми би можеле да се надминат доколку податоците од секој клас се снимени на блоковски вериги. На тој начин, никој нема да може да манипулира со тие податоци. Може во секое време да се провери кој наставник и кога предавал и кој ученик присуствувал на часот. Ако учениците знаат дека сè е снимено и податоците не можат да се сменат, тие сигурно ќе посетуваат настава поредовно. И тоа е најважниот услов за да го подобрат својот успех. Подобрениот успех на учениците е најголемото задоволство за наставниците. Родителите, исто така, би биле позадоволни од таквото однесување на своите деца. Целото општество би имало најголема корист од добро образование.

2.4.16 Споделување знаење

Еврипедија (Everipedia), првата енциклопедија во светот што користи систем од блоковски вериги, објави планови за изградба на нова вики-мрежа со отворен код, која ќе ја децентрализира базата на знаења на Википедија, овозможувајќи секој уредник да стане веб-администратор. Според Лари Сангер, еден од основачите на Википедија, а сега вршител на должноста директор за информации на Еврипедија, оваа енциклопедија ќе се потпира на блоковска верига која им овозможува на корисниците да се однесуваат поодговорно. Уредниците во Еврипедија ќе можат да заработуваат IQ „токени“ врз основа на своите корисни придонеси, што ќе ги претставуваат виртуелните акции на платформата [41]. Способноста на поединецот да биде акционер во енциклопедијата, што тој/таа ја уредува и за возврат да добие вистинска парична вредност е привлечна идеја. Клиентите ќе треба да платат депозит пред да дадат придонеси. Ако нивните промени се сметаат за неточни, тие ќе го изгубат токенот, а оние чии промени се точни ќе добијат оригинален депозит и дополнителни токени како награда. Теодор Форселиус, исто така, наведува две дополнителни придобивки од преместувањето на Еврипедија во блоковски вериги. Првата придобивка е што податоците повеќе нема да се складираат на централизиран сервер, што значи дека ќе преживеат дури и ако централната организација, Еврипедија, престане да постои. Втората придобивка е што ќе биде невозможно да се цензурираат податоците, што значи дека владите што моментално ја цензурираат Википедија нема да можат да ги спречат корисниците да придонесат за платформата.

2.4.17 Осигурување

Технологијата на блоковски вериги има примена во осигурителниот сектор бидејќи секторот се базира на договори и доверба меѓу две страни. Употребата на блоковски вериги во овој поглед би значела дека и договорот за осигурување и личните информации на потрошувачот може да се чуваат во дистрибуираната евиденција, додека потрошувачот контролира кој има пристап. Податоците остануваат на уредот на

корисникот и тоа може да ја елиминира потребата за брокери и други посредници меѓу осигурителните компании и потрошувачите. Присуството на паметни договори веќе се чувствува во осигурителниот сектор. Осигурителната компанија AXA работеше со веб–страница наречена Fizzy, која им обезбеди на потрошувачите да не се одложуваат летови за два или повеќе часа [47]. Fizzy ја евидентира набавката на авионски билети на етереум блоковска верига и ги поврзува паметните договори што произлегуваат од глобалните бази на податоци за воздушниот сообраќај. Доколку се забележи доволно задоцнување, надоместокот се плаќа автоматски. Друг пример за користење паметни договори во осигурителниот сектор е да се обештетат земјоделците во случај на суша или друга катастрофа што им наштетува на нивниот имот. Другата примена на блоковски вериги во осигурителната индустрија е нејзината улога во потенцијално намалување на измамите [47]. Во овој случај, блоковската верига ги евидентира сите осигурителни полиси и сите побарувања во еден дистрибуиран регистар.

2.4.18 Прехранбена индустрија

Технологијата на блоковски вериги е потенцијално корисна за сите во прехранбената индустрија. Кога се јавува болест предизвикана од храна, ресторaните што служат храна или продавниците за храна често имаат потешкотии да го откријат изворот на зараза. Следењето со блоковски вериги ќе помогне веднаш да се следат засегнатите производи и нивното потекло, брзо да се лоцира проблемот, така што контаминираните производи може да се отстранат од менијата, полиците и синџирите за снабдување [44]. Стандардите и репутацијата на продавачите базирани на блоковски вериги ќе обезбедат интегритет на маркетинг-тврдењата. Сертификатите и извештаите за постојната ревизија на постојните објекти ќе бидат регистрирани на страницата за да се докажат горенаведените тврдења. Ако сите добавувачи во синџирот на снабдување ги следат овие правила и ги запишуваат податоците за потеклото на храната во децентрализирани системи за мониторинг, оние што даваат лажни тврдења или погрешно го прикажуваат потеклото на своите производи ќе бидат уништени. Технологијата на блоковски вериги им овозможува на земјоделците и на производителите пристап во реално време до цените на суровините и пазарните податоци. На овој начин, земјоделците имаат подобри информации за пазарот и можат да бидат поконкурентни и попродуктивни. Сè повеќе луѓе сакаат да знаат што содржат производите што ги консумираат. Тие сакаат да можат да направат сигурен избор на храна за себе, за своите семејства и за своите заедници. Технологијата на блоковски вериги ќе помогне да се изгради таа доверба.

2.4.19 Сметководство и ревизија

Бидејќи промената на трансакциите или цели блокови се речиси невозможни во блоковски вериги, употребата на оваа технологија го олеснува докажувањето на интегритетот на електронските датотеки. Еден пристап е да се генерираат хеш-низи за постоечките докази, како што се фактурите. Ова множество на хеш-низи е дигитален

отпечаток. Понатаму, овој отпечаток е непроменлив и е снимен на блоковски вериги преку трансакција. Во секое време е можно да се докаже интегритетот на таа датотека со повторно создавање на дигитален отпечаток и споредување со дигиталниот отпечаток зачуван во блокот. Во случај дигиталните отпечатоци да бидат идентични, докажано е дека документот останал непроменет од првото поставување блок во блоковската верига [41]. Компаниите ќе имаат корист од овој тип на имплементација на блоковски вериги. Стандардизацијата ќе им овозможи на финансиските ревизори автоматски да потврдат голем дел од најважните податоци за финансиското известување. Трошоците и времето потребно за извршување на ревизијата ќе бидат значително намалени, а времето може да се искористи за проверка на многу сложени трансакции или механизми за внатрешна контрола.

2.4.20 Прекугранични плаќања

Прекуграничните плаќања, генерално, се однесуваат на транснационален и трансрегионален трансфер на средства меѓу две или повеќе земји или територии преку меѓународна трговија, меѓународни инвестиции и други меѓународни побарувања и долгови, користејќи одредени инструменти за порамнување и платежни системи. Традиционалното прекугранично плаќање се базира на банкарскиот систем којшто ги има следниве карактеристики: одзема многу време, има висока цена, зафаќа повеќе средства и има ниска безбедност. Сепак, сите овие тесни грла можат ефикасно да се надминат со примена на блоковски вериги за реконструкција на кредитниот систем и проширување на границата за плаќање. Истражувачите истакнуваат дека примената на технологијата на блоковски вериги за прекугранично плаќање има висок потенцијален ефект. Технологијата на блоковски вериги ќе го подобри платниот систем преку обезбедување солидна структура за прекугранични трансакции и отстранување на скапите посреднички трошоци и постепено ќе го ослабне или измени деловниот модел на постојните платежни индустрии [55]. Yao и Zhu предлагаат технологијата на блоковски вериги да се усвои за прекугранично плаќање врз основа на примена на блоковските вериги на Visa и Swift. Конзорциумот R3 работи со 22 од банките членки за да изгради решение за прекугранични плаќања во реално време на Корда, што е распределена книга на конзорциумот „инспирирана од блоковски вериги“ [56].

2.4.21 SARS-CoV-2 вирус наспроти Црвениот крст: подобри решенија преку блоковски вериги и вештачка интелигенција

За време на ковид-кризата во Вухан, кинеската влада нареди сите јавни донации да бидат испратени до пет добротворни организации поддржани од владата. Ова е враќање назад, пред да биде воведен Законот за добротворно организирање во Кина во 2016 година со цел да се овозможи основање приватни добротворни организации. Законот за добротворни цели беше наменет за развој на добротворното поле и заштита на интересите на релевантните засегнати страни [57]. Иако од сите добротворни организации во Кина се бара да имаат добри структури за внатрешно управување, се

претпоставува дека петте добротворни организации поддржани од владата се соодветни и подобро оспособени да управуваат со тековната криза. Таа претпоставка може да биде во спротивност со историските и поновите докази, кои сугерираат дека организациите одговорни за одговарање на кризи се чини дека се борат да управуваат со своите основни одговорности [57]. Во овој случај, и не за прв пат, Црвениот крст во Кина беше во центарот на лутината на јавноста. „Една од научените лекции беше дека одговорот на итни случаи мора подобро да се развие на локално ниво“. Ова го кажа Црвениот крст во 2017 година на десетгодишнината од смртоносниот земјотрес во провинцијата Сечуан во западна Кина. Милијарди долари биле донирани по земјотресот во Сечуан, но биле „погрешно ракувани“. Јавноста во Кина повторно била лута поради погрешното ракување со донациите, а тоа влијаело на подготвеноста за донирање. Технологијата на блоковски вериги и вештачката интелигенција сега се во честа употреба од глобалните технолошки компании и претставуваат алатки што можат да се користат за подобро управување со кризи. Приватната блоковска верига ќе овозможи снимање и следење на сè што е донирано, од парични донации до маски N95. Исто така, се создаваат јасни точки во кои е можно да се бара одговорност од лице или организација, од вчитување донации за испорака до конечна крајна употреба. Важно е дека на блоковските вериги, исто така, може да им се даде јавна видливост, обезбедувајќи транспарентност за сите засегнати страни – донатори и корисници, како и тела за јавен надзор. Секој може да го следи напредокот и користењето на својата донација.

2.4.22 Е-гласање

Употребата на технологијата на блоковски вериги би спречила секој учесник да може да мами, од гласачите до бројачите на гласови во спроведувањето на изборите. Технологијата на блоковски вериги ќе се погрижи поединецот да не може да гласа неколку пати затоа што постои непроменлив запис за нивниот глас и идентитет. Исто така, никој не може да ги избрише гласовите, бидејќи, како што беше кажано, податоците запишани во блоковската верига се непроменливи. Одговорните за пребројување на гласовите ќе имаат конечен запис за бројот на гласови што регулаторите или ревизорите можат да го контролираат во секое време. Резултатите може да се шифрираат, што би ја зголемило транспарентноста, додека го одржува клучното чувство за приватност. Резултатите внесени и зачувани во блоковски вериги не се само непроменливи и транспарентни, туку и достапни. Ова значи дека гласањето со блоковски вериги е поефикасно од традиционалното гласање [45]. Оваа технологијата може да се користи и за подобрување на процесите на гласање во јавни и приватни компании и организации.

3 Машинско учење

3.1 Опис

Машинско учење е научна област што им дава на компјутерите можност да учат без да бидат експлицитно програмирани [97]. За компјутерска програма се вели дека учи од искуството *E* во однос на одредена задача *T* и одредена мерка за успешност *P*, доколку нејзината успешност *P* на задачата *T* се подобрува со искуството *E* [98].

Алпајдин (2004) го дефинира машинското учење како процес на програмирање компјутери за оптимизирање на перформансите на критериумите со користење на стекнати податоци или искуство. Машинското учење им овозможува на машините, најчесто компјутерите, да учат, односно да стануваат сè подобри и подобри при решавање на одредени проблеми во однос на „искуството“. Со други зборови, машината учи од податоците, и колку повеќе податоци има на располагање машината, толку е поголема шансата да го реши проблемот со подобра прецизност или точност.

На пример, филтерот за несакана пошта е програма за машинско учење што може да научи да обележува „спам“ пораки. Примерите што системот ги користи за да научи се наречени множество за обука. Секој пример за обука се нарекува примерок. Во овој случај, задачата *T* е да означува спам за нова е-пошта, искуството *E* се податоците што се користат за обука и треба да се дефинира мерката за перформанси *P*, на пример, може да се користи односот на правилно класифицирана е-пошта спрема вкупно добиената пошта. Оваа конкретна мерка за изведба се нарекува точност и често се користи при класификација на задачи [58].

Спам-филтерот базиран на техниките за машинско учење, автоматски учи кои зборови и фрази се добри индикатори за несакана пошта, откривајќи невообичаени чести обрасци на зборови во примерите со несакана пошта во споредба со примерите со посакувана пошта. Програмата е многу пократка, полесна за одржување и најверојатно поточна во споредба со програмата со традиционално програмирање.

Машинското учење е наредниот предизвик кај технологијата на блоковски вериги [59]. Програмскиот интерфејс на машинското учење не е комплексен и лесно е разбирлив за секој и најмногу се користи кај развивачите на апликации коишто сакаат да се занимаваат со податочно рударење, кориснички дизајн, експериментирање и анализа на податоци.

Во претходните децении, алгоритмите за машинско учење и нивните технологии, најчесто биле користени при напредна анализа на податоци. Денес, неколку организации го користат програмскиот интерфејс на машинското учење со цел негово популаризирање и доближување кон обичните корпорации. Интерфејсот е креиран со цел лесна имплементација на машинското учење врз податоците при поставувањето на одредени претпоставки во апликациите што обработуваат податоци.

Машинското учење го подобрува квалитетот на животот на луѓето преку бројни апликации. Меѓу апликациите на машинско учење во областа на здравството, науката,

индустријата и сл. е навременото откривање на болести како што се канцер, глауком и други болести коишто одземаат човечки животи со голема брзина. Друга примена е визуелизацијата на паметни автомобили, ефикасното пребарување на веб, што ги олесни пребарувањата на Интернет, јазичните преводи кои неизмерно помагаат во светските комуникации и ја ограничуваат големата јазична бариера меѓу земјите, реализацијата на системите за откривање измама и препознавање на лица.

Машинското учење е аспект на компјутерската наука што им овозможува на компјутерите да извршуваат одредена задача со учење. Преку учењето, системите се способни да се прилагодат да извршуваат слични или сродни задачи без да бидат експлицитно програмирани за тие задачи користејќи го претходното искуство. Машинското учење користи податоци и разни алгоритми со цел да се постигне процесот на учење користејќи го претходното искуство. Познати алгоритми за машинско учење се: алгоритамот k-најблиски соседи, вештачки невронски мрежи, случајни шуми, машини со носечки вектори, наивниот Бајесов класификатор и други [58].

Класификацијата на системите базирани на машинско учење може да се сврсти во неколку категории. Првата категорија зема предвид дали се обучени со надзор или не. Според ова, системот е поделен на: надгледувано, ненадгледувано, полунадгледувано и принудно учење. Втората категорија ги класифицира системите според тоа дали тие се способни постепено да учат или да учат исклучиво одеднаш. Третата поделба се заснова на тоа дали системот може да создаде модел за предвидување базиран на примероци во податоци за обука или, едноставно, да споредува нови податоци со познати податоци. Според оваа поделба, постојат системи базирани на примери и системи базирани на модели [58].

3.2 Преглед на алгоритми за машинско учење

Алгоритмите за машинско учење доаѓаат во многу форми и можат да се класифицираат според количината и видот на надзор што го добиваат за време на обуката [108].

3.2.1 Надгледувано учење

Во **надгледуваното учење,** посакуваниот излез за моделот е веќе познат. Тој е претставен само со влезен пример и треба да научи да го произведува предвидениот излез [68].

Кај овој тип на алгоритми, учењето се прави така што се користат влезни податоци за кои е позната излезната вредност. Алгоритамот учи така што за дадените влезни податоци го споредува излезот што се генерирал (реалниот излез) со оној што треба да се добие (очекуваниот излез), пресметувајќи ја грешката што ја направил. Врз основа на грешката се прават соодветно корекции во функцијата за евалуација. Преку повторување на процесот алгоритамот учи со цел кога ќе добие влезни податоци за кои

не е познат резултатот (овие податоци кај ваквите алгоритми ќе се добиваат после завршувањето на процесот на тренирање), да може да даде што е можно попрецизен резултат до вистинскиот. Учењето кај овие алгоритми застанува кога ќе биде достигнато прифатливо ниво на перформанси. Алгоритамот за влезните податоци со доволен степен на прецизност ќе ја пресметува или предвидува соодветната излезна вредност. Во зависност од проблемот кој го решаваат, овие алгоритми може да се поделат во две групи:

- **класификациски** - за даден предмет (влезни податоци) одредуваат на која група (класа) ѝ припаѓа. Спам-филтерот е добар пример за учење со надгледување каде се користи класификација: тој е обучен со многу примери на е-пошта заедно со нивната класа (спам или посакувана порака) и мора да научи како да ги класифицира новите пораки.
- **регресивни** - за разлика од класификациските алгоритми кои предвидуваат класа, регресивните алгоритми предвидуваат вредност за функции со континуирана вредност. Во некои случаи, земајќи ги предвид предвидените вредности, може да се извлече линеарна зависност меѓу одредени атрибути (влезни податоци).

Регресивните алгоритми се обидуваат да добијат одредена нумеричка вредност во зависност од односите помеѓу својствата на податоците. Системот се обидува да го открие односот помеѓу својствата и со тоа да се предвиди потребната нумеричка вредност.

Типична регресиона задача е да се предвиди целната нумеричка вредност, како што е цената на недвижностите, со оглед на множеството карактеристики (година на изградба на недвижноста, големина на просторот, позиција, итн.), кои се нарекуваат индикатори. Целта на овој систем е да се предвиди цената на претстојните недвижнини со регресија. За да се обучи системот, треба да му се дадат многу примери на недвижнини, вклучувајќи ги и нивните индикатори и нивните ознаки (т.е. цените).

Алгоритмите како линеарна регресија, логистичка регресија, дрва за одлучување, наивен Бајесов класификатор и машини со носечки вектори, спаѓаат во групата на алгоритми за надгледувано учење.

Некои од алгоритмите што учат со надгледување, а се однесуваат на проблеми со класификација и регресија се: линеарна регресија, логистичка регресија, k-најблиски соседи (KNN), машините со носечки вектори (SVM), дрво за одлучување, случајни шуми и невронски мрежи (NN) [58].

Линеарна регресија

Линеарна регресија се користи за проучување на линеарната врска меѓу зависната променлива Y и една или повеќе независни променливи X. Зависната променлива Y мора да биде континуирана, додека независните променливи можат да бидат или

континуирани (на пример возраст), бинарни (на пример пол) или категориски (на пример боја на очи). Првичната проценка на можната врска меѓу две континуирани променливи секогаш треба да се прави врз основа на графички приказ.

Непроменлива линеарна регресија ја проучува линеарната врска меѓу зависната променлива Y и единствената независна променлива X. Моделот на линеарна регресија ја опишува зависната променлива со права линија што е дефинирана со равенката:

$$Y = a + bX$$

Прво, параметрите a и b на регресивната линија се проценуваат од вредностите на зависната променлива Y и независната променлива X со помош на статистички методи. Регресивната линија овозможува да се предвиди вредноста на зависната променлива Y од таа на независната променлива X. Така, на пример, откако ќе се изврши линеарна регресија, ќе може да се процени тежината на една личност (зависна променлива) од неговата/нејзината висина (независна променлива).

Регресијата, во основа, претставува моделирање на врските меѓу променливите што се итеративно преработени преку користење на метод на грешка во претпоставките кои се креирани во моделот. Методите за регресија се во основа статистички методи.

Алгоритам k-најблиски соседи (KNN)

Алгоритамот k-најблиски соседи (KNN, k-nearest neighbors) се користи во две форми: класификација за дискретни податоци и регресија за континуирани податоци. Принципот кај методите на најблискиот сосед е да се најде предефиниран број примероци за обука што се наоѓаат најблиску до растојанието до новата точка и да се предвиди ознаката од нив. Бројот на примероци може да биде дефиниран од корисникот, константа (k-учење на најблискиот сосед), или да варира врз основа на локалната густина на точките (радиус-базирано соседно учење). Општо, растојанието може да биде која било метричка мерка: стандардното евклидово растојание е најчестиот избор. Методите базирани на соседите се познати како методи на генерализирање на машинско учење, бидејќи тие едноставно ги „паметат“ сите негови податоци за обука, евентуално трансформирани во структура за брзо индексирање.

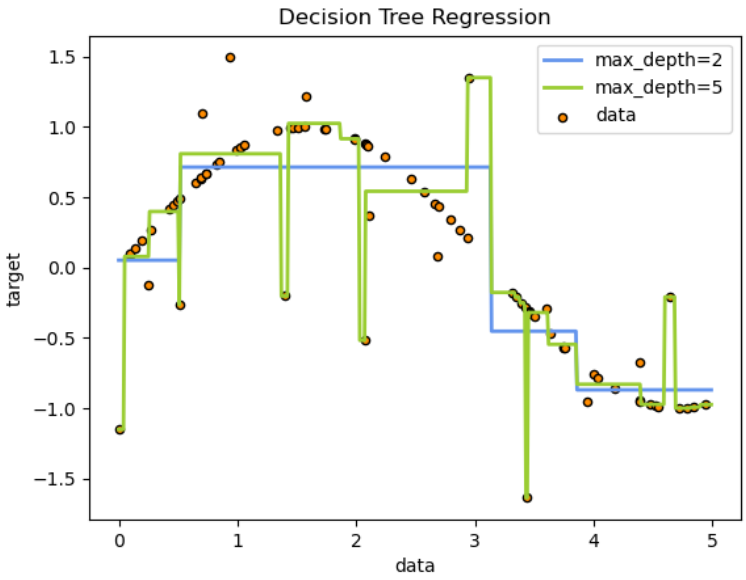
Алгоритми за машини со носечки вектори (SVMs)

Целта на алгоритмите за машини со носечки вектори (SVMs, Support Vector Machines) е наоѓање на оптимална линеарно раздвојувачка хиперрамнина низ базата на податоци со цел да ги класифицираат податоците во две групи. Тоа претставува линеарен сепаратор за која било димензија; тоа може да биде линија (2D), рамнина (3D) и хиперрамнина (4D +) [63].

Најдобра хиперрамнина е онаа што ја максимизира маргината. Маргината е растојание помеѓу хиперрамнината и неколку блиски точки. Овие блиски точки се вектори за поддршка затоа што тие ја контролираат хиперрамнината [63].

Дрва за одлучување

Дрвата за одлучување (DT, Decision Tree) се непараметарски надгледуван метод на учење што се користи за класификација и регресија. Целта е да се создаде модел што ја предвидува вредноста на целната променлива со учење едноставни правила за одлука што се изведува од карактеристиките на податоците. Секое дрво може да се смета како дел од постојаното приближување кон целта. Во примерот подолу, дрвата за одлучување учат од податоците како да ја приближат синусната крива со збир на правила за одлука од типот „ако-тогаш-друго“ (if-then-else). Колку е „подлабоко“ дрвото, толку покомплексни се правилата за одлука, а моделот е подобар, слика 21 [62].



Слика 21. Регресија со помош на дрво на одлуки [62].

Предности на дрвата за одлучување се:

- едноставност за разбирање и толкување,
- дрвата можат лесно да се визуелизираат,
- потребна е мала подготовка на податоци,
- другите техники често бараат нормализирање на податоците, треба да се создадат лажни променливи и да се отстранат празни вредности.

Овој модул не поддржува вредности што недостасуваат (missing values). Цената за користење на дрвото, т.е. предвидување на податоци, има логаритамска комплексност во зависност од бројот на податочни точки што се користат за обука на

дрвото. Овој алгоритам е способен да управува и со нумерички и со категорични податоци. Другите техники обично се специјализирани за анализа на збирки на податоци што имаат само еден вид на променлива. Овој алгоритам е способен да се справи со проблеми со повеќе излези. Користи модел на бела кутија. Ако дадена ситуација може да се забележи во модел, објаснувањето за состојбата лесно се објаснува со булова логика. Спротивно на тоа, во моделот на црна кутија (на пример, кај вештачките невронски мрежи), резултатите можат да бидат потешки за толкување. Можно е да се потврди моделот со користење на статистички тестови. Тоа овозможува да се земе предвид сигурноста на моделот.

Слабата страна на дрвата за одлучување е тоа што при учењето можат да се создадат премногу комплексни дрва кои не ги генерализираат добро податоците. Ова се нарекува прекумерно тренирање (overfitting). За да се надмине овој проблем, потребно е да се применат соодветни механизми, како што е кастрење, поставување на минимален број примероци потребни на јазол на лист или поставување на максимална длабочина на дрвото. Дрвата за одлучување може да бидат нестабилни затоа што малите варијации во податоците може да резултираат во создавање на сосема поинакво дрво. Овој проблем се ублажува со користење на дрва на одлука во рамките на еден ансамбл. Предвидувањата на дрвата за одлука не се ниту глатки, ниту континуирани, туку одделни постојани приближувања, како што се гледа на слика 21. Затоа, тие не се добри при екстраполацијата. Следствено, алгоритмите за практично учење на дрвата на одлучување се темелат на хевристички алгоритми, како што е алчниот алгоритам, каде локално оптималните одлуки се донесуваат на секој јазол. Таквите алгоритми не можат да гарантираат враќање на глобално оптималното дрво на одлуки. Ова може да се ублажи со обука на повеќе дрва кај ученик на ансамбл, каде што од карактеристиките и примероците се земаат примероци по случаен избор со замена. Постојат концепти кои е тешко да се научат бидејќи дрвата за одлучување не ги изразуваат лесно, како што е операторот XOR, проблеми со паритет или мултиплексер [62]. При процесот на учење, кај дрвата за одлучување се создаваат пристрасни дрва ако доминираат некои податоци. Затоа се препорачува да се балансира базата на податоци пред да се вклопи во дрвото на одлучување.

Случајни шуми (random forests)

Класификаторот на случајни шуми е алгоритам за надгледувано учење, што може да се користи за проблеми со регресија и класификација. Тој е меѓу најпопуларните алгоритми за машинско учење поради неговата висока флексибилност и лесната имплементација.

