

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА
АКАДЕМІЯ»

Кафедра інформатики факультету інформатики

**Механізми алготрейдингу криптоактивами зі штучним
інтелектом**

**Текстова частина до курсової роботи
за спеціальністю «Системний аналіз» - 124**

Керівник курсової роботи
Канд. економ. наук, ст. викладач
Невмержицький Є. І.

(підпис)

“ ____ ” _____ 2021 р.

Виконав студент СА-1

Чумак В.В.

“ ____ ” _____ 2021 р.

Київ 2021

Календарний план виконання курсової роботи

Тема: Механізми алготрейдингу криптоактивами зі штучним інтелектом

Календарний план виконання роботи:

№ п/п	Назва етапу роботи	Термін виконання	Примітка
1	Отримання завдання	20.10.2020	
2	Огляд літератури за темою	20.12.2020	
3	Аналіз технології блокчейн	15.02.2021	
4	Написання алгоритмів прогнозування курсу криптовалюти	20.03.2021	
5	Написання текстової частини роботи	07.04.2021	
6	Розробка практичної частини	25.04.2021	
7	Створення слайдів для доповіді	05.05.2021	
8	Остаточне оформлення роботи та слайдів	10.05.2021	

Студент _____

Керівник _____

“ _____ ” _____

Зміст

Зміст	3
Анотація.....	4
Вступ.....	5
1. Принципи роботи технології блокчейн.....	6
1.1. Що таке блокчейн?	6
1.2. Біткоїн – перше застосування блокчейну	7
2. Поняття криптовалют і криптоактивів.....	11
2.1. Що таке криптовалюта?.....	11
2.2. Методи захисту системи.....	12
3. Поняття трейдингу та волатильності	15
3.1. Трейдинг: основні поняття	15
3.2. Волатильність: поняття та способи обрахунку	17
4. Алгоритми прогнозування курсу криптовалют	20
4.1. Опис засобів програмування	20
4.1.1. Обрані технології (середовище, мови)	20
4.1.2. Опис використаних “3-rd party services”	20
4.2. Опис та реалізація алгоритмів.....	21
4.2.1. Прогнозування курсу на основі новин з використанням NLP	21
4.2.1.1. Опис алгоритму	21
4.2.1.2. Результат алгоритму	23
4.2.2. Прогнозування за допомогою рекурентної нейронної мережі.....	24
4.2.2.1. Опис алгоритму	24
4.2.2.2. Результат алгоритму	26
Висновки.....	28
Список використаних джерел.....	30

Анотація

У цій роботі описуються способи аналізу ринку для вдалого трейдингу. Робота присвячена дослідженню відносно нової галузі – блокчейну та криптовалют, вивченню поняття трейдингу та способам досягання успіху в цій справі.

До уваги беруться автоматизовані способи аналізу ринку, алгоритми вивчення та передбачення ціни на криптовалюту, для прикладу розглядається біткоїн. В роботі було описано, реалізовано та протестовано два алгоритми прогнозування курсу біткоїна: перша – модель, яка будує передбачення на основі новин, для цього було використано сервіс OpenBlender, який надає можливість зв'язати зміну курсу зі світовими новинами, які потім розбиваються на дві підгрупи: позитивні (курс пішов вгору), негативні (курс не піднімався); другий – 4-шарова нейронна мережа формату LSTM з використанням TensorFlow, в якій інформація про валюту зберігається у вигляді послідовностей, на основі яких і робляться прогнози.

Результати першого алгоритму – дні, коли алгоритм «вгадав» (передбачив) підйом курсу та відсоток точності такого прогнозу. Результати другого – прогнози на тестовій вибірці, яка була відсічена від загальної у ході побудови, візуальне відображення – графіки ціни, яку передбачала мережа та справжньої ціни для порівняння.

Один з алгоритмів було використано для трейдингу валютою Dogecoin, яка дуже залежить від новин, особливо твітів (меседжі в соціальній мережі Twitter), результатом став приріст початкових вкладів на 132%.

Вступ

Криптовалюти все більше стають невід'ємною частиною сучасного життя кожної людини. Щодня з'являються новини про різку зміну курсу біткоїна(BTC), про створення нових криптовалют, які, за думкою експертів, перевершать попередні.

Такі поняття як «блокчейн», «криптоактив», «трейдинг» вже не здаються чимось новим, а навпаки все більше стають буденністю. Новини про твіти Ілона Маска, які піднімають курс Dogecoin в 10 разів, про інвестування компанією Tesla в біткоїн – все це вже реальність, якою б далекою для деяких людей вона не була.

Світ рухається вперед, і, якщо раніше люди заробляли на обміні валют, то зараз це переходить в криптовалюту – так зване купи за 100, а продай за 200. Чи можна бути впевненим, що купивши за 100, ти точно знайдеш момент і зможеш продати за 200? Звісно ж ні, інвестування – це ризик, але його можна і потрібно мінімізувати. Маючи достатню базу знань в цій сфері, можна проаналізувати ринок і знайти найкращий варіант вкладення своїх грошей.

Для аналізу ринку необов'язково сидіти і вивчати графіки, читати наукові статті, адже в епоху цифрових технологій на допомогу приходить комп'ютер, який в правильних руках зробить більше, ніж будь-який фінансовий аналітик. Алгоритми для прогнозування курсів криптовалют оперують величезними обсягами даних, порівнюють попередні Open Price та Close Price, створюють модель, яка б змогла передбачити ріст тієї чи іншої криптовалюти.

Ця робота присвячена реалізації та порівнянню таких алгоритмів, аналізу переваг та недоліків кожного з них. Для цього було розв'язано наступні задачі:

- 1) Пошук вже існуючих систем прогнозування та їх аналіз
- 2) Написання алгоритмів прогнозування з використанням Open Data
- 3) Аналіз роботи та порівняльна характеристика алгоритмів

1. Принципи роботи технології блокчейн

1.1. Що таке блокчейн?

Технологія блокчейн – це, мабуть, найкращий винахід Інтернету. Вона дозволяє обмінюватися цінною інформацією без необхідності довіри або наглядача вище. Уявіть, два користувачі ставлять 50 доларів на завтрашню погоду в Києві. Перший впевнений, що буде сонячно, другий – чекає дощу. У них є три варіанти управління цією транзакцією:

- Довіра одне одному. Переможений зобов'язується віддати переможцю 50 доларів. Якщо користувачі друзі, це може бути хорошим способом управління цим. Однак, якщо ні, немає гарантій, що виграш буде виплачений.
- Перетворення ставки на контракт . За умови укладеного контракту обидві сторони будуть більш схильні платити. Однак, якщо хтось із двох вирішить не платити, переможець повинен буде заплатити додаткові гроші для покриття судових витрат, і судовий процес може зайняти багато часу. Особливо для невеликої суми готівки це не здається оптимальним способом управління транзакцією.
- Залучення нейтральної третьої сторони. Кожен з користувачів віддає 50 доларів третій стороні, яка виплатить загальну суму переможцю. В цьому випадку ніхто не гарантує, що третя сторона не зникне з усіма грошима, залишивши учасників суперечки з пустими руками. Отож залишається вибрати один із перших двох варіантів: довіру чи контракт.

Ні довіра, ні договір не є оптимальним рішенням: не можна довіряти незнайомцям, а виконання контракту вимагає часу та грошей. Технологія блокчейну цікава тим, що пропонує третій варіант, який є безпечним, швидким і дешевим.[3]

Вона дозволяє написати програму, в яку обидва учасники суперечку зможуть надіслати свої ставки, а після оголошення результатів, переможець автоматично отримає свій виграш. Для ставок в 50 доларів це можливо «вистріл з гармати по горобцях», але якщо мова йде про продаж будинку, бізнесу – це швидкий, дешевий і найбезпечніший спосіб.

Блокчейн – розподілена база даних, що зберігає впорядкований ланцюг записів(блоків), що постійно збільшується. Кожен з блоків містить часову мітку, хеш попереднього запису та дані транзакцій, подані як хеш-дерево.

