

УНИВЕРЗИТЕТ У БЕОГРАДУ
ФАКУЛТЕТ ОРГАНИЗАЦИОНИХ НАУКА

ЗАВРШНИ РАД

**Тема: Предвиђање цене Bitcoin-а помоћу
неуронских мрежа на основу тржишног
сентимента**

Ментор:

Доц. др Ивана Драговић

Студент:

Матија Милекић 161/14

Београд, 2018. године

Изјава захвалности:

Белику захвалност дугујем Александру Ракићевићу, асистенту на Катедри за управљање системима, на Факултету организационих наука, који ми је помогао својим руководством приликом писања овог завршног рада, неким чудом имао стрпљења до самог краја за сва моја питања и захтеве док сам програмирао скрипту.

Поред тога бих хтео да се захвалим ауторима Ракићевић, А., Миловановић А., и Аничич Р. који су ми омогућили приступ кодовима из њиховог истраживања „*An application of neural networks and fundamental analysis for automated trading: Belgrade stock exchange case*“ објављеног на конференцији „*Symorg 2016: Reshaping the Future Through Sustainable Business Development and Entrepreneurship*“

Такође, посебну захвалност дугујем и својој тетки, која ми је урадила целу граматичку рецензију овог завршног рада како би други рецензенти могли да се посвете само на суштину, а не на моје правописне и академске грешке.

Апстракт

Предмет овог рада је креирање система за алгоритамско трговање на тржишту криптовалута. Систем је базиран на вештачким неуронским мрежама, техничкој и сентимент анализи. Циљ овог рада је да се истражи да ли постоји веза између сентимента откривеног на друштвеној мрежи Reddit и промене цене код Bitcoin-а. Систем ће покушати да покаже везу ова два и то на периоду од годину дана (01.08.2017.-01.08.2018.), на основу дневних података који су прикупљени са интернета. Симулације трговања су урађене више пута како би се добило кредибилно решење. У резултатима се може видети да је систем успео да оствари позитиван просечан принос што нам може донекле потврдити нашу претпоставку да је психологија тржишта битан фактор на овом тржишту.

Креирање система, обучавање и тестирање мреже су извршени у развојном окружењу и програмском језику МАТЛАБ, док је скрипта за извлачење података писана уз помоћ Python програмског језика. Подаци су преузети са друштвене мреже Reddit и сајта Coinmarketcap.

Кључне речи: Вештачке неуронске мреже, предвиђање, алгоритамско трговање, анализа сентимента, Bitcoin, Reddit

Садржај

1. Увод.....	1
2. Теоријске основе.....	3
2.1. Неуронске мреже.....	3
2.1.1. Увод у неуронске мреже.....	3
2.1.2. Структура неуронских мрежа.....	3
2.1.3. Типови архитектуре неуронских мрежа.....	4
2.1.4. Примена неуронских мрежа.....	7
2.2. Сентимент анализа.....	9
2.2.1. Основна идеја сентимент анализе.....	9
2.2.2. Фазе сентимент анализе.....	10
2.2.3. Типови сентимент анализе.....	10
2.2.4. Данашњи допринос.....	11
2.3. Bitcoin.....	12
2.3.1. Историјат.....	12
2.3.2. Флукутација вредности.....	14
2.3.3. Експертска анализа.....	20
2.3.4. Генерално запажање.....	22
3. Предложени систем.....	23
3.1. Предпроцесирајућа компонента.....	24
3.2. Компонента сентимент анализе.....	27
3.3. Компонента логике трговања.....	29
4. Резултати симулације.....	30
4.1. Подаци.....	30

4.2. Параметри система.....	32
4.2.1. Параметри стратегије.....	32
4.2.2. Параметри система.....	32
4.3. Мере перформанси.....	33
4.3.1. Резултати добијени коришћењем улазног скупа 1: сентимент анализа.....	34
4.3.2. Резултати добијени коришћењем улазног скупа 2: сентимент анализа и Bitcoin доминација.....	35
4.3.3. Резултати добијени коришћењем улазног скупа 3: сентимент анализа, Bitcoin доминација и техничка анализа.....	37
4.3.4. Резултати добијени коришћењем улазног скупа 4: упросечена сентимент анализа.....	39
4.3.5. Резултати добијени коришћењем улазног скупа 5: упросечена сентимент анализа и просечна Bitcoin доминација.....	40
4.3.6. Резултати добијени коришћењем улазног скупа 6: упросечена сентимент анализа , просечна Bitcoin доминација и упросечена техничка анализа.....	42
4.4. Упоредна анализа резултата.....	44
5. Закључак.....	46
6. Литература.....	48
Додатак.....	51
Увоз библиотека које су потребне да би ова скрипта могла да ради.....	51
Дефинисање почетка и краја првог дана од ког се креће анализа.....	51
Иницијализација свих листи које су потребне за чување података.....	51
Увоз података из текстуалних датотека.....	51
Кодирање петље у којој се шаље парамтризован захтев ка сервису, након чега се одвајају подаци из одговора самог сервиса.....	52
Сентимент анализа података на основу речи из датотека.....	53

Трансформација датума из формата <i>timestamp</i> у <i>datetime</i>	54
Попуњавање листи са одговарајућим подацима.....	54
Обавештавање корисника да је завршена анализа за тај дан и итерирање на следећи.....	54
Извоз података из скрипте у табелу.....	55

1. Увод

Промене вредности Bitcoin-а у последњих годину дана су привукле велику пажњу како инвеститора тако и академске заједнице. Основно питање којим се аналитичари баве јесте да ли ће вредност биткоина пасти или порасти, односно, да ли треба купити или продати токене. Постоје бројна истраживања, а на неким се и данас ради, која покушавају да одговоре на ово питање. Велика волатилности и сама флукуација вредности Bitcoin-а створила је изузетну заинтересованост према овом тржишту. Због своје непредвидивости и немогућности да се нађе директна веза између промене цене и дешавања на тржишту, као нови аспект у овој анализи додали смо поверење друштвене заједнице у ову криптовалуту. Управо том аспектима ћемо посветити посебну пажњу у овом завршном раду. Овај рад представља проширење и надоградњу на [1], изложеном на конференцији Symorg. Разлика у односу на тај рад је то што је сада проширено истраживање са анализе друштвеног сентимента са једне главне објаве у том дану на тридесет најбољих. Поред тога, урађена је и сентимент анализа квалитативних података како би се боље схватиле објаве на друштвеној мрежи Reddit.

Трговање криптовалутама је постала веома актуелна тема у академској заједници, где велики број истраживача покушава да применом низа модерних метода рачунарске интелигенције и машинског учења предвиди будуће цене Bitcoin-а. С једне стране, стручњаци покушавају да истраже везу између великих трансакција појединаца и реакција активних инвеститара [2], а са друге, утицај расправа на друштвеним мрежама на формирање вредности саме криптовалуте [3-5]. У потрази за одговорима улази се у дубоку анализу коментара на друштвеним мрежама [6], али се праве и класе индикатора социјалних мрежа који се баве предвиђањем цене [7].

У овом раду коришћени су подаци преузети са друштвене мреже Reddit, над којима смо спровели сентимент анализу и на тај начин попунили податке који су у вези са окружењем. Тржишне податке смо проследили систему за трговање заједно са квантитативним и квалитативним подацима са друштвене мреже. На основу добијених података покушаћемо да одговоримо на питање да ли постоји веза између сентимента на друштвеним мрежама и будуће цене Bitcoin-а.

У другом поглављу овог рада су дате теоријске основе неуронских мрежа, сентимент анализе и криптовалуте Bitcoin који представљају окосницу овог рада. У делу који се бави неуронским мрежама представили смо увод у вештачке неуронске мреже, њену структуру, типове архитектуре неуронске мреже и саму примену у данашње време. Затим, приказали смо и објаснили шта је сентимент анализа, како она ради, фазе израде сентимент анализе и њене типове. Додатно је објашњено зашто нам је битна и шта нам може донети ова анализа. На крају теоријског дела је објашњен кратак историјат криптовалуте и њена флукуација кроз време. Пружена је експертска анализа и дато мишљење о самој будућности Bitcoin-а.

Треће поглавље обухвата сам изглед система, односно поделу и начин функционисања сваког његовог дела. Ту је представљена слика самог система и кораци који се извршавају како би се на самом крају аутоматски трговало.

Четврто поглавље садржи резултате свих симулације које су извршене при тестирању овог система и покушај решавања проблема. Решавању тог проблема је приступљено упоредном анализом свих резултата.

На самом крају дали смо закључак и предлоге за будуће анализе самог система. Где је објашњено шта је наш систем успео да покаже и које су потенцијалне даље смернице.

У шестом поглављу дат је преглед коришћене литературе. У додатку рада представљени су кодови скрипте која је коришћена за прикупљање података.

2. Теоријске основе

У овом поглављу ћемо представити кратак преглед теоријских основа које се тичу неуронских мрежа, сентимент анализе и криптовалуте Bitcoin.

2.1. Неуронске мреже

2.1.1. Увод у неуронске мреже

Када помињемо неуронске мреже, пре свега мислимо на вештачке неуронске мреже. Захваљујући неуронима, људски мозак је способан да изводи велики број сензорних, моторних, механичких и интелектуалних активност много брже него што то раде најбржи компјутери на свету. Рад на вештачким неуронским мрежама подстакнут је проучавањима способности биолошких неуронских мрежа у људском мозгу, са жељом да се те способности репродукују, односно да се направи машина која ће да моделује по принципима које примењује људски мозак приликом брзог увиђања и решавања проблема[8].

Неуронска мрежа је паралелно дистрибуирани систем направљен од једноставних процесирајућих јединица који имају могућност да сачувају обрађене податке и да их омогуће за даље коришћење[8].