Името „шума“ доаѓа од фактот дека овој алгоритам се состои од неколку таканаречени дрва на одлучување. Дрвото на одлучување го претставува начинот на кој компјутерот доаѓа до заклучок. Сè започнува од коренскиот јазол. Штом алгоритамот ќе ги добие податоците со кои ќе работи, постапката што следи се сведува

на поставување на вистинските прашања. Одговорите на прашањата мора да бидат во форма на точно или неточно, односно 0 или 1. Значи, за секое прашање, секој од јазлите создава свои 2 нови подјазли. Поставувањето доволно вистински прашања води до последниот јазол во дрвото. Изборот на вистинските прашања, односно изборот на атрибутите што треба да се тестираат во одреден јазол, се врши со пресметување на информациската добивка (information gain). Информациската добивка го претставува очекуваното намалување на ентропијата предизвикано од разделување на јазлите. Поголемото намалување значи поголема добивка од информации.

Општиот алгоритам за изградба на дрво за одлуки е следниот [120]:

Чекор 1. Се пресметува ентропијата на коренскиот јазол.

Чекор 2. Се пресметува информациската добивка за сите атрибути и се избира оној атрибут што дава најголема информациска добивка.

Чекор 3. Се гради следното ниво на дрвото каде што претходно избраниот атрибут сега е коренот на дрвото.

Чекор 4. Се повторуваат чекорите од 1 до 3 сè додека ентропијата не достигне вредност 0.

Така, дрвото учи да поставува прашања што ќе го решат проблемот најбрзо и најпрецизно. По обуката, може да му се дадат податоци од ист формат и тип без резултати и да се побара од дрвото да го предвиди резултатот. Резултатот од ваквото дрво кое било обучено на еден начин и со една група податоци најчесто нема да биде многу прецизно ако дрвото има мала длабочина, а ако има голема, тоа е предмет на шум во податоците (прекумерно тренирање). Тука се доаѓа до идејата за случајни шуми. Идејата е да има повеќе такви дрва кои учат од случајно избрани податоци, па оттука и името „случајни“, за да ги направат дрвата што е можно поразновидни. Дрвата може да имаат и поголема длабочина, бидејќи зголемувањето на бројот на дрвата го намалува влијанието на шумот, односно го компензира прекумерно тренирање на поединечните дрва за одлучување. Исто така, поради големата длабочина на дрвото, т.е. комплексноста, решен е проблемот со недоволно добро толкување на врските помеѓу променливите во рамките на искористеното множество на податоци (недоволно тренирање). Високата отпорност на недоволно тренирање и прекумерно тренирање е една од најголемите предности на алгоритамот за случајни шуми.

Овој алгоритам наоѓа примена во бројни апликации, како што се избирачи на карактеристики, системи за препораки и класификатори на слики. Некои од реалните апликации вклучуваат откривање измама, класификација на апликации за заем и предвидување на болести. Овој алгоритам претставува основа за алгоритамот борута (Boruta), кој избира витални карактеристики во базата на податоци.

Принципот на работа на алгоритамот борута е следен: се претпоставува дека базата на податоци има m карактеристики, случајната шума по случаен избор ќе избере k од нив, каде k<m. Следно, алгоритамот ќе го пресмета коренскиот јазол помеѓу k-те одличја со избирање на јазол што има најголема информациска добивка. По тоа, алгоритамот го дели јазолот во јазли деца и го повторува овој процес n пати. Вредноста на k се одржува константна за време на растењето на шумата. Секое дрво расти во најголема можна мера. Нема кастрење на дрвата. Се добива шума со дрва. Следно, се комбинираат резултатите од сите дрва на одлука присутни во шумата. Тоа е, секако, еден од најсофистицираните алгоритми бидејќи се базира на функционалноста на дрвата на одлука. Технички, тој е ансамбл-алгоритам. Алгоритамот генерира индивидуални дрва на одлучување преку индикација за избор на атрибут. Секое дрво се базира на независен случаен примерок. Во проблем со класификација, секое дрво гласа и најпопуларната класа е крајниот резултат. Од друга страна, во проблем со регресија, се пресметува просекот на сите излези на дрвата и тоа ќе биде краен резултат [62].

Стапката на грешка во шумите зависи од две работи [119]:

- Корелацијата помеѓу кои било две дрва во шумата. Зголемувањето на корелацијата ја зголемува стапката на грешка во шумите.
- Јачината на секое поединечно дрво во шумата. Дрвото со мала стапка на грешка е силен класификатор. Зголемувањето на јачината на поединечните дрва ја намалува стапката на грешка во шумите.

Намалувањето на k ја намалува и корелацијата и јачината. Со негово зголемување се зголемуваат и двете. Некаде помеѓу е „оптимален“ опсег за k - обично доста широк. Користејќи ја стапката на грешка надвор од торбата (out of bag error), може брзо да се најде вредноста за k во опсегот. Ова е единствениот прилагодлив параметар на кој случајните шуми се донекаде чувствителни.

Карактеристики на алгоритамот случајни шуми [119]:

- ненадминат е по точност меѓу сегашните алгоритми;
- работи ефикасно со големи множества на податоци;
- може да се справи со илјадници влезни променливи без бришење на променливите;
- дава проценки за тоа кои променливи се важни во класификацијата;
- генерира внатрешна непристрасна проценка на грешката во генерализацијата како што напредува градбата на шумите;
- има ефективен метод за проценка на податоците што недостасуваат и ја одржува точноста кога недостасуваат голем дел од податоците;
- има методи за балансирање на грешки во неизбалансирани збирки на податоци од популацијата во класите;

- генерираните шуми може да се зачуваат за идна употреба на други податоци;
- се пресметуваат прототипови кои даваат информации за односот помеѓу променливите и класификацијата;
- ја пресметува близината помеѓу паровите случаи што може да се користат при групирање, лоцирање на оддалечени или (со скалирање) давање интересни прикази на податоците;
- способностите на горенаведеното може да се прошират на неозначени податоци, што доведува до ненадгледувано групирање, прегледи на податоци и откривање на оддалеченост и
- нуди експериментален метод за откривање на променливи интеракции.

Наивен Бајесов класификатор

Наивниот Бајесов класификатор се користи за класифицирање на објектите во класи според најголемата веројатност на припаѓање. Методот ја применува Бајесовата теорема поради што и го носи името Бајесов класификатор, а се нарекува наивен бидејќи претпоставува дека карактеристиките на објектот се независни една од друга. Иако оваа претпоставка за независноста на карактеристиките речиси секогаш е неточна, сепак, методот се покажува како доста ефективен во голем број реални и комплексни проблеми. Друга корист од оваа претпоставка е тоа што се поедноставуваат пресметките.

За дадена класна променлива y и зависно својство на зависност во форма на вектор x_1 до x_n, Бајесовата теорема е [63]:

$$P(y|x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n|y)}{P(x_1, \dots, x_n)}$$

Со користење на условена независна претпоставка дека

$$P(x_i|y, x_1, \dots x_{i-1}, x_{i+1}\ x_n) = P(x_i|y)$$

за сите i, оваа релација се поедноставува како:

$$P(y|x_1, \dots, x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x_1, \dots, x_n)}$$

Бидејќи $P(x_1, \dots, x_n)$ е константно со оглед на влезот, може да се користи следново правило за класификација [63]:

$P(y|x_1, \dots, x_n) \propto P(y)\prod_{i=1}^{n} P(x_i|y)$ следи

$$\hat{y} = \arg\underbrace{max}_{y} P(y)\prod\nolimits_{i=1}^{n} P(x_i|y)$$

Во овој израз може да се користи MAP (Maximum A Posteriori) проценка со цел оценување на $P(y)$ и $P(x_i|y)$; првата претпоставка е релативна фреквенција на класата y во тренирачкото множество.

И покрај едноставните претпоставки на овој алгоритам, класификаторите најдобро функционираат во реални ситуации, како што се класификација на документи и филтрирање на пораки со спам-содржина. Класификаторите најчесто работат со мало количество на тренирачки податоци за проценка на потребните параметри [63].

Невронски мрежи

Пронаоѓачот на првиот неврокомпјутер, д-р Роберт Нехт-Нилсен, невронската мрежа ја дефинира како пресметувачки систем изграден од голем број на едноставни, брзи, внатрешно поврзани елементи за процесирање, кои што процесираат информации со помош на нивните динамички состојби на одговор во однос на надворешен влез [65].

Вештачката невронска мрежа е алгоритам за учење, инспириран од структурата и функционалниот аспект на биолошка невронска мрежа. Пресметувањата се структурирани во однос на меѓуповрзани групи на вештачки неврони, кои се поделени по слоеви и секој слој врши одредена измена на влезните параметри. Постојат три поголеми групи на неврони: влезни неврони, скриени неврони и излезни неврони [65].

Модерните невронски мрежи се алатки за нелинеарно статистичко податочно моделирање. Тие најчесто се користат за моделирање на комплексни врски меѓу влезот и излезот, за изнаоѓање шеми во податоците и карактеризирање на статистичката структура за непозната заедничка веројатност на распределба на повеќе набљудувани променливи.

Постојат следните видови на вештачки невронски мрежи:

Перцептрон: Наједноставниот и најстариот модел на вештачка невронска мрежа, перцептрон, е линеарен класификатор што се користи за бинарни предвидувања [65].

Повеќеслојна вештачка невронска мрежа: Пософистицирана од перцептронот, повеќеслојната вештачка невронска мрежа (на пр.: конволуциска невронска мрежа, повторувачка невронска мрежа итн.) е способна да реши покомплексни задачи за класификација и регресија благодарение на нејзиниот скриен слој [65].

Функцијата за активирање на јазол во вештачка невронска мрежа го дефинира излезот на тој јазол во однос на влезот или збирот на влезовите. Најпознати функции за активирање на јазолот се: сигмоидната функција [66], функцијата tanh (хиперболичен тангенс) [66], функцијата softmax [67], додека пак најновите вештачки невронски мрежи користат корегирани линеарни единици за скриените слоеви (ReLU, Rectified Linear Unit) [67] и Leaky ReLU [67].

Невронските мрежи наоѓаат практична примена во: машинската перцепција, компјутерската визија, процесирањето на природните јазици, препознавањето на синтаксички форми, пребарувачите, медицинската дијагностика, биоинформатиката, интеракцијата мозок : машина, хемоинформатиката, препознавањето на лажни платежни картички, анализата на берзи, класификацијата на ДНК-секвенци, секвенциското рударење, препознавањето на говор и ракопис, препознавањето на форми во компјутерска визија, индустријата за игри, софтверското инженерство, динамичките адаптивни веб-страници, локомоторниот систем кај роботите, финансиските пресметки, набљудувањето на здравјето кај пациенти, сентименталната анализа.

3.2.2 Ненадгледувано учење

Во ненадгледуваното учење, од системите се очекува да учат правилно од дадени влезови [70]. Системот треба многу добро да ги истражи влезните податоци, да ги идентификува моделите и да произведе излез од некој вид. Во методите на учењето без надзор спаѓаат групирањето, откривањето на нови вредности, смалување на димензионалноста, откривање на врски меѓу примероците и др.

Задачата за групирање е да се поделат примероците на голем број класи врз основа на меѓусебната сличност. Во зависност од тоа што го знаеме за податоците, задачата за групирање може да биде да ги поделиме примероците во однапред дефинирани класи, или може да биде да го одредиме потребниот број на класи и кој примерок во која класа припаѓа. За да може да се групираат примероците, потребно е да може да го пресметаме нивното растојание, односно потребни се одредени метрички вредности.

Задачата на постапките за намалување на димензионалноста е да се намали бројот на атрибути со кои се опишуваат примероците. Едноставен пример: постои збир на примероци што се точки во тродимензионален простор и сите лежат во иста рамнина. Бидејќи точките лежат во иста рамнина може да се замисли да се нацрта нов координатен систем во таа рамнина: тоа би бил дводимензионален координатен систем и секој модел би можел да се опише користејќи два броја, претставувајќи точки во дводимензионален простор.

Овој процес на учење работи добро на податоците за трансакции.

3.2.3 Принудно учење

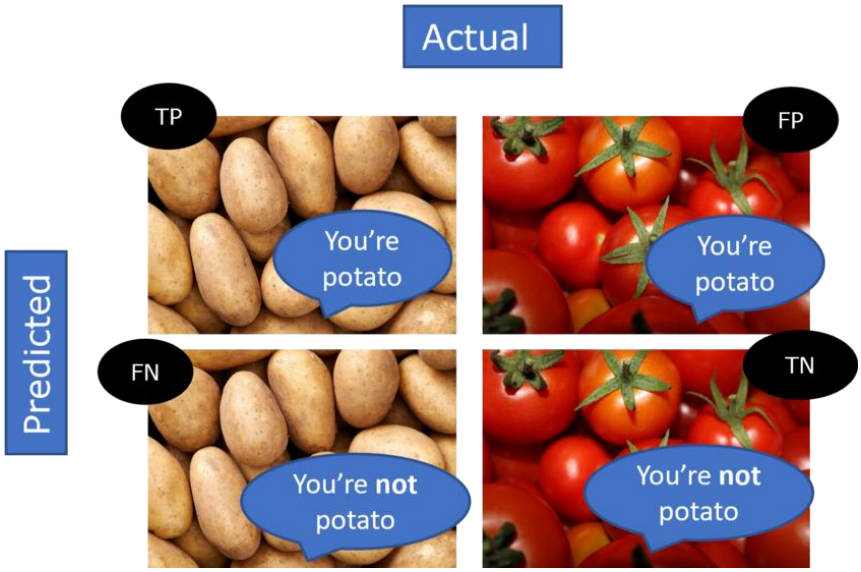
Принудното учење најчесто се користи во апликации за игри каде што се доделуваат награди или казни на агент врз основа на неговите постапки. Агентот извршува дејства засновани на информации од околината, и како одговор на секое дејство од околината добива награда или казна. Задачата на принудното учење е да се развие „системот за управување“ на агентот, односно да се открие оптималната стратегија на неговото однесување, така што агентот ги максимизира наградите што ги

добива „на долг рок“ [70]. Принудното учење обезбедува и квалитативни и квантитативни рамки за разбирање и моделирање на адаптивното одлучување во услови на награди и казни. Ова учење може да се примени во широк спектар на задачи, како на пример [58]: програма што контролира робот што оди, програма што го контролира движењето на Pac-Man во играта на Atari, програма што ја игра популарната игра Go, програма што ги набљудува цените на берза и одлучува колку да се купи или продаде и др.

4 Модел со кој се предвидуваат временските серии на реализирана променливост на пазарната цена на биткоинот

4.1 Оценување на моделот за класификација

Откако ќе се изгради модел за предвидување, од клучно значење е да се оцени моделот за да се разбере колку се добри предвидувањата и како може дополнително да се подобрат. За оценување на моделите за класификација, има неколку метрики, како што се точност, специфичност, чувствителност, прецизност итн. За да се објаснат овие метрики, прво треба да се разберат следните термини користејќи ја сликата 22. Во оваа илустрација, нека компирот одговара на позитивна класа, а доматот одговара на негативна класа [76].



Слика 22. Пример за компир : домат за илустрација на TP/FP/TN/FN [76].

Точно позитивно (*TP*) – ако се предвидува дека излезот ќе биде позитивна класа и тој е позитивен. На сликата, се предвидува дека класата ќе биде компир и навистина е компир. Лажно позитивно (*FP*) – ако се предвидува дека излезот ќе биде позитивна класа, но, всушност, тоа е негативна класа. На сликата, се предвидува дека класата ќе биде компир, но тоа е домат. Во статистиката, ова се нарекува и грешка од тип I. Лажно негативно (*FN*) – ако се предвидува дека излезот ќе биде негативна класа, но, всушност, тоа е позитивна класа. На сликата, се предвидува дека класата не е компир, а е компир. Во статистиката, ова се нарекува и грешка од тип II. Точно негативно (*TN)* – ако се предвидува дека излезот ќе биде негативна класа и тој е, всушност, негативен. На сликата, се предвидува дека класата не е компир и навистина не е компир.

Ако горната слика се замени со матрица од броеви, се добива т.н. конфузиона матрица (confusion matrix).

Точноста мери колку често класификаторот правилно предвидува (позитивно или негативно). Дефиниција за точност:

$$Tochnost = \frac{TP + TN}{TP + FP + TN + FN}$$

Модел со поголема точност не мора да значи модел со добри перформанси. Пример сценарио каде што постои множество податоци со 95 компири и 5 домати, а направени се 100 набљудувања. Во овој случај, моделот најверојатно ќе предвиди дека излезот ќе биде 100 % компир. Ова би резултирало со точност од 95 %, но моделот воопшто не прави добра работа со предвидувањето на доматите. Ова е причината поради која треба да се испитаат и други метрики.

Прецизноста мери колку добро моделот ги предвидува случаите што припаѓаат на позитивната класа. Поточно, прецизноста мери колкав дел од позитивните страни моделот ги предвидува правилно. Прецизноста е корисна во случаи кога FP е повисок од *FN*. Оваа метрика е од суштинско значење во областите како препораки за видео/музика, веб-локации за е-трговија итн., каде што помала прецизност ќе ги наведе клиентите да се префрлат на производите или услугите на конкурентите.

$$Preciznost = \frac{TP}{TP + FP}$$

Чувствителноста (осетливоста), позната и како „*Recall*“ или „Sensitivity“, мери колку позитивни случаи моделот можел правилно да предвиди. Ова се користи кога *FN* е повисоко од *FP*. Оваа метрика е критична во медицинскиот домен каде што и покрај лажните аларми, важно е позитивните случаи да не останат неоткриени.

$$Chuvstvitelnost = \frac{TP}{TP + FN}$$

Специфичноста е како чувствителноста, но за негативна класа. Во медицинска смисла, ја мери способноста на моделот да предвиди колку пациенти биле без болест.

$$Specifichnost = \frac{TN}{TN + FP}$$

F1 резултат: Тоа е комбинација (хармонична средина) на прецизност и чувствителност (Recall). Тоа е многу корисно кога *FP* и *FN* се приближно поедначкви и кога *TN* е многу висок (на пр. повеќе здрави индивидуи отколку пациенти).

$$F1 = 2\frac{Preciznost \cdot Recall}{Preciznost + Recall}$$

4.1.1 ROC крива и AUC-вредност

Receiver Operator Characteristic (ROC) кривите се популарен начин за визуелизација на компромисот меѓу чувствителноста и специфичноста кај бинарните класификатори. Овие криви можат да се објаснат со едноставен пример преку „поглед од желка“ (turtle's eye view): класификатор се користи за сортирање случаи по редослед од најголема до најмала веројатност да бидат позитивни, а желка слична на лого маршира по оваа низа случаи. Желката смета дека сите случаи што ги поминала биле позитивни. Во зависност од нивната вистинска класа, тие се или лажно позитивни (FP) или точно позитивни (TP); ова е еквивалентно на прилагодување на прагот на резултат. Кога желката поминува низ TP, таа прави чекор нагоре по y-оската, а кога поминува FP, прави чекор надесно на x-оската. Големините на чекорите се обратно пропорционални со бројот на точно позитивни (во насока y) или негативни (во насока x), така што патеката секогаш завршува на координатите (1, 1). Резултатот е график на вистинска позитивна стапка (TPR, или чувствителност) наспроти лажно позитивна стапка (FPR, или 1-специфичност), што е сè што е ROC-кривата. Пресметувањето на површината под кривата (AUC, Area under the ROC Curve) е еден начин да се сумира во една вредност. Еден начин на толкување на AUC е како веројатност дека моделот повисоко рангира случаен позитивен пример отколку случаен негативен пример [84].

4.1.2 Прекумерно тренирање (overfitting)

Добриот модел може да ја научи шемата од податоците за обука и потоа да ја генерализира на нови податоци (од слична дистрибуција). Прекумерно вклопување е кога моделот е способен речиси совршено да ги собере вашите податоци за обука, но слабо работи на новите податоци. Моделот претерано ќе се вклопи кога ја учи многу специфичната шема и шумот од податоците за обуката. Овој модел не може да ја извлече „големата слика“ ниту општата шема од податоците. Оттука, според новите и различни податоци, перформансите на пренаменетиот модел ќе бидат слаби.

Прекумерното вклопување се случува кога моделот има премногу слобода да ги собере податоците. Потоа, лесно е моделот совршено да ги собере податоците за обуката (и да ја минимизира функцијата за загуба). Оттука, посложените модели се со поголема веројатност да имаат прекумерно вклопување. На пример, линеарна регресија со разумен број на променливи никогаш нема прекумерно да се вклопи. Моделот е едноставен и ограничен на линеарни односи меѓу променливите. Од друга страна, случајната шума или невронската мрежа лесно можат прекумерно да се вклопат. Тие имаат многу параметри што можат да ја минимизираат функцијата на загуба.

За да се открие прекумерно вклопување, треба да се види како еволуира грешката на тестот. Сè додека грешката на тестот се намалува, моделот е добар. Од друга страна, зголемувањето на грешката во тестот покажува дека веројатно има прекумерно вклопување.

За да се избегне прекумерно вклопување, се додаваат повеќе ограничувања на моделот:

- вкрстена валидација ги проценува перформансите на моделот на независно множество на податоци;
- ласо и регулација на гребенот додаваат казна за параметрите кои се големи или премногубројни;
- раното запирање го запира моделот кога грешката на тестот ќе почне да расте;
- напуштање, додавање шум на влезот [77].

4.1.3 Недоволно тренирање (underfitting)

Кога моделот не ги научил добро обрасците во податоците за обука и не може добро да ги генерализира новите податоци, тоа е познато како недоволно вклопување. Моделот со недоволно вклопување има слаби перформанси на податоците за обука и ќе резултира со несигурни предвидувања. Недоволното вклопување се јавува поради голема пристрасност и мала варијанса. Причини за недоволно тренирање [82]:

- податоците што се користат за обука не се чистат и во нив има шум;
- моделот има голема пристрасност;
- големината на користената база на податоци за обука не е доволна;
- моделот е премногу едноставен.

Начини за справување со недоволно вклопување:

- зголемување на бројот на карактеристики во базата на податоци;
- зголемување на сложеноста на моделот;
- намалување на шумот во податоците;
- зголемување на времетраењето на обуката на податоците.

4.2 Развој на вистинскиот модел на регресија

Во докторскава дисертација се прави предвидување на остварената променливост на биткоинот. Остварената променливост е мерење на променливоста на приносот за инвестициски производ со испитување на неговите историски вредности во одреден временски период. Се добива од реализираната варијанса, а воведена е од Барндорф-Нилсен и Шефард (2002). Евалуација на степенот на несигурност и/или можна финансиска загуба/добивка од инвестицијата во бизнис може да се пресмета со користење на нестабилност/променливост во цените на акциите на ентитетот. Најчестиот метод за проценка на променливоста во статистиката е со пресметување на стандардното отстапување, т.е. варијација во вредностите од средната вредност.

Секојдневните податоци со висока промена се користат од аналитичарите за да се проценат нивоата на променливост на час/ден/неделно/месечно. Податоците потоа може да се користат за да се процени променливото движење на продажбата. При

анализата, земени се податоци чија фреквенција е 1 час од платформата Џемини (Gemini) [73] и потоа, користејќи ги тие податоци, се пресметува остварената променливост со дневна фреквенција. На берзата Џемини се следат и се креираат датотеки за дневни, часовни и минутни податоци за цените на временските серии за физичкиот пазар за парови од американски долар (УСД) и најпопуларните криптовалути, како што се: биткоин, етериум, лајткоин и др. Секоја датотека може да се преземе во CSV-формат. Во секоја датотека, постојат податоци за цените на OHLC (Open/High/Low/Close) кои се ажурираат на дневна основа. За потребите на овој труд земени се грануларни часовни податоци враќајќи се до 2015 година за биткоин/долар и податоците ги имаат следните полиња:

- временска ознака во Unix-формат (Unix time или Epoch time – број на секунди од 00:00 часот на 1 јануари 1970 год.),
- датум – овој временски печат е временска зона UTC,
- симбол – симболот за кој се однесуваат податоците од временските серии,
- отворено – ова е почетната цена во временскиот период,
- висока – ова е највисоката цена во временскиот период,
- ниска – ова е најниската цена во временскиот период,
- затвори – ова е цената на затворањето во временскиот период,
- волумен - за БТК/УСД, ова е во износ на биткоини.

Проценката на променливоста се пресметува со мерење на стандардното отстапување од просечната цена на набљудуваниот објект во даден временски период. Бидејќи променливоста е нелинеарна, реализираната варијанса прво се мери со преведување на вредностите земени од берза во логаритамски вредности и потоа се пресметува стандардното отстапување (стандардна девијација). Варијансите во дневните вредности се пресметуваат како што е прикажано подолу:

$$r_t = log(P_t) - log(P_{t-1}) \tag{4.1}$$

каде што P е пазарната цена на биткоин, а t е временскиот период.

Во следниот чекор, реализираната варијанса се пресметува со пресметување на збирот на квадратите од стандардната девијација:

$$Realized\ Variance_t = \sum_{i=1}^{n} r_t^2 \tag{4.2}$$

каде n е бројот на набљудувања, кој во овој случај е 24 (број на часови во еден ден), а t е временскиот период (се пресметува за секој ден).

Следен чекор е да се пресмета реализираната променливост, која е квадратен корен од реализираната варијанса.

$$Realized\ Volatility_t = \sqrt{\sum_{i=1}^{n} r_t^2} \tag{4.3}$$

За да се пресмета остварената променливост на биткоин, креирана е апликација во програмскиот јазик R [1], во која најпрво се вчитуваат часовните податоци за пазарната цена на биткоин преземени од платформата Џемини, користејќи ја функцијата *read.csv*(). Во следниот чекор се подредуваат податоци користејќи ја функцијата *order()*, така што првото набљудување е најстаро и последното набљудување е најново.