1.2. Біткоїн – перше застосування блокчейну

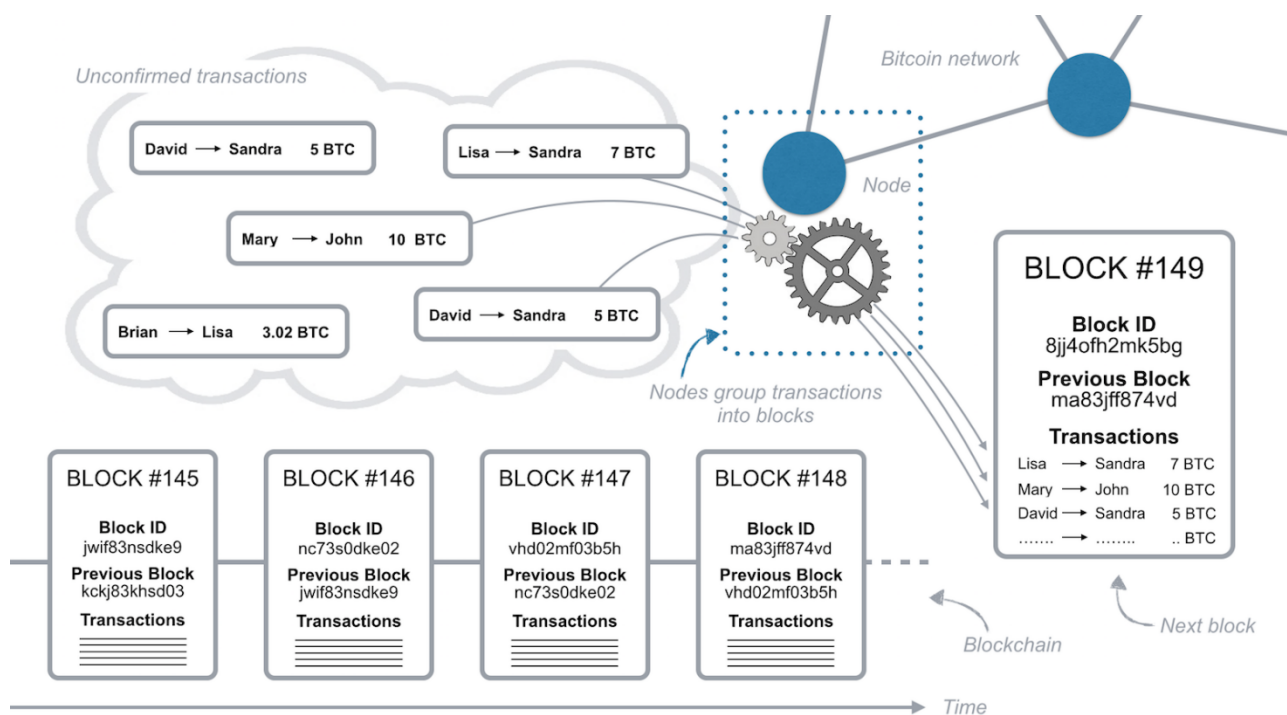
Одна з головних переваг блокчейну – децентралізація. Розглянемо це поняття на прикладі роботи біткоїну – найбільш відомому застосуванні технології блокчейн. Уявіть, що компанія володіє сервером, що складається з 10 000 комп'ютерів, у базі даних зберігається вся інформація про рахунок клієнта. Ця компанія має склад, що містить усі ці комп'ютери під одним дахом, і має повний контроль над кожним із цих комп'ютерів та всією інформацією, що міститься в них.[1] Подібним чином, біткоїн складається з тисяч комп'ютерів, але кожен комп'ютер або група комп'ютерів, що містять свій блокчейн, знаходяться в іншому географічному розташуванні, і всі вони експлуатуються окремими особами або групами людей. Ці комп'ютери, що складають мережу біткоїнів, називаються вузлами.

У блокчейні кожен вузол має повний запис даних, які зберігались у блокчейні з моменту його створення. Якщо один вузол має помилку у своїх даних, він може використовувати тисячі інших вузлів як орієнтир, щоб виправити себе. Таким чином, жоден вузол у мережі не може змінити інформацію, що зберігається в ній. Через це історія транзакцій у кожному блоці, що становить блокчейн біткоїна, незворотна. Якщо один користувач втручається в запис транзакцій, всі інші вузли перетинають один одного і легко визначають

вузол з неправильною інформацією. Ця система допомагає встановити точний і прозорий порядок подій. [2]

Для біткоїна ця інформація є переліком транзакцій, але також можливо, щоб блокчейн містив будь-яку інформацію, таку як юридичні контракти, державні ідентифікаційні дані або інвентаризація товарів компанії.

Для відстеження кількості біткоїнів, якими володіє кожен користувач, блокчейн використовує книгу – цифровий файл, який відстежує всі операції з біткоїнами. Неформально блокчейн можна визначити як систему, яка дозволяє групі підключених комп'ютерів вести єдину оновлену та безпечну книгу.[3]



Зображення 1

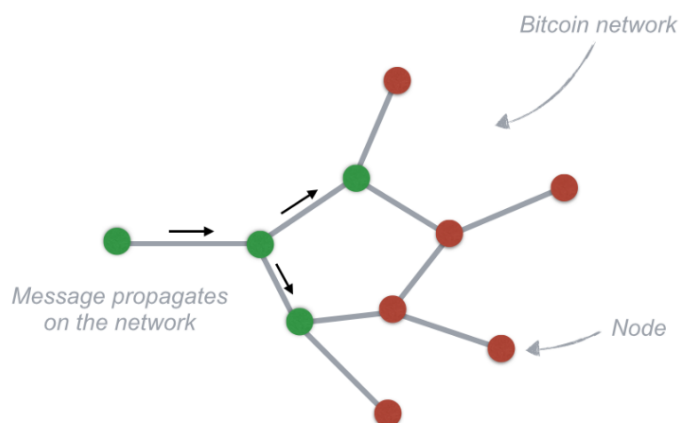
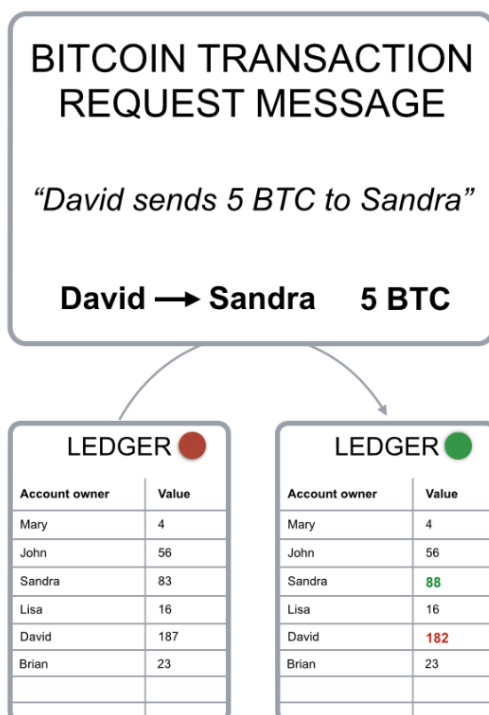
Те, що книга ведеться групою підключених комп'ютерів, а не централізованою організацією, такою як банк, має кілька наслідків:

- У банківській системі користувач знає лише власні транзакції та залишки на рахунках; на блокчейні кожен може бачити транзакції всіх інших.

- Хоча ви, як правило, можете довіряти своєму банку, мережа біткоїнів розподіляється, і якщо щось піде не так, немає служби довідки, яку можна зателефонувати, або когось, щоб подати на нього позов.
- Система блокчейну спроектована таким чином, що не потрібна довіра, як така, адже безпека та надійність отримуються за допомогою спеціальних математичних функцій та алгоритмів.