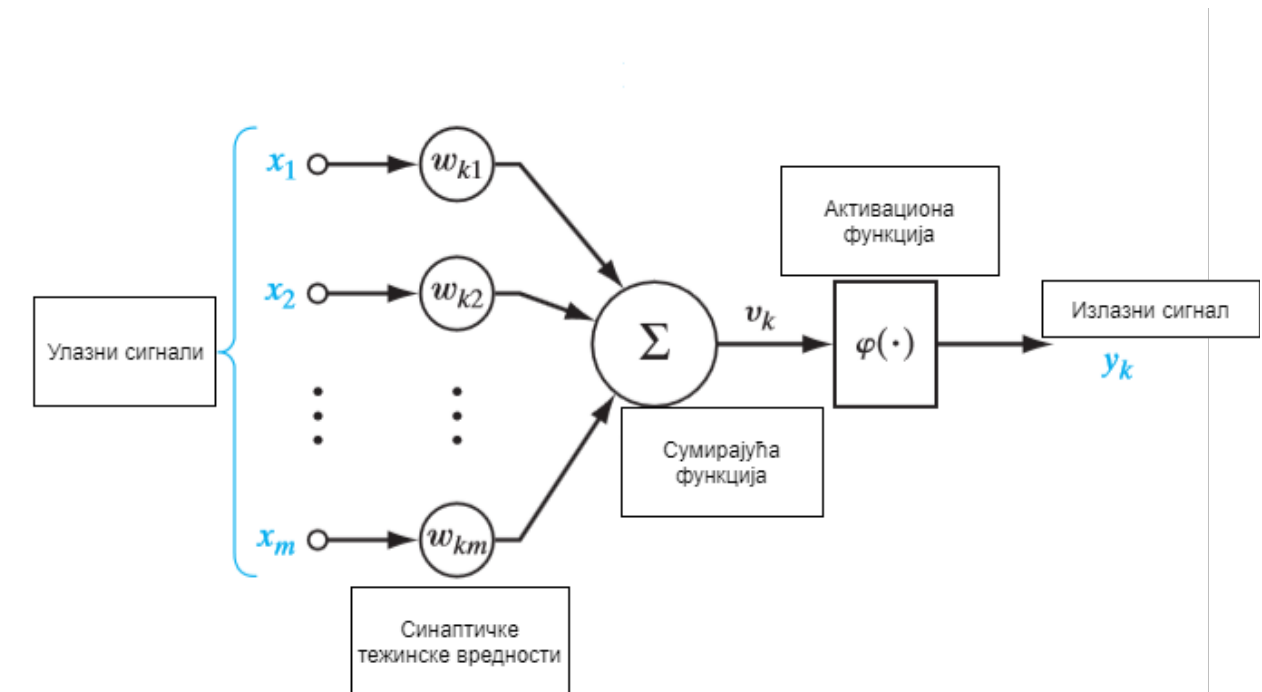
Вештачке неуронске мреже нису толико комплексне као наш мозак, али постоје две кључне сличности између биолошких и вештачких неуронских мрежа: а) Главни чворови обе мреже су једноставни рачунарски уређаји који су повезани на високом нивоу; б) Конекције међу неуронима су те које одређују шта ће бити главна функција саме неуронске мреже[9].

2.1.2. Структура неуронских мрежа

Фундаментални чинилац вештачке неуронске мреже је неурон. Неурон је јединица за процесирање информација која омогућава функционисање саме неуронске мреже. Постоји неколико главних чинилаца неуронског модела[8] који су приказани на Слици 1.:

- скуп синапси или повезујућих веза, које су описане тежинским факторима. Улазни параметар се множи са синаптичком тежином и предаје даље.

- сабирач - сабира улазне сигнале који су измерени на основу синпатичке тежинске вредности самог неурона.
- активациона функција - лимитира величину излаза самог неурона. Сврха активационе функције је да сам излазни сигнал стави у неки опсег који је предвиђен овом неуронском мрежом. У већини ситуација ова вредност варира или између 0 и 1 или између -1 и 1.



Слика 1: Структура неуронске мреже

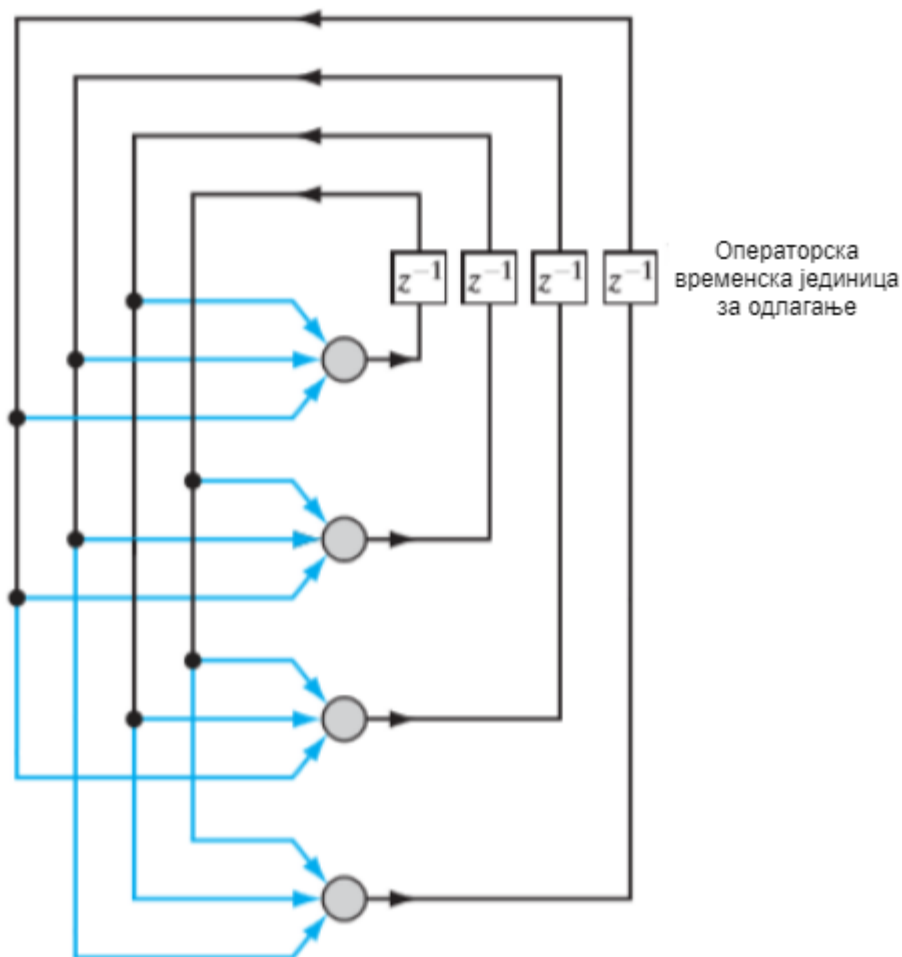
2.1.3. Типови архитектуре неуронских мрежа

Постоји велики број типова неуронских мрежа који се деле на основу алгоритама који користе у својој обради података. Генерална подела неуронских мрежа је на: а) нерекурентне мреже и б) рекурентне мреже.

- Нерекурентне мреже (енг. *Feed-forward*) – мрежа која садржи улазе, излазе и скривене слојеве. Сигнали у овој мрежи путују само у једном смеру. Улазни подаци су предати процесуирајућем елементу који изводи одређену рачуницу над њима. Сваки елемент ради своје калкулације на основу измерених сума својих улаза. Након што се заврши са овим улазом, нова израчуната вредност

постаје нови улаз за следећи слој и тако докле год није прошла кроз све слојеве и није одредила све излазе[10].

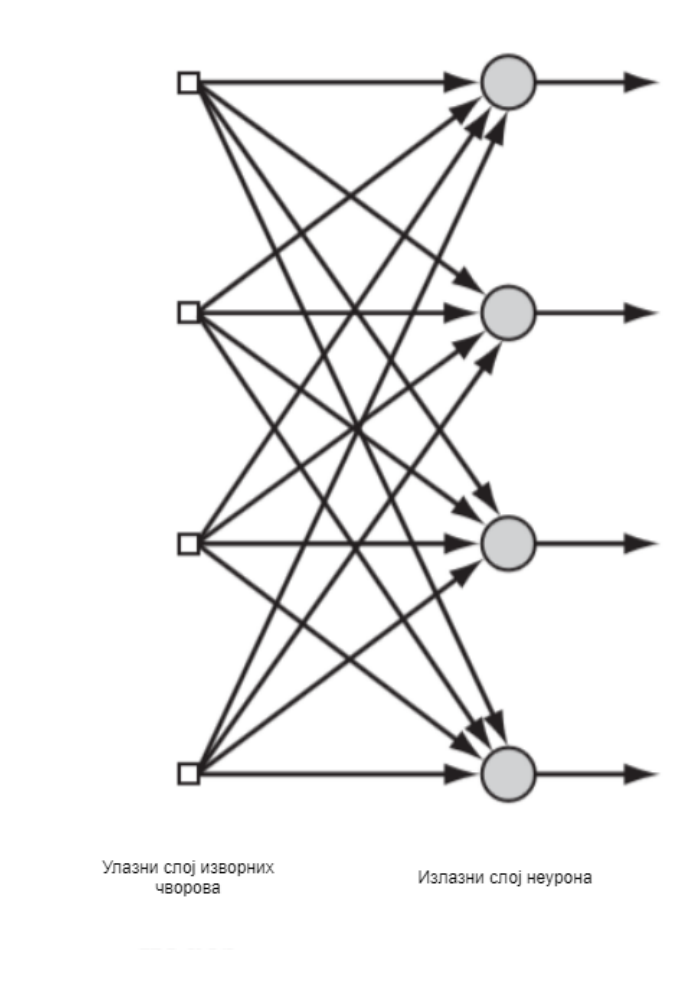
- Рекурентне мреже (енг. *Feed-back*) – мрежа која може да има сигнале који путују у оба смера. Све могуће конекције између неурона су могуће. Због приступачности петљама које ова мрежа има, овај систем се претвара у линеарни динамички систем који се мења константно, док се не достигне жељено стање[10]. На Слици 2 се тачно може видети итераторска могућност коју рекурентна мрежа има.



Слика 2: Рекурентна неуронска мрежа

Поред те поделе, рекурентне архитектуре можемо даље поделити на једнослојне или вишеслојне рекурентне мреже.

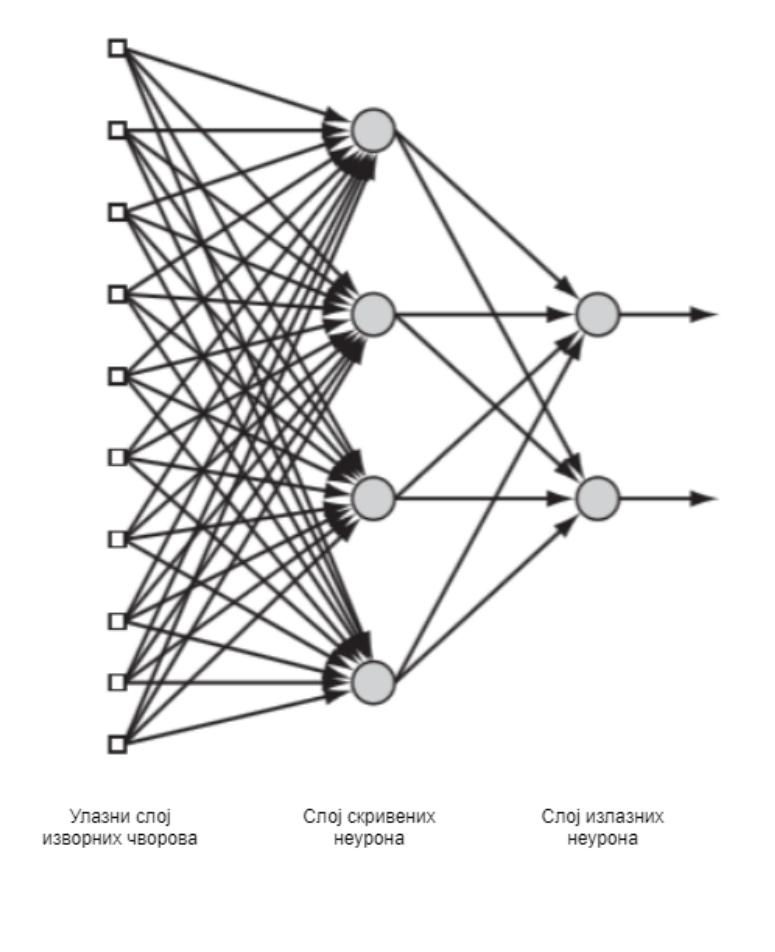
- Једнослојна рекурентна мрежа (енг. *Single-Layer Feed-forward Network*) – представља рекуренту мрежу која има само један слој односно нема скривене слојеве и то се најбоље може видети на Слици 3. Ово представља најједноставнију форму неуронске мреже где се улазни подаци директно пројектују на излазне чворове[8].