Потоа е додадена секвенца за датум користејќи ја функцијата seq, дефинирајќи ја почетната и крајната временска точка со фреквенцијата од 1 час. За почетен датум во временската низа е земен „08-10-2015 13:00:00“, а за краен датум е земен „12-01-2022 12:00:00“. При пресметките се зема цената на биткоинот при затворање, а другите податоци се изоставени. Се пресметуваат логаритамските вредности на цената на биткоин за резултатите да бидат попрецизни, а, исто така, се креира вектор полн со NA-вредности со иста должина како и податоците за пазарна цена на биткоинот, користејќи ја функцијата *rep()*. Тој вектор се пополнува со стандардното отстапување во секој час прикажано во равенка (4.1). Потоа се пресметуваат квадратите на стандардното отстапување во нов вектор и се спојува со другите податоци. Користејќи ја библиотеката *library(dplyr)* се создава нова колона која има единствен датум за секој ден. Тоа е колона за идентификација, па подоцна, користејќи ги функциите *group_by, sum и arrange*, се сумираат квадратите на стандардното отстапување што одговараат на еден датум за еден ден, така што ја добиваме реализираната варијанса (4.2). Користејќи ја функцијата *sqrt()* се пресметува остварената променливост во дневна фреквенција (4.3). За потребите на истражувањето преземени се слободни податоци од веб-страницата https://www.blockchain.com/charts во *.csv* format за последните 3 години за карактеристиките прикажани во табела 4.

Сите податоци за карактеристиките со командата *read.csv()* се вчитани во посебни променливи во програмскиот пакет R. Потоа, сите податоци се вчитани во табела со податоци со командата *data.frame()*. Во тие податоци додадена е и променливоста на биткоинот што претходно е пресметана со податоци преземени од платформата за менување на криптовалути www.gemini.com.

За да може програмскиот јазик *R* да го препознае временскиот период, додадена е и променлива Датум. Потоа, користејќи функција *write.csv()*, се создава нова *.csv*-датотека која ги содржи сите карактеристики преземени од blockchain.com, реализираната променливост на биткоин и променлива Датум.

Откако е комплетирана базата на податоци, за да се започне со анализа на податоците потребно е да се направат некои трансформации за да се балансираат податоците. За таа цел треба да се примени логаритам во секоја зависна карактеристика на биткоинот. Со оваа трансформација секој вектор *x* се заменува со

[1] За потребната анализа користен е пакетот RStudio Version 1.4.1717 за Edubuntu.

неговиот природен логаритам со функцијата $\log(x)$. Оваа трансформација се прави за да може резултатите од статистичката анализа да станат попрецизни.

Табела 4. Карактеристики на биткоинот преземени од https://www.blockchain.com/charts.

Променлива	**Опис**
cost_per_transaction	Приходите на рударите поделени со бројот на трансакции.
cost_per_transaction_percent	Приходите на рударите како процент од обемот на трансакцијата.
estimated_transaction_volume_usd	Вкупната проценета вредност во американски долари на трансакциите на блоковските вериги.
miners_revenue	Вкупната вредност во УСД на наградите за блок и таксите за трансакции што им се платени на рударите.
n_transactions	Вкупниот број на потврдени трансакции дневно.
n_transactions_per_block	Просечниот број на трансакции по блок во изминатите 24 часа.
output_volume	Вкупната вредност на сите излезни трансакции дневно.
trade_volume	Вкупната вредност во американски долари на обемот на тргување на главните берзи за биткоини.
transaction_fees_usd	Вкупната вредност во американски долари на сите такси за трансакции што им се платени на рударите.

За да се подготват податоците за машинско учење се врши нормализација на податоците. Целта на нормализацијата е да се прилагодат вредностите на нумеричките колони во база на податоци до стандардна скала, без искривување на варијациите во опсегот на вредности. Нормализација на податоците е потребна само кога карактеристиките во една база на податоци имаат различни опсези. Постојат различни начини за нормализирање на податоците. Избран е методот на нормализација Min-Max. Податоците се нормализираат за да имаат вредности меѓу 0 и 1. Формулата на min-max е дадена како:

$$x_i' = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

Кога имаме некои карактеристики кои имаат многу различни опсези, многу е вообичаено карактеристиките со повисоки опсези да имаат тенденција суштински да влијаат на резултатот повеќе поради нивната поголема вредност, но тоа не значи дека оваа карактеристика е поважна од другите, и затоа е потребно да се направи нормализација на податоците.

За да се предвиди остварената променливост на биткоин се користи алгоритам за машинско учење. Избран е алгоритамот случајни шуми, којшто е вклучен во методите за учење на ансамблот. Учењето на ансамблот е тип на техника на учење под надзор каде основниот концепт е да се генерираат неколку модели за обука и потоа едноставно да се комбинираат нивните излезни правила или нивната H_x хипотеза, да се конструира силен модел кој работи многу добро, не се тренира прекумерно и, исто така, ја балансира пристрасноста и варијансата Bias-Variance Tradeoff.

Алгоритамот случајни шуми е алгоритам за надгледувано учење кој се користи и за регресија и за класификација. Случајната шума е метод за учење што работи со конструирање на повеќе дрва на одлуки на примероци од податоци. Случајна шума е класификатор што се состои од збирка независни дрва за одлучување, секое дрво претставува еден глас во мнозинството одлуки. Овој метод создава голем број на дрва за одлучување за време на обуката. Од креираното множество на дрва за одлучување, методот на случајни шуми добива предвидување од секое дрво поединечно и на крајот го избира најдоброто решение со гласање. Целта на овој метод е да се намали варијансата на конечниот модел. Предноста на овој метод е во тоа што не бара интервалидација, и случајните шуми не се прилагодуваат на податоците за учење.

Идејата е дека наместо да се создадава единствен компликуван и сложен модел, кој би можел да има голема варијанса што доведува до прекумерно вклопување, или да биде премногу едноставен и да има голема пристрасност, што води до недоволно вклопување, може да се генерираат многу модели за обука во тренинг-множеството и на крајот да се комбинираат. Таквата техника е случајна шума, која е вообичаена техника за здружување што се користи за подобрување на предвидливиот резултат на дрвата на одлучување преку нивно просечно мерење за да се намали варијансата во дрвата. Дрвата за одлучување се сметаат за многу јасни и лесни за интерпретација и разбирливи техники за моделирање. Но, главниот недостаток кај нив е тоа што имаат ниска предиктивна ефикасност и лоша генерализација на тест-множеството. Тие се познати и како слаби ученици, кои сè уште имаат помалку шанси и имаат грешка помала од 50 %. Исто така, голема разлика е тоа што сметаме само случајно подмножество на m предвидувачи секогаш кога правиме поделба на тренинг и тест-множество. Бројот на случајно избрани променливи за креирање на секое дрво е главниот параметар за нагодување во случајните шуми. Може да се менува бројот на

дрва, но обично не е важно за предвидувањето [75]. Исклучувањето на некои од предвидувачите сега изгледа како лош избор, но има повеќе смисла бидејќи резултатот од тоа би бил дека секое дрво користи различни предвидувачи за да ги подели податоците во различни времиња. Ова имплицира дека 2 дрва генерирани на исти податоци за обука ќе имаат случајно различни променливи избрани при секоја поделба, така што дрвата ќе бидат декорелирани и независни едно од друго. Заклучниот резултат од моделот на ансамблот се одредува со пребројување на мнозинството гласови од сите дрва за одлука. Овој концепт е познат како полнење на торби (bagging). Бидејќи секое дрво за одлука зема различно множество на податоци за обука како влез, отстапувањата во оригиналната база на податоци за обука не влијаат на конечниот резултат добиен од собирањето на одлуките од секое дрво. Затоа, полнењето на торби како концепт ја намалува варијансата без да ја менува пристрасноста на комплетниот ансамбл.

Случајните шуми се засноваат на едноставна идеја: „мудроста на толпата“. Агрегат на резултатите од повеќе предвидувачи дава подобро предвидување од најдобриот индивидуален предвидувач. Група на предиктори се нарекува ансамбл. Затоа, оваа техника се нарекува учење на ансамблот. За да се подобри техниката на бинарни предвидувања со дрвата на одлуки, може да се обучи група класификатори, секој на различно случајно подмножество од множеството за тренирање. За да се направи предикција, се земаат предикциите на сите поединечни дрва, а потоа се предвидува класата што ќе добие најмногу гласови [74].

Клучни придобивки од користењето на случајна шума се:

- лесно користење,
- ефикасност,
- точност,
- разновидност.

Може да се користи за класификација или регресија. Поприфатлива е за почетници, споредено со слично точни алгоритми како невронските мрежи.

Секоја алатка има свои недостатоци. Бидејќи случајната шума користи многу дрва за одлучување, може да бара многу меморија за поголеми проекти. Затоа може да се случи овој алгоритам да биде побавен од некои други поефикасни алгоритми. Понекогаш, бидејќи ова е метод заснован на дрво на одлуки и дрвата за одлучување често страдаат од преоптоварување, овој проблем може да влијае на целокупната шума. Овој проблем обично стандардно го спречува случајната шума, бидејќи користи случајни подмножества од карактеристиките и гради помали дрва со тие подмножества. Ова може да ја забави брзината на обработка, но да ја зголеми точноста.

Случајната шума била предложена од Хо (1995) [72]. Катлер и Брејман развиле продолжување на алгоритамот кој комбинира торбичка и случаен избор на

карактеристики за да се изгради колекција од дрва за одлучување со што се намалува варијансата [106]. Кога се користи истото множество на предвидувачи за да се создаде секое од дрвата, многу од дрвата ќе бидат премногу слични (или исти), што ќе резултира со ограничено пребарување на просторот за предвидување. Комбинирањето во торби ја намалува варијансата на поединечни дрва, но пристрасноста е приближно иста бидејќи се користат истите дрва. Дрвата од комбинирањето во торби обично имаат високи корелации. Случајната шума како алгоритам била предложена за да се задржи малата варијанса при користење на метод на торба, но, дополнително, да се намали пристрасноста со декорелација на дрвата. Ова се постигнува со случаен избор на подмножество од променливите за креирање на секое дрво [75]. Спротивно на функционалноста на случајни шуми, базирана на Џанг и Ма (2012), за случаен вектор $X = (X1, \dots, Xp)^T$, чија димензија е p и ги претставува реалните вреднувани влезни или предвидувачки променливи и случајна променлива Y која го претставува вистинскиот вреднуван одговор, претпоставуваме $P_{XY}(X, Y)$ непозната заедничка дистрибуција. Целта е да се најде функција за предвидување *f(X)* за да се предвиди *Y*. Функцијата за предвидување е специфицирана со функција на загуба $L\big(Y, f(x)\big)$ и поставена да ја намали вредноста на очекуваната загуба [71]:

$$E_{XY}\left(L\big(Y, f(X)\big)\right)$$

каде што долните индекси *x* и *y* укажуваат на очекувањата со почит кон заедничката распределба на X и Y. Функцијата на загуба, $L\big(Y, f(X)\big)$ е начин да се измери колку е блиску *f(X)* до *Y*. Се казнуваат вредности на *f(X)* кои се далеку од *Y*. За регресија класичниот избор на *L* се квадратни загуби на грешка:

$$L\big(Y, f(X)\big) = \big(Y - f(X)\big)^2$$

Резултатот е дека за квадратна загуба на грешка, минимизирањето на $E_{XY}\left(L\big(Y, f(X)\big)\right)$ го дава хипотетичкото очекување:

$$f(X) = E(Y|X = x)$$

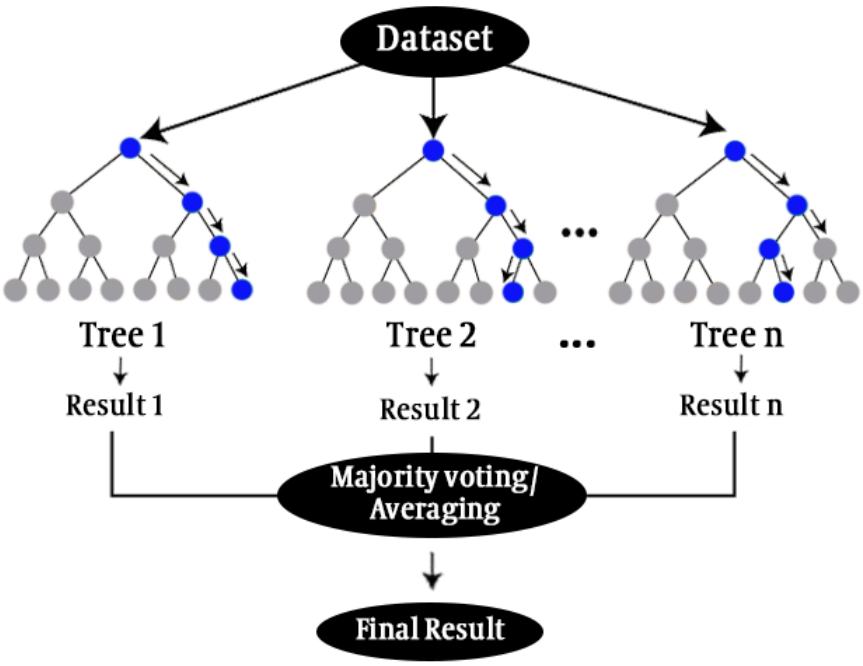
позната и како регресивна функција.

Во збирката на базни ученици *h1(x), ..., hJ(x),* ансамблот f е конструиран комбинирајќи ги овие базни ученици, така што како резултат се добива ансамблот предвидувач *f(X)*. Во случај на регресија, се зема просек од овие ученици [75]:

$$f(x) = \frac{1}{J}\sum_{j=1}^{J} h_j(x)$$

каде што *J* е бројот на базни ученици што се користат во регресионата функција *f(X)*. Секој основен ученик е дрво. Притоа, *j*-тото дрво е означено како *hj(X*, Φ_j*)*, каде

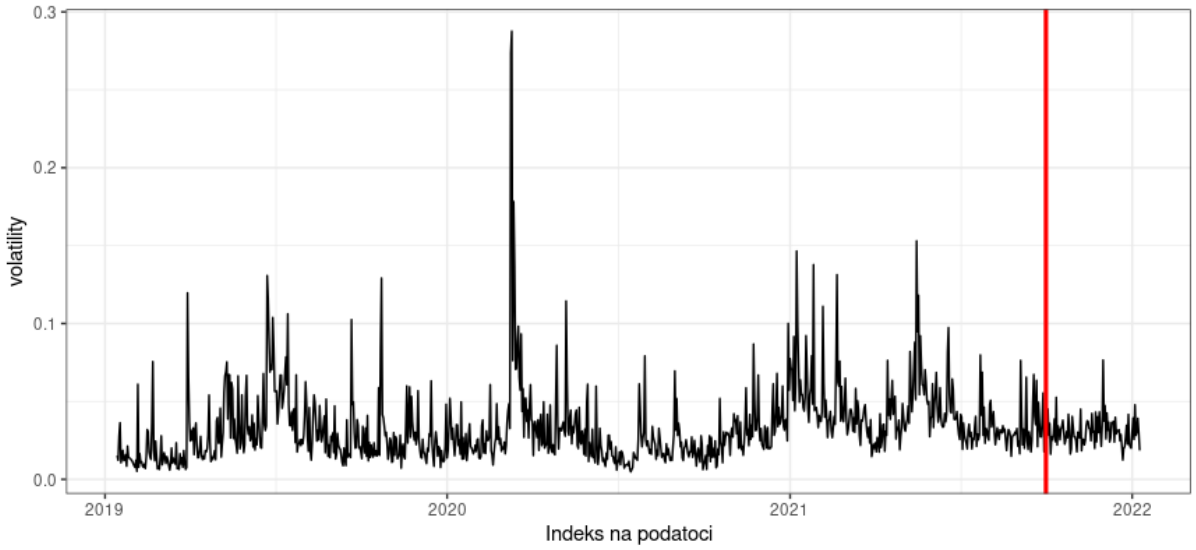
што Φ_j е збирка од случајни променливи - карактеристики кои се независни за *j = 1, ... J*, прикажано на слика 23.



Слика 23. Поедноставена случајна шума [71].

Во алгоритамот за предвидување на променливоста на биткоинот се користи пакетот *forecastML* во програмскиот јазик *R*. Овој пакет обезбедува соодветни функции, бројки и визуелизација на предвидувањата на временските серии, користејќи алгоритми за машинско учење. За анализата се користи случајната шума, како алгоритам за машинско учење.

При користење на алгоритмите за машинско учење, прво се генерира моделот користејќи податоци за обука (тренирање), а потоа се предвидуваат вредности за податоците од множеството за тестирање. За да се направат предвидувања, се користи функцијата *predict()*. По извршеното претпроцесирање на податоците (пронаоѓање и пополнување на отсутни вредности, нормализација и логаритмирање), следен чекор во истражувањето е поделбата на податоците во множество за тренирање и множество за тестирање. Податоците што се користат за да се предвиди променливоста на биткоинот содржат 1095 набљудувања, почнувајќи од 14.01.2019 до 12.01.2022 година. За множеството за тренирање се земаат првите 995 набљудувања, а останатите 100 набљудувања се користат како тест множество, што е прикажано на графикот кој е генериран од програмата во програмскиот пакет *R* на слика 24.



Слика 24. Поделба на податоците во множество за тренирање и множество за тестирање генерирани од *R*.

Во методот на предвидување се користат три различни хоризонти за предвидување. Овие различни хоризонти се користат за да може да се предвидува за кратко и долго време, со цел да се комбинираат предвидувањата во конечната прогноза и на тој начин да се минимизира грешката. Потоа се дефинира функцијата randomForest со нејзините аргументи. Дел од вредностите што се обработуваат и кои претходно се собрани и пресметани се прикажани на слика 25.

Show 10 entries Search:

	volatility	cost.per.transaction	cost.per.transaction.percent	estimated.transaction.volume.usd	miners.revenue
1	0.0152567607713393	0.046778508992879	0.322541642513931	0.862495579164528	0.0593492053914017
2	0.0122673031198655	0.0384146193482597	0.103213590081294	0.895570138639151	0.0267646814609875
3	0.0277644058257748	0.0339371751511753	0.27417378825243	0.864815255002254	0.0199890670023295
4	0.0362739361102721	0.028156985945227	0.32536317321839	0.856386312238102	0.00839817407494214
5	0.0106804876649871	0.0466460062878096	0.388509729943156	0.851205227374965	0.0343155053014304
6	0.0143153418154273	0.0713737190716976	0.530328849235241	0.835314002537559	0.0371716805978449
7	0.0185998403219624	0.103767783503376	0.691559543413382	0.824080266513529	0.0819591803924355
8	0.0120142465964583	0.0557364330249999	0.39021745019011	0.854703361396329	0.0672828233892861
9	0.0123670504560558	0.0317402150660793	0.291909348975558	0.860636162103619	0.00536812987663187
10	0.0158477758391569	0.0759512827814911	0.346992963697467	0.858926934821584	0.0563252830749133

n.transactions	n.transactions.per.block	output.volume	trade.volume	transaction.fees.usd
0.630929753571458	0.968053814306652	0.198143960649361	0.0502014758039836	0.0631335132798605
0.630929753571458	0.974066575868498	0.293055509784198	0.0633472095158617	0.0794747789889414
0.630929753571458	0.974579845655724	0.212274347740891	0.0633472095158617	0.103370923126063
0.630929753571458	0.976826929287596	0.176718296758542	0.0502014758039836	0.0946757066007581
0.630929753571458	0.969597402194412	0.183767644160766	0.0502014758039836	0.0727604863701678
0.630929753571458	0.963516412991385	0.100906590470636	0.0502014758039836	0.0502768022180346
0.630929753571458	0.946309302831116	0.115237968411214	0.0502014758039836	0.00780403254665578
0.630929753571458	0.962183091066399	0.320990077762146	0.0502014758039836	0.0651738803858468
0.630929753571458	0.97329409073816	0.250853449454731	0.0502014758039836	0.0696549507769284
0.630929753571458	0.956055925430112	0.252705717805423	0.0502014758039836	0.0855749637891286

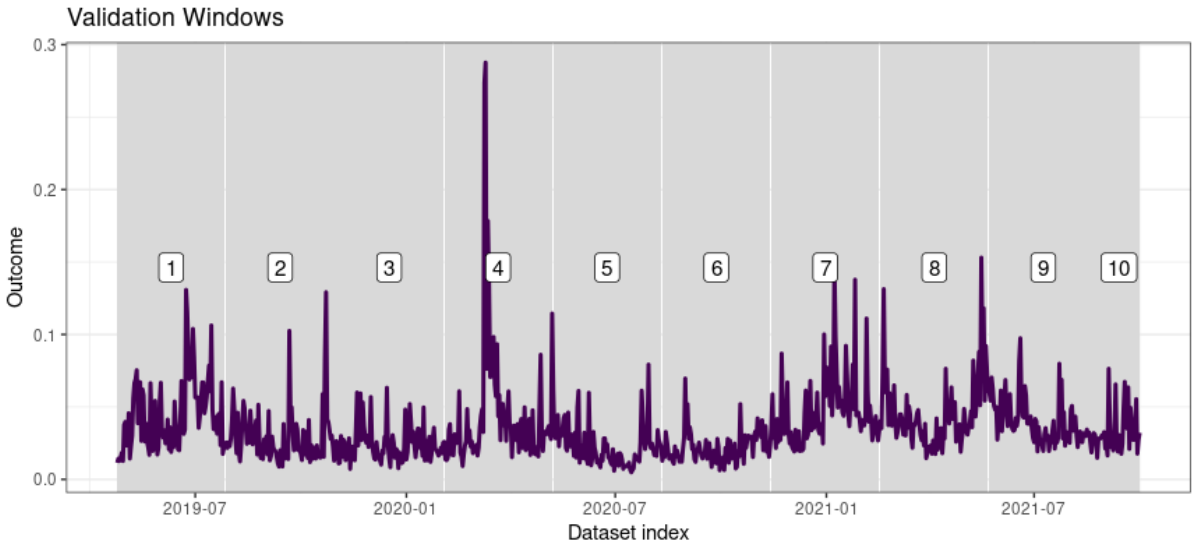
Слика 25. Дел од податоците од припремената база на податоци генерирани од *R*.

Првиот чекор во процесот на предвидување е да се создадат некои прозорци за валидација за да се изврши вгнездено поставување на вкрстена валидација. Следно, се обучува моделот и се презентираат предвидувањата, остатоците и некои метрики за грешки. Потоа се прогнозира во тест-множеството користејќи ги прозорците за валидација и се прикажуваат реалните наспроти предвидените вредности.

На почеток, се дефинира големината на секој хоризонт за предвидување. Хоризонт е аргумент во функцијата create_lagged_df() со која се создава модел на обука и множество на податоци за предвидување. Хоризонт претставува нумерички вектор на еден или повеќе прогнозирани хоризонти, што се мерат во редови на податоци. Ако се дадени датуми, хоризонтот со вредност 1 би бил еднаков на 1*фреквенција во календарско време [103].

Во моделот е избрано да се користат различни хоризонти за предвидување коишто на крајот се комбинираат за да зголеми ефикасноста при предвидувањето. Избрано е првиот хоризонт да биде 20 чекори напред, вториот да биде 50 чекори напред и последниот да биде 100 чекори напред. Се избираат хоризонтите што треба да се предвидат, а, исто така, се избира погледот низ одредени временски чекори во минатото.

Потоа, со помош на *create_lagged_df()*, заостанатите карактеристики се креираат врз основа на аргументот за поглед во минатото (lookback). Потоа, со функцијата *create_windows()*, се создаваат индекси за партиционирање (поделба) на базата на податоци за обука. Збирките на податоци за валидација се креирани во соседни блокови со должина на прозорец 95, за да имитираат предвидувања преку хоризонтите за прогноза во повеќе чекори, наспроти случајно избраните редици. Креирани се 10 прозорци за валидација, како што е претставено на слика 26:



Слика 26. Поделба на податоците во 10 прозорци за валидација генерирани од *R*.

Во овој вгнезден систем за вкрстена валидација, моделот се обучува со податоци од 9 прозорци и точноста на предвидувањето се оценува во десеттиот прозорецот. Тоа значи дека за секој хоризонт на директна прогноза, ќе треба да обучиме 10 модели, секој теоретски избира различни оптимални хиперпараметри и има различни коефициенти од внатрешниот процес на вкрстена валидација. Проценката на варијациите помеѓу овие модели е добар начин да се процени стабилноста под динамиката на различни временски серии во даден метод на моделирање.

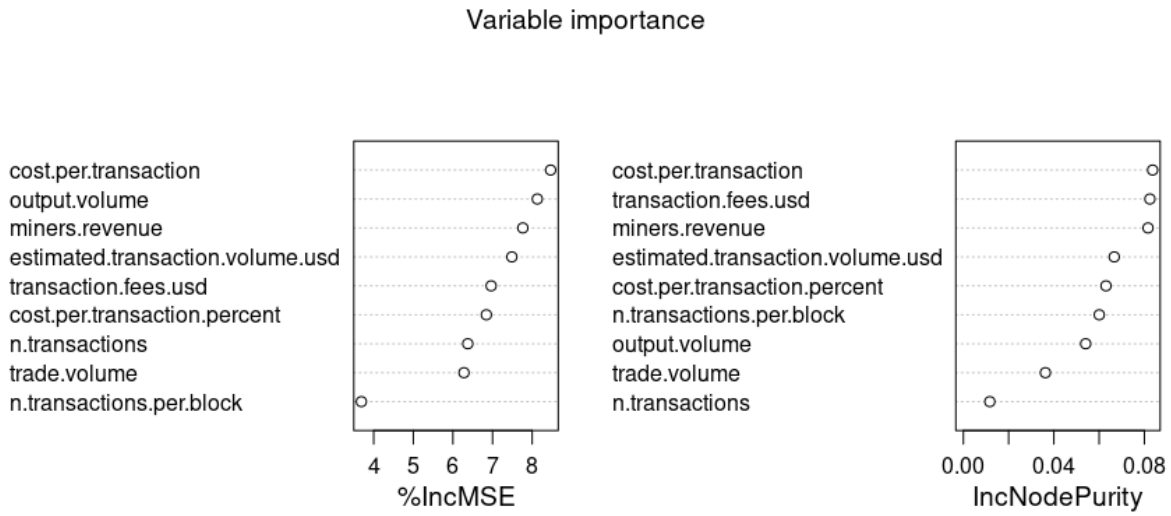
Потоа се дефинира функцијата *randomForest()*. Во аргументите се користи бројот на дрва и тој се поставува на 200. За почеток се поставуваат голем број дрва за да може да се одлучи за точниот број на дрва што ќе се користи во конечната функција и се користи важноста на променливите за да може да се разбере кои променливи влијаат на променливоста на биткоинот.