Для здійснення транзакцій на блокчейні потрібен гаманець – програма, яка дозволяє зберігати та обмінювати свої біткоїни. Оскільки витратити свої збереження може лише сам користувач, кожен гаманець захищений спеціальним криптографічним методом, який використовує унікальну пару різних, але пов'язаних ключів: приватний і відкритий ключі. Якщо повідомлення зашифровано за допомогою певного відкритого ключа, лише власник парного приватного ключа може його розшифрувати та прочитати, і навпаки, якщо користувач зашифрує повідомлення своїм приватним ключем, розшифрувати його зможе лише парний відкритий ключ.

Приклад біткоїн транзакції[3]:



Each *node* receives the transaction request message, updates its own copy of the *ledger* and passes on the message to the nearby *nodes*.

Девід надсилає 5 BTC (біткоїнів) Сандрі: повідомлення про цю транзакцію розходить по всій мережі, кожен вузол отримує його, оновлює свою копію книги і надсилає повідомлення до сусідніх вузлів.

Ще одна перевага такої системи – анонімність. Користувачам дозволено створювати скільки завгодно гаманців і виконувати транзакції, саме тому неможливо відслідкувати справжнього власника цих гаманців. Біткоїн-адреси генеруються за допомогою RIPE-MD160 (RACE Integrity Primitives Evaluation Message Digest) хеш функції, яка створює 160-розрядне хеш-значення для вхідного повідомлення. Отже, загальна кількість можливих біткоїн-адрес становить 2^{160} . За останніми підрахунками в мережі біткоїн зараз лише близько 500 мільйонів адрес.[4]

Отже, блокчейн - вибудований за певними правилами безперервний послідовний ланцюжок блоків, що містять інформацію. Система складається з вузлів, які містять копії ланцюжків блоків, що дозволяє перевіряти правдивість інформації на окремих вузлах. Технологія з'явилась у 2009 році й запропонувала новий принцип реєстрації транзакцій, у 2013 платформа Ethereum ввела в обіг смарт-контракти, які дозволяють автоматично виконати операцію, за умови дотримання певних умов, при цьому зберігаючи їх в блокчейні. Таким чином виключається ризик третьої сторони, зменшуються посередники, витрати. Великі компанії все більше вивчають можливості технології та її застосування, серед них IBM, Microsoft, Samsung.

Це швидка, дешева та надійна система для управління та передачі інформації. Наразі широко використовується для створення нових криптовалют, для безпечної передачі даних, управління ідентифікаційною інформацією, міжнародні платежі, захист авторського права, електронне голосування, анонімна передача повідомлень. [5]

2. Поняття криптовалют і криптоактивів

2.1. Що таке криптовалюта?

Криптовалюта – різновид цифрової валюти, захищеної криптографією, що робить підробку даних майже неможливим, емісія та облік якої виконується децентралізовано та в автоматичному режимі.[6] Визначальна особливість криптовалют – це те, що вони, як правило, не випускаються жодним центральним органом влади, що робить їх незалежними від втручання уряду.

Криптовалюти – аналоги традиційних (фіатних) валют. Згідно з визначенням Bitcoin Security: криптовалюта – швидка та надійна система платежів і грошових переказів, заснована на новітніх технологіях і непідконтрольна жодному уряду. Офіційний сайт Bitcoin визначає криптовалюти як новий вид грошей та інноваційну систему платежів, побудовану на P2P-технології, в якій емісія та обробка транзакцій відбуваються колективно, тобто зусиллями мережі.

Рейтинг криптовалют за рівнем капіталізації (станом на 6 травня 2021 року):

Назва криптовалюти	Поточна ціна (USD \$)	Ринкова капіталізація млрд. USD \$
Bitcoin	\$58,349.29	1087.5
Ethereum	\$3,501.15	404.3
Binance Coin	\$649.85	99.6
Dogecoin	\$0.6032	79.2
XRP	\$1.67	75.9

Таблиця 1

Все більш популярною останні місяці стає криптовалюта Dogecoin, яка була створена ще в кінці 2013 року, але нещодавно набрала популярності завдяки твітам Ілона Маска, який назвав її своєю улюбленою криптовалютою. Маючи ціну менше 1го долара, вона посідає 4те місце у світі за капіталізацією, що

пов'язано з великим запасом випущених одиниць, їх у Doge – майже 130 млрд., тоді як у біткоїна – майже 19млн.

Найпопулярніший та на даний момент найцінніший представник – біткоїн – криптовалюта, заснована в 2009 році групою людей під псевдонімом Сатосі Накамото. Більшість сучасних криптовалют – клони або форки (forks) біткоїна, деякі ж – побудовані з нуля. Саме на прикладі біткоїну будуть розглянуті такі поняття як трейдинг та можливі варіанти прогнозування курсу.

2.2. Методи захисту системи

Криптовалюти покликані спростити переказ коштів безпосередньо між двома сторонами без необхідності довіреної третьої сторони, такої як банк або компанія, що видає кредитні картки. Натомість ці передачі забезпечуються використанням відкритих та приватних ключів та різних форм заохочувальних систем, таких як Proof of Work або Proof of Stake .

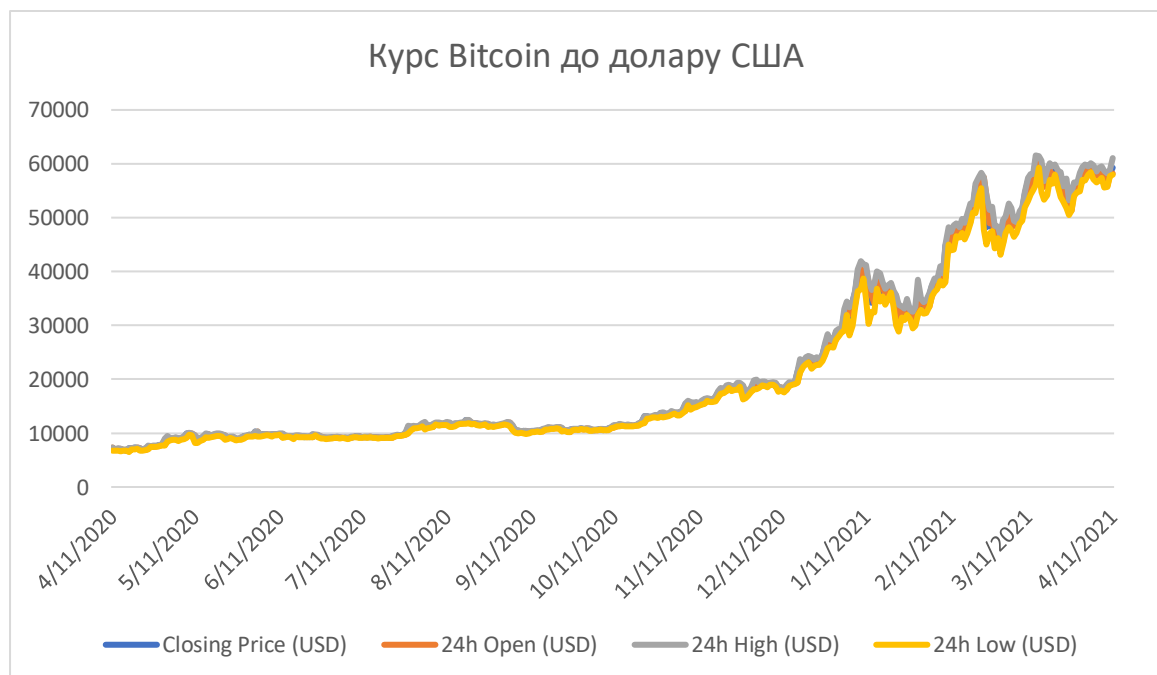
Proof of Work – це принцип захисту систем від зловживання послугами, заснований на необхідності виконання стороною, яка робить запит (клієнтом) деякої досить складної тривалої, результат якої легко і швидко перевіряється стороною, що обробляє запит (сервером). У системі Bitcoin використовується багаторівневе хешування — хеш попереднього блоку стає частиною наступного, таким чином неможливо змінити блок без зміни хешів у всіх наступних блоках, а для цього потрібен перерахунок кожної хеш суми, що робить цю задачу дуже затратною.

