

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**
Кафедра обчислюваної математики

Кваліфікаційна робота бакалавра
за спеціальністю 113 Прикладна математика
на тему:
**"ЧИСЕЛЬНИЙ ПРОГНОЗ КРИПТОВАЛЮТИ ЗА ДАНИМИ
ОСТАННІХ РОКІВ"**

Виконав студент 4-го курсу
Тархуні Анатолій Алекс



Науковий керівник:
Асистент кафедри обчислювальної математики,
доктор філософії
Тимошенко Андрій Анатолійович



Засвідчую, що в цій роботі немає
запозичень з праць інших авторів без
відповідних посилань.

Студент



Роботу розглянуто й допущено до
захисту на засіданні кафедри
обчислювальної математики
«29» травня 2023р., протокол № 8
Завідувач кафедри

Ляшко Сергій Іванович



Зміст

Вступ.....	3
Розділ 1 Теорії економічних процесів фінансового ринку.....	4
1.1 Теорія ринкової рівноваги.....	5
1.2 Теорія ймовірності.....	9
1.3 Теорія рефлексивності.....	11
1.4 Теорія ефективності ринку.....	13
1.5 Теорія раціональності очікувань.....	15
1.6 Теорія Ігор.....	17
1.7 Теорія Хаосу.....	19
1.8 Теорія випадкових блукань.....	21
Розділ 2 Методи прогнозування біткоїна.....	22
2.1 Технічний аналіз.....	23
2.2 Фундаментальний аналіз.....	25
2.3 Машинне навчання.....	27
2.4 Аналіз настроїв.....	30
2.5 Кількісний Аналіз.....	32
Розділ 3 Обробка даних та підходи до моделювання.....	35
3.1 Дані.....	35
3.2 Лінійна регресія.....	39
3.3 Поліноміальна регресія.....	40
3.4 В-сплайн регресія.....	44
3.5 N-сплайн регресія.....	46
3.6 Гладкий кубічний сплайн який базується на крос-валідації.....	48
3.7 Підводимо підсумки.....	50
Висновок.....	51
Список використаних джерел.....	52

ВСТУП

Валютний ринок є найбільшим і найліквіднішим фінансовим ринком у світі з обсягом торгів понад 5 трильйонів доларів на день. Прогнозування валют - це процес, який передбачає майбутні рухи різних валют одна відносно однієї. Це життєво важливий інструмент для інвесторів, трейдерів і підприємств, які займаються міжнародними операціями. Точність валютних прогнозів може мати значні наслідки для прибутковості та успіху цих суб'єктів.

Існує кілька методів прогнозування валютних курсів, включаючи технічний аналіз, фундаментальний аналіз і аналіз методами машинного навчання. Технічний аналіз передбачає використання графіків і моделей для виявлення тенденцій у русі валют. Фундаментальний аналіз, з іншого боку, передбачає аналіз економічних і політичних факторів, які можуть вплинути на вартість валюти. Метод машинного навчання передбачає використання математичних моделей для прогнозування майбутніх рухів валют. Ці методи часто використовуються в комбінації для підвищення точності валютних прогнозів.

Незважаючи на різні методи, що використовуються, прогноз валюти є складною галуззю. Однією з головних проблем є непередбачуваність глобального економічного та політичного ландшафту. Економічна політика, політичні події і навіть стихійні лиха можуть мати значний вплив на вартість валют, що ускладнює точне прогнозування майбутніх рухів. Крім того, взаємопов'язаність глобальних економік означає, що на вартість валюти впливають фактори, які можуть бути неочевидними.

Не дивлячись на ці проблеми, прогнозування валютних курсів має важливе значення для бізнесу та інвесторів, які мають справу з міжнародними операціями. Точні валютні прогнози можуть допомогти компаніям мінімізувати ризик втрат через коливання валютних курсів. Наприклад, якщо компанія знає, що певна валюта, ймовірно, зросте в ціні, вона може купити цю валюту заздалегідь, зафіксувавши вигідний обмінний курс. Так само інвестори можуть використовувати валютні прогнози для виявлення вигідних можливостей на міжнародних ринках.

РОЗДІЛ 1

Теорії економічних процесів фінансового ринку

Історія економічних процесів починається з часів появи перших людей. У давні часи люди займалися збиранням рослин і полюванням на тварин для отримання їжі та інших ресурсів. З часом, коли людська цивілізація розвивалася, виникали нові форми економічної діяльності.

У давнину торгівля була важливим економічним процесом. Люди обмінювалися товарами, щоб отримати необхідні ресурси. Наприклад, китайці в обмін на шовк отримували західні товари, такі як скло та метали.

У середньовіччі розвивалася феодальна економіка, коли земля була в основі економічних відносин. Феодали володіли землею та наймали робітників для обробки її. У кінці середньовіччя виникла мануфактура, коли ремісники об'єднувалися, щоб виробляти товари на продаж.

З початком промислової революції в XVIII-XIX століттях економіка почала швидко розвиватися. З'явилася нова форма власності - капітал, який використовувався для виробництва товарів на масову продаж. У цей період виникли нові види підприємств - фабрики, залізниці та інші, що використовували нові технології та машини.

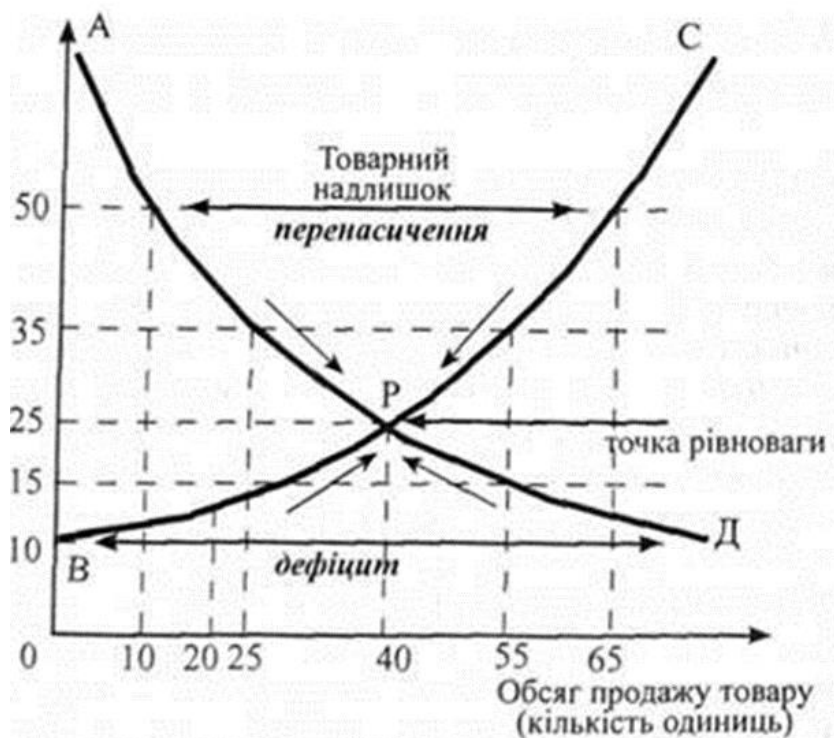
У XX столітті економіка почала ще швидше розвиватися, зокрема завдяки розвитку науково-технічної революції та збільшенню міжнародної торгівлі. Розвиток інформаційних технологій та глобалізація також суттєво вплинули на економічні процеси.

Основні теорії:

1. Теорія ринкової рівноваги
2. Теорія ефективності ринку
3. Теорія ймовірності
4. Теорія рефлексивності
5. Теорія раціональних сподівань учасників ринку
6. Теорія ігор
7. Теорія хаосу
8. Теорія випадкових блукань

1.1 Теорія ринкової рівноваги

Для аналізу ринку варто враховувати рівень попиту та пропозиції. Рис. 1 демонструє приклад пошуку точки рівноваги, де вказано кількість та обсяги продажу, як впливає рівень попиту та пропозиції на ситуацію.



Мал. 1 - Взаємодія попиту і пропозиції

Крива пропозиції біткоїна показує взаємозв'язок між ціною біткоїна і кількістю, яку майнери (ті хто створює криптовалюту за рахунок проведення цифрових операцій) готові запропонувати. Майнери отримують винагороду в біткоїнах за перевірку транзакцій в мережі біткоїн. Ця винагорода має двоелементну природу.

- Перше - фіксована винагорода за блок, яку отримує майнер, який першим розв'язує складну математичну задачу.
- Друге - комісія за транзакції, які майнер перевіряє. Крива пропозиції для біткоїна є гнучкою, і може змінюватися в залежності від того, як змінюється складність майнінгу, кількість майнерів та винагороди.

Крива попиту на біткоїн показує взаємозв'язок між ціною біткоїна і кількістю, яку готові купити інвестори. Інвестори вимагають біткоїн з різних причин, в тому числі як засіб збереження вартості, як

платіжний засіб і як інвестиції. Якщо інвестори вважають, що біткоїн є цінним активом, вони будуть готові платити за нього більше.

Коли крива пропозиції і крива попиту перетинаються, ринок біткоїна знаходиться в стані рівноваги. У цей момент ціна біткоїна не є ані надто високою, ані надто низькою, і у майнерів чи інвесторів немає стимулів змінювати свою поведінку. Однак, коли спостерігається нерівновага між кривими попиту і пропозиції, може виникнути надмірна кількість запитів на купівлю або продаж, що призведе до зміни ціни біткоїна.

Аналізуючи криві попиту і пропозиції, можна робити прогнози щодо майбутньої ціни біткоїна. Наприклад, якщо крива пропозиції зміщується вправо, це означає, що буде більше біткоїнів, доступних для купівлі. Це, швидше за все, призведе до зниження ціни біткоїна. І навпаки, якщо крива попиту зміщується вправо, це означає, що буде більше попиту на біткоїн. Це, швидше за все, призведе до зростання ціни біткоїна.

Важливо відзначити, що теорія ринкової рівноваги базується на ряді припущень, і ці припущення не завжди можуть бути виконані в реальному світі. Наприклад, припущення, що майнери та інвестори є раціональними суб'єктами, не завжди може бути вірним. У деяких випадках шахтарі та інвестори можуть приймати рішення на основі емоцій або інших факторів, які не є раціональними. Тому, при аналізі кривих попиту та пропозиції, варто досліджувати також поведінку реальних людей на ринку, а не лише теоретичні передбачення.

Крім того, теорія ринкової рівноваги припускає, що не існує ринкових провалів. Провали ринку виникають тоді, коли ринок не розподіляє ресурси ефективно. Наприклад, провал ринку може статися, якщо існує негативний зовнішній фактор, такий як забруднення

навколишнього середовища. В цьому випадку ринок не буде виробляти соціально оптимальну кількість біткоїнів.

В цілому, ця теорія є корисним інструментом для розуміння того, як працює ринок біткоїнів. Однак, важливо усвідомлювати припущення, які лежать в основі теорії, і враховувати можливість ринкових збоїв. Економісти та інвестори використовують теорію ринкової рівноваги для розуміння тенденцій ринку, а також для прийняття рішень щодо інвестування та торгівлі біткоїном.

Ось деякі з факторів, які можуть впливати на попит і пропозицію біткоїна і, відповідно, на рівноважну ціну біткоїна:

1. Кількість майнерів. Чим більше майнерів, тим більше біткоїнів буде запропоновано.
2. Складність майнінгу. Чим складніше видобувати біткоїн, тим менше його буде запропоновано.
3. Кількість інвесторів. Чим більше інвесторів, тим більше буде попит на біткоїн.
4. Сприйнята цінність біткоїна. Якщо інвестори вважають, що біткоїн є цінним активом, вони будуть готові платити за нього більше.
5. Державне регулювання. Сприятливе для біткоїна державне регулювання може призвести до зростання ціни на біткоїн. Несприятливе для біткоїна державне регулювання може призвести до зниження ціни біткоїна.
6. Відомість та сприйняття біткоїна в глобальному суспільстві. Чим більше людей знає про біткоїн та вірить в його перспективи, тим більше буде попит на цей актив.
7. Конкуренція від інших криптовалют. Якщо на ринку з'являється новий біткоїн-конкурент, це може призвести до зниження ціни на біткоїн через зменшення попиту.

