

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО–НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ

Кафедра комп'ютерної інженерії

Пояснювальна записка

до магістерської роботи
на ступінь вищої освіти
магістр

на тему: «**ВДОСКОНАЛЕННЯ МЕТОДУ ПРОГНОЗУВАННЯ**
ФІНАНСОВИХ ЧАСОВИХ РЯДІВ НА ОСНОВІ МЕРЕЖІ РАДІАЛЬНО-
БАЗИСНИХ ФУНКЦІЙ»

Виконав: студент 6 курсу, групи ПДМ–61
спеціальності

121 _____ Інженерія програмного забезпечення
(шифр і назва спеціальності)

_____ Кривошопка Д.О.

(прізвище та ініціали)

Керівник _____ Трінтіна Н.А.

(прізвище та ініціали)

Рецензент _____

(прізвище та ініціали)

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра Інженерія програмного забезпечення

Ступінь вищої освіти - «Магістр»

Напрямок підготовки - 121 «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ

Завідувач кафедри
Комп'ютерної інженерії

_____ О.В. Негоденко

— ___ || _____ 2022 року

ЗАВДАННЯ
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Кривошاپка Дмитро Олександрович

(прізвище, ім'я, по батькові)

1. Тема роботи: «Вдосконалення методу прогнозування фінансових часових рядів на основі мережі радіально базисних функцій»

Керівник роботи к.т.н., доц. Трінтіна Наталія Альбертівна,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затвержені наказом вищого навчального закладу від — 12 жовтня року №122.

2. Строк подання студентом

роботи 31.12.2022

3. Вихідні дані до роботи: Матеріали переддипломної практики, методи та моделі прогнозування, мережі прогнозування.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити):

4.1 Огляд предметної області;

4.2 Розробка моделей та методів;

4.3 Розробка програмного забезпечення моделей;

4.4 Проведення моделювання та аналіз отриманих результатів

5. Перелік графічного матеріалу.

5.1 Мета, об'єкт та предмет дослідження

5.2 Порівняння існуючих моделей та методів прогнозування

5.3 Структура нейромережі

5.4 Математична модель радіально базисної функції

5.5 Діаграма роботи системи

5.6 Різниця точності прогнозування в залежності від кількості епох навчання нейронної мережі

5.7 Результати прогнозування

6. Дата видачі завдання 14» жовтня 2022 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Отримання завдання на магістерську роботу	14.10.2022	
2	Огляд предметної області	16.10.2022	
3	Аналіз моделей, функцій та методів прогнозування часових рядів	07.11.2022	
4	Розробка функцій прогнозування	10.11.2022	
5	Розробка програмного забезпечення функцій	11.11.2022	
6	Тестування програмного забезпечення	12.11.2022	
7	Моделювання та аналіз результатів	13.11.2022	
8	Написання та оформлення пояснювальної записки	14.11.2022	
9	Розробка графічних та презентаційних матеріалів	08.12.2022	
10	Захист магістерської роботи	17.01.2022	

Студент _____ Кривошапка Д.О.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Н.А. Трінтіна
підпис) (прізвище та ініціали)

РЕФЕРАТ

Текстова частина магістерської роботи : 89 с., 3 табл., 28 рис., 1 дод., 23 джерела.

РАДІАЛЬНО БАЗИСНА ФУНКЦІЯ, МОДЕЛЬ ДЛЯ АНАЛІЗУ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ PYTHON

Об'єкт дослідження – прогнозування фінансових часових рядів.

Мета роботи – підвищення ефективності прогнозування фінансових часових рядів для біржових торгів за допомогою радіально базисних функцій.

Методи дослідження — методи прогнозування часових рядів, методи проектування та розробки програмного забезпечення, методи роботи з нейронними мережами

Проведено дослідження моделей та методів аналізу часових рядів. Визначено, найзручніше використовувати експонціанальну систему для аналізу часових рядів так як вона найбільш ефективна в даній сфері.

Виявлено що найдоцільніше використовувати радіально базисні функції у поєднанні з експонціанальною моделлю так як вона найсумісніша з даною моделлю.

На основі результатів виконаних досліджень запропоновано до реалізації універсальний алгоритм для пошуку закономірностей у часових рядах та прогнозування на їх основі майбутніх значень.

Отримані результати дозволяють з'ясувати з визначеною вірогідністю майбутній рух ринку.

ЗМІСТ

ВСТУП	10
1 ПОСТАНОВКА ЗАВДАННЯ ТА ОГЛЯД МОДЕЛЕЙ ПРОГНОЗУВАННЯ ТИМЧАСОВИХ РЯДІВ	13
1.1 Обґрунтування актуальності тематики дослідження та постановка задачі	13
1.2 Огляд моделей прогнозування	14
1.2.1 Регресійні моделі	15
1.2.2 Авторегресійні моделі	17
1.2.3 Моделі експонентного згладжування	19
1.2.4 Нейромережеві моделі	21
1.2.5 Моделі на базі класифікаційно-регресійних дерев	26
1.2.6 Моделі на основі радіально-базисних функцій	31
1.2.7 Інші моделі та методи прогнозування	32
1.3 Порівняння моделей прогнозування	34
1.4 Комбіновані методи	41
2 МОДЕЛІ ЕКСТРАПОЛЯЦІЇ ТИМЧАСОВИХ РЯДІВ З ВИБІРКИ МАКСИМАЛЬНОГО ПОДОБИ	44
2.2 Модель без урахування зовнішніх факторів	44
2.1.1 Вибірки часового ряду	44
2.1.2. Апроксимація вибірки	45
2.1.3. Подібність вибірок	48
2.1.4. Опис моделі екстраполяції	52
2.2. Модель з урахуванням зовнішніх факторів	55

2.2.1. Вибірки часових рядів	55
2.2.2. Апроксимація вибірки	55
2.2.3. Подібність вибірок	57
2.2.4. Опис моделі	57
2.3. Варіанти моделей щодо вибірки максимального подоби	60
2.4. Висновки	63
3 МЕТОД ПРОГНОЗУВАННЯ НА МОДЕЛІ ЕКСТРАПОЛЯЦІЇ З ВИБІРКИ МАКСИМАЛЬНОГО ПОДОБИ	64
3.1. Алгоритм екстраполяції тимчасового ряду без урахування зовнішніх факторів	64
3.2. Алгоритм ідентифікації моделі	70
3.2.1 Опис алгоритму	70
3.2.2. Розпаралелювання обчислень	72
3.2.3. Набори моделей	73
3.3.4. Оцінка часу ідентифікації	75
3.3. Висновки	76
4 РОЗРОБКА ПРОГРАМНОГО ЗАСТОСУВАННЯ ПРОГНОЗУВАННЯ ФІНАНСОВИХ ТИМЧАСОВИХ РЯДІВ НА ОСНОВІ ВИКОРИСТАННЯ РАДІАЛЬНО БАЗИСНИХ ФУНКЦІЙ	77
4.1 Програмна реалізація	77
4.2 Опис результатів виконання прогнозування даних	79
ВИСНОВКИ	86

ВСТУП

Актуальність теми. Завдання прогнозування майбутніх значень тимчасового ряду на основі його історичних значень є основою для фінансового планування в економіці та торгівлі, планування, управління та оптимізації обсягів виробництва, складського контролю [1,2]. В нинішній час компанії накопичують історичні значення економічних та фізичних показників у базах даних, що суттєво збільшує обсяги вхідної інформації для Завдання прогнозування. Разом з тим, розвиток апаратних та програмних коштів потребує все більші обчислювальні платформи завдяки яким можлива реалізація складних алгоритмів прогнозування. Крім того, нинішні методи та підходи до технічного та економічного управління пред'являють дедалі жорсткіші вимоги до точності прогнозування.

Отже, завдання прогнозування часових рядів більш ускладнюється одночасно із розвитком інформаційних технологій. В даний час завдання прогнозування різних часових рядів актуальна і є невід'ємною частиною щоденної роботи багатьох компаній.

Завдання прогнозування часового ряду вирішується на основі створення моделі прогнозування, що адекватно описує досліджуваний процес. На сьогоднішній день існує безліч моделей прогнозування тимчасових рядів: регресійні та авторегресійні моделі, нейромережеві моделі, моделі експоненційного згладжування, моделі на основі ланцюгів Маркова, класифікаційні моделі та інші.

Популярними та часто використовуваними є авторегресійні та нейромережеві класи моделей [3]. Істотним недоліком авторегресійного класу є велика кількість вільних параметрів, ідентифікація яких неоднозначна та ресурсомістка [4]. Істотним недоліком класу нейромережевих моделей є недоступність проміжних обчислень, виконуються в «чорній скриньці», і, як наслідок, складність інтерпретація результатів моделювання. Крім того, ще одним недоліком даного класу моделей є складність вибору алгоритму навчання нейронної мережі [5].

Магістерська робота присвячена розробці нової радіально-базисної моделі прогнозування, яка має порівнянну з іншими моделями ефективність

прогнозування різних часових рядів і при цьому усуває основний і найбільш істотний недолік авторегресійного класу моделей – велика кількість вільних параметрів.

Метою роботи є розробка нової моделі та відповідного їй методу прогнозування, що відноситься до класу радіально-базисних функцій та усуває основний недолік даного класу моделей велика кількість вільних властивостей. Нова модель та відповідний їй метод повинен мати високу швидкість обчислення прогнозних значень та порівнянну з іншими моделями точність прогнозування різних тимчасових рядів. Для досягнення цієї мети були поставлені та вирішені наступні завдання:

1. Здійснити огляд моделей та методів прогнозування тимчасових рядів, виявити переваги та недоліки кожного класу моделей. Виявити найбільш використовувані класи моделей прогнозування та їх основні недоліки, визначити перспективні підходи, що дозволяють усунути недоліки авторегресійного класу моделей.

2. Розробити нову модель прогнозування часових рядів, усуває зазначений недолік авторегресійного класу моделей.

3. Розробити новий метод прогнозування на підставі запропонованої моделі та виконати програмну реалізацію алгоритмів.

4. Оцінити ефективність запропонованої моделі прогнозування під час вирішення завдання прогнозування різних часових рядів.

Методи дослідження. При вирішенні поставлених завдань у роботі використано методи математичного моделювання, аналіз тимчасових рядів, регресійний аналіз, методи об'єктно-орієнтованого програмування.

Наукова новизна. У магістерській роботі отримано такі основні результати, що виносяться на захист.

1. Модель екстраполяції часових рядів за вибіркою максимальної подоби, що відноситься до класу авторегресійних моделей і має єдиний параметр.

2. Метод прогнозування часових рядів на підставі розробленої радіально-базисної моделі, що містить набір алгоритмів для екстраполяції тимчасових рядів, ідентифікації моделі та побудови довірчого інтервалу прогнозних значень.

Практична цінність роботи. Розроблена модель та метод прогнозування щодо вибірки максимальної подоби можуть застосовуватися для прогнозування часових рядів різних предметних областей

Розроблені алгоритми екстраполяції часових рядів з урахуванням та без урахування зовнішніх факторів наочні для програмної реалізації. Швидкість обчислення прогнозних значень під час використання моделі висока. Завдання ідентифікація моделі спрощена в порівнянні з іншими моделями авторегресійного класу

1 ПОСТАНОВКА ЗАВДАННЯ ТА ОГЛЯД МОДЕЛЕЙ ПРОГНОЗУВАННЯ ТИМЧАСОВИХ РЯДІВ

1.1 Обґрунтування актуальності тематики дослідження та постановка задачі

Слово прогноз виникло від грецького πρόβωσις, що означає передбачення, передбачення. Під прогнозуванням розуміють пророцтво майбутнього з допомогою наукових методів. Часовий ряд включає два обов'язкові елементи - мітку часу та значення показника серії, отримані тим чи іншим способом і згідно з окремим способом і відповідно до вказаної відмітки часу.

Кожен часовий ряд розглядається як вибірка розгортання з нескінченної популяції, створеної стохастичним процесом, на який впливає багато факторів [1]. На рисунку 1.1 наведено приклад часового ряду цін на електроенергію.

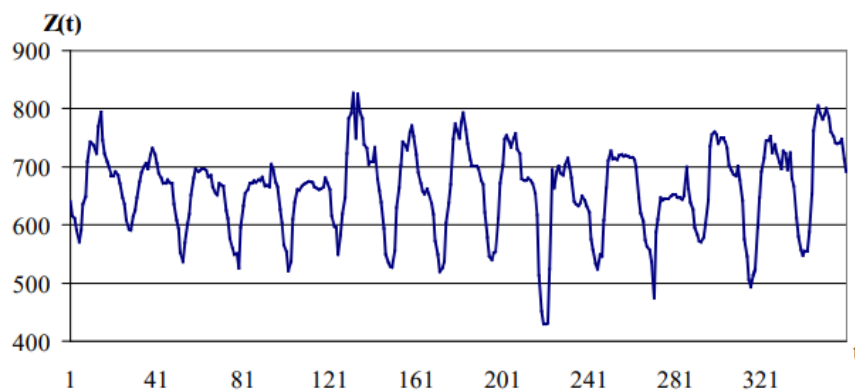


Рисунок 1.1 - Приклад часового ряду цін на електроенергію

Одну з класифікацій часових рядів наведено в роботі [6]. Відповідно часові ряди відрізняються за своїм визначенням.

Крім того, метод прогнозування містить дії щодо оцінки якості прогнозних значень. Загальний ітеративний підхід до побудови моделі прогнозування складається з наступних кроків [1].

Крок 1. На першому кроці на підставі попереднього власного або стороннього досвіду вибирається загальний клас моделей для прогнозування тимчасового ряду на заданий обрій.

Крок 2. Певний загальний клас моделей є широким. Для безпосереднього припасування до вихідного тимчасового ряду, розвиваються грубі методи ідентифікації підкласів моделей. Такі методи ідентифікації використовують якісні оцінки часового ряду.

Крок 3. Після визначення підкласу моделі необхідно оцінити її параметри, якщо модель містить параметри, або структуру, якщо модель відноситься до категорії структурних моделей (розділ 1.3). На цьому етапі зазвичай використовуються ітеративні способи, коли проводиться оцінка ділянки (або всього) тимчасового ряду при різних значеннях змінних величин. Як правило, цей крок є найбільш трудомістким у зв'язку з тим, що часто враховуються всі доступні історичні значення тимчасового ряду.

1.2 Огляд моделей прогнозування

Перед тим як перейти до огляду моделей, слід зазначити, що назви моделей та відповідних методів як правило збігаються. Наприклад, роботи [1,11-13] присвячені одній з найпоширеніших моделей прогнозування авторегресія проінтегрованого ковзного середнього з урахуванням зовнішнього фактора. Інтуїтивні методи прогнозування не передбачають розробку моделей прогнозування та відображають індивідуальні міркування фахівців щодо перспектив розвитку процесу. Крім того, в даний час широко поширене застосування експертних систем, зокрема з використанням нечіткої логіки [15]. У статті [16] докладно описано інтуїтивні методи прогнозування

Крім того, слід зазначити, що для вузькоспеціалізованих задач іноді застосовуються спеціальні моделі прогнозування. Так наприклад, для завдання прогнозування рівня цукру крові людини застосовуються моделі з урахуванням диференціальних рівнянь [8]. Для завдання прогнозування транспортного потоку, що в останні

кілька років актуальна для мегаполісів, використовуються гідродинамічні моделі [17].

Для прогнозування природних явищ, таких як землетруси, застосовується, наприклад, модель, основою якої покладено нелінійні клітини (або стільники), що знаходяться під впливом зовнішнього поля, і які мають внутрішній стан, що змінюється в часі під впливом цього поля [18]. Аналогічні моделі розробляються і застосовуються для спеціальних процесів та систем.

1.2.1 Регресійні моделі

Модель регресії надає функцію, яка описує зв'язок між однією або декількома незалежними змінними та змінною відповіді, залежною чи цільовою змінною.

Наприклад, залежність між зростом і вагою можна описати моделлю лінійної регресії. Регресійний аналіз є основою для багатьох типів прогнозів і для визначення впливу на цільові змінні. Коли ви чуєте про дослідження в новинах, які розповідають про економію палива, чи причину забруднення, чи вплив часу, проведеного за екраном, на навчання, для підтвердження своїх тверджень часто використовують регресійну модель.

Типи регресії:

1. Лінійна

Лінійна регресія – це модель, у якій зв'язок між входами та результатами є прямою лінією. Це найпростіше концептуалізувати і навіть спостерігати в реальному світі. Навіть коли зв'язок не дуже лінійний, наш мозок намагається побачити закономірність і приєднати до цього зв'язку елементарну лінійну модель.

Одним із прикладів може бути кількість відгуків на маркетингову кампанію. Якщо ми надішлемо 1000 електронних листів, ми можемо отримати п'ять відповідей. Якщо цю залежність можна змоделювати за допомогою лінійної регресії, ми очікуємо отримати десять відповідей, коли надішлемо 2000 електронних листів. Ваша діаграма може відрізнятись, але загальна ідея полягає в тому, що ми пов'язуємо предиктор і ціль і припускаємо зв'язок між ними.

Використовуючи модель лінійної регресії, ми хочемо оцінити кореляцію між кількістю надісланих електронних листів і рівнем відповідей. Іншими словами, якщо лінійна модель достатньо добре відповідає нашим спостереженням, ми можемо оцінити, що чим більше електронних листів ми надішлемо, тим більше відповідей ми отримаємо.

Коли висуваються подібні твердження, незалежно від того, пов'язані вони з фізичними вправами, щастям, здоров'ям чи будь-якими іншими твердженнями, за лаштунками зазвичай існує регресійна модель, яка підтверджує це твердження.

Крім того, підгонку моделі можна описати за допомогою середньоквадратичної помилки. По суті, це дає нам число, яке точно показує, наскільки добре підходить лінійна модель.

Серйозніші приклади лінійної регресії включають прогнозування тривалості перебування пацієнта в лікарні, співвідношення між доходом і злочинністю, освітою і народжуваністю, продажами і температурою.

2. Множинна

Множинна регресія вказує на те, що існує більше ніж одна вхідна змінна, яка може вплинути на результат або цільову змінну. Для нашого прикладу електронної кампанії ви можете включити додаткову змінну з кількістю електронних листів, надісланих за останній місяць.

Розглядаючи обидві вхідні змінні, починає з'являтися чіткіша картина про те, що спонукає користувачів реагувати на кампанію та як оптимізувати час і частоту надсилання електронних листів. Хоча концептуалізація моделі стає складнішою з більшою кількістю вхідних даних, зв'язок може залишатися лінійним.

Для цих моделей важливо точно розуміти, який вплив має кожне вхідне значення та як вони поєднуються для отримання кінцевих результатів цільової змінної.

3. Нелінійна

Лінійна модель може працювати для деяких частин маркетингового прикладу вище. Однак ми знаємо, що в міру збільшення кількості електронних листів у певній кампанії кількість відповідей починає зменшуватися порівняно з кількістю

надісланих електронних листів. Щоб змодельовати це, нам потрібна модель нелінійної регресії.

Ця модель спочатку показуватиме позитивний зв'язок між кількістю електронних листів і відповіддю, але зі збільшенням кількості електронних листів модель згладжується та стає майже постійною. Важливим висновком тут є те, що важливо розуміти, коли модель потенційно може бути нелінійною. Якщо ви робите припущення на основі лінійної моделі, ви можете отримати результати, які дуже відрізняються від очікувань.

4. Покрокове регресійне моделювання

У той час як інші елементи, про які ми говорили досі, є конкретними типами моделей, поетапна регресія є скоріше технікою. Якщо модель включає багато потенційних вхідних даних, аналітик може почати з найбільш безпосередньо корельованої вхідної змінної для побудови моделі. Коли це буде зроблено, наступним кроком буде зробити модель більш точною.

Для цього додаткові вхідні змінні можна додавати до моделі по черзі в порядку значущості результатів. Використовуючи наш приклад електронного маркетингу, початкова модель може базуватися лише на кількості надісланих електронних листів. Тоді ми додамо щось на зразок середнього віку одержувача електронного листа. Після цього ми додамо середню кількість електронних листів, які кожен одержувач отримав від нас. Кожна додаткова змінна додасть невелику додаткову точність моделі.

Цей процес є поетапним підходом до моделювання для отримання найточнішої моделі. Як альтернатива, аналітик може почати з більшого набору вхідних змінних, а потім поступово видалити найменш значущі, щоб дістатися до бажаної моделі.

1.2.2 Авторегресійні моделі

У моделі множинної регресії ми прогнозуємо цікаву змінну за допомогою лінійної комбінації предикторів. У моделі авторегресії ми прогнозуємо цікаву

змінну, використовуючи лінійну комбінацію минулих значень змінної. Термін авторегресія вказує на те, що це регресія змінної відносно самої себе.

Таким чином, авторегресійна модель порядку P можна записати як:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t, \quad (1.1)$$

де ε_t це білий шум. Це схоже на множинну регресію, але з відсталими значеннями y_t як провісники. Як правило її називають $AR(p)$ модель, авторегресійна модель порядку p .

Авторегресійні моделі надзвичайно гнучкі в обробці широкого діапазону різних шаблонів часових рядів. Дві серії на рисунку _ показують серії з моделі $AR(1)$ і моделі $AR(2)$. Зміна параметрів $\varphi_1, \dots, \varphi_p$ призводить до різних моделей часових рядів. Дисперсія терміну помилки ε_t змінить лише масштаб серії, а не візерунки.

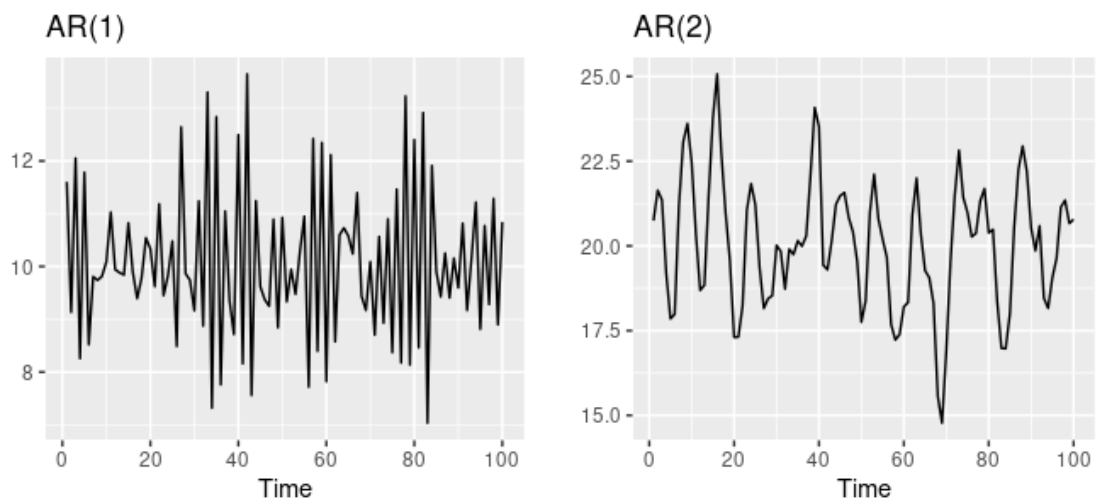


Рисунок 1.2 Два приклади даних авторегресійних моделей з різними параметрами

Для моделі $AR(1)$:

1. коли $\varphi_1=0$, y_t еквівалентно білому шуму;
2. коли $\varphi_1=1$ і $c=0$, y_t еквівалентно випадковому блуканню;

3. коли $\phi_1=1$ і $c \neq 0$, y_t еквівалентно випадковому блуканню з дрейфом;
4. коли $\phi_1 < 0$, y_t має тенденцію коливатися навколо середнього значення.

1.2.3 Моделі експонентного згладжування

Моделі експоненційного згладжування розроблені в середині XX ст. століття і до сьогодні є широко поширеними в силу їх простоти та наочності.

Модель експонентного згладжування (exponential smoothing, ES) використовується для фінансово-економічного моделювання процеси [24]. Основою експоненціального згладжування є ідея постійна перевірка прогнозних значень при отриманні поточний. Модель ES призначає експоненціально зменшуючі ваги. спостереження в літньому віці. Отже, останні доступні спостереження мають більший вплив на прогнозну цінність, ніж старіші Моніторинги.