Proof of Stake – це метод захисту, заснований на необхідності доказу зберігання певної кількості коштів на рахунку. Використання цього методу означає, що алгоритм криптовалюти задля підтвердження чергового блоку в ланцюжку вибере обліковий запис з великою кількістю коштів на рахунку. Цей принцип використовується як альтернатива методу Proof of work, в якому більшу ймовірність бути обраним задля підтвердження блоку має обліковий запис з великими обчислювальними потужностями.

У сучасних криптовалютних системах «гаманець» користувача або адреса рахунку має відкритий ключ, тоді як приватний ключ відомий лише власнику і використовується для підписання транзакцій. Перекази коштів здійснюються з мінімальними комісіями за обробку, що дозволяє користувачам уникати високих стягнень, які зазвичай застосовуються банками та фінансовими установами за банківські перекази .

Однак існує й інша сторона: криптовалюти часто використовуються для протизаконного бізнесу, оскільки вони майже повністю анонімні і відслідкувати власників гаманців практично неможливо. Саме це робить їх придатними для таких дій як відмивання грошей, ухилення від податків торгівля незаконними товарами.

Ще одна суперечність – нестабільний курс криптовалюти. Для когось – це шанс заробити, а хтось, маючи криптовалюту, втрачає на змінах курсу. На графіку задано курс криптовалюти біткоїн до долару США за останній рік[7].



Діаграма 1

Аналізуючи графік, можна помітити, що загалом тренд курсу – це більше про зростання, але чітко видно наскільки він не стабільний останні декілька

місяців, це і визначає ще одну проблему криптовалют – нестабільність курсу. Для розуміння порівняємо це з курсом реальної валюти (гривні) до долару США за той самий період[8].



Діаграма 2

З першого погляду здається, що графік більш стрибавий, ніж попередній, але якщо розглянути різницю між найбільшим і найменшим значенням за рік, то у випадку реальної валюти вона становить майже 2 гривні, у випадку криптовалют – більше 50000 тис. доларів США. Звісно, таке порівняння не зовсім коректне, адже це дві різні речі, але воно наявно показує одне – нерегульована валюта дуже не стабільна і як сказав би «справжній криптотрейдер»: « Цим потрібно користуватися.»

Отже, криптовалюта – наразі основне застосування технології блокчейн. Останні підрахунки показали, що капіталізація усіх криптовалют світу вже перевищила 2трлн. доларів США, і ця цифра була досягнута всього за три місяці після здобуття 1трлн. доларів США, що показує наскільки це поняття стає актуальним та популярним у сучасному світі.

3. Поняття трейдингу та волатильності

3.1. Трейдинг: основні поняття

Трейдинг – участь у торгах на валютній, фондовій, товарній біржі. Інтернет-трейдинг – це спосіб доступу до даних торгів за допомогою інтернету як засобу зв'язку. Суть трейдингу полягає в найважливішому правилі будь-якої торгівлі: купувати дешевше, а продавати дорожче.

Для того, щоб заробляти на трейдингу, потрібно вміти шукати правильний момент для входу на ринок та виходу з нього. Проте 80% трейдерів відкривають угоди, базуючись на власних емоціях та інтуїції, що на руку великим біржовим гравцям, які заробляють саме на невдачах таких новачків.

Секрет успішних угод криється у виключенні впливу інтуїції та, особливо, емоцій і у використанні аналітичних даних. Виділяють два основні типи аналізу: фундаментальний – вивчення економіки в цілому; технічний – базується на вивченні графіку курсу фінансового інструменту, вважається, що саме історія зміни вартості відображає психологію натовпу.

Технічний аналіз базується на трьох основних постулатах, сформованих Чарльзом Доу ще в 1900 році:

1) Ринок враховує все

Зміна ціни – наслідок величезної кількості чинників (економічних, політичних, психологічних), які впливають на неї.

2) Рух ринку підкоряється тенденціям

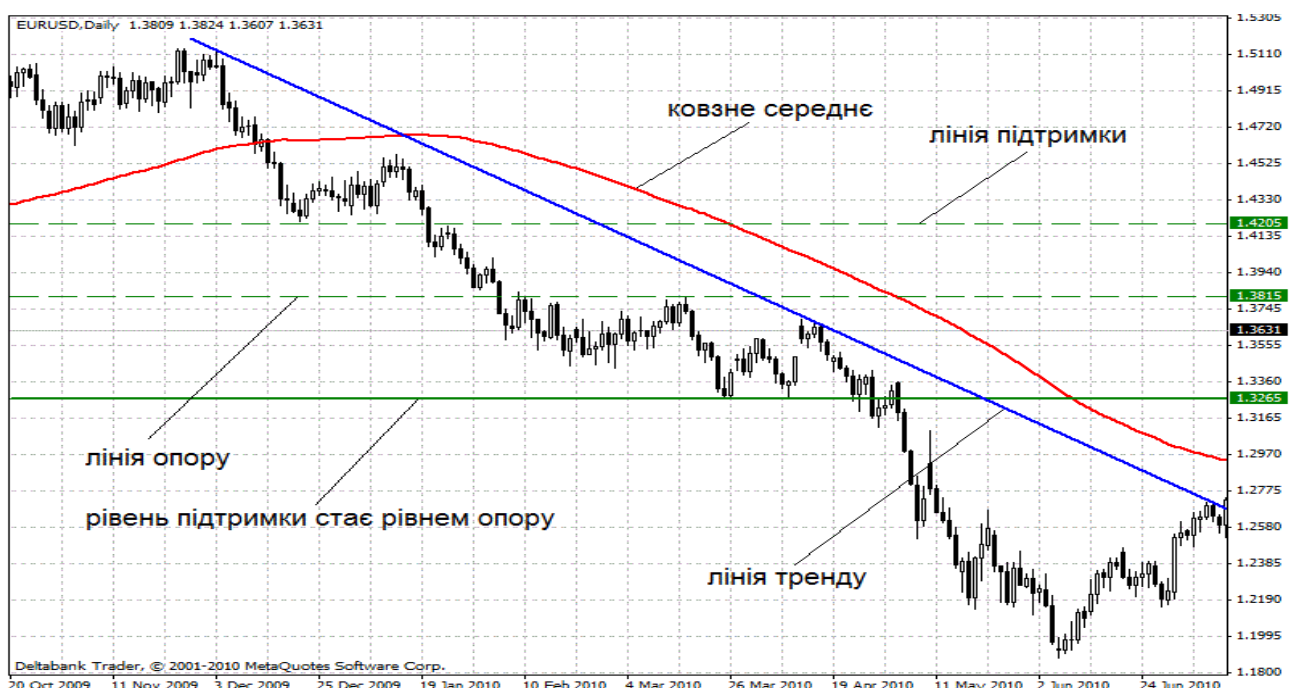
Життя ринку ділиться на періоди зростання та спадання цін, а в середині них відбувається розвиток пануючої тенденції, яка існує до точки розвороту ринку.

3) Історія повторюється

Визначені конфігурації на графік цін мають властивість повторюватись з часом і багаторазово.

Для технічного аналізу ринку використовується купа різних індикаторів та показників, до найбільш важливих відносяться[9]:

- Початкова інформація:
 - Ціна здійснення операції
 - Об'єм торгів
 - Попит і пропозиція ринку.
- Інерційна інформація:
 - «Японські свічки» – показник розмаху ціни (волатильності)
 - Ковзні середні – напрям довгострокового тренду
- Швидкість і прискорення ринку
 - Моментум – аналіз швидкості і напрямку зміни ціни (перша похідна)
 - Стохастичний індикатор – аналіз прискорення зміни ціни (2-га похідна)
 - Індекс відносної сили RSI – порівняння швидкостей росту та падіння ціни у певний проміжок часу
- Інші індикатори
 - Лінії тренду, Болінджера, підтримки та опору
 - Спрямованість ціни та об'єму ринку
 - Сходження та розбіжність ковзних середніх



3.2. Волатильність: поняття та способи обрахунку

Волатильність – це статистичний показник, який характеризує тенденцію зміни ціни. Волатильність є найважливішим фінансовим показником в управлінні фінансовими ризиками, де представляє собою міру ризику використання фінансового інструменту за заданий проміжок часу, та виражається в абсолютному або у відносному від початкової вартості значенні.[10]

Наприклад, за останній місяць ціна української гривні відносно долару США змінилась на 7%, саме ці 7% і називаються волатильністю гривневої валюти за останній місяць. Розрізняють два види волатильності:

- Історична волатильність – це величина, що дорівнює стандартному відхиленню вартості фінансового інструменту за заданий проміжок часу, розрахованому на основі історичних даних про його вартість.
- Очікувана волатильність – величина, обчислена на основі поточної вартості фінансового інструменту в припущенні, що ринкова вартість фінансового інструменту відображає очікувані ризики.