Слика 3: Једнослојна рекурентна неуронска мрежа

- Вишеслојна рекурентна мрежа (енг. *Multi-Layer Feed-forward Network*) – разликује се од једнослојне по присутности једног или више скривених

слојева. Назив скривена потиче од чињенице да одређен део неуронске мреже уопште не може директно да се види од стране улазних или излазних чворова мреже. Сврха скривених слојева је да се утиче на спољашње улазе и излазе мреже[8]. На Слици 3. приказана је вишеслојна нерекурентна неуронска мрежа са 4 скривена слоја.



Слика 4: Вишеслојна нерекурентна неуронска мрежа

2.1.4. Примена неуронских мрежа

Неуронска мрежа је први пут примењена шездесетих година и од тада је коришћена у различитим областима[11]: телекомуникацијама, контроли звука и вибрација, при контролеру акцелерометра за ласер и многим другим.

Због способности да прихвати и процесуира велику количину података у кратком периоду, у данашње време је добила много ширу примену у сфери пословања и академских дисциплина, као што су:

- Компресија слика [14]
- Предикција тржишних цена [15]
- Медицина [16]
- Препознавање лица [13]
- Архитектура [12]

Поред наведених области се не сме изоставити трговање као једна од изузетно битних области примене неуронских мрежа. Примену је нашла у предвиђању тренда или будуће цене.

2.2. Сентимент анализа

2.2.1. Основна идеја сентимент анализе

Сентимент анализа је аутоматски процес разумевања значења неког текста. На основу овог процеса писани и изговорени текст се интерпретира као позитиван, негативан или неутралан за тему коју истражујемо. Сентимент анализа нам помаже да квалитативне податке претворимо у квантитативне, тако што ћемо одређене речи, слоге, реченице пондерисати на основу наше теме и самим тим им доделити одређену вредност. У тренутку када речи постану бројеви, тада можемо радити анализу текста и анализирати га као подобног или неподобног мишљења[17][19][20].

Постоји неколико метода и алгоритама на основу којих може да се имплементира сентимент анализа. Сви постојећи начини могу да се групишу у три генерална типа[20]:

1. Алгоритми који функционише на основу јасно дефинисаних сетова правила. Ово представља класичан систем који на основу одређених улаза даје излазе на основу претходно постављених правила. Систем овакве природе уме да буде поприлично непрецизан јер не узима у обзир начин на који су речи постављене у реченици и самим тим може погрешно протумачити и дати супротан резултат од очекиваног. Поред тога, није скалабилан, што значи да додавање нових правила може представљати суков са пређашњим правилима и као резултат давати нежељене резултате.
2. Алгоритми који се ослања на технике машинског учења да би научио из података и самим тим постављао правила на основу тога. За сентимент анализу заснованих на оваквим алгоритмима морамо имати процес тренинга и процес предвиђања. У првом процесу се моделу дају подаци на основу којих ће научити наш систем да одређене улазе повезује са одговарајућим излазима. Онда у другом процесу, систем узима до сада не виђене улазе који су убачени у модел који на основу наученог генерише предвиђена решења[21].

3. Хибридни алгоритам - комбинује претходно наведене алгоритме. Употребом оваквог система узима се најбоље ставке из оба и на тај начин се унапређује тачност и прецизност резултата нашег система.

2.2.2. Фазе сентимент анализе

Сентимент анализу можемо поделити у четири фазе[20]:

1. Фаза прикупљања података – представља фазу у којој се прикупљају подаци са одређеног извора. Начин прикупљања може бити кроз скрипту као у овом раду или простим скидањем података са интернета;
2. Фаза тренирања података – представља фазу у којој модел покушава да учи на основу задатих улаза који је његов одговарајући излаз. У овој фази долази до упаривања одређених улаза са са излазима како би се лакше предвиђало касније.
3. Фаза тестирања података – представља фазу у којој се модел који је научио у претходној фази пушта кроз исти скуп података да би прорачунали колико је добро научио на задатом скупу. Уколико смо незадовољни са резултатима враћамо се на другу фазу.
4. Фаза предвиђања – представља фазу у којој се научени модел пушта да ради предвиђање на новом скупу података који има исте улазе и даје одређене резултате. На основу ових излаза се даље раде анализе тј доносе одлуке о будућим потезима.

2.2.3. Типови сентимент анализе

Постоји више типова сентимент анализе. Поделе се врше на основу тога шта желимо да докажемо тј. какав излаз желимо да добијемо на основу наших улаза. Неки од најважнијих су:

1. Анализа мишљења – представља анализу поларитета мишљења у односу на то да ли је оно позитивно, негативно или неутрално. Поред тога, уместо поделе на само ове три категорије, постоје могућности где мишљења можемо да категоришемо на[20]:

- веома позитивна
 - позитивна
 - неутрална
 - негативна
 - веома негативна
2. Детекција емоција – има за циљ да открије позитивну или негативну емотивну реакцију људи на неку тему. Различите теме изазивају код људи супротне реакције што отежава анализу[20].
 3. Аспектна анализа – покушава да из излаза који је добила на основу анализе закључи на који аспект истраживане теме се односи. Она се углавном користи за производе где покушава да се утврди који део производа се корисницима највише свиђа односно не свиђа[22].
 4. Анализа намере – представља најтежи задатак који сентимент анализа решава. Првествено треба схватити да ли је текст имао позитиван или негативан поларитет, а након тога закључити која је намера особе која је то изговорила односно написала. Ово изискује додатно знање које треба да се научи машина да би могла доћи близу онога што је је била намера самог аутора текста[23].

2.2.4. Данашњи допринос

У данашње време, када се пословања, информација, маркетинга и других битних ставки, пребацује на интернет, компанијама које нуде своје услуге није довољно само да знају на који начин се њихов производ перципира на тржишту, већ покушавају да схвате и осећања потенцијалних купаца[8]. Управо сентимент анализа омогућава овакву информацију у чему и лежи њен велики значај. Компанија захваљујући овој анализи може да делује превентивно или тренутно, јер има увид у промену односа купаца према производу. Да бисмо боље разумели околину морамо да схватимо њихова осећања. То нам сентимент анализа омогућава и зато је јако битна. Тренутак када ми знамо да је дошло до промене у јавном мњењу, можемо да будемо ти који ће одреаговати на те промене превентивно[19].

Међу главним користима сентимент анализе могу да се нађу:

1. Приоритизација акција од стране фирме када је у питању доношење одлука који се тичу окружења.
2. Праћење трендова током времена.
3. Укључење у конкретан период у времену. Конкретан догађај који је обележио тај дан и какво је мишљење креирао код људи.
4. Праћење конкуренције на друштвеним мрежама и поређење реакција у односу на одређене производе.
5. Увид у веру власника и будућих купаца криптовалута према њиховом власништву и каква ће бити њихова следећа акција.

2.3. Bitcoin

У овом поглављу ћемо објаснити главни део теме који се анализира. Овде ће бити објашњено шта је Bitcoin, његов кратак историјат, флукуација његове вредности и сумирани коментари стрчњака у вези његове будућности. Предвиђање цене акција, валута, берзанских роба итд увек је представљало посебан изазов за људе. Криптовалуте су посебно интересантне с обзиром да њихове цене имају велику флукуацију што отвара могућност за остваривање високих добитака на тржишту. Овај рад преставља покушај да се на систематски начин провери способност неуронских мрежа, да се уоче обрасци у кретању цена Bitcoin-а и евентуално искористе високе флукуације.

2.3.1. Историјат

Давне 2008. године, особа која се представила као Сатоши Накамото издао/ла је рад на тему „*A Peer-to-Peer Electronic Cash System*“. Овај рад је подразумевао методе коришћења *peer-to-peer* мреже за генерисање система, који је био описан као систем за електронске трансакције, који не мора да се заснива на поверењу учесника система[26].

Након тога, Сатоши Накамото креира први изворни блок и мрежа почиње да ради. Као и свака технологија или новитет, нарочито када је реч о финансијским системима, било је потребно време да ова мрежа буде прихваћен и почне да се користи.

До прве трансакције у систему дошло је годину дана након постављања и она се догодила између Сатоши Накамотоа и програмера Хала Финија. Фини је скинуо софтвер за Bitcoin оног дана када је био објављен и за то је био награђен са 10 јединица ове криптовалуте. Наравно, осим њега и други појединци су дали допринос развоју овог система[25].

Током наредних неколико година, почела је да расте вредност биткоина, али су истовремено настајале и друге криптовалуте. Bitcoin-у је нарочито ојачавало вредност то што су неке фирме почеле да га прихватају као платежно средство, чиме је расло и поверење, а самим тим и број корисника, односно купаца[27].

Bitcoin веома брзо достиже вредности од 40 до 60 долара, да би после краће корекције цена наставила вртоглави раст до нивоа 200 до 400 долара. Као последица овог вртоглавог успона долази до поделе јавности на оне који су сматрали да Bitcoin представља будућност финансијског платног система и добру прилику за улагање. И на оне који су то сматрали једном врстом преваре(због непостојања неке суштинске вредности иза Bitcoin-а). Неке интересантне чињенице везане за Bitcoin у периоду од 2013. до 2017. су[28]:

- број академских чланак, истраживања и радова о Bitcoin-у је растао експоненцијално са 83 у 2009. на 424 у 2012. до чак 3580 у 2016.;
- одређени број држава почиње да уврштава криптовалуте у категорију правог новца и даје му платежну функцију;
- број продаваца који је почео да прихвата Bitcoin је премашио цифру од 100.000.

Свака од ових чињеница које су горе набројане представља одређен подстицај већем поверењу у саму криптовалуту, а оно што ће овај рад покушати да покаже јесте

да је са овим догађајима и са оваквим начином ширења новости о биткоину формира и сама вредност криптовалуте.

2.3.2. Флуктуација вредности

Још од настанка Bitcoin-а волатилност ове криптовалуте је био опште познат. Разлика у односу на данашњу вредност промене вредности је то што не причамо више о скоковима са 10\$ на 20\$ који се сматрао великим коракном, већ о данашњим скоковима са 15 000\$ на 16 000\$. Када овако посматрамо скокове морамо да будемо пажљиви из ког угла сагледавамо промену вредности јер првенствен скок је значио промену вредности за 100%, док је други био једва 6%. Чак и када и овако гледамо морамо узети још један параметар у обзир, а то је колико јединица криптовалуте купац поседује.

Почетна призната вредност Bitcoin-а 2010. године је износила 0.003 долара. Пре тога, Bitcoin није имао вредност, јер нису постојале никакве трансакције на тржишту, те вредност није могла ни да се израчуна. Највиша вредност коју је Bitcoin икада достигао је износила 19 783,06 долара и то се догодило 17. децембра 2017. године.



Посматрајући Сliku 5, не можемо много тога да закључимо о његовој флуктуацији током времена, јер једина чинјеница коју можемо да запазимо јесте то

да су се од средине 2017. до почетка 2018. дешавале велике промене. Како бисмо боље анализирали његову волатилност, морамо да поделимо целокупан период на неколико под периода.

1. На Слици 6 приказано је кретање цене у периоду 28.4. до 31.12.2013.