5 Резултати

Важноста на променливите што се користат во предвидувањето е претставена во табела 5 и на слика 35, каде *%IncMSE* означува процентуално зголемување на средната квадратна грешка, а *IncNodePurity* го означува зголемувањето на чистотата на јазолот. Процентуалното зголемување на средната квадратна грешка *(%IncMSE*) е најефективниот метод за идентификување важни променливи во алгоритамот случајни шуми. Поголема вредност на *%IncMSE* претставува поголема важност на променливта. Втората важна мерка, *IncNodePurity* се однесува на функцијата на загуба, која е избрана од најдобрите поделби. Функцијата на загуба е *MSE* за регресија и *Gini-impurity* за класификација [103].

Табела 5. Важност на променливите кои се користат во предвидувањето.

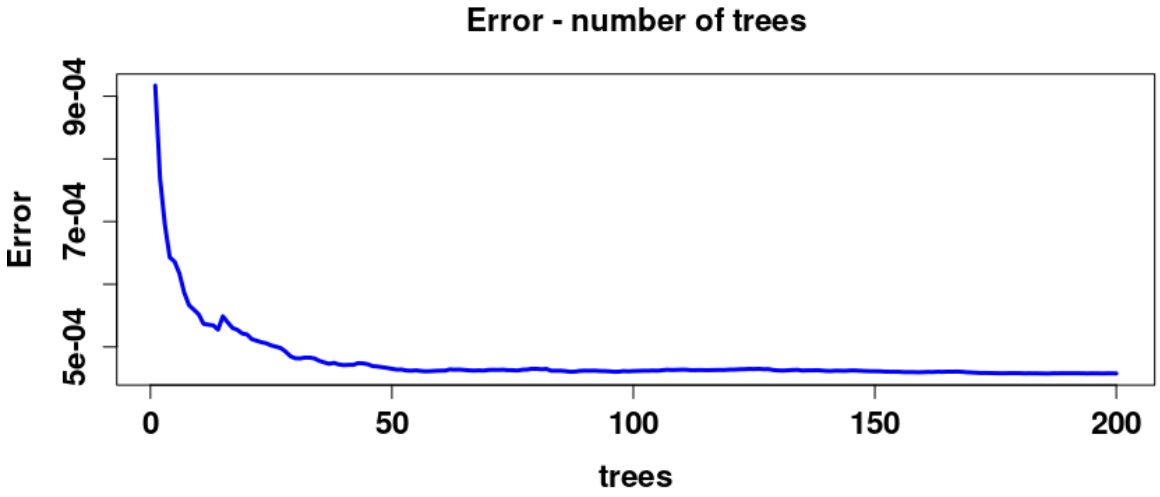
Променлива	%In	IncNode
cost_per_transaction	8.46	0.083652
cost_per_transaction_perce	6.84	0.063034
estimated.transaction.volu	7.48	0.066687
miners_revenue	7.76	0.081617
n_transactions	6.37	0.011682
n_transactions_per_block	3.68	0.060066
output_volume	8.12	0.054055
trade_volume	6.27	0.036256
transaction_fees_usd	6.96	0.082443



Слика 27. Важност на променливите генериран од *R*.

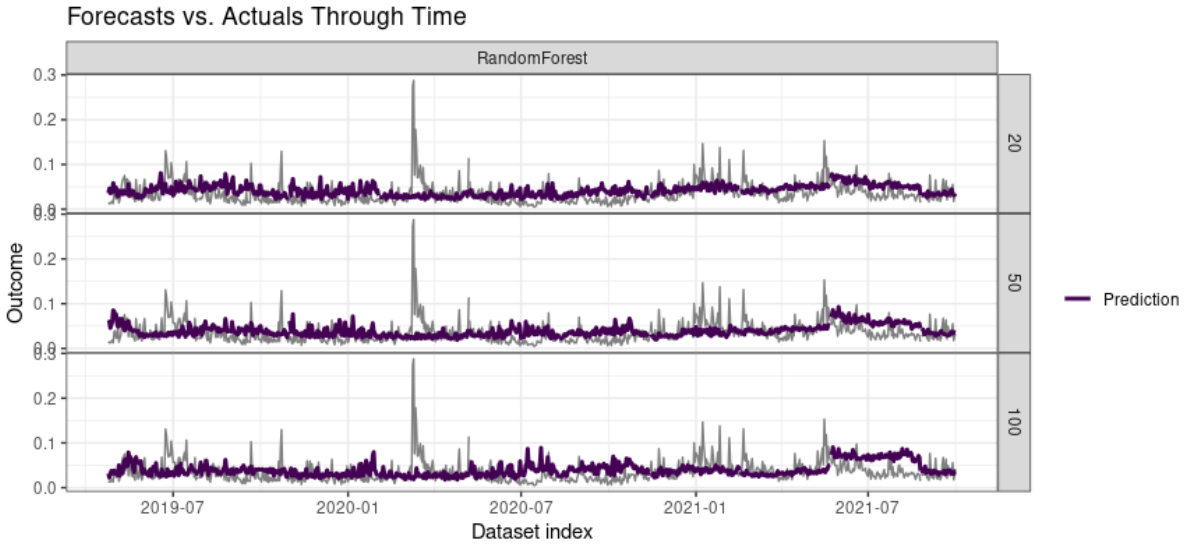
Од слика 27, може да се забележи дека од избраните карактеристики, најважна е cost.per.transaction кај процентуалното зголемување на средната квадратна грешка и кај зголемувањето на чистотата на јазолот.

На слика 28 е прикажана грешката на секој модел во зависност од бројот на дрва. Кривата на графикот е во форма на лакт. Со функцијата *which.min(model_imp$mse)* се пресметува дека најмала грешка се добива за 186 дрва [104]. Како што се забележува од графикот, на почетокот грешката опаѓа многу брзо како што бројот на дрва се зголемува, но по приближно 50 дрва намалувањето на грешката е речиси незначително. За да се избегне преоптоварување со користење на повеќе дрва се избира да се користат 50 дрва во даденото испитување.



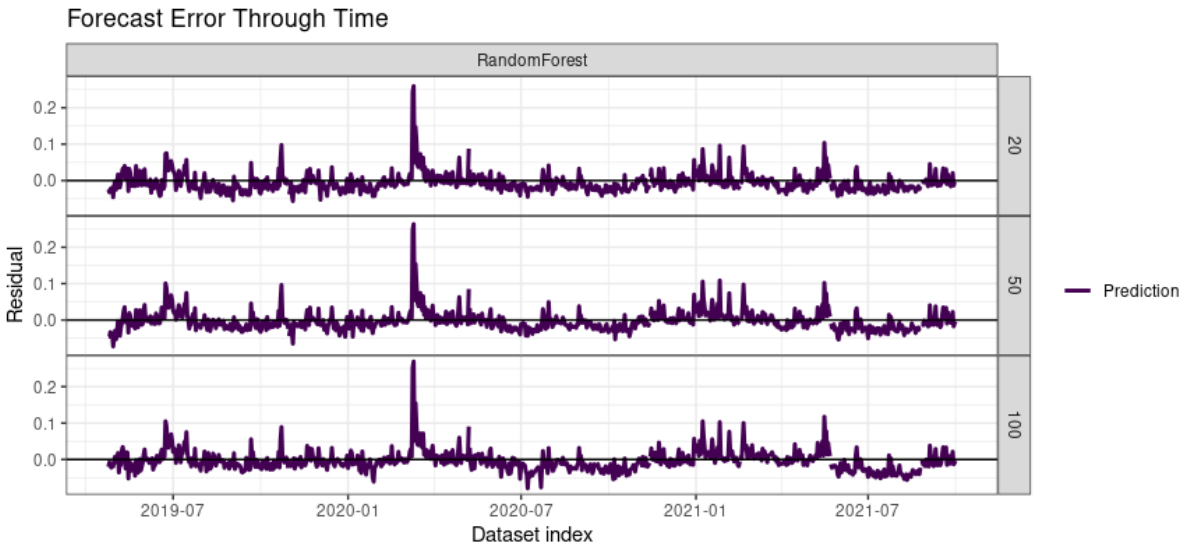
Слика 28. Графички приказ на грешката во зависност од бројот на дрва генерирани од *R* .

Следно, се користи функцијата *train_model()* со која ќе бидат обучени вкупен број од 30 модели (10 прозорци за валидација * 3 хоризонти на прогноза, бидејќи се користат прозорците за валидација). За секој предикциски хоризонт се добиваат 10 предвидувања. Потоа се дефинира функција за предвидување користејќи *predict*(). На слика 29 се претставени вгнездените прогнози за вкрстена валидација за секој прозорец за валидација и хоризонт на прогноза врз основа на моделот.



Слика 29.Вгнездените прогнози за вкрстена валидација за секој прозорец за валидација и за секој хоризонт генерирани од *R*.

На слика 30 претставени се остатоците од предвидените вредности.



Слика 30. Остатоците од предвидените вредности генерирани од *R*.

Од сликата може да се забележи дека во периоди со поголема променливост грешката и остатоците се поголеми.

Средната апсолутна грешка (*MAE, Mean Absolute Error*) е мерка за грешки помеѓу спарените набљудувања кои го изразуваат истиот феномен. Примерите за *Y* наспроти *X* вклучуваат споредби на предвиденото наспроти набљудуваното, последователното време наспроти почетното време и една техника на мерење наспроти алтернативна техника на мерење. *MAE* се пресметува според равенката [77]:

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n}$$

Таа е аритметички просек на апсолутните грешки $|e_i| = |y_i - x_i|$, каде што y_i е предвидување, а x_i вистинската вредност.

Средната квадратна грешка (RMSE, Root Mean Square Error) е стандардното отстапување на грешките во предвидувањето. Остатоците се мерка за тоа колку се оддалечени точките на податоци од регресионата линија. RMSE е мерка за тоа колку се распространети овие остатоци. Коренот на средната квадратна грешка (RMSE) најчесто се користи во климатологијата, прогнозирањето и регресивната анализа за да се потврдат експерименталните резултати. Се пресметува со следното равенство [107]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}|y_i - x_i|^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}|e_i|^2}{n}}$$

каде $|e_i| = |y_i - x_i|$ се апсолутни грешки, y_i е предвидување, а x_i вистинската вредност.

Просечната апсолутна процентуална грешка (*MAPE, Mean Absolute Percentage Error)*, позната и како средна апсолутна процентуална девијација (*MAPD,* Mean Absolute Percentage Deviation), е мерка за точноста на предвидување на методот на предвидување во статистиката. Обично ја изразува точноста како сооднос дефиниран со формулата [78]:

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

каде A_t е вистинската вредност и F_t е прогнозираната вредност.

Симетрична средна апсолутна процентуална грешка (SMAPE или sMAPE, Symmetric Mean Absolute Percentage Error) е мерка за точност базирана на процентуални (или релативни) грешки. Обично се дефинира на следниов начин [78]:

$$SMAPE = \frac{100\%}{n}\sum_{t=1}^{n}\frac{|F_t - A_t|}{\frac{|F_t| + |A_t|}{2}}$$

каде A_t е вистинската вредност и F_t е прогнозираната вредност.

На слика 31 се претставени стандардните грешки во множеството за тренирање:

Show 10 entries Search:

	model	window_start	window_stop	mae	mape	smape
1	RandomForest	2019-04-24	2019-04-24	0.017	68.013	49.191

Слика 31. Стандардни грешки во множеството за тренирање генерирани од *R*.

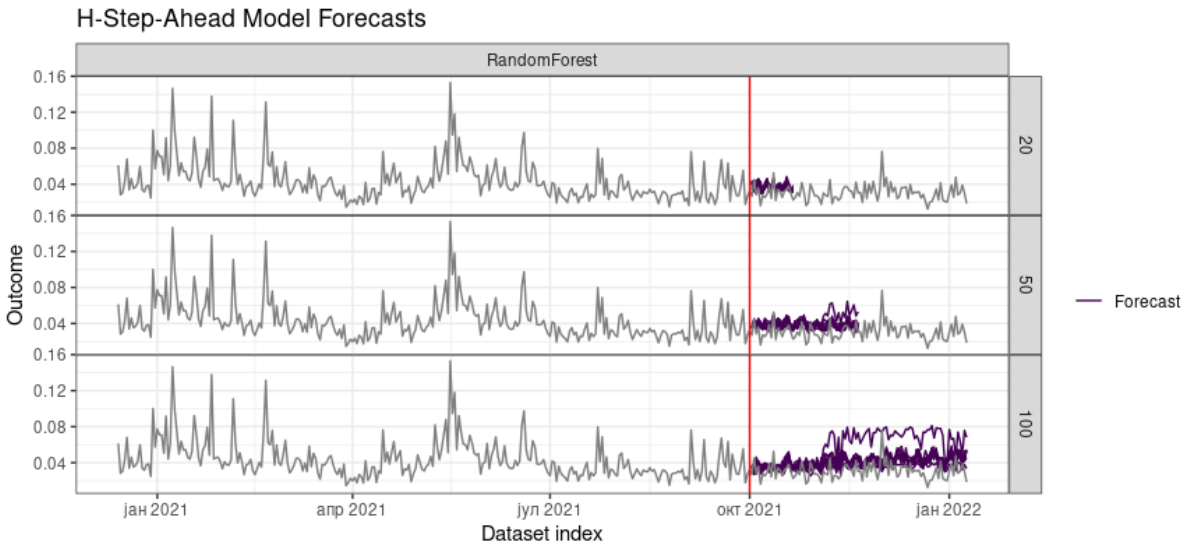
Следниот чекор е да се направи предвидување кога ќе се примени тест-множеството. Се користи функцијата *create_lagged_df()* со *type = "forecast"*. Користејќи ја функцијата *predict()* се прават предвидувања користејќи го истиот модел кој беше обучен претходно, но овој пат предвидените вредности одговараат на тест-множеството, прикажани на слика 32.

Show 10 entries Search:

cast_horizon	horizon	window_length	window_number	forecast_period	volatility_pred
20	1	95	1	2021-10-02	0.0413470958116285
20	2	95	1	2021-10-03	0.0352421085899088
20	3	95	1	2021-10-04	0.0321380088997544
20	4	95	1	2021-10-05	0.0456273624643628
20	5	95	1	2021-10-06	0.0349029077748487
20	6	95	1	2021-10-07	0.0320323093621885
20	7	95	1	2021-10-08	0.0418710840180415
20	8	95	1	2021-10-09	0.046009059391244
20	9	95	1	2021-10-10	0.0382462341271478
20	10	95	1	2021-10-11	0.0382064526818429

Слика 32. Дел од предвидените вредности кои одговараат на тест-множеството по хоризонти генерирани од *R*.

На слика 33 се претставени предвидувањата во тест-множеството. Има вкупно 30 прогнози, по 10 за секој хоризонт. Секоја прогноза одговара на секој прозорец за валидација кој претходно бил обучен за вкрстена валидација. На овој начин можеме да имаме поглед на стабилноста на прогнозите. Од графиците може да се забележи дека нема случаи со големо отстапување.



Слика 33. Предвидувањата во тест-множеството по хоризонти, генерирани во *R*.

Метриката на грешки за тест множеството по хоризонти, предвидените вредности наспроти вистинските вредности, се прикажани на слика 34 и табела 6. Средните просечни грешки се претставени во табела 7.

Show 10 entries Search:

	model	model_forecast_horizon	mae	mape	mdape	smape
1	RandomForest	20	0.009	32.73	25.074	26.351
2	RandomForest	50	0.009	37.902	31.734	29.358
3	RandomForest	100	0.012	48.656	37.467	35.971

Слика 34. Стандардни грешки во множеството за тестирање генерирани од *R*.

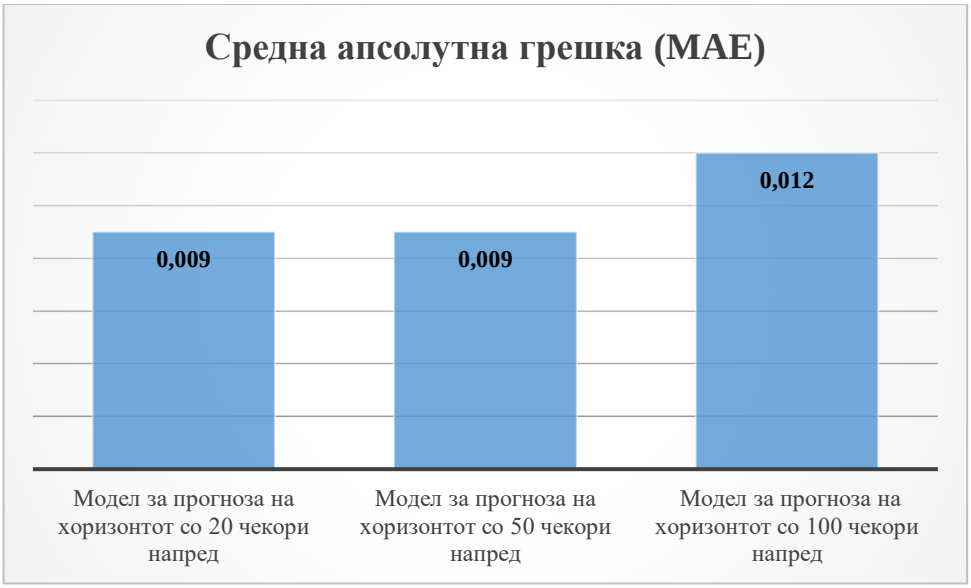
Табела 6. Стандардни грешки во множеството за тестирање со користење на различни хоризонти и прозорци за валидација.

Хоризонт за предвидување во моделот	Средна апсолутна грешка *(MAE)*	Апсолутната процентуална грешка (*MAPE)*	Просечната апсолутна процентна грешка (MDAPE)	Симетричната средна апсолутна процентуална грешка (sMAPE)
20	0.009	32.73	25.074	26.351
50	0.009	37.902	31.734	29.358
100	0.012	48.656	37.467	35.971

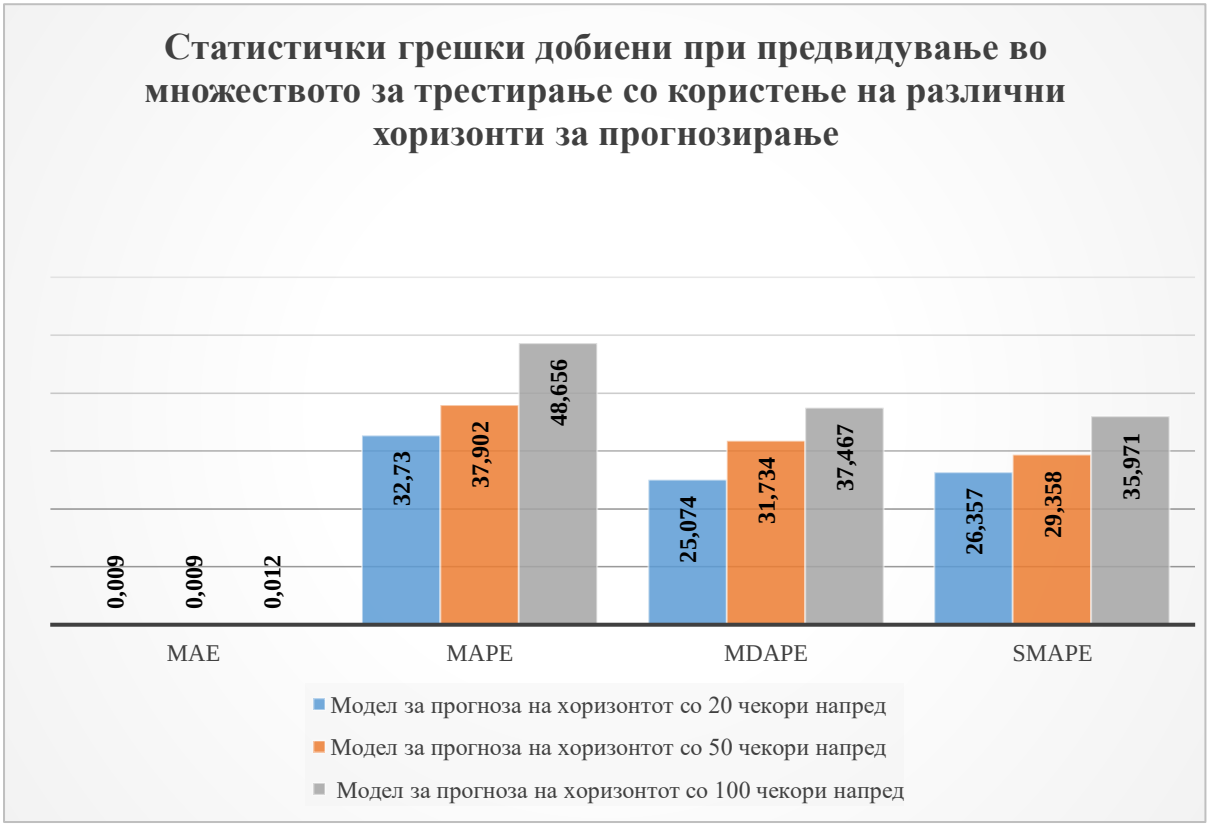
Табела 7. Средна вредност на стандардни грешки во множеството за тестирање со користење на различни хоризонти и прозорци за валидација во множеството за тестирање.

Средна вредност MAE	**Средна вредност MAPE**	**Средна вредност sMAPE**
0.009	37.826	29.358

Податоците за метриките за грешки од табела 6 визуелно се прикажани на слика 35 и слика 36.



Слика 35. Графички приказ на средната апсолутна грешка во множеството за тестирање.

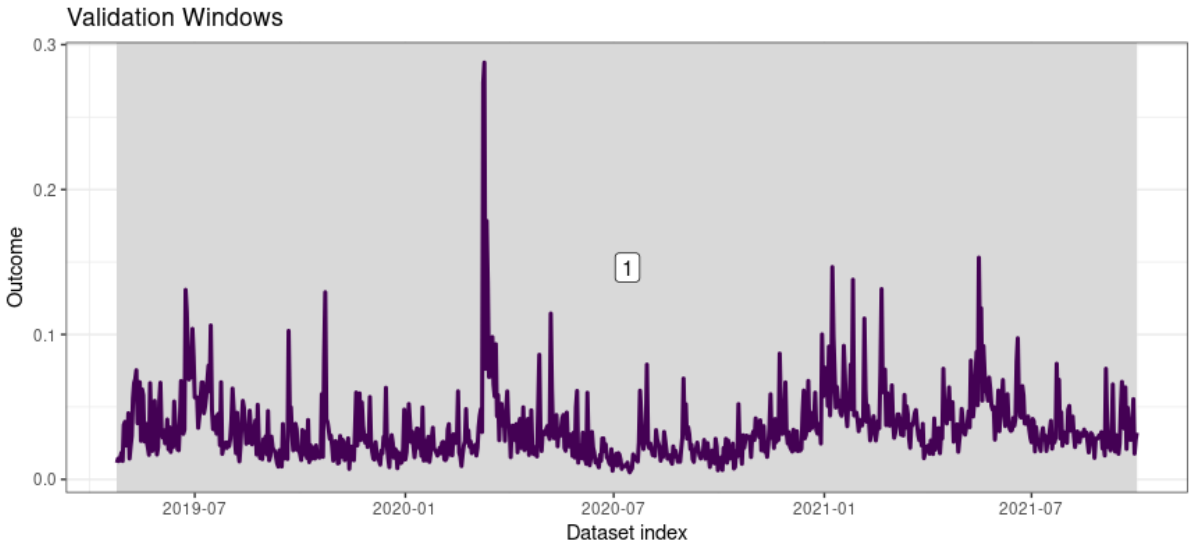


Слика 36. Графички приказ на стандардни грешки во множеството за тестирање.

Во споредба со метриката за грешки на множеството за тренирање, метриката на грешки на тест-множеството е помала. Од прикажаните резултати може да се забележи дека грешката е помала за пократок хоризонт за предвидување што е очекувано. Вредностите на грешките се задоволителни што укажува на фактот дека моделот е добар.