Волатильність показує наскільки сильно ціна активу коливається навколо середньої ціни, тому більш волатильні (нестабільні) активи вважаються ризиковими, оскільки ціна стає менш передбачуваною.

Найпростіший підхід до визначення волатильності цінного паперу - це обчислення стандартного відхилення його цін протягом певного періоду часу.

Нехай аналітику потрібно порахувати волатильність акцій ABC Corp. за останні 4 дні. Початкові ціни наступні: 10, 12, 9, 14 доларів відповідно. Для цього потрібно виконати наступне[10]:

- 1) Пошук середньої ціни: $(10+12+9+14)/4 = 11.25$
- 2) Пошук різниць початкових цін з середньою у квадратичній формі:
 1. $(10-11.25)^2 = 1.56$
 2. $(12-11.25)^2 = 0.56$
 3. $(9-11.25)^2 = 5.06$
 4. $(14-11.25)^2 = 7.56$
- 3) Пошук суми квадратних різниць: $1.56+0.56+5.06+7.56 = 14.75$
- 4) Пошук дисперсії: $14.75/4 = 3.69$
- 5) Пошук стандартного відхилення: $\sqrt{3.69} = 1.92$

Стандартне відхилення вказує на те, що ціна акцій ABC Corp. зазвичай відхиляється від середньої ціни акцій на \$1,92.

Більш складний підхід до обчислення стабільності активу – використання Бети (β) – показника волатильності активу відносно всього ринку. Компанія з вищою бета-версією має більший ризик, а також більший очікуваний прибуток. Бета-коефіцієнт можна інтерпретувати так:

- $\beta = 1$ компанія точно така ж нестабільна, як і ринок
- $\beta > 1$ більш волатильний, ніж ринок
- $0 < \beta < 1$ менш волатильний, ніж ринок
- $\beta = 0$ не пов'язано з ринком
- $\beta < 0$ негативно корелює з ринком

Бета-показник цінних паперів обчислюється шляхом ділення добутку *коваріації* прибутків цінних паперів та прибутковості ринку на *дисперсію* прибутковості ринку протягом певного періоду.

$$\beta = \frac{\text{Коваріація}(R_e, R_m)}{\text{Дисперсія}(R_m)}$$

R_e – рентабельність активу

R_m – рентабельність ринку

Коваріація – міра того, як змінюються прибутки акцій пов'язані зі змінами прибутковості ринку.

Розрахунок бета-показника використовується, щоб допомогти інвесторам зрозуміти, чи рухається акція в тому ж напрямку, що і решта ринку. Він також надає уявлення про те, наскільки волатильні або наскільки ризиковані акції щодо решти ринку. Щоб бета-показник міг надати корисну інформацію, ринок, який використовується як еталон, повинен бути пов'язаний з активом.

4. Алгоритми прогнозування курсу криптовалют

4.1. Опис засобів програмування

4.1.1. Обрані технології (середовище, мови)

Для реалізації обох алгоритмів було використано мову програмування Python версії 3.8 x64. Python – інтерпретована мова програмування зі строгою динамічною типізацією, часто використовується для роботи з Big Data, Machine Learning, Neural Networks, адже містить широкий спектр потрібних бібліотек для обрахунку та візуалізації результатів. Мова відноситься до мов з «низьким порогом входу», адже вона дуже просто для розуміння та використання.

Для програмування на Python було використане середовище розробки PyCharm від JetBrains – одне з найпопулярніших, та найзручніших середовищ. Компанія надає студентську підписку на час навчання, що дозволяє безкоштовно і в повному обсязі використовувати їх середовища на час навчання в університеті. Для Могилянців така підписка працює по корпоративній пошті.

4.1.2. Опис використаних “3-rd party services”

TensorFlow – відкрита бібліотека для машинного навчання, побудови та тренування нейронних мереж, розроблена компанією Google. В алгоритмі прогнозування використовуються можливість створення, компіляції, тренування та тестування моделі.

CoinDesk – головний дата-провайдер в області криптовалют, використовується для збору даних про ціну криптовалют в часі.

OpenBlender – один з найбільших порталів датасетів для машинного навчання: велика кількість потрібної інформації, зручний спосіб поєднання таблиць, легка обробка та збір даних. [13]

4.2. Опис та реалізація алгоритмів

4.2.1. Прогнозування курсу на основі новин з використанням NLP

4.2.1.1. Опис алгоритму

Новини – один з факторів зміни ціни криптовалют, саме тому їх варто використовувати для моделі, яка прогнозуватиме подальший курс. Для прикладу буде розглянуто біткоїн, а за допомогою сервісу OpenBlender взято інформацію про ціну криптовалюти впродовж потрібного періоду. Алгоритм буде передбачати зростання або падіння ціни за наступний торговий день, базуючись на попередніх цінах та новинах, які їм передували[11].

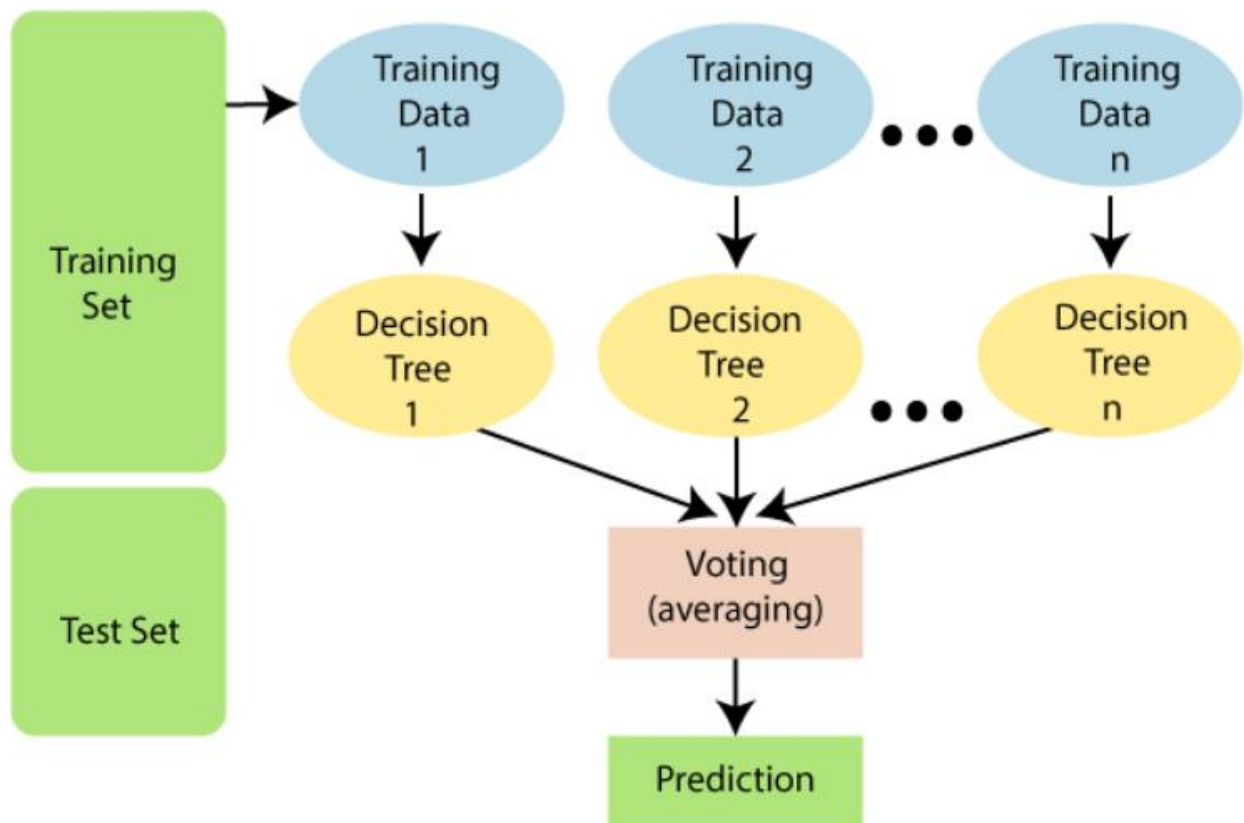
Для обробки новин використовується NLP (Natural Language Processing) – спосіб обробки природної мови машинами, ця технологія вже використовується в медицині, в зупинці спаму на поштових скриньках, в індексуванні фальшивих (замовних) новин, і, звичайно, відмінно спрацює на аналізі коментарів, новин, чуток при прогнозуванні курсу криптовалюти.