Слика 6: Вредност Bitcoin-а у периоду: 1. 11.4.2013.-31.12.2013.

На овој слици већ можемо да видимо до каквих је промена дошло у том периоду. Када бисмо упоредили ову слику са сликом целог тока вредности криптовалуте, овај период нам не би много значео. Због тога, када гледамо из ове сужене перспективе, можемо да издвојимо неколико битних догађаја за које се сматра да су утицали на вредност биткоина:

- Дошло је до прве трансакције, односно првог коришћења ове валуте за плаћање неке услуге;
- Месец дана након ове трансакције вредност Bitcoin-а је скочила за 900%, у року од 5 дана;
- Вредност се у једном тренутку изједначава са вредношћу од 1 долара;
- Тренутак када је вредност досегла 31 долара, што је изазвало страх од изградње финансијског балона;

- Након 2 месеца вредност пада на 2 долара. Због овако великог пада се сматрало да је то крај Bitcoin-а;
- Након дугог и напорног периода, Bitcoin-а успева да постигне три пута своју највећу вредност и да поново падне испод исте;
- Највећа позната вредност у овом периоду је била 1 242 долара, она је постигнута 29.11.2013. године након што је пар дана пре тога била свега између 150/200\$.

2. На Слици 7 приказано је кретање цене у периоду од 31.12.2013. до 01.05.2014.



Слика 7: Вредност Bitcoin-а у периоду: 31.12.2013.-01.05.2014.

Разлог зашто је овај период изабран је тај што је након своје највеће вредности икада Bitcoin био у констатном паду, као што графикон показује:

- Директан пад на 600 долара, а затим поновни раст на 1000 долара, што је доста узбуркавало јавност;
- Пад на 550 долара и нагло подизање на 750 долара;
- Стабилизација на 400 долара;

- Крајем априла Bitcoin је достигао најнижу вредност још од прве крипто финансијске кризе која се догодила 2012-2013. године.

3. На Слици 8 приказано је кретање цене у периоду од 01.05.2014. до 01.02.2017.



Слика 8: Вредност Bitcoin-а у периоду: 01.05.2014.-01.02.2017.

Трећи период је познат по томе што је испуњен променама, али на начин на који Bitcoin то до сада није показивао, а то су постепен и сигуран раст или пад вредности. Да би на самом крају тог периода достигли нову вредност која је била приближна некадашњој највећој и износила 1150 долара. Наставак пада вредности до марта 2015. где се коначно стабилизовао на 200/300 доалра

- Следи константан и стабилан раст вредности;
- У новембру је порастао на 500\$, у јуну 2016. на 750 долара;
- На крају је у јануару 2016. године успео да стигне своју некадашњу вредност и приђе на 1150 долара.

4. На Слици 9 приказано је кретање цене у периоду од 01.02. до 01.09.2017.



Слика 9: Вредност Bitcoin-а у периоду: 01.02.2017.-01.09.2017.

Овај период власници биткоина памте као највеће скокове икада. Након што је изгубио 30% вредности, биткоин почиње да се враћа вредност:

- јануар – са 750 долара на нови врхунац од 1290 долара;
- мај – Bitcoin први пут прешлази вредност од 2000 долара;
- крајем јуна тај рекорд је опет оборен са вредношћу од преко 3000 долара;
- у августу и септембру вредност прелази 4 и 5 хиљада долара.

5. На Сlici 10 приказано је кретање цене у периоду од 01.09. до 17.12.2017.



Слика 10: Вредност Bitcoin-а у периоду: 01.09.2017.-17.12.2017.

Након драстичног пада који се догодио средином септембра, расправе међу власницима криптовалуте о дефинитивном крају Bitcoin-а. Ипак, испоставило да ће ово бити најуспешнији период за Bitcoin, где ће се достићи вредност од невероватних 17.900\$.

6. На Слици 10 приказано је кретање цене у периоду од 17.12.2017. до 14.09.2018.



Слика 11: Вредност Bitcoin-а у периоду: 17.12.2017.-14.09.2018.

Оно што овај график показује је последњи позитиван скок који је Bitcoin имао, а то је скок на 19 783.06\$. Од тог тренутка почиње константан пад вредности који већина тумачи као „пуцање балона“ односно „пропадање“ Bitcoin-а.

Период који је овај рад обухвата три фазе развоја вредности. То су IV, V и VI, а све их карактерише пад и раст вредности Bitcoin-а. Поред тога, још један од разлога зашто је баш овај период изабран за анализу је и начин на који су померања вредности коментарисана на друштвеним мрежама, првенствено на Reddit-у и Twitter-у. Из тог разлога сматрамо да овде постоји основа за истраживање где ћемо покушати да покажемо везу између реакција људи на интернету и вредности Bitcoin-а.

2.3.3. Експертска анализа

У овом поглављу ћемо изнети мишљења економиста о самој криптовалуту и како њена вредност настаје и флукутира. Та мишљења ћемо обрадити кроз неке од проблематичних страна Bitcoin-а и шта стручњаци сматрају да су његове мане:

- Чување вредности - када валута има функцију чувања вредности, власник те валуте је добија куповином, разменом добара или услуга. Када је валута

потрошена, власник очекује да ће добити исту економску вредност за коју је платио. Када је биткоин у питању то се ретко дешава, посебно у последњих годину дана када му се вредност мењала и по 20-40% у периоду од неколико до недељу дана[29];

- Упоређивање вредности производа на основу вредности криптовалуте – Да би валута имала ову функцију морали бисмо да имамо могућност да је третирамо као бројилац. На пример, када бисмо узели кафу у *Кафетерији* и платили је 240 динара, а кафа у локалу који је изван центра града или ван Београда кошта 120 динара, лако бисмо схватили да је кафа дупло скупља. Проблем који биткоин има када је ова функција у питању јесте тај што је некада једна пица могла да се купи за 10.000BTC, а сада се ретко ко може похвалити да има приближно исту уштеђевину ове валуте. Због екстремне волатилности, њему се ова функција не може доделити. Можда најозбиљнија препрека коју биткоин има, а која се често заобилази као тема разговора, јесте чињеница да цена једног биткоина је много већа у односу на већину производа[29];
- Средство размене- с обзиром да се Bitcoin-ова вредност не може израчунати на основу фундаменталне анализе саме фирме, њених производа и на основу тога израчунати колико вреди. Његова вредност се рачуна на основу коришћења у потрошачкој економији. Доказ за такво рачунање вредности Bitcoin-а можемо видети на основу одређеног броја људи који живе на основу трошења само криптовалуте или одређеног броја бизниса који прихвата Bitcoin. Баш због оваквих догађаја долазимо до ситуације да већина трансакција која се дешавају коришћењем биткоина не подразумевају куповину производа или коришћење неких услуга. Поред тога, још једна препрека на коју биткоин наилази да би постао средство размене је стварање нових токена ове криптовалуте, осим уколико је купац власник неколико ригова на којима може да рудари ову криптовалюту[29];

- Ефикасност самог Bitcoin-а – стручњаци сматрају да ће ефикасност Bitcoin-а расти са повећањем броја инвеститора. По њиховом мишљењу, тек када буде постао стабилан биће могућа анализа и лакше предвиђање[30];
- Недостаци Bitcoin-а – неки од недостатака који се набрајају су нецентрализација, односно недостатак неког централног тела које ће нагледати саму валуту, као и одсуство потврде идентитета приликом трансакција, што код потенцијалних корисника ствар сумњу. Као последњи недостатак, али не најмање битан, истиче се то да су трансакције ове криптовалуте неповратне, што значи да ако дође до неке грешке, због комплексности самог система, повратак новца није толико једноставан[31];

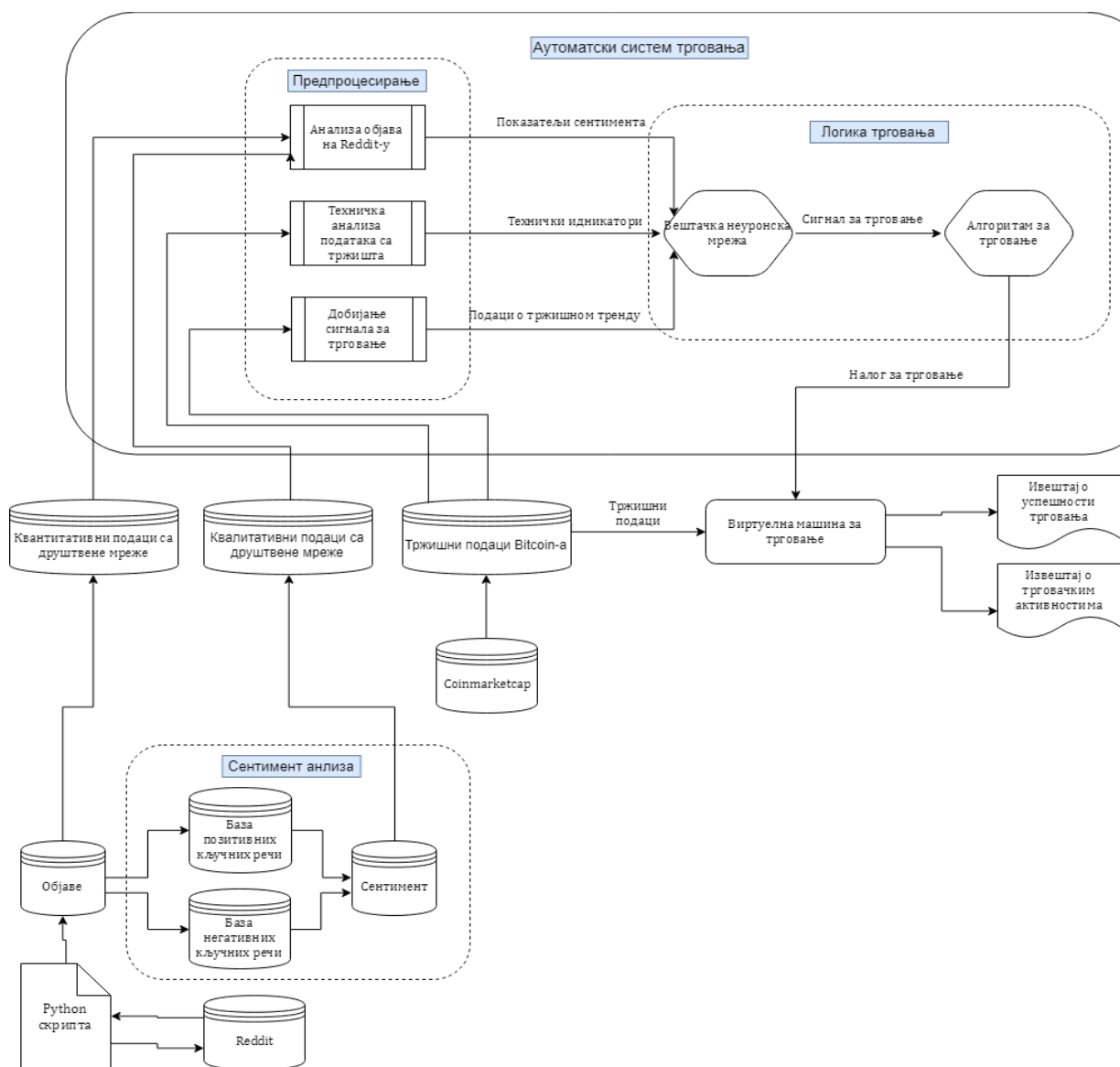
2.3.4. Генерално запажање

Постоји неколико могућих праваца у којима би криптовалута могла да се креће:

- грађење поверење инвеститора - истраживања су показала да су флукуације биткоина формиране на основу психолошких фактора који су утицали на инвеститоре, а не на основу фундаменталне вредности саме криптовалуте. Тиме је доказано да начин кретања вредности Bitcoin-а је у вези са перцепцијом инвеститора, односно колико они имају поверења у будућу вредност[32];
- привремено стабилизовање вредности - као што смо раније поменули, Bitcoin је у више ситуација привремено стабилизован, након чега је уследио поновни раст.
- званична валута - Bitcoin се у потпуности прихвата као званична валута и почиње да се користи у свакодневном животу.