Функція моделі ES має вигляд

$$Z(t) = S(t) + \varepsilon_t, \quad (1.2)$$

$$S(t) = a * Z(t - 1) + (1 + a) * S(t - 1), \quad (1.3)$$

де a - коефіцієнт згладжування, $0 < a < 1$; початкові умови визначаються як $S(1)=Z(0)$. У цій моделі кожне наступне згладжене значення $S(t)$ є виваженим середнім між попереднім значенням часового ряду $Z(t)$ та попереднього згладженого значення $S(t-1)$.

Модель Хольта або подвійне експонентне згладжування застосовується для моделювання процесів, що мають тренд. В цьому випадку у моделі необхідно розглядати дві складові: рівень та тренд [24].

Рівень та тренд згладжуються окремо

$$Z(t) = S(t) + \varepsilon_t; \quad (1.4)$$

$$S(t) = a * Z(t - 1) + (1 - a) * (S(t - 1) - B(t - 1)); \quad (1.5)$$

$$B(t) = y * (S(t - 1) - S(t - 2)) + (1 + y) * B(t - 1). \quad (1.6)$$

Тут a — коефіцієнт згладжування рівня, як і моделі, y — коефіцієнт згладжування тренду.

Модель Хольта-Вінтерса або потрійне експоненційне згладжування застосовується для процесів, які мають тренд та сезонну складову.

$$Z(t) = (R(t) + G(t)) * S(t). \quad (1.7)$$

Тут $R(t)$ — згладжений рівень без урахування сезонної складової

$$R(t) = \frac{a * Z(t-1)}{S(t-L)} + (1 + a) * (R(t - 1) + G(t - 1)), \quad (1.8)$$

$G(t)$ - згладжений тренд

$$G(t) = \beta * (S(t - 1) - S(t - 2)) + (1 - \beta) * G(t - 1), \quad (1.9)$$

а $S(t)$ — сезонна складова

$$S(t) = \frac{y * Z(t-1)}{S(t-L)} + (1 - y) * S(t - L). \quad (1.10)$$

Розмір L визначається довжиною сезону досліджуваного процесу.

Моделі експоненційного згладжування найбільш популярні довгострокового прогнозування.

Моделі штучних нейронних мереж лежать в основі багатьох найскладніших застосувань машинного навчання. Класифікація, проблеми регресії та аналіз настроїв – це лише деякі способи використання штучних нейронних мереж сьогодні. Існує багато різних типів штучних нейронних мереж. Вони відрізняються

з різних причин, таких як складність, архітектура мережі, щільність і потік даних. Але різні типи мають спільну мету моделювання та спроб відтворити поведінку нейронів для покращення машинного навчання.

1.2.4 Нейромережеві моделі

Штучні нейронні мережі мають широкий спектр використання в машинному навчанні. Кожен тип моделі штучної нейронної мережі має різні переваги та варіанти використання. Загалом, вони в основному використовуються для вирішення більш складних проблем, ніж це було б можливо за допомогою більш традиційних методів машинного навчання. Приклади можуть включати складну обробку природної мови та мовний переклад із можливостями машинного навчання, які покладаються на штучні нейронні мережі. Повторювані нейронні мережі також часто використовуються для аналізу настроїв або перекладу тексту. Глибина та масштаб нейронної архітектури означає, що можна досягти нелінійного процесу прийняття рішень.

Штучні нейронні мережі використовуються у формі глибокого навчання машинного навчання. Це називається глибоким навчанням, оскільки моделі використовують «глибоку» багаторівневу архітектуру штучної нейронної мережі. Оскільки кожен рівень штучної нейронної мережі може обробляти дані, моделі можуть будувати абстрактне розуміння даних. Ця архітектура означає, що моделі можуть виконувати дедалі складніші завдання, наприклад, розуміння природної мови або категоризацію складних типів файлів.

Штучні нейронні мережі вже використовуються в машинному навчанні для:

1. Системи рекомендацій для клієнтів, користувачів і споживачів таких продуктів, як потокові послуги або електронна комерція.
2. Для забезпечення віртуальної допомоги та програмного забезпечення для розпізнавання мовлення.
3. Складні моделі класифікації зображень, аудіо та документів, наприклад, у програмному забезпеченні розпізнавання обличчя.

4. У автоматичному вилученні ознак із необроблених даних без міток.

Існують різні типи штучних нейронних мереж, які відрізняються за складністю. У цьому посібнику розглядаються різні типи штучних нейронних мереж, зокрема, що вони собою являють і як вони використовуються.

Існує багато різних типів штучних нейронних мереж, що відрізняються за складністю. Вони поділяють заплановану мету віддзеркалення функції людського мозку для вирішення складних проблем або завдань. Структура кожного типу штучної нейронної мережі певним чином відображає нейрони та синапси. Однак вони відрізняються за складністю, варіантами використання та структурою. Відмінності також включають спосіб моделювання штучних нейронів у кожному типі штучної нейронної мережі та зв'язки між кожним вузлом. Інші відмінності включають те, як дані можуть проходити через штучну нейронну мережу та щільність вузлів.

5 прикладів різних типів штучних нейронних мереж включають:

1. Штучні нейронні мережі прямого зв'язку
2. Нейронні мережі персептрон і багатошаровий персептрон
3. Радіальна основа функції штучних нейронних мереж
4. Рекурентні нейронні мережі
5. Модульні нейронні мережі
6. Штучні нейронні мережі прямого зв'язку

Як впливає з назви, штучна нейронна мережа Feedforward — це коли дані переміщуються в одному напрямку між вхідним і вихідним вузлами. Дані переміщуються вперед через шари вузлів і не повертаються назад через ті самі шари. Хоча може бути багато різних рівнів з багатьма різними вузлами, односторонній рух даних робить нейронні мережі Feedforward відносно простими. Моделі штучної нейронної мережі з упередженим зв'язком в основному використовуються для спрощених задач класифікації. Моделі виходять за рамки традиційної моделі машинного навчання, але не відповідають рівню абстракції, який є в моделі глибокого навчання.

Нейронні мережі персептрон і багатошаровий персептрон

Персептрон — одна з найперших і найпростіших моделей нейрона. Модель Perceptron — це двійковий класифікатор, який розділяє дані на дві різні класифікації. Як лінійна модель це один із найпростіших прикладів типу штучної нейронної мережі.

Багатошарові штучні нейронні мережі Perceptron додають складності та щільності, завдяки можливості для багатьох прихованих шарів між вхідним і вихідним шарами. Кожен окремий вузол на певному рівні з'єднаний з кожним вузлом на наступному рівні. Це означає, що багатошарові моделі Perceptron є повністю підключеними мережами, і їх можна використовувати для глибокого навчання.

Вони використовуються для більш складних проблем і завдань, таких як складна класифікація або розпізнавання голосу. Через глибину та складність моделі обробка та обслуговування моделі можуть потребувати ресурсів і часу.

Нейронні мережі радіальної базисної функції зазвичай мають вхідний рівень, рівень із вузлами радіальної базисної функції з різними параметрами та вихідний рівень. Моделі можна використовувати для класифікації, регресії для часових рядів і для керування системами. Функції радіального базису обчислюють абсолютне значення між центром і заданою точкою. У випадку класифікації радіальна базисна функція обчислює відстань між вхідними даними та вивченою класифікацією. Якщо введення найближче до певного тегу, воно класифікується як таке.

Загальне використання нейронних мереж з радіальною базисною функцією — це керування системою, наприклад системи, які контролюють відновлення живлення після відключення електроенергії. Штучна нейронна мережа може зрозуміти пріоритетний порядок відновлення живлення, віддаючи пріоритет ремонту найбільшій кількості людей або основним службам.

Рекурентні нейронні мережі

Рекурентні нейронні мережі є потужними інструментами, коли модель розроблена для обробки послідовних даних. Модель буде переміщати дані вперед і повертати їх назад до попередніх кроків у штучній нейронній мережі, щоб

найкраще виконати завдання та покращити прогнози. Шари між вхідним і вихідним рівнями повторюються, оскільки відповідна інформація повертається назад і зберігається. Пам'ять виходів із рівня повертається до входу, де вона зберігається, щоб покращити процес для наступного входу.

Потік даних подібний до штучних нейронних мереж Feedforward, але кожен вузол зберігатиме інформацію, необхідну для вдосконалення кожного кроку. Завдяки цьому моделі можуть краще зрозуміти контекст вхідних даних і уточнити прогноз вихідних даних. Наприклад, система інтелектуального тексту може використовувати пам'ять попереднього слова в рядку слів, щоб краще передбачити результат наступного слова. Повторювана штучна нейронна мережа краще підійде для розуміння настрою цілого речення порівняно з більш традиційними моделями машинного навчання.

Рекурентні нейронні мережі також використовуються в моделях послідовності, які використовуються для обробки природної мови. У цих моделях використовуються дві рекурентні нейронні мережі, які складаються з одночасного кодера та декодера. Ці моделі використовуються для реактивних чат-ботів, мови перекладу або для узагальнення документів.

Модульні нейронні мережі

Модульна штучна нейронна мережа складається з ряду мереж або компонентів, які працюють разом (хоча і незалежно) для досягнення завдання. Таким чином, складне завдання можна розбити на більш дрібні компоненти. У разі застосування до обробки даних або обчислювального процесу швидкість обробки буде збільшена, оскільки менші компоненти можуть працювати в тандемі.

Кожен компонент мережі виконує інше підзавдання, яке в поєднанні завершує загальні завдання та результат. Цей тип штучної нейронної мережі є вигідним, оскільки він може зробити складні процеси ефективнішими та може застосовуватися до низки середовищ.

Хоча існує величезний потенціал для використання штучних нейронних мереж у машинному навчанні, цей підхід пов'язаний із деякими труднощами. Моделі складні, і може бути важко пояснити аргументацію рішення, що в багатьох

випадках є операцією чорної скриньки. Це робить питання пояснюваності серйозним викликом і розглядом.

Для всіх типів моделей машинного навчання точність остаточної моделі значною мірою залежить від кількості та якості доступних навчальних даних. Модель, побудована за допомогою штучної нейронної мережі, потребує навіть більше даних і ресурсів для навчання, ніж традиційна модель машинного навчання. Це означає мільйони точок даних на відміну від сотень тисяч, необхідних для традиційної моделі машинного навчання.

Найскладніші штучні нейронні мережі часто називають глибокими нейронними мережами, посиляючись на багаторівневу мережеву архітектуру. Моделі глибокого навчання зазвичай навчаються з використанням позначених навчальних даних, тобто даних із визначеним входом і виходом. Це відоме як кероване машинне навчання, на відміну від неконтрольованого машинного навчання, яке використовує нерозмічені необроблені навчальні дані. Модель вивчатиме особливості та закономірності в позначених навчальних даних і навчатиметься виконувати задумане завдання за допомогою прикладів у навчальних даних. Штучні нейронні мережі потребують величезної кількості даних для навчання, більше ніж для більш традиційних алгоритмів машинного навчання. Це в сфері великих даних, тому може знадобитися багато мільйонів точок даних.

Потреба в такому великому масиві маркованих якісних даних є обмежуючим фактором для можливості розробки моделей штучних нейронних мереж. Тому організації обмежені тими, хто має доступ до необхідних великих даних. Найпотужніші моделі штучних нейронних мереж мають складну багаторівневу архітектуру. Ці моделі потребують величезної кількості ресурсів і потужності для обробки наборів даних. Для цього потрібні потужні, ресурсомісткі графічні процесори та системна архітектура. Знову ж таки, рівень необхідних ресурсів є обмежуючим фактором і проблемою для організацій.

Метод трансферного навчання часто використовується для зниження ресурсомісткості. У цьому процесі наявні знання з інших моделей і існуючих штучних нейронних мереж можуть бути передані або адаптовані при розробці

нової моделі. Це спрощує розробку, оскільки моделі не створюються щоразу з нуля, а можуть бути створені з елементів існуючих моделей.

1.2.5 Моделі на базі класифікаційно-регресійних дерев

Деревоподібні моделі використовують низку правил «якщо-тоді» для створення прогнозів на основі одного або кількох дерев рішень. Усі моделі на основі дерева можна використовувати або для регресії (прогнозування числових значень), або для класифікації (прогнозування категоріальних значень). Найчастіше виділяють три типи моделей на основі дерева:

1. Моделі дерева рішень, які є основою всіх моделей на основі дерева.
2. Моделі випадкового лісу, метод «ансамбля», який будує багато дерев рішень паралельно.
3. Моделі посилення градієнта, метод «ансамбля», який будує багато дерев рішень послідовно.

Модель дерева рішень можна використовувати для візуального представлення «рішень» або правил «якщо-тоді», які використовуються для створення прогнозів. Ось приклад дуже простої моделі дерева рішень (Рисунок 1.3):

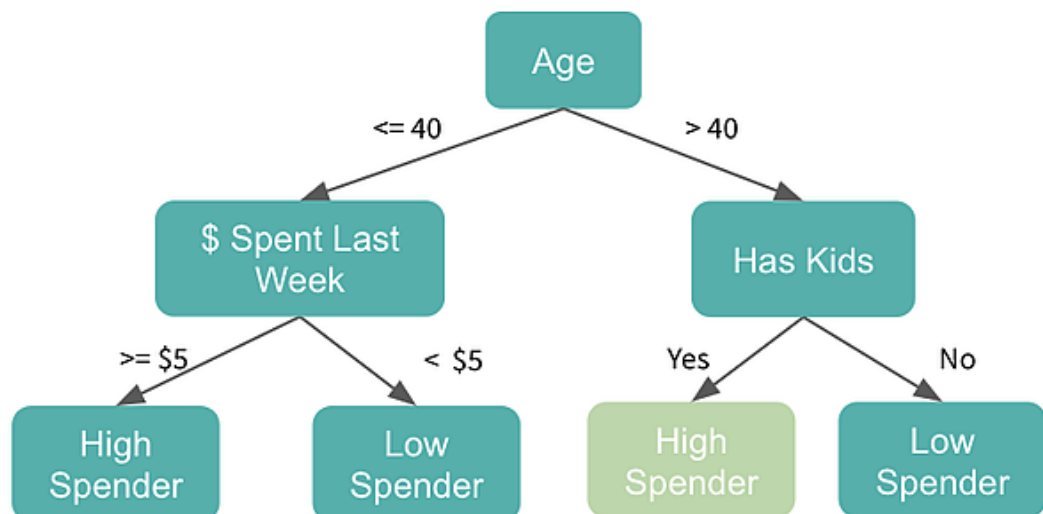


Рисунок 1.3 – Приклад моделі дерева рішень

Ми переглянемо кожне запитання «так» чи «ні» або **вузол рішень** у дереві та будемо рухатися вниз по дереву відповідно, доки не досягнемо остаточних прогнозів. Наше перше запитання, яке називається нашим **кореневим вузлом**, полягає в тому, чи Джорджу більше 40, і оскільки йому так, ми перейдемо до вузла «Має дітей». Оскільки відповідь ствердна, ми передбачимо, що цього тижня він буде багато витратити на Willy Wonka Candy.

Ще одне зауваження: тут ми намагаємося передбачити, чи буде Джордж витратити багато коштів, тому це приклад класифікаційного дерева, але ми можемо легко перетворити це на дерево регресії, передбачивши фактичні витрати Джорджа в доларах. Процес залишиться тим самим, але кінцеві вузли будуть числовими прогнозами, а не категоричними. По суті, є два ключових компоненти для побудови моделі дерева рішень: визначення того, які функції потрібно розділити, а потім рішення, коли припинити розділення.

Визначаючи, на які функції потрібно розділити, мета полягає в тому, щоб вибрати функцію, яка створить найбільш однорідні кінцеві набори даних. Найпростіший і найчастіше використовуваний метод зробити це шляхом мінімізації **ентропії**, вимірювання випадковості в наборі даних, і максимізації **отримання інформації**.

Хоча скорочення є хорошим методом покращення ефективності прогнозування моделі дерева рішень, одна модель дерева рішень зазвичай не дає надійних прогнозів сама по собі. Щоб покращити прогностичну силу нашої моделі, ми можемо побудувати багато дерев і поєднати передбачення, що називається **ансамблем**. Об'єднання фактично відноситься до будь-якої комбінації моделей, але найчастіше використовується для позначення моделей на основі дерева.

Ідея полягає в тому, щоб багато слабких припущень об'єдналися, щоб створити одне вагоме припущення. Ви можете подумати, що ансамбль запитує аудиторію «*Хто хоче стати мільйонером?*» Якщо запитання справді складне, учасник може віддати перевагу об'єднанню багатьох припущень, а не лише власним припущенням.

Щоб глибше розібратися в цій метафорі, конкурентом буде одна модель дерева рішень. Одне окреме дерево може не бути чудовим прогнозом, але якщо ми побудуємо багато дерев і об'єднаємо всі передбачення, ми отримаємо досить хорошу модель! Два з найпопулярніших ансамблевих алгоритмів — випадковий ліс і градієнтне посилення, які є досить потужними та зазвичай використовуються для розширених програм машинного навчання.

Моделі випадкових лісів просувають концепцію Пакетування на крок далі. Окрім побудови багатьох дерев із вибіркового набору даних, кожен вузол може розділятися лише на випадковий вибір функцій моделі.

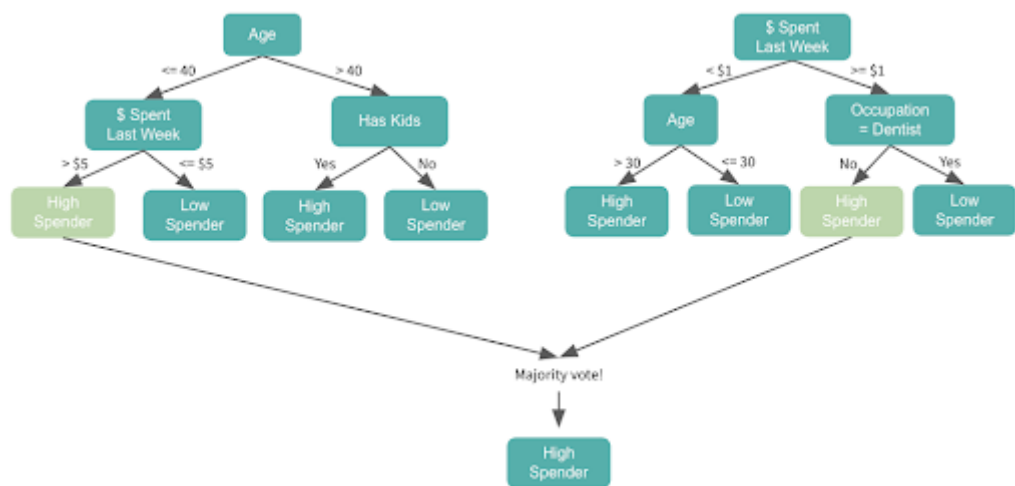


Рисунок 1.4 – Схема моделі випадкових лісів

Наприклад, уявіть, що кожен вузол може відокремлюватися від різних випадкових трьох функцій із нашого набору функцій. Дивлячись на наведене вище, ви можете помітити, що два дерева починаються з різними характеристиками — перше починається з віку, а друге — з витрачених доларів. Це тому, що, незважаючи на те, що вік може бути найважливішою характеристикою в наборі даних, він не був обраний у групі з трьох ознак для другого дерева, тому для початку модель мала використовувати наступну найважливішу характеристику, витрачені долари.

Кожен наступний вузол також буде розділено на випадковий вибір із трьох функцій. Припустімо, що наступна група функцій у наборі даних «витрачено

менше ніж 1 долар минулого тижня» включала вік, і цього разу порогове значення 30 років призвело до найвищого приросту інформації серед усіх функцій, вік більше або менше 30 років буде наступним розкол.

Ми побудуємо наші два дерева окремо й отримаємо більшість голосів. Зауважте, що якби це була задача регресії, ми б отримали середнє значення.

Ось короткий підсумок випадкових лісів:

Переваги:

1. Добре справляється зі складними, нелінійними зв'язками
2. Добре обробляти набори даних з високою розмірністю (багато функцій).
3. Добре обробляйте відсутні дані
4. Вони потужні та точні
5. Їх можна швидко навчити. Оскільки дерева не залежать одне від одного, їх можна навчати паралельно.

Недоліки:

1. Менш точні для проблем регресії, оскільки вони мають тенденцію до перепідбору

Підвищення — це метод ансамблевого дерева, який будує послідовні невеликі дерева — часто лише один вузол — з кожним деревом, зосередженим на виправленні сумарної помилки попереднього дерева. Отже, ми розділимо наше перше дерево за найбільш прогнозованою ознакою, а потім оновимо вагові коефіцієнти, щоб переконатися, що наступні розбиття дерева за будь-якою ознакою дозволять правильно класифікувати точки даних, які були неправильно класифіковані у початковому дереві. Наступне дерево буде зосереджено на правильній класифікації помилок із цього дерева тощо. Остаточний прогноз — це зважена сума всіх індивідуальних прогнозів.

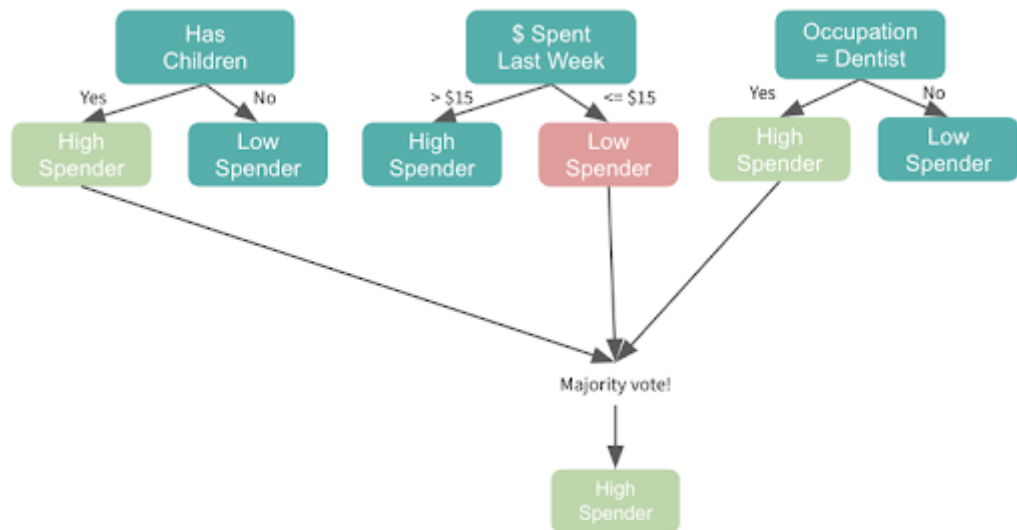


Рисунок 1.5 – Схема методу підвищення

Градiєнтне посилення є найпопулярнішим розширенням посилення, яке використовує алгоритм градiєнтного спуску для оптимізації.

Ось короткий підсумок посилення градiєнта:

Переваги:

1. Вони потужні та точні, у багатьох випадках навіть більше, ніж випадковий ліс
2. Добре справляється зі складними, нелінійними зв'язками
3. Вони добре справляються з незбалансованими даними

Недоліки:

1. Повільніше навчатися, оскільки дерева потрібно будувати послідовно
2. Схильність до переобладнання, якщо дані мають шум
3. Важче налаштувати гіперпараметри

Деревоподібні моделі дуже популярні в машинному навчанні. Модель дерева рішень, основу моделей на основі дерева, досить проста для інтерпретації, але, як правило, є слабким прогнозом. Ансамблеві моделі можна використовувати для створення точніших прогнозів на основі багатьох дерев, причому випадковий ліс і посилення градiєнта є двома найпопулярнішими. Усі моделі, засновані на дереві,

можна використовувати для регресії або класифікації та можуть досить добре обробляти нелінійні зв'язки.

1.2.6 Моделі на основі радіально-базисних функцій

Мережа радіальних базисних функцій є штучною нейронною мережею мережа, яка використовує радіальні базисні функції як активацію функції. Це лінійна комбінація радіальних базисних функцій. Вони використовуються в апроксимації функцій, часових рядів прогнозування та контроль.