Суть алгоритму:

- 1) Збір даних: щодобові відкриваючі та закриваючі ціни, відсоток зміни ціни, найвище та найнижче значення за добу.
- 2) Обчислення логарифмічної різниці між цінами закриття та відкриття для кожної доби.
- 3) Визначення цільової змінної: 1, якщо логарифмічна різниця додатня (ціна за добу зросла); 0, якщо – недодатня (ціна за добу не зросла).
- 4) Збір даних по новинах: зв'язка по часу (timestamp), поділ на «позитивні» (такі, після яких ціна зросла) та «негативні» (такі, після яких ціна не зросла) новини.
- 5) Поділ загальної вибірки на тренувальний та тестовий сети (80% та 20% відповідно).

б) Застосування Random Forest Regressor для прогнозування.

Random Forest – популярний алгоритм машинного навчання, який належить до контрольованої техніки навчання (навчання з учителем). Він може бути використаний як для проблем класифікації, так і для регресії в ML. Він базується на концепції ансамблевого навчання, що являє собою процес поєднання декількох класифікаторів для вирішення складної проблеми та поліпшення роботи моделі. Це класифікатор, який містить ряд дерев рішень для різних підмножин даного набору даних і бере середнє значення для підвищення точності прогнозування цього набору даних". Замість того, щоб покладатися на одне дерево рішень, випадковий ліс бере прогноз з кожного дерева на основі більшості голосів прогнозів і передбачає кінцевий результат[12].



Зображення 4

4.2.1.2. Результат алгоритму

Після тестування алгоритму на даних, отриманих за період 01.09.2020 – 25.04.2021, було отримано наступні результати:

```
[[11 20]
 [50 58]]
Accuracy Score:
0.49640287769784175
Precision Score:
0.7435897435897436
```

Зображення 5

Точність – 49%, алгоритм виявив майже половину змін ціни, на основі новин.

Оцінка точності – 74%, алгоритм правильно класифікував ріст або падіння ціни в 3 з 4 випадків. Алгоритм передбачив ріст ціни 78 разів, з них насправді сталися 58 підйомів, тобто 74%.

Іншими словами, незважаючи на лише половину знайдених змін ціни, алгоритм досить точно передбачав її підвищення, тобто, якщо першочерговою задачею стоїть визначення підвищень ціни для подальшого проведення торгів та уникнення падінь, модель справляється досить точно, навіть якщо доводиться жертвувати половиною випадків зміни ціни.

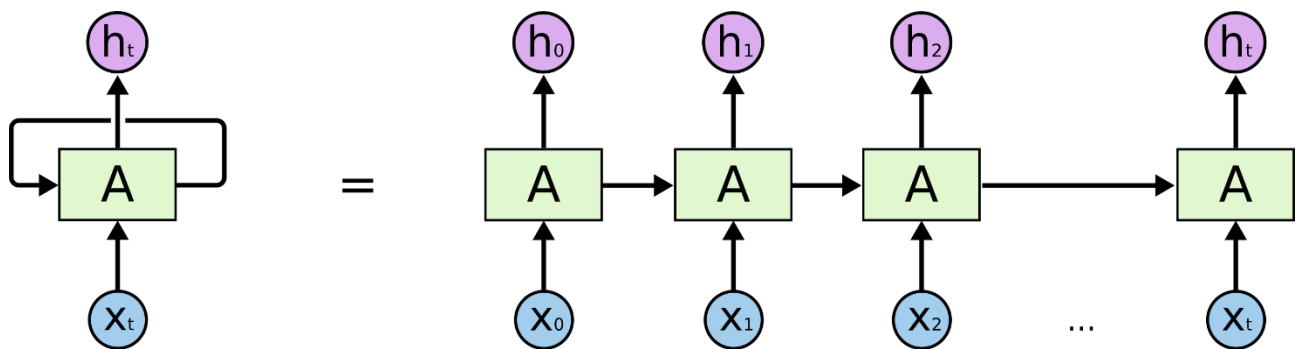
Отже, алгоритм досить точно знайшов дні росту ціни, хоч і втратив деяку частину з них. Якщо використовувати це в трейдингу – така модель допоможе «залишитись в плюсі», не зважаючи на те, що це буде досить довгострокова подія. Звичайно, не можна забувати про вплив «великих» факторів, які не передбачить жоден алгоритм (політичні явища, заяви публічних осіб, вихід великих компаній на ринок криптовалют, тощо), такі події моментально змінюють саму ціну та її тренд. Чого тільки вартує приклад з Ілоном Маском та криптовалютою Dogecoin, ціна на яку, після публічної підтримки останнім, зросла в сотні разів.

4.2.2. Прогнозування за допомогою рекурентної нейронної мережі

4.2.2.1. Опис алгоритму

RNN (рекурентні нейронні мережі) – клас нейронних мереж, які дозволяють використовувати попередні виходи як входи, маючи приховані стани.[14]

Людина не починає думати з нуля щосекунди, читаючи якусь книгу, дивлячись фільм, вона розуміє кожне слово, опираючись на знання здобуті на мить, хвилину, годину раніше. Звичайні нейронні мережі не можуть цього робити, і це є серйозним недоліком. Рекурентні нейронні мережі покликані вирішити цю проблему, завдяки циклам в середині них, які зберігають інформацію.



Зображення 6

Розглядаючи схему, наведену вище, нейронна мережа знаходиться в стані A , розглядає вхід X_t і повертає значення h_t . Цикл дозволяє передавати інформацію від одного кроку мережі до іншого[15].

Рекурентну нейронну мережу можна розглядати як кілька копій однієї і тієї ж мережі, кожна з яких передає повідомлення наступнику. Ці ланцюжки показують, що рекурентні нейронні мережі тісно пов'язані з послідовностями та списками. Вони є природною архітектурою нейронної мережі, яка використовується для таких даних. За останні кілька років був неймовірний успіх у застосуванні RNN для різноманітних проблем: розпізнавання мови, моделювання мови, переклад, субтитри до зображень. Важливим для цих успіхів

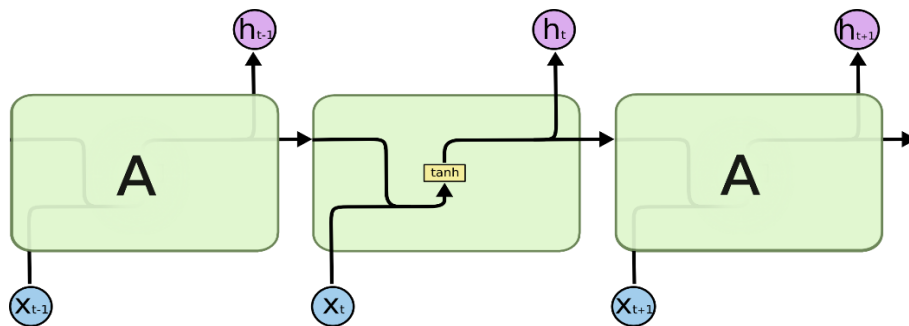
є використання “LSTM”, особливого виду періодичної нейронної мережі, яка працює для багатьох завдань набагато краще, ніж у стандартній версії.

