Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Нейромережевий підхід до прогнозування фінансового ринку та побудови інвестиційного портфеля
(тема)

Виконав:
студент 2 курсу, групи СШІм-20-2
Пахомов І. Ю.
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
(повна назва спеціалізації)

Керівник проф. Рябова Н. В.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ В.О. Філатов
(підпис) (прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Штучного інтелекту
(повна назва)

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«» _____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Пахомову Івану Юрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Нейромережевий підхід до прогнозування фінансового ринку та побудови інвестиційного портфеля
затверджена наказом університету від 24 березня 20 22 р. № 414 Ст
2. Термін подання студентом роботи до екзаменаційної комісії 12 травня 20 22 р.
3. Вихідні дані до роботи адаптована нейронна мережа на основі довготривалої короткочасної пам'яті (LSTM), алгоритм для короткострокового прогнозування цінового тренду на фондовому ринку, модель прогнозування цінового тренду з використанням LSTM, набір даних з міжнародного ринку для тренування та тестування системи
4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз предметної галузі, аналіз існуючих методів та підходів, постановка задачі, проектування та реалізація технічного проекту, підготовка набору даних, розробка алгоритму, розробка нейронної мережі на основі довготривалої короткочасної пам'яті та її навчання, експериментальна оцінка ефективності методів та порівняння з аналогічними роботами

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Рисунок 1 – Модель імітування ринку, Рисунок 2 – Високорівнева архітектура пропонованого рішення, Рисунок 3 – Структура даних для витягнутого набору даних, Рисунок 4 – Таблиця огляду набору даних з різними категоріями і підмножинами полів, Рисунок 5 – Детальний технічний дизайн пропонованого рішення, Рисунок 7 – Опис алгоритмів, Рисунок 8 – Матриця плутанини, Рисунок 9 – Взаємозв'язок між кількістю головних компонент і ефективністю навчання, Рисунок 10 – Крива навчання пропонованого рішення, Рисунок 11 – Взаємозв'язок між кількістю ознак і часом навчання, Рисунок 12 – Матриці порівняння-конфузії передбачень моделей, Рисунок 13 – Порівняння пропонованого рішення з аналогічними роботами, Рисунок 14 – Порівняння продуктивності пропонованої моделі – з PCA та без нього.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	28.03.2022	виконано
2	Аналіз предметної області та постановка задачі	01.04.2022 – 03.04.2022	виконано
3	Дослідження методів прогнозування фінансового		
	ринку	04.04.2022 – 08.04.2022	виконано
4	Розробка системи прогнозування фінансового		
	ринку	09.04.2022 – 15.04.2022	виконано
5	Аналіз та тестування розробленої системи	16.04.2022 – 21.04.2022	виконано
6	Написання пояснювальної записки	22.04.2022 – 29.04.2022	виконано
7	Нормоконтроль	30.04.2022 – 03.04.2022	виконано
8	Перевірка на академічний плагіат	04.05.2022 – 05.04.2022	виконано
9	Підготовка презентації та доповіді	06.05.2022	виконано
10	Попередній захист	07.05.2022	виконано
11	Рецензування	09.05.2022	виконано
12	Захист перед ЕК	12.05.2022	

Дата видачі завдання 28 березня 20 22 р.

Студент _____
(підпис)

Керівник роботи _____ _____
(підпис) (підпис)

РЕФЕРАТ

Пояснювальна записка: 69 с., 22 рис., 1 дод., 20 джерел.

АНАЛІЗ ДАННИХ, ГЛИБИННЕ НАВЧАННЯ, ДОВГОТРИВАЛА КОРОТКОЧАСНА ПАМ'ЯТЬ, ІНВЕСТИЦІЇ, НЕЙРОННА МЕРЕЖА, ПРОГНОЗУВАННЯ, ФІНАНСОВИЙ РИНОК, ЧАСОВІ РЯДИ

Об’єкт дослідження – система прогнозування фінансового ринку та побудови інвестиційного портфеля на основі нейронних мереж.

Мета даної роботи – дослідження та аналіз існуючих методів побудови системи прогнозування фінансового ринку, а також реалізація нового методу на основі проведених досліджень.

Предмет дослідження – методи побудови систем прогнозування фінансового ринку.

В результаті проведених досліджень, а також використанні, комбінуванні та адаптації методів попередньої обробки даних, розробки ознак і глибокого навчання, було вирішено задачу прогнозуванні короткострокових цінових тенденцій. Отриманні результати використовуються у побудові системи прогнозування фінансового ринку за допомогою нейромережевого підходу, а саме моделі LSTM.

Запропонована система є актуальною та може бути корисною при вирішенні задач в багатьох галузях, де використовуються прогнозування часових рядів та короткострокових цінових тенденцій.

ABSTRACT

Explanatory note: 69 p., 22 fig., 1 an., 20 sources.

DATA ANALYSIS, DEEP LEARNING, LONG SHORT TERM MEMORY, STOCKS, NEURAL NETWORK, FORECASTING, FINANCIAL MARKET, TIME SERIES

The object of the research is the system of forecasting the financial market and building an investment portfolio based on neural networks.

The goal of this work is to investigate and analyze existing methods of creating a system of forecasting the financial market, as well as the implementation of a new method on the basis of the conducted research.

The subject of research – methods of creating systems of forecasting financial market.

As a result of the research, as well as using, combining and adapting methods of advanced data processing, development of indicators and extensive training, the problem of predicting short-term price trends was solved. The obtained results are used to build a system of financial market forecasting based on the neuro-mechanical approach, namely the LSTM model.