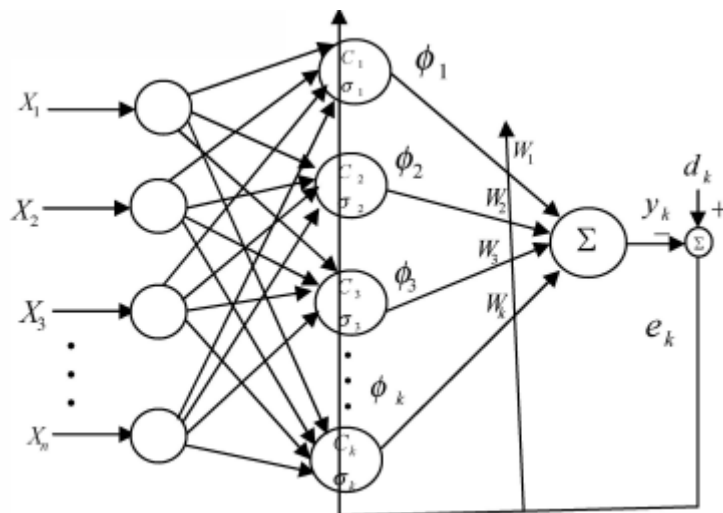


Рисунок 1.6 – Схема роботи радіально-базисних функцій

Мережі радіальної базисної функції (RBF) зазвичай мають три шари:

1-й Вхідний шар: вхідний шаблон входить у цей шар і є піддається функції прямої передачі, тобто виходу нода дорівнює його входу.

2-й Прихований шар: вузли прихованого шару, які називаються кластером центри радіально симетричні, і ці вузли включають такі важливі параметри:

- Центральний вектор c_k : цей вектор, який зазвичай зберігається як вагові коефіцієнти від вхідного до прихованого шару є а багатовимірний вектор у вхідному просторі.

- Евклідова відстань: це міра відстані, яка визначає, наскільки далеко знаходиться вхідний вектор X з елементами X_i від центрального вектора C_k з

елементами c_{ik} ($i=1$ до N , $k=1$ до L , де N і L є кількістю вузлів у вхідний і прихований шари відповідно).

$$I_k = ||x_i - c_k|| = \sqrt{(E(x_i - c_{ik}))^2} \quad (1.11)$$

• Базисна функція: передаточна функція, наприклад Гаусса функція, перетворює евклідове підсумовування I_k , щоб отримати вихід для кожного вузла. Ця функція теж має під собою основу ширина σ_k і його вихід Φ_k надається:

$$\varphi_k = \exp(-I_k^2 / \sigma_k^2) \quad (1.12)$$

Вихідний шар: отримує вихідні дані цього шару через зважену суму результатів прихованого шару використовуючи:

$$y_k = \varphi_k \times w_k^T \quad (1.13)$$

Такі параметри мережі, як вага, центр і дисперсія оновлюється за допомогою (4) - (8).

$$w(n+1) = w(n) + \mu_w e(n) \varphi(n) \quad (1.14)$$

$$c_k(n+1) = c_k(n) + \mu_c \frac{e(n)w_k(n)}{\sigma_k^2(n)} \Phi[x(n), c_k(n), \sigma_k |x(n) - c_k(n)|] \quad (1.15)$$

$$\sigma_k(n+1) = \sigma_k(n) + \mu_\sigma \frac{e(n)w_k(n)}{\sigma_k^3(n)} \Phi[x(n), c_k(n), \sigma_k] |x(n) - c_k(n)|^2 \quad (1.16)$$

$$\varphi(n) = [\Phi[x(n), c_1, \sigma_1], \Phi[x(n), c_2, \sigma_2], \dots, \Phi[x(n), c_k, \sigma_k]]^T \quad (1.17)$$

$$e(n) = y(n) - y_d(n) \quad (1.18)$$

де $y(n)$ є результатом моделі, а $y_d(n)$ є бажаним вихід.

1.2.7 Інші моделі та методи прогнозування

Крім класів моделей прогнозування, розглянутих вище, існують менш поширені моделі та методи прогнозування. Головним недоліком моделей та

методів, згаданих у цьому розділі, є недостатня методологічна база, тобто недостатньо докладне опис можливостей як моделей, так і шляхів визначення їх параметрів. Крім того, у відкритому доступі можна знайти лише невелике кількість статей, присвячених застосуванню даних методів.

Support Vector Machine (SVM) — це контрольований алгоритм машинного навчання, який використовується як для класифікації, так і для регресії. Мета алгоритму SVM — знайти гіперплощину в N -вимірному просторі, яка чітко класифікує точки даних. Розмірність гіперплощини залежить від кількості ознак. Якщо кількість вхідних елементів дорівнює двом, то гіперплощина є просто лінією. Якщо кількість вхідних елементів дорівнює трьом, то гіперплощина стає двовимірною площиною. Стає важко уявити, коли кількість ознак перевищує три.

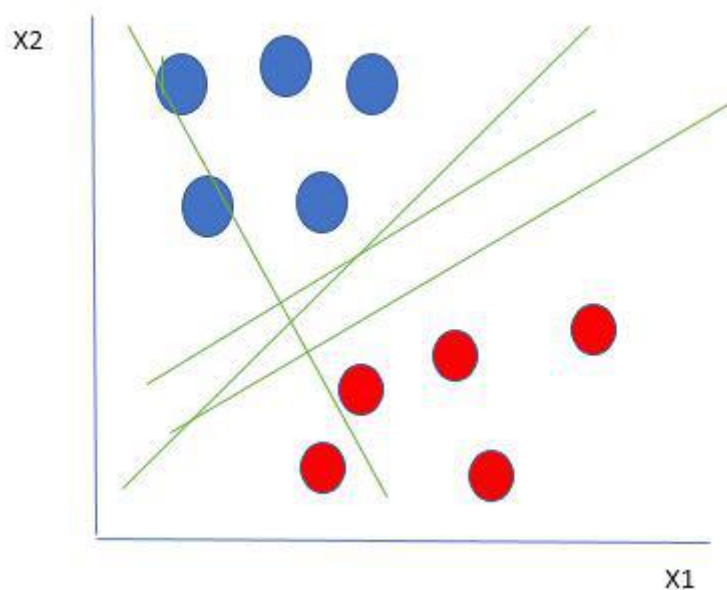


Рисунок 1.7 – Схема роботи SVM

Генетичні алгоритми (GA) — адаптивні евристичні алгоритми пошуку, які належать до більшої частини еволюційних алгоритмів. Генетичні алгоритми базуються на ідеях природного відбору та генетики. Це інтелектуальне використання випадкового пошуку з історичними даними для спрямування пошуку в область кращої продуктивності в просторі рішень. Вони зазвичай

використовуються для створення високоякісних рішень для проблем оптимізації та пошуку.

Генетичні алгоритми моделюють процес природного відбору, тобто ті види, які можуть адаптуватися до змін у навколишньому середовищі, здатні виживати, розмножуватися та переходити до наступного покоління. Простими словами, вони імітують «виживання найбільш пристосованих» серед індивідів послідовних поколінь для вирішення проблеми. Кожне покоління складається з популяції індивідів, і кожна індивідуум представляє точку в просторі пошуку та можливого рішення. Кожна особа представлена у вигляді рядка символів/цілих чисел/плаваючої величини/біт. Цей рядок є аналогом хромосоми.

1.3 Порівняння моделей прогнозування

У попередньому розділі наведено огляд існуючих моделей прогнозування. У цьому розділі розглянуто переваги і недоліки як описаних вище моделей, а й методів. Говорячи про переваги та недоліки моделей прогнозування, необхідно брати до уваги та відповідні їм методи

Регресійні моделі та методи.

Переваги регресійної моделі

1. Регресійні моделі легко зрозуміти, оскільки вони побудовані на основних статистичних принципах, таких як кореляція та похибка найменшого квадрата.
2. Результатом регресійних моделей є алгебраїчне рівняння, яке легко зрозуміти та використовувати для прогнозування.
3. Сила (або відповідність) регресійної моделі вимірюється за допомогою коефіцієнтів кореляції та інших пов'язаних статистичних параметрів, які добре зрозумілі.
4. Прогнозна потужність регресійних моделей збігається з іншими прогнозними моделями та іноді є кращою, ніж конкурентні моделі.
5. Регресійні моделі можуть включати всі змінні, які потрібно включити в модель.

6. Інструменти регресійного моделювання є поширеними. Майже всі пакети інтелектуального аналізу даних містять статистичні пакети, які включають інструменти регресії. Електронні таблиці MS Excel також можуть надавати можливості простого регресійного моделювання.

Недоліки регресійної моделі

1. Регресійні моделі не можуть працювати належним чином, якщо вхідні дані містять помилки (тобто дані низької якості). Якщо попередня обробка даних не виконана належним чином для видалення відсутніх значень або зайвих даних, викидів або незбалансованого розподілу даних, валідність регресійної моделі страждає.

2. Регресійні моделі чутливі до колінеарних проблем (тобто існує сильна лінійна кореляція між незалежними змінними). Якщо незалежні змінні сильно корельовані, то вони з'їдають прогностичну силу одна одної, і коефіцієнти регресії втрачуть свою жорсткість.

3. Зі збільшенням кількості змінних надійність регресійних моделей знижується. Регресійні моделі працюють краще, якщо у вас є невелика кількість змінних.

4. Регресійні моделі автоматично не враховують нелінійність. Користувачеві потрібно уявити, які додаткові умови можуть знадобитися додати до моделі регресії, щоб покращити її відповідність.

5. Регресійні моделі працюють із наборами даних, що містять числові значення, а не з категоріальними змінними. Існують способи роботи з категоріальними змінними шляхом створення кількох нових змінних зі значенням «так/ні».

Експоненціальне згладжування – це техніка, яка використовується для виявлення значних змін у даних шляхом розгляду останніх даних. Також відомий як усереднення, цей метод використовується для складання короткострокових прогнозів.

Список переваг експоненціального згладжування

1. Його легко вивчити та застосувати для методів експоненціального згладжування потрібні лише три частини даних. Поперше, йому потрібен прогноз на останній період часу. Подруге, потрібно фактичне значення для цього періоду часу. І потрете, йому потрібне значення константи згладжування, ваговий коефіцієнт, який відображає вагу, надану останнім значенням даних.

2. Виробляє точні прогнози.

Експоненціальний метод згладжування дає прогноз на один період вперед. Використовуючи техніку прогнозування трендів, можна генерувати прогнози на більше періодів вперед. Прогноз вважається точним, оскільки він враховує різницю між фактичними прогнозами та тим, що насправді сталося.

3. Це надає більше значення останнім спостереженням.

Дані спостережень – це сума двох або більше компонентів, один з яких є випадковою похибкою, тобто різницею між спостережуваним значенням і справжнім значенням. У техніці згладжування випадковою варіацією нехтують. Таким чином, набагато легше побачити основне явище.

Список недоліків експоненціального згладжування

1. Він створює прогнози, які відстають від фактичної тенденції. Відставання є побічним ефектом процесу згладжування. У назві цього методу є причина «згладжування», оскільки він нехтує підйомами та падіннями, пов'язаними з випадковою варіацією. Таким чином, побачивши це на графіку, ви побачите більш гладку лінію або криву. Але ігнорування випадкової варіації також дозволяє побачити основне явище, що допомагає при представленні даних і прогнозуванні майбутніх значень.

2. Він не впорається з тенденціями. Експоненціальне згладжування найкраще використовувати для прогнозів, які є короткостроковими та за відсутності сезонних або циклічних коливань. Як наслідок, прогнози є неточними, якщо наявні дані з циклічними або сезонними коливаннями. Таким чином, такий вид усереднення не працюватиме належним чином, якщо в ряді є тенденція.

Подібні методи є точними лише тоді, коли можна припустити розумну кількість безперервності між минулим і майбутнім. Таким чином, він найкраще

підходить для короткострокового прогнозування, оскільки припускає, що майбутні моделі та тенденції виглядатимуть як поточні моделі та тенденції. Хоча таке припущення може здатися розумним у короткостроковій перспективі, воно створює проблеми, чим далі просувається прогноз.

Штучні нейронні мережі – це найпростіше моделювання людського мозку де будівельними блоками є нейрони. У мозку людини близько 100 мільярдів нейрон. Кожен нейрон має точку з'єднання від 1000 до 100 000. У людському мозку інформація зберігається таким чином, щоб її розподіляли. Ми можемо одночасно витягувати з пам'яті більше одного фрагмента цієї інформації, якщо це необхідно. Ми не помилимося, коли скажемо, що мозок людини складається з тисяч дуже і дуже потужних паралельних процесорів. У багат шарових штучних нейронних мережах також є нейрони розміщені подібно до людського мозку. Кожен нейрон з'єднаний з іншими нейронами з певними коефіцієнтами. Під час навчання інформація розповсюджується на ці точки підключення, щоб мережа вивчалася.

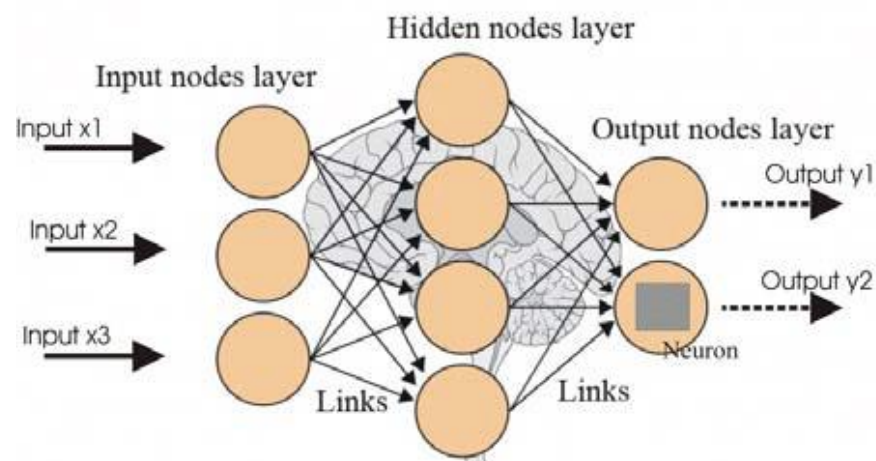


Рисунок 1.8 - Рівні штучної нейронної мережі

Як показано на рисунку 1, нейронна мережа складається з трьох рівнів: вхідного рівня, проміжного рівня та вихідного рівня. Показані тут прямокутники представляють нейрони, а стрілки — точки з'єднання. Набір даних, підготовлений для навчання на вхідному рівні, показується в мережі. Мережа призначає ваги

подій, які вивчає, точкам з'єднання на проміжному рівні. Не кожна точка має бути значенням, а деякі бали можуть дорівнювати нулю. Між шарами додається граничне значення, щоб нульові значення в точках з'єднання не стали нульовими.

Переваги штучних нейронних мереж:

1. Зберігання інформації у всій мережі: інформація, як у традиційному програмуванні, зберігається у всій мережі, а не в базі даних. Зникнення кількох фрагментів інформації в одному місці не перешкоджає функціонуванню мережі.
2. Наявність відмовостійкості: пошкодження однієї чи кількох елементів ШНМ не перешкоджає створенню вихідних даних. Ця функція робить мережу відмовостійкою.
3. Можливість паралельної обробки: штучні нейронні мережі мають числову потужність, яка може виконувати більше ніж одну роботу одночасно.

Недоліки штучних нейронних мереж

1. Залежність від апаратного забезпечення: Штучні нейронні мережі потребують процесорів з паралельною обчислювальною потужністю відповідно до їхньої структури.
2. Незрозуміла поведінка мережі: це найважливіша проблема ШНМ. Коли мережа створює тестове рішення, воно не дає підказки про те, чому і як.
3. Визначення належної структури мережі: Не існує спеціального правила для визначення структури штучної нейронної мережі. Відповідна мережева структура досягається шляхом досвіду та методом проб і помилок.

Дерево рішень(Decision Tree) — дуже популярний алгоритм машинного навчання. Дерево рішень вирішує проблему машинного навчання шляхом перетворення даних у представлення дерева. Кожен внутрішній вузол подання дерева позначає атрибут, а кожен листовий вузол позначає мітку класу.

Алгоритм дерева рішень може бути використаний для вирішення проблем як регресії, так і класифікації.

Переваги:

1. У порівнянні з іншими алгоритмами дерева рішень вимагають менше зусиль для підготовки даних під час попередньої обробки.
2. Дерево рішень не потребує нормалізації даних.
3. Дерево рішень також не потребує масштабування даних.
4. Відсутні значення в даних також НЕ впливають значною мірою на процес побудови дерева рішень.
5. Модель дерева рішень дуже інтуїтивно зрозуміла, і її легко пояснити технічним командам, а також зацікавленим сторонам.

Недоліки:

1. Невелика зміна в даних може спричинити значні зміни в структурі дерева рішень, викликаючи нестабільність.
2. Для дерева рішень іноді обчислення можуть бути набагато складнішими порівняно з іншими алгоритмами.
3. Дерево рішень часто вимагає більше часу для навчання моделі.
4. Навчання дереву рішень є відносно дорогим, оскільки складність і час вимагають більше.
5. Алгоритм дерева рішень не підходить для застосування регресії та прогнозування безперервних значень.

Таблиця 1.1 Порівняння моделей та методів прогнозування

Модель та метод	Переваги	Недоліки
Регресійні моделі та методи	Простота, прозорість моделювання, гнучкість, однаковість аналізу та проектування	складність визначення функціональною залежності; трудомісткість знаходження коефіцієнтів залежності; відсутність можливості моделювання

Таблиця 1.1 Порівняння моделей та методів прогнозування

Авторегресійні моделі та методи	простота, прозорість моделювання; однаковість аналізу та проектування; безліч прикладів застосування	трудомісткість та ресурсомісткість ідентифікації моделей; неможливість моделювання нелінійностей; низька адаптивність
Моделі та методи експоненційного згладжування	простота моделювання; однаковість аналізу та проектування	Недостатня гнучкість; вузька сфера використання моделей
Нейромережеві моделі та методи	нелінійність моделей; масштабованість, висока адаптивність; однаковість аналізу та проектування; безліч прикладів застосування	відсутність прозорості; складність вибору архітектури; жорсткі вимоги до навчальної вибірки; складність вибору алгоритму навчання; ресурсомісткість процесу навчання
Моделі на базі дерев рішень	масштабованість; швидкість та простота процесу навчання; можливість враховувати категоріальні змінні	неоднозначність алгоритму побудови дерева; складність питання зупинки

1.4 Комбіновані методи

Комбіновані методи – це методи машинного навчання, які можуть забезпечити кращу точність шляхом поєднання кількох базових моделей замість використання однієї моделі. Використовуючи методи ансамблю, модель машинного навчання може давати значно точніші результати.

Зазвичай ансамблеві методи є ідеальними, коли вам доводиться мати справу з регресією та класифікацією, оскільки вони можуть підвищити точність вашої моделі та зменшити її зміщення.

Види Комбінованих Методів(Рис. 1.9). Найпопулярнішими методами комбінування є Bagging і Boosting, які належать до парадигм під назвами Parallel і Sequential відповідно.

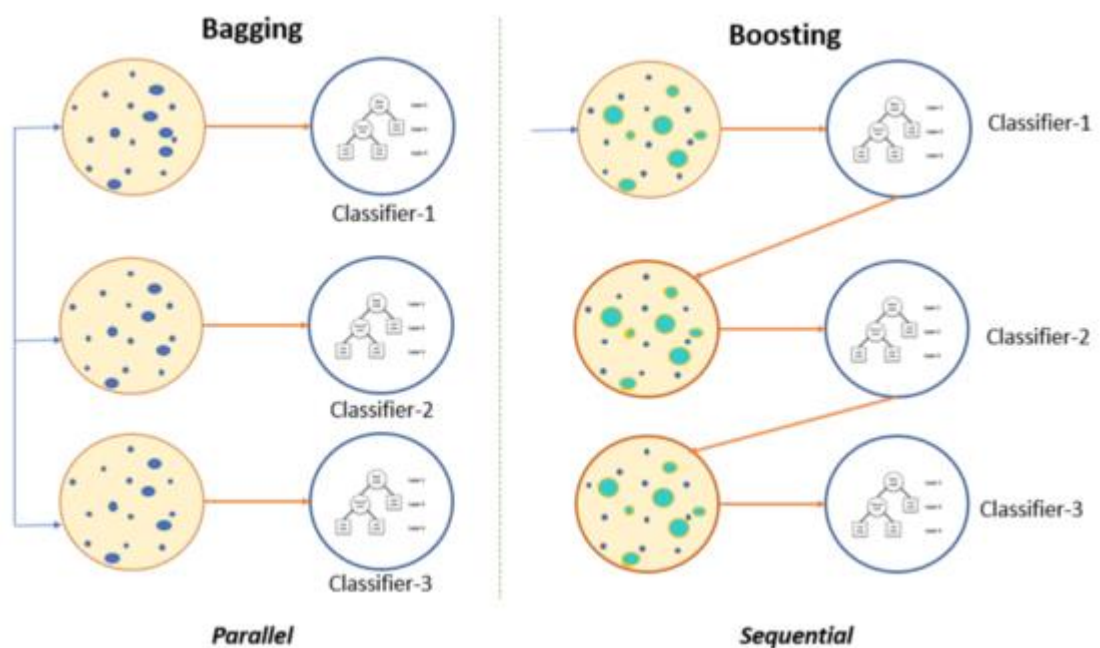


Рисунок 1.9 – Схеми комбінованих методів

Об'єднання в пакети (Bagging)(Рис. 1.10). Об'єднання в пакети, також відоме як Bootstrap Aggregating, відноситься до категорії методу паралельного ансамблю, це спосіб зменшити дисперсію моделі машинного навчання шляхом генерації додаткових даних на етапі навчання.

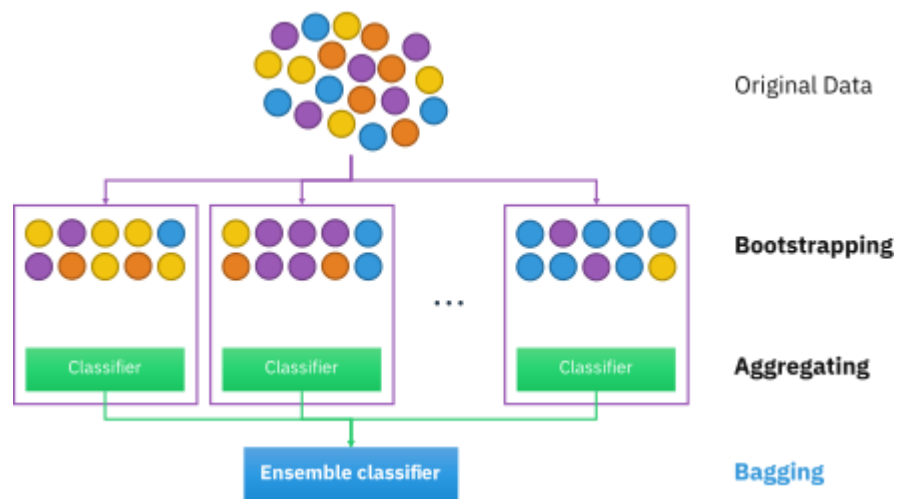


Рисунок 1.10 – Схема методу «Об’єднання в пакети»

Як згадувалося вище, пакетування також відоме як Bootstrap Aggregating, тому Bootstrapping і агрегація є двома окремими методами.

Бутстрапінг – це техніка вибірки, яка створює багато змодельованих вибірок з одного набору даних, тоді як агрегація включає в себе всі можливі результати прогнозу та рандомізує результат. Однак, якщо техніка агрегування пропущена, прогнози будуть неточними, оскільки всі результати вашої моделі не будуть враховані.

Значне зменшення дисперсії, що призведе до усунення переобладнання, отже, ми можемо мати набагато кращу точність прогнозної моделі за допомогою Bagging.

Підвищення — це техніка методу послідовного ансамблю, яка вчиться на попередніх помилках прогнозування, щоб робити кращі прогнози в майбутньому (Рис. 1.11).