3. Предложени систем

У овом раду смо направили и тестирали аутоматски систем за трговање. Овај систем је био намењен да тргује са Bitcoin криптовалutom користећи вештачку неуронску мрежу као компоненту за машинско учење. Предложени систем за трговање је наставак на претходне радове[1][33].

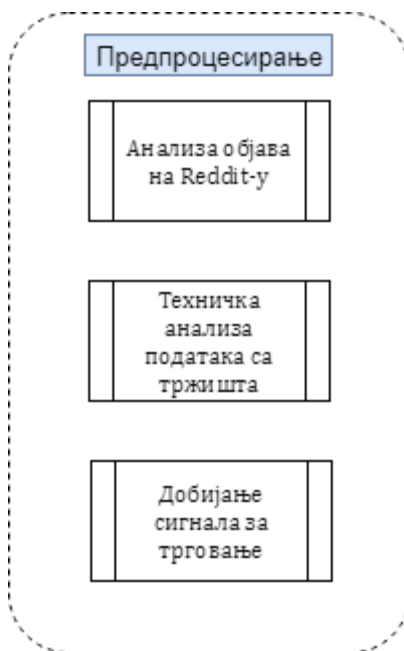


Слика 12: Предложени систем

Предложени систем који можемо видети на Слици 12 се састоји од три компоненте: а)предпроцесирајућа компонента, б)компонента логике трговања и в)компонента сентимент анализе

3.1. Предпроцесирајућа компонента

Предпроцесирајућа компонента је сачињена од три модула који функционишу паралелно. У процесу добијања сигнала за трговање се користе подаци везани за цене криптовалута како би се израчунали сигнали за трговање који ће се користити при тренирању вештачке неуронске мреже. У модулу за анализу објава са Reddit-а се узимају у обзир квалитативни и квантитативни подаци преузети са друштвене мреже и креирају сентимент индикатори. И у последњем модулу за техничку анализу података са тржишта се прикупљају подаци са тржишта криптовалута и рачунају различити технички индикатори.



Слика 13: Предпроцесирајућа компонента

На Слици 13 можемо видети сва три модула која се налазе у овој компоненти, а то су:

Модул за анализу објава на Reddit-y – који узима квалитативне и квантитативне податке са друштвене мреже и над њима ради анализу. Подаци који се овде обрађују су: Број коментара на 30 објава, Број реакција на 30 објава, Позитивни сентимент и Негативни сентимент. Прва два податка представљају квантитативне податке, док су друга два квалитативна. Од ових података даљом анализом добијамо

улазне параметре за неуронску мрежу, а то су: Просечан број коментара по објави, Просечан број реакција по објави, Сентимент и Количник сентимента.

- Просечан број коментара по објави (енг. *CommentsPerPost*) – добија се дељењем суме свих коментара тог дана са бројем постови које смо узели.
- Просечан број реакција по објави (енг. *ScorePerPost*) – добија се дељењем суме свих реакција тог дана са бројем објава које смо узели.
- Сентимент (енг. *SentimentScore*) – добија се одузимањем позитивног сентимента од негативног сентимента[20]:

$$\text{SentimentScore} = \text{PositiveSentiment} - \text{NegativeSentiment} \quad (1)$$

- Количник сентимента (енг. *SentimentRatio*) – добија се дељењем позитивног сентимента у односу на негативни[20]:

$$\text{SentimentRatio} = \frac{\text{PositiveSentiment}}{\text{NegativeSentiment}} \quad (2)$$

Модул за техничку анализу података са тржишта се служи подацима који су прикупљени са тржишта. Ти подаци подразумевају ценовне податке саме криптовалуте у том дану, обим трговања, укупну тржишну вредност и процентуалну заступљеност Bitcoin-а. Који се директно прослеђују неуронској мрежи као улазни подаци.

Модул за добијање сигнала за трговање се служи техничким индикаторима који анализирају цену криптовалуте. На основу тих индикатора се добијају сигнали за трговање који могу предвиђати да аутоматски систем за трговање треба да купи, задржи или прода Bitcoin.

- Однос вредности трговања и тржишне капитализације криптовалуте (енг. *Volume2Mcap*) – добија се простим дељењем ове две вредности[34]:

$$\text{Volume2Mcap} = \frac{\text{tradingVolume} \times \text{marketPrice}}{\text{MarketCap}} \quad (3)$$

- Релативни индекс снаге [RSI] – индикатор импулса који служи како би се измерила величина скорашњих цена на основу којих би се анализирале прекомерне куповне или продајне ситуације[34]:

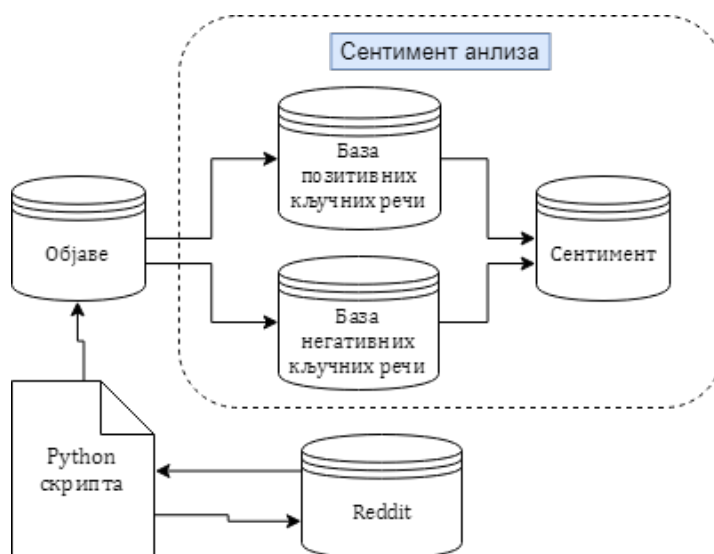
$$RSI = 100 - \frac{100}{1 + RS} \quad (4)$$

- Индикатор брзине промене [ROC] – индикатор који мери процентуалну промену цене између два задата периода[34]:

$$ROC = \left[\frac{Close_X - Close_{X-n}}{Close_{X-n}} \right] * 100 \quad (5)$$

3.2. Компонента сентимент анализе

Компонента сентимент анализе је најбоље анализирана у коду који је објашњен у додатку уз помоћ ког се добијају подаци са друштвене мреже Reddit. Квантитативни подаци се одмах чувају у бази за квантитативне податке са друштвене мреже. Квалитативни подаци пролазе кроз процес сентимент анализе тако што свака реч која је поменута у наслову или тексту пролази кроз базу кључних позитивних и негативних речи и на основу поклапања јој се додељује одређена вредност. Након пондерисања свих објава подаци се спуштају у базу квалитативних података са друштвене мреже.



Слика 14: Компонента сентимент анализе

На Слици 14 је описан цео процес који се дешава пре сентимент анализе и како се она одиграва када добије конкретне податке. Детаљан опис скрипте се налази у додатку, али оно што је кључно навести овде је да се уз помоћ ње прикупљају сви подаци који су нам потребни за предвиђање на основу сентимента тржишта. Извлачењем података са друштвене мреже Reddit се преузимају објаве у протеклих годину дана. Након преузимање објаве, анализира се састав сваког наслова и сваког текста који је објављен у току тог дана. Свака реч пролази кроз позитивну и негативну базу кључних речи. На основу поклапања са одређеним речима долази до додељивања тежишних вредности на одређене објаве. Након што је свака објава прошла кроз анализу садржаја долази до сумације свих тежишних вредности како би

се одредила тежишна вредност тог дана. Након рачунања сентимента за тај дан подаци се даље деле на квалитативне и квантитативне и такви се спуштају у одговарајуће базе података. Те базе података се даље користе као извори улазних података за претпроцесирајућу анализу и даље неуронску мрежу.

3.3. Компонента логики трговања

Компонента логики трговања је задужена за логику трговања кроз два модула. Први модул се односи на учење који је имплементиран у виду вештачке неуронске мреже, док се други модул односи на креирање алгоритма за трговање. Други модул користи предвиђања неуронске мреже како би одлучио да ли да купи, прода или да задржи Bitcoin.



Слика 15: Компонента логики трговања

На Слици 15 су приказани улазни подаци који улазе у модул логики трговања која на основу наученог даје као излаз налог за трговање на основу ког ће виртуелна машина трговати. Вештачка неуронска мрежа користи ове улазне податке како би на првих 100 дана учила како даље да тргује. Након тог периода, неуронска мрежа почиње да предвиђа односно да даје сигнале за трговање који су представљени као улазни подаци алгоритма за трговање. Алгоритам за трговање користи ове податке како би на основу задатих улаза коначно виртуелној машини дао налог за трговање који може обухватати куповину, продају или задржавање самог Bitcoin-а. Након завршетка трговања подаци о успешности трговања и извештај о трговачким активностима се чувају како би се анализирали резултати.

4. Резултати симулације

4.1. Подаци

Подаци који су коришћени у доле наведеним симулацијама су сви преузети са интернета и представљени су у изворном облику без додатних дотеривања како се не би утицало на истраживање.

Табела је сачињена од 365 редова и 14 колона. Сваки ред представља један дан и обухваћен је период од 01.08.2017. до 01.08.2018. године.

Табела 1: Улазни подаци

Датум	Рб	Отварајућа	Највиша	Најнижа	Затварајућа	Вредност трговања	Укупна тржишна вредност	Процентуална заступљеност Bitcoin-а	Број коментара на 30 објава	Број реакција на 30 објава	Позитивни сентимент	Негативни сентимент
8/1/2017	0	2,871.30	2,921.35	2,685.61	2,718.26	1,324,669,952	47,321,755,648	51.35	4,454	12,031	191	66
8/2/2017	1	2,727.13	2,762.53	2,668.59	2,710.67	1,094,950,016	44,950,835,200	45.01	3,283	10,361	37	31
8/3/2017	2	2,709.56	2,813.31	2,685.14	2,804.73	804,796,992	44,666,400,768	44.51	3,580	7,789	57	22
8/4/2017	3	2,806.93	2,899.33	2,743.72	2,895.89	1,002,120,000	46,276,210,688	45.41	2,688	11,201	44	66
8/5/2017	4	2,897.63	3,290.01	2,874.83	3,252.91	1,945,699,968	47,778,197,504	47.04	3,657	21,412	53	33
8/6/2017	5	3,257.61	3,293.29	3,155.60	3,213.94	1,105,030,016	53,720,891,392	47.71	2,365	6,794	65	32

Подаци који су приказани у Табели 1 се састоје од следећих:

- Датум (енг. *Date*) представља датум односно дан на који се тај ред у табели односи. Почетна вредност је 01.08.2017. године, а последња 01.08.2018. године.
- Рб (енг. *No*) представља редни број реда у табели. Ова вредност се касније у неуронској мрежи гледа као идентификациони број.
- Цена на отварању (енг. *Open*) представља отварајућу вредност Bitcoin-а у том дану. То је вредност са којом је започето трговање у том дану.
- Највиша цена (енг. *High*) представља највишу вредност Bitcoin-а у том дану. То је вредност преко које криптовалута није прелазила у том дану.
- Најнижа цена (енг. *Low*) представља најнижу вредност Bitcoin-а у том дану. То је вредност испод које криптовалута није прелазила у том дану.
- Цена на затварању (енг. *Close*) представља затварајућу вредност Bitcoin-а у том дану. То је вредност са којом је завршено трговање у том дану.