1.2 Теорія ймовірності[5]

Теорія ймовірностей може бути використана для прогнозування ціни біткоіна. Це розділ математики, який вивчає ймовірність настання подій. Ймовірність - це міра того, наскільки можливо відбувається певна подія. Вона може бути використана для прогнозування майбутнього, призначаючи ймовірності різним можливим результатам.

Існує кілька різних способів використання теорії ймовірностей для прогнозування ціни біткоіна. Один з поширених підходів полягає у використанні історичних цінових даних для виявлення закономірностей. Після того, як ці закономірності виявлені, їх можна використовувати для прогнозування майбутніх цінових рухів.

Інший підхід полягає у використанні статистичних моделей. Статистичні моделі - це математичні рівняння, які можна використовувати для прогнозування поведінки системи. Вони можуть бути використані для прогнозування ціни біткоіна, враховуючи різноманітні фактори, такі як новини, економічні дані та технологічні розробки.

Прогнозування ціни біткоіна з використанням теорії ймовірностей може бути важливим інструментом для інвесторів, які хочуть зробити прибуток на ринку криптовалют. На жаль, немає одного досконалого методу для прогнозування ціни біткоіна.

Найкращим підходом є використання комбінації методів та усвідомлення обмежень кожного з них.

Ось декілька факторів, які можна використовувати для прогнозування ціни біткоіна за допомогою теорії ймовірності:

1. Історичні цінові дані. Історичні цінові дані можуть бути використані для виявлення закономірностей в ціні біткоїна. Ці закономірності можуть бути використані для прогнозування майбутніх цінових рухів.
2. Статистичні моделі. Статистичні моделі можна використовувати для прогнозування ціни біткоїна, враховуючи різноманітні фактори, такі як новини, економічні дані та технологічні розробки.
3. Думки експертів. Думки експертів можна використовувати для того, щоб отримати уявлення про те, як, на їхню думку, поводитиметься ціна біткоїна в майбутньому.

Проте, необхідно зазначити, що є деякі обмеження використання теорії ймовірності для прогнозування ціни біткоїна, які не можна ігнорувати:

1. Ринок непередбачуваний.
2. Ринок біткоїна є дуже волатильним, а це означає, що важко передбачити майбутній рух ціни.
3. Ринок не є ефективним. Ринок біткоїнів не є ефективним ринком, а це означає, що ціни не завжди відображають всю доступну інформацію.

Немає гарантії успіху. Немає жодних гарантій того, що будь-який метод прогнозування ціни біткоїна буде успішним. Тому при прийнятті рішень про інвестування в біткоїн, потрібно уважно вивчати ринок та використовувати комбінацію методів для прогнозування майбутньої ціни.

1.3 Теорія рефлексивності[6]

Теорія рефлексивності в прогнозуванні біткоїна стверджує, що на ціну біткоїна впливають очікування учасників ринку, що створює складність у точному прогнозуванні майбутньої ціни біткоїна. Це означає, що на ціну біткоїна може впливати те, що люди думають про нього. Так, якщо люди вважають, що ціна біткоїна зростатиме, вони з більшою ймовірністю купуватимуть біткоїн, що призведе до зростання ціни, а якщо люди вважають, що ціна біткоїна знизиться, вони з більшою ймовірністю продаватимуть біткоїн, що призведе до падіння ціни.

Однак, теорія рефлексивності може бути використана для того, щоб зробити деякі прогнози щодо майбутньої ціни біткоїна. Наприклад, якщо з'являється багато позитивних новин про біткоїн, це може призвести до зростання ціни біткоїна, а якщо з'являється багато негативних новин про біткоїн, це може призвести до зниження ціни біткоїна.

Теорія рефлексивності говорить про те, що ціна біткоїна є складним і непередбачуваним активом. Важливо пам'ятати про теорію рефлексивності при прийнятті інвестиційних рішень. Це означає, що якщо інвестори не враховують цю теорію, їхні інвестиції можуть призвести до великих втрат.

Основними факторами, які можуть впливати на очікування учасників ринку і, відповідно, на ціну біткоїна є:

1. Новини. Позитивні новини, такі як оголошення про нове партнерство або запуск нового продукту, можуть призвести до зростання ціни біткоїна. Негативні новини, такі як порушення безпеки або репресії з боку уряду, можуть призвести до зниження ціни біткоїна.

2. Економічні дані. Позитивні економічні дані, такі як низький рівень безробіття і високі темпи зростання ВВП, можуть призвести до зростання ціни біткоїна. Негативні економічні дані, такі як висока інфляція та низькі процентні ставки, можуть призвести до зниження ціни біткоїна.
3. Технологічні розробки. Технологічні розробки, які роблять біткоїн безпечнішим або простішим у використанні, можуть призвести до зростання ціни біткоїна. Технологічні розробки, які роблять біткоїн менш захищеним або більш складним у використанні, можуть призвести до зниження ціни біткоїна.
4. Державне регулювання. Сприятливе для біткоїна державне регулювання може призвести до зростання ціни біткоїна. Несприятливе для біткоїна державне регулювання може призвести до зниження ціни біткоїна.

Важливо відстежувати ці фактори, щоб приймати обґрунтовані інвестиційні рішення, і не забувати про теорію рефлексивності. Останнім часом, біткоїн став об'єктом значного інтересу з боку інвесторів і фінансових аналітиків. Цікавою властивістю біткоїна є те, що він не залежить від жодної централізованої влади, що робить його унікальним активом для інвестування.

1.4 Теорія ефективності ринку[7]

Теорія ефективності ринку є однією з найбільш відомих та важливих понять в економічній науці. Вона стверджує, що вся доступна інформація вже відображена в ціні цінного паперу, що робить неможливим постійне обігравання ринку за допомогою технічного аналізу або інших методів. Теорія ефективності ринку базується на

припущенні, що всі учасники ринку є раціональними та мають доступ до однакової інформації. На практиці це припущення не завжди виконується, оскільки завжди є учасники ринку, які мають більше інформації, ніж інші.

Існує три рівні ефективності ринку, які відрізняються за ступенем відображення інформації в ціні:

1. Ефективність слабкої форми. На слабо ефективному ринку вся минула інформація про ціну вже відображена в поточній ціні. Це означає, що неможливо передбачити майбутній рух цін, дивлячись на минулі цінові дані.
2. Напівсильна форма ефективності. На ринку з напівсильною ефективністю вся загальнодоступна інформація вже відображена в поточній ціні. Це означає, що неможливо передбачити майбутній рух цін, вивчаючи новини, економічні дані або іншу загальнодоступну інформацію.
3. Ефективність сильної форми. На ринку з високою ефективністю вся інформація, в тому числі приватна, вже відображена в поточній ціні. Це означає, що неможливо постійно обігравати ринок, навіть якщо у вас є доступ до приватної інформації.

Ринок біткоїна, як правило, вважається напівсильно ефективним. Це означає, що можна зробити деякі прогнози щодо майбутнього руху ціни, дивлячись на новини та економічні дані. Однак, неможливо постійно перемагати ринок, використовуючи технічний аналіз або інші методи.

Існує ряд факторів, які можуть ускладнити прогнозування ціни біткоїна, в тому числі:

1. Висока волатильність. Ціна біткоїна є дуже волатильною, що означає, що вона може змінюватися швидко і непередбачувано. Це ускладнює складання точних прогнозів.
2. Обмежена ліквідність. Ринок біткоїнів є відносно неліквідним, а це означає, що може бути важко купити або продати велику кількість біткоїнів без зміни ціни. Це може ускладнити виконання угод і призвести до збитків.
3. Шахрайство і маніпуляції. Ринок біткоїнів є відносно новим і нерегульованим ринком. Це робить його вразливим до шахрайства і маніпуляцій, що також може ускладнити складання точних прогнозів.

Загалом, теорія ринкової ефективності припускає, що важко постійно перемагати ринок, прогнозуючи ціну біткоїна. Однак, все ж таки можна зробити деякі прогнози щодо майбутніх цінових рухів, дивлячись на новини та економічні дані.

Важливо пам'ятати, що ринок біткоїна є динамічним та змінюється з часом. Тому, навіть якщо ви не можете постійно перемагати ринок, варто продовжувати навчатися та розвивати свої знання про нього. Це допоможе вам зробити кращі інвестиційні рішення в майбутньому.

1.5 Теорія раціональності очікувань[16]

Теорія раціональних очікувань (RE) стверджує, що учасники ринку, такі як інвестори, формують свої очікування щодо майбутніх економічних змінних на основі всієї доступної інформації. Це означає, що їхні очікування не є упередженими і що вони є найкращим здогадом про майбутнє, враховуючи наявну інформацію.

Теорія RE може бути застосована до прогнозування цін на біткоїн. Якщо учасники ринку раціональні, вони будуть використовувати всю доступну інформацію для формування своїх очікувань щодо майбутніх цін на біткоїн. Ця інформація включає минулі ціни на біткоїн, економічні дані, новини та інші фактори, які можуть вплинути на ціну біткоїна.

Якщо учасники ринку раціональні, то їх очікування щодо майбутніх цін на біткоїн повинні бути неупередженими. Це означає, що їх очікування не повинні бути систематично вищими або нижчими за фактичну ціну біткоїна. Однак існують докази того, що учасники ринку можуть не бути раціональними, коли справа доходить до прогнозування цін на біткоїн.

Наприклад, дослідження Чикагського університету показало, що учасники ринку схильні надмірно реагувати на новини, які впливають на ціну біткоїна. Ця надмірна реакція може призвести до утворення бульбашок і обвалів на ринку біткоїна. Надмірна реакція на новини є проявом того, що учасники ринку можуть мати певну необ'єктивність та перебувати під впливом емоцій, таких як страх і жадібність, при прийнятті інвестиційних рішень.

Інше дослідження, проведене Кембриджським університетом, показало, що учасники ринку більш схильні купувати біткоїн, коли ціна зростає, і продавати біткоїн, коли ціна падає. Це свідчить про те, що учасники ринку можуть перебувати під впливом емоцій та не завжди раціонально оцінювати ситуацію, що впливає на ціни біткоїна.

В цілому, дані свідчать про те, що учасники ринку можуть не бути раціональними, коли справа доходить до прогнозування цін на біткоїн. Це означає, що теорія RE не може бути точним способом прогнозування цін на біткоїн, і що ціни на біткоїн будуть

волатильними та важко прогнозувати. Однак, теорія RE все ще може бути корисним інструментом для розуміння факторів, які впливають на ціни біткоїна.

Ось деякі з наслідків теорії ВРЦ для прогнозування цін на біткоїн:

1. Ціни на біткоїн будуть волатильними, оскільки учасники ринку можуть мати необ'єктивність та перебувати під впливом емоцій при прийнятті інвестиційних рішень.
2. Учасники ринку можуть сприймати інформацію необ'єктивно, і це може впливати на їхні прогнози щодо майбутніх цін на біткоїн.
3. Ціни на біткоїн будуть залежати від різних факторів, включаючи новини, економічні дані, інформацію про майбутні події та настрої учасників ринку.

Теорія RE - це складна теорія, і вона все ще обговорюється економістами. Однак вона є корисним інструментом для розуміння факторів, які впливають на ціни біткоїна та може допомогти інвесторам зробити більш обдумані інвестиційні рішення.