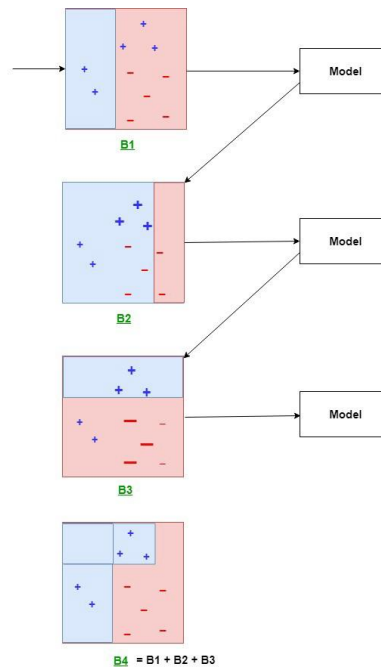


Рисунок 1.11 – Діаграма роботи методу підвищення

Основна ідея техніки Boosting полягає в тому, щоб об'єднати кілька слабких базових учнів для створення одного сильного учня, отже створивши чудову прогностичну модель машинного навчання.

Прискорення можна досягти за допомогою дуже відомого алгоритму, відомого як адаптивне прискорення (AdaBoost). Техніка AdaBoost полягає в тому, що ваги повторно призначаються кожному екземпляру, причому вищі ваги призначаються неправильно класифікованим екземплярам.

Підвищення поступово створює ансамбль шляхом навчання кожної моделі з тим самим набором даних, але де ваги екземплярів коригуються відповідно до помилки останнього передбачення, таким чином зменшуючи зміщення та дисперсію та забезпечуючи модель з підвищеною точністю.

2 МОДЕЛІ ЕКСТРАПОЛЯЦІЇ ТИМЧАСОВИХ РЯДІВ З ВИБІРКИ МАКСИМАЛЬНОГО ПОДОБИ

2.2 Модель без урахування зовнішніх факторів

2.1.1 Вибірки часового ряду

Нехай заданий часовий ряд $Z(t) = Z(1), Z(2), \dots, Z(T)$. Набір послідовних значень $Z_t^M = Z(t), Z(t+1), \dots, Z(t+M-1)$, що лежить всередині вихідного часового ряду, назвемо вибіркою довжини M з моментом початку звіту t ; $M \in \{1, 2, \dots, T\}, t \in \{1, 2, \dots, T-M+1\}$. Фактично вибіркою є шматочок тимчасового ряду, що має точку початку відліку та довжину. Дві вибірки однакової довжини, що належать одному тимчасовому ряду, позначимо через тимчасову затримку k : $Z_t^M = Z(t), \dots, Z(t+M-1)$ та $Z_{t-k}^M = Z(t-k), \dots, Z(t-k+M-1)$; $k \in \{1, 2, \dots, t-1\}$. На малюнках 2.1 - 2.4 показані вибірки різних часових рядів. У роботі [40] сформульовано підхід до моделювання тимчасових рядів за допомогою вибірок: «Pattern modelling refers to the process of describing time series as a series of patterns». (Моделювання тимчасових рядів при допомозі вибірок засноване на припущенні, що тимчасовий ряд є послідовністю вибірок.)

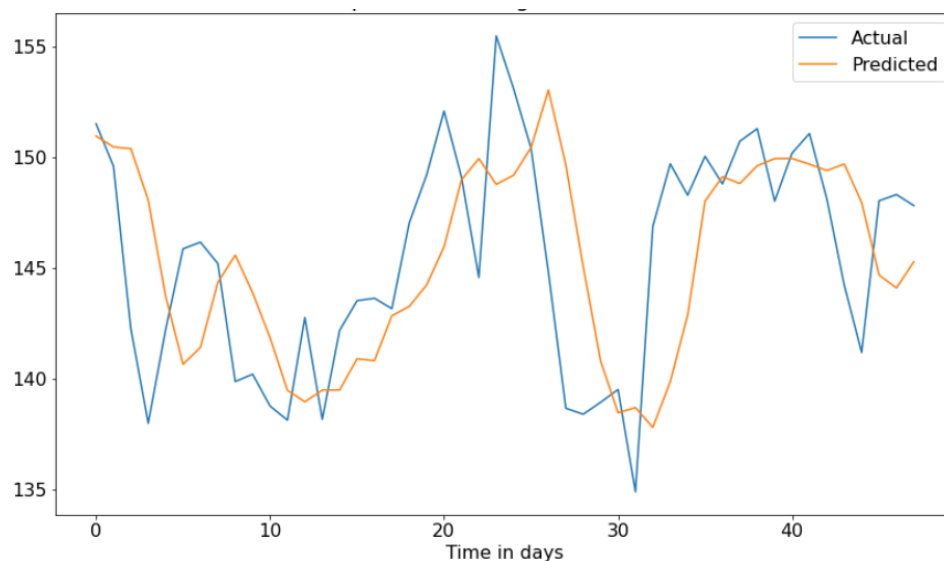


Рисунок 2.1 – Часові ряди акцій компанії Apple (15.10.2021-15.10.2022)

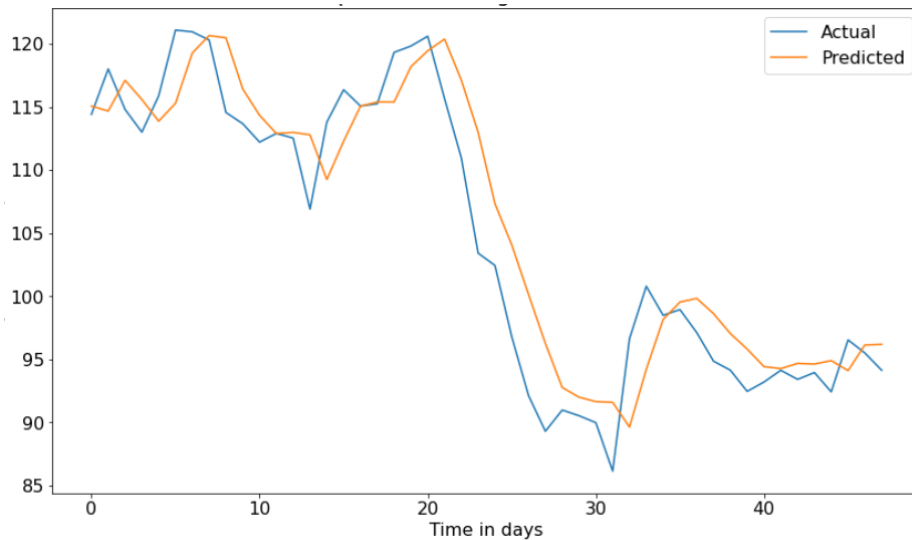


Рисунок 2.2 – Часові ряди компанії Amazon (15.10.2021-15.10.2022)

На малюнках 2.1 - 2.2 зображені вибірки часових рядів, які ілюструють властивість вибірок, сформульовану у роботі [42]: «Pieces of time series in the past might have resemblance to pieces in the future». (Фактичні вибірки тимчасового ряду можуть мати подібність до майбутніх вибірок). Вказана властивість вибірок буде використана для побудови моделі прогнозування. У цьому розділі спочатку розглядається завдання апроксимації однієї вибірки за допомогою іншої, а потім результати апроксимації застосовуються для побудови моделі прогнозування часового ряду.

2.1.2. Апроксимація вибірки

Для розрахунків перейдемо до векторного позначення вибірки $Z_t^M = (Z(t), Z(t+1), \dots, Z(t+M-1))^T$ та тимчасового ряду $Z_1^T = (Z(1), Z(2), \dots, Z(T))^T$. Тут і далі, говорячи про вибірки тимчасового ряду будемо використовувати позначення Z_t^M , говорячи про вектори, відповідних зазначеним вибіркам, писатимемо жирну Z_t^M .

Використовуючи властивість вибірок повторюватися, апроксимуємо вибірку Z_t^M за допомогою вибірки Z_{t-k}^M наступним чином:

$$Z_t^M = \alpha_1 Z_{t-k}^M + \alpha_0 I^M + E^M. \quad (2.1)$$

Тут α_1 і α_0 - коефіцієнти, I^M - одиничний вектор, E^M — вектор значення помилок апроксимації. Вираз (2.1) можна переписати як

$$Z_t^M = \alpha_1 Z_{t-k}^M + \alpha_0 I^M. \quad (2.2)$$

У формулі (2.2) Z_t^M - апроксимовані значення вибірки Z_t^M . Тут і далі, говорячи про обчислювані (модельні) значення вибірки Z_t^M використовуємо позначення Z_t^M (з кришечкою). Постановка задачі апроксимації вибірки. Нехай дана функціональна залежність вибірок (2.2). Необхідно визначити такі значення α_1 і α_0 , щоб квадрат відхилень модельних значень вибірки Z_t^M від фактичних Z_t^M був мінімальний

$$\sigma^2 = \sum_{i=0}^{M-1} (Z(t+i) - Z(t+i))^2 \rightarrow \min \quad (2.3)$$

Після знаходження коефіцієнтів α_1 і α_0 необхідно оцінити вектор помилок E^M .

Рішення. Нехай дана лінійна залежність (2.1), тоді функція помилки апроксимації S_k^M для вибірок Z_t^M та Z_{t-k}^M із затримкою k має вигляд

$$S_k^M(\alpha_1, \alpha_0) = \sum_{i=0}^{M-1} \sigma_i^2 = \sum_{i=0}^{M-1} (Z(t+i) - \alpha_1 Z(t-k+i) - \alpha_0)^2. \quad (2.4)$$

Функція S_k^M називається функцією суми квадратів (sum of squares function) [19]. Завдання полягає в тому, щоб підібрати такі значення α_1 і α_0 , щоб при підстановці їх (2.4) було отримано мінімальне можливе значення $S_k^M(\alpha_1, \alpha_0)$. Рисунок 2.5 ілюструє визначення функції суми квадратів.

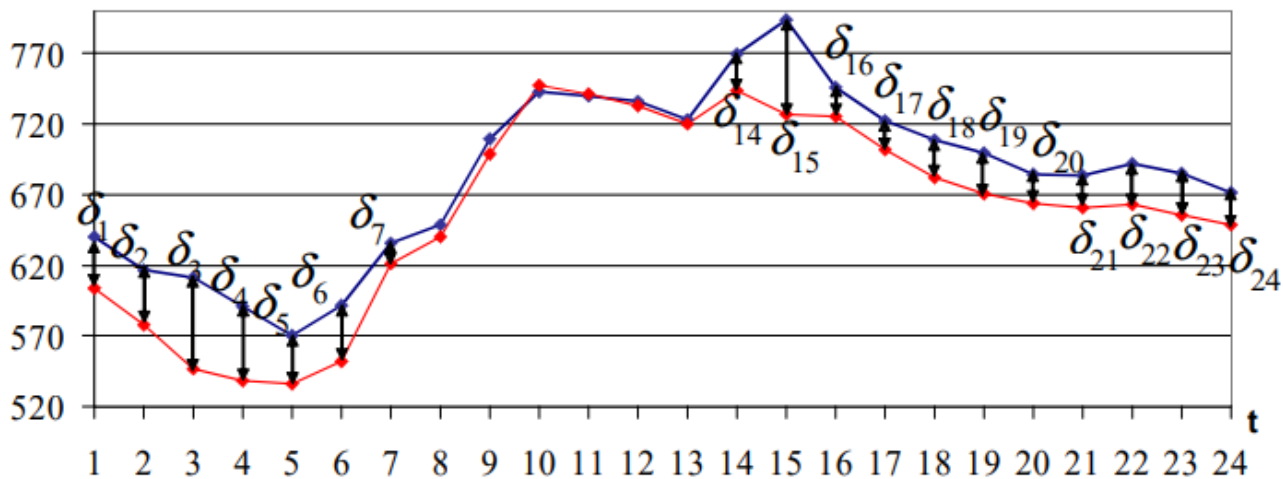


Рисунок 2.3 - Визначення функції суми квадратів S_k^M за виразом (2.4)

Розв'яжемо задачу за допомогою методу найменших квадратів, докладно описаного у роботі [19]. Після низки перетворень завдання зводиться до розв'язання матричного рівняння

$$Z_X * A = Z_Y, \text{ де } A = \begin{bmatrix} \alpha_1 \\ \alpha_0 \end{bmatrix}. \quad (2.5)$$

Вирішивши матричне рівняння (2.5), визначаємо значення коефіцієнтів α_1 і α_0 , що відповідають мінімуму функції $S_k^M(\alpha_1, \alpha_0)$. Рівняння (2.5) може бути вирішено будь-яким відомим методом. Вихідні матриці є квадратними і рішення може бути знайдено, наприклад, за допомогою зворотної матриці

$$Z_X^{-1} Z_X A = Z_X^{-1} Z_Y, \quad (2.6)$$

$$A = Z_X^{-1} Z_Y. \quad (2.7)$$

Оцінка помилки апроксимації. Помилка визначається за формулою

$$E^M = Z_t^M - Zx_t^M = Z_t^M - (\alpha_1 Z_{t-k}^M + \alpha_0 I^M). \quad (2.8)$$

В даний час точність моделювання часових рядів E^M прийнято оцінювати за допомогою двох показників [28]:

- Середня абсолютна помилка (mean absolute error, MAE)

$$MAE = \frac{1}{M} \sum_{i=t}^{t+M-1} |Z(i) - Zx(i)|. \quad (2.9)$$

- Середня абсолютна помилка у відсотках, (mean absolute percentage error, MAPE)

$$MAPE = \frac{1}{M} \sum_{i=t}^{t+M-1} \frac{|Z(i) - Zx(i)|}{Z(i)} * 100\%. \quad (2.10)$$

Тут і далі, говорячи про точність моделювання часових рядів (апроксимації, прогнозування) будемо застосовувати значення показників MAE (2.9) та MAPE (2.10). У цьому розділі було розглянуто апроксимацію однієї вибірки тимчасового ряду за допомогою іншої, що лежить на осі часу на k звітів ліворуч, тобто раніше. Ця властивість представлення нових вибірок тимчасового ряду за допомогою старих буде використано для визначення моделі екстраполяції.

2.1.3. Подібність вибірок

Властивість двох вибірок, що полягає в тому, що одна вибірка може бути виражена через іншу за допомогою лінійної залежності (2.1), назвемо подобою двох вибірок. Покажемо, що для загального випадку лінійної регресії мінімум помилки регресії відповідає максимуму лінійної кореляції Пірсона. Нехай дана модель

$$Yx = \alpha_1 X + \alpha_0 I. \quad (2.11)$$

Тоді функція суми квадратів визначається як різниця модельних $Yx(i)$ та фактичних значень $Y(i)$ деяких спостережень [19]

$$S_{reg} = \sum_{i=1}^M (Yx(i) - Y(i))^2 \quad (2.12)$$

Позначимо \bar{Y} - середнє значення модельних та фактичних спостережень, які за якістю лінійної регресії рівні

$$\bar{Y} = \frac{1}{M} \sum_{i=1}^M Y(i) = \frac{1}{M} \sum_{i=1}^M \hat{Y}(i) \quad (2.13)$$

З книги [19] відомо, що сума квадратів відхилень досліджуваних спостережень $Y(i)$ від середнього значення \bar{Y} складається із суми квадратів відхилень модельних значень $\hat{Y}(i)$ від \bar{Y} та суми квадратів помилок регресії, визначеної у виразі (2.12). Таким чином має місце співвідношення

$$\sum_{i=1}^M (Y(i) - \bar{Y})^2 = \sum_{i=1}^M (\hat{Y}(i) - \bar{Y})^2 + \sum_{i=1}^M (Y(i) - \hat{Y}(i))^2. \quad (2.14)$$

Виразимо помилку регресії $\sum_{i=1}^M (Y(i) - \hat{Y}(i))^2$ і отримаємо

$$\begin{aligned} \sum_{i=1}^M (Y(i) - \hat{Y}(i))^2 &= \sum_{i=1}^M (Y(i) - \bar{Y}(i))^2 - \sum_{i=1}^M (\hat{Y}(i) - \bar{Y}(i))^2 \rightarrow S_{reg} = \\ \sum_{i=1}^M (Y(i) - \hat{Y}(i))^2 &= \sum_{i=1}^M (Y(i) - \bar{Y}(i))^2 * \left(1 - \frac{\sum_{i=1}^M (Y(i) - \bar{Y}(i))^2}{\sum_{i=1}^M (Y(i) - \hat{Y}(i))^2} \right). \end{aligned} \quad (2.15)$$

Величина

$$R^2 = \frac{\sum_{i=1}^M (Y(i) - \bar{Y}(i))^2}{\sum_{i=1}^M (Y(i) - \hat{Y}(i))^2} \in [0,1] \quad (2.16)$$

називається квадратом множинного коефіцієнта кореляції. Іноді даний коефіцієнт називаються коефіцієнтом детермінації

Варто звернути увагу, що при $Y(i)=const$ для $i \in \{1,2,3, \dots\}$ знаменник R^2 звертається в нуль. Проте за якістю регресії [19], для такого випадку модельні значення також будуть постійними $\hat{Y}=const$ для $i \in \{1,2,3, \dots\}$ та $R^2 = 1$. Відомо, що на практиці такий випадок неможливий, зв'язку з тим, що значення $Y(i)$, як правило, є результатами вимірів. Перетворивши (2.15) отримаємо вираз

$$S_{reg} = \sum_{i=1}^M (Y(i) - \hat{Y}(i))^2 * (1 - R^2). \quad (2.17)$$

При цьому сума квадратів відхилень досліджуваних спостережень $Y(i)$ від середнього значення \bar{Y} є величиною незмінною та характеризує властивість спостережуваного процесу. Таким чином, із рівності (2.17) очевидно, що при $R^2 \rightarrow \max, S_{reg} \rightarrow \min$.

Далі розглянемо коефіцієнт лінійної кореляції Пірсона ρ , визначаємий виразом

$$\rho(X, Y) = \frac{\sum_{i=1}^M (X(i) - \bar{X})(Y(i) - \bar{Y})}{\sqrt{\sum_{i=1}^M (X(i) - \bar{X})^2 * \sum_{i=1}^M (Y(i) - \bar{Y})^2}} \in [-1,1]. \quad (2.18)$$

Згідно з аналізом, представленим у книзі [19], зв'язок двох аналізованих коефіцієнтів має такий вигляд

$$\rho^2(\hat{Y}, Y) = R^2 \quad (2.19)$$

$$\rho(\hat{Y}, Y) = \text{sign}(\alpha_1) \sqrt{R^2}. \quad (2.20)$$

Таким чином, відомо, що модуль величини $\rho(\hat{Y}, Y)$ дорівнює модулю величини R , а отже, можна сформулювати наступну властивість

$$R^2 \rightarrow \max, S_{reg} \rightarrow \min \quad (2.21)$$

$$|\rho(\hat{Y}, Y)| \rightarrow \max, S_{reg} \rightarrow \min. \quad (2.22)$$

Усі подробиці регресійного аналізу, а також зв'язки аналізованих коефіцієнтів наведено у книзі [19].

Повернімося до досліджуваних вибірок часового ряду. Нехай даний тимчасовий ряд Z_1^T , для деякої вибірки Z_t^M , що належить даному тимчасовому ряду, визначимо всі значення $S_k^M(\alpha_1, \alpha_0)$ для $k \in \{1, 2, \dots, t-1\}$, $M = \text{const}$. Далі, у безлічі значень S_k^M знайдемо мінімальне

$$S_{kmin}^M = \min(S_1^M, S_2^M, \dots, S_{t-1}^M) \quad (2.23)$$

Відповідно до властивості (2.20) мінімум помилки регресії S_{kmin}^M відповідає максимуму модуля коефіцієнта лінійної кореляції (2.18). Тобто якщо для $k \in \{1, 2, \dots, t-1\}$ і $M = \text{const}$ визначити множину значень модуля кореляції

$$\rho_k^M = |\rho(\hat{Z}_t^M, Z_t^M)| = \frac{|\sum_{i=1}^M (\hat{Z}(t+i) - \bar{Z})(Z(t+i) - \bar{Z})|}{\sqrt{\sum_{i=1}^M (\hat{Z}(t+i) - \bar{Z})^2 * \sum_{i=1}^M (Z(t+i) - \bar{Z})^2}} \in [0, 1], \quad (2.24)$$

а після визначити максимальне значення отриманого множника

$$\rho_{kmax}^M = \max(\rho_1^M, \rho_2^M, \dots, \rho_{t-1}^M), \quad (2.25)$$

то затримка $kmin$ з виразу (2.21) та затримка $kmax$ з виразу (2.23) будуть рівні між собою, тобто $kmin = kmax$. Проведені розрахунки підтверджують це твердження.

Визначену в (2.21) та (2.23) затримку, що відповідає мінімуму помилки регресії S_{kmin}^M та максимуму модуля кореляції ρ_{kmax}^M позначимо $kmax$, а вибірку Z_{t-kmax}^M назовемо вибіркою максимальної подоби (most similar pattern). Вибір максимальної подібності Z_{t-kmax}^M є вибіркою, яка при підстановці рівняння (2.2), дає в результаті значення вибірки \hat{Z}_t^M , яка максимально точно описує вихідну вибірку Z_t^M . При реалізації обчислень визначення вибірки максимальної схожості Z_{t-kmax}^M можна використовувати як значення помилки S_k^M , так і значення модуля кореляції ρ_k^M . У наведеному в дисертації прикладі використовувалось значення коефіцієнта лінійної кореляції (розділ 3.1).

Гіпотеза подібності. Якщо вихідна вибірка Z_t^M та модельна вибірка \hat{Z}_t^M отримана на підставі (2.2) з використанням вибірки Z_{t-kmax}^M , мають значення величини ρ_{kmax}^M , близьке до одиниці, то для деяких значень P та вибірок Z_{t-kmax}^M , Z_t^{M+P} значення величини ρ_{kmax}^{M+P} також близьке до одиниці.

Аналогічним чином можна сформулювати гіпотезу подібності до u разі врахування помилок регресії S_k^M : якщо вихідна вибірка Z_t^M та модельна вибірка \hat{Z}_t^M , отримана за формулою (2.2) на підставі вибірки Z_{t-kmax}^M , мають мінімальне значення помилки $S_{kmin}^M = \min(S_1^M, S_2^M, \dots, S_{t-1}^M)$, то для деяких значень P та вибірок Z_{t-kmin}^{M+P} , Z_t^{M+P} величини помилки регресії $S_{kmin}^{M+P} \rightarrow \min(S_1^{M+P}, S_2^{M+P}, \dots, S_{t-1}^{M+P})$. Далі в роботі використано перше формулювання гіпотези подібності.

Представлені у четвертому розділі дисертації результати прогнозування підтверджують справедливість гіпотези для досліджуваних тимчасових рядів. Для тимчасових рядів з інших предметних областей справедливість гіпотези необхідно перевіряти.

2.1.4. Опис моделі екстраполяції

Нехай дано тимчасовий ряд Z_1^T . Для цього часового ряду потрібно визначити значення Z_{T+1}^P . Використовуючи властивість вибірок, сформульована у розділі 2.1.3., виразимо вибірку Z_{T+1}^P через деяку вибірку Z_T^P , яка лежить всередині вихідного часового ряду Z_1^T

$$\widehat{Z}_{T+1}^P = \alpha_1 Z_\tau^P + \alpha_0 I^P$$

Алгоритм визначення вибірки Z_τ^P складається із двох кроків.

- Визначення вибірки Z_{kmax}^M .
- Обчислення Z_τ^P .

Розглянемо детально кожен крок

Визначення вибірки Z_{kmax}^M . На цьому кроці для вибірки Z_{T-M+1}^M , яка містить значення процесу безпосередньо перед моментом прогнозу, знаходимо вибірку максимальної подоби Z_{kmax}^M . Пошук вибірки максимальної подоби здійснюємо перебором усіх можливих значень затримок $k \in \{1, 2, \dots, T - M - 1\}$. Для кожного значення k із зазначеного діапазону вирішуємо задачу апроксимації (розділ 2.1.2.), в результаті якої визначаємо коефіцієнти α_1 і α_0 , відповідні k . Далі для знайденої пари коефіцієнтів визначаємо значення модельної вибірки \widehat{Z}_{T-M+1}^M , на основі яких обчислюємо значення ρ_k^M (2.22). Після того, як безліч значень ρ_k^M для $k \in \{1, 2, \dots, T - M - 1\}$ отримано, визначаємо значення ρ_{kmax}^M за виразом (2.23) та відповідну вибірку $Z_{T-M+1-kmax}^M$. Для спрощення позначимо затримку $kmax^* = T - M + 1 - kmax$ та вибірку максимальної подоби Z_{kmax}^M (SimilarHistory).