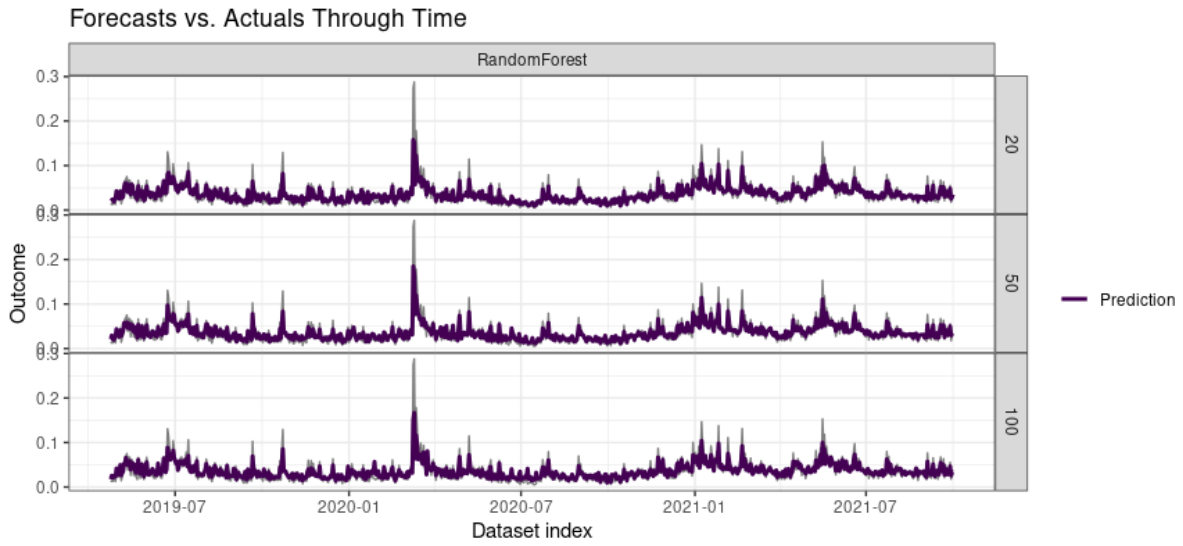
Во вториот дел од анализата потребно е да се обучи моделот низ целата база на податоци за обука без вгнездена вкрстена валидација. Без вгнездени вкрстени валидации и прозорци за задржување, графикот на предвидување во основа, одговара на моделот. Потоа се прикажува графикот на реални наспроти предвидени вредности и се пресметуваат метриките за грешки. Следниот чекор е да се предвиди во тест-множеството и повторно да се прикажат реалните наспроти предвидените вредности и метрика за грешки за тест-множеството. Потоа ги комбинираме предвидувањата на секој хоризонт на предвидените вредности на тест-множеството. Конечниот дел е да се предвиди надвор од примерокот за секој хоризонт со користење на множествата за тренирање и тестирање и повторно да се комбинираат прогнозите надвор од примерокот за секој хоризонт за да се прикаже конечната комбинирана прогноза.

Следно, во *create_windows()* се поставува *window_length = 0* за да се произведе база на податоци без вгнездена вкрстена валидација, само со еден прозорец. На слика 37 е прикажан графикот за предвидување без вгнездена вкрстена валидација, кој одговара на моделот.



Слика 37. График за предвидување без вгнездена вкрстена валидација генериран од *R*.

Потоа повторно се обучува нов модел користејќи само еден прозорец за валидација. Се користи *create_lagged_df()* и *type = "train"* за прво да се прогнозира во множеството за обука. Со користење на *train_model()* и *predict()*, повторно се тренира моделот и се прави предвидување, но овој пат без прозорци за валидација. Прогнозите во множеството за обука во различни хоризонти се претставени на слика 38.



Слика 38. Прогнози во секој хоризонт со користење на еден прозорец за валидација во множеството за тренирање генерирани од *R*.

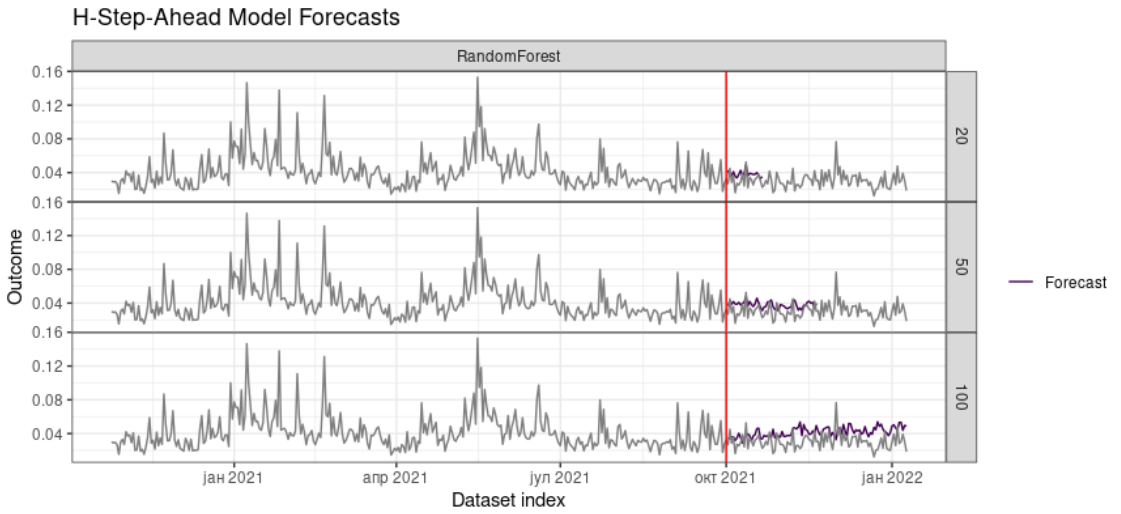
Стандарните грешки при предвидувањето во множеството за тренирање со еден прозорец за валидација се прикажани на слика 39.

Show 10 entries Search:

	model	window_start	window_stop	mae	mape	mdape	smape
1	RandomForest	2019-04-24	2019-04-24	0.006	21.983	14.368	18.672

Слика 39. Стандардни грешки во множеството за тренирање со еден прозорец за валидација генерирани од *R*.

По прогнозата во множеството за тренирање, се променува *type = "forecast"* во *create_lagged_df()*, така што се предвидува во тест-множество користејќи го моделот обучен без прозорци за валидација. Прогнозите на тест-множеството се претставени на слика 40, а во табелата од слика 41 се прикажани метриките за грешки на предвидените вредности во тест-множеството без прозорци за валидација.



Слика 40. Прогнози во секој хоризонт со користење на еден прозорец за валидација во множеството за тестирање генерирани од *R*.

Show 10 entries Search:

	model	mae	mape	mdape	smape
1	RandomForest	0.009	36.725	30.604	28.638

Слика 41. Стандардни грешки во множеството за тестирање без вгнездена вкрстена валидација, генерирани од *R*.

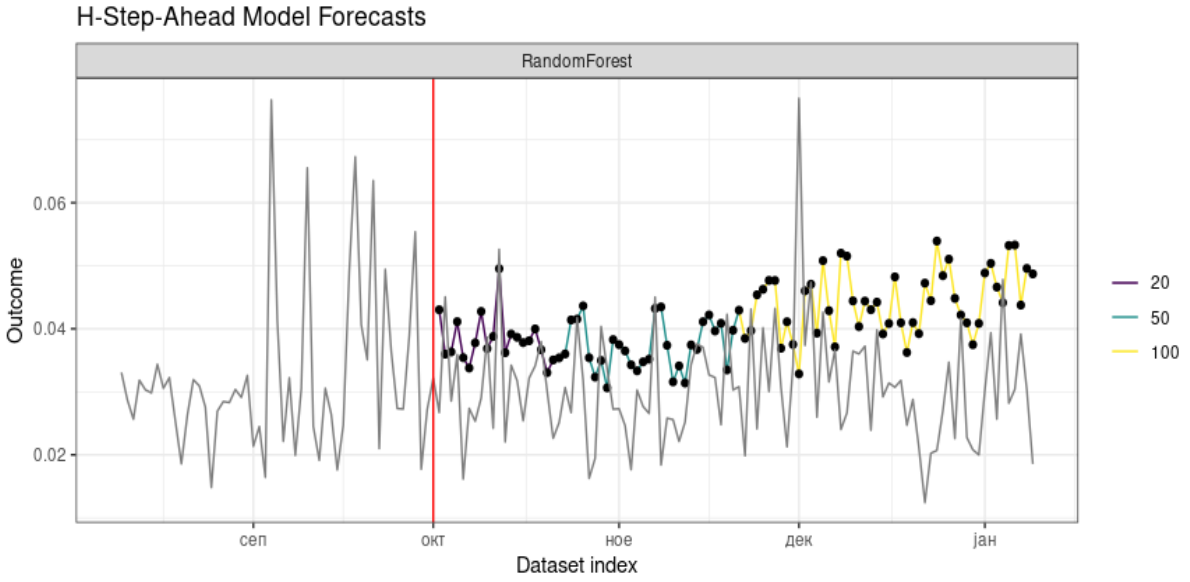
Табела 8. Средна вредност на стандардни грешки во множеството за тестирање, без вгнездена вкрстена валидација.

Средна вредност MAE	**Средна вредност MAPE**	**Средна вредност sMAPE**
0.009	36.725	28.638

Ако се споредат податоците во табела 7 и 8, се забележува дека грешката на тест-множеството е поголема во споредба со множеството за тренирање. Тој резултат е очекуван, бидејќи податоците од тест-множеството не се користат кога се тренира моделот. Фактот што грешката во тест-множеството не станува екстремно поголема, укажува на тоа дека моделот предвидува добро. Средната апсолутна грешка во множеството за тренирање е 0,006, а во тест множеството е 0,009, што докажува дека моделот има добро предвидување.

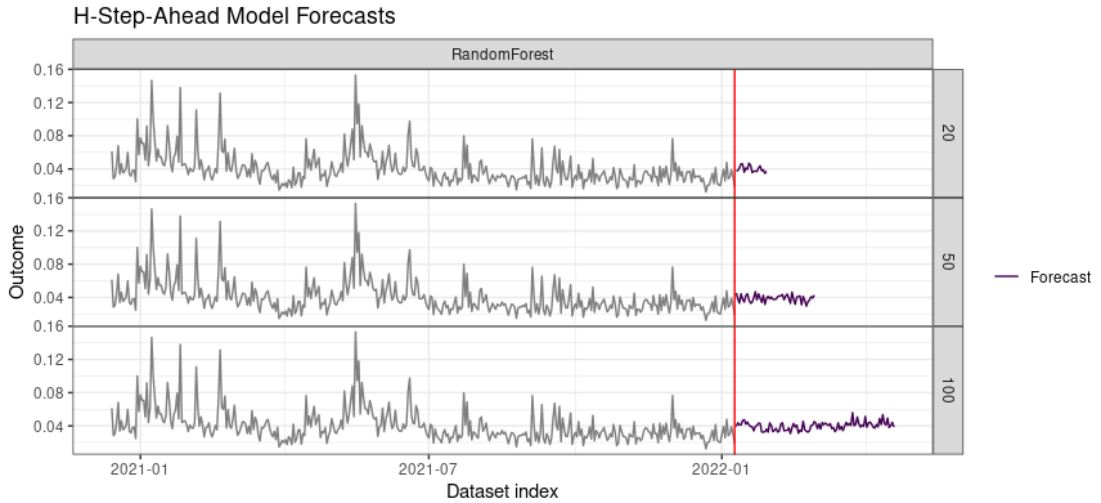
Според алгоритамот случајни шуми, следниот чекор е да се комбинираат предвидувањата на секој хоризонт на директна прогноза користејќи ја функцијата *combine_forecasts()*. Добиената комбинирана шема за прогноза е претставена на слика 42. Со наредбата *combine_forecasts()* се прогнозира 100 чекори напред, каде што ќе се комбинираат различни прогнози за да се добие подобар резултат. Прогнозите се

комбинираат на тој начин што краткорочните прогнози се произведени од краткорочните модели, а долгорочните предвидувања од долгорочниот модел. Крајната прогноза за 100 чекори напред е добиена на следниот начин: за првите 20 чекори напред се зема резултатот за предвидување што го дава првиот модел, за предвидување од 21 до 50 чекор што следат се земаат податоци од вториот модел и за предвидување од 51 до 100 чекор напред се земаат податоци од третиот модел.



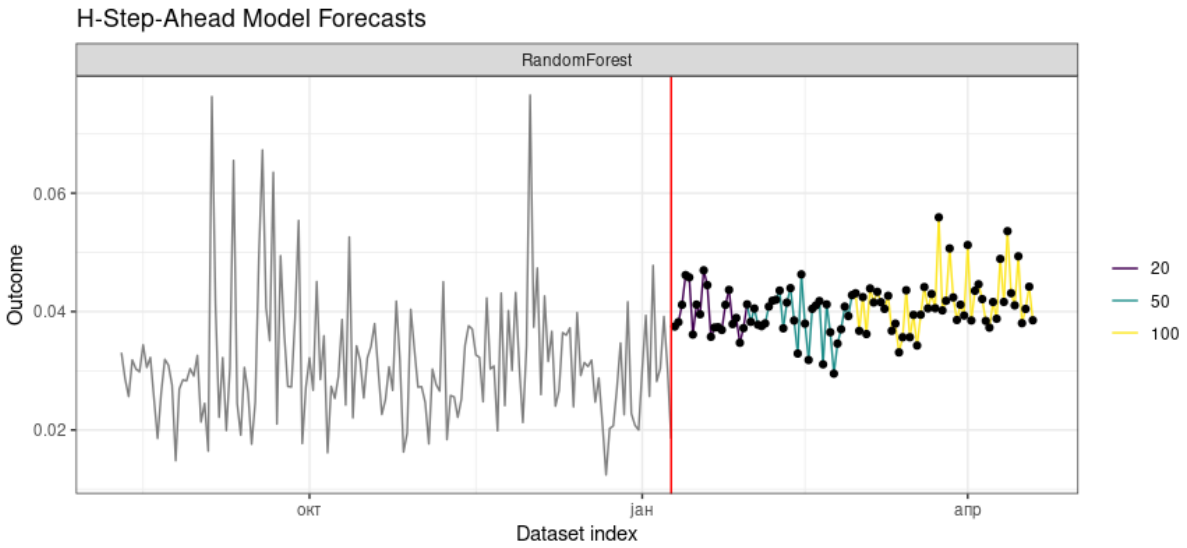
Слика 42. Комбинираната шема за прогноза во тест-множеството генерирана од *R*.

Следен чекор според алгоритамот за случајна шума е да се оствари предвидување надвор од примерокот. Во функцијата *create_lagged_df()* како аргументи се задаваат податоците од двете множества и за тренирање и за тестирање и тип, *type = "forecast"* и со користење на функцијата *predict()* се прогнозира надвор од примерокот, што е претставено на слика 43.



Слика 43. Предвидување надвор од примерокот генерирано од *R*.

Повторно се комбинираат предвидувањата на секој хоризонт надвор од примерокот користејќи ја функцијата *combine_forecasts()*. Добиената комбинирана шема за прогноза надвор од примерокот за 100 чекори напред е претставена на слика 44. Со наредбата *combine_forecasts()* се прогнозира 100 чекори напред, каде што се комбинираат различни прогнози за да се добие подобар резултат.



Слика 44. Комбинираната шема за прогноза надвор од примерокот генерирана од *R*.

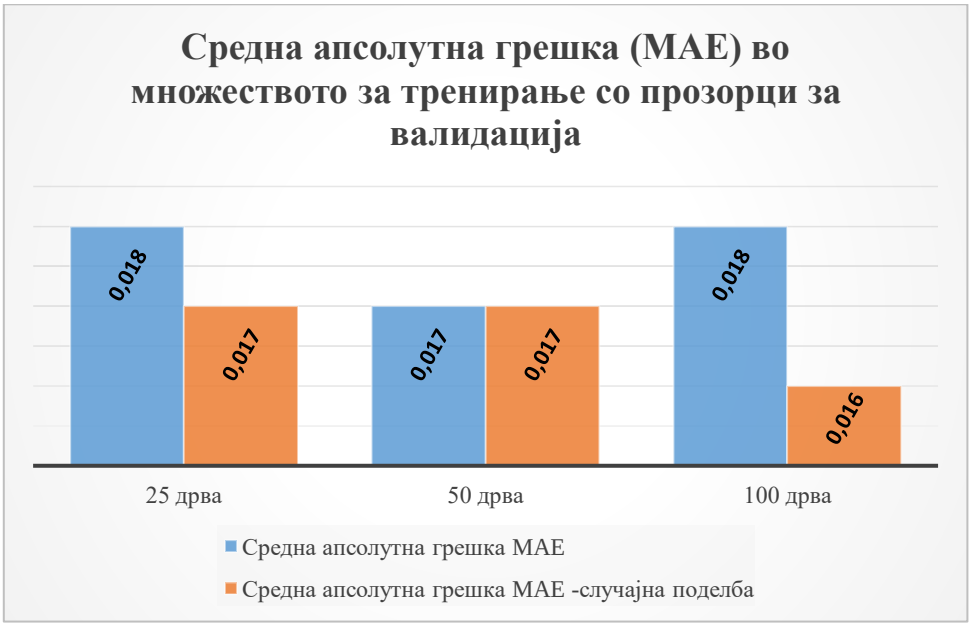
6 Дискусија за добиените резултати

Во докторскиов труд направена е анализа на резултатите што се добиени при користење на направениот алгоритам во програмскиот јазик *R* за предвидување на променливоста на биткоинот. Моделот што е креиран со користење на алгоритамот од машинско учење случајни шуми е испитуван во неколку различни ситуации.

6.1 Анализа на резултатите добиени со користење на различен број на дрва во алгоритамот

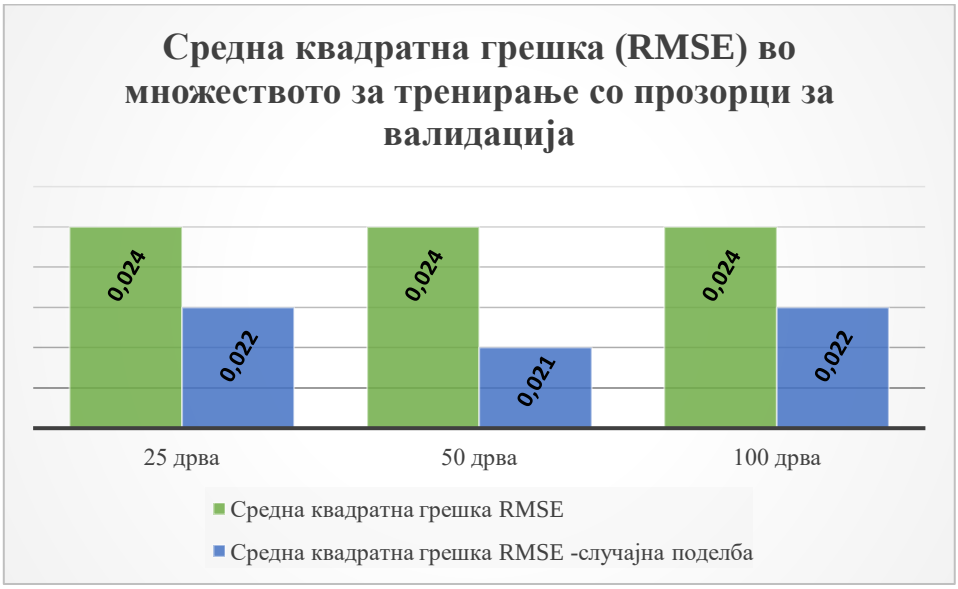
Направена е анализа на резултатите добиени со користење на различен број на дрва во алгоритамот и тоа за 25, 50 и 100. Направени се испитувања во ситуација кога множеството за тренирање и тестирање е поделено со точно одреден датум на временската серија и втора ситуација кога случајно се избира кои податоци ќе се земат во множеството за тренирање и тестирање со функцијата *sample.split()* од библиотеката *caTools*, при предвидување на променливоста на пазарната цена на биткоинот. При ова испитување користена е поделба на множеството за тренирање и тестирање во соднос 90 % : 10 %.

На графиконот на слика 45 прикажана е споредба на средната апсолутна грешка *MAE* со користење на различен број на дрва и различна поделба на множеството за тренирање и тестирање, во множеството за тренирање со користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 45. Средна апсолутна грешка (*MAE*) во множеството за тренирање со прозорци за валидација во ситуации со користење на 25, 50 и 100 дрва.

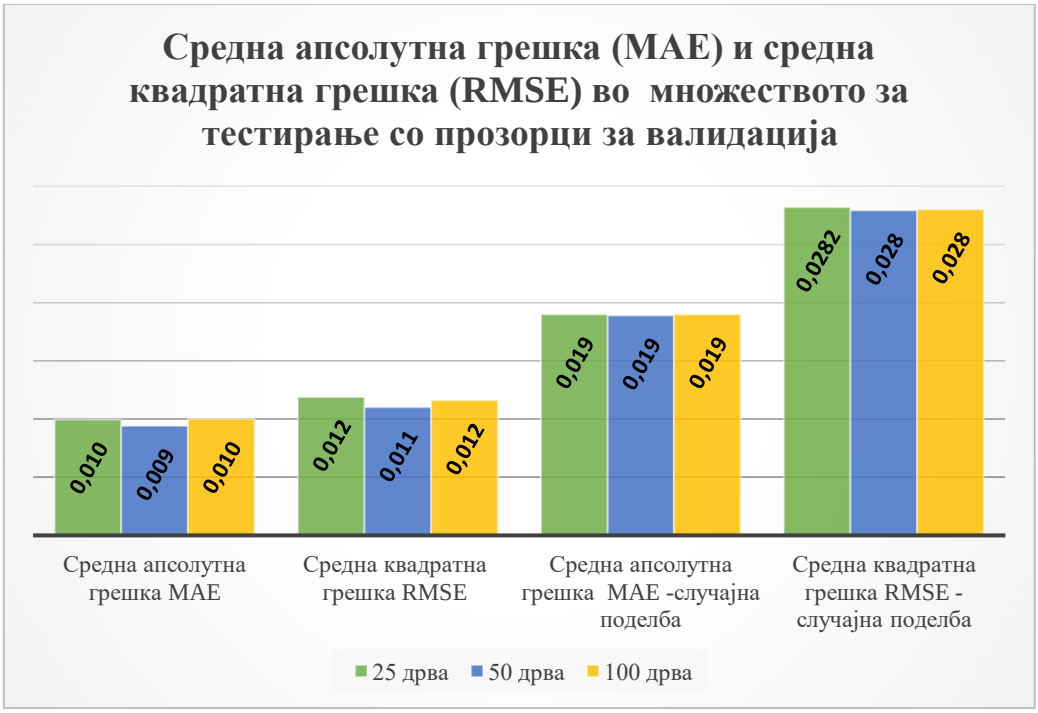
На графиконот на слика 46 е прикажана споредбата на средната квадратна грешка *RMSE,* која се користи за оценка на моделот за предвидување со користење на различен број на дрва и различна поделба на множествата за тренирање и тестирање, во множеството за тренирање со користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 46. Средна квадратна грешка (*RMSE*) во множеството за тренирање со прозорци за валидација во ситуации со користење на 25, 50 и 100 дрва.

Од прикажаните графикони може да се забележи дека најмала средна апсолутна грешка (*MAE*) во множеството за тренирање со прозорци за валидација при предвидувањето е добиена кога во алгоритамот за предвидување се користат 100 дрва и случајна поделба на множествата за тренирање и тестирање со вредност 0,016. А средната квадратна грешка *RMSE* која, исто, е мерка за евалуација на моделот за предвидување, најмала вредност има при користење на 50 дрва и случајна поделба на множествата за тренирање и тестирање со вредност 0,021.

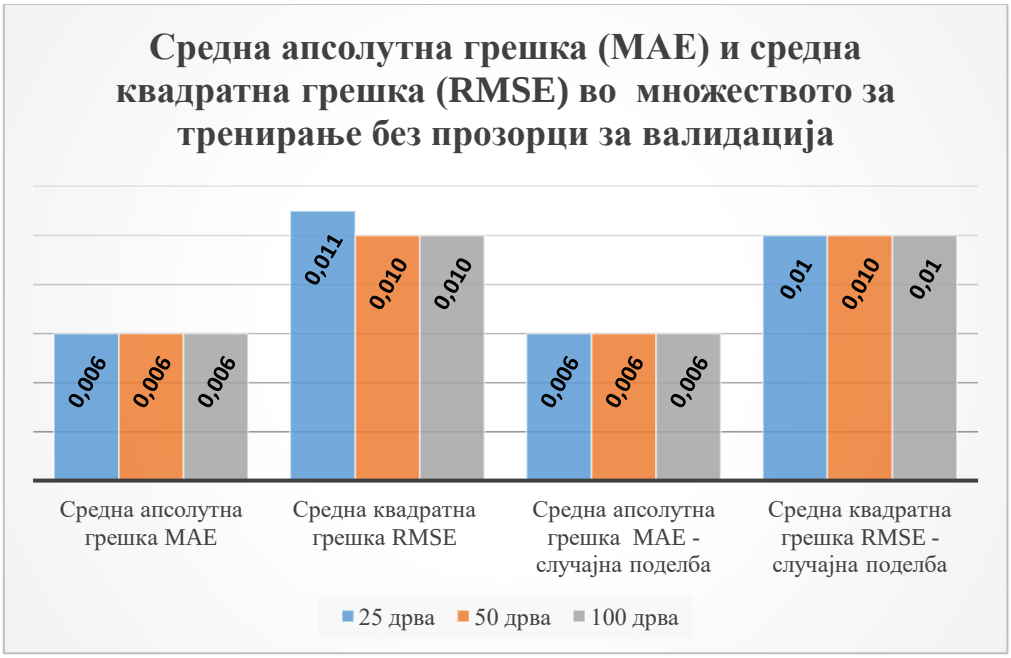
На графиконот на слика 47 прикажана е споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка (*RMSE)* со користење на различен број на дрва и различна поделба на множествата за тренирање и тестирање, во множеството за тестирање со користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 47. Средната апсолутна грешка (*MAE*) и средната квадратна грешка *(RMSE*) во множеството за тестирање со прозорци за валидација во ситуации со користење на 25, 50 и 100 дрва.

Од графиконот може да се заклучи дека средната апсолутна грешка *(MAE)* и средната квадратна грешка (*RMSE)* добиле најмала вредност при користење на 50 дрва и фиксна поделба на множеството за тренирање и тестирање со вредност од 0,009 и 0,011, соодветно.