- Обим трговања (енг. *Volume*) представља вредност трговања тог дана, односно производ свих трансакција тог дана са њиховим ценама у тренутку трансакција.
- Укупна тржишна вредност (енг. *Market cap*) представља укупну тржишну вредност Bitcoin-а у доларима тог дана у односу на тренутну цену криптовалуте. Ова вредност се добија множењем тренутног број токена на тржишту са тренутном вредношћу.
- Процентуална заступљеност Bitcoin-а (енг. *Bitcoin dominance*) представља процентуалну заступљеност Bitcoin-а у односу на све остале криптовалуте које се тренутно налазе на тржишту.
- Број коментара на 30 објава представља број коментара које су имале 30 најбољих објава у том дану. То су биле објаве које су имале највећу посећеност и највећу интеракцију са корисницима и пратиоцима ове друштвене мреже.
- Број реакција на 30 објава представља суму свих позитивних и негативних реакција које су 30 најбољих објава имале у том дану. То су биле објаве које су имале највећу посећеност и највећу интеракцију са корисницима и пратиоцима ове друштвене мреже.
- Позитивни сентимент (енг. *Positive sentiment*) представља вредност која је израчуната на основу сентимент анализе најбољих 30 објава у том дану.
- Негативни сентимент (енг. *Negative sentiment*) представља вредност која је израчуната на основу сентимент анализе најбољих 30 објава у том дану.

Подаци су у већини случаја прикупљени из два извора која тренутно пружају највећу количину података:

- Цена на отварању, затварању, највиша, најнижа цена, обим трговања и процентуална заступљеност Bitcoin-а су преузети са сајта www.coinmarketcap.com уз помоћ њиховог јавног доступног сервиса за приступ овим подацима.

- Све објаве су преузете са друштвене мреже Reddit уз помоћ јавно доступног сервиса <https://pushshift.io>. Начин на који су преузете објаве са ове мреже су показани у додатку где је објашњен цео код и параметри који су коришћени да би се извукло 30 најбољих објава у том дану.
- Позитивни и негативни сентимент су израчунати на основу сентимент анализе свих објава које су претходно биле преузете са интернета. Првенствено је израчунат позитивни и негативни сентимент сваке објаве засебно, а на крају су сабрани сви бројачи који су дали позитиван и негативан сентимент за тај дан

4.2. Параметри система

У овом делу четвртог поглаља, представимо параметре система.

4.2.1. Параметри стратегије

Параметри за подешавање стратегије су:

- Почетан баланс на рачуну (на енг. *Initial account balance*) – представља стање на рачуну на почетку симулације, изражено у доларима.
- Процентуална величина лота (на енг. *Lot Percentage*) – представља проценат баланса са рачуна који ће бити уложен при новој куповној трансакцији.

4.2.2. Параметри система

Параметре за оптимизацију стратегије чине:

- Почетна величина узорка за тренирање (на енг. *Initial Training Sample Size*) – представља колико ће дуго неуронска мрежа тренирати пре него што крене да тргује;
- Узастопни губитнички праг (на енг. *Consecutive Losses Threshold*) – представља број узастопних трансакција након којих, уколико стратегија губи паре, треба да обустави трговање и поново почне обучавање неуронске мреже за трговање;

- Зауоставни губитнички проценат (на енг. *Stop Loss Percentage*) – представља процентуални ниво губитка који је стратегија спремна да прихвати у односу на отворену позицију. Када ниво пређе дефинисани зауоставни губитнички проценат тада стратегија аутоматски затвара отворену позицију;
- МаПериод (на енг. *MaPeriod*) – представља временски прозор на основу ког се рачунају вредности техничких индикатора;
- Праг сигнала за трговање (на енг. *Trading Signal Threshold*) – уколико сигнал за трговање пређе дефинисану вредност прага, систем ће покренути процес трговања.

4.3. Мере перформанси

За евалуацију успешности система за трговање одабране су следеће мере:

- Укупан број трговања – представља колико је пута неуронска мрежа куповала односно продавала Bitcoin када је добијала сигнал за неку од ове две акције.
- Проценат позитивних трансакција – представља проценат позитивних трговања у односу на укупан број трговања за задати период[34]:

$$\text{PositiveTradeRatio} = \frac{\text{PositiveTrades}}{\text{PositiveTrades} + \text{NegativeTrades}} \quad (6)$$

- Принос на инвестицију [ROI] (енг. *Return on investment*) представља процентуални запис који пореди ефикасност свих инвестиција у том периоду у односу на релативну цену те инвестиције[34]:

$$\text{ROI} = \frac{\text{GainFromInvestment} - \text{CostOfInvestment}}{\text{CostOfInvestment}} \quad (7)$$

- Профитни фактор [PF] (енг. *Profit factor*) представља процентуални запис који дели суму свих профитних трансакција са сумом свих губитних трансакција тако да уколико смо изгубили новац вредност ће бити између 0 и 1. Свака вредност која се налази преко 1, представља позитивну симулацију односно стратегију[34]:

$$\text{ProfitFactor} = \frac{\text{GrossProfit}}{\text{GrossLoss}} \quad (8)$$

- Број добитних односно негативних симулација по стратегији.

4.3.1. Резултати добијени коришћењем улазног скупа 1: сентимент анализа

У овом делу ћемо представити резултате трговања добијене коришћењем првог улазног скупа података који обухвата:

1. *commentsPerPost*,
2. *scorePerPost*,
3. *sentimentScore*,
4. *sentimentRatio*,
5. *volume2mcap*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата

Табела 2: Резултати добијени коришћењем улазног скупа 1

	No	Улазни подаци 1			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	30	46.6667	-0.0355	0.9642
	2	27	37.0370	-0.1243	0.8766
	3	31	45.1613	0.0897	1.1042
	4	17	29.4118	-0.3514	0.4941
	5	29	37.9310	-0.5322	0.4708
	6	22	59.0909	-0.2996	0.5989
	7	27	33.3333	-0.5725	0.4542
	8	17	35.2941	-0.2668	0.6670
	9	22	31.8182	0.0183	1.0230
	10	11	54.5455	0.1743	1.5467
	AVG:	23.3	41.03	-0.19	0.82
	MIN:	11	29.41	-0.57	0.45
	MAX:	31	59.09	0.17	1.55
				P:	3.00
				N:	7.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи први улазни скуп, износи 23,3 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан.

У просеку систем остварује 41% позитивних трансакција, губи 19% инвестиције, док је просечна вредност профитног фактора 0,82. Што нам говори да се на сваки изгубљени 1 долар заради 0,82 долара, што је лош показатељ.

Поред тога можемо увидети да је систем имао три добитне симулације и седам губитних. Када боље погледамо резултате можемо видети да су добитне симулације доносиле мали профит, док су губитне имале велике губитке. То нам говори да овај улазни скуп не може да реши посао.

Симулације које морамо издвојити и упоредити их су шеста и десета. Обе симулације су се показале са високим процентом позитивних трансакција. Оно што можемо извући из тога је да је у шестој систем лоше научио и да када је добијао прерано је излазио из позиције, а када је губио предуго остајао у позицији.

4.3.2. Резултати добијени коришћењем улазног скупа 2: сентимент анализа и Bitcoin доминација

У овом делу ћемо представити резултате трговања добијене коришћењем другог улазног скупа података који обухвата:

1. *commentsPerPost*,
2. *scorePerPost*,
3. *sentimentScore*,
4. *sentimentRatio*,
5. *volume2mcap*
6. *dominance*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата.

Табела 3: Резултати добијени коришћењем улазног скупа 2

	No	Улазни подаци 2			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	22	31.8182	0.0352	1.0458
	2	19	36.8421	-0.0494	0.9315
	3	32	56.2500	1.0597	1.7962
	4	36	50.0000	-0.4913	0.5330
	5	21	42.8571	0.8952	2.0157
	6	20	30.0000	-0.7022	0.1385
	7	33	36.3636	-0.5100	0.5001
	8	28	32.1429	-0.5602	0.5177
	9	29	34.4828	-0.1366	0.8704
	10	17	29.4118	0.2295	1.1572
	AVG:	25.7	38.02	-0.02	0.95
	MIN:	17	29.41	-0.7	0.14
	MAX:	36	56.25	1.06	2.02
				P:	4.00
				N:	6.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи други улазни скуп, износи 25,7 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан. Да се додавањем новог улазног параметра повећао проценат у односу на први улазни скуп

У просеку систем остварује 38.02% позитивних трансакција што је мање у односу на први улазни скуп, губи 2% инвестиције што је боље од првог скупа, док је просечна вредност профитног фактора 0,95. Што нам говори да се на сваки изгубљени 1 долар заради 0,95 долара, што је лош показатељ, али ипак бољи у односу на први улазни скуп.

Поред тога можемо увидети да је систем имао четири добитне симулације и шест губитних. Када боље погледамо резултате можемо видети да су добитне

симулације доносиле мали профит или имали један екстремни случај, док су губитне имале велике губитке. То нам говори да овај улазни скуп не може да реши посао.

Симулације које морамо издвојити су трећа и десета. Трећа симулације се издвојила као најдинамичнија симулација од свих и поред тога показала са високим процентом позитивних трансакција. Док код десете морамо нагласити да без обзира што је била мање динамична показала као добитна.

4.3.3. Резултати добијени коришћењем улазног скупа 3: сентимент анализа, Bitcoin доминација и техничка анализа

У овом делу ћемо представити резултате трговања добијене коришћењем трећег улазног скупа података који обухвата:

1. *commentsPerPost*,
2. *scorePerPost*,
3. *sentimentScore*,
4. *sentimentRatio*,
5. *volume2mcap*
6. *dominance*
7. *SSTOK*
8. *ROC*
9. *RSI*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата.

Табела 4: Резултати добијени коришћењем улазног скупа 3

	No	Улазни подаци 3			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	33	42.4242	0.3875	1.1858
	2	33	42.4242	-0.3588	0.7166
	3	40	42.5000	-0.5727	0.4406
	4	38	50.0000	-0.2795	0.7497
	5	33	42.4242	-0.3588	0.7166
	6	31	64.5161	1.3542	2.6325
	7	30	53.3333	0.5126	1.2460
	8	28	42.8571	0.1659	1.0834
	9	41	53.6585	-0.2987	0.8165
	10	26	34.6154	-0.5747	0.4585
	AVG:	33.3	46.88	0	1
	MIN:	26	34.62	-0.57	0.44
	MAX:	41	64.52	1.35	2.63
				P:	4.00
				N:	6.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи трећи улазни скуп, износи 33,3 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан. Да се додавањем новог улазног параметра повећао проценат у односу на други улазни скуп.