Обчислення Z_τ^P . Відповідно до гіпотези подібності, сформульованої в розділі 2.1.3., як вибірка Z_τ^P використовуємо вибірку $Z_{kmax^*+M}^P$, тобто вибірку, розташовану на осі часу відразу за вибіркою максимальної подоби. Вибірку Z_τ^P назвемо базовою історією (BaseHistory). На малюнку 2.6 представлено розташування всіх розглянутих вибірок.

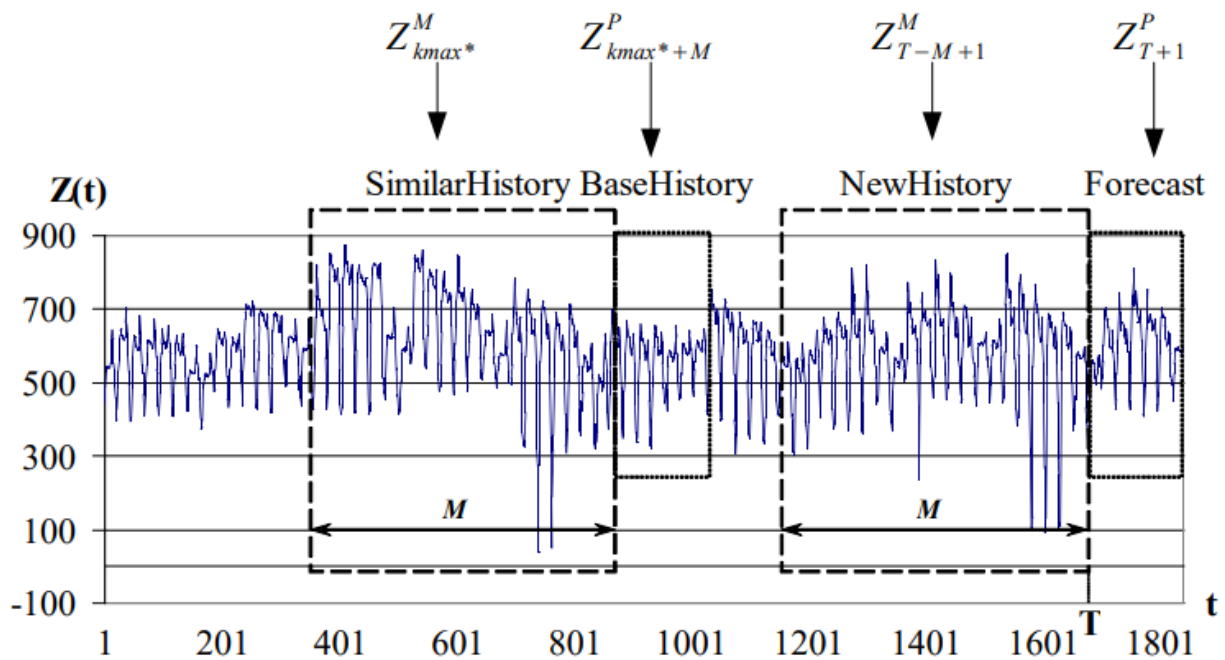


Рисунок 2.4 - Положення вибірок Z_{kmax}^M , Z_{kmax+M}^P , Z_{T-M+1}^M та Z_{T+1}^P

Екстрапольовані значення вибірки \hat{Z}_{T+1}^P (Прогноз, Forecast) визначається за формулою

$$\hat{Z}_{T+1}^P = \alpha_1 Z_{kmax+M}^P + \alpha_0 I^P = EMMSP(M), \quad (2.26)$$

яка являє собою модель екстраполяції часових рядів по вибірці максимальної подоби (extrapolation model on most similar pattern, далі EMMSP). У роботі [40] використовується термін closest pattern, у роботі [42] - closest piece, у роботі [41] - most similar vector.

Особливості EMMSP:

- Модель відноситься до класу авторегресійних моделей прогнозування;
- Модель може працювати з нерівновідданими тимчасовими рядами;
- модель працює зі стаціонарними та нестаціонарними тимчасовими рядами;
- модель має один параметр M , визначення якого докладно розглянуто у третьому розділі дисертації;
- Екстраполяція P значень тимчасового ряду проводиться за одну ітерацію.

2.2. Модель з урахуванням зовнішніх факторів

2.2.1. Вибірки часових рядів

Нехай дано тимчасовий ряд $Z(t)$ та зовнішні фактори, представлені у вигляді тимчасових рядів $X_1(t), \dots, X_S(t)$, що відповідають вихідному ряду за відмітками часу. Потрібно побудувати модель прогнозування вихідного часового ряду $Z(t)$, яка враховуватиме вплив значень зовнішніх факторів $X_1(t), \dots, X_S(t)$.

Якщо зовнішні фактори $X_1(t), \dots, X_S(t)$ мають тимчасовий дозвіл, відмінне від дозволу вихідного ряду $Z(t)$, необхідно зробити додаткові перетворення та привести зовнішні фактори в відповідність з вихідним часовим рядом $Z(t)$ за відмітками часу.

В основу моделі екстраполяції з урахуванням зовнішнього фактора покладено припущення про повторюваність вибірок часового ряду (2.1.1.). Крім того, відомо, що для обліку зовнішніх факторів в авторегресійних моделях вводяться додаткові регресори (розділ 1.3.2).

У цьому розділі спочатку розглядається завдання апроксимації однієї вибірки за допомогою іншої з урахуванням вибірок зовнішніх факторів, а потім результати апроксимації застосовуються для побудови моделі прогнозування часового ряду.

2.2.2. Апроксимація вибірки

Нехай задана вибірка вихідного часового ряду Z_t^M та вибірки $X_{(1)t}^M, \dots, X_{(S)t}^M$ є вибірками зовнішніх факторів, відповідних Z_t^M за позначками часу. Апроксимуємо вибірку Z_t^M з урахуванням вибірок $X_{(1)t}^M, \dots, X_{(S)t}^M$ за формулою

$$Z_t^M = \alpha_{S+1} Z_{t-k}^M + \alpha_S X_{(S)t}^M + \dots + \alpha_1 X_{(1)t}^M + \alpha_0 I^M + E^M$$

Тут $\alpha_{S+1}, \alpha_S, \dots, \alpha_0$ — коефіцієнти, що визначаються за методом найменших квадратів. Вектор E^M - вектор значень помилок апроксимації, I^M - поодинокий вектор. Вираз (2.26) можна переписувати у вигляді

$$\hat{Z}_t^M = \alpha_{S+1} Z_{t-k}^M + \alpha_S X_{(S)t}^M + \dots + \alpha_1 X_{(1)t}^M + \alpha_0 I^M$$

Постановка задачі апроксимації вибірки. Нехай дана функціонально залежність (2.27). Необхідно визначити такі значення коефіцієнтів $\alpha_{S+1}, \alpha_S, \dots, \alpha_0$, щоб квадрат відхилень модельних значень вибірки \hat{Z}_t^M від фактичних Z_t^M був мінімальний

$$\sigma^2 = \sum_{i=0}^{M-1} (Z(t+i) - \hat{Z}(t+i))^2 \rightarrow \min \quad (2.27)$$

Після знаходження коефіцієнтів $\alpha_{S+1}, \alpha_S, \dots, \alpha_0$ необхідно оцінити вектор помилок E^M .

Рішення. Нехай дана функціональна залежність (2.26), тоді функція суми квадратів має вигляд

$$S_k^M(\alpha_{S+1}, \dots, \alpha_0) = \sum_{i=0}^{M-1} (Z(t+i) - (\alpha_{S+1} Z(t-k+i) + \alpha_S X_S(t+i) + \dots + \alpha_1 X_1(t+i) + \alpha_0))^2 \quad (2.28)$$

Повторюючи міркування, наведені у розділі 2.1.2., наведемо завдання до матричного рівняння

$$Z_X * A = Z_Y, \text{ де } A = \begin{bmatrix} \alpha_{S+1} \\ \alpha_S \\ \dots \\ \alpha_1 \\ \alpha_0 \end{bmatrix} \quad (2.29)$$

Вирішуючи матричне рівняння (2.30), визначаємо значення коефіцієнтів $\alpha_{S+1}, \alpha_S, \dots, \alpha_0$, що відповідають мінімуму функції $S_k^M(\alpha_{S+1}, \alpha_S, \dots, \alpha_0)$. Рішення, як і в попередньому випадку, буде знайдено за допомогою зворотної матриці

$$Z_X^{-1} * Z_X * A = Z_X^{-1} * Z_Y, \quad (2.30)$$

$$A = Z_X^{-1} * Z_Y. \quad (2.31)$$

Оцінка вектора помилок апроксимації E^M описана в розділі 2.1.2.

2.2.3. Подібність вибірок

Нехай дана вибірка вихідного часового ряду Z_t^M та вибірки зовнішніх факторів $X_{(1)t}^M, \dots, X_{(S)t}^M$. Модельна вибірка \hat{Z}_t^M визначається лінійною залежністю (2.27). Обчислимо всі значення $S_k^M(\alpha_{S+1}, \dots, \alpha_0)$ для затримок $k \in \{1, 2, \dots, t-1\}$ і $M = const$. Далі, у безлічі значень S_k^M знайдемо мінімальне за виразом (2.21). У разі множинної регресії (2.27) рівність мінімуму помилки регресії (2.21) та максимуму модуля лінійної кореляції (2.23) не виконується, а тому в цьому випадку слід розглядати лише помилку регресії $S_k^M(\alpha_{S+1}, \dots, \alpha_0)$.

Визначену у виразі (2.21) затримку, що відповідає мінімуму помилки регресії S_{kmin}^M позначимо за аналогією з попереднім випадком $kmax$, а вибірку Z_{t-kmax}^M назовемо вибіркою максимальної подоби (Most similar pattern). Вибір максимальної подібності Z_{t-kmax}^M є вибіркою, яка при підстановці рівняння (2.27), дає в результаті значення вибірки \hat{Z}_t^M , яка максимально точно описує вихідну вибірку Z_t^M з урахуванням вибірок зовнішніх факторів $X_{(1)t}^M, \dots, X_{(S)t}^M$. Величину мінімальної помилки S_{kmin}^M , відповідну затримку $kmax$ будемо далі позначати S_{kmax}^M .

Гіпотеза подібності. Якщо вихідна вибірка Z_t^M та модельна вибірка \hat{Z}_t^M отримана на підставі (2.27) з використанням вибірки Z_{t-kmax}^M вибірок зовнішніх факторів $X_{(1)t}^M, \dots, X_{(S)t}^M$, мають мінімальне значення величини S_{kmax}^M , то для деяких значень P та вибірок $Z_{t-kmax}^M, X_{(1)t}^{M+P}, \dots, X_{(S)t}^{M+P}, Z_t^{M+P}$ значення величини S_{kmax}^{M+P} також прагне мінімального.

Представлені у четвертому розділі дисертації результати прогнозування підтверджують справедливість гіпотези для досліджуваних тимчасових рядів. Для тимчасових рядів з інших предметних областей справедливість гіпотези необхідно перевіряти додатково.

2.2.4. Опис моделі

Нехай дано вихідний часовий ряд Z_1^T та зовнішні фактори $X_{(1)1}^{T+P}, \dots, X_{(S)1}^{T+P}$. Для вихідного часового ряду потрібно визначити значення Z_{T+1}^P , враховуючи доступні значення Z_1^T та $X_{(1)T+1}^P, \dots, X_{(S)T+1}^P$. Використовуючи властивість вибірок,

сформульоване у розділі 2.2.3., виразимо вибірку Z_{T+1}^P через деяку вибірку Z_τ^P , що лежить усередині вихідного часового ряду Z_1^T , і вибірки $X_{(1)T+1}^P, \dots, X_{(S)T+1}^P$ наступним чином

$$\hat{Z}_{T+1}^P = \alpha_{S+1} Z_\tau^P + \alpha_S X_{(S)T+1}^P + \alpha_0 I^P. \quad (2.32)$$

Алгоритм визначення вибірки Z_τ^P складається з двох кроків:

- Визначення вибірки Z_{kmax}^M ;
- Обчислення Z_τ^P .

Визначення вибірки Z_{kmax}^M . На цьому кроці, як у випадку екстраполяції без урахування зовнішніх факторів для вибірки Z_{T-M+1}^M , містить значення процесу безпосередньо перед моментом прогнозу, знаходимо вибірку максимальної подоби Z_{kmax}^M . Пошук вибірки максимальної подоби здійснюємо перебором усіх можливих значень затримок $k \in \{1, 2, \dots, T - M - 1\}$. Для кожного значення k із зазначеного діапазону вирішуємо задачу апроксимації (2.2.2.), в результаті якої визначаємо коефіцієнти $\alpha_{S+1}, \alpha_S, \dots, \alpha_0$, відповідні k . Далі для обчислених коефіцієнтів визначаємо значення модельної вибірки, на підставі яких обчислюємо значення помилки регресії S_k^M (2.22). Після того, як безліч значень S_k^M для $k \in \{1, 2, \dots, T - M - 1\}$ отримано, визначаємо значення S_{kmax}^M (2.23) та відповідну вибірку $Z_{T-M+1-kmax}^M$. Як і раніше позначимо затримку $kmax^* = T - M + 1 - kmax$, а вибірку максимальної подоби Z_{kmax}^M (SimilarHistory).

Обчислення Z_τ^P . Відповідно до гіпотези подібності, сформульованої в розділі 2.2.3., як вибірка Z_τ^P у виразі (2.33) використовуємо вибірку $Z_{kmax^*+M}^P$, тобто вибірку, розташовану на осі часу відразу за вибіркою максимальної подоби. Для випадку обліку одного зовнішнього чинника положення вибірок показано малюнку 2.7

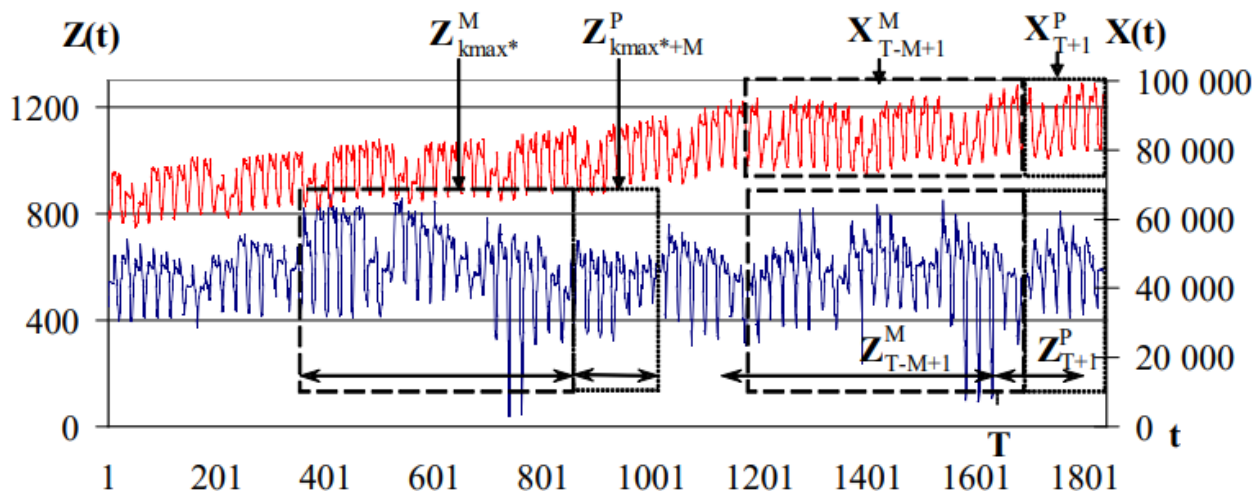


Рисунок 2.5 - Положення вибірок на осі часу

Для зручності назвемо вибірки таким чином:

- вибір нової історії Z_{T-M+1}^M (NewHistory);
- вибірка максимальної подоби Z_{kmax*}^M (SimilarHistory);
- вибір базової історії Z_T^P (BaseHistory);
- вибір історії зовнішнього фактора X_{T-M+1}^M (FactorHistory);
- вибір прогнозу зовнішнього фактора X_{T+1}^P (FactorForecast).

Екстраполовані значення вибірки \hat{Z}_{T+1}^P (Forecast) визначаємо за формулою

$$\hat{Z}_{T+1}^P = \alpha_{S+1} Z_{kmax*+M}^P + \alpha_S X_{(S)T+1}^P + \dots + \alpha_1 X_{(1)T+1}^P + \alpha_0 I^P = EMMSPX(M),$$

яка являє собою розширену модель екстраполяції тимчасових рядів щодо вибірки максимальної подоби (extrapolation model on most similar pattern extended, далі EMMSPX).

Слід зазначити, що зовнішні чинники, зазвичай, надають вплив на досліджуваний процес у відповідний час. Наприклад, з роботи [9] відомо, що на енергоспоживання чинить істотний вплив температура навколишнього середовища, а саме: при різких змін температури стрибкоподібно змінюється енергоспоживання.

На завершення цього розділу необхідно відзначити особливості EMMSPX:

- Модель відноситься до класу авторегресійних моделей прогнозування, що враховують дискретні зовнішні чинники;

- Модель може враховувати кілька дискретних зовнішніх факторів;
- модель ефективно працює з нерівновідданими тимчасовими рядами;
- модель працює зі стаціонарними та нестаціонарними тимчасовими рядами;
- модель має один параметр M , визначення якого докладно розглянуто у третьому розділі дисертації;
- Екстраполяція P значень тимчасового ряду проводиться за одну ітерацію

2.3. Варіанти моделей щодо вибірки максимального подоби

Модель (2.34) можна розділити на дві частини – модель авторегресії та модель зовнішнього фактора

$$\hat{Z}_{T+1}^P = \alpha_2 Z_{kmax*+M}^P + \alpha_0 I^P + \alpha_1 X_{T+1}^P. \quad (2.33)$$

Модель авторегресії та модель зовнішнього фактора можуть бути модифіковані для підвищення точності прогнозування.

Авторегресійна EMMSP з безліччю вибірок. Прогнозні значення \hat{Z}_{T+1}^P обчислюються як лінійна комбінація кількох вибірок базової історії з різними затримками

$$\hat{Z}_{T+1}^P = \alpha_q Z_{\tau_q}^P + \alpha_{q-1} Z_{\tau_{q-1}}^P + \dots + \alpha_1 Z_{\tau_1}^P + \alpha_0 I^P. \quad (2.34)$$

В рамках дисертації проводились дослідження ефективності збільшення кількості вибірок, що приймаються до уваги. На практиці кількість вибірок вихідного часового ряду, що використовуються, не перевищує двох, тобто прогнозні значення \hat{Z}_{T+1}^P обчислюються як лінійна комбінація двох вибірок базової історії з різними затримками

$$\hat{Z}_{T+1}^P = \alpha_2 Z_{\tau_2}^P + \alpha_1 Z_{\tau_1}^P + \alpha_0 I^P. \quad (2.35)$$

Однак не виключено, що в ряді завдань корисним буде використання трьох-чотирьох вибірок. Ця модель (2.37) застосовувалася для прогнозування тимчасових рядів ціни електроенергії [45].

Авторегресійна EMMSP з використанням q-того ступеня значень вибірок. Модель прогнозування є лінійною комбінацією ступенів вибірки максимальної подоби

$$\hat{Z}_{T+1}^P = \alpha_q (Z_\tau^P)^q + \alpha_{q-1} (Z_\tau^P)^{(q-1)} + \dots + \alpha_1 Z_\tau^P + \alpha_0 I^P \quad (2.36)$$

Тут $(Z_\tau^P)^n$ - вибірка, значеннями якої є n-ні ступені значень тимчасового ряду

$$(Z_\tau^P)^n = \begin{bmatrix} Z^n(\tau) \\ \dots \\ Z^n(\tau + P) \end{bmatrix} \quad (2.37)$$

В рамках дисертації проводились дослідження ефективності підвищення ступеня вибірки максимальної подоби, які показали, що точність прогнозування для деяких часових рядів підвищується при використанні другого ступеня, тобто моделі

$$\hat{Z}_{T+1}^P = \alpha_2 (Z_\tau^P)^2 + \alpha_1 Z_\tau^P + \alpha_0 I^P \quad (2.38)$$

Подальше збільшення ступеня вибірки максимальної подоби ефективності моделі не підвищує.

EMMSPX з використанням q-того ступеня значень зовнішніх факторів.
У разі доступності даних лише по одному зовнішньому фактору можлива модифікація моделі аналогічно (2.38)

$$\hat{Z}_{T+1}^P = \alpha_1 Z_\tau^P + \alpha_0 I^P + \beta_q (X_T^P)^q + \beta_{q-1} (X_T^P)^{(q-1)} + \dots + \beta_1 X_T^P. \quad (2.39)$$

При цьому значення вибірки $(X_T^P)^n$ визначаються як n -ні ступені значень зовнішнього чинника (2.39)

Важливо, що при побудові моделей (2.36) - (2.41) застосовується одноманітний підхід до аналізу та проектування моделей, описаний у розділі 2.2. справжньої роботи.

Обмеження застосування моделі екстраполяції часових рядів за вибіркою максимальної подоби. Запропоновані у цьому розділі моделі EMMSP та EMMSPX можуть застосовуватися для прогнозування тимчасових рядів на значення P вперед при виконанні набору умов.

- Довжина часового ряду становить не менше ніж $500 P - 700 P$.
- Тимчасовий ряд рівновіддалений; у разі нерівновіддаленого тимчасового ряду застосування моделі можливе лише за умови, що його рівновіддалені відрізки повинні містити не менше, ніж $100 P - 150 P$ відліків.
- Тимчасовий ряд відноситься до класу тимчасових рядів із довгою пам'яттю.
- Завдання прогнозування на P значень уперед відноситься до класу короткострокового або середньострокового прогнозування даного типу часового ряду. Не рекомендується використовувати розроблену модель для довгострокового прогнозування
- У разі обліку набору дискретних зовнішніх факторів, їх тимчасовий дозвіл має бути наведений до дозволу вихідного часового ряду. Довжина вихідного часового ряду та тимчасових рядів зовнішніх факторів може не збігатися.

Крім того, слід зазначити, що для вузькоспеціалізованих задач іноді застосовуються спеціальні моделі прогнозування. Так наприклад, для завдання прогнозування рівня цукру крові людини застосовуються моделі з урахуванням диференціальних рівнянь [8].

Для прогнозування природних явищ, таких як землетруси, застосовується, наприклад, модель, основою якої покладено нелінійні клітини (або стільники), що знаходяться під впливом зовнішнього поля, і які мають внутрішній стан, що

змінюється в часі під впливом цього поля [18]. Аналогічні моделі розробляються і застосовуються для спеціальних процесів та систем.

2.4. Висновки

1) У цьому розділі запропоновані моделі екстраполяції тимчасових рядів з вибірки максимальної подібності з урахуванням та без урахування зовнішніх факторів.

2) Запропоновані моделі відносяться до класу авторегресійних моделей прогнозування і мають всі переваги, характерні для цього класу.

3) Запропоновані моделі усувають суттєвий недолік вказаного класу - велика кількість вільних параметрів, що вимагають ідентифікації. Обидві моделі екстраполяції щодо вибірки максимального подоби мають єдиний параметр.

4) Розроблено варіанти моделей з вибірки максимального подібності, використання яких може призводити до підвищення точності прогнозування часового ряду.

3 МЕТОД ПРОГНОЗУВАННЯ НА МОДЕЛІ ЕКСТРАПОЛЯЦІЇ З ВИБІРКИ МАКСИМАЛЬНОГО ПОДОБИ

Метод прогнозування на моделі екстраполяції часових рядів по вибору максимальної подоби містить такі алгоритми.

- 1) Алгоритм екстраполяції тимчасового ряду без урахування зовнішніх факторів.
 - 2) Алгоритм екстраполяції часового ряду з урахуванням зовнішніх факторів.
 - 3) Алгоритм ідентифікації моделі.
 - 4) Алгоритм побудови довірчого інтервалу для прогнозних значень.
- Далі докладно розглянуті усі перелічені вище алгоритми.

3.1. Алгоритм екстраполяції тимчасового ряду без урахування зовнішніх факторів

Алгоритм екстраполяції складається з наступних кроків.

- 1) Визначити вибірку нової історії.
- 2) Визначити вибірку максимальної подоби.
- 3) Визначити вибірку базової історії.
- 4) Обчислити прогнозні значення.