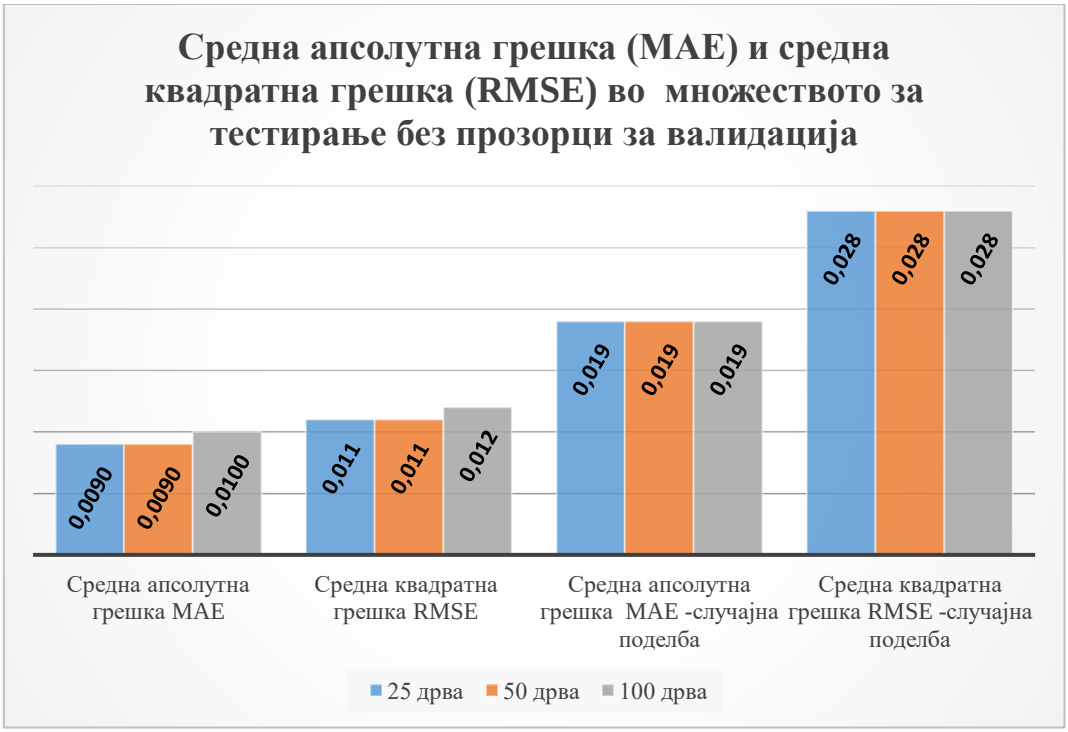
Споредбата на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различен број на дрва и различна поделба на множеството за тренирање и тестирање, во множеството за тренирање без користење на прозорци за валидација при предвидување на променливоста на биткоинот е прикажана на слика 48.



Слика 48. Средна апсолутна грешка *(MAE)* и средна квадратна грешка (*RMSE*) во множеството за тренирање без прозорци за валидација во ситуации со користење на 25, 50 и 100 дрва.

Од графиконот може да се заклучи дека вредноста на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* е речиси идентична при користење на 25, 50 или 100 дрва без разлика на начинот на поделба на множествоата за тренирање и тестирање со вредност од 0,006 и 0,010, соодветно.

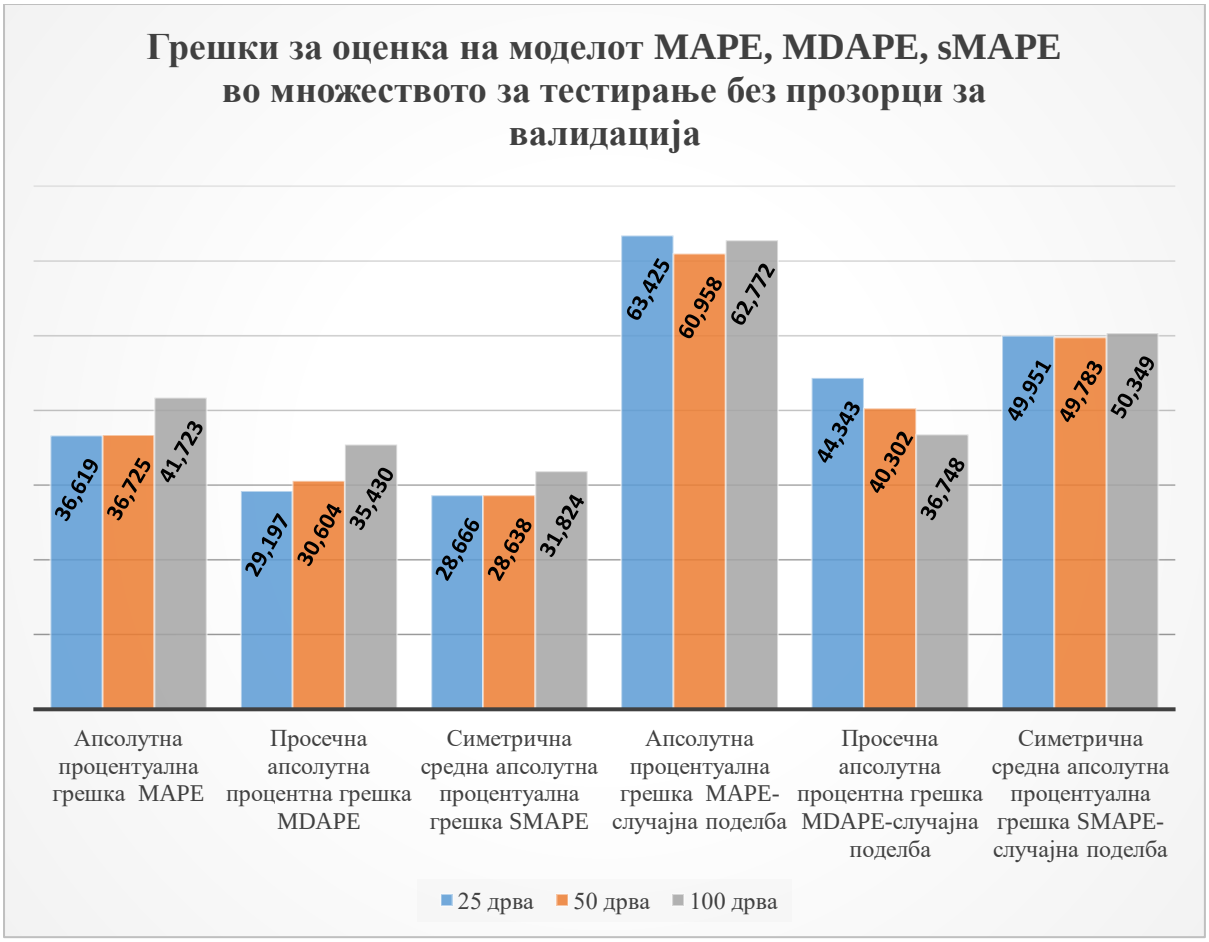
На графиконот од слика 49 е прикажана споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различен број на дрва и различна поделба на множествата за тренирање и тестирање, во множеството за тестирање без користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 49. Средна апсолутна грешка (*MAE*) и средна квадратна грешка (*RMSE*) во множеството за тестирање без прозорци за валидација во ситуации со користење на 25, 50 и 100 дрва.

Од графиконот на слика 49 може да се забележи дека бројот на дрва не влијае многу на прикажаните грешки. Дополнително, може да се заклучи дека помала грешка за моделот се добива ако се користи фиксна поделба на множеството за тестирање и тренирање, и тоа *MAE* = 0.009, а *RMSE* = 0.011.

На слика 50 графиконот прикажува споредба во множеството за тестирање без користење на прозорци за валидација на апсолутната процентуална грешка (*MAPE*), просечната апсолутна процентна грешка (*MDAPE*) и симетричната средна апсолутна процентуална грешка *(sMAPE)* со користење на различен број на дрва и различна поделба на множествата за тренирање и тестирање, при предвидување на променливоста на биткоинот.



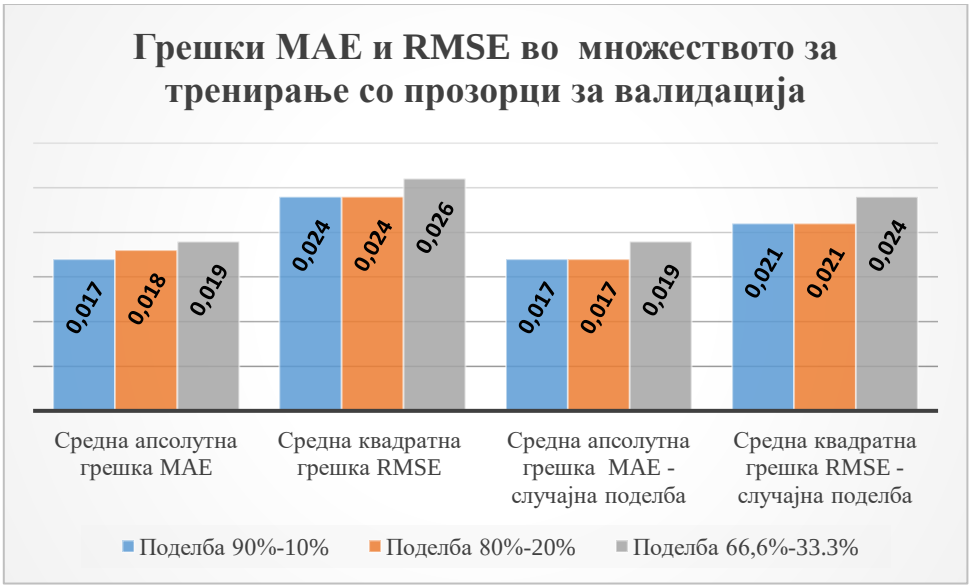
Слика 50. Апсолутната процентуална грешка (*MAPE*), просечната апсолутна процентна грешка (*MDAPE*) и симетричната средна апсолутна процентуална грешка (*sMAPE*) со користење на различен број на дрва.

Од графиконот на слика 50 може да се заклучи дека апсолутната процентуална грешка (*MAPE*), просечната апсолутна процентна грешка *(MDAPE)* и симетричната средна апсолутна процентуална грешка *(sMAPE)* имаат значително помали вредности кога се користи поделба на множествата за тренирање и тестирање во точно одредена точка на временската серија. Се заклучува дека тој модел за предвидување е подобар.

6.2 Анализа на резултатите добиени со користење на различна поделба на множествата за тренирање и тестирање

Во овој труд направена е анализа на резултатите добиени со користење на различна поделба на множествата за тренирање и тестирање, и тоа за сооднос 90 % : 10 %, 80 % : 20 % и 66,6 % : 33,3 %. Направени се испитувања во ситуација кога множествата за тренирање и тестирање се поделени со точно одреден датум на временската серија и во ситуација кога случајно се избираат кои податоци од базата ќе се земаат во множествата за тренирање и тестирање со функцијата *sample.split()* од библиотеката *caTools*, при предвидување на променливоста на пазарната цена на биткоинот. Визуелна споредба на средната апсолутна грешка (*MAE)* и средната

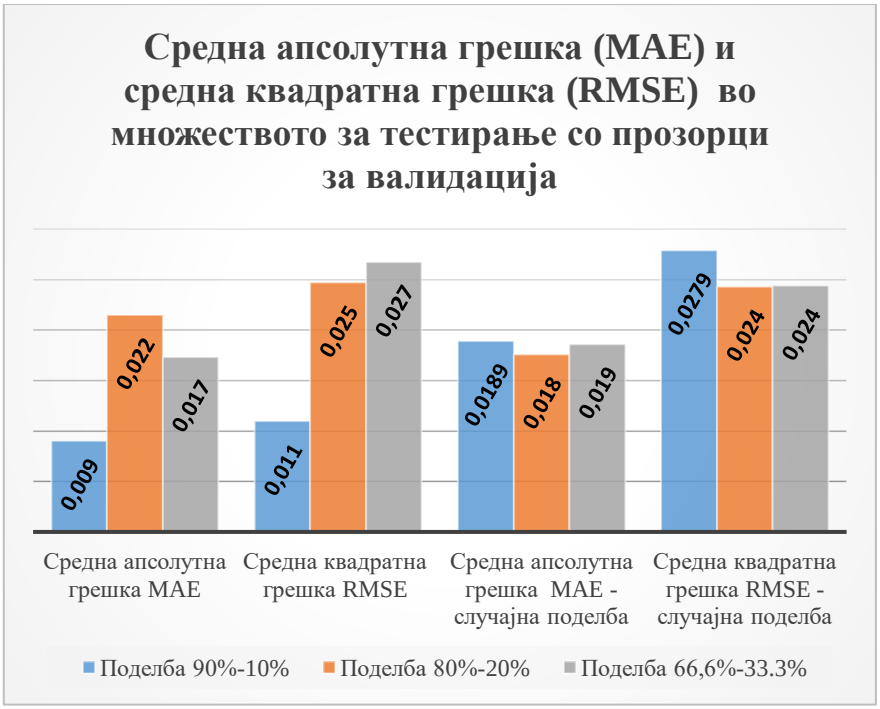
квадратна грешка (*RMSE*) во множеството за тренирање со прозорци за валидација, прикажана е на слика 51.



Слика 51. Средна апсолутна грешка (*MAE)* и средна квадратна грешка (*RMSE*) во множеството за тренирање со прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

Од графиконот може да се заклучи дека вредноста на средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) не се менуваат значително со менување на соодносот на поделбите на множеството за тренирање и тестирање. Најмала вредност за *MAE* = 0,017, а за *RMSE* = 0,021 при случајна поделба на податоците.

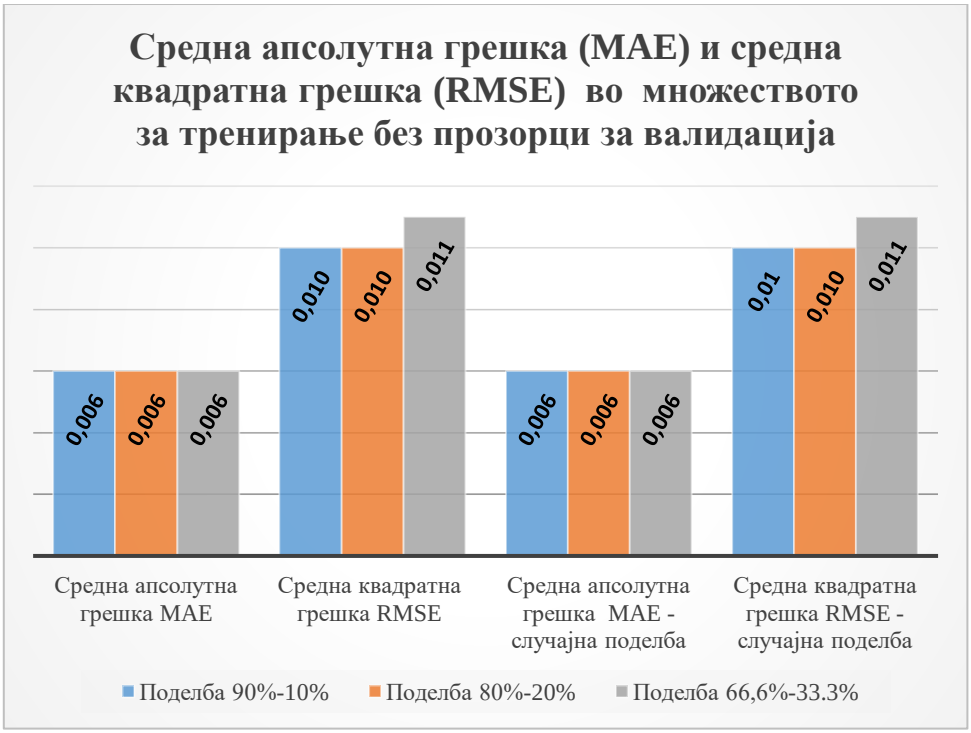
Следно се анализирани грешките во множеството за тестирање со користење на прозорци за валидација при предвидување на променливоста на биткоинот, прикажани на слика 52. На графиконот прикажана е споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различна поделба на множествата за тренирање и тестирање и различен избор на начинот на поделба.



Слика 52. Средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) во множеството за тестирање со прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

Од визуелниот приказ на слика 52 може да се заклучи дека средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) добиваат најмали вредности кога сообносот на поделба на множеството за тренирање и тестирање е 90 % : 10 % со фиксна поделба на податоците од временската серија *MAE* = 0,009, а *RMSE* = 0,011.

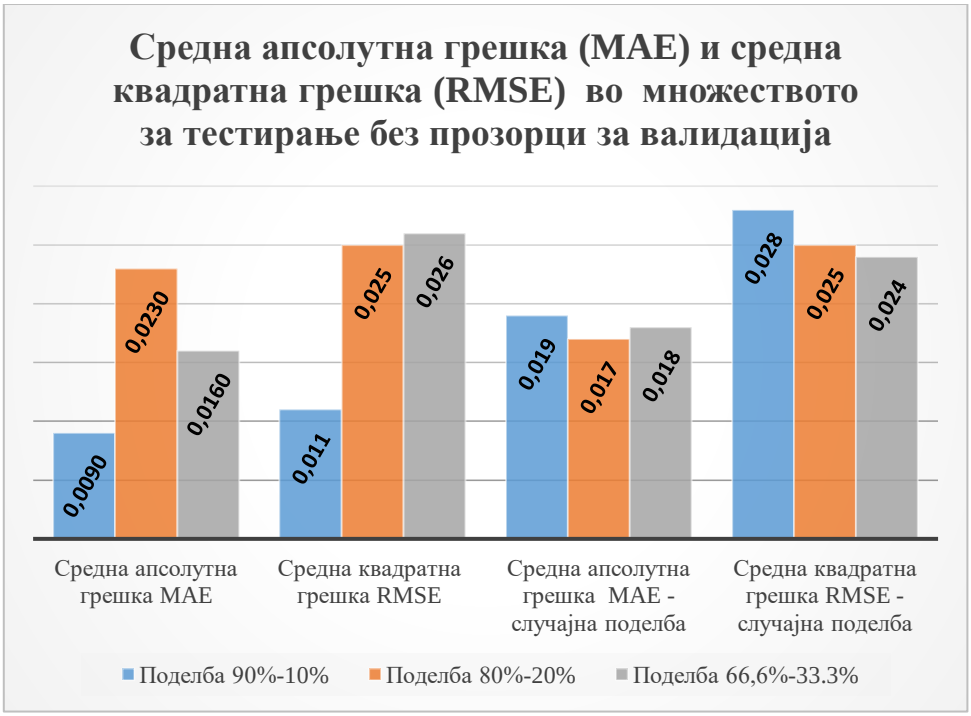
Грешките во множеството за тренирање без користење на прозорци за валидација при предвидување на променливоста на биткоинот прикажани се на слика 53. На графиконот прикажана е споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различна поделба на множествата за тренирање и тестирање и различен сооднос на поделба.



Слика 53. Средната апсолутна грешка (MAE) и средната квадратна грешка (RMSE) во множеството за тренирање без прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

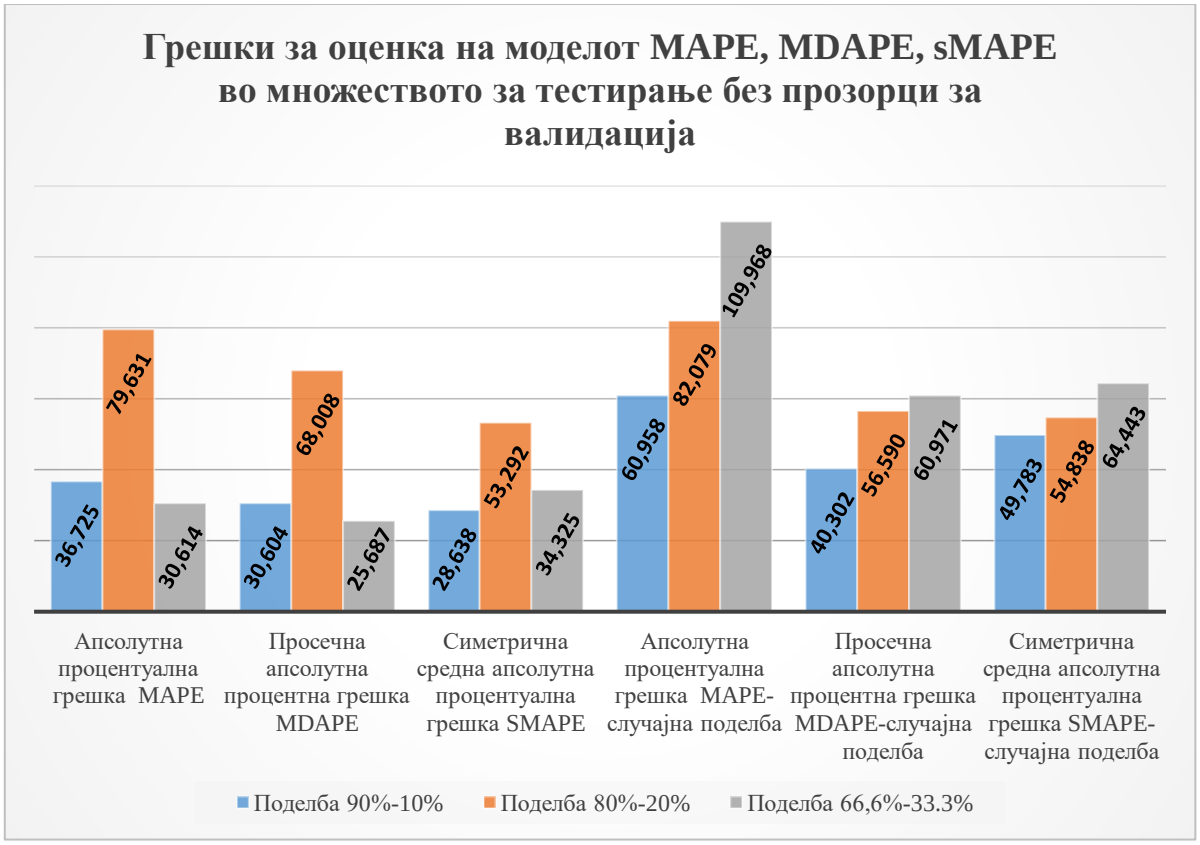
Од графиконот прикажан на слика 53, може да се заклучи дека вредноста на средната апсолутна грешка (*MAE)* и средната квадратна грешка (*RMSE*) е речиси идентична при користење на различна поделба на множеството за тренирање и тестирање, и тоа за сооднос 90 % : 10 %, 80 % : 20 % и 66,6 % : 33,3 %, без разлика на начинот на поделба на множествата за тренирање и тестирање со вредност од 0,006 и 0,010, соодветно.

На слика 54 визуелно е прикажана споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различна поделба на множествата за тренирање и тестирање и различен избор на начинот на поделба, во множеството за тестирање без користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 54. Средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) во множеството за тестирање без прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

Од графиконот на слика 54 може да се заклучи дека средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) најмали вредности имаат кога соодносот на поделба на множеството за тренирање и тестирање е 90 % : 10 % со фиксна поделба на податоците од временската серија *MAE* = 0,009, а *RMSE* = 0.011.



Слика 55. Апсолутната процентуална грешка (*MAPE*), просечната апсолутна процентна грешка (*MDAPE*) и симетричната средна апсолутна процентуална грешка (*sMAPE*) во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

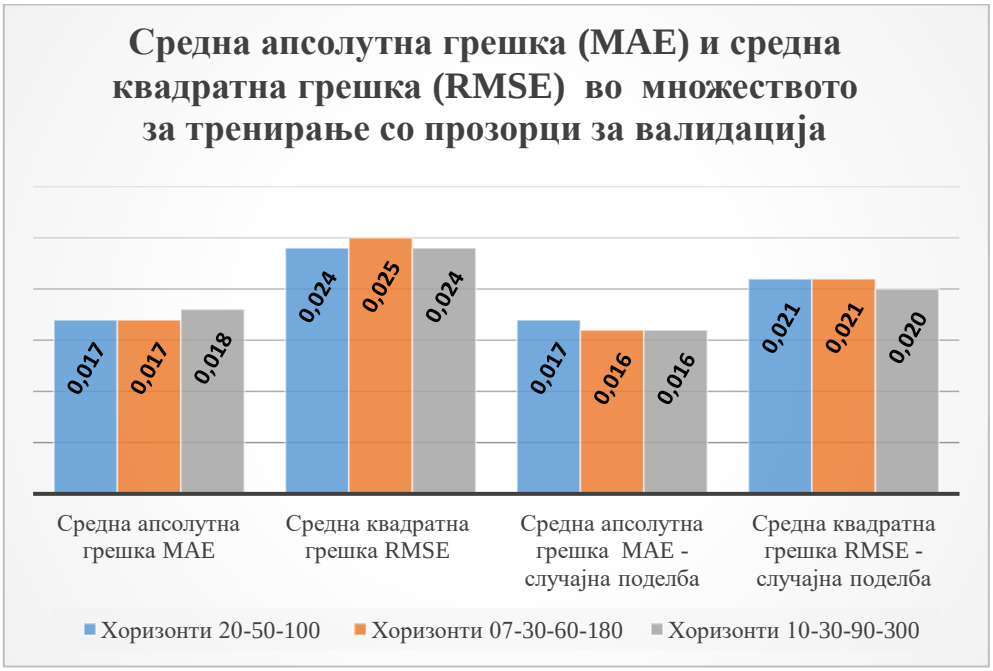
Од графиконот на слика 55, може да се заклучи дека апсолутната процентуална грешка (*MAPE*) и просечната апсолутна процентна грешка *(MDAPE)* имаат значително помали вредности кога се користи фиксна поделба на множествата за тренирање и тестирање и избран сооднос 66,6 % : 33.3 % и со што се заклучува дека тој модел за предвидување е подобар.

6.3 Анализа на резултатите добиени со користење на различни хоризонти за предвидување

Направена е анализа на резултатите добиени со користење на различни хоризонти за предвидување и тоа со 20 – 50 – 100 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 90 % : 10 %. Потоа, анализирана е ситуација со 7 – 30 – 60 – 180 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 83,5 % : 16,5 %. Следно е разгледана трета ситуација со 10 – 30 – 90 – 300 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 73 % : 27 %. Направени се испитувања во ситуација кога множеството за тренирање и тестирање се поделени со точно одреден датум на временската серија и во ситуација кога случајно се избираат кои податоци ќе се земат во множеството за тренирање и тестирање со функцијата *sample.split()* од

библиотеката *caTools*, при предвидување на променливоста на пазарната цена на биткоинот.