The proposed system is relevant and can be useful for solving problems in many industries that use

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНИХ ТЕРМІНІВ

ANN – Artificial Neural Network – штучна нейронна мережа;

CNN – Convolutional Neural Network – конволюционная нейронна мережа;

DNN – Deep Neural Network – глибока нейронна мережа;

DTW – Dynamic Time Warping – динамічне викривлення часу;

LSTM – Long Short Time Memory – довготривала короткочасна пам'ять;

MLP – Many Layer Perceptron – багатошаровий персептрон;

PCA – Principal components analysis – аналіз головних компонент;

ReLU – Rectified Linear Units – випрямлена лінійна одиниця;

RFE – Recursive Feature Elimination – рекурсивне усунення ознак**;**

RNN – Recurrent neural network – рекурентні нейронні мережі;

SGD – Stochastic Gradient Descent – стохастичний градієнтний спуск;

SVM – Support Vector Machine – машина опорних векторів.

ВСТУП

Людина завжди прагне полегшити своє життя. Переважне уявлення у суспільстві є те, що багатство приносить комфорт та розкіш, тому не дивно, що було зроблено так багато роботи над способами прогнозування ринків. Різні технічні, фундаментальні та статистичні індикатори були запропоновані та використовуються з різними результатами. Проте жодна техніка чи комбінація технік була досить успішною, щоб послідовно обіграти ринок.

Нейронні мережі – це метод штучного інтелекту для моделювання складних цільових функцій. Для певних типів завдань, таких як навчання інтерпретації складних даних датчиків реального світу, штучні нейронні мережі (ІНС) є одними з найефективніших методів навчання, відомих у цей час. Протягом останнього десятиліття вони широко застосовувалися у сфері прогнозування фінансових тимчасових рядів, та його значення у цій галузі постійно зростає.

Метою даної роботи є аналіз нейронних мереж для прогнозування фінансового ринку. Розглянути існуючі аналітичні та комп'ютерні підходи для прогнозування, а також, як саме нейромережі вирішують проблеми у перерахованих вище методах. Також побудувати сучасну модель прогнозування для прогнозування цінового тренду, яка фокусується на короткостроковому прогнозуванні цінового тренду.

Прогнозування фінансових часових рядів, як відомо, є непростим завданням через загальноприйняту форму ефективності ринку та високий рівень шуму. Ще в 2003 році Ванг застосував штучні нейронні мережі для прогнозування цін на фондовому ринку і зосередився на обсязі, як специфічній характеристиці фондового ринку. Одним з ключових висновків, зроблених ними, було те, що обсяг не був визнаний ефективним для поліпшення прогнозування на наборах даних, що використовуються, а саме S&P 500 і DJI. Інші розробники націлилися на короткострокове

прогнозування і застосували модель машини опорних векторів (SVM) для прогнозування цін на акції. Їхній основний внесок полягає в проведенні порівняння між багатошаровим перцептроном (MLP) і SVM, після чого було встановлено, що в більшості сценаріїв SVM перевершує MLP, при цьому результат також впливають різні торгові стратегії.

Методи оптимізації, такі як аналіз основних компонентів (PCA), також застосовувався для короткострокового прогнозування цін на акції. Протягом багатьох років дослідники не тільки фокусувалися на аналізі цін на акції, але й намагалися аналізувати операції на фондовому ринку, такі як обсяг та ризики, що розширює область дослідження аналізу фондового ринку та показує, що ця область досліджень усе ще має великий потенціал.

У міру розвитку методів штучного інтелекту в останні роки багато запропонованих рішень намагалися об'єднати методи машинного навчання та глибокого навчання на основі попередніх підходів, а потім запропонували нові метрики, які служать як навчальні характеристики. Цей тип попередніх робіт відноситься до галузі інженерії ознак і може розглядатись як джерело натхнення для ідей розширення ознак у нашому дослідженні. У цій роботі дослідники запропонували згорткову нейронну мережу (CNN), а також модель на основі нейронної мережі з довготривалою пам'яттю (LSTM) для аналізу різних кількісних стратегій на фондових ринках. CNN служить для стратегії вибору акцій, автоматично отримує ознаки на основі кількісних даних, а потім слідує за LSTM, щоб зберегти ознаки тимчасового ряду для підвищення прибутку.

На підставі всіх вищезгаданих робіт ми можемо зробити три ключові внески до нашого проекту:

1) новий набір даних, витягнутий та очищений;

2) комплексна розробка ознак;

3) адаптована нейрона мережа на основі довготривалої короткочасної пам'яті (LSTM).

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Вступ до фінансового прогнозування

Фінансове прогнозування – це процес, обробка, оцінка або прогноз майбутніх показників бізнесу. За допомогою фінансового прогнозу ви намагаєтеся передбачити, як бізнес виглядатиме у фінансовому плані в майбутньому.

Найпоширенішим прикладом складання фінансових прогнозів є прогнозування доходів компанії. Показники продажів зрештою визначають, в якому етапі перебуває (комерційна) організація. Тому вони є важливими показниками до прийняття правильних рішень, які сприяють досягненню організаційних цілей.

Іншими важливими аспектами фінансового прогнозування є прогнозування інших доходів, майбутніх постійних та змінних витрат, а також капіталу.

Для складання прогнозів застосовуються історичні дані про результати діяльності. Вони допомагають передбачити майбутні тенденції. Компанії та підприємці використовують фінансове прогнозування, щоб визначити, як розподілити свої ресурси, чи будуть очікувані витрати на певний період.