1.6 Теорія Ігор[8]

Теорія ігор - це розділ математики, який вивчає прийняття стратегічних рішень. Її можна використовувати, щоб зрозуміти, як різні суб'єкти будуть поводитися в певній ситуації, і передбачити результат цієї ситуації.

Теорію ігор можна використовувати для прогнозування ціни біткоїна, розуміючи, як поводитимуться різні учасники ринку. Наприклад, якщо інвестори вважають, що ціна біткоїна зростатиме,

вони з більшою ймовірністю купуватимуть біткоїн, що призведе до зростання ціни. І навпаки, якщо інвестори вважають, що ціна біткоїна знизиться, вони з більшою ймовірністю продаватимуть біткоїн, що призведе до падіння ціни.

Теорія ігор також може бути використана для розуміння того, як різні уряди будуть регулювати біткоїн. Наприклад, якщо уряд вважає, що біткоїн є загрозою для його фінансової системи, він, швидше за все, регулюватиме біткоїн таким чином, щоб ускладнити його використання. І навпаки, якщо уряд вважає, що біткоїн може бути використаний для сприяння економічному зростанню, він, швидше за все, регулюватиме біткоїн таким чином, щоб зробити його використання простим.

В цілому, теорія ігор може бути корисним інструментом для прогнозування ціни біткоїна. Однак важливо зазначити, що теорія ігор не є досконалою наукою. Існує багато факторів, які можуть вплинути на ціну біткоїна, і неможливо з упевненістю передбачити, як саме той чи інший фактор вплине на ціну.

Ось деякі з факторів, які можна використовувати для прогнозування ціни біткоїна за допомогою теорії ігор:

1. Очікування учасників ринку. Тобто таких як інвестори і трейдери, можуть мати значний вплив на ціну біткоїна. Якщо учасники ринку вважають, що ціна біткоїна зросте, вони з більшою ймовірністю будуть купувати біткоїн, що призведе до зростання ціни. І навпаки, якщо учасники ринку вважають, що ціна біткоїна знизиться, вони з більшою ймовірністю будуть продавати біткоїн, що призведе до зниження ціни.
2. Дії урядів. Політика має значний вплив на ціну біткоїна. Наприклад, якщо уряд забороняє біткоїн, це, швидше за все,

призведе до зниження ціни на біткоїн. І навпаки, якщо уряд приймає біткоїн як законний платіжний засіб, це, швидше за все, призведе до зростання ціни біткоїна.

3. Розвиток нових технологій. Розвиток нових технологій також може мати значний вплив на ціну біткоїна. Наприклад, якщо буде розроблена нова технологія, яка полегшить використання біткоїна, це, швидше за все, призведе до зростання ціни біткоїна. І навпаки, якщо буде розроблена нова технологія, яка ускладнить використання біткоїна, це, швидше за все, призведе до зниження ціни біткоїна.

Важливо відзначити, що жоден окремий фактор не може з упевненістю передбачити ціну біткоїну. Найкращим підходом є використання комбінації факторів і усвідомлення обмежень кожного з них.

1.7 Теорія Хаосу[3][4]

Теорія хаосу - це розділ математики, який вивчає поведінку складних систем. Основою її є ідея, що невеликі зміни в початкових умовах системи можуть призвести до великих змін у поведінці системи з часом.

Теорію хаосу можна використовувати для прогнозування ціни біткоїна, розуміючи, як невеликі зміни на ринку можуть призвести до значних змін ціни біткоїна з часом. Наприклад, якщо невелика кількість інвесторів почне купувати біткоїн, це може призвести до значного зростання ціни біткоїна. І навпаки, якщо невелика кількість інвесторів почне продавати біткоїн, це може призвести до значного зниження ціни біткоїна.

Проте не завжди легко користуватися теорією хаосу для прогнозування ціни біткоїна. Вона може бути корисним інструментом для розуміння динаміки ринку біткоїна і прогнозування майбутньої ціни біткоїна.

Ось декілька факторів, які можна використовувати для прогнозування ціни біткоїна за допомогою теорії хаосу:

1. Кількість користувачів біткоїна. Це ключовий фактор, що може вплинути на ціну біткоїна. Якщо кількість користувачів біткоїна збільшується, це може призвести до зростання ціни біткоїна. І навпаки, якщо кількість користувачів біткоїна зменшується, це може призвести до зниження ціни біткоїна.
2. Кількість видобутого біткоїна. Це ще один ключовий фактор, що може вплинути на ціну біткоїна. Якщо кількість видобутого біткоїна зменшується, це може призвести до зростання ціни біткоїна. І навпаки, якщо кількість видобутого біткоїна збільшується, це може призвести до зниження ціни біткоїна.
3. Волатильність ринку біткоїнів. Це ще один ключовий фактор, що може вплинути на ціну біткоїна. Якщо ринок біткоїна нестабільний, це може призвести до значних коливань ціни біткоїна. І навпаки, якщо ринок біткоїна стабільний, це може призвести до менших коливань ціни біткоїна.

Важливо відзначити, що жоден окремий фактор не може з упевненістю передбачити ціну біткоїну. Найкращим підходом є використання комбінації факторів та усвідомлення обмежень кожного з них.

1.8 Теорія випадкових блукань

Теорія випадкових блукань підтверджується тим фактом, що ціна біткоіна була дуже волатильною в минулому. Наприклад, у 2017 році ціна біткоіна зросла більш ніж на 1000%, але потім знизилася більш ніж на 80% у 2018 році. Така волатильність ускладнює точне прогнозування ціни біткоіна. Проте, це не означає, що неможливо зробити жодних прогнозів щодо ціни біткоіна. Деякі прогнози все ще можна зробити, дивлячись на історичні цінові дані та визначаючи тенденції. Однак важливо усвідомлювати обмеження теорії випадкових блукань і не надто покладатися на якийсь один прогноз.

Ціна біткоіна є результатом взаємодії різних чинників, таких як ринкові тенденції, рівень попиту та пропозиції, наявність конкурентів та рівень прийняття на ринку. Основних факторів, що впливають на ціну біткоіна, не так багато, але кожен з них може мати значний вплив на ціну.

1. Новинні події. Позитивні новини, такі як оголошення про нове партнерство або запуск нового продукту, можуть призвести до зростання ціни біткоіна. Негативні новини, такі як порушення безпеки або репресії з боку уряду, можуть призвести до зниження ціни біткоіна.
2. Економічні дані. Позитивні економічні дані, такі як низький рівень безробіття і високі темпи зростання ВВП, можуть призвести до зростання ціни біткоіна. Негативні економічні дані, такі як висока інфляція і низькі процентні ставки, можуть призвести до зниження ціни біткоіна.
3. Технологічні розробки є ще одним важливим фактором, що впливає на ціну біткоіна. Технологічні розробки, які роблять біткоїн безпечнішим або простішим у використанні, можуть призвести до зростання ціни біткоіна. Технологічні розробки, які

роблять біткоїн менш захищеним або більш складним у використанні, можуть призвести до зниження ціни біткоїна.

4. Державне регулювання є ще одним фактором, на який впливає ціна біткоїна. Сприятливе для біткоїна державне регулювання може призвести до зростання ціни біткоїна. Несприятливе для біткоїна державне регулювання може призвести до зниження ціни біткоїна.

РОЗДІЛ 2

Методи прогнозування біткоїна

Існує різний методи, які можна використовувати для прогнозування ціни біткоїна. Деякі з найпоширеніших методів включають:

1. Технічний аналіз: Цей метод використовує історичні цінові дані для виявлення закономірностей, які можуть передбачити майбутній рух цін.
2. Фундаментальний аналіз: Цей метод вивчає такі фактори, як попит і пропозиція, економічні умови і регуляторні зміни, щоб зробити прогнози щодо майбутньої ціни біткоїна.
3. Машинне навчання: Цей метод використовує алгоритми для навчання на основі історичних даних та прогнозування майбутніх цінових змін.

2.1 Технічний аналіз[11]

Технічний аналіз - це метод прогнозування майбутньої ціни цінного паперу шляхом аналізу його історичних цінових коливань. Технічні

аналітики вважають, що минулі цінові зміни можуть бути використані для прогнозування майбутніх цінових змін.

Існує багато різних інструментів та індикаторів технічного аналізу, які можуть бути використані для прогнозування ціни біткоїну, в тому числі і для прогнозування ціни біткоїну:

Рухомі середні: Рухомі середні згладжують цінові дані і визначають тенденції. Найпоширенішими типами є просте рухоме середнє (SMA) і експоненціальне рухоме середнє (EMA).

1. Смуги Боллінджера: Смуги Боллінджера - це індикатор волатильності, який може виявити перекупленість і перепроданість.
2. Індекс відносної сили (RSI): RSI - це індикатор імпульсу, який може виявити перекупленість і перепроданість.
3. Рівні корекції Фібоначчі: Рівні відновлення Фібоначчі визначають потенційні рівні підтримки і опору.
4. Свічкові діаграми: Свічкові діаграми використовують свічки для відображення руху цін і виявлення моделей, таких як бичачі моделі поглинання і ведмежі моделі поглинання.

Технічні аналітики використовують ці інструменти та індикатори для виявлення закономірностей в ціні біткоїна з метою прогнозування майбутніх цінових рухів. Однак важливо зазначити, що технічний аналіз не є досконалою наукою, і немає ніяких гарантій, що минулі цінові рухи повторяться. Ціна біткоїна є волатильною і може швидко змінюватися, тому прогнози слід сприймати з певною часткою обережності.

Існують деякі обмеження технічного аналізу:

1. Минулі показники не є показником майбутніх результатів: Ціна біткоїна є волатильним активом, який може швидко змінюватися під впливом новин, економічних умов і регуляторних змін.
2. Технічний аналіз не є досконалою наукою: Технічний аналіз використовує минулі рухи цін для прогнозування майбутніх, але немає ніяких гарантій, що ці моделі будуть повторюватися.
3. Технічний аналіз може зайняти багато часу: Збір історичних цінових даних, виявлення закономірностей і прогнозування майбутніх цінових рухів може зайняти багато часу.

Незважаючи на ці обмеження, технічний аналіз може бути корисним інструментом для прогнозування ціни біткоїна. Важливо використовувати його в поєднанні з іншими методами, такими як фундаментальний аналіз і машинне навчання, для найбільш точного прогнозу.

2.2 Фундаментальний аналіз[13]

Фундаментальний аналіз - це метод прогнозування майбутньої ціни цінного паперу шляхом аналізу його фундаментальних основ. Фундаментальні аналітики вважають, що ціна цінного паперу визначається його внутрішньою вартістю, яка базується на таких факторах, як фінансові показники компанії, її конкурентна позиція і перспективи зростання.

Існує багато різних інструментів фундаментального аналізу та індикаторів, які можна використовувати для прогнозування ціни біткоїна. Деякі з найпоширеніших інструментів включають:

1. Попит і пропозиція: Фундаментальні аналітики вважають, що ціна біткоїна визначається попитом і пропозицією. Пропозиція

біткоїнів обмежена, оскільки коли-небудь буде видобуто лише 21 мільйон біткоїнів. Попит на біткоїн визначається різними факторами, в тому числі зростанням кількості підприємств, які приймають біткоїн в якості оплати, зростанням кількості людей, які вважають, що біткоїн є засобом збереження вартості, а також зростанням кількості людей, які вважають, що біткоїн є проривною технологією.