Основною проблемою рекурентних нейронних мереж є труднощі доступу до давньої інформації, але LSTM мережі вирішують цю проблему.

Мережі довготривалої пам’яті – зазвичай їх просто називають “LSTM” – це особливий вид RNN, здатний вивчати довгострокові залежності. LSTM чітко розроблені, щоб уникнути проблеми довгострокової залежності. Запам’ятовувати інформацію протягом тривалого періоду часу – це практично їх поведінка за замовчуванням, а не те, чому вони намагаються навчитися.

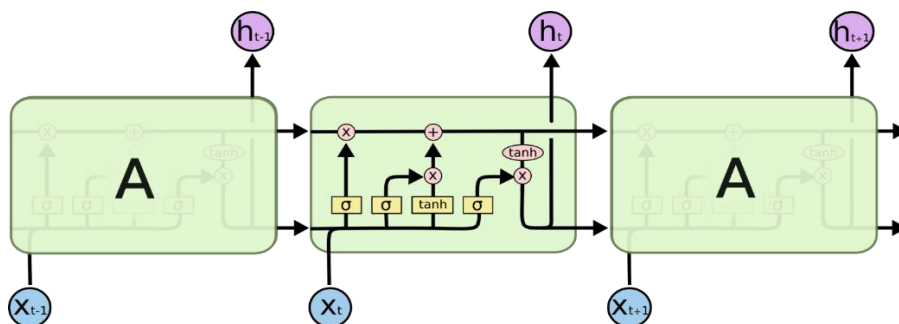
Звичайні RNN мережі мають структуру ланцюжка і лише один повторюваний модуль всередині кожного шару. LSTM також мають цю ланцюгоподібну структуру, але повторюваний модуль має іншу структуру. Замість того, щоб мати один шар нейронної мережі, існує чотири.

Схема звичайного RNN[15]:



Зображення 7

Схема LSTM[15]:



Зображення 8

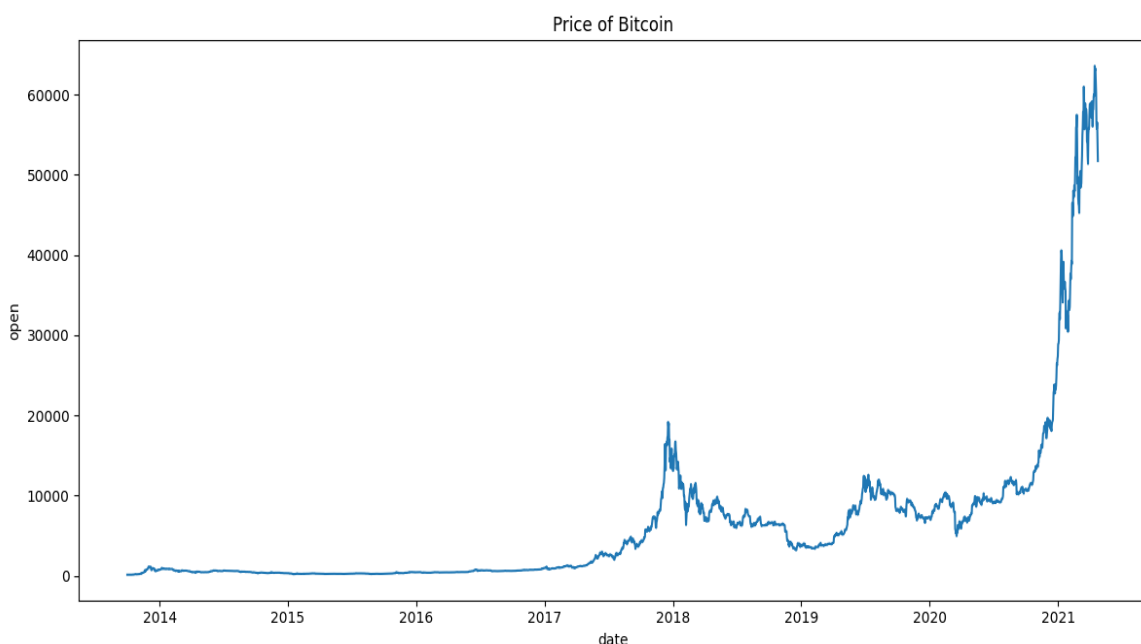
Ідея алгоритму:

Використовуючи історичні дані криптовалюти (дані по ціні за певний період), створити рекурентну нейронну мережу із використанням LSTM шарів для прогнозування тренду ціни криптовалюти. План роботи алгоритму:

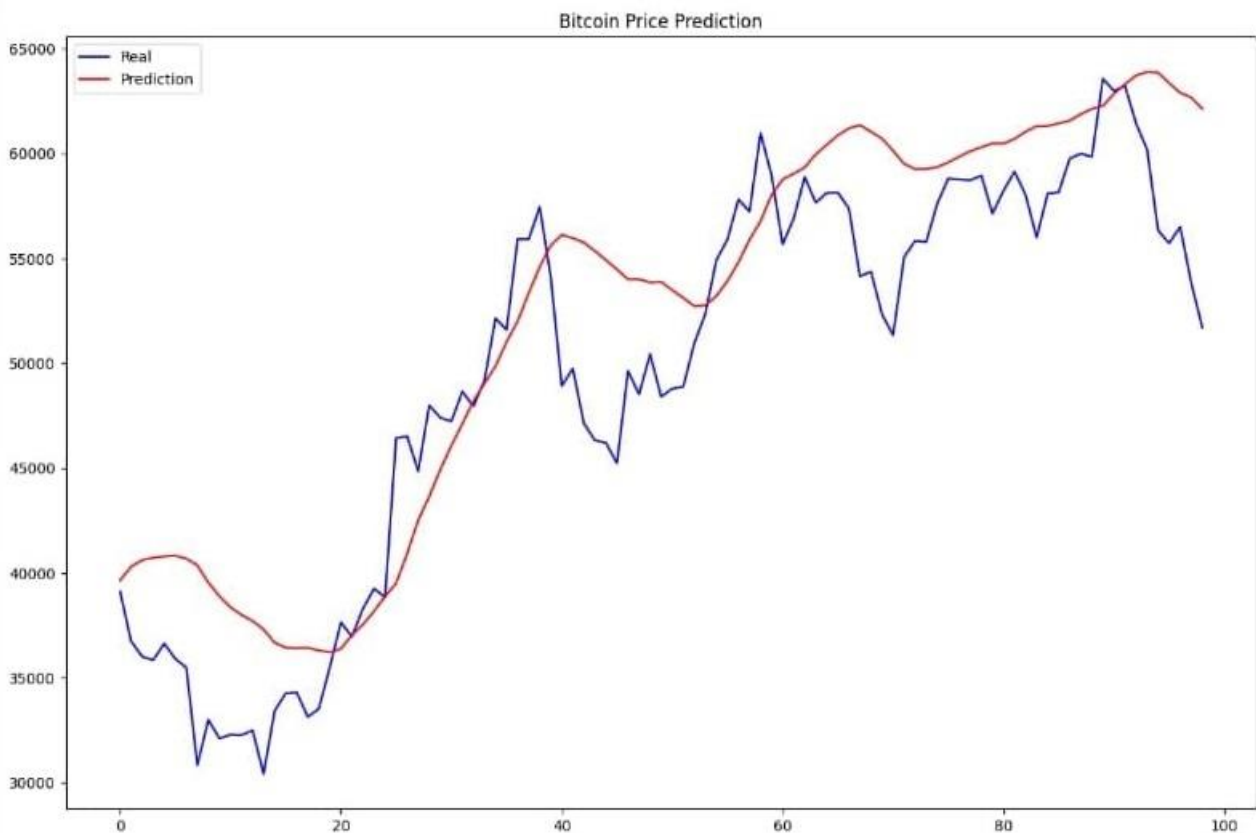
- 1) Збір даних: щодобові відкриваючі та закриваючі ціни, відсоток зміни ціни, найвище та найнижче значення за добу.
- 2) Визначення довжини послідовності та попередня обробка даних.
- 3) Нормалізація даних (стискання до відрізка $[0, 1]$).
- 4) Побудова моделі: 3 LSTM шари та 1 вихідний (dense) шар.
- 5) Компіляція моделі в TensorFlow.
- 6) Тренування початкових даних, визначення кількості «епох»
- 7) Створення прогнозів за допомогою навченої моделі.
- 8) Візуалізація результатів.