Інвестори використовують фінансове прогнозування, щоб визначити, чи вплинуть певні події на акції компанії. Інші аналітики використовують прогнози для екстраполяції того, як зміняться такі тенденції, як ВНП чи безробіття наступного року. Що далі за часом, то менш точним буде прогноз.

У 1996 році Ясер С. Абу-Мостафа представив короткий вступ у прогнозування на фінансових ринках з акцентом на товарні ф'ючерси та іноземну валюту. Він пояснив основи прогнозування та шумну природу фінансових даних.

Дані фінансового ринку дуже галасливі. За словами Яссера, розглядайте ринок як систему, яка приймає велику кількість інформації (фундаментальні показники, події новин, чутки, хто колись що купив і т.д.) і виробляє вихідну інформацію. Модель, така як нейронна мережа, намагається імітувати ринок, але вона приймає x, що є лише невеликим підмножиною інформації.

На рисунку 2.1 зображена модель імітування ринку.

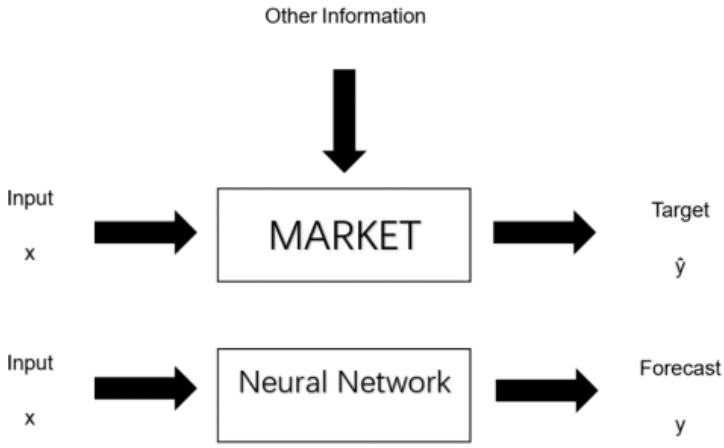


Рисунок 2.1 – Модель імітування ринку

Ясер запропонував один із можливих способів мінімізації шуму – використання дуже великого навчального набору даних. Проте, через не стаціонарність фінансових даних, старі дані можуть бути зовсім інші моделі, ніж нові дані. Щоб подолати цю труднощі, зокрема, збільшуючи розмір набору даних, але обмежуючи тимчасову довжину у встановлених межах.

1.2 Типи фінансового прогнозування

Кількісні прогнози використовують для аналізу великої кількості історичних даних виявлення тенденцій і закономірностей. Кількісні

прогнози, як правило, менш схильні до спотворень, ніж спекулятивні прогнози.

Однак якщо історичних даних не так багато, кількісний метод стає менш ефективним. Тому кількісні та спекулятивні прогнози часто використовуються у тандемі.

Нижче ми розглянемо приклади кількісних методів прогнозування.

Фінансова звітність – в основному використовують показники продажу та очікувані витрати попередніх років як основу для складання прогнозів.

Аналіз часових рядів – це популярний метод кількісного прогнозування, який передбачає збирання даних за певний період з метою виявлення тенденцій. Аналіз часових рядів є одним із найпростіших способів і може бути досить точним, особливо в короткостроковій перспективі.

Причина-наслідок – прогнозист шукає причинно-наслідкові зв'язки змінних з іншими змінними, такими як зміна доходу споживачів, рівень споживчої довіри, процентні ставки, рівень безробіття і т.д. Цей метод використовує тимчасові ряди з минулого для багатьох відповідних змінних, на основі яких будується прогноз.

Якісні прогнози (спекуляція) – це те, що робиться на основі інтуїції та досвіду. Людський розум здатний бачити зв'язки між подіями та розуміти контекст так, як це не можуть зробити комп'ютери.

Однак люди також схильні до певних упереджень, які ускладнюють обробку та аналіз великої кількості даних. Спекулятивні прогнози найкраще використати у малому бізнесі з невеликою кількістю або повною відсутністю історичних даних.

Нижче ми розглянемо приклади якісних методів прогнозування.

Експертні думки та бачення – при використанні цього методу для складання прогнозу збираються думки та думки ключових співробітників таких відділів, як виробництво, продаж, закупівля та операційна діяльність.

Референтні прогнози – цей метод полягає у прогнозуванні результатів запланованих дій на основі аналогічних сценаріїв з інших часових періодів чи місць. Ці прогнози ґрунтуються виключно на людському судженні.

Метод Дельфі – для методу Дельфі створюється серія анкет, які заповнюються групою експертів незалежно друг від друга. Після збору результатів першої анкети створюється друга, що базується на результатах першої. Другий документ знову надається експертам, яких просять заново оцінити відповіді, які вони дали в першій анкеті. Цей процес повторюється доти, доки дослідники не дійдуть загального списку загальноприйнятих думок.

Дослідження споживачів – компанії часто проводять маркетингові дослідження серед споживачів. Збір даних здійснюється, наприклад, за допомогою дзвінків, інтерв'ю, анкетування або вибіркового тестування. Величезний обсяг інформації, отриманий у результаті, піддається аналізу складання прогнозів.

Сценарні прогнози – у цьому методі прогнозист генерує різні результати, засновані на результатах різних сценаріїв. Керівництву компанії залишається вирішити, який із безлічі сценаріїв є найімовірнішим.

1.3 Мотиви прогнозування фінансового ринку

Існує кілька мотивів для прогнозування ціни на фондовому ринку. Найголовніша з них – це фінансова вигода. Будь-яка система, яка може послідовно вибирати переможців на динамічному ринку, зробить власника системи дуже багатим. Тому багато людей, включаючи дослідників, професіоналів у галузі інвестицій, і звичайних інвесторів постійно шукають цієї чудової системи, яка принесе їм високий дохід.