2. Економічні умови: Фундаментальні аналітики також вважають, що на ціну біткоїна впливають економічні умови. Наприклад, якщо економіка зростає, попит на біткоїн, ймовірно, буде більшим. І навпаки, якщо економіка перебуває в стані рецесії, попит на біткоїн, швидше за все, буде меншим.
3. Регуляторні зміни: Фундаментальні аналітики також вважають, що на ціну біткоїна впливають регуляторні зміни. Наприклад, якщо уряд приймає біткоїн як законний платіжний засіб, це, ймовірно, призведе до збільшення попиту на біткоїн. І навпаки, якщо уряд забороняє біткоїн, це, швидше за все, призведе до зниження попиту на біткоїн.

Фундаментальні аналітики використовують ці інструменти та індикатори для визначення факторів, які можуть вплинути на ціну біткоїна. Вони вважають, що ці фактори можуть бути використані для прогнозування майбутніх цінових змін.

Однак важливо зазначити, що фундаментальний аналіз не є досконалою наукою. Немає жодних гарантій, що фактори, які аналізуються, матимуть бажаний вплив на ціну біткоїна. Крім того, ціна біткоїна є волатильною і може швидко змінюватися, тому будь-який прогноз слід сприймати з певною часткою обережності.

Ось деякі з обмежень фундаментального аналізу:

1. Майбутнє є невизначеним: Фундаментальні аналітики роблять прогнози щодо майбутньої ціни біткоїна на основі аналізу поточних та історичних даних. Однак майбутнє є невизначеним, і немає ніяких гарантій, що фактори, які аналізуються, матимуть бажаний вплив на ціну біткоїна.
2. Біткоїн - це новий клас активів: Біткойн - це новий клас активів, і історичні дані для аналізу обмежені. Це ускладнює для фундаментальних аналітиків складання точних прогнозів щодо майбутньої ціни біткоїна.
3. Фундаментальний аналіз може зайняти багато часу: Фундаментальний аналіз може бути трудомістким процесом. Аналітикам потрібно збирати дані, аналізувати їх і робити прогнози щодо майбутнього руху цін.

Незважаючи на ці обмеження, фундаментальний аналіз може бути корисним інструментом для прогнозування ціни біткоїна. Однак важливо використовувати його в поєднанні з іншими методами, такими як технічний аналіз і машинне навчання, щоб отримати найбільш точний прогноз.

Ось деякі з найкращих джерел інформації для фундаментального аналізу біткоїна:

1. Новини: Новини можуть надати інформацію про фактори, які можуть вплинути на ціну біткоїна, такі як нові партнерства, нові продукти і регуляторні зміни.
2. Урядові звіти: Урядові звіти можуть надавати інформацію про економічні умови, такі як інфляція, безробіття і зростання ВВП.
3. Фінансова звітність компаній: Фінансова звітність компанії може надавати інформацію про фінансові показники компанії, такі як дохід, прибуток і грошові потоки.

Важливо відзначити, що це лише деякі з багатьох джерел інформації, які можна використовувати для фундаментального аналізу біткоїна. Найкращі джерела інформації будуть відрізнятися в залежності від цілей і толерантності до ризику.

2.3 Машинне навчання

Моделі машинного навчання можуть передбачити ціну біткоїна, навчаючись на історичних даних про ціни. Існує кілька моделей, які можуть бути використані, зокрема:

1. Лінійна регресія: Ця проста модель прогнозує ціну біткоїна, знаходячи лінійну залежність між ціною та іншими факторами, такими як кількість користувачів біткоїна або кількість видобутого біткоїна.
2. Машини опорних векторів: Ця складніша модель прогнозує ціну біткоїна, знаходячи гіперплощину, яка розділяє дані на два класи, наприклад, "вгору" і "вниз".
3. Випадкові ліси: Цей тип ансамблю моделей поєднує в собі прогнози декількох дерев рішень. Древа рішень прогнозують ціну біткоїна, розбиваючи дані на все менші і менші групи, поки кожна група не буде містити лише одну точку даних.

Моделі машинного навчання можуть навчатися на історичних цінових даних, щоб визначити закономірності, які визначають ціни на біткоїн. Після навчання моделі можуть передбачати майбутні цінові рухи.

Однак моделі машинного навчання не є досконалыми. Не гарантується, що модель машинного навчання буде точно прогнозувати ціну біткоїна. Крім того, ціна біткоїна є волатильною і

може швидко змінюватися, тому будь-який прогноз слід сприймати з певною долею обережності.

Існує кілька обмежень у використанні моделей машинного навчання для прогнозування біткоїна:

1. Моделі машинного навчання не є досконалими: Моделі машинного навчання навчаються на історичних даних, і немає жодних гарантій, що закономірності, які визначали ціни на біткоїн у минулому, будуть продовжувати визначати ціни на біткоїн у майбутньому.
2. Ціна біткоїна нестабільна: Ціна біткоїна є нестабільним активом, який може швидко змінюватися під впливом різних факторів, в тому числі новин, економічних умов та регуляторних змін. Це ускладнює для моделей машинного навчання точне прогнозування ціни біткоїна.
3. Моделі машинного навчання можуть бути упередженими: Моделі машинного навчання навчаються на даних, зібраних з реального світу. Ці дані можуть бути упередженими, і ця упередженість може відобразитися на прогнозах моделі машинного навчання.

Незважаючи на ці обмеження, моделі машинного навчання можуть бути корисними для прогнозування ціни біткоїна. Однак для підвищення точності їх слід використовувати разом з іншими методами, такими як технічний і фундаментальний аналіз.

Ось кілька способів підвищити точність моделей машинного навчання для прогнозування біткоїна:

1. Використовуйте більше даних: Чим більше даних, на яких навчається модель машинного навчання, тим точнішими будуть прогнози.

2. Використовуйте різні джерела даних: Моделі машинного навчання можна навчати на даних з різних джерел, таких як новини, урядові звіти та фінансова звітність компаній. Це може допомогти зменшити упередженість прогнозів моделі.
3. Використовуйте різні моделі машинного навчання: Жодна модель машинного навчання не є ідеальною. Використання різноманітних моделей може допомогти підвищити точність прогнозів.
4. Використовуйте валідаційний набір: Валідаційний набір - це набір даних, які не використовуються для навчання моделі. Потім модель тестується на валідаційному наборі, щоб побачити, наскільки точними є прогнози. Це може допомогти уникнути надмірного пристосування моделі до навчальних даних.

Важливо відзначити, що це лише кілька способів підвищення точності моделей машинного навчання для прогнозування біткоіна. Найкращі методи залежать від конкретної моделі та наявних даних.

2.4 Аналіз Настроїв[14]

Аналіз настроїв щодо криптовалют передбачає використання методів обробки природної мови (NLP) для аналізу соціальних мереж і новинних стрічок, пов'язаних з криптовалютами, з метою визначення настрою повідомлень, якими діляться користувачі. Настрої можуть бути позитивними, негативними або нейтральними і можуть дати уявлення про те, як люди ставляться до різних криптовалют і пов'язаних з ними тем. Аналіз настроїв щодо криптовалют - це зростаюча сфера, яку вивчають як дослідники, так і інвестори.

Існує кілька типів моделей аналізу настроїв:

1. Моделі на основі правил: Ці моделі використовують заздалегідь визначені правила для ідентифікації певних слів, пов'язаних з позитивними або негативними настроями. Наприклад, слово "щасливий" може бути визначене як позитивне, а слово "сумний" як негативне, і ці правила можуть визначити загальний настрій фрагмента тексту.
2. Моделі машинного навчання: Ці моделі використовують алгоритми для навчання на основі маркованих даних, таких як відгуки клієнтів, які були вручну позначені як позитивні або негативні. Потім моделі можуть робити прогнози щодо настрою нового тексту на основі того, що вони дізналися.
3. Моделі на основі лексики: Ці моделі використовують заздалегідь створений словник слів і пов'язані з ними оцінки настрою для визначення загального настрою фрагмента тексту. Наприклад, слово "любов" може мати високу позитивну оцінку, тоді як слово "ненависть" може мати високу негативну оцінку.
4. Гібридні моделі: Ці моделі поєднують кілька підходів, таких як навчання на основі правил і машинне навчання, для підвищення точності. Наприклад, гібридна модель може використовувати правила для визначення певних слів або фраз, які тісно пов'язані з позитивними чи негативними настроями, а потім використовувати алгоритми машинного навчання для уточнення своїх прогнозів на основі додаткових даних.
5. Моделі глибокого навчання: Ці моделі використовують нейронні мережі для аналізу великих обсягів даних і виявлення шаблонів у тексті, які асоціюються з позитивними чи негативними настроями. Незважаючи на свою потужність, ці моделі вимагають великих обсягів навчальних даних і обчислювальних ресурсів.
6. Ансамблеві моделі: Ці моделі об'єднують прогнози кількох моделей, щоб забезпечити точніші прогнози, ніж будь-яка окрема модель

могла б досягти самостійно. Наприклад, ансамблева модель може поєднувати прогнози моделі на основі правил, моделі машинного навчання та моделі на основі лексикону, щоб отримати остаточний прогноз щодо настрою фрагмента тексту. Поєднуючи сильні сторони декількох моделей, ансамблеві моделі часто можуть досягти вищої точності, ніж будь-яка окрема модель самостійно.

Приклади кожної з них включають:

1. Моделі на основі правил: Бібліотека TextBlob на мові Python, яка ідентифікує настрої в тексті за допомогою набору попередньо визначених правил.
2. Моделі машинного навчання: Модель VADER (Valence Aware Dictionary and Sentiment Reasoner), яка прогнозує настрої в тексті за допомогою лексикону та алгоритмів машинного навчання.
3. Лексикологічні моделі: Лексикон SentiWordNet, який присвоює словам позитивні та негативні оцінки на основі їхньої семантичної спрямованості.
4. Гібридні моделі: Стенфордський інструмент аналізу настроїв CoreNLP, який поєднує підходи на основі правил і машинного навчання для досягнення високої точності.
5. Моделі глибокого навчання: Нейронна мережа LSTM (Long Short-Term Memory), яку можна навчати на великих обсягах даних для прогнозування настроїв у тексті.
6. Ансамблеві моделі: Класифікатор випадкових лісів, який поєднує кілька дерев рішень для прогнозування настроїв у тексті.

2.5 Кількісний Аналіз[15]

Кількісний аналіз дозволяє спрогнозувати майбутні цінові рухи біткоїна, аналізуючи історичні дані та визначаючи закономірності і

тенденції. Це передбачає використання статистичних моделей і математичних алгоритмів для аналізу великих обсягів даних, таких як обсяг торгів, ринкова капіталізація, складність майнінгу та інші фактори, які можуть вплинути на ціну біткоїна.

Використовуючи такий аналіз, трейдери та інвестори можуть визначити потенційні можливості для купівлі або продажу на основі історичних тенденцій та закономірностей. Наприклад, вони можуть використовувати технічний аналіз для визначення рівнів підтримки і опору на графіку ціни біткоїна або використовувати регресійний аналіз для виявлення кореляції між рухом ціни біткоїна та іншими факторами, такими як обсяг торгів або складність видобутку.

Кількісний аналіз також може розробляти прогнозні моделі, які передбачають майбутні зміни цін на основі історичних даних. Ці моделі можуть використовувати алгоритми машинного навчання або інші статистичні методи для виявлення закономірностей в даних, які передбачають майбутній рух цін.

В цілому, кількісний аналіз може надати цінну інформацію про рух цін на біткоїн і допомогти трейдерам та інвесторам приймати більш обґрунтовані інвестиційні рішення. Однак важливо зазначити, що жодна модель не може передбачити майбутнє з упевненістю, тому завжди важливо проявляти обережність при прийнятті інвестиційних рішень.

Кількісні моделі для прогнозування:

1. Stock to Flow (S2F) model: Ця модель оцінює майбутню вартість біткоїна на основі співвідношення його поточного запасу (загальної пропозиції) до річного потоку (нових видобутих монет). Чим вище співвідношення S2F, тим дефіцитнішим є актив і тим вищою очікується його вартість.