У просеку систем остварује 46,88% позитивних трансакција што је више у односу на први и други улазни скуп, губи 0% инвестиције што је боље у односу на оба, док је просечна вредност профитног фактора 1,00. Што нам говори да се на сваки изгубљени 1 долар заради 1 долар, што није лош показатељ, али ипак бољи у односу на прва два улазна скупа.

Поред тога можемо увидети да је систем имао четири добитне симулације и шест губитних. Када боље погледамо резултате можемо видети да су добитне симулације доносиле мали профит или имали један екстремни случај, док су губитне имале велике губитке. То нам говори да овај улазни скуп не може да реши посао.

Симулације које морамо издвојити су шеста и девета. Шеста симулација се издвојила као најпрофитабилнија симулација од свих и поред тога показала са

највишим процентом позитивних трансакција. Док код девете морамо нагласити да без обзира што је имала више од 50% позитивних трансакција се идаље показала као губитна.

4.3.4. Резултати добијени коришћењем улазног скупа 4: упросечена сентимент анализа

У овом делу ћемо представити резултате трговања добијене коришћењем четвртог улазног скупа података који обухвата први скуп података који је упросечен:

1. *commentsPerPost_ema*,
2. *scorePerPost_ema*,
3. *sentimentScore_ema*,
4. *sentimentRatio_ema*,
5. *volume2mcap_ema*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата.

Табела 5: Резултати добијени коришћењем улазног скупа 4

	No	Улазни подаци 4			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	25	48.0000	0.0901	1.0705
	2	25	36.0000	-0.5526	0.5356
	3	30	46.6667	-0.1886	0.7749
	4	32	43.7500	0.1671	1.1072
	5	28	42.8571	-0.0316	0.9791
	6	28	39.2857	0.8119	1.3638
	7	40	50.0000	0.2788	1.3035
	8	29	37.9310	0.6198	1.2909
	9	21	52.3810	0.0327	1.0396
	10	32	40.6250	-0.4865	0.5831
	AVG:	29	43.75	0.07	1
	MIN:	21	36	-0.55	0.54
	MAX:	40	52.38	0.81	1.36
				P:	6.00
				N:	4.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи четврти улазни скуп, износи 29 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан. Да се упросечењем првог улазног скупа повећала динамика трговања.

У просеку систем остварује 43,75% позитивних трансакција што је више у односу на први и други, али мањи од трећег улазног скупа, добија 0.07% инвестиције што је боље у односу на све до сада улазне скупове, док је просечна вредност профитног фактора 1,00. Што нам говори да се на сваки изгубљени 1 долар заради 1 долар, што није лош показатељ, али ипак бољи у односу на све остале улазне скупове због већег повратка на инвестиције.

Поред тога можемо увидети да је систем имао шест добитних симулација и четири губитне. Када боље погледамо резултате можемо видети да су и добитне и губитне симулације приближне профитном фактору 1. То нам говори да овај улазни скуп не може да донесе велики профит, а ни да изгуби велику количину капитала.

Не постоји ниједна конкретна симулација која се може издвојити за дубљу анализу због релативно подједнаких резултата.

4.3.5. Резултати добијени коришћењем улазног скупа 5: упросечена сентимент анализа и просечна Bitcoin доминација

У овом делу ћемо представити резултате трговања добијене коришћењем петог улазног скупа података који обухвата други скуп података који је упросечен:

1. *commentsPerPost_ema*,
2. *scorePerPost_ema*,
3. *sentimentScore_ema*,
4. *sentimentRatio_ema*,
5. *volume2mcap_ema*
6. *dominance_ema*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата.

Табела 6: Резултати добијени коришћењем улазног скупа 5

	No	Улазни подаци 5			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	32	43.7500	0.2166	1.2234
	2	22	31.8182	-0.5881	0.4560
	3	30	43.3333	1.2378	1.7533
	4	27	44.4444	-0.1437	0.8480
	5	26	38.4615	-0.3572	0.5980
	6	22	36.3636	-0.1613	0.8333
	7	18	44.4444	0.6508	2.0851
	8	40	52.5000	0.1791	1.1287
	9	27	44.4444	0.0320	1.0242
	10	24	37.5000	-0.4933	0.5635
	AVG:	26.8	41.71	0.06	1.05
	MIN:	18	31.82	-0.59	0.46
	MAX:	40	52.5	1.24	2.09
				P:	5.00
				N:	5.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи пети улазни скуп, износи 26,8 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан. Да се упросечењем првог улазног скупа повећала динамика трговања, али се додавањем новог улазног параметра смањила у односу на претходни скуп података.

У просеку систем остварује 41,71% позитивних трансакција што је мање у односу на претходни скуп података, али идаље веће у односу на проценат позитивних трансакција код друго улазног скупа. Добија 0.06% инвестиције што је приближно у односу на претходни улазни скуп, док је просечна вредност профитног фактора 1,05. Што нам говори да се на сваки изгубљени 1 долар заради 1,05 долара, што је добар показатељ који нам говори да је ово профитабилан улазни скуп.

Поред тога можемо увидети да је систем имао пет добитних симулација и пет губитних. Када боље погледамо резултате можемо видети да су и добитне и губитне симулације имале своје екстремне случајеве. Што нам не олакшава упоредну анализу засебних симулација, али свакако показује као профитабилан систем.

Симулацију коју треба издвојити је седма, јер се показала као симулација која је имала проценат профитних трансакција испод 50%, а поред тога показала као најпрофитабилнија симулација.

4.3.6. Резултати добијени коришћењем улазног скупа 6: упросечена сентимент анализа , просечна Bitcoin доминација и упросечена техничка анализа

У овом делу ћемо представити резултате трговања добијене коришћењем шестог улазног скупа података који обухвата трећи скуп података који је упросечен:

1. *commentsPerPost_ema*,
2. *scorePerPost_ema*,
3. *sentimentScore_ema*,
4. *sentimentRatio_ema*,
5. *volume2mcap_ema*
6. *dominance_ema*
7. *SSTOK_ema*
8. *ROC_ema*
9. *RSI_ema*

Да би унапредили валидност добијених резултата, за сваки улазни скуп података урађено је 10 симулација. Потом су израчунате средња, максимална и минимална вредност добијених мера перформанси, на основу којих је извршена анализа и дискусија резултата.

Табела 7: Резултати добијени коришћењем улазног скупа 6

	No	Улазни подаци 6			
		Укупно трговања	Проценат позитивних трговања	Повратак на инвестицију	Профитни фактор
Стратегије	1	31	54.8387	1.1333	1.3099
	2	28	50.0000	0.8169	1.4770
	3	28	53.5714	0.7110	1.3026
	4	27	48.1481	0.8487	1.3858
	5	29	41.3793	-0.0413	0.9739
	6	28	35.7143	-0.0399	0.9773
	7	39	38.4615	-0.0015	0.9992
	8	37	59.4595	0.1768	1.1451
	9	28	39.2857	-0.3120	0.7901
	10	36	58.3333	0.0604	1.0490
	AVG:	31.1	47.92	0.34	1.14
	MIN:	27	35.71	-0.31	0.79
	MAX:	39	59.46	1.13	1.48
				P:	6.00
				N:	4.00
				T:	10.00

Просечан број трансакција које је систем извршио у 10 симулација, користећи шести улазни скуп, износи 31,1 у посматраном периоду од 12 месеци. Можемо закључити да се систем показао као динамичан. Да се упросечењем трећег улазног скупа није драстично променила динамика трговања.

У просеку систем остварује 47,92% позитивних трансакција што је највише у односу на све остале улазне скупове. Добија у просеку 0.34% инвестиције што је веће у односу на претходни улазни скуп, док је просечна вредност профитног фактора 1,14. Што нам говори да се на сваки изгубљени 1 долар заради 1,14 долара, што је до сада најбољи показатељ који нам говори да је ово најпрофитабилнији улазни скуп.

Поред тога можемо увидети да је систем имао шест добитних симулација и четири губитне. Када боље погледамо резултате можемо видети да су и добитне и губитне симулације осцилирале између 0,79 до 1,48. Што нам говори да би нам овај улазни скуп или осигурао профит или изгубио мали проценат капитала.

Не постоји ниједна конкретна симулација која се може издвојити за дубљу анализу због релативно подједнаких резултата.

4.4. Упоредна анализа резултата

Табела 8: Сумирани резултати

	No	Укупно трговања			Проценат позитивних трговања			Повратак на инвестиције			Профитни фактор		
		AVG	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX
Улазни подаци	1	23.3	11	31	41.03	29.41	59.09	-0.19	-0.57	0.17	0.82	0.45	1.55
	2	25.7	17	36	38.02	29.41	56.25	-0.02	-0.7	1.06	0.95	0.14	2.02
	3	33.3	26	41	46.88	34.62	64.52	0	-0.57	1.35	1	0.44	2.63
	4	29	21	40	43.75	36	52.38	0.007	-0.55	0.81	1	0.54	1.36
	5	26.8	18	40	41.71	31.82	52.5	0.06	-0.59	1.24	1.05	0.46	2.09
	6	31.1	27	39	47.92	35.71	59.46	0.34	-0.31	1.13	1.14	0.79	1.48

У Табели 7 можемо јасно да видимо да су се најбоље показале трећа и шеста стратегија. То нам говори да су стратегије које су користиле све податке и све техничке индикаторе најбоље предвиђале и да је систем давао најбоље сигнале када треба трговати.

Поред тога треба нагласити да без обзира на то што је трећа била динамичнија у односу на шесту и имала веће максимуме у неким колонама, шеста ипак боља стратегија. То нам говори да уколико нормализујемо податке на периоду од три дана, као што је у нашој ситуацији, добијамо најбољи увид у тренутне податке.

Треба издвојити и упоредити први и други експеримент. Други експеримент је имао мањи проценат позитивних трансакција од првог, а опет је имао виши принос на инвестиције и виши профитни фактор. На истом нивоу можемо упоредити и четврти и пети експеримент који су се показали подједнако као прва два.

Оно што је најбитније да се издвоји је упоредна анализа прве три стратегије са друге три стратегије. Главни закључак који се може извести на основу Табеле 7 је чињеница да када су коришћене упросеचेњени показатељи систем се показивао боље. Имао је већи просечан број укупног броја трговања, већи проценат позитивних трговања, виши проценат принос на инвестиције и виши профитни фактор.

Овакви резултази отварају врата за даљој дискусији. Питања која се могу поставити су: Да ли даљим додавањем техничких индикатора можемо добити боље резултате? Да ли анализом садржаја свих коментара те објаве можемо добити бољи увид у сентимент анализу? Да ли ширењем на више друштвених мрежа можемо боље предвидети будућу вредност?

5. Закључак

Циљ овог рада је покушај да се увиди веза између сентимента на друштвеним мрежама и вредности Bitcoin-а. При овом покушају, направљен је систем за аутоматизовано трговање уз помоћ вештачких неуронских мрежа. У предложеном систему, вештачка неуронска мрежа има задатак да обучи стратегију за трговање користећи показатеље који су добијени на основу података са Reddit друштвене мреже, али и додатно, показатеље техничке анализе.