У дослідницькому та фінансовому співтовариствах існує й друга мотивація. Вона була запропонована в Гіпотезі ефективного ринку (ГЕР), за якою ринки є ефективними в тому сенсі, що можливості для отримання

прибутку виявляються настільки швидко, що перестають бути можливостями. ГЕР фактично стверджує, що жодна система не може постійно вигравати у ринку, тому що якщо ця система стане загальнодоступною, всі будуть використовувати її, зводячи нанівець потенційну вигоду. З приводу обґрунтованості ГЕР постійно точаться суперечки, і деякі дослідники намагалися використовувати нейронні мережі для підтвердження своїх тверджень. Єдиної думки про достовірність ГЕР немає, але багато спостерігачів ринку схильні вірити в її слабкі форми і тому часто не бажають ділитися власними системами інвестування.

Нейронні мережі використовуються для прогнозування цін на фондовому ринку, оскільки вони здатні вивчати нелінійні відображення між входами та виходами. На противагу ГЕР деякі дослідники стверджують, що фондовий ринок та інші складні системи демонструють хаос. Хаос – це нелінійний детермінований процес, який здається випадковим лише тому що він не може бути легко виражений. Завдяки здатності нейронних мереж до навчання нелінійних, хаотичних систем, можливо, вдасться перевершити традиційний аналіз та інші комп'ютерні методи. Крім прогнозування фондового ринку, нейронні мережі були навчені для виконання різних фінансових завдань, пов'язаних з фінансами. Існують експериментальні та комерційні системи, що використовуються для відстеження товарних ринків та ф'ючерсів, торгівлі іноземною валютою, фінансового планування, стабільності компанії та прогнозування банкрутства. Банки використовують нейронні мережі для сканування заявок на кредити та позики, щоб оцінити ймовірність банкрутства, у той час як грошові менеджери можуть використовувати нейронні мережі для планування та побудови прибуткових портфелів у режимі реального часу.

2 АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА ПІДХОДІВ ДО ПРОГНОЗУВАННЯ ФІНАНСОВОГО РИНКУ

2.1 Аналітичні та комп'ютерні методи

До епохи комп'ютерів люди торгували акціями та товарами, покладаючись переважно на інтуїцію. У міру того як рівень інвестування та торгівлі зростав, люди шукали інструменти та методи, які дозволили б їм збільшити прибуток та одночасно мінімізації ризику. Статистика, технічний аналіз, фундаментальний аналіз та лінійна регресія – всі вони використовуються для того, щоб спробувати передбачити напрямок руху ринку та отримати з цього вигоду. Жоден із цих методів не довів, що є незмінно вірним інструментом прогнозування, і багато аналітиків сперечаються про користь багатьох із них. Тим не менш, ці методи представлені в тому вигляді, в якому вони зазвичай використовуються на практиці, і є базовим стандартом, за яким нейронні мережі повинні перевершувати інші. Крім того, багато з цих методів використовуються для попередньої обробки вихідних даних, а їх результати подаються в нейронні мережі як вхідні дані.

2.1.1 Технічний аналіз

Ідея технічного аналізу полягає в тому, що ціни на акції рухаються відповідно до тенденцій, продиктованих постійно мінливими відносинами інвесторів у відповідь на різні фактори. Використовуючи статистику цін, обсягів та відкритого інтересу, технічний аналітик технічний аналітик використовує графіки для прогнозування майбутнього руху акцій. Технічний аналіз ґрунтується на припущенні, що історія повторюється і що майбутній напрямок ринку можна визначити, вивчивши ціни минулих років.

Таким чином, технічний аналіз є спірним та суперечить гіпотезі ефективного ринку. Тим не менш, він використовується приблизно на 90% великих біржових трейдерів.

Незважаючи на широке поширення, технічний аналіз піддається критиці, оскільки він дуже суб'єктивний. Різні люди можуть інтерпретувати графіки по-різному. Графіки цін застосовуються виявлення тенденцій. Передбачається, що тенденції засновані на попиті та пропозиції, які часто мають циклічний або помітний характер.

Існує безліч технічних індикаторів, отриманих в результаті аналізу графіків, які можуть бути формалізовані в торгові правила або використані як вхідні дані для нейронних мереж. Деякі категорії технічних індикаторів включають індикатори фільтрації, індикатори імпульсу, аналіз ліній тренду, теорію циклів, індикатори обсягу, хвильовий аналіз та аналіз патернів. Індикатори можуть надавати короткострокову або довгострокову інформацію, допомагати виявляти тенденції чи цикли на ринку або вказувати на силу ціни акції за допомогою рівнів підтримки та опору.

Прикладом технічного індикатора є ковзна середня. Змінна середня ціни на акції за певний період часу, що дозволяє краще побачити тенденції. Було розроблено кілька торгових правил, що належать до ковзної середньої. Наприклад, «коли ціна закриття рухається вище ковзної середньої, генерується сигнал купівлі». На жаль, ці індикатори часто дають помилкові сигнали та відстають від ринку.

Тобто, оскільки ковзна середня – це оцінка минулого, технічний трейдер часто втрачає великий потенціал руху акцій до того, як буде згенерований відповідний торговий сигнал.

Таким чином, хоча технічний аналіз може дати уявлення про ринок, його вкрай суб'єктивний характер і властива йому тимчасова затримка не роблять його ідеальним для сучасних швидких та динамічних торгових ринків.