2. Scarcity model: Ця модель передбачає, що вартість біткоїна прямо пропорційна його дефіциту, який визначається обмеженою пропозицією і зростаючим попитом.
3. Metcalfe's Law: Ця модель припускає, що вартість мережі (такої як Біткоїн) пропорційна квадрату кількості користувачів у цій мережі. Очікується, що зі збільшенням кількості людей, які приймають біткоїн, його вартість буде зростати.
4. Hypecycle Theory: Ця модель прогнозує, що біткоїн переживе серію бульбашок і падінь, перш ніж врешті-решт досягне стабільного рівня цін, заснованого на його фундаментальній вартості.
5. Bayesian Regression Analysis: Ця статистична модель використовує історичні дані про ціну біткоїна та інші фактори, такі як складність майнінгу та обсяг транзакцій, для прогнозування майбутнього руху цін на основі байєсівської теорії ймовірності.
6. NVT Ratio: Ця модель використовує відношення вартості мережі до транзакцій (NVT) для оцінки справедливої вартості біткоїна. Коефіцієнт NVT розраховується шляхом ділення ринкової капіталізації біткоїна на щоденний обсяг транзакцій. Високий показник NVT свідчить про те, що біткоїн переоцінений, в той час як низький показник NVT свідчить про те, що він недооцінений.
7. Puell Multiple: Ця модель використовує відношення щоденної вартості емісії біткоїна до його 365-денної ковзної середньої як міру його справедливої вартості. Коли це співвідношення високе, це свідчить про те, що біткоїн переоцінений, тоді як низьке співвідношення свідчить про його недооціненість.
8. Mayer Multiple: Ця модель використовує відношення поточної ціни біткоїна до його 200-денної ковзної середньої як міру його

справедливої вартості. Коли це співвідношення високе, це означає, що біткоїн переоцінений, тоді як низьке співвідношення свідчить про його недооціненість.

РОЗДІЛ 3

Обробка даних та підходи до моделювання

3.1 Дані

Однією з переваг використання датасетів з Kaggle є те, що вони часто проходять перевірку спільнотою та відбираються на основі якості, достовірності та релевантності. Багато датасетів мають детальний опис, зазначаються джерела даних та відповідність загальним стандартам.

Популярність датасету на Kaggle може бути обумовлена кількістю завантажень, коментарів та високими оцінками користувачів. Це може свідчити про його корисність для аналітики, машинного навчання або досліджень відповідної галузі.

Вибір добре оціненого датасету може забезпечити надійну та якісну основу для проведення аналізу даних, розробки моделей прогнозування або виконання досліджень. Однак, перед використанням датасету варто оцінити його відповідність конкретним потребам і завданням дослідження, а також провести необхідні перевірки та попередню обробку.

Ось частина нього

	Date	Open	High	Low	Close	Adj Close	Volume
1996	2020-03-06	9078.308594	9167.695313	9032.079102	9122.545898	9122.545898	40826885651
2515	2021-08-07	42832.796875	44689.859375	42618.566406	44555.800781	44555.800781	40030862141
1495	2018-10-22	6486.049805	6543.799805	6462.979980	6487.160156	6487.160156	3672860000
423	2015-11-15	333.050995	334.661987	317.489990	320.165985	320.165985	44213100
36	2014-10-24	358.591003	364.345001	353.304993	358.345001	358.345001	15585700

Я буду працювати з даними 'Date' та 'Close'. Тож спочатку зробимо обробку датасету.

Я поділив мій датасет на два: 0.8 (від основного) тренувальний і 0.2 (від основного) тестовий. Також ми додаємо до них колонки 'Date' та 'Close' також я їх переіменував на 'date' та 'price'.

	date	price
90	2014-12-17	319.776001
70	2014-11-27	369.670013
837	2017-01-02	1021.750000
741	2016-09-28	604.728027
956	2017-05-01	1421.599976

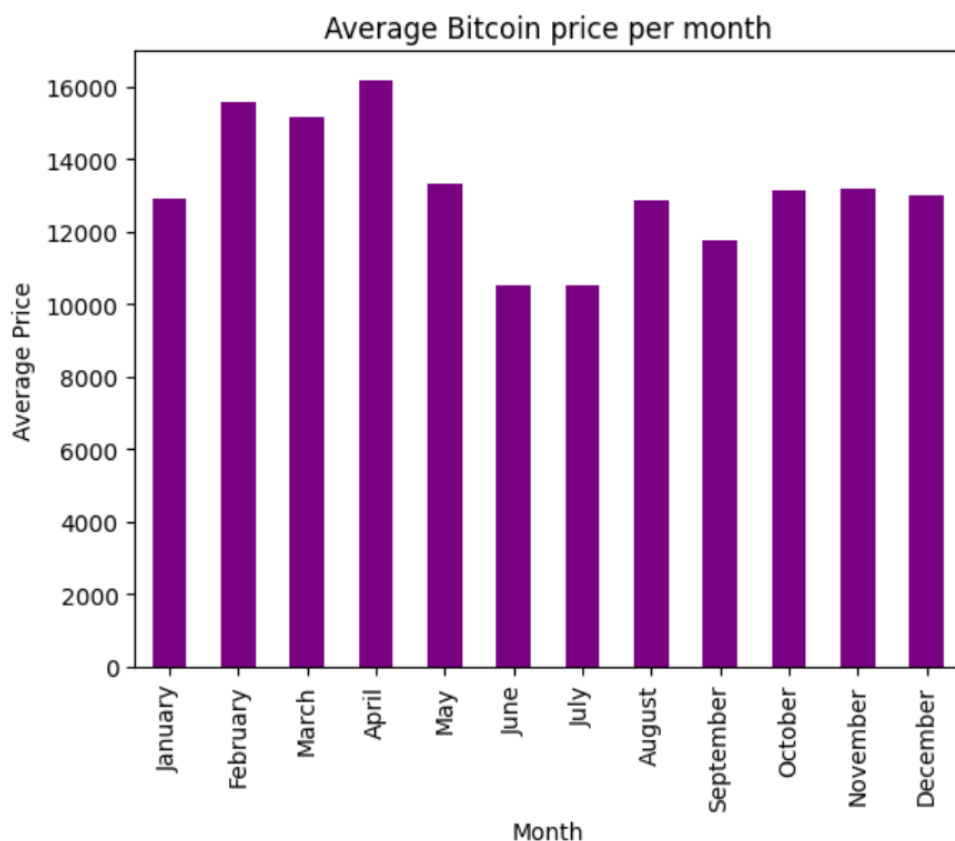
Train датасет має 2502 рядків та 2 колонок
 Test датасет має 626 рядків та 2 колонок

	price		price
count	2502.000000	count	626.000000
mean	13324.253265	mean	12748.617402
std	16089.964744	std	15866.759849
min	178.102997	min	208.097000
25%	751.406494	25%	642.554260
50%	7403.709717	50%	7274.663818
75%	19140.484375	75%	18829.000488
max	67566.828125	max	65992.835938

Далі я почав обробляти дату для того щоб з нею було краще працювати.

	date	price	month	month_name	weekday	weekday_name
1158	2017-11-19	8036.490234	11	November	6	Sunday
1498	2018-10-25	6476.290039	10	October	3	Thursday
1911	2019-12-12	7243.134277	12	December	3	Thursday
2495	2021-07-18	31796.810547	7	July	6	Sunday
2403	2021-04-17	60683.820313	4	April	5	Saturday

Далі для цікавості я вивів середню ціну біткоіна по місяцям.

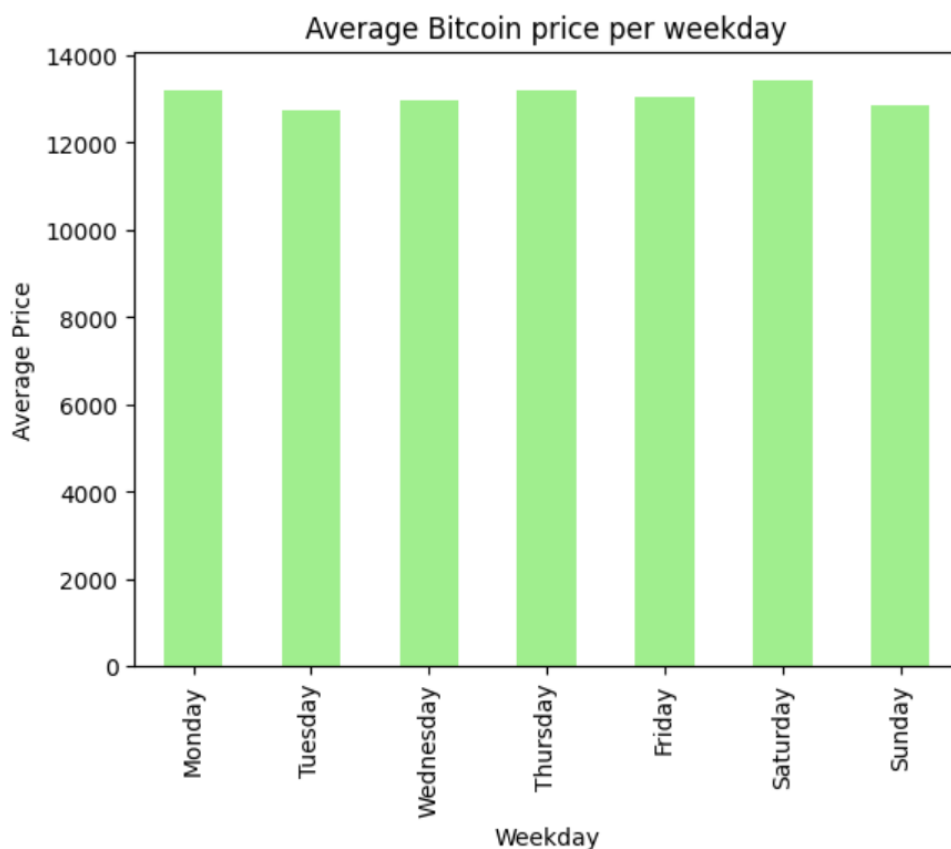


Можемо побачити як середня ціна біткоіна змінюється можемо побачити приблизні паттерни. Тобто з початку року ціна росте чим і набуває максимуму в березні і потім вона падає до серпня фактори які впливають на ціну біткоіна влітку будуть наведені нижче. Настрій ринку: Влітку настрої ринку та поведінка інвесторів можуть бути під впливом різноманітних факторів.

1. Адаптація складності майнінгу: Складність майнінгу Bitcoin, яка впливає на швидкість створення нових Bitcoin, коригується приблизно кожні два тижні.
2. Маніпуляція ринком: Ринки криптовалют все ще відносно молоді і піддаються маніпуляції.
3. Поведінка інвесторів: Поведінка учасників ринку, така як фіксація прибутку, перебалансування портфеля або зміна рівня ризику, може впливати на ціну Bitcoin.

Можемо спостерігати в кінці майже кожного сезону різке збільшення ціни.

Далі, побачивши такий графік, мені стало цікаво чи є якась залежність між днями тижнями.



Для мене на цьому графіку не бачу ніякої, навіть примарної залежності ціни від дня тижня.

Далі я замінив дату в нормальному вигляді на кількість днів. Тобто перший день це перша дата в нашому датасеті і далі по дням до останнього дня.

Для оцінки результату я буду використовувати коефіцієнт $R^2 = 1 - \frac{RSS}{TSS}$ Де $RSS = \sum_{i=1}^n (y_i - v_i)^2$ - сума квадратів залишків

$$TSS = \sum_{i=1}^n (y_i - \bar{v})^2$$
 - загальна сума квадратів

Чим ближче R^2 до 1 тим точніше прогнозовані значення.