Далі наведемо опис кожного вказаного вище кроку, ілюструючи розрахунки розв'язанням наступного завдання. Нехай дані значення тимчасового ряду цін на акції APPLE (№1 у таблиці 7) з 22.09.2019 до 22.09.2022; довжина часового ряду дорівнює 24 624. Позначимо часовий ряд $Z(t)$. Потрібно визначити 24 значення часового ряду цін за 23.09.2022. Вважаємо параметр моделі $M = 216$ заданим.

1) Визначити вибірку нової історії.

Вибіркою нової історії є вибірка часового ряду, значення якої передують моменту прогнозу T . У поточній постановці задачі вибірка нової історії дорівнює $Z_{T-M+1}^M = Z_{24409}^{216}$.

2) Визначити вибірку максимальної подоби.

Для визначення вибірки максимальної подоби необхідно визначити значення модуля лінійної кореляції ρ_k^M для вибірки Z_{T-M+1}^M і всіх вибірок із затримкою $k \in \{1, 2, \dots, T-M-1\}$. При цьому для кожного значення k із зазначеного діапазону потрібно вирішити задачу апроксимації вибірки Z_{T-M+1}^M за допомогою вибірки $Z_{T-M+1-k}^M$. Позначимо момент часу $T-M+1=t$ і розв'яжемо це завдання.

Обчислимо апроксимовані значення вибірки

$$\hat{Z}_{T-M+1}^M = \alpha_1 Z_{T-M+1-k}^M + \alpha_0 I^M, \quad (3.1)$$

а з урахуванням позначення

$$\hat{Z}_t^M = \alpha_1 Z_{t-k}^M + \alpha_0 I^M. \quad (3.2)$$

Згідно з методом найменших квадратів, коефіцієнти апроксимації визначимо, виходячи з рівняння

$$Z_X * A = Z_Y, \quad (3.3)$$

де значення елементів матриці Z_X та Z_Y визначаються наступним чином:

$$Z_X = \begin{bmatrix} \sum_{i=0}^{M-1} Z^2(k+i) & \sum_{i=0}^{M-1} Z(k+i) \\ \sum_{i=0}^{M-1} Z(k+i) & M \end{bmatrix}; \quad (3.4)$$

$$Z_Y = \begin{bmatrix} \sum_{i=0}^{M-1} Z(k+i) * Z(T-M+1+i) \\ \sum_{i=0}^{M-1} Z(T-M+1+i) \end{bmatrix}. \quad (3.5)$$

Знайдені коефіцієнти апроксимації підставимо в (3.1), а далі визначимо значення модуля кореляції ρ_k^M за виразом (2.22). Повторюючи обчислення для кожного значення k із зазначеного діапазону, визначимо безліч значень $\rho_1^M, \rho_2^M, \rho_3^M, \dots, \rho_{T-M-1}^M$. Значення ρ_k^M для $k \in \{1, 2, \dots, 1000\}$ наведено малюнку 3.1.

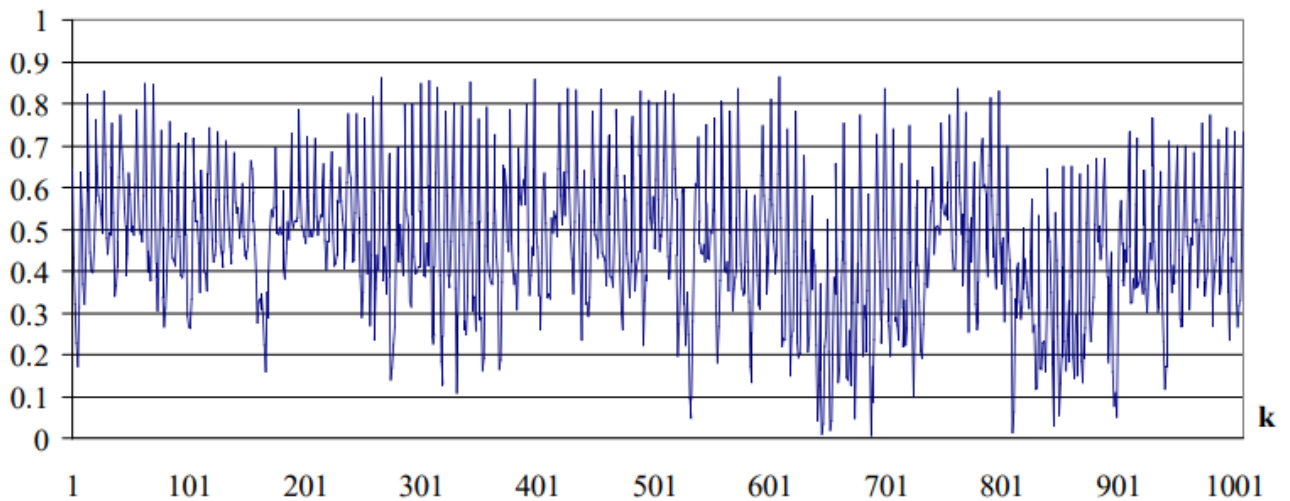


Рисунок 3.1 - Значення міри подібності ρ_k^M для $k \in \{1, 2, \dots, 1000\}$

Далі на підставі (2.23) визначимо значення максимуму кореляції ρ_{kmax}^M та відповідну затримку $kmax$. Для розв'язуваного завдання значення максимуму кореляції $\rho_{kmax}^M = 0.862$ і відповідає затримці $kmax = 9817$. Для затримки $kmax$ рівняння апроксимації має вигляд

$$\hat{Z}_{24409}^{216} = 0.7759Z_{14592}^{216} + 172.7604I^{216} \quad (3.6)$$

Результати апроксимації представлені малюнку 3.2

- Вибірка максимальної схожості, 14-22.10.2020
- Вибірка нової історії, 14-22.06.2022
- Апроксні значення вибірки нової історії, 14-22.06.2022

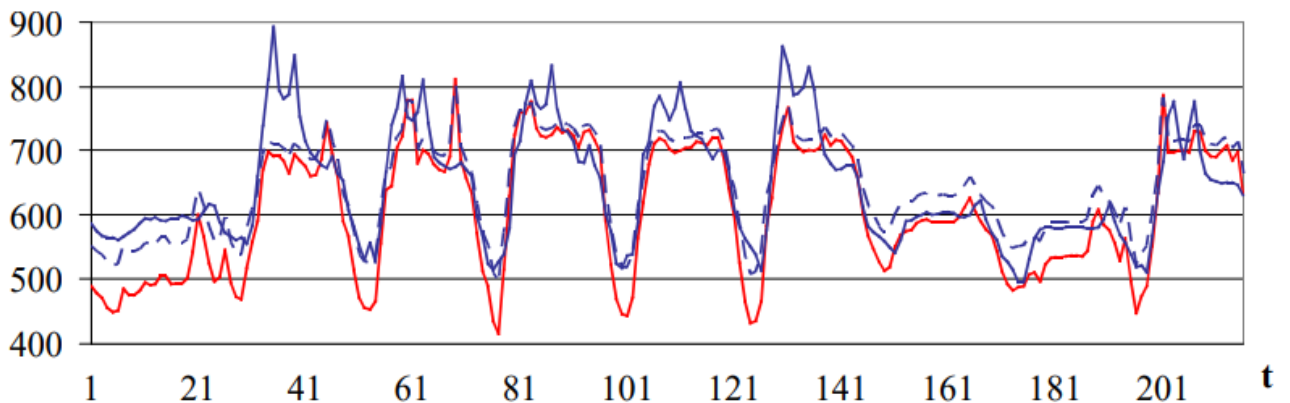


Рисунок 3.2 - Приклад апроксимації

3) Визначити вибірку базової історії.

Згідно з гіпотезою подібності (2.1.3.), як вибірка базової історії Z_{τ}^P беремо вибірку наступну за вибіркою максимальної подоби Z_{14592}^{216} , тобто вибірка базової історії дорівнює $Z_{\tau}^P = Z_{kmax*+M+1}^P = Z_{14809}^{24}$.

4) Обчислити прогнози значення.

Обчислимо значення вибірки \hat{Z}_{24625}^{24} відповідно до залежності

$$\hat{Z}_{24625}^{24} = 0.7759 Z_{14809}^{24} + 172.7604 I^{24} \quad (3.7)$$

Результат екстраполяції представлений малюнку 3.3

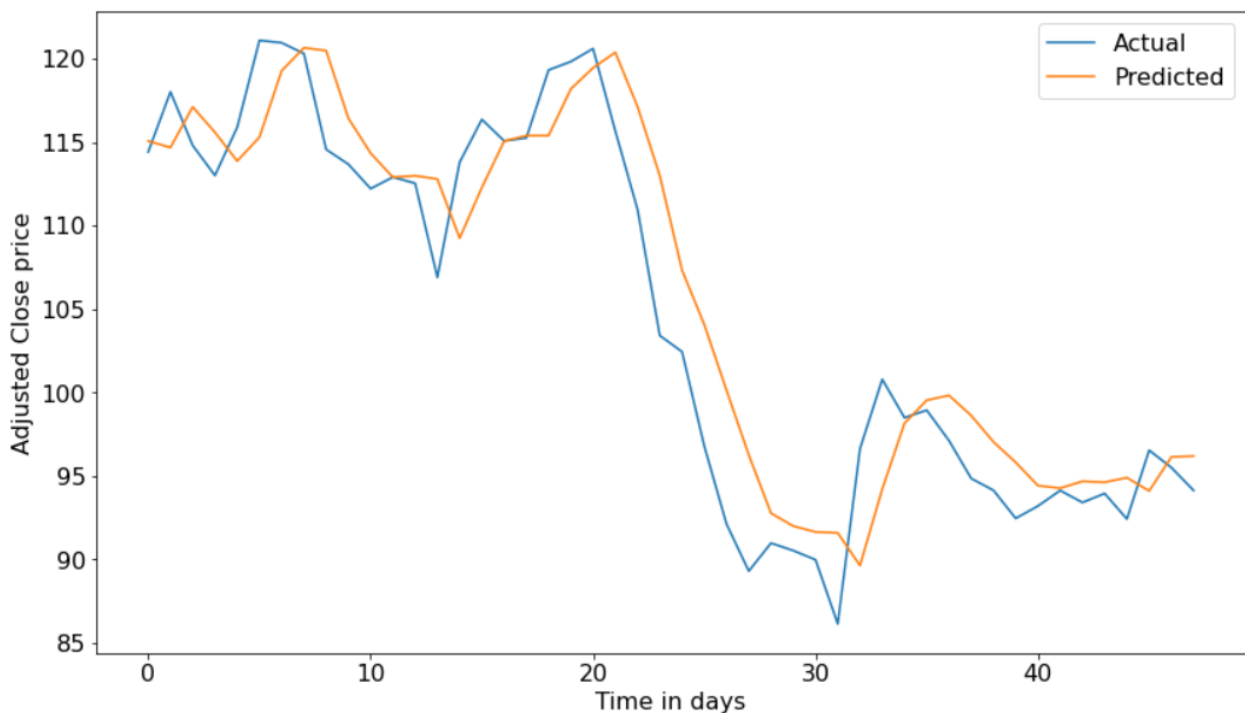


Рисунок 3.3 - Результат екстраполяції \hat{Z}_{24625}^{24} , вибірка базової історії Z_{14808}^{24} і фактичні значення цін на електроенергію

Результати показують, що помилка апроксимації близька, але не дорівнює помилці екстраполяції. Залежність помилок екстраполяції та апроксимації розглянуто у розділі 3.4.

Після опису кроків алгоритму екстраполяції необхідно провести оцінку часу обчислень прогнозних значень за його програмної реалізації.

Оцінка часу розрахунку прогнозних значень. Розроблений алгоритм екстраполяції реалізований за допомогою програмного комплексу MATLAB [46]. Експерименти проводились на персональному комп'ютері наступної модифікації:

- Процесор AMD Ryzen 5 3600 3.60 ГГц, 4ГБ DDR2,
- материнська плата Asus Pro Max Gaming.

Оцінка продуктивності даного персонального комп'ютера за допомогою тесту Java Micro Benchmark складає 828 одиниць. Оцінки продуктивності комп'ютерів та серверів за допомогою даного тесту коливаються у широкому діапазоні від 95 до 22 054 одиниць [47].

Час розрахунку прогнозних значень часового ряду t_p залежить від довжини часового ряду T та продуктивності комп'ютера. Одною з особливостей моделі екстраполяції є те, що P прогнозних значень визначаються за один пробіг алгоритму, наприклад час розрахунку одного значення вперед дорівнює часу розрахунку 24 значень уперед. Експериментальна залежність часу розрахунку t_p від довжини тимчасового ряду T для зазначеного персонального комп'ютера представлено на малюнку 3.4, значення представлені у таблиці 3.1

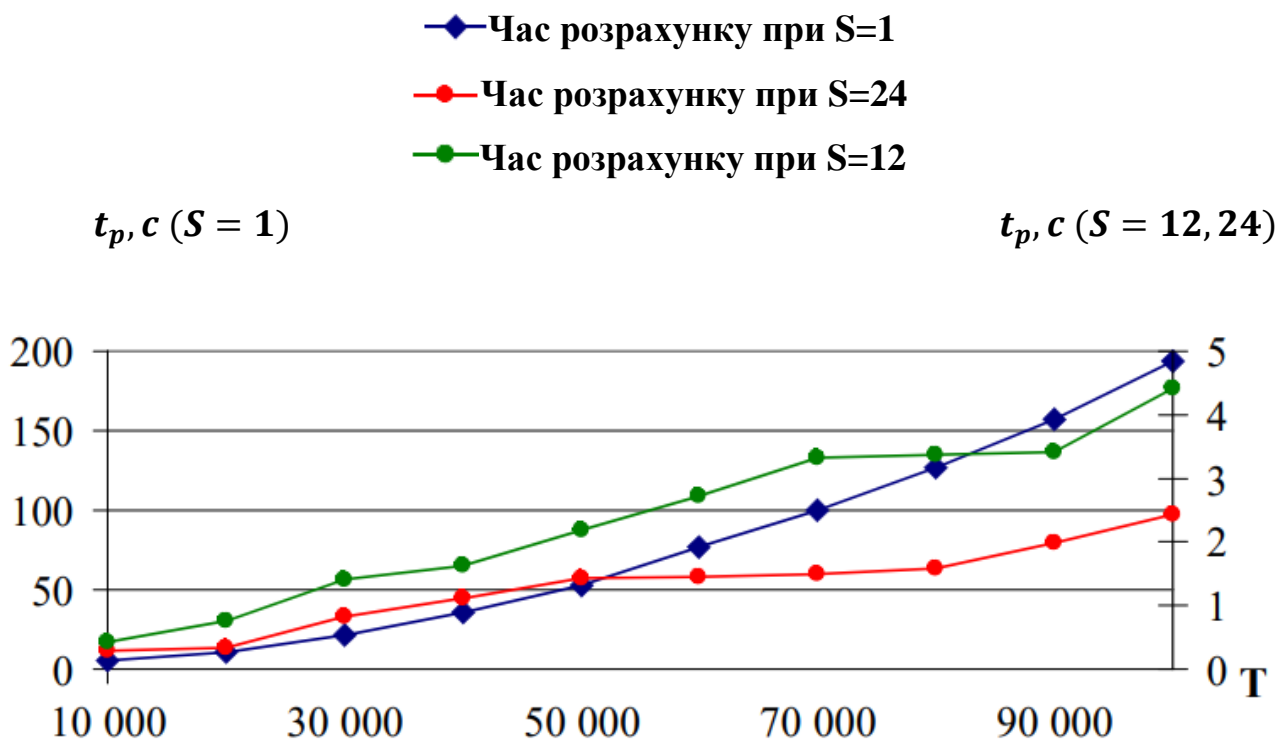


Рисунок 3.4 - Експериментальна залежність часу розрахунку t_p від довжини тимчасового ряду T

На малюнку 3.4 наведено три залежності часу розрахунку на 24 значення вперед t_p від довжини часового ряду T за умови, що безліч значень модуля кореляції ρ_k^M визначається з кроком

- 1) $S=1$, тобто вираховуються значення $\rho_1^M, \rho_2^M, \rho_3^M, \dots, \rho_{T-M-1}^M$;
- 2) $S=12$, тобто вираховуються значення $\rho_{12}^M, \rho_{24}^M, \rho_{36}^M, \dots, \rho_{T-M-1}^M$;
- 3) $S=24$, тобто вираховуються значення $\rho_{24}^M, \rho_{48}^M, \rho_{72}^M, \dots, \rho_{T-M-1}^M$.

Таблиця 3.1 - Значення часу розрахунку за різної довжини часового ряду

Довжина часового ряду T	Час розрахунку t_p , с		
	S=1	S=12	S=24
10000	5.08	0.43	0.29
20000	11.12	0.76	0.33
30000	21.80	1.41	0.83
40000	35.73	1.63	1.11
50000	52.87	2.19	1.43
60000	77.17	2.72	1.45
70000	99.61	3.32	1.49
80000	126.68	3.38	1.59
90000	156.85	3.42	1.99
100000	193.90	4.42	2.44

Згідно з оцінкою [48], високою вважається швидкість, за якої обчислення 24 прогнозних значень займає трохи більше 20 хвилин. Час розрахунку 24 значень тимчасового ряду довжиною 100 000 значень та переборі значень ρ_k^M з кроком $S=1$ становить близько 200 секунд на вказаному персональний комп'ютер. При кроці $S=12$ аналогічний час розрахунків не перевищує 5 секунд; при $S = 24$ - 2.5 секунд (таблиця 2). Дані оцінки підтверджують високу швидкість обчислень

запропонованого алгоритму екстраполяції часового ряду без урахування зовнішніх факторів.

3.2. Алгоритм ідентифікації моделі

3.2.1 Опис алгоритму

У цьому розділі запропоновано алгоритм ідентифікації моделей EMMSP (2.25) та EMMSPX (2.34). Завдання ідентифікації полягає у знаходженні параметра моделі M , який визначає довжину вибірок нової історії та максимальної подоби (рис. 2.6, 2.7). Виконаємо ідентифікацію обох моделей за одним алгоритмом.

- 1) Визначити тестовий та контрольний періоди часового ряду.
- 2) Визначити час попередження P та діапазон можливих значень параметра M .
- 3) Прогнозувати тестовий період на P значень уперед при всіх значення параметра M із встановленого діапазону.
- 4) Побудувати залежність помилки прогнозування від M .
- 5) Експертно визначити остаточне значення параметра M .

Розглянемо ці кроки ідентифікації докладніше.

- 1) Визначити тестовий та контрольний періоди тимчасового ряду.

На цьому кроці вихідний часовий ряд $Z(t)$ розділимо на три частини пропорції, наприклад, 1:1:1. Отримані частини назвемо базовий період (33%), тестовий період (33%) та контрольний період (34%) тимчасового ряду

- 2) Визначити час попередження P та діапазон можливих значень параметра M .

Далі, виходячи з розв'язуваного завдання прогнозування часового ряду, потрібно визначити час попередження P , а також діапазон можливих значень параметра M . Спочатку рекомендуємо брати широкий діапазон можливих значень M , наприклад, $M \in \{2P, \dots, 15P\}$. Значення M можна перебирати з кроком $S=1$, тобто $M \in \{P, P \square 1, P \square 2, \dots, 15P\}$. Однак практика вирішення цього завдання показала, що зручніше перебирати значення параметра M з кроком $S=P$ чи $S=0.5P$, тобто. M

$\in \{P, 1.5 P, 2 P, \dots, 15 P\}$. Цей підхід скорочує час ідентифікації, суттєво не впливаючи на точність наступного прогнозування.

3) Прогнозувати тестовий період на P значень уперед при всіх значеннях параметра M із встановленого діапазону.

Для кожного значення параметра M із встановленого діапазону, прогнозуємо значення тимчасового ряду на P значень уперед усередині тестовий період. Рекомендуємо встановлювати довжину тестового періоду в діапазоні від $150 P$ до $300 P$.

За результатами прогнозування для кожного значення M визначаємо середню абсолютну помилку прогнозу для всього тестового періоду

$$MAE(M) = \frac{1}{K} \sum_{t=Q_{test}}^{Q_{test}+K} |\hat{Z}(t) - Z(t)|, \quad (3.8)$$

де K - кількість значень тимчасового ряду, що потрапили всередину тестового періоду, Q_{test} - початок тестового періоду на осі часу (рис. 3.9), $\hat{Z}(t)$ - прогнозні значення, отримані під час прогнозування з параметром моделі M .

4) Побудувати залежність помилки прогнозування від M .

Будуємо графік залежності $MAE(M)$ для тестового періоду і визначаємо діапазон значень M , що відповідає стійкому мінімуму $MAE(M)$.

Розглянемо як приклад залежність $MAE(M)$ для тимчасового ряду акцій APPLE, представленого на малюнку 3.5 Спочатку було обрано діапазон $M \in \{36, 48, \dots, 360\}$, усередині якого виділено діапазон стійкого мінімуму помилки $M \in \{168, 180, \dots, 240\}$.

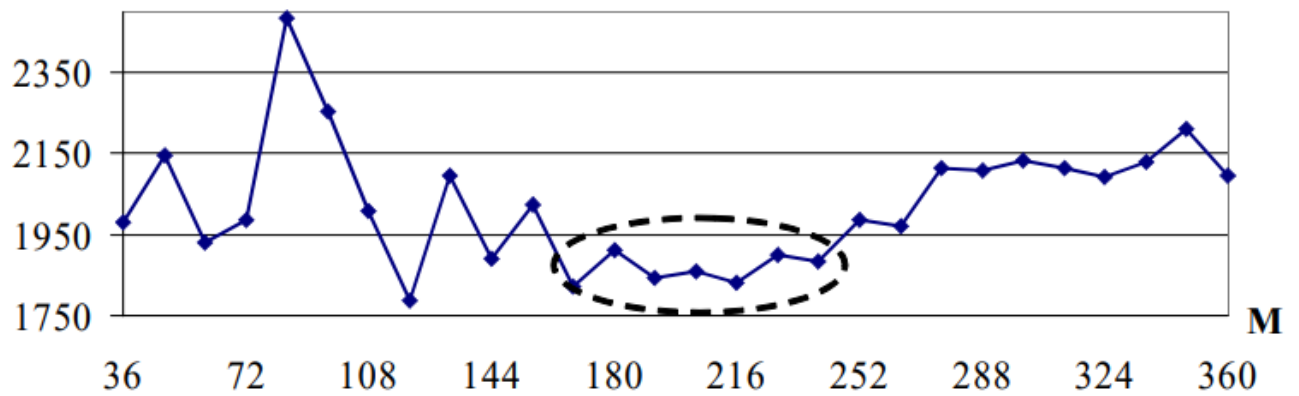


Рисунок 3.5 - Залежність $MAE(M)$ для $M \in \{36, 48, \dots, 360\}$

На малюнку 3.5 представлено типовий вид залежності $MAE(M)$, при малих і великих значеннях M з діапазону $\{1.5 P, 2 P, \dots, 15 P\}$ помилка, як правило, велика. Однак існує стабільний мінімум значень $MAE(M)$ для деяких проміжних значень M . У наведеному прикладі значення параметра M перебиралися з кроком $S = 0.5 P$.

5) Експертно визначити значення параметра M .

На останньому кроці експертом вибирається значення параметра моделі M у середині діапазону стійкого мінімуму.

Рекомендуємо вибрати значення M таким чином, щоб значення $MAE(M)$ для сусідніх точок приблизно збігалися зі значенням $MAE(M)$ для вибраного значення M . У разі, якщо стійкий мінімум значень $MAE(M)$ лежить близько до лівої межі досліджуваного діапазону M , то рекомендуємо брати найбільше значення з стійкого діапазону мінімуму.

Після виконання всіх кроків алгоритму завдання ідентифікації вирішена і модель може застосовуватися для прогнозування контрольного періоду.

3.2.2. Розпаралелювання обчислень

Найбільш ресурсомісткий крок у сенсі часу та обчислювальних потужностей у запропонованому алгоритмі ідентифікації моделі складається у прогнозуванні тестового періоду при різних значеннях параметра моделі M із встановленого діапазону (крок 3). В зв'язку з тим що результати прогнозування при одному значенні параметра M не залежать від результатів прогнозування при іншому значенні M даний процес може бути розпаралелений наступним чином.