2.1.2 Фундаментальний аналіз

Фундаментальний аналіз включає глибокий аналіз ефективності і прибутковості компанії для визначення ціни її акцій. Вивчаючи загальні економічні умови, конкуренцію компанії та інші фактори, можна визначити очікувану прибутковість та внутрішню вартість акцій. Даний вид аналізу передбачає, що поточна (і майбутня) ціна акції залежить від її внутрішньої вартості та очікуваного прибутку від інвестицій. У міру появи нової інформації про стан компанії, очікувана прибутковість акцій зміниться, що вплине на ціну акцій.

Перевагами фундаментального аналізу є його систематичний підхід та здатність передбачати зміни до того, як вони з'являться на графіках. Компанії порівнюються одна з одною, і перспективи їхнього зростання співвідносяться з поточною економічною ситуацією. Це дозволяє інвестору краще пізнати компанію. На жаль, стає складніше формалізувати всі ці знання з метою автоматизації, а інтерпретація цих знань може бути суб'єктивною. Також важко визначити час ринку за допомогою фундаментального аналізу. Хоча видатна інформація може вимагати рух акцій, фактичний рух може бути відкладено через невідомі фактори або до тих пір, поки решта ринку інтерпретує інформацію таким же чином. По суті, фундаментальний аналіз передбачає, що інвестори на 90% логічні, детально вивчаючи свої інвестиції, тоді як технічний аналіз передбачає, що інвестори на 90% психологічні, реагують на зміни у ринковому середовищі передбачуваним чином.

2.1.3 Традиційне прогнозування часових рядів

Прогнозування часових рядів аналізує минулі дані та прогнозує оцінки майбутніх значень даних. По суті, цей метод намагається

змоделювати нелінійну функцію за допомогою рекурентної залежності, отриманої з попередніх значень.

Отримана рекурентна залежність може бути використана для прогнозування нових значень тимчасового ряду, які, як можна сподіватися, будуть кращими за апроксимацію фактичних значень.

Результати, отримані за допомогою цих моделей, часто порівнюються з результатами нейронних мереж. Існує два основних типи прогнозування часових рядів: одновимірний та багатовимірний.

Одновимірні моделі, такі як модель Бокса-Дженкінса, містять лише одну змінну в рекурентному рівнянні. Модель Бокса-Дженкінса – це складний процес припасування даних до відповідних параметрів моделі.

Рівняння, що використовуються в моделі, містять попередні значення ковзних середніх та цін. Модель Бокса-Дженкінса хороша для короткострокового прогнозування, але вимагає великої кількості даних, і це складний процес визначення відповідних рівнянь та параметрів моделі.

Багатовимірні моделі – це одномірні моделі, розширені для «виявлення випадкових факторів, що впливають на поведінку даних». Як випливає з назви, ці моделі містять більш ніж одну змінну у своїх рівняннях. Регресійний аналіз – це багатовимірна модель, яку часто порівнюють із нейронними мережами. У цілому нині, прогнозування часових рядів забезпечує прийнятну точність на коротких проміжках часу, але точність прогнозування часових рядів різко знижується зі збільшенням тривалості прогнозу.

2.1.4 Гіпотеза ефективного ринку

Гіпотеза ефективного ринку (ГЕР) стверджує, що будь-якої миті часу ціна акції повністю відображає всю відому інформацію про акцію. Оскільки вся відома інформація використовується учасниками ринку оптимально,

коливання ціни мають випадковий характер, оскільки нова інформація з'являється випадково.

Таким чином, ціни на акції роблять «випадкову прогулянку», і інвестор не може обіграти ринок.

Незважаючи на досить сильне твердження, яке на практиці здається неправдою, було отримано непереконливі дані, які спростовують ГЕР. Різні дослідження дійшли висновку, що ГЕР можна прийняти чи відкинути. Багато цих досліджень використовували нейронні мережі для обґрунтування своїх тверджень.

Однак, оскільки нейронна мережа хороша лише так, наскільки вона була навчена, важко стверджувати про прийняття чи відхилення гіпотези, ґрунтуючись лише на продуктивності нейронної мережі.

2.1.5 Теорія хаосу

Щодо нового підходу до моделювання нелінійних динамічних систем, таких як фондовий ринок, є теорія хаосу. Теорія хаосу аналізує процес у припущенні, що частина процесу є детермінованою, а частина – випадковою.

Хаос – це нелінійний процес, який видається випадковим. Різні теоретичні тести були розроблені для перевірки того, чи система є хаотичною (чи має вона хаос у своїх тимчасових рядах).

Теорія хаосу – це спроба показати, що порядок існує в випадковості. Припускаючи, що фондовий ринок є хаотичним, а чи не просто випадковим, теорія хаосу суперечить ГЕР.

По суті, хаотична система є комбінацією детермінованого та випадкового процесів. Детермінований процес може бути охарактеризований за допомогою регресійного припасування, у той час як випадковий процес може бути охарактеризований за допомогою статистичними параметрами функції розподілу. Отже, використання лише

детерміністичних чи статистичних методів зможе повністю передати природу хаотичної системи. Здатність нейронних мереж вловлювати як детерміновані, і випадкові особливості, робить її ідеальною для моделювання хаотичних систем.

2.1.6 Інші комп'ютерні методи

Багато інших комп'ютерних методів було використано для прогнозування ринку. Вони варіюються від програм побудови графіків до складних експертних систем. Також використовувалася нечітка логіка. Експертні системи послідовно обробляють знання та формулюють їх у вигляді правил. Вони можуть бути використані для формулювання торгових правил на основі технічних індикаторів. У цією ролі експертні системи можна використовувати разом із нейронними мережами для прогнозування ринку. У такій комбінованій системі нейронна мережа може виконувати свої прогнозування, а експертна система може підтвердити прогноз з урахуванням своїх відомих торгових правил. Перевага експертних систем полягає в тому, що вони можуть пояснити, як вони отримують свої результати.