3.2 Лінійна Регресія

Лінійна регресія - це статистичний метод моделювання, який використовується для встановлення зв'язку між залежною змінною та однією або кількома незалежними змінними. Вона передбачає лінійний зв'язок між змінними, що означає, що залежна змінна може бути передбачена як лінійна комбінація незалежних змінних.

Метою лінійної регресії є знаходження найкращої апроксимаційної прямої, яка мінімізує різницю між передбачуваними значеннями та фактичними значеннями залежної змінної. Ця пряма називається регресійною прямою або прямою найкращого прилягання. Вона визначається рівнянням:

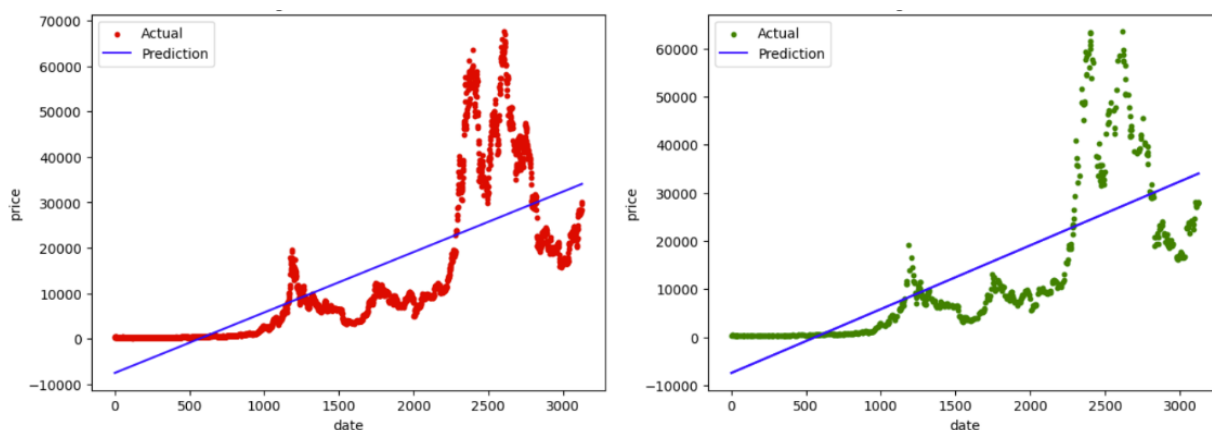
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

де:

- y - залежна змінна, яку ми хочемо передбачити,
- β_0 - точка перетину з осю y , яка представляє значення y , коли всі незалежні змінні дорівнюють нулю,

- $\beta_1, \beta_2, \dots, \beta_n$ - коефіцієнти або нахил, які представляють зміну у при одиничній зміні кожної незалежної змінної,
- x_1, x_2, \dots, x_n - незалежні змінні,
- ε - помилка, яка враховує варіативність, яка не пояснюється моделлю.

Коефіцієнти ($\beta_1, \beta_2, \dots, \beta_n$) оцінюються за допомогою методу найменших квадратів (OLS), який мінімізує суму квадратів різниць між передбаченими та фактичними значеннями. Після отримання оцінених коефіцієнтів їх можна використовувати для прогнозування нових даних.



$$R^2 = 0.547600802995594$$

Мал. 2 - Лінійна регресія на трен(зліва) та на тест(справа) та коефіцієнт детермінації

3.3 Поліноміальна регресія

Формула: У поліноміальній регресії використовується модель, що має наступний вигляд:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1^3 + \dots + \beta_n x_1^n + \varepsilon$$

де:

- y - залежна змінна, яку ми хочемо передбачити,

- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ - коефіцієнти, які представляють вплив незалежних змінних на залежну змінну,
- x_1 - незалежна змінна,
- n - ступінь поліному,
- ε - помилка, яка враховує варіативність, яка не пояснюється моделлю.

Дані: Для виконання поліноміальної регресії потрібні дані, які включають значення залежної змінної (y) і відповідні значення незалежної змінної (x_1).

Важливо пам'ятати, що поліноміальна регресія може бути чутлива до вибору ступеня поліному, і недостатній або надмірний ступінь може призвести до некоректних результатів. Також, як і в інших методах регресії, слід перевірити виконання необхідних припущень, таких як нормальність помилок і лінійність у параметрах.

Ось кілька прикладів, де поліноміальна регресія може бути використана:

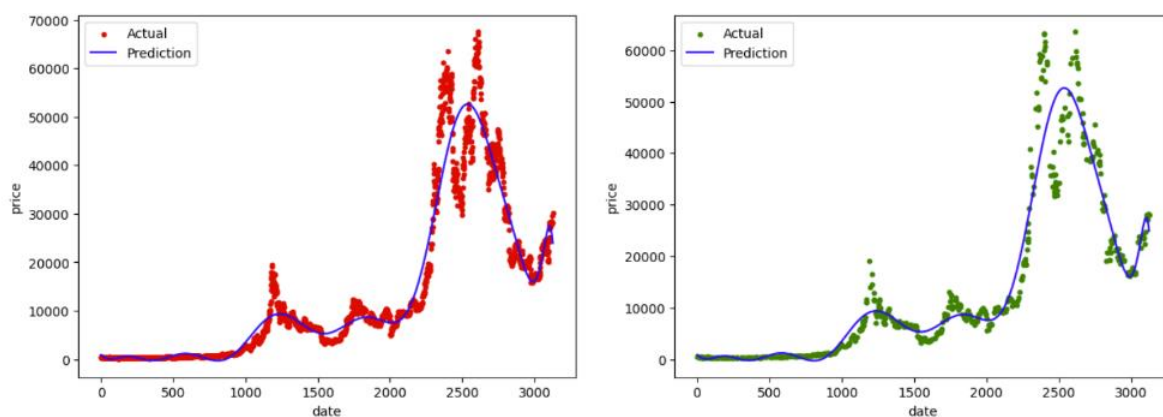
1. Економічний аналіз: В поліноміальній регресії можна моделювати взаємозв'язок між економічними факторами і показниками, такими як ВВП, споживча витрата, інвестиції тощо. Наприклад, можна використовувати поліноміальну регресію для аналізу впливу інфляції на економічне зростання, де ступені поліному можуть відображати неоднорідність залежності у різних періодах.
2. Медична дослідження: В поліноміальній регресії можна досліджувати залежність між вхідними факторами і медичними показниками. Наприклад, можна використовувати поліноміальну регресію для дослідження впливу віку на серцеві

захворювання, де ступені поліному можуть відображати нелінійність цієї залежності.

- Маркетингові дослідження: Поліноміальна регресія може бути використана для аналізу взаємозв'язку між маркетинговими зусиллями і продажами. Наприклад, можна використовувати поліноміальну регресію для вивчення впливу рекламних витрат на продажі, де ступені поліному можуть відображати нелінійну залежність між цими змінними.

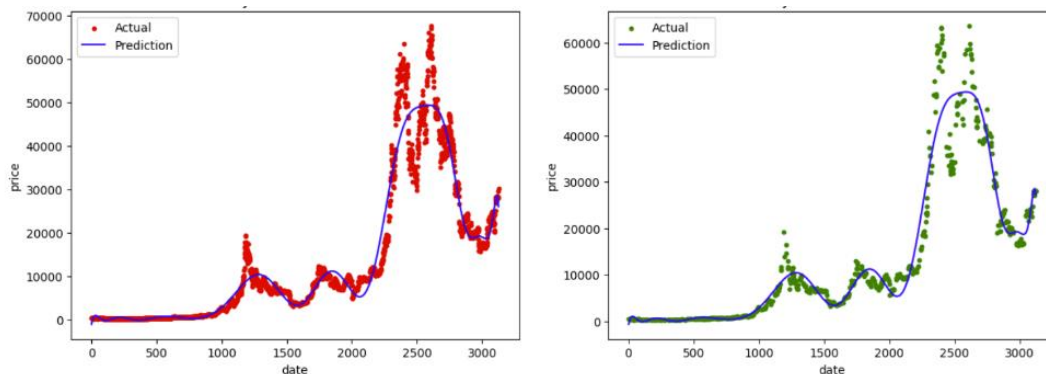
Це лише кілька прикладів використання поліноміальної регресії. Застосування можуть бути широкими і залежати від конкретного контексту та даних, які вивчаються.

Тож я взяв 22, 44, 66, 88 порядків регресій



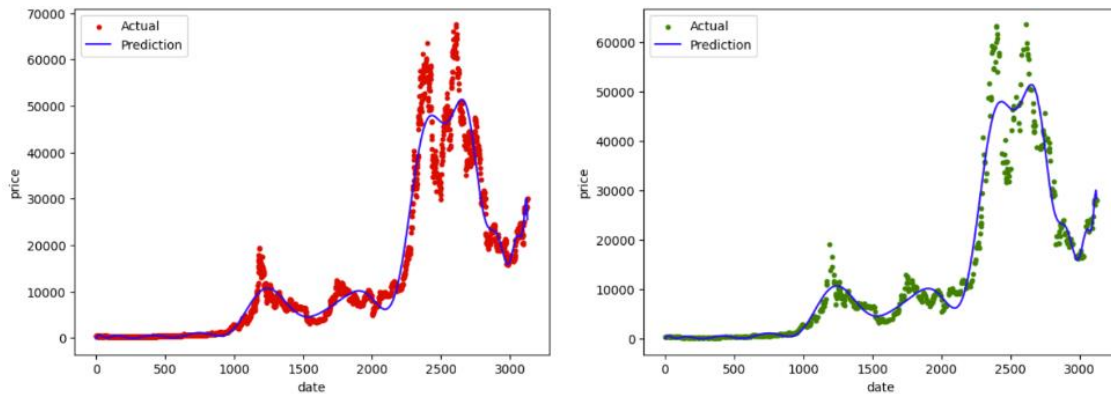
$$R^2 = 0.921847973471223$$

Мал. 3 - Поліноміальна регресія 22-порядку на трен та на тест та коефіцієнт детермінації



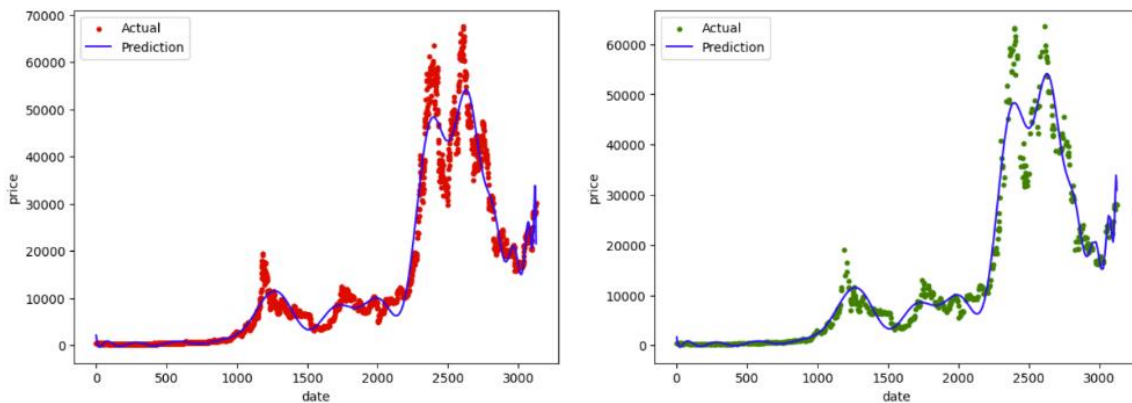
$$R^2 = 0.9312422647780334$$

Мал. 4 - Поліноміальна регресія 44-порядку на трен, та на тест та коефіцієнт детермінації



$$R^2 = 0.9380333081827431$$

Мал. 5 - Поліноміальна регресія 66-порядку на трен, та на тест та коефіцієнт детермінації



$$R^2 = 0.9469603716408488$$

Мал. 6 - Поліноміальна регресія 88-порядку на трен та на тест та коефіцієнт детермінації

3.4 В-сплайн функція

Формула: В-сплайн функція може бути представлена наступною формулою:

$$y = \beta_0 + \beta_1 B_1(x) + \beta_2 B_2(x) + \dots + \beta_n B_n(x)$$

де:

- y - залежна змінна, яку ми намагаємося передбачити,
- x - незалежна змінна,
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ - коефіцієнти регресії,
- $B_1(x), B_2(x), \dots, B_n(x)$ - базисні функції б-сплайнів.