1) Розділити тестовий період тимчасового ряду на непересічні підперіоди, об'єднання яких дає тестовий період.

2) Проводити паралельно обчислення прогнозних значень для підперіодів при різних значеннях параметра M .

Таблиця 3.2 – Схема розпаралелювання обчислень

Значення параметра M	Тестовий підперіод №1	Тестовий підперіод №2	...	Значення $MAE(M)$
M_1	$MAE_{№1}(M_1)$	$MAE_{№2}(M_1)$...	$MAE(M_1)$
M_2	$MAE_{№1}(M_2)$	$MAE_{№2}(M_2)$...	$MAE(M_2)$
M_3	$MAE_{№1}(M_3)$	$MAE_{№2}(M_3)$...	$MAE(M_3)$
...

Для значення параметра моделі M_1 проводиться прогноз усередині тестового підперіоду №1 та визначається величина помилки $MAE_{№1}(M_1)$. Паралельно з даними обчисленнями проводяться оцінки інших величин $MAE_{№q}(M_i)$. Після завершення всіх обчислень підсумкові значення $MAE(M_1)$ визначаються як середні величини від множини значень оцінок помилки $MAE_{№1}(M_1), MAE_{№2}(M_1), \dots$, відповідних розглянутого значення параметра M_1 .

У разі посилення вимог щодо швидкості ідентифікації моделі кількість тестових підперіодів має бути збільшено.

3.2.3. Набори моделей

Рішенням задачі ідентифікації моделі екстраполяції є значення параметра M . Якщо сформована модель $EMMSP(M)$ або $EMMSPX(M)$ має недостатню точність прогнозування, можливе підвищення точності за рахунок використання різних моделей для різних відрізків часового ряду.

Нехай потрібно прогнозувати тимчасовий ряд $Z(t)$. При цьому шуканий часовий ряд $Z(t)$ можна явно розбити на деякі чергуючі відрізки, наприклад, тимчасовий ряд, що має погодинний дозвіл, може бути розбитий по днях тижня,

місяцях і т.д. Тоді у процесі ідентифікації моделі необхідно визначати параметр моделі M для кожного встановленого відрізка окремо. Ілюстрація такого набору моделей для прогнозування часового ряду, що має погодинний дозвіл, є таблиця 3.3

Таблиця 3.3 Набір моделей для прогнозування часового ряду, що має погодинний дозвіл

День тижня	Параметр моделі
Понеділок	M_1
Вівторок	M_2
Середа	M_3
Четвер	M_4
П'ятниця	M_5
Субота	M_6
Неділя	M_7

У цьому прикладі кожному дню тижня відповідає власне значення параметра M : при прогнозі значень понеділка використовується модель $EMMSP(M_1)$, при прогнозі вівторка — $EMMSP(M_2)$ і т.д.

Отримати набір моделей для тимчасового ряду можна на підставі результатів ідентифікації моделі в такий спосіб.

1) Розбити результати прогнозування тестового періоду при різних значеннях параметра M відповідно до встановлених заздалегідь відрізками часового ряду (наприклад, днями тижня).

2) Визначити залежність $MAE(M)$ для кожного відрізка.

3) Визначити діапазон стійкого мінімуму та остаточне значення M для кожного відрізка окремо.

Набори моделей застосовувалися для прогнозування часових рядів акцій на ринку. Результати прогнозування цін, подані у розділах 4.1.2. та 4.1.3. дисертації,

підтверджують, що застосування наборів підвищує точність прогнозування(таблиці 10 і 14)

3.3.4. Оцінка часу ідентифікації

Ідентифікація моделі, запропонована у цьому розділі дисертації, має наступну оцінку часу обчислень на кроці №3 (Інші кроки не вимагають оцінки часу)

$$t_{id} = \frac{T_{TEST}}{P} * t_p * N_M * \frac{828}{PC}. \quad (3.9)$$

Тут t_{id} – час ідентифікації за секунди, T_{TEST} – довжина тестового періоду часового ряду, P – час попередження, N_M – кількість можливих значень параметра M , PC – оцінка продуктивності комп'ютера з тесту Java Micro Benchmark [47]. Час розрахунку t_p визначається за таблицями 2, 3 і 4. Якщо довжина часового ряду перевищує діапазони, задані в таблицях, необхідно лінійно екстраполювати значення t_p за відповідного кроку перебору значень затримки k щодо безлічі значень модуля кореляції.

Для прикладу $T_{TEST} = 250$; в зв'язку з тим що ідентифікації виконувалася на вказаному персональному комп'ютері, оцінка його продуктивності $PC = 828$. Час попередження $P = 24$ кількість значень параметра M в діапазоні, що оцінюється $N_M = 28$ (рис. 3.10). Початок тестового періоду визначається $Q_{test} = 1258 - 250 \cdot 2 = 758$, таким чином прогнозування часового ряду проводиться у діапазоні $\{758, \dots, 1008\}$. Перебір значень затримки k виконується з кроком $S=1$. Звернувшись до таблиці 2, знаходимо оцінку значення $t_p = 5.08$ отримуємо

$$t_{id} = \frac{250}{24} * 5.08 * 28 * \frac{828}{828} = 1481 = 24 \text{ хв.} \quad (3.10)$$

Таким чином, для ідентифікації тимчасового ряду необхідно 4.63 години роботи зазначеного у розділі 3.1. типу персонального комп'ютера.

Якщо проведена оцінка часу ідентифікації занадто висока, то суттєве скорочення часу може бути досягнуто застосуванням паралельних обчислень (розділ 3.2.2).

3.3. Висновки

1) Розроблено метод прогнозування на базі моделі екстраполяції за вибіркою максимальної подоби.

2) Розроблено алгоритми екстраполяції часового ряду з урахуванням та без урахування зовнішніх факторів. Дослідження показали високу швидкість обчислень прогнозних значень.

3) Запропоновано алгоритм ідентифікації параметра моделі М. Алгоритм ідентифікації містить обчислення, які можуть виконуватись паралельно за запропонованою схемою.

4) Проведено оцінку часу послідовних обчислень для розв'язання задачі ідентифікації моделі.

4 РОЗРОБКА ПРОГРАМНОГО ЗАСТОСУВАННЯ ПРОГНОЗУВАННЯ ФІНАНСОВИХ ТИМЧАСОВИХ РЯДІВ НА ОСНОВІ ВИКОРИСТАННЯ РАДІАЛЬНО БАЗИСНИХ ФУНКЦІЙ

4.1 Програмна реалізація

Для створення робочої системи для прогнозування часових рядів було поставлено завдання реалізації алгоритмів прогнозування як самостійного серверного додатка, здатного працювати без втручання експертів. На сьогоднішній день переважна більшість серверів використовують сімейство операційних систем UNIX. При виборі мови програмування та системи управління базами даних (СУБД) бралися до розрахунку такі вимоги:

- Розробка та експлуатація серверного додатку під керуванням широкого набору операційних систем;

- Наявність готових бібліотек, що містять реалізацію відомих математичних функцій, а також бібліотек, що реалізують взаємодія програми з різними джерелами даних, без ліцензійних обмежень;

- Наявність бібліотек для звернення до сервера системи управління базами даних без ліцензійних обмежень.

Для розробки було обрано компіювану мову програмування Python [50], яка

- працює під управлінням найбільшої кількості операційних систем,

- широко застосовується для створення серверних програм та має у відкритому доступі набір необхідних документованих бібліотек без ліцензійних обмежень,

- Найпопулярніша на даний момент мова програмування

Як сервер СУБД був обраний MySQL [51], який також працює з управлінням широкого набору операційних систем та має високу продуктивність. На сьогоднішній день MySQL є найбільш поширеною сервером СУБД, який не має ліцензійних обмежень.

Відповідно до вимог описаних в розділі 2 було створено програмний комплекс, що складається з функціональних блоків, представлених на малюнку 4.1.

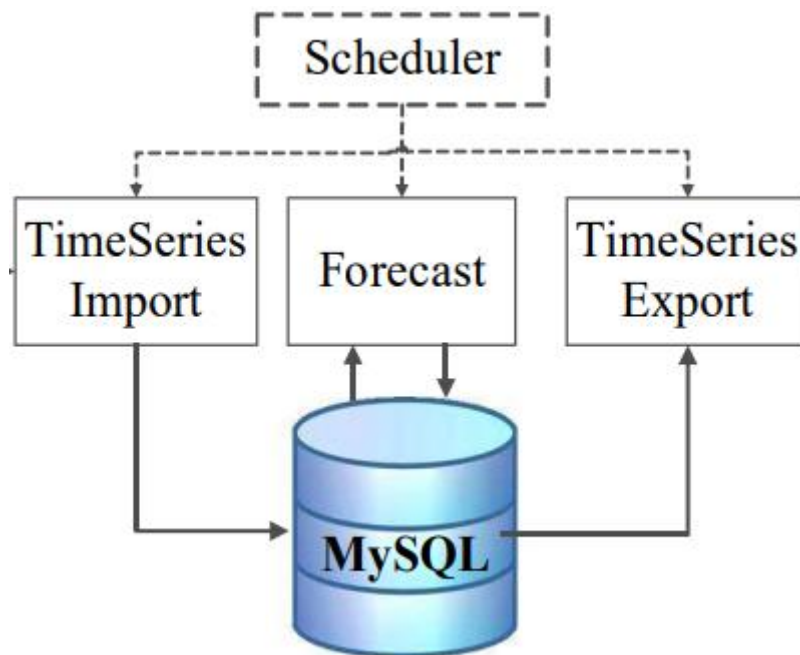


Рисунок 4.1. Структура програмного комплексу для прогнозування часових рядів

Блок «Scheduler» виконує керування процесом щоденного прогнозування. Для кожного часового ряду та кожного часу попередження (добу, тиждень та місяць наперед) створюються три програмні завдання:

- Імпорт фактичних даних тимчасового ряду за минулі добу з XML-макетів із сайту біржі;
- прогнозування часового ряду на моделі авторегресивній моделі за допомогою радіально базисних функцій;
- Експорт прогнозних значень ряду у формат CSV.

Блок «Scheduler» керує запуском за часом та перевіркою коректності завершення кожного завдання в поточній добі, а також створює нові завдання для розрахунку наступної доби. Створена система прогнозування не має графічного інтерфейсу, працює автономно та не вимагає втручання експерта.

Блок "TimeSeries Import" витягує з XML формату фактичні значення для всіх часових рядів енергоспоживання та цін на електроенергію, та завантажує отримані значення тимчасових рядів у базу даних.

Блок "Forecast" виконує прогнозування часових рядів показників за допомогою моделі EMMSP. В рамках цього модуля реалізовано алгоритми

екстраполяції часових рядів з урахуванням та без урахування зовнішніх факторів (розділи 3.1., 3.2.).

4.2 Опис результатів виконання прогнозування даних

Імперично встановлено, що навчання на подібних даних доцільно проводити не менше 30-60 епох.

Це пов'язано з тим, що в разі успішного навчання на малому числі епоху не буде достовірним з точки зору виявлення прихованих патернів і залежностей в даних.

Занадто висока точність на сирих даних на зразок ціни закриття і простих алгоритмах швидше за все буде говорити про перенавчання при підготовці навчальної вибірки.

Проведемо навчання моделі для завдання класифікації і подивимося на графіки епох навчання та середніх похибок (рис.4.1 та 4.2).

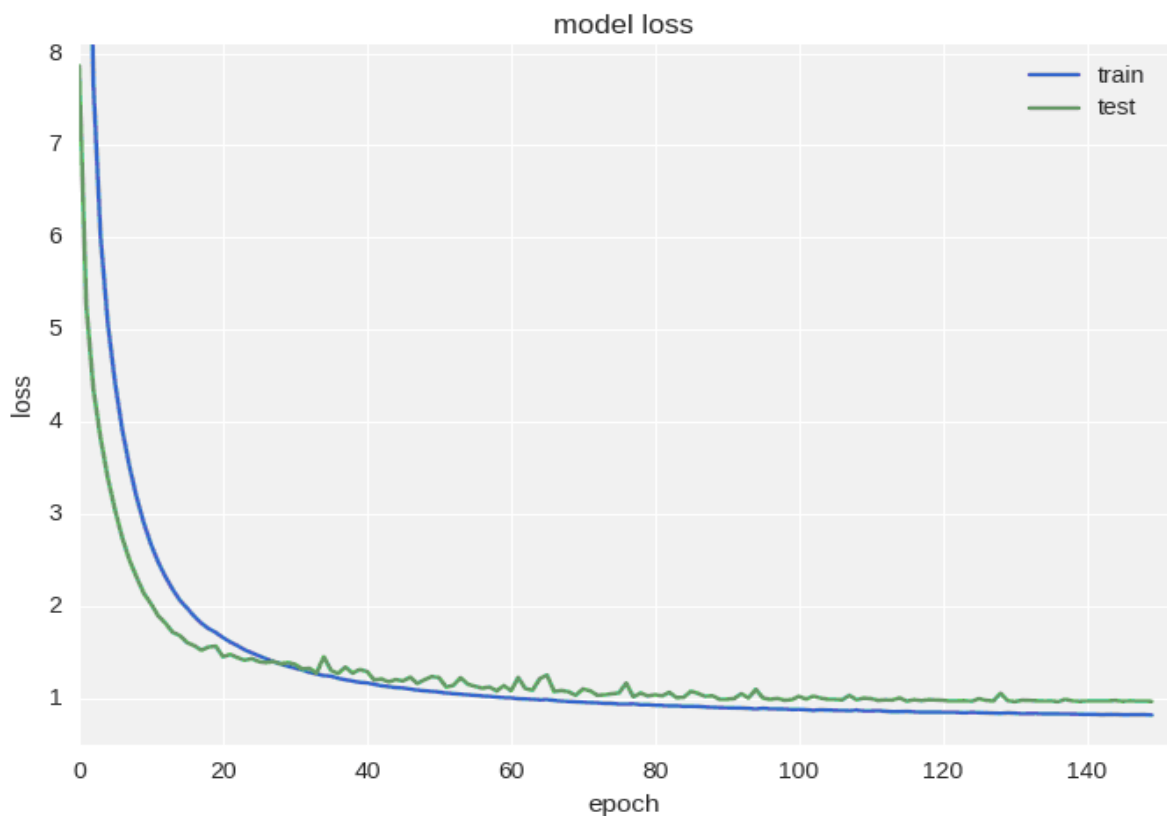


Рисунок 4.1 – Графік відображення помилки у процесі навчання по епохам

Як можна побачити з рисунків вище, помилка і точність для тестової вибірки весь час залишається практично одному значенні, а помилка для тренувальної вибірки знижується, точність при цьому зростає, що є свідченням перенавчання ІНС. Змінимо модель на більш глибоку, з двома шарами(Додаток Б):

В процесі регуляризації ми накладаємо певні обмеження на ваги нейронної мережі, щоб не було великого розкиду в значеннях і не дивлячись на велику кількість ваг мережі частина з них перетворити на нуль для спрощення відповідним чином.

Для цього додамо до функції помилки додаткової складової з L2 нормою за сумою ваг, в Keras це робиться за допомогою `keras.regularizers.activity_regularizer` (Додаток Б).

Після введення даної функції модифікована нейронна мережа проходить процес навчання трохи краще з точки зору функції помилки, урівень точності при цьому є все ще занадто низьким.

Отримані результати побудови графіка відображення помилки у процесу навчання модифікованої мережі з регуляризацією по епохам наведено на рис. 4.2

Результати побудови графіка відображення точності процесу навчання модифікованої мережі з регуляризацією по епохам наведено на рис.4.3 відповідно.

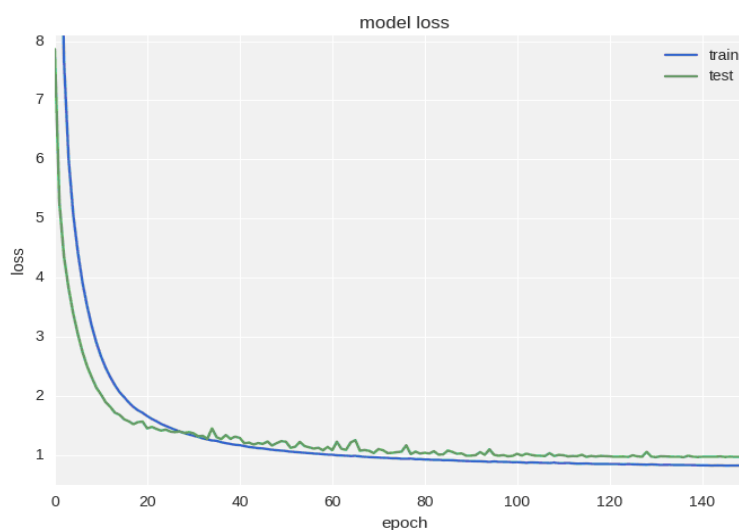


Рисунок 4.2 – Графік відображення помилки у процесу навчання модифікованої ШНМ з регуляризацією по епохам

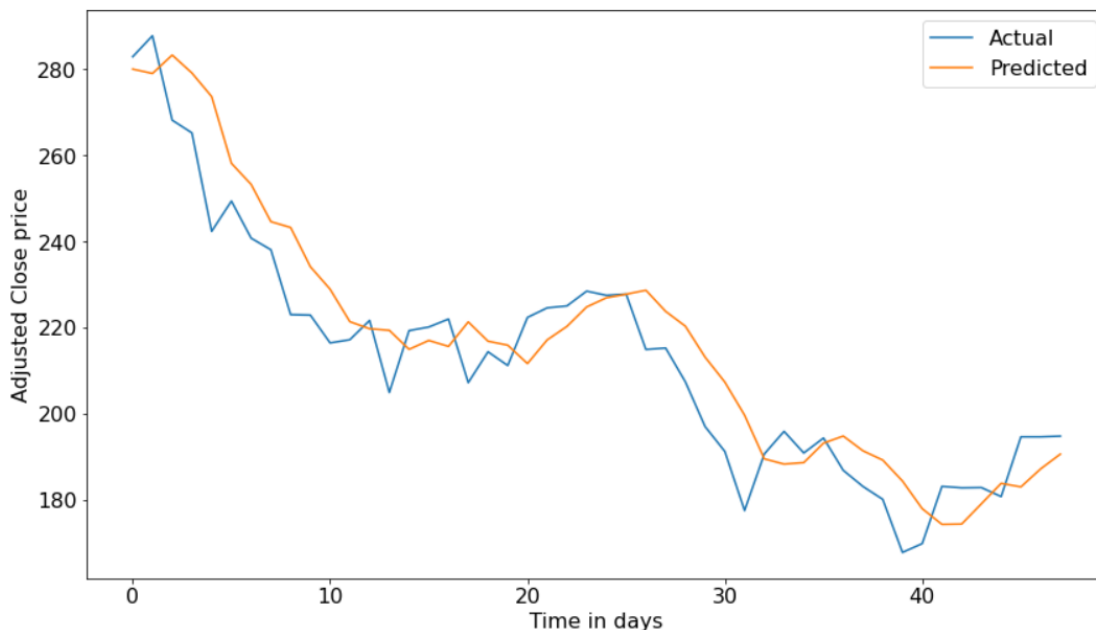


Рисунок 4.3 – Графік відображення точності процесу навчання модифікованої радіально базисної мережі з регуляризацією по епохам

Подібний ефект як зменшення помилки навчання ІНС без зниження точності може зустрічатися при роботі з даними високого ступеня зашумленості або випадкової природи.

Багато в чому, це пояснюється тим, що рівень помилки обчислюється на базі значення cross-entropy, яке може зменшуватися під час того, як точність - це індекс нейрона з правильною відповіддю, який навіть при зміні помилки може залишатися неправильним.

У зв'язку з вищесказаним слід додати ще більше регуляризації в розробленій моделі ІНС шляхом застосування методики Dropout, що є формою випадкового ігнорування значень деяких ваг в процесі навчання, щоб уникнути ко-адаптації нейронів (щоб вони не вивчали однакові ознаки). Доданий програмний код виглядає наступним чином(ДОДАТОК Б).

Згідно із зазначеним значенням параметра методу Dropout між двома прихованими шарами будуть ігноруватися зв'язки під час навчання з імовірністю 50% для кожної ваги. Dropout при цьому не слід додавати між вхідним шаром і першим прихованим, так як в цьому випадку навчання буде проводитися лише на

зашумлених даних. Даний метод також не є ефективним при його використанні перед виходом і не застосовується під час тестування ШНМ.

Графік відображення помилки у процесу навчання модифікованої нейронної мережі на базі використання Dropout по епохам наведено на рис.4.4. Графік відображення точності процесу навчання модифікованої ШНМ на базі використання Dropout по епохам наведено на рис.4.4.

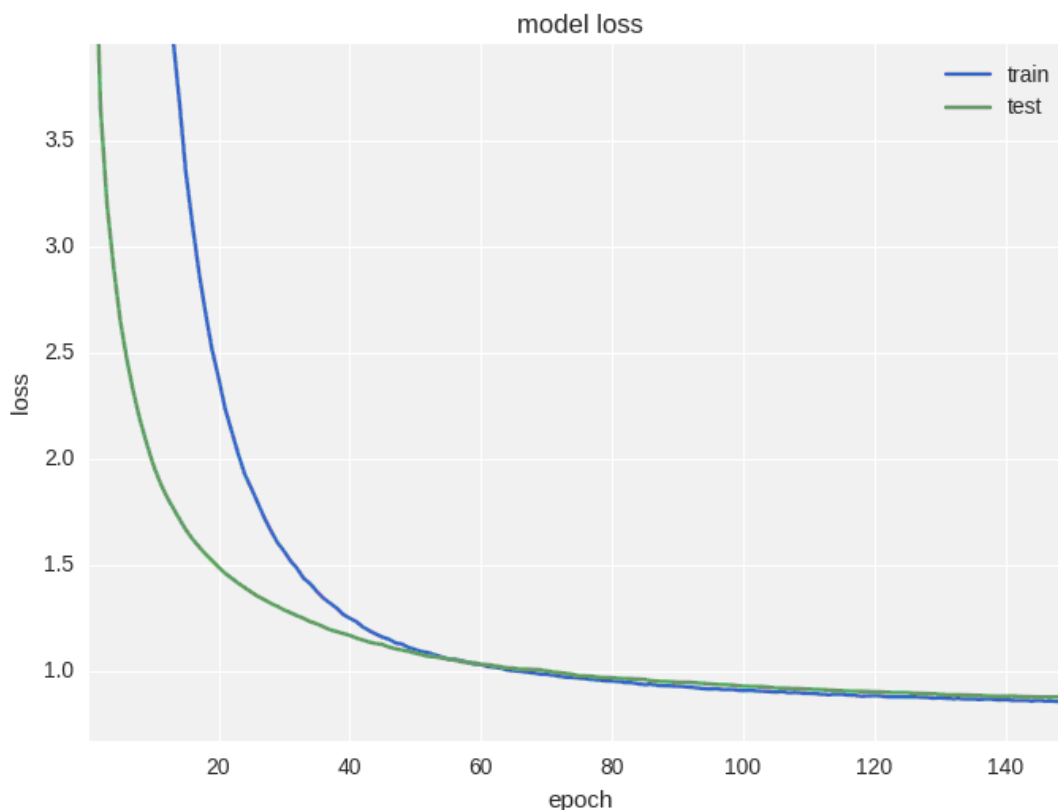


Рисунок 4. 4 – Графік відображення помилки у процесу навчання модифікованої ШНМ на базі використання Dropout по епохам

Як можна побачити з наведених графіків, значення помилки і точності адекватні, якщо зупинити навчання мережі трохи раніше точки перенавчання, то можливо отримати 78% точності прогнозу руху ціни, що є більш точним. Не дивлячись на те, що коливання в наступний день має випадкову природу, аналізуючи графіки-свічки, можливо детектувати досить явний тренд на наступні 5-7 днів. Для прикладу спрогнозуємо рух ціни через 5 днів з останньої вдалої

архітектурою і навчимо ІНС на більшій кількості епох. Графік відображення помилки у процесі прогнозування наведено на рис.4.5.