У разі нейронних мереж важко проаналізувати важливість вхідних даних та те, як мережа отримала свої результати. Однак нейронні мережі працюють швидше, оскільки виконуються паралельно та більш стійкі до збоїв. Основною проблемою застосування експертних систем на фондовому ринку є складність формулювання знань про ринки, оскільки ми не до кінця їх розуміємо. Нейронні мережі мають перевагу перед експертними системами, оскільки вони можуть отримувати правила без їхньої явної формалізації. У такому хаотичному і лише частково зрозумілому середовищі, як фондовий ринок, це є важливим чинником. Важко отримати інформацію від експертів і формалізувати її таким чином, щоб її можна було використовувати в експертних системах.

Експертні системи ефективні тільки в межах своєї галузі знань і не працюють добре, коли є недостатня або неповна інформація. Нейронні мережі краще справляються з динамічними даними, можуть узагальнювати та робити «обґрунтовані припущення». Таким чином, нейронні мережі більше підходять для роботи на фондовому ринку, ніж експертні системи.

2.2 Опис існуючих підходів до прогнозування фінансового ринку

Kim та Han [4] побудували модель у вигляді комбінації штучних нейронних мереж (ІНС) та генетичних алгоритмів (ГА) з дискретизацією ознак для прогнозування індексу цін на акції. Дані, використані в їх дослідженні, включають технічні індикатори, а також напрямок зміни щоденного індексу цін на акції Кореї (KOSPI). Вони використовували дані, що містять вибірки з 2928 торгових днів, починаючи з січня 1989 по грудень 1998 року, і навели обрані ними ознаки і формули. Вони також застосували оптимізацію дискретизації ознак як техніку, яка схожа на зменшення розмірності.

Сильною стороною їхньої роботи є те, що вони ввели GA для оптимізації ANN. По-перше, кількість вхідних ознак та елементів обробки у прихованому шарі дорівнює 12 і не регулюється. Іншим обмеженням є процес навчання ІНС, і автори зосередилися лише на двох факторах під час оптимізації. Хоча вони, як і раніше, вважають, що ГА має великий потенціал для оптимізації дискретизації ознак. Наш ініціалізований пул ознак відноситься до вибраних ознак. Дослідники також представили рішення для прогнозування напряму японського фондового ринку на основі оптимізованої моделі штучної нейронної мережі. У цій роботі автори використовують генетичні алгоритми разом із моделями на основі штучних нейронних мереж та називають її гібридною моделлю GA-ANN.

Piramuthu [5] провів ретельну оцінку різних методів відбору ознак додатків інтелектуального аналізу даних. Він використовував такі набори

даних, як дані про схвалення кредитів, дані про неповернення кредитів, дані про веб-трафіку, дані про там і дані про кіанг, і порівняв, як різні методи відбору ознак оптимізують роботу дерева рішень. Методи вибору ознак, які він порівнював, включали імовірнісну міру відстані: міру Бхаттачар'ї, міру Матусити, міру дивергенції, міру відстані Махаланобіса та міру Патріка-Фішера. Для міжкласових заходів відстані: міра відстані Мінковського, міра відстані між міськими кварталами, евклідова відстань, міра відстані Чебичова та нелінійний (Парзен та гіперсферичне ядро) міра відстані. Сильною стороною даної є те, що автор оцінив як ймовірнісні методи вибору ознак на основі відстані, так і кілька міжкласових методів. Крім того, автор проводив оцінку на основі різних наборів даних, що посилює переваги даної роботи. Проте алгоритм оцінки був лише дерево рішень. Ми не можемо зробити висновок про те, чи будуть методи відбору ознак працювати так само ефективно на більшому наборі даних або більш складної моделі.

Hassan та Nath [6] використали приховану марківську модель (HMM) для прогнозування фондового ринку за цінами акцій чотирьох різних авіакомпаній. Вони скоротили стан моделі до чотирьох станів: ціна відкриття, ціна закриття, найвища ціна та найнижча ціна. Сильною стороною цієї роботи є те, що підхід не потребує експертних знань для побудови моделі прогнозування. Хоча ця робота обмежена рамками галузі авіакомпаній і оцінена дуже невеликому наборі даних, вона може призвести до створення універсальної моделі прогнозування. Для порівняння можна було б використати один із підходів у роботах, пов'язаних із прогнозуванням фондового ринку. Автори вибрали максимум 2 роки як діапазон дат для тренувального та тестового наборів даних, що надало нам посилання на діапазон дат для нашої частини оцінки.

Lei L. [7] використав вейвлетову-нейронну мережу (WNN) для прогнозування тенденцій зміни цін на акції. Автор також застосував грубе безліч (Rough Set, RS) для скорочення атрибутів як оптимізацію. Грубе

безліч було використано зменшення розмірів ознак тренду ціни акції. Воно також використовувалося визначення структури вейвлет-нейронной мережі. Набір даних у даній роботі складається з п'яти відомих індексів фондового ринку, а саме: (1) індекс SSE Composite (Китай), (2) індекс CSI 300 (Китай), (3) індекс All Ordinaries (Австралія), (4) індекс Nikkei 225 (Японія) та (5) індекс Dow Jones (США). Оцінка моделі проводилася з урахуванням різних індексів ринку, і результат був переконливим з погляду спільності. Використання грубої множини для оптимізації розмірності ознак перед обробкою знижує обчислювальну складність. Проте автор лише підкреслив налаштування параметрів щодо обговорення, але не вказав слабкі сторони самої моделі. Тим часом, ми також виявили, що оцінки проводилися на індексах, і та сама модель може не мати такої ж ефективності, якщо її застосувати до конкретних акцій.