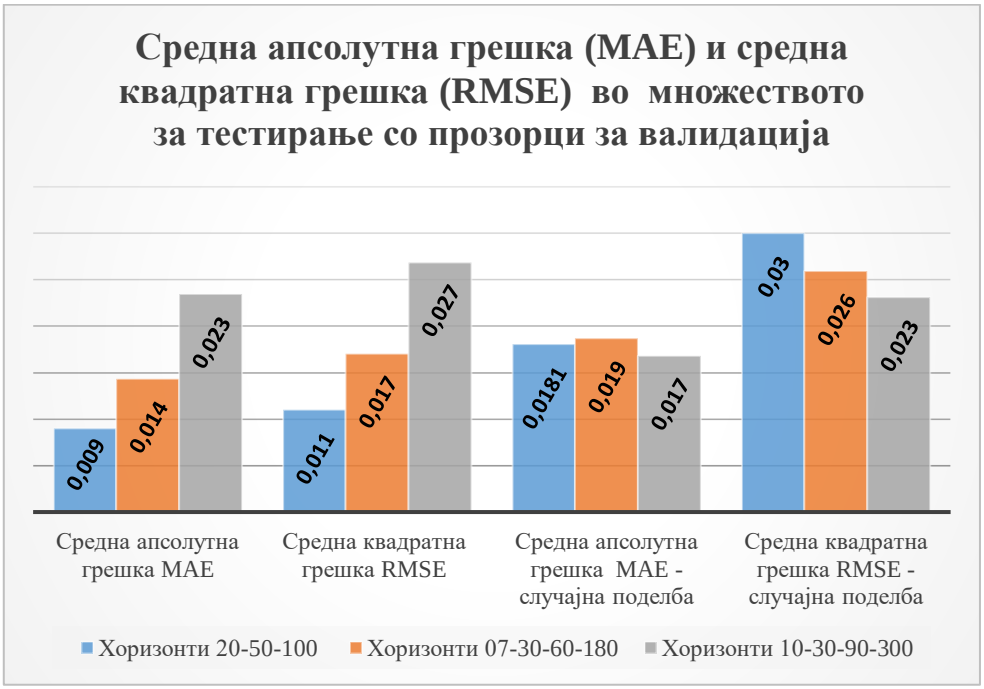
На следниот графикон од слика 56, прикажана е споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка (*RMSE*) во множеството за тренирање со користење на прозорци за валидација, со користење на различен број на хоризонти за предвидување, различен сооднос и различен начин на поделба на множествата за тренирање и тестирање, при предвидување на променливоста на биткоинот.



Слика 56. Средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) во множеството за тренирање со прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

Од графиконот на слика 56 може да се заклучи дека вредноста на средната апсолутна грешка (*MAE)* и средната квадратна грешка (*RMSE*) многу малку се променува при користење на различни хоризонти за предвидување, без разлика на начинот на поделба на множествата за тренирање и тестирање. Најмала вредност за *MAE* = 0.016, а за *RMSE* = 0.020, кои се добиени при случајна поделба на множествата за тренирање и тестирање во сооднос 73 % : 27 % и хоризонти за предвидување 10 – 30 – 90 – 300 .

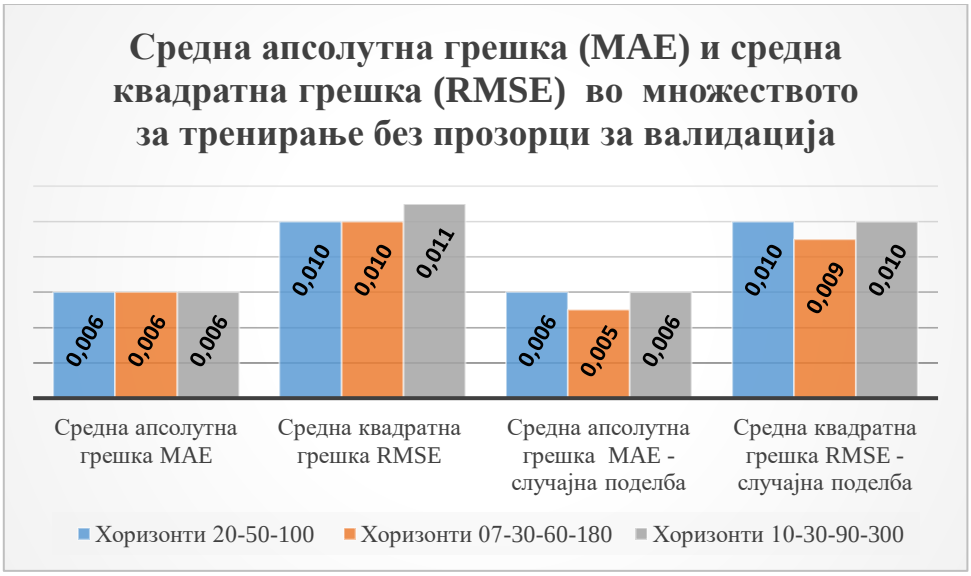
Следно, се прават анализи за тестирање со прозорци за валидација. Од графиконот на слика 57 може да се направи споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка (*RMSE*) во множеството за тестирање со прозорци за валидација со користење на различен број на хоризонти за предвидување и различна поделба на множествата за тренирање и тестирање.



Слика 57. Средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) во множеството за тестирање со прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

Од графиконот на слика 57 може да се заклучи дека вредноста на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* е помала при фиксна поделба на податоците од временската серија и тоа кога се испитуваат хоризонтите 20 – 50 – 100 при поделба на множествата со сооднос 90 % : 10 % и ги имаат следниве вредности: *MAE* = 0.009, а *RMSE* = 0.011.

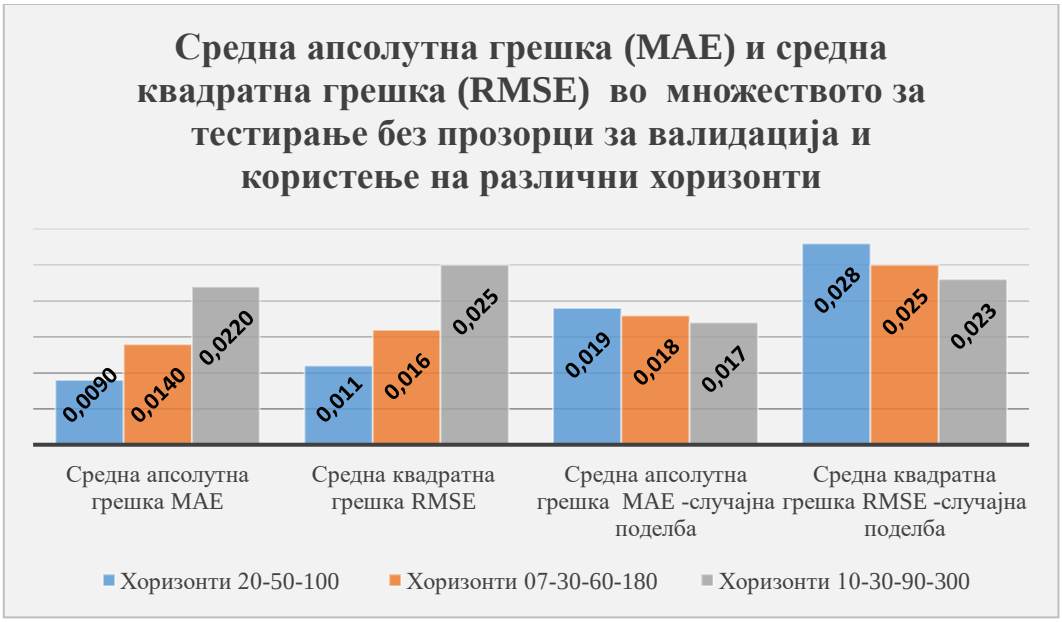
На слика 58 визуелно е прикажана споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различен број на хоризонти за предвидување и различна поделба на множествата за тренирање и тестирање, во множеството за тренирање без користење на прозорци за валидација при предвидување на променливоста на биткоинот.



Слика 58. Средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE)* во множеството за тренирање без прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање.

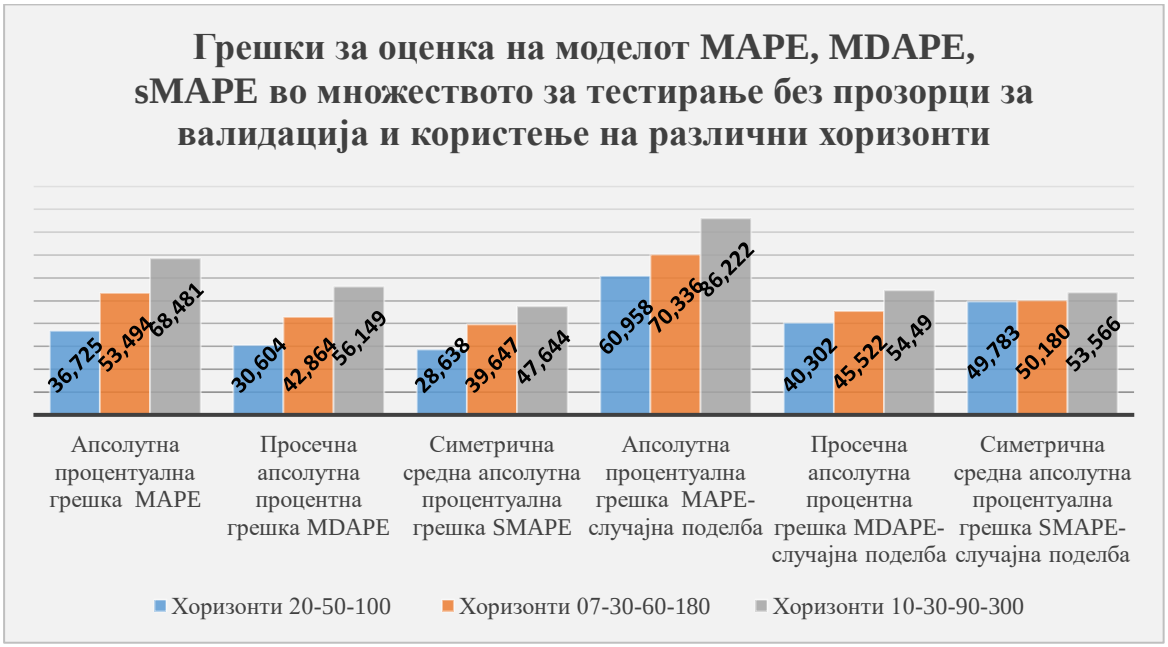
Од графиконот на слика 58 може да се заклучи дека вредноста на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* во множеството за тренирање без прозорци за валидација се речиси идентични при користење на различни хоризонти за предвидување без разлика на начинот на поделба на множествата за тренирање и тестирање, со најмала вредност од 0,005 и 0,009 при случајна поделба на податоците за множество за тренирање и множество за тестирање.

Анализата продолжува на множеството за тестирање без користење на прозорци за валидација. На слика 59 прикажана е споредба на средната апсолутна грешка (*MAE)* и средната квадратна грешка *(RMSE)* со користење на различен број на хоризонти за предвидување и различна поделба и различен сооднос на поделба на множествата за тренирање и тестирање, при предвидување на променливоста на биткоинот.



Слика 59. Средната апсолутна грешка (MAE) и средната квадратна грешка (RMSE) во множеството за тестирање без прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање и различни хоризонти.

Од графиконот на слика 59 може да се заклучи дека вредностите на грешките со кои го оценуваме моделот малку се разликуваат. Во овој случај помали грешки се добиени кога поделбата на податоци во множествата за тренирање и тестирање е направена на фиксен датум. Најмали вредности се добиени кога се користени хоризонти за предвидување 20 – 50 – 100 и сооднос на поделба 90 % : 10 % и тоа *MAE* = 0.009, а *RMSE* = 0.011.



Слика 60. Апсолутната процентуална грешка (MAPE), просечната апсолутна процентна грешка (MDAPE) и симетричната средна апсолутна процентуална грешка (sMAPE) во множеството за тестирање без прозорци за валидација во ситуации со користење на различна поделба на множествата за тренирање и тестирање и различни хоризонти.

Од графиконот на слика 60, може да се заклучи дека апсолутна процентуална грешка (*MAPE),* просечната апсолутна процентна грешка *(MDAPE)* и симетричната средна апсолутна процентуална грешка (sMAPE), имаат значително помали вредности кога се користи фиксна поделба на множествата за тренирање и тестирање. Исто така, од визуелниот приказ може да се заклучи дека грешките се помали кога се користат помали хоризонти за предвидување и тоа 20 – 50 – 100 чекори напред, со што се заклучува дека тој модел за предвидување е подобар.

6.4 Резиме на добиените резултати

Во докторската дисертација направена е анализа на резултатите добиени со испитувања во случај кога множествата за тренирање и тестирање се поделени на точно одреден датум на временската серија, и втор случај кога случајно се избира кои податоци ќе се земат во множеството за тренирање и тестирање со функцијата *sample.split()* од библиотеката *caTools*, при предвидување на променливоста на пазарната цена на биткоинот.

Во првиот случај испитувани се резултатите при користење на различен број на дрва во алгоритамот и тоа за 25, 50 и 100. При ова испитување користена е поделба на множеството за тренирање и тестирање во соднос 90 % : 10 %.

Од графиконот на слика 49 може да се забележи дека бројот на дрва не влијае многу на прикажаните грешки. Дополнително, може да се заклучи дека помала грешка за моделот се добива ако се користи фиксна поделба на множеството за тестирање и тренирање, и тоа *MAE* = 0.009, а *RMSE* = 0.011.

Од визуелниот приказ на слика 50 може да се заклучи дека апсолутната процентуална грешка (*MAPE),* просечната апсолутна процентна грешка *(MDAPE)* и симетричната средна апсолутна процентуална грешка *(sMAPE)* имаат значително помали вредности кога се користи поделба на множествата за тренирање и тестирање во точно одредена точка на временската серија. Се заклучува дека тој модел за предвидување е подобар.

Во следната ситуација направена е анализа на резултатите добиени со користење на различна поделба на множествата за тренирање и тестирање, и тоа за сооднос 90 % : 10 %, 80 % : 20 % и 66,6 % : 33,3 %.

Од графиконот на слика 54 се заклучува дека средната апсолутна грешка (*MAE*) и средната квадратна грешка (*RMSE*) најмали вредности имаат кога сообносот на поделба на множеството за тренирање и тестирање е 90 % : 10 % со фиксна поделба на податоците од временската серија *MAE* = 0,009, а *RMSE* = 0.011.

Од податоците прикажани со графиконот на слика 55, може да се заклучи дека апсолутната процентуална грешка (*MAPE)* и просечната апсолутна процентна грешка *(MDAPE)* имаат значително помали вредности кога се користи фиксна поделба на

множествата за тренирање и тестирање и избран сооднос 66,6 % : 33.3 % и со што се заклучува дека тој модел за предвидување е подобар.

Исто така, анализирани се резултатите добиени со користење на различни хоризонти за предвидување и тоа со 20 – 50 – 100 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 90 % : 10 %. Потоа, анализирана е ситуација со 7 – 30 – 60 – 180 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 83,5 % : 16,5 %. Следно е разгледана трета ситуација со 10 – 30 – 90 – 300 чекори напред со користење на поделба на множествата за тренирање и тестирање во сооднос 73 % : 27 %.

Од визуелно прикажаните податоци на слика 59 може да се заклучи дека вредностите на грешките со кои го оценуваме моделот малку се разликуваат. Во овој случај помали грешки се добиени кога поделбата на податоци во множествата за тренирање и тестирање е направена на фиксен датум. Најмали вредности се добиени кога се користени хоризонти за предвидување 20 – 50 – 100 и сооднос на поделба 90 % : 10 % и тоа *MAE* = 0.009, а *RMSE* = 0.011.

Од графиконот на слика 60, може да се заклучи дека апсолутна процентуална грешка (*MAPE),* просечната апсолутна процентна грешка *(MDAPE)* и симетричната средна апсолутна процентуална грешка (*sMAPE*), имаат значително помали вредности кога се користи фиксна поделба на множествата за тренирање и тестирање. Исто така, од визуелниот приказ може да се заклучи дека грешките се помали кога се користат помали хоризонти за предвидување и тоа 20 – 50 – 100 чекори напред, со што се заклучува дека тој модел за предвидување е подобар.

Скоро во сите анализирани ситуации помали грешки се добиваат кога базата на податоци што се користи како влез, е поделена во точно одреден датум на временската серија за да ги добиеме множеството за тренирање и множеството за тестирање.

7 Заклучок

Технологијата на блоковски вериги обезбедува нов начин на размислување за тоа како да се постигне консензус во различни ситуации. За прв пат, повеќе недоверливи страни можат да создадат и договорат еден извор на вистина, без учество на посредници. Напредокот на оваа технологијата тврди и има потенцијал да револуционизира многу области на човековата активност, особено во финансискиот и деловниот свет.

Со користење на технологијата на блоковски вериги се предлага систем за електронски трансакции без посредници. Традиционалниот систем е нецелосен бидејќи не постои начин да се спречи двојното трошење. За да се реши ова, постои мрежа на рамноправен систем (peer-to-peer), која користи алгоритам „доказ за работа“ за чување на јавна историја на трансакции. За напаѓачот да може да ги промени јазлите е пресметковно речиси невозможно ако искрените јазли контролираат мнозинство од мрежата. Сите јазли работат истовремено со мала координација. Јазлите можат да ја напуштат и повторно да се приклучат на мрежата по желба, прифаќајќи ја целосната блоковска верига како доказ за тоа што се случило додека ги немало. Тие гласаат со моќта на процесорот, изразувајќи го своето прифаќање на валидни блокови, работејќи на нивно проширување и отфрлајќи ги неважечките блокови со одбивање да работат на нив. Сите потребни правила и стимулации може да се применат со овој механизам за консензус.

Ако се погледне како се развивала технологијата во изминативе петнаесет години, се забележува дека оваа технологија е спротивна на логиката на обработката во облак (cloud computing). Кај обработката во облак може да пристапат повеќе јазли кон една база на податоци. Овие јазли не мора да имаат свој приватен примерок од базата на податоци. Понатаму, јазлите кои чуваат копии од блоковската верига добиваат постојани надградби. Овие јазли се дистрибуираат низ целиот свет. Поради тоа, блоковските вериги имаат висока латентност (латентноста го прикажува времето што е потребно за пренос на податоци преку мрежата). Затоа се јавуваат проблеми со скалирање. Биткоин може да процесира околу 7 трансакции во секунда. Етериум излегува со околу 20 трансакции во секунда, додека пак, Виза може да процесира до 1700 трансакции во секунда [81].

Технологијата на блоковски вериги би можела да ја револуционизира основната технологија на платниот систем и кредитните информациски системи во банките, со што ќе ги надградат и трансформираат. Апликациите за оваа технологија, исто така, промовираат формирање на повеќецентрични, слабо посредувани сценарија, кои ќе ја подобрат ефикасноста на банкарската индустрија. Вреди да се напомене дека проблемите со регулативата, ефикасноста и безбедноста отсекогаш предизвикувале опсежна дебата во процесот на секоја нова финансиска иновација. Историјата не е запрена од сегашните пречки, бидејќи техничките, регулаторните и другите проблеми

на технологијата на блоковски вериги на крајот ќе бидат решени. Оттука, изгледите за интегрирање на оваа технологијата во банкарската индустрија најверојатно ќе се појават наскоро. Оваа технологија може да помогне на многу начини преку намалени трансакциски трошоци со користење на паметен договор кој е вграден за општа цел за поедноставување на процедурите, намалување на административните оптоварувања и отстранување на посредниците.

Биткоинот е популарна криптовалута и е широко истражувана во областа на економијата и компјутерската наука. Во докторскава дисертација се користеше алгоритам за машинско учење случајни шуми (random forests) за да се предвидат временските серии на реализирани флуктуации на берзанската цена на биткоинот и да се испита дали информациите за блоковска верига може да се користат за предвидување на променливоста и цената на биткоинот. Многу луѓе во светот го користат биткоинот како инвестиција поради неговата висока променливост и на овој начин можат да добијат огромни профити, но и загуби за кратко време.

Од гледна точка на машинско учење со консензус, алгоритамот случајна шума се смета за еден од најдобрите алгоритми што предвидува многу добро со висока точност. Во R-пакетот во моделот што е направен во оваа дисертација е користена библиотеката *forecastML*, која содржи многу визуелизации на предвидувањата на моделите на временски серии.

Во емпириската постапка оваа дисертација, во основа, е поделена на три главни дела. Во првиот дел се користи вгнездена вкрстена валидација, користејќи 10 прозорци за валидација, предвидување за множество за обука и предвидување за множество за тестирање. Од пресметаните грешки прикажани во табелите 6 и 7 може да се заклучи дека грешката не е висока и покрај фактот што се користат прозорци за валидација. Исто така, се заклучува дека како што хоризонтот станува подолг, така се зголемува и грешката. Тој резултат е сосема нормално да се добие, но може да се забележи и дека дополнителната грешка не е висока, што значи дека може да се направат точни предвидувања и за подолги хоризонти.

Во вториот дел од испитувањата не се користат прозорци за валидација, па моделот се обучува користејќи го целото множество за тренирање. На овој начин, всушност, се добива моделот што одговара. Тоа се користи за да се комплетира обуката за моделот и конечното предвидување. Метриките на грешки при користење на податоците во множеството за обука и податоците во множеството за тестирање се прикажани на слика 40 и слика 42, соодветно. Се забележува дека точноста е поголема во случај кога не се користат прозорци за валидација. Тоа се очекува, точноста е многу поголема бидејќи моделот е трениран со целиот примерок за обука, така што способноста за предвидување е подобра. Грешките во споредба со метриката за грешки од првиот дел, каде што се користат прозорци за валидација, се значително помали. Понатаму, грешката на множеството за тестирање во споредба со множеството за

обука не е многу поголема, така што алгоритамот има способност точно да предвидува податоци што не се вклучени при обуката. Конечно, алгоритамот има способност да ги комбинира прогнозните хоризонти за да има подобар конечен резултат и, особено, за долгите хоризонти, каде грешката е поголема.

Во емпириската постапка, на почетокот избрани се 200 дрва, како максимален број дрва што се користат во алгоритамот при обуката. Од слика 29 може да се заклучи дека колку е поголем бројот на дрва, грешката станува помала. Но, грешката не станува значително помала во споредба со бројот на дрва што визуелно е прикажано на слика 29. Најмала грешка се добива ако се користат 186 дрва, но од графиконот кој има облик на лакт се заклучува дека по 50 дрва грешката се намалува многу малку, па затоа во ова истражување е избрано бројот на дрва да биде 50 [104]. На тој начин се намалува грешката, но, во исто време, нема голема сложеност во алгоритамот и процедурата за обука.

За предвидувањето се употребуваат методи за машинско учење за да има подобри и попрецизни резултати во споредба со традиционалното прогнозирање како авторегресивни, векторски авторегресивни методи итн. Тоа се докажува и со ниските грешки што се добиваат во работениот модел во тренинг- и тест-множеството. Во оваа дисертација се користени податоци за карактеристиките на биткоинот од http://www.blockchain.com/charts за последниве 3 години. Добиените емпириски резултати покажуваат мала грешка при предвидувањето, со што се заклучува дека е оправдано да се користат за испитување на цената и променливоста на биткоинот.

Според литературата, колку е помала просечната апсолутна грешка, толку е подобар моделот на предвидување. Во нашиот модел средната апсолутна грешка *(MAE)* во множеството за тренирање со користење на 10 прозорци за валидација е 0,017, а во тест-множеството, просечната апсолутна грешка од различни хоризонти е 0,009. Средната апсолутна грешка (*MAE)* во множеството за обука без прозорци за валидација е 0,006, а во множеството за тестирање е 0,009. Може да заклучи дека ако во реалните податоци променливоста на биткоинот е голема, тогаш грешката во предвидувањето ќе биде поголема. Од прикажаниот пример може да се заклучи дека грешката во предвидувањето на податоците е помала кога хоризонтот на прогнозата е пократок и кога променливоста е помала.

Во третиот дел од испитувањата направени се анализи на различни ситуации во кои дополнително се мери средната квадратна грешка *(RMSE)* за правилно да се оцени точноста на моделот за предвидување. Прво се испитувани ситуации во кои се земаат различен број на дрва и тоа 25, 50 и 100. Секој пример е испитуван прво во фиксна поделба на временската серија, а потоа со случајна поделба на множеството за тренирање и тестирање. Испитувани се статистичките грешки *MAE, RMSE, MAPE, MDAPE, sMAPE* за сите 4 случаи: множество за тренирање со прозорци за валидација, множество за тестирање со прозорци за валидација, множество за тренирање без

прозорци за валидација и множество за тестирање без прозорци за валидација. Од прикажаните резултати се заклучува дека *MSE* и *RMSE* не се големи и малку се разликуваат при испитување со користење на различен број на дрва. Исто така, се заклучува дека помала грешка за моделот се добива ако се користи фиксна поделба на множествата за тестирање и тренирање, и тоа *MAE* = 0.009, а *RMSE* = 0.011, при користење на 25 и 50 дрва.

Следен чекор во испитувањата се ситуации во кои се зема различен сооднос при поделбата на множествата за тренирање и тестирање и тоа 90 % : 10 %, 80 % : 20 % и 66,6 % : 33,3 %. Секој пример е испитуван прво во фиксна поделба на временската серија, а потоа со случајна поделба на множествата за тренирање и тестирање. Испитувани се статистичките грешки *MAE, RMSE, MAPE, MDAPE, sMAPE* за сите 4 случаи, множества со прозорци за валидација и множества без прозорци за валидација. Од прикажаните резултати се заклучува дека *MSE* и *RMSE* имаат мали вредности, што укажува на фактот дека моделот за предвидување на променливоста на биткоинот е добар. Од визуелниот приказ може да се заклучи дека според вредноста на грешките за евалуација на моделот најдобро е да се користи поделба на множеството податоци во сооднос 90 % : 10 % и тоа со фиксна поделба на временската серија.