Дані: Для застосування В-сплайн функції потрібні наступні дані:

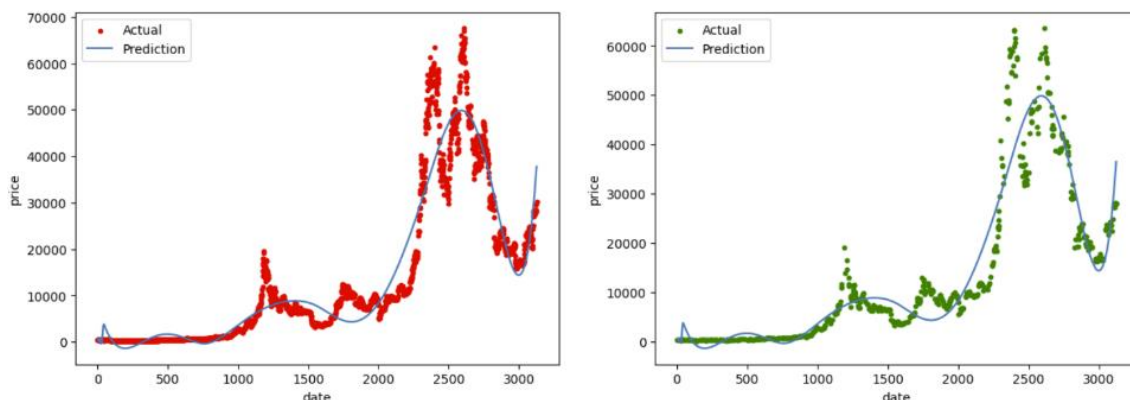
- Залежна змінна (y): це значення, яке ми намагаємося передбачити або пояснити.
- Незалежна змінна (x): це значення, яке використовується для передбачення залежної змінної.
- Вузли (nodes): це точки, на яких будуть розташовуватися базисні функції В-сплайнів. Вони можуть бути рівномірно розподілені або розміщені згідно з особливими потребами аналізу.

Опис методу: В-сплайн функція є методом апроксимації даних за допомогою функції, складеної з сегментів, відомих як В-сплайни. Кожен В-сплайн є кусково-поліноміальною функцією, яка з'єднується з іншими функціями на вузлах. Коефіцієнти регресії ($\beta_0, \beta_1, \beta_2, \dots, \beta_n$) контролюють поведінку цих функцій.

Оцінка коефіцієнтів В-сплайн регресії здійснюється шляхом мінімізації суми квадратів різниць між передбачуваними значеннями залежної змінної та фактичними значеннями. При цьому застосовуються обмеження на похідні функцій В-сплайнів у вузлах для забезпечення гладкості апроксимуючої функції.

В-сплайн функція використовується для апроксимації та аналізу даних з нелінійними та неоднорідними залежностями. Вона є потужним методом, який може бути використаний в різних галузях для моделювання та передбачення різноманітних явищ.

В моєму випадку перший раз я поділив на квантилі 25, 50, 75 і отримав.



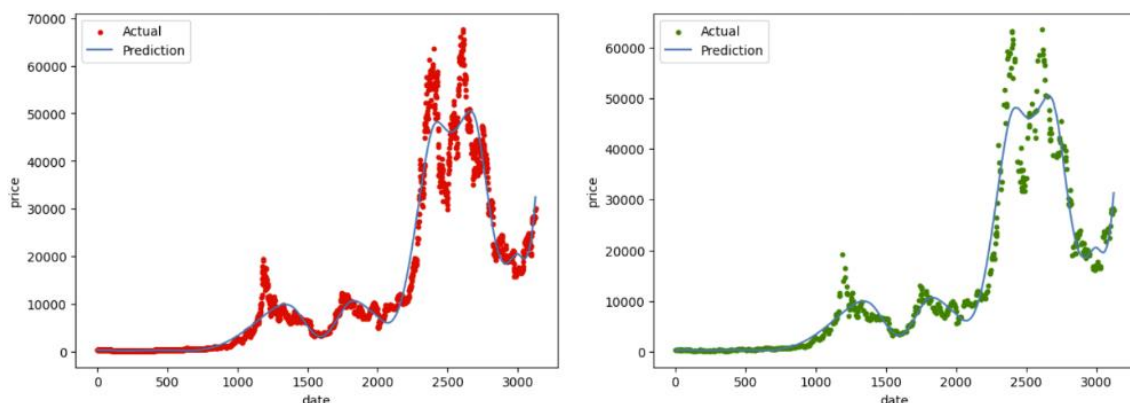
$$R^2 = 0.8971612833195418$$

Мал. 7 - В-сплайн регресія квантилі 25, 50, 75 на трен та на тест та коефіцієнт детермінації

Можемо побачити що ситуація непогана, але не дуже й гарна. Потім я ввів свої вузли.

```
knots=[150,1000,1400,1600,1800,2000, 2400, 2500, 2700, 3000 ]
```

І отримав вже таку точність:



$$R^2 = 0.9345904642361558$$

Мал. 8 - B-сплайн регресія зі своїми вузлами на трен та на тест та коефіцієнт детермінації

І отримав вже точність 0.93 що має непоганий скачок з попереднього результату. Але як 1 з найкращих функцій в аналізі даних її результат не дотягує в основному через мене я не зміг підібрати правильну кількість та самі вузли.

3.5 N-сплайн регресія

Формула: Натуральний сплайн регресії може бути представлений наступною формулою:

$$y = \beta_0 + \beta_1x + \beta_2x^2 + \dots + \beta_nx^n$$

де:

- y - залежна змінна, яку ми намагаємося передбачити або пояснити,
- x - незалежна змінна,
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ - коефіцієнти регресії.

Дані: Для застосування натурального сплайн регресії потрібні наступні дані:

- Залежна змінна (y): це значення, яке ми намагаємося передбачити або пояснити.
- Незалежна змінна (x): це значення, яке використовується для передбачення залежної змінної.

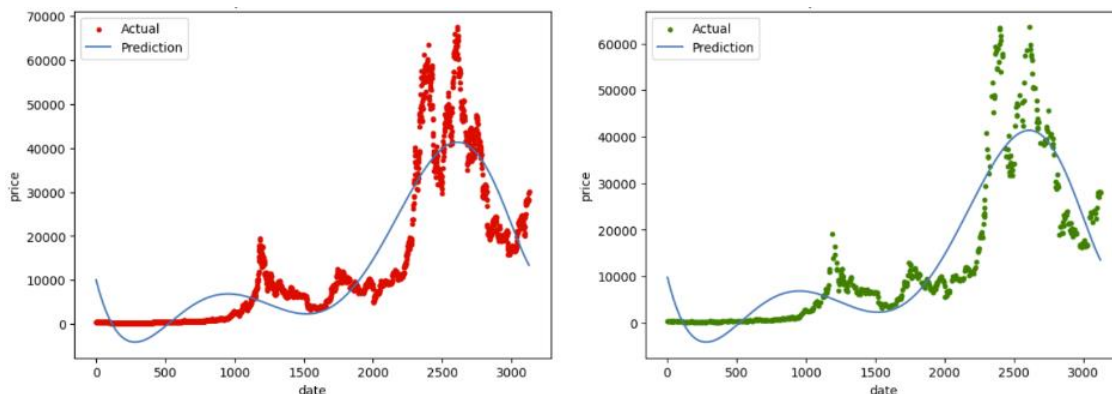
Опис методу: Натуральний сплайн регресії є методом апроксимації даних, який використовує поліноміальні функції для

моделювання залежності між змінними. Основна ідея полягає в тому, що функція апроксимації складається з сегментів, кожен з яких є поліноміальною функцією степеня n .

Метод натурального сплайну регресії вимагає розв'язання системи лінійних рівнянь, щоб знайти коефіцієнти регресії $(\beta_0, \beta_1, \beta_2, \dots, \beta_n)$. Ці коефіцієнти визначають форму та поведінку апроксимуючої функції.

Головна особливість натурального сплайну регресії полягає в тому, що на кінцях діапазону значень незалежної змінної застосовується умова нульової другої похідної. Це забезпечує гладкість та неперервність функції апроксимації на всьому діапазоні.

Натуральний сплайн регресії часто використовується в статистичному аналізі та машинному навчанні для моделювання складних залежностей та передбачення значень залежної змінної. Він може бути застосований у різних сферах, таких як економіка, фінанси, медицина, соціальні науки та багато інших.



$$R^2 = 0.7872911042486117$$

Мал. 9 - N-сплайн регресія на трен та на тест та коефіцієнт детермінації

3.6 Гладкий кубічний сплайн який базується на крос-валідації

Гладкий сплайн є методом апроксимації функції за допомогою кусково-поліноміальних сегментів. Він може бути представлений наступною формулою:

$$S(x) = \sum_{i=1}^n y_i B_i(x)$$

де:

- $S(x)$ - гладкий сплайн,
- x - вхідний параметр,
- y_i - значення залежної змінної на i -му вузлі сплайна,
- $B_i(x)$ - базисна функція Б-сплайна на i -му вузлі.

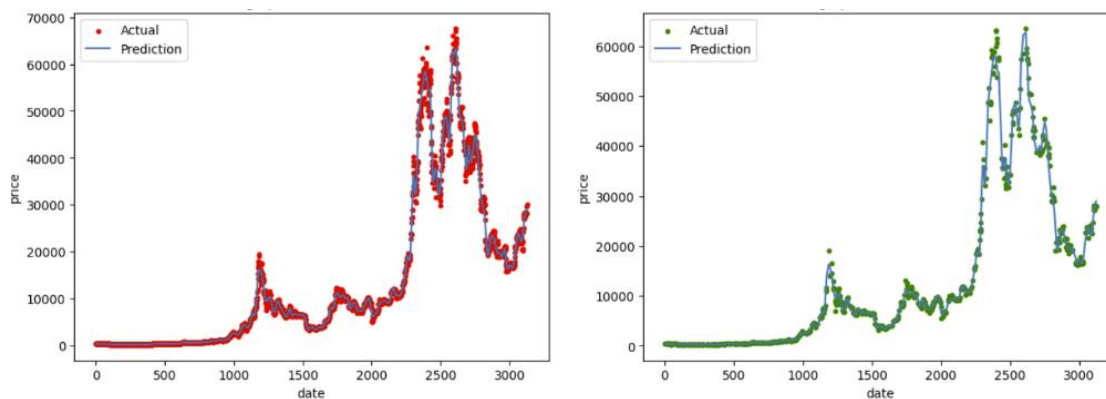
Дані:

- Тренувальний набір даних, що складається з пар (x_i, y_i) , де x_i - вхідний параметр, y_i - значення залежної змінної.

Опис методу:

1. Розділення тренувального набору даних на підвибірки.
2. Для кожної підвибірки виконується побудова поліноміального сплайна.
3. Використовується метод гладкої апроксимації, який забезпечує гладкість сплайна на перетинах сегментів.
4. Параметри сплайна визначаються шляхом розв'язування системи лінійних рівнянь або іншими методами оптимізації.
5. Побудований сплайн може бути використаний для прогнозування значень залежної змінної на нових вхідних параметрах.