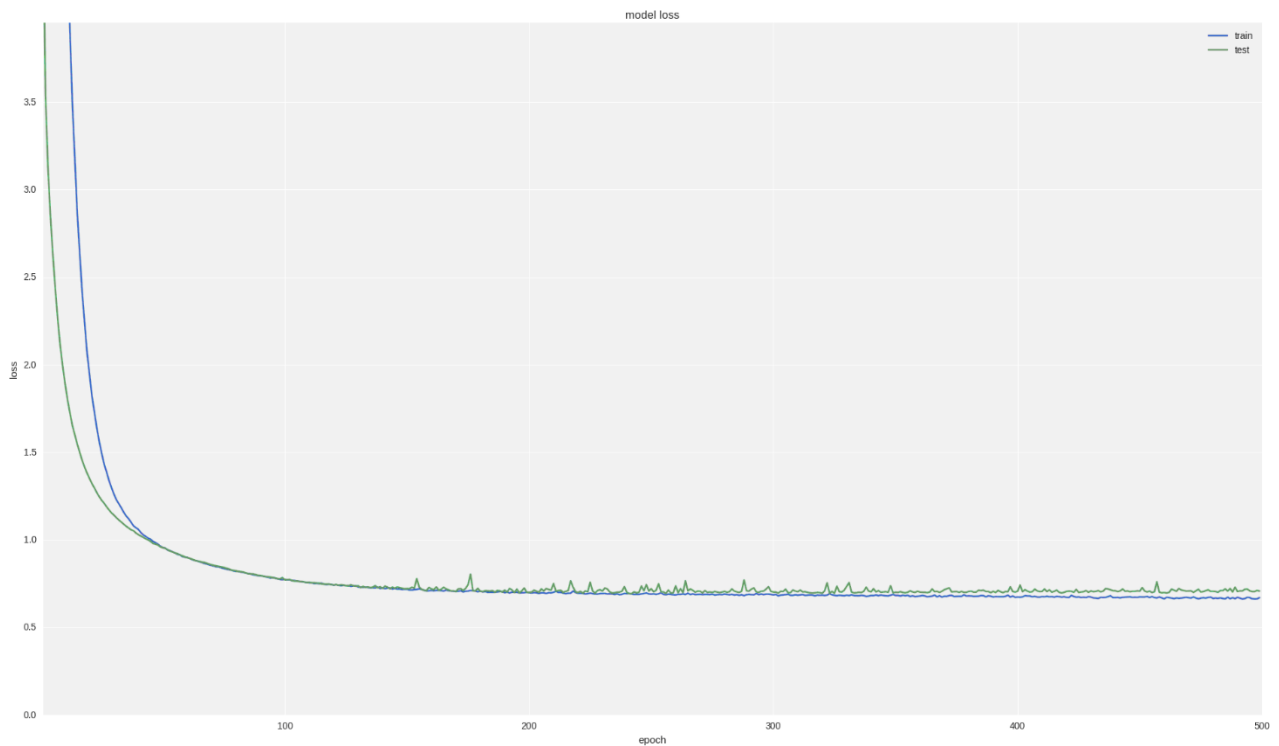


Рисунок 4.5 – Графік відображення помилки у процесі прогнозування

Графік відображення точності у процесі прогнозування наведено на рис.4.6. Як бачимо, якщо ми зупинимо навчання досить рано (з часом все одно настає оверфітінг), то можемо отримати 82% точності.

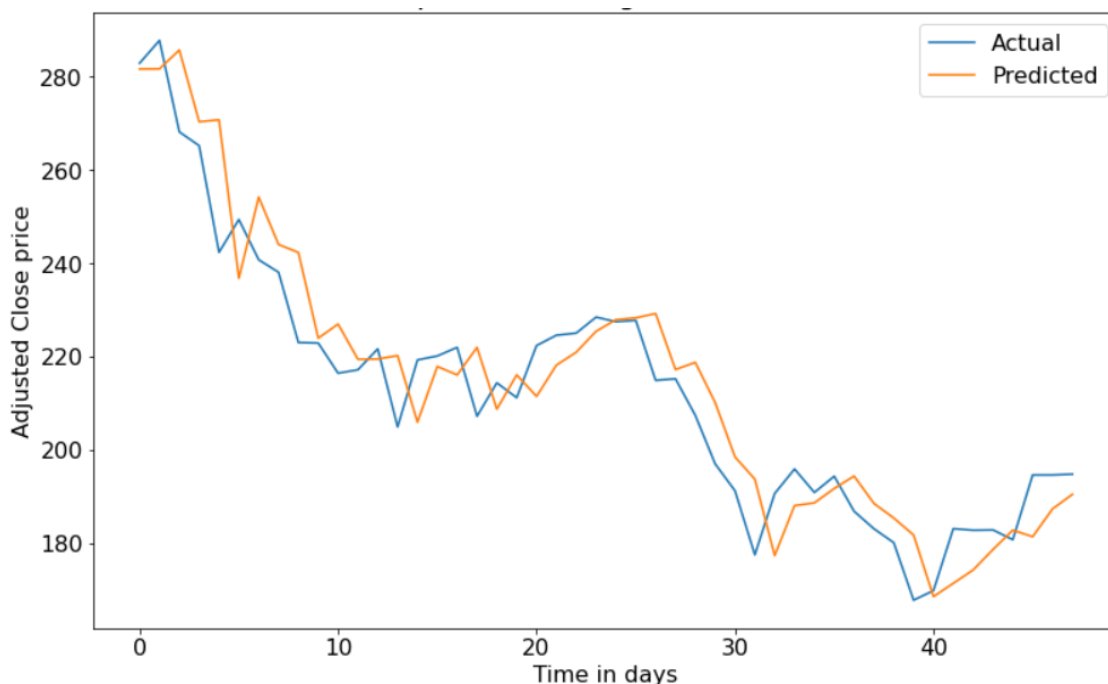


Рисунок 4.6 – Графік відображення точності у процесі прогнозування

На перший погляд, результати класифікації не є високими, але ми навчили найпростіший вид нейронної мережі на одновимірних даних без особливої попередньої обробки. У якості подальшого розвитку дослідження доцільним є виконання ряду кроків, які дозволяють вивести точність на рівень 80-85%:

- Навчати ШНМ на високочастотних даних (кожну годину, кожні п'ять хвилин), тому що більшу кількість даних дозволить сформувати більше патернів і знизити рівень перенавчання.
- Використовувати більше складні архітектури нейронних мереж, які призначені для роботи з послідовностями, наприклад, convolutional neural networks, recurrent neural networks.
- Крім ціни закриття використовувати всі дані з аналізованого набору (high, low, open, close, volume), тобто в кожен момент часу звертати увагу на всю доступну інформацію.
- Оптимізувати процес підбора параметрів, тобто розмір вікна, кількість нейронів в прихованих шарах, крок навчання повинні вибиратися шляхом проведення додаткових оціночних процедур. Це дозволить, наприклад, з'ясувати,

що, можливо, необхідно формувати вікно на 45 днів і проводити навчання ІНС з меншим кроком.

– Використовувати більш підходящі для нашої задачі функції втрат (наприклад, для прогнозування зміни ціни ми могли б штрафувати нейронну за неправильний знак, звичайна MSE до знаку числа інваріантна).

ВИСНОВКИ

1. Здійснено аналіз можливостей та переваг використання радіально-базисної функції для прогнозування часових рядів.
2. Проаналізовано сутності моделі та методів прогнозування часових рядів.
3. Сформовано процес прогнозування та математичну модель вирішення завдання.
4. Проведено аналіз логіко-математичного апарату радіально-базисних функцій для прогнозування часових рядів.
5. Проведено практичне тестування та описано результати навчання штучної нейромережі.
6. Встановлено, що за допомогою радіально-базисних функцій підвищується точність фінансових часових до 82%, що можна побачити в порівнянні рисунку 4.3 та 4.6

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Gheyas I.A., Smith L.S. A Neural Network Approach to Time Series Forecasting // Proceedings of the World Congress on Engineering, London, 2009, Vol 2 [Електронний ресурс]. Р. 1292 – 1296. Режим доступу: www.iaeng.org/publication/WCE2009/WCE2009_pp1292-1296.pdf
2. Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting. 2009, No. 4. P. 213 – 223.
3. Mazengia D.H. Forecasting Spot Electricity Market Prices Using Time Series Models: Thesis for the degree of Master of Science in Electric Power Engineering. Gothenburg, Chalmers University of Technology, 2008. 89 p.
4. Parzen E. Long memory of statistical time series modeling // NBER-NSF Time Series Conference, USA, Davis, 2004 [Електронний ресурс]. 10 p. Режим доступу: <http://www.stat.tamu.edu/~eparzen/Long%20Memory%20of%20Statistical%20Time%20Series%20Modeling.pdf>
5. Jingfei Yang M. Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universität, 2006. 139 p.
6. AAPL Historical Prices | Apple Inc. Stock - Yahoo Finance. – [Електронний ресурс] – Режим доступу: <https://finance.yahoo.com/quote/AAPL/history?period1=1104534000&period2=1491775200&interval=1d&filter=history&frequency=1d>.
7. Norizan M., Maizah Hura A., Zuhaimy I. Short Term Load Forecasting Using Double Seasonal ARIMA Model // Regional Conference on Statistical Sciences, Malaysia, Kelantan, 2010. P. 57 – 73.
8. Collantes-Duarte J., Rivas-Echeverriat F. Time Series Forecasting using ARIMA, Neural Networks and Neo Fuzzy Neurons // WSEAS International Conference

on Neural Networks and Applications, Switzerland, 2002 [Электронный ресурс]. 6 p.
Режим доступа: www.wseas.us/e-library/conferences/switzerland2002/papers/464.pdf

9. Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models / A.J. Conejo [at al.] // IEEE transaction on power systems. 2005, Vol. 20, No. 2. P. 1035 – 1042.

10. Armstrong J.S. Forecasting for Marketing // Quantitative Methods in Marketing. London: International Thompson Business Press, 1999. P. 92 – 119.

11. Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting / J.B. Rundle [at al.] // Colloquium of the National Academy of Sciences, Irvine, USA, 2002. P. 2514 – 2521.

12. Prajakta S.K. Time series Forecasting using Holt-Winters Exponential Smoothing // Kanwal Rekhi School of Information Technology Journal [Электронный ресурс]. 2004. 13 p. Режим доступа:http://www.it.iitb.ac.in/~praj/acads/seminar/04329008_ExponentialSmoothing.pdf

13. Pradhan R.P., Kumar R. Forecasting Exchange Rate in India: An Application of Artificial Neural Network Model // Journal of Mathematics Research. 2010, Vol. 137 2, No. 4. P. 111 – 117.

14. Yildiz B., Yalama A., Coskun M. Forecasting the Istanbul Stock Exchange National 100 Index Using an Artificial Neural Network // An International Journal of Science, Engineering and Technology. 2008, Vol. 46. P.36 – 39.

15. An Artificial Neural Network Approach for Day-Ahead Electricity Prices Forecasting / J. Catalao [at al.] // 6th WSEAS international conference on Neural networks, USA, Stevens Point, 2005. P. 80 – 83.

16. Kumar M. Short-term load forecasting using artificial neural network techniques: Thesis for Master of Science degree in Electrical Engineering. India, Rourkela, National Institute of Technology, 2009. 48 p.

17. Zhu J., Hong J., Hughes J.G. Using Markov Chains for Link Prediction in Adaptive Web Sites // 1st International Conference on Computing in an Imperfect World, UK, London, 2002. P. 60 – 73.
18. Hannes Y.Y., Webb P. Classification and regression trees: A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity // International Food Policy Research Institute [Электронный ресурс]. 1999. 59 p. Режим доступа: http://www.fao.org/sd/erp/toolkit/BOOKS/classification_and_regression_trees_intro.pdf
19. Huang W., Nakamoria Y., Wang S. Forecasting stock market movement direction support vector machine // Elsevier: computers and operation research. 2005, Vol. 32. P. 2513 – 2522.
20. Mahfoud S., Mani G. Financial Forecasting Using Genetic Algorithms // Applied Artificial Intelligence. 1996, Vol. 10, No.6. P. 543 – 565.
21. Nogales F.J., Conejo A.J. Electricity price forecasting through 138 transferfunction models // Journal of the Operational Research Society. 2006, Vol. 57, No. 4. P. 350 – 356.
22. Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. 2002, Vol 33. P. 23 – 34.
23. Hinman J., Hickey E. Modeling and forecasting short term electricity load using regression analysis // Journal of Institute for Regulatory Policy Studies [Электронный ресурс]. 2009. 51 p. Режим доступа: <http://www.irps.ilstu.edu/research/documents/LoadForecastingHinman-HickeyFall2009.pdf>
24. Fogler H.R. A pattern recognition model for forecasting // Management science. 1974, No.8. P. 1178 – 1189.
25. Discovering Patterns in Electricity Price Using Clustering Techniques / F. Martínez Álvarez [et al.] // ICREPQ International Conference on Renewable Energies

and Power Quality, Spain, Sevilla, 2007 [Электронный ресурс]. 8 p. Режим доступа: <http://www.icrepq.com/icrepq07/245-martinez.pdf>

26. Singh S. Pattern Modelling in Time-Series Forecasting // Cybernetics and Systems-AnInternational Journal. 2000, Vol. 31, No. 1. P. 49 – 65. 41. Scherer Perlin M. Nearest neighbor method // Revista Eletrônica de Administração [Электронный ресурс]. 2007, Vol. 13, No. 2. 15 p. Режим доступа: http://read.adm.ufrgs.br/edicoes/pdf/artigo_495.pdf

27. Fernández-Rodríguez F., Sosvilla-Rivero S., Andrada-Félix J. NearestNeighbour Predictions in Foreign Exchange Markets // Fundacion de Estudios de Economia Aplicada [Электронный ресурс]. 2002, No.5. 36 p. Режим доступа: <http://www.fedea.es/pub/Papers/2002/dt2002-05.pdf>

28. Mishra S. Short term load forecasting using computation intelligence methods: Thesis for the degree of Master of technology electronics and communication engineering. India, Rourkela, National Institute Of Technology, 2008. 89 p.

29. Oliva R., Watson N. Managing Functional Biases in Organizational Forecasts: A Case Study of Consensus Forecasting in Supply Chain Planning // HBS Working Paper [Электронный ресурс]. 2006, No.10. P. 7 – 24. Режим доступа: <http://www.hbs.edu/research/pdf/07-024.pdf>

30. Reinaldo C. Garcia A. GARCH Forecasting Model to Predict Day-Ahead Electricity Prices // Workshop of Applied Infrastructure, Germany, Berlin, 2003 141 [электронный ресурс]. 14 p. URL: http://www.wip.tu-berlin.de/typo3/fileadmin/documents/infraday/2003/papers/Contreras-Garcia-Garcia2003-paperGarch_Models_to_Predict_Electricity_Prices.pdf

31. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices / R.C. Garcia [at al.] // IEEE Transactions on Power Systems. 2005, Vol. 20, No. 2. P. 867 – 874.

32. Day-ahead electricity prices forecasting based on time series models: a comparison / R. Espinola [at al.] // 14th Power Systems Computation Conference, Spain, Sevilla, 2002, Session 15, Paper 6 [Электронный ресурс]. 8 p. Режим доступа: http://www.psc-central.org/uploads/tx_ethpublications/s15p06.pdf

33. Zareipour H., Bhattacharya K., Canizares C.A. Forecasting the Hourly Ontario Energy Price by Multivariate Adaptive Regression Splines // IEEE Power Engineering Society General Meeting, Canada, Montreal, 2006. 7 p.

34. Zareipour H. Price Forecasting and Optimal Operation of Wholesale Customers in a Competitive Electricity Market: Thesis for Ph.D degree. Canada, Ontario, 2006. 201 p.

35. Bunnoon P., Chalermyanont K., Limsakul C. A Computing Model of Artificial Intelligent Approaches to Mid-term Load Forecasting: a state-of-the-artsurvey for the researcher // IACSIT International Journal of Engineering and Technology. 2010, No.1. P. 94 – 100.

36. Basaran Filik U.,Kurban M. A New Approach for the Short-Term Load Forecasting with Autoregressive and Artificial Neural Network Models // International Journal of Computational Intelligence Research. 2007, No.3. P. 66 – 71.

37. Taylor J.W., McSharry P.E. Short-Term Load Forecasting Methods: An Evaluation Based on European Data // IEEE Transactions on Power Systems 142 2008, Vol.22. P. 2213 – 2219.

Додаток А



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ



Кафедра інженерії програмного забезпечення

МАГІСТЕРСЬКА РОБОТА «РОЗРОБКА МЕТОДУ ПРОГНОЗУВАННЯ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ НА ОСНОВІ МЕРЕЖІ РАДІАЛЬНО-БАЗИСНИХ ФУНКЦІЙ»

Виконав: студент групи ПДМ – 61, Кривошапка Дмитро Олександрович

Керівник: к.т.н., доц. Трінтіна Наталія Альбертівна

Київ - 2022

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

2

Мета роботи: підвищення ефективності прогнозування фінансових часових рядів для біржових торгів за допомогою радіально базисних функцій

Об'єкт дослідження: прогнозування фінансових часових рядів

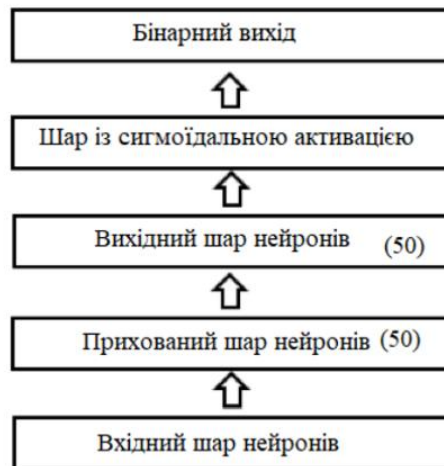
Предмет дослідження: мережі радіально-базисних функцій та моделі експоненційного згладжування

Порівняння існуючих моделей та методів прогнозування

3

Модель та метод	Переваги	Недоліки
Регресійні моделі та методи	Простота, прозорість моделювання, гнучкість, однаковість аналізу та проектування	складність визначення залежності, трудомісткість знаходження коефіцієнтів залежності
<u>Авторегресійні моделі та методи</u>	простота, однаковість аналізу та проектування; безліч прикладів застосування	<u>Ресурсомісткість</u> неможливість моделювання <u>нелінійностей</u> , низька адаптивність
<u>Моделі та методи експоненційного згладжування</u>	простота моделювання; однаковість аналізу та проектування	Недостатня гнучкість, вузька сфера використання моделей

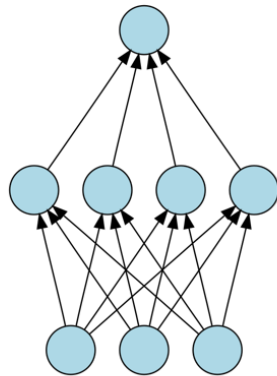
Структура неймережі



4

Математична модель радіально базисної мережі

Лінійний вихідний рівень



Вихід y

Лінійні ваги

Радіальні базисні функції

Ваги

Вхід x

Вхід можна моделювати як вектор дійсних чисел x . Вихід мережі тоді, є скалярною функцією вхідного вектора φ та має вигляд:

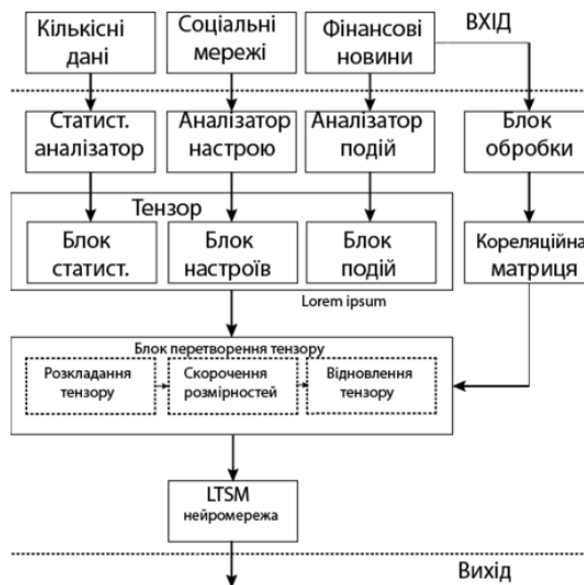
$$\varphi(x) = \sum_{i=1}^N a_i \rho(\|x - c_i\|)$$

де N - кількість нейронів у прихованому шарі, c_i є центральним вектором для нейрона i , та a_i - це вага нейрона i в лінійному виході нейронів.

Вхідний шар

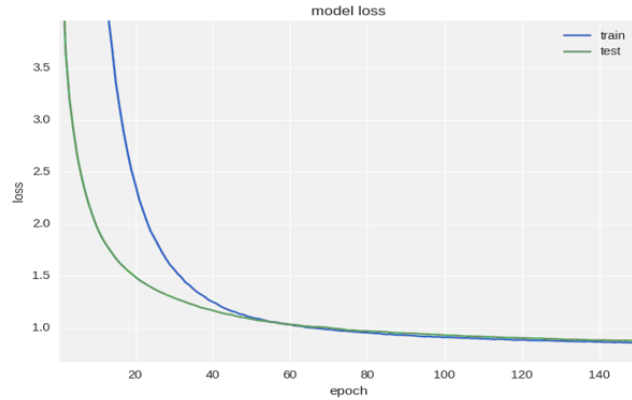
5

Діаграма роботи системи



6

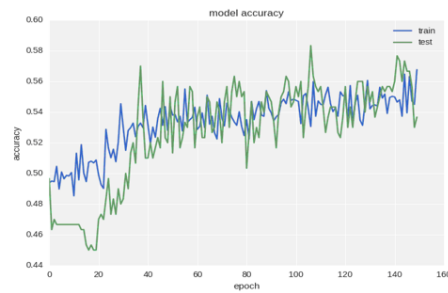
Різниця точності прогнозування в залежності від кількості епох навчання нейронної мережі



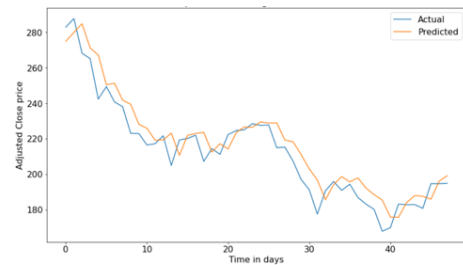
На графіку зображена точність прогнозування в залежності від кількості пройдених епох навчання

РЕЗУЛЬТАТИ ПРОГНОЗУВАННЯ

9



Графік компанії Amazon (20.01.2022 – 20.10.2022)



Графік компанії Amazon (20.01.2022 – 20.10.2022)

ВИСНОВКИ

10

1. Здійснено аналіз можливостей та переваг використання радіально-базисної функції для прогнозування часових рядів.
2. Проаналізовано сутності моделі та методів прогнозування часових рядів.
3. Сформовано процес прогнозування та математичну модель вирішення завдання.
4. Проведено аналіз логіко-математичного апарату радіально-базисних функцій для прогнозування часових рядів.
5. Проведено практичне тестування та описано результати навчання штучної нейромережі.
6. Встановлено, що за допомогою радіально-базисних функцій підвищується точність фінансових часових

ПУБЛІКАЦІЇ ТА АПРОБАЦІЯ РОБОТИ

11

Тези доповідей на конференціях:

1. Кривошапка Д.О. ПЕРЕВАГИ ТА НЕДОЛКИ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ // XV НАУКОВО-ТЕХНІЧНА КОНФЕРЕНЦІЯ «СУЧАСНІ ІНФОКОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ». – Київ: ДУТ, 2022. (робота ще не опублікована, але подана до друку)
2. Кривошапка Д.О. МЕТОДИ ПОКРАЩЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ «ПРОБЛЕМИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ». – Київ: ДУТ, 2022. (робота ще не опублікована, але подана до друку)

ДЯКУЮ ЗА УВАГУ!

ДОДАТОК Б

```
model = Sequential()
model.add(Dense(64, input_dim=30))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dense(16))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dense(2))
model.add(Activation('softmax'))
```

```
model = Sequential()
model.add(Dense(64, input_dim=30,
activity_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dense(16,
activity_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dense(2))
model.add(Activation('softmax'))
```

```
model = Sequential()
model.add(Dense(64, input_dim=30,
activity_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dropout(0.5))
model.add(Dense(16,
activity_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(LeakyReLU())
model.add(Dense(2))
model.add(Activation('softmax'))
```