Lee MC [8] використав машину опорних векторів (SVM) разом із гібридним методом вибору ознак прогнозування тенденцій ринку акцій. Набір даних у цьому дослідженні є піднабіром даних індексу NASDAQ у базі даних Тайванського економічного журналу (TEJD) за 2008 рік. У частині відбору ознак використовувався гібридний метод, у ролі обгортки виступав підтримуваний послідовний прямий пошук (SSFS). Ще однією перевагою даної роботи є те, що в ній була розроблена докладна процедура налаштування параметрів із зазначенням продуктивності при різних значеннях параметрів. Чітка структура моделі відбору ознак є також евристикою для первинного етапу структурування моделі. Одним з обмежень було те, що продуктивність SVM порівнювалася лише з нейронною мережею із зворотним розповсюдженням (BPNN) і не порівнювалася з іншими алгоритмами машинного навчання.

Sirignano та Cont [9] використали рішення глибокого навчання, навчене на універсальному наборі ознак фінансових ринків. Набір даних, що використовується, включав записи всіх угод на купівлю та продаж, а також скасування ордерів для приблизно 1000 акцій NASDAQ через книгу ордерів

біржі. NN складається з трьох шарів з блоками LSTM і шару feed-forward з випрямленими лінійними блоками (ReLU) в кінці зі стохастичним алгоритмом градієнтного спуску (SGD) в якості оптимізації. Їх універсальна модель була здатна узагальнювати та охоплювати акції, відмінні від тих, що були у навчальних даних. Хоча вони відзначили переваги універсальної моделі, вартість навчання все ще була дорогою. Тим часом через неявне програмування алгоритму глибокого навчання неясно, чи є марні ознаки, забруднені при подачі даних у модель. Автори виявили, що було б краще, якби вони виконали частину відбору ознак перед навчанням моделі, і вважають це ефективним способом зниження обчислювальної складності.

Gao YZ та ін. [10] передбачали тенденції зміни цін на акції за допомогою SVM і виконували фрактальний відбір ознак для оптимізації. Як набор даних вони використовували композитний індекс Шанхайської фондової біржі (SSECI), а як ознак – 19 технічних індикаторів. Перед обробкою даних оптимізували вхідні дані, виконавши відбір ознак. При пошуку найкращої комбінації параметрів вони також використовували метод пошуку по сітці, який є k-крос-валідацією. Крім того, проведено всебічну оцінку різних методів відбору ознак. Як зазначили автори у висновку, вони розглядали лише технічні показники, але не макро- та мікрофактори у фінансовій сфері. Джерело даних, яке використовували автори, було аналогічне нашому набору даних, що робить результати їх оцінки корисними для нашого дослідження. Вони також згадали метод під назвою k-перехресна валідація під час тестування комбінацій гіперпараметрів.

McNally та ін. [11] використовували RNN і LSTM для прогнозування ціни біткоїну, оптимізували їх за допомогою алгоритму Борута для підбору ознак, і він працює аналогічно класифікатору random forest. Крім вибору ознак, вони також використовували оптимізацію Байєса для вибору параметрів LSTM. Набір даних Bitcoin від 19 серпня 2013 року до 19 липня 2016 року. Використовували кілька методів оптимізації для покращення

продуктивності методів глибокого навчання. Основна проблема їх роботи – надмірне припасування. Дослідницьке завдання прогнозування динаміки ціни біткоїну має деякі подібності із прогнозуванням цін на фондовому ринку. Приховані особливості та шуми, вбудовані в дані про ціни, є загрозою для цієї роботи. Автори розглядали питання дослідження як проблему тимчасової послідовності. Найкраща частина цієї роботи – розробка та оптимізація функцій; ми могли б повторити методи, які вони використовували у нашій попередній обробці даних.

Weng та Lu [12] зосередилися на короткостроковому прогнозуванні цін на акції за допомогою ансамблевих методів чотирьох відомих моделей машинного навчання. Набір даних для цього дослідження складається із п'яти наборів даних. Вони отримали ці набори даних із трьох API з відкритим вихідним кодом та пакету R під назвою TTR. Вони використовували такі моделі машинного навчання: (1) ансамбль нейромережної регресії (NNRE), (2) Random Forest з необрізаними деревами регресії як базові учні (RFR), (3) AdaBoost з необрізаними деревами регресії як базові учні (BRT) та (4) комплекс регресії вектора підтримки (SVRE). Детальне дослідження ансамблевих методів, визначених для короткострокового прогнозування ціни акції. Маючи базові знання, автори вибрали вісім технічних індикаторів для дослідження, а потім провели ретельну оцінку п'яти наборів даних. Основний внесок даної роботи полягає в тому, що вони розробили платформу для інвесторів з використанням R, яка не вимагає від користувачів введення власних даних, а викликає API для отримання даних із прямого онлайн-джерела. З погляду дослідження, вони оцінювали лише прогнозування ціни на термін від 1 до 10 днів уперед, але не оцінювали більш тривалі терміни, ніж два торгові тижні або коротший термін, ніж 1 день. Основним обмеженням дослідження є те, що вони проаналізували лише 20 американських акцій, модель не може бути узагальнена на інші фондові ринки або потребує подальшої ревалідації, щоб побачити, чи не страждає вона від проблем із припасуванням.