Скуп коришћених података је подељен на два дела: податке за обучавање и податке за тестирање система. У оквиру рада, приказани су резултати шест симулационих експеримената са предложеним системом у којима су коришћени различити улазни показатељи за обучавање стратегије. У три експеримента коришћене су сирове вредности показатеља, док су у преостала три коришћене њихове упросечене вредности. Излаз из неуронске мреже коришћен је као управљачка променљива система, на основу које доноси одлуку о томе да ли ће трговати и у ком смеру. Евалуација система је извршена коришћењем различитих мера перформанси: проценат профитабилних трансакција, приноса на инвестиције и профитног фактора.

На основу шест симулационих експеримената можемо донети одређене закључке везане за наш систем. Прва ствар која се показала је то да се увећавањем броја улазних променљивих и њихова разноврсност повећавала и успешност самог система. Упросечењем вредности улазних показатеља доводи до још бољих односно успешнијих резултата неуронске мреже. Увођењем свих изабраних улазних показатеља у систем и њиховим упросечивањем, приказани систем је успео да достигне ниво профитабилности гледано са аспекта просечних вредности посматраних мера успешности на 10 симулација система.

Поређењем резултата овог завршног рада са резултатима рада који је претходио овој студији „*Neural networks in market sentiment analysis for automated trading: The case of Bitcoin*“ можемо закључити да се повећањем разноврсности улазних променљивих, како са аспекта анализе тржишног сентимента тако и са

аспекта техничке анализе, може повећати успешност система и постићи његова профитабилност.

Највеће ограничење предложеног система уграђено је кроз једну битну карактеристику вештачких неуронских мрежа – њихова особина да се приликом понављања симулације са истим подацима оне обуче на потпуно другачији начин и да дају другачије излазне вредности, које се потом пресликавају у различите нивое успешности посматраног система. Другим речим, неуронска мрежа ће се приликом сваке нове симулације понашати другачије и произвести другачије резултате.

Иако су приказани резултати добијени као просечни на основу 10 изведених симулација, они нам не могу бити гаранција за будуће успешно коришћење овог система.

Овај завршни рад представља наставак истраживања покренутог током основних студија. Ауторова жеља је да овај рад постане база за будуће истраживање, чији ће фоку бити на детаљнијој анализи података доступних на друштвеној мрежи Reddit (нпр. детаљна анализа текста свих коментара на одређеној објави), као и ширење на више друштвених мрежа или проширење базе техничких показатеља који се могу користити као улазне променљиве система.

6. Литература

- [1] Matija Milekić, Aleksandar Rakićević, Pavle Milošević (2018) Neural networks in market sentiment analysis for automated trading: The case of Bitcoin. In Symposium proceedings-XVI International symposium Symorg 2018: Doing Business in the Digital Age: Challenges, Approaches and Solutions.
- [2] Gandal, N., Hamrick, J. T., Moore, T., & Oberman, T. (2018). Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics*, in press.
doi:10.1016/j.jmoneco.2017.12.004
- [3] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1), 1-8.
- [4] Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface*, 11(99), 20140623. doi:10.1098/rsif.2014.0623
- [5] Garcia, D., & Schweitzer, F. (2015). Social signals and algorithmic trading of Bitcoin. *Royal Society open science*, 2(9), 150288. doi:10.1098/rsos.150288
- [6] Kim, Y. B., Kim, J. G., Kim, W., Im, J. H., Kim, T. H., Kang, S. J., & Kim, C. H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PloS one*, 11(8), e0161197. doi:10.1371/journal.pone.0161197
- [7] Phillips, R. C., & Gorse, D. (2017, November). Predicting cryptocurrency price bubbles using social media data and epidemic modelling. In *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on* (pp. 1-7). IEEE.
- [8] Haykin, S. S., Haykin, S. S., Haykin, S. S., & Haykin, S. S. (2009). *Neural networks and learning machines* (Vol. 3). Upper Saddle River, NJ, USA: Pearson.
- [9] Demuth, H. B., Beale, M. H., De Jess, O., & Hagan, M. T. (2014). *Neural network design*. Martin Hagan.
- [10] Artificial Neural Network, An Introduction to Data Mining (2016). University of Toronto, Faculty of Applied Science & Engineering.

http://chemeng.utoronto.ca/~datamining/dmc/artificial_neural_network.htm Приступ 15.09.2018.

[11] Widrow, B., Rumelhart, D. E., & Lehr, M. A. (1994). Neural networks: applications in industry, business and science. *Communications of the ACM*, 37(3), 93-106.

[12] Aydinalp, M., Ugursal, V. I., & Fung, A. S. (2002). Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks. *Applied energy*, 71(2), 87-110.

[13] Steffens, J. B., Elagin, E. V., Nocera, L. P. A., Maurer, T., & Neven, H. (2001). *U.S. Patent No. 6,301,370*. Washington, DC: U.S. Patent and Trademark Office.

[14] Mohan, R., Smith, J. R., & Li, C. S. (1999). Adapting multimedia internet content for universal access. *IEEE Transactions on Multimedia*, 1(1), 104-114.

[15] Yoo, P. D., Kim, M. H., & Jan, T. (2005, November). Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. In *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on* (Vol. 2, pp. 835-841). IEEE.

[16] Lisboa, P. J., & Taktak, A. F. (2006). The use of artificial neural networks in decision support in cancer: a systematic review. *Neural networks*, 19(4), 408-415.

[17] https://en.wikipedia.org/wiki/Sentiment_analysis Приступ 15.09.2018.

[18] <https://stackoverflow.com/questions/24605702/good-dataset-for-sentiment-analysis> Приступ 15.09.2018.

[19] <https://www.clarabridge.com/sentiment-analysis/> Приступ 15.09.2018.

[20] <https://monkeylearn.com/sentiment-analysis/> Приступ 15.09.2018.

[21] Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*(pp. 79-86). Association for Computational Linguistics.

- [22] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
- [23] Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- [24] Liu, B., Hu, M., & Cheng, J. (2005, May). Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web* (pp. 342-351). ACM.
- [25] <https://www.investopedia.com/terms/b/bitcoin.asp> Приступ 15.09.2018.
- [26] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.
- [27] https://en.wikipedia.org/wiki/History_of_bitcoin Приступ 15.09.2018.
- [28] Chohan, U. W. (2017). A History of Bitcoin.
- [29] Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In *Handbook of digital currency* (pp. 31-43).
- [30] Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, 148, 80-82.
- [31] Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2), 213-38.
- [32] Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130, 32-36.
- [33] Rakićević, A., Milovanović, A., & Aničić, R. (2016, June). An application of neural networks and fundamental analysis for automated trading: Belgrade stock exchange case. In *Symposium proceedings-XV International symposium Symorg 2016: Reshaping the Future Through Sustainable Business Development and Entrepreneurship* (p. 626-634). University of Belgrade, Faculty of Organizational Sciences.
- [34] Achelis, S. B. (2001). *Technical Analysis from A to Z*. New York: McGraw Hill.

Додатак

Увоз библиотека које су потребне да би ова скрипта могла да ради

#Import of required libraries for this project

```
import requests
import json
import time
import re
import pandas
import openpyxl
```

Дефинисање почетка и краја првог дана од ког се креће анализа

#Defining the start and the end of the first day that we want to analyze

```
start = 1501545600
end = 1501631999
```

Иницијализација свих листи које су потребне за чување података

#Initialization of every list

```
dates = []
comments = []
scores = []
positive_sentiment = []
negative_sentiment = []
positive_slash_negative = []
sentiment_score = []
```

Увоз података из текстуалних датотека

#Import of positive and negative pools of words

```
p = open("positive.txt", 'r')
n = open("negative.txt", 'r')

x = p.read().split("\n")
y = n.read().split("\n")
```

**Кодирање петље у којој се шаље параметризован захтев ка сервису,
након чега се одвајају подаци из одговора самог сервиса**

```
i = 1
number_of_days = 365
while i < number_of_days:
    #URL creation
    url = 'https://api.pushshift.io/reddit/submission/search/?
after='+str(start)+'&before='+str(end)
+'&sort_type=score&sort=desc&subreddit=Bitcoin&category=best&size=30'
    r = requests.get(url)

    #Data export from the Response
    responseData = r.json()
    data = responseData['data']

    #Reseting all the counters and scores to 0 for the new posts
    counter = 0
    positive_score = 0
    negative_score = 0
    total_score = 0
    total_comments = 0
    for post in data:
        #Data export from the post
        post_id = post['id']
        title = post['title']
        comment = post['num_comments']
        score = post['score']
        if 'selftext' not in post:
            text = ""
        else:
            text = post['selftext']

        total_score = total_score + score
        total_comments = total_comments + comment

    positive_counter=0
    negative_counter=0
```

Сентимент анализа података на основу речи из датотека

```
#TITLE ANALYSIS
```

```
title = re.sub('[, .! ? ; :]', ' ', title)
words_title = re.findall(r"[\w']+", title)
for word in words_title:
    for positive_word in x:
        if positive_word == word.lower():
            positive_counter = positive_counter + 1
            break
    for negative_word in y:
        if negative_word == word.lower():
            negative_counter = negative_counter + 1
            break
```

```
#IF TEXT AVAILABLE TEXT ANALYSIS
```

```
if text != "":
    text = re.sub('[, .! ? ; :]', ' ', text)
    words_text = re.findall(r"[\w']+", text)

    positive_counter=0
    negative_counter=0
    for word in words_text:
        for positive_word in x:
            if positive_word == word.lower():
                positive_counter = positive_counter + 1
                break
        for negative_word in y:
            if negative_word == word.lower():
                negative_counter = negative_counter + 1
                break
    if text == "":
        counter = counter + 1
    positive_score = positive_score + positive_counter
    negative_score = negative_score + negative_counter
```

Трансформација датума из формата *timestamp* у *datetime*

```
#Date transformation from timestamp to datetime  
date = time.strftime('%Y-%m-%d', time.localtime(start))
```

Попуњавање листи са одговарајућим подацима

```
#Adding all the elements to their lists  
dates.append(date)  
comments.append(str(total_comments))  
scores.append(str(total_score))  
positive_sentiment.append(str(positive_score))  
negative_sentiment.append(str(negative_score))  
positive_slash_negative.append(str(positive_score/negative_score))  
sentiment_score.append(str(positive_score-negative_score))
```

Обавештавање корисника да је завршена анализа за тај дан и итерирање на следећи

```
#Writing out to the console that the data for this date have been finished  
day = {"date":date,"positive score":positive_score,"negative  
score":negative_score}  
print(day)  
  
#Changing the iterator for the next day  
i=i+1  
start+=86399  
end+=86399
```

Извоз података из скрипте у табелу

#Data export to datasheet

```
new_dataframe = pandas.DataFrame(  
    {  
        "Date": dates,  
        "Num of comments on 30 posts": comments,  
        "Score on 30 posts": scores,  
        "Positive Sentiment": positive_sentiment,  
        "Negative Sentiment": negative_sentiment,  
        "P/N": positive_slash_negative,  
        "Sentiment Score": sentiment_score  
    }  
)  
writer = pandas.ExcelWriter('output.xlsx')  
new_dataframe.to_excel(writer, 'Sheet1')  
writer.save()
```