4.2.2.2. Результат алгоритму

Початкову вибірку даних було поділено на тренувальну (95%) та тестову (5%), тестова частина склала останніх 100 днів із вибірки, а тренувальна – попередні 2000 днів. Курс валюти біткоїн, побудований з вибірки даних:



Результат роботи мережі – прогноз (графік курсу) для сотні тестових днів з вибірки (синій графік – реальна ціна в той період, червоний – прогноз мережі):



Зображення 10

Дивлячись на графік, чітко помітно, що модель хоч і не спрогнозувала точну ціну щодня, але «потрапила в тренд» і ловила підйоми та падіння, не дивлячись на те, що курс біткоїну в той період був дуже нестабільним через прихід компанії Ілона Маска на ринок криптовалюти. Дана модель більше про напрямок руху курсу, ніж про точкове значення, оскільки глобальні тренди зміни ціни вона виявила точно.

В практичному випадку, для торгів ця модель гарно прогнозує тренд валюти, на відмінну від попередньої, яка могла швидко передбачити падіння або підйом ціни наступного дня. Цей алгоритм більше підходить для довгострокових вкладень, якщо спиратись на нього і купувати валюту на «низах» і чекати підйомів, які передбачила нейронна мережа.

Висновки

Після вивчення теми блокчейну та криптовалют, стає зрозумілим чому ці поняття стають все більш популярними. Це нові технології, які мають великі перспективи. Разом з тим все більше розростається трейдинг на крипто біржах. Для когось це хобі, для когось – робота, а хтось стабільно заробляє на цьому, вважаючи це своєю забавою.

Як і все у нашому світі, трейдинг одразу хочеться перевести в автоматичний режим: створити алгороботів для прибуткових торгів, розробити алгоритми та моделі, які будуть передбачувати курси валют, розробити системи для швидкого технічного аналізу, вирахування індикаторів, тощо. Більшість з цього вже доступно, і допомагає досягати високих прибутків на біржі.

Проведені експерименти показали, що передбачити курс алгоритмічно – це не такий вже й міф, звісно ж 100% гарантії досягти неможливо, але й прогноз погоди не завжди справджується, а люди досі ним користуються.

Алгоритми показали надійний результат, і один з них було використано для торгів криптовалютою Dogecoin. В результаті початковий депозит було збільшено на 123%, що можна вважати прекрасним результатом.

Проведена робота показала, що з достатніми знаннями у сфері машинного навчання та нейронних мереж, можна створювати і тестувати різноманітні моделі, для прогнозування ціни і використання цих результатів у трейдингу.

З кожним роком світ все більше переходить в Інтернет, і заробляти у всесвітній павутині – це вже не новина, але заробляти на трейдингу, користуючись не лише власним чуттям, знаннями чи обрахунками, а й результатами правильно побудованої моделі машинного навчання – це те, що не так часто зустрічається в теперішньому світі.

Існує багато інших способів реалізувати прогнозування курсу криптовалют: лінійна регресія, метод К найближчих сусідів, Байєсівський

метод, звичайні дерева рішень, тощо. Всі вони обчислюють те саме, але різними способами. Деякі дослідження показують, що звичайна лінійна регресія в багатьох випадках передбачає значення ціни краще, ніж складні, багатосарові моделі. [18]

Ще один відомий спосіб прогнозування курсу – використання глибокої нейронної мережі прямого розповсюдження (Deep Feed Forward Neural Network).[17] Це протилежний вид нейронних мереж до рекурентних, які були використані у п. 4.2.2. Інформація в DFFNN переміщується від вхідного шару через приховані шари до вихідного, в мережі немає зворотного зв'язку або циклу, що в основному робить його глибоким і швидким. Навчання DFFNN виконується шляхом оновлення зміни його ваг на кожній ітерації згідно з трьома різними числовими алгоритмами: спряжений градієнт із перезавантаженнями Пауелл-Біла, еластичний алгоритм та алгоритм Левенберга-Марквардта.

В даній роботі були розглянуті два кардинально різні алгоритми – один зі звичайним прогнозуванням, але за допомогою рекурентної нейронної мережі типу LSTM, інший – алгоритм з наповненням, який використовує світові новини, та, базуючись на них, визначає вектор зміни курсу в певній точці часу.

Список використаних джерел

1. Що таке блокчейн? Основні поняття [Електронний ресурс] – Режим доступу до ресурсу: <https://www.investopedia.com/terms/b/blockchain.asp>
2. Як працює технологія блокчейн? [Електронний ресурс] – Режим доступу до ресурсу: <https://builtin.com/blockchain>
3. Про блокчейн простими словами. [Електронний ресурс] – Режим доступу до ресурсу: <https://onezero.medium.com/how-does-the-blockchain-work-98c8cd01d2ae>
4. Біткоїн адреси [Електронний ресурс] – Режим доступу до ресурсу: <https://steadyblockchain.com/exactly-how-many-possible-bitcoin-addresses-could-be-created/>
5. Практичне використання технології блокчейн [Електронний ресурс] – Режим доступу до ресурсу: <https://decenter.org/ru/primenenie-blokcheina>
6. Що таке криптовалюта? [Електронний ресурс] – Режим доступу до ресурсу: <https://www.investopedia.com/terms/c/cryptocurrency.asp>
7. Інформація про ціну біткоїна [Електронний ресурс] – Режим доступу до ресурсу: <https://www.coindesk.com/price/bitcoin>
8. Курс гривні [Електронний ресурс] – Режим доступу до ресурсу: <https://bank.gov.ua/ua/markets/exchangerate-chart>
9. Технічний аналіз [Електронний ресурс] – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D1%85%D0%BD%D1%96%D1%87%D0%BD%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7
10. Поняття волатильності [Електронний ресурс] – Режим доступу до ресурсу: <https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/volatility-vol/>
11. Метод прогнозування курсу криптовалют [Електронний ресурс] – Режим доступу до ресурсу: <https://towardsdatascience.com/predict-any-cryptocurrency-applying-nlp-with-global-news-e938af6f7922>

12. Random Forest Algorithm – деталі [Електронний ресурс] – Режим доступу до ресурсу: <https://www.javatpoint.com/machine-learning-random-forest-algorithm>
13. OpenBlender – документація [Електронний ресурс] – Режим доступу до ресурсу: https://www.openblender.io/#/api_documentation
14. LSTM – покрове пояснення [Електронний ресурс] – Режим доступу до ресурсу: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>
15. Розуміння рекурентних нейронних мереж [Електронний ресурс] – Режим доступу до ресурсу: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
16. Laura Alessandretti, Abeer ElBahrawy, Luca Maria Aiello, Andrea Baronchelli, "Anticipating Cryptocurrency Prices Using Machine Learning", Complexity, vol. 2018, Article ID 8983590, 16 pages, 2018.
<https://doi.org/10.1155/2018/8983590>
17. Lahmiri, S., Bekiros, S. Deep Learning Forecasting in Cryptocurrency High-Frequency Trading. Cogn Comput 13, 485–487 (2021).
<https://doi.org/10.1007/s12559-021-09841-w>
18. Lekkala Sreekanth Reddy, Dr.P. Sriramya “A Research On Bitcoin Price Prediction Using Machine Learning Algorithms”, INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH (2020)
<http://www.ijstr.org/final-print/apr2020/A-Research-On-Bitcoin-Price-Prediction-Using-Machine-Learning-Algorithms.pdf>