Kara та ін. також використовували ANN та SVM для прогнозування руху індексу цін на акції. Набір даних, що використовується ними, охоплює період часу з 2 січня 1997 року по 31 грудня 2007 року на Стамбульській фондовій біржі. Головною сильною стороною цієї роботи є докладний опис процедур налаштування параметрів. Слабкими сторонами даної роботи є те, що ні технічний показник, ні структура моделі не мають новизну, і автори не пояснили, чим їхня модель краща за інші моделі в попередніх роботах.

Таким чином, допомогли б додаткові роботи щодо валідації на інших наборах даних. Вони пояснили, як ANN та SVM працюють з характеристиками фондового ринку, а також записали налаштування параметрів. Реалізаційна частина нашого дослідження могла б отримати користь із цієї попередньої роботи.

Jeon та ін. [13] провели дослідження великого набору даних на основі мілісекундних інтервалів, використовуючи відстеження графа патернів для виконання завдань прогнозування цін на акції. Вони використовували набір даних, заснований на мілісекундних інтервалах, великий набір історичних даних про ціни на акції від KOSCOM, з серпня 2014 до жовтня 2014 року, ємністю 10G-15G. Для розпізнавання образів автор застосував евклідову відстань, динамічну деформацію часу (DTW). Для відбору ознак використовувалася покрокова регресія. Автори виконали завдання прогнозування за допомогою ANN та Hadoop та RHive для обробки великих даних. Перед обробкою даних з дискретних даних отримано агреговані дані з інтервалом в 5 хвилин. Основною сильною стороною цієї роботи є очевидна структура всієї процедури реалізації. Хоча вони використовували відносно стару модель, іншою слабкою стороною є те, що загальний часовий проміжок навчального набору даних є надзвичайно коротким. У реальному житті важко отримати доступ до даних, заснованих на мілісекундних інтервалах, тому модель не така практична, як модель, заснована на щоденних даних.

Fischer та Krauss [14] застосували модель fuzzy-GA для вирішення завдання вибору акцій. Вони використовували ключові акції з 200 найбільших за ринковою капіталізацією, перерахованих як інвестиційний всесвіт на Тайванській фондовій біржі. Крім того, дані річного фінансового звіту та прибутковість акцій були взяті з бази даних Taiwan Economic Journal (TEJ) на сайті www.tej.com.tw/ за період з 1995 до 2009 року. Вони провели нечітку функцію членства з параметрами моделі, оптимізованими за допомогою GA, та отримали ознаки для оптимізації скорингу акцій. Автори запропонували оптимізовану модель для вибору та оцінки акцій. На відміну від моделі прогнозування, автори більше зосередилися на ранжируванні, виборі та оцінці ефективності акцій. Їхня структура більш практична для інвесторів. Але в частині перевірки моделі вони порівнювали модель не з існуючими алгоритмами, а зі статистикою еталона, що ускладнює визначення того, чи GA перевершить інші алгоритми.

Fischer та Krauss [14] використали довготривалу пам'ять (LSTM) для прогнозування фінансового ринку. Як набору даних вони використовували дані про складові індексу S&P 500 від Thomson Reuters. Вони отримали всі списки складових індексу S&P 500 на кінець місяця з грудня 1989 по вересень 2015 року, потім об'єднали списки в бінарну матрицю, щоб усунути упередження тих, хто вижив. Як оптимізатор автори також використовували RMSprop, який є міні-пакетною версією rprop. Головною сильною стороною цієї роботи є те, що автори використали новітню техніку глибокого навчання для виконання прогнозів. Вони покладалися на техніку LSTM, не маючи фонових знань у фінансовій галузі. Хоча LSTM перевершила стандартні алгоритми DNN та логістичної регресії, автор не згадав про зусилля щодо навчання LSTM із тривалими залежностями.

Tsai та Hsiao [15] запропонували рішення у вигляді комбінації різних методів відбору ознак прогнозування акцій. Як джерело даних вони використовували базу даних Taiwan Economic Journal (TEJ). У їхньому аналізі використовувалися дані з 2000 по 2007 рік. У своїй роботі вони

застосували аналіз основних компонентів (PCA) для зниження розмірності, генетичні алгоритми (GA) та дерева класифікації та регресії (CART) для відбору важливих характеристик. Вони не покладалися лише на технічні індекси. Натомість вони також включили в аналіз фундаментальні та макроекономічні індекси. Автори також здійснили порівняння методів відбору ознак. Валідацію було проведено шляхом об'єднання статистики ефективності моделі зі статистичним аналізом.

Nekoeiqachkanloo та ін. [16] запропонували систему з двома різними підходами для інвестування в акції. Сильні сторони запропонованого рішення очевидні. По-перше, це комплексна система, що складається з попередньої обробки даних та двох різних алгоритмів для пропозиції найкращих інвестиційних порцій. По-друге, у систему вбудований компонент прогнозування, який також зберігає особливості часового ряду. І останнє, але не менш важливе: їх вхідні характеристики є сумішшю фундаментальних характеристик і технічних індексів, які покликані заповнити прогалину між фінансовою та технічною областями. Проте їхня робота має недолік у частині оцінки. Замість оцінити запропоновану систему на великому наборі даних, вони обрали 25 відомих акцій. Існує велика ймовірність, що відомі акції потенційно можуть мати деякі загальні приховані характеристики.

Одним з основних недоліків, виявлених у відповідних роботах, є обмеженість побудованих та використовуваних механізмів попередньої обробки даних. Технічні роботи переважно зосереджені на побудові моделей прогнозування. Роботи, що стосуються інвестиційної сфери, виявляють більший інтерес