Во следниот дел од емириската постапка правени се испитувања каде се земаат различни вредности за хоризонтите за предвидување со вредности: 20 – 50 – 100, следно 7 – 30 – 60 – 180 и трета група 10 – 30 – 90 – 300 за сите 4 случаи, множества со прозорци за валидација и множества без прозорци за валидација. Од добиените резултати се заклучува дека сите грешки за евалуација имаат задоволителни вредности. Помали вредности се добиени со помали вредности на хоризонтите, што е очекувано. И во овој случај помали се грешките добиени кога се користи поделба на временската серија на множество за тренирање и тестирање со фиксен датум и затоа се препорачува да се користи таков модел при регресиона анализа. Со направените испитувања може да се каже дека технологијата на блоковски вериги успешно може да се користи за подобрување на финансиските сервиси, како што е предвидувањето на променливоста на биткоинот, за што секојдневно се интересираат сè повеќе луѓе што сакаат да инвестираат во оваа криптовалута.

Во оваа дисертација моделирна е променливоста на пазарната цена на биткоинот како основа за мерење на факторот на ризик кај финансиските сервиси со употреба на технологијата на блоковски вериги. Со предвидувањето на промената на вредноста на биткоинот се подобрува работењето на финансиските сервиси, се намалува факторот на ризик при инвестирање, работа на берзи, штедење и сл.

Овој модел може да биде корисен и за откривање на аномалии и измамнички активности во финансиското работење. Кога вистинското однесување на цената на криптовалутата значително се менува од моделираното однесување, тоа може да укаже

на ефектот на надворешни фактори како што се големите глобални настани, како и измамнички активности.

Во понатамошно истражување може да се испита дали има некои макроекономски или финансиски променливи и индекси, кои влијаат на променливоста на биткоинот. Во овој труд избран е еден специфичен алгоритам за машинско учење, случајни шуми, за да се предвидат временските серии на реализираната променливост на биткоинот. Истата постапка може да се направи со користење на друг алгоритам за машинско учење, како што се невронските мрежи, машините со носечки вектори, логистичка регресија, ласо, регресија на *k*-најблиски соседи итн. Во понатамошните истражувања, може да се испита кој од овие алгоритми предвидува со поголема точност. Различни типови на променливост може да се испитаат како зависни променливи на моделот, или различни видови методологија во која предвидувањето нема да биде временска серија, т.е. регресија, туку класификација каде што предвидувањето се прави со користење на категорична променлива за зголемување или намалување.

Ова истражување претставува само почеток на низа истражувања во областа на технологијата на блоковски вериги во комбинација со алгоритмите за машинско учење за подобрување на перформансите на финасиските сервиси.

Користена литература

[1] Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." *Decentralized Business Review* (2008): 21260.

[2] Dragana Tadić Živković, blockchain technology: opportunity or a threat to the future development of banking, Proceedings of Ekonbiz, 2018.

[3] Stefano Tempesta, Decentralized Applications with Azure Blockchain as a Service, [Internet] URL:https://msdn.microsoft.com/magazine/mt847188?MC=MSAzure&MC=CloudDev&MC=IoT&MC=MachLearn&MC=Vstudio (пристапено на 19.07.2022).

[4] Ghosh, Debraj. "How the Byzantine General Sacked the Castle: A Look Into Blockchain, (2016).", [Internet] URL: https://medium.com/@DebrajG/how-the-byzantine-general-sacked-the-castle-a-look-into-blockchain-370fe637502c (пристапено на 19.07.2022).

[5] Aleksandar Matanović, "Osnove kriptovaluta i blokčein tehnologije", [Internet] URL: http://fzp.singidunum.ac.rs/demo/wp-content/uploads/Osnove-kriptovaluta-i-blok%C4%8Dein-tehnologije.pdf, (пристапено на 06.06.2022).

[6] Yevgeniy Brikman, Bitcoin by analogy, [Internet] URL: https://www.ybrikman.com/writing/2014/04/24/bitcoin-by-analogy (пристапено на 19.07.2022).

[7] Jonathan Waldman, Blockchain - Blockchain Fundamentals, [Internet] URL: https://msdn.microsoft.com/magazine/mt845650 (пристапено на 19.07.2022).

[8] Rauchs, Michel, Andrew Glidden, Brian Gordon, Gina C. Pieters, Martino Recanatini, François Rostand, Kathryn Vagneur, and Bryan Zheng Zhang. "Distributed ledger technology systems: A conceptual framework." Available at SSRN 3230013 (2018).

[9] Lutpin, Dos and Don'ts of Peer-to-Peer Trading, [Internet] URL: https://www.crypto-news.net/dos-and-donts-of-peer-to-peer-trading (пристапено на 19.07.2022).

[10] Tasatanattakool, Pinyaphat, and Chian Techapanupreeda. "Blockchain: Challenges and applications." In 2018 International Conference on Information Networking (ICOIN), pp. 473-475. IEEE, 2018.

[11] Capgemini, Blockchain and Capgemini SAP Leonardo – part 1, [Internet] URL: https://www.capgemini.com/gb-en/2018/01/blockchain-and-capgemini-sap-leonardo-part-1 (пристапено на 19.07.2022).

[12] Hozjan, Domina. "Blockchain." PhD diss., University of Zagreb. Faculty of Science. Department of Mathematics, 2017.

[13] Szabo, Nick. "Smart contracts: building blocks for digital markets." EXTROPY: The Journal of Transhumanist Thought, (16) 18, no. 2 (1996): 28.

[14] Lukić, Velimir. "POTENCIJALI I OGRANIČENJA PRIVATNIH DIGITALNIH VALUTA/ POTENTIALS AND LIMITS OF PRIVATE DIGITAL CURRENCIES.", [Internet] URL: http://www.ekof.bg.ac.rs/wp-content/uploads/2016/03/Seminar-katedre-2017-Potencijali-i-ograni%C4%8Denja-privatnih-digitalnih-valuta-PDF.pdf, (пристапено на 19.07.2022).

[15] MUTAWAKKIL, Online MD5 Hash Generator & SHA1 Hash Generator, [Internet] URL: https://onlinemd5-com.mutawakkil.com (пристапено на 19.07.2022).

[16] CBINSIGHTS, What is blockchain technology, [Internet] URL: https://www.cbinsights.com/research/what-is-blockchain-technology (пристапено на 19.07.2022).

[17] Prirodoslovno-matematički fakultet, matematički odsjec, Sveučilište u Zagrebu, Osnovni pojmovi, Definicija eliptičke krivulje, [Internet] URL: https://web.math.pmf.unizg.hr/~duje/ecc/elipdef.html, (пристапено на 19.07.2022).

[18] Rainie, Lee, and Maeve Duggan "Privacy and information sharing." (2016).

[19] Zyskind, Guy, and Oz Nathan. "Decentralizing privacy: Using blockchain to protect personal data." In 2015 IEEE Security and Privacy Workshops, pp. 180-184. IEEE, 2015.

[20] Karafiloski, Elena, and Anastas Mishev. "Blockchain solutions for big data challenges: A literature review." In IEEE EUROCON 2017-17th International Conference on Smart Technologies, pp. 763-768. IEEE, 2017.

[21] Chakravorty, Antorweep, and Chunming Rong. "Ushare: user controlled social media based on blockchain." In Proceedings of the 11th international conference on ubiquitous information management and communication, pp. 1-6. 2017.

[22] McConaghy, Trent, and David Holtzman. "Towards an ownership layer for the internet." ascribe GmbH (2015).

[23] Rosenzweig, Roy. "The road to Xanadu: Public and private pathways on the history web." The Journal of American History 88, no. 2 (2001): 548-579.

[24] Dimitri de Jonghe, "SPOOL Protocol", [Internet] URL: https://github.com/ascribe/spool, (пристапено на 19.07.2022).

[25] McConaghy, Trent, Rodolphe Marques, Andreas Müller, Dimitri De Jonghe, Troy McConaghy, Greg McMullen, Ryan Henderson, Sylvain Bellemare, and Alberto Granzotto. "Bigchaindb: a scalable blockchain database." white paper, BigChainDB (2016).

[26] Swan, Melanie. Blockchain: Blueprint for a new economy. " O'Reilly Media, Inc.", 2015.

[27] "PRS for Music takes legal action against SoundCloud streaming service", The Guardian, 2015.

[28] Hall, Christopher, and Casey Alt. "Lê Quý Quốc Cường, and Sean Moss-Pultz (2016)." Bitmark: The property system for the digital environment.

[29] Lin, Tzu-Yun, Yu-Chiang Frank Wang, and Sean Moss-Pultz. "ObjectMinutiae: Fingerprinting for Object Authentication." In Proceedings of the 23rd ACM international conference on Multimedia, pp. 815-816. 2015.

[30] Federal Trade Commission. "Internet of things: Privacy & security in a connected world." Washington, DC: Federal Trade Commission (2015).

[31] Pureswaran, Veena, and Paul Brody. "Device democracy: Saving the future of the Internet of Things." IBM Corporation (2015): 23.

[32] Buterin, Vitalik. "A next-generation smart contract and decentralized application platform." white paper 3, no. 37 (2014): 2-1.

[33] Jentzsch, Christoph. "Decentralized autonomous organization to automate governance." White paper, November (2016).

[34] Dorri, Ali, Salil S. Kanhere, and Raja Jurdak. "Blockchain in internet of things: challenges and solutions." arXiv preprint arXiv:1608.05187 (2016).

[35] Tal Rapke, MD “Blockchain Technology & the Potential for Its Use in Healthcare”, 2016

[36] Gupta, Nitesh, Anand Jha, and Purna Roy. "Adopting blockchain technology for electronic health record interoperability." Cognizant Technol. Solutions, Teaneck, NJ, USA, White Paper (2016).

[37] Linn, Laure A., and Martha B. Koo. "Blockchain for health data and its potential use in health it and health care related research." In ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST, pp. 1-10. 2016.

[38] Ekblaw, Ariel, Asaph Azaria, John D. Halamka, and Andrew Lippman. "A Case Study for Blockchain in Healthcare: “MedRec” prototype for electronic health records and medical research data." In Proceedings of IEEE open & big data conference, vol. 13, p. 13. 2016.

[39] Windley, Phillip J. *Digital Identity: Unmasking identity management architecture (IMA)*. " O'Reilly Media, Inc.", 2005.

[40] TechGenix (2018). “Blockchain technology: Why it will change the world”, [Internet], URL: http://techgenix.com/blockchain-technology, (пристапено на 06.06.2022).

[41] Širić, Mario. "BLOCKCHAIN TEHNOLOGIJA I NJEN UTJECAJ NA SVIJET." PhD diss., University of Split. Faculty of economics Split, 2018.

[42] Object Computing (2017). “8 ways blockchain is changing the world”, [Internet], URL: https://objectcomputing.com/news/2017/12/20/8-ways-blockchain-changing-world, (пристапено на 06.06.2022).

[43] Blockchain Expo (2018). “How will blockchain impact the insurance sector?”, [Internet], URL: https://www.blockchain-expo.com/2018/02/blockchain/blockchain-insurance, (пристапено на 06.06.2022).

[44] Willie, Paul. "Can all sectors of the hospitality and tourism industry be influenced by the innovation of blockchain technology?" Worldwide Hospitality and Tourism Themes (2019).

[45] Roopak, T. M., and R. Sumathi. "Electronic voting based on virtual id of aadhar using blockchain technology." In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pp. 71-75. IEEE, 2020.

[46] Kripto-online, Šta-je-blockchain, [Internet], URL: https://kripto-online.info/blockchain/sta-je-blockchain, (пристапено на 06.06.2022).

[47] AXA, AXA goes blockchain with fizzy, [Internet], URL: https://www.axa.com/en/magazine/axa-goes-blockchain-with-fizzy, (пристапено на 06.06.2022).

[48] Gu, Jingjing, Binglin Sun, Xiaojiang Du, Jun Wang, Yi Zhuang, and Ziwang Wang. "Consortium blockchain-based malware detection in mobile devices." IEEE Access 6 (2018): 12118-12128.

[49] Androulaki, Elli, Artem Barger, Vita Bortnikov, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart et al. "Hyperledger fabric: a distributed operating system for permissioned blockchains." In Proceedings of the thirteenth EuroSys conference, pp. 1-15. 2018.

[50] Aggarwal, Shubhani, and Neeraj Kumar. "Hyperledger." In Advances in computers, vol. 121, pp. 323-343. Elsevier, 2021.

[51] Minović, Miroslav. "Blockchain technology: usage beside cripto currencies." In Infotech 2017: ICT conference & exhibition, Arandjelovac, Serbia, June, pp. 7-8. 2017.

[52] Vujičić, Dejan, Dijana Jagodić, and Siniša Ranđić. "Blockchain technology, bitcoin, and Ethereum: A brief overview." In 2018 17th international symposium infoteh-jahorina (infoteh), pp. 1-6. IEEE, 2018.

[53] Mijoska, Mimoza, and Blagoj Ristevski. "Blockchain Technology and its Application in the Finance and Economics." (2020): 197-202.

[54] Deloitte, Key challenges, [Internet], URL: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/Innovation/deloitte-uk-blockchain-key-challenges.pdf, (пристапено на 06.06.2022).

[55] Holotiuk, Friedrich, Francesco Pisani, and Jürgen Moormann. "The impact of blockchain technology on business models in the payments industry." (2017).

[56] Crosman, P. "R3 to take on Ripple with cross-border payments blockchain." American Banker (2017).

[57] Johnstone, Syren. "A viral warning for change. COVID-19 versus the red cross: Better solutions via blockchain and artificial intelligence." COVID-19 Versus the Red Cross: Better Solutions Via Blockchain and Artificial Intelligence (February 3, 2020). University of Hong Kong Faculty of Law Research Paper 2020/005 (2020).

[58] Géron, Aurélien. "Hands-on machine learning with scikit-learn and tensorflow: Concepts." Tools, and Techniques to build intelligent systems (2017).

[59] Molly Galetto, Machine learning and big data analytics the perfect marriage, [Internet], URL: http://www.ngdata.com/machine-learning-and-big-data-analytics-the-perfect-marriage (пристапено на 19.07.2022).

[60] Schneider, Astrid, Gerhard Hommel, and Maria Blettner. "Linear regression analysis: part 14 of a series on evaluation of scientific publications." Deutsches Ärzteblatt International 107, no. 44 (2010): 776.

[61] Tandel, Aakash. "Support Vector Machines—A Brief Overview–Towards Data Science." Retrieved May 26 (2017): 2019.

[62] Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.

[63] Pedregosa, e. a. (2011). Scikit-learn: Machine Learning in Python. Retrieved from Naive Bayes, [Internet], URL: http://scikit-learn.org/stable/modules/naive_bayes.html, (пристапено 19.07.2022).

[64] Ali, Jehad, Rehanullah Khan, Nasir Ahmad, and Imran Maqsood. "Random forests and decision trees." International Journal of Computer Science Issues (IJCSI) 9, no. 5 (2012): 272.

[65] Zaderej, Victor Vasyl. "The use of neural networks to reduce process variability." PhD diss., Quinnipiac University, 1995.

[66] Yuen, Brosnan, Minh Tu Hoang, Xiaodai Dong, and Tao Lu. "Universal activation function for machine learning." *Scientific reports* 11, no. 1 (2021): 1-11.

[67] Pratiwi, Heny, Agus Perdana Windarto, S. Susliansyah, Ririn Restu Aria, Susi Susilowati, Luci Kanti Rahayu, Yuni Fitriani, Agustiena Merdekawati, and Indra Riyana Rahadjeng. "Sigmoid activation function in selecting the best model of artificial neural networks." In *Journal of Physics: Conference Series*, vol. 1471, no. 1, p. 012010. IOP Publishing, 2020.

[68] Liu, Xinyu, and Xiaoguang Di. "TanhExp: A smooth activation function with high convergence speed for lightweight neural networks." *IET Computer Vision* 15, no. 2 (2021): 136-150.

[69] Kotsiantis, Sotiris B., Ioannis D. Zaharakis, and Panayiotis E. Pintelas. "Machine learning: a review of classification and combining techniques." Artificial Intelligence Review 26, no. 3 (2006): 159-190.

[70] Marko Čupić, Umjetna inteligencija, Uvod u strojno učenje, 2020, [Internet], URL: http://java.zemris.fer.hr/nastava/ui/ml/ml-20200410.pdf, (пристапено 19.07.2022).

[71] Αριστείδου, Χριστόφορος. "Study of the volatility of bitcoin cryptocurrency using machine learning methods: an implementation in R." (2020).

[72] Ho, Tin Kam. "Random decision forests." In Proceedings of 3rd international conference on document analysis and recognition, vol. 1, pp. 278-282. IEEE, 1995.

[73] Gemini, Gemini Exchange Data, [Internet], URL: https://www.cryptodatadownload.com/data/gemini, (пристапено 06.06.2022).

[74] Daniel Johnson, R Random Forest Tutorial with Example, [Internet], URL: https://www.guru99.com/r-random-forest-tutorial.html, (пристапено 06.06.2022).

[75] Ross Jacobucci, Random Forests, University of Notre Dame, [Internet], URL: https://statisticalhorizons.com/wp-content/uploads/2021/11/Advanced-Machine-Learning.pdf, (пристапено 06.06.2022).

[76] Srinath Sridharan, Predictive modeling with Machine Learning in R, [Internet], URL: https://srinath-sridharan.medium.com/predictive-modeling-with-machine-learning-in-r-part-2-evaluation-metrics-for-classification-d6591749a6, (пристапено 19.07.2022).

[77] Willmott, Cort J., and Kenji Matsuura. "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance." *Climate research* 30, no. 1 (2005): 79-82.

[78] Khair, Ummul, Hasanul Fahmi, Sarudin Al Hakim, and Robbi Rahim. "Forecasting error calculation with mean absolute deviation and mean absolute percentage error." In *Journal of Physics: Conference Series*, vol. 930, no. 1, p. 012002. IOP Publishing, 2017.

[79] Kreinovich, Vladik, Hung T. Nguyen, and Rujira Ouncharoen. "How to estimate forecasting quality: A system-motivated derivation of symmetric mean absolute percentage error (SMAPE) and other similar characteristics." (2014).

[80] R-bloggers, Machine learning explained overfitting, [Internet], URL: https://www.r-bloggers.com/2017/06/machine-learning-explained-overfitting, (пристапено 06.06.2022).

[81] Kai Sedgwick, No visa doesn't handle 24000 tps and neither does your pet blockchain, [Internet], URL: https://news.bitcoin.com/no-visa-doesnt-handle-24000-tps-and-neither-does-your-pet-blockchain, (пристапено 19.07.2022).

[82] Esteban Ordano et al., Decentraland, A blockchain-based virtual world [Internet], URL: https://docs.decentraland.org/decentraland/whitepaper (пристапено 06.06.2022).

[83] Avijeet Biswal, The Complete Guide on Overfitting and Underfitting in Machine Learning, [Internet], URL: https://www.simplilearn.com/tutorials/machine-learning-tutorial/overfitting-and-underfitting#what_is_overfitting, (пристапено 19.07.2022).

[84] Horton, Bob. "Calculating auc: the area under a roc curve." (2016).

[85] ISG, Blockchain: Success Lies in Identifying the Right Use Cases, [Internet], URL: https://isg-one.com/research/articles/full-article/blockchain-success-lies-in-identifying-the-right-use-cases, (пристапено 19.07.2022).

[86] Zhang, Li, Yongping Xie, Yang Zheng, Wei Xue, Xianrong Zheng, and Xiaobo Xu. "The challenges and countermeasures of blockchain in finance and economics." Systems Research and Behavioral Science 37, no. 4 (2020): 691-698.

[87] Zheng, Zibin, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang. "An overview of blockchain technology: Architecture, consensus, and future trends." In 2017 IEEE international congress on big data (BigData congress), pp. 557-564. Ieee, 2017.

[88] James A. Chambers, My Early Bitcoin ASIC Miners (2013-2014) – Pictures / History, [Internet], URL: https://jamesachambers.com/early-bitcoin-asic-miner-pictures-history, (пристапено 19.07.2022).

[89] Delton Rhodes, What is Delegated Proof of Stake? An Overview of DPoS Blockchains, [Internet], URL: https://komodoplatform.com/en/academy/delegated-proof-of-stake, (пристапено 19.07.2022).

[90] MinerGate, Blockchain Consensus Types: Proof of Elapsed Time, Proof of Authority, Proof of Bandwidth, [Internet], URL: https://minergate.com/blog/blockchain-consensus-types-proof-of-elapsed-time-proof-of-authority-proof-of-bandwidth, (пристапено 19.07.2022).

[91] Steven Buchko, What Is Blockstack (STX)? | The First SEC-Qualified Token Offering, [Internet], URL: https://coincentral.com/blockstack-stx, (пристапено 19.07.2022).

[92] Wikipedia, Доменски именски систем, [Internet], URL: https://mk.wikipedia.org/wiki/Доменски_именски_систем, (пристапено 19.07.2022).

[93] Etutorials, Microsoft Products, Lesson 6: Understanding Active Directory Concepts, [Internet], URL: https://etutorials.org/Microsoft+Products/microsoft+windows+xp+professional+training+kit/Chapter+5+-

+Using+the+DNS+Service+and+Active+Directory+Service/Lesson+6nbspUnderstanding+Active+Directory+Concepts/ (пристапено 19.07.2022).

[94] De Filippi, Primavera, and Benjamin Loveluck. "The invisible politics of bitcoin: governance crisis of a decentralized infrastructure." Internet policy review 5, no. 4 (2016).

[95] Venkata Marella, Bitcoin: A Social Movement Under Attack, [Internet], URL: https://core.ac.uk/download/pdf/301373906.pdf, (пристапено 19.07.2022).

[96] Tashevski, Panche, et al. "Application of the Blockchain Technology in Medicine and Healthcare." (2021): 65-69.

[97] Samuel, Arthur L. "Some studies in machine learning using the game of checkers. II—Recent progress." IBM Journal of research and development 11.6 (1967): 601-617.

[98] Mitchell, Tom M., and Tom M. Mitchell. Machine learning. Vol. 1. No. 9. New York: McGraw-hill, 1997.

[99] Aponte-Novoa, Fredy Andres, Ana Lucila Sandoval Orozco, Ricardo Villanueva-Polanco, and Pedro Wightman. "The 51% attack on blockchains: A mining behavior study." IEEE Access 9 (2021): 140549-140564.

[100] Chen, Guang, Bing Xu, Manli Lu, and Nian-Shing Chen. "Exploring blockchain technology and its potential applications for education." Smart Learning Environments 5, no. 1 (2018): 1-10.

[101] Hayes, Adam S. "Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin." Telematics and informatics 34, no. 7 (2017): 1308-1321.

[102] Er-Rajy, L., A. El Kiram My, M. O. H. A. M. E. D. El Ghazouani, and O. Achbarou. "Blockchain: Bitcoin wallet cryptography security, challenges and countermeasures." Journal of Internet Banking and Commerce 22, no. 3 (2017): 1-29.

[103] Dewi, Christine, and Rung-Ching Chen. "Random forest and support vector machine on features selection for regression analysis." Int. J. Innov. Comput. Inf. Control 15, no. 6 (2019): 2027-2037.

[104] RDRR, forecastML: Time Series Forecasting with Machine Learning Methods, [Internet], URL: https://rdrr.io/cran/forecastML/man/create_lagged_df.html, (пристапено 19.07.2022).

[105] Zach, How to Build Random Forests in R (Step-by-Step), [Internet], URL: https://www.statology.org/random-forest-in-r, (пристапено 19.07.2022).

[106] Leo Breiman, Random Forest, Statistics Department University of California, Berkeley, 2001.

[107] Chai, Tianfeng, and Roland R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)." Geoscientific Model Development Discussions 7, no. 1 (2014): 1525-1534.

[108] Folkers, Casper. "Scalable machine learning algorithms on a big data infrastructure." (2016).

[109] Cash, Bitcoin. "Bitcoin cash." Development 2 (2019).

[110] Kriskó, Andrea. "Crypto currencies–currencies governed by belief Bitcoin, Piggycoin, Monero, Peercoin, Ethereum and the rest." In Conference book Konferenciakötet, p. 283. 2013.

[111] King, Sunny. "Primecoin: Cryptocurrency with prime number proof-of-work." July 7th 1, no. 6 (2013).

[112] Ramos, Daniel, and Gabriel Zanko. "A Review of Zcash as a Cryptocurrency Platform Aimed Towards Maintaining Privacy Between All Parties." (2021).

[113] Duffield, Evan, and Daniel Diaz. "Dash: A privacycentric cryptocurrency." (2015): 7- 8.

[114] Deniz, Asena, and Dilek Teker. "Determinants of Cryptocurrency Market: An Analysis for Bitcoin, Ethereum and Ripple." International Journal of Business and Social Science 11, no. 11 (2020).

[115] Siddiqui, Shiza. "Long Term Comovement among Cryptocurrencies: an Application of Cointegration Analysis." (2020).

[116] Pîrjan, Alexandru, Dana-Mihaela PETROŞANU, Mihnea Huth, and Mihaela NEGOIȚĂ. "RESEARCH ISSUES REGARDING THE BITCOIN AND ALTERNATIVE COINS DIGITAL CURRENCIES." Journal of Information Systems & Operations Management 9, no. 1 (2015).

[117] Jacobs, Frederic. "Providing better confidentiality and authentication on the Internet using Namecoin and MinimaLT." arXiv preprint arXiv:1407.6453 (2014).

[118] Kasireddy, Preethi. "How does Ethereum work, anyway." Medium (2017).

[119] Leo Breiman, Adele Cutler, Random forests, [Internet], URL: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

[120] Branko Žitko, Upotreba stabla odlučivanja kod testiranja znanja metodom kviza, [Internet], URL: https://bib.irb.hr/datoteka/145853.Upotreba_stabla_odlucivanja_u_testiranju_znanja_pomocu_kviza.pdf