Гладкий сплайн дозволяє апроксимувати функцію з високою точністю, забезпечує гладкість і безшовність на перетинах сегментів, що дозволяє отримати більш натуральну і плавну апроксимацію

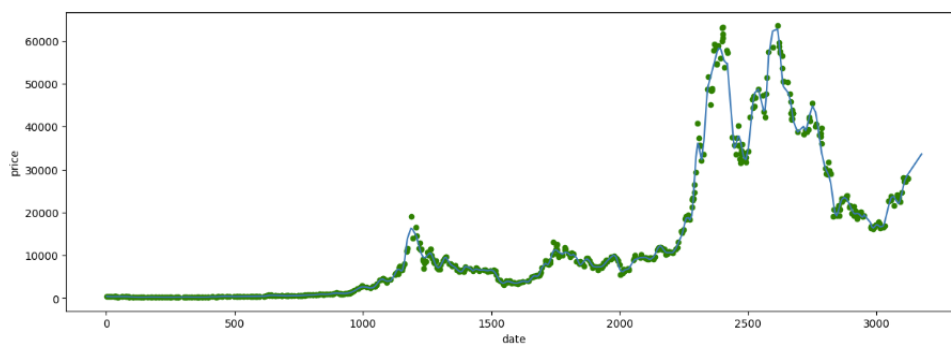


$$R^2 = 0.9949724359380482$$

Мал. 10 - В-сплайн регресія квантилі 25, 50, 75 на трен та на тест та коефіцієнт детермінації

Згладжуючий сплайн - лямбда, знайдена за допомогою перехресної перевірки: $1.4662757817507642e - 09$.

Можемо побачити наскільки точніше гладкий сплайн, ніж всі інші функції. Далі я зробив прогноз найточнішим методом



Мал. 11 - Прогнозування майбутньої ціни

Тож в кінці я вивів коефіцієнт R^2 всіх методів щоб порівняти їх.

3.7 Підводимо підсумки

	R2
Linear Regression	0.568840
22-order Polynomial	0.928341
44-order Polynomial	0.939202
66-0order Polynomial	0.946904
88-0order Polynomial	0.956095
B-spline	0.902780
Cubic B-spline, knots manually chosen	0.944814
n-spline	0.811086
Smoothing spline	0.995926

Мал. 12 - Таблиця Коефіцієнтів для порівняння R^2

І ми можемо побачити, наскільки гладкий сплайн є кращим за інші, зокрема, під час крос-валідації. Крім того, слід пам'ятати, що В-сплайн з обраними точками може мати потенціал стати таким же, або навіть кращим, ніж гладкий сплайн.

ВИСНОВОК

За отриманими результатами обчислення коефіцієнта детермінації R^2 , можна зробити наступні висновки:

1. Лінійна регресійна модель демонструє низьку придатність у поясненні варіації вихідної змінної з R^2 рівним 0.569.

2. Зі збільшенням степеня полінома в поліноміальній регресійній моделі значення R^2 зростає. Наприклад, поліном 88-го порядку досягає високого значення R^2 у 0.956, що свідчить про високу придатність моделі у поясненні даних.
3. Модель В-сплайну демонструє задовільну здатність у поясненні даних з R^2 рівним 0.903. Цей метод апроксимації використовується для гнучкого моделювання залежностей, особливо коли вихідні дані мають складну структуру.
4. Кубічний В-сплайн з ручним вибором вузлів досягає високого значення R^2 у 0.945, що підтверджує його високу точність у відтворенні залежності між змінними.
5. Модель N-сплайну здатна пояснити лише невелику частину варіації даних, що відображається в низькому значенні R^2 рівним 0.811. Цей метод апроксимації не є оптимальним для даного набору даних.
6. Згладжувальний сплайн демонструє найвищу придатність у прогнозуванні значень залежної змінної з високим значенням R^2 у 0.996. Цей метод використовується для створення гладкої апроксимації, зменшення впливу шуму та покращення узагальнюючих здатностей моделі.

Список використаних джерел

1. Гаврилюк І.П., Макаров В.Л. Методи обчислень: Підручник: У 2 частинах. - К.: Вища школа, 1995р.
2. Голубєва К.М. Денисов С.В., Кашпур О.Ф., Ключин Д.А., Риженко А.І. Чисельні методи Інтегрування Голубєва К.М., Денисов С.В., Кашпур О.Ф., Ключин Д.А., Риженко А.І. Чисельні методи

інтегрування. – Київ: Київський національний університет імені Тараса Шевченка, 2019 – 55 с.

3. Hayward, T., & Preston, J. Chaos theory, economics and information: the implications for strategic decision-making. *Journal of Information Science*, Vol. 25 №3. 1999. P. 173–182.
<https://doi.org/10.1177/016555159902500301>
4. Faggini Marisa, Parziale Anna. The Failure of Economic Theory. Lessons from Chaos Theory. *Modern Economy*. Vol. 3, №1. 2012. P. 1-10. doi: [10.4236/me.2012.31001](https://doi.org/10.4236/me.2012.31001).
5. Enrico Scalas¹, Introduction to Probability Theory for Economists. 2008. 58 P.
URL: <https://www.lem.sssup.it/phd/documents/probpisanew.pdf>
6. Lecture by George Soros, transcript: General Theory of Reflexivity.
URL:
<https://www.opensocietyfoundations.org/uploads/9ae17912-2262-4646-8ffc-d01afc934c36/george-soros-general-theory-of-reflexivity-transcript.pdf>
7. Burton G. Malkiel, The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, Vol. 17, №1. 2003. P. 59-82.
<https://doi.org/10.1257/089533003321164958>
8. Eric Van Damme, Game Theory and the Market. URL:
https://www.researchgate.net/publication/226708757_Game_Theory_and_the_Market
9. Wahba, Spline Models for Observational Data. CBMS-NSF Regional Conference Series in Applied Mathematics.
<https://doi.org/10.1137/1.9781611970128>
10. Putri, K. S., & Halim, S. Currency movement forecasting using time series analysis and long short-term memory. *International Journal of Industrial Optimization*, Vol. 1 №2. 2020. P. 71–80.
<https://doi.org/10.12928/ijio.v1i2.2490>

11. Varsity by Zerodha, Technical Analysis. URL: <https://zerodha.com/varsity/module/technical-analysis/>
12. CA Ashok Ajmera-CMD, Ajcon Global Services Ltd. Fundamental Analysis. URL: <https://ru.scribd.com/document/641723689/Untitled>
13. Varsity by Zerodha, Fundamental Analysis. URL: <https://zerodha.com/varsity/module/fundamental-analysis/>
14. Ali Raheman, Social Media Sentiment Analysis for Cryptocurrency Market Prediction. URL: <https://arxiv.org/ftp/arxiv/papers/2204/2204.10185.pdf>
15. Daniele Bernardi, Ruggero Bertelli. Bitcoin price forecast using quantitative models. 2021. URL: https://www.researchgate.net/publication/353377031_Bitcoin_price_forecast_using_quantitative_models
16. John F. Muth. Rational Expectations and the Theory of Price Movements URL: https://www.fep.up.pt/docentes/pcosme/s-e-1/se1_trab_0910/se1.pdf

ВІДГУК
на випускню кваліфікаційну роботу бакалавра
«Чисельний прогноз криптовалюти за даними останніх років»
студента 4-го курсу кафедри обчислювальної математики
факультету комп'ютерних наук та кібернетики
Київського національного університету імені Тараса Шевченка
Тархуні Анатолія Алекса

Робота присвячена аналізу курсу біткоіна, використовуючи: лінійну регресію, поліноміальну регресію, B-сплайн інтерполяцію, N-сплайн інтерполяцію, гладкий кубічний сплайн на основі крос-валідації. Автор розглядає основні теоретичні ідеї та співвідношення для кожного методу, супроводжує отримані результати графіками коливань курсу та їх наближенням. Виконується прогноз курсу за відомими значеннями, та порівнюється з фактично отриманим показником. Для аналізу якості побудованої апроксимації, для кожного випадку перевірено чисельний параметр відхилення, отримані результати є достатньо точними.

Методи побудови апроксимації, наведені у роботі, дозволяють виконати аналіз коливань курсу та обрати набір функцій, що апроксимує курс біткоіна.

Студент самостійно виконав роботу та досяг гарних результатів кількома підходами. Вважаю, що він продемонстрував належний рівень знань та кваліфікації. Роботу можу оцінити на «відмінно».

Асистент кафедри обчислювальної
математики факультету комп'ютерних
наук та кібернетики
Київського національного університету
імені Тараса Шевченка,
доктор філософії



Андрій
ТИМОШЕНКО

РЕЦЕНЗІЯ

на кваліфікаційну роботу бакалавра:

«Чисельний прогноз криптовалюти за даними останніх років»

студента 4-го курсу кафедри обчислювальної математики

факультету комп'ютерних наук та кібернетики

Київського національного університету імені Тараса Шевченка

Тархуні Анатолія Алекса

Рецензована робота присвячена аналізу курсу криптовалюти біткоїн за допомогою лінійної регресії, поліноміальної регресії та кількох видів інтерполяції. Дослідження супроводжується графічним представленням отриманих результатів та оцінкою різниці між фактичними і апроксимованими показниками для кожного методу. Використовуючи актуальні дані, було встановлено корисність усереднення даних та порівняно ефективність запропонованих підходів.

Робота відображає поєднання теоретичного опису основної ідеї кожного методу, його практичне застосування та аналіз отриманої точності, демонстрацію порівняння справжніх значень з функціональним наближенням.

Студент виконав поставлене завдання та заслуговує на оцінку «відмінно», а також на присвоєння кваліфікації бакалавра.

Професор кафедри обчислювальної
математики факультету комп'ютерних
наук та кібернетики
Київського національного університету
імені Тараса Шевченка,
доктор фізико-математичних наук



Дмитро КЛЮШИН

Ім'я користувача:
Оноцький В'ячеслав ФКомпНаук

ID перевірки:
1015426592

Дата перевірки:
05.06.2023 11:40:04 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
05.06.2023 11:50:30 EEST

ID користувача:
100002816

Назва документа: ТархуніАнатолійАлекс

Кількість сторінок: 52 Кількість слів: 8020 Кількість символів: 60976 Розмір файлу: 1.29 MB ID файлу: 1015088719

Виявлено модифікації тексту (можуть впливати на відсоток схожості)

0.8% Схожість

Найбільша схожість: 0.44% з джерелом з Бібліотеки (ID файлу: 1014448825)

0.42% Джерела з Інтернету

52

Сторінка 54

0.69% Джерела з Бібліотеки

46

Сторінка 54

0% Цитат

Вилучення цитат вимкнене

Вилучення списку бібліографічних посилань вимкнене

0% Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи

13

Підозріле форматування

8
сторінок

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

СИСТЕМА ЗАПОБІГАННЯ ТА ВИЯВЛЕННЯ АКАДЕМІЧНОГО ПЛАГІАТУ

Довідка про оригінальність кваліфікаційної роботи за освітнім рівнем бакалавр

Експертна оцінка роботи науковим керівником :

Робота студента 4-го курсу Тархуні Анатолія Алекса «Чисельний прогноз криптовалюти за даними останніх років» виконана самостійно, при цьому обсяг цитувань та запозичень становить 0.8% та не перевищує норму.

Науковий керівник:

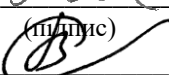


(підпис)

Тимошенко А.А.

(ПБ)

Оператор:



(підпис)

Оноцький В.В.

(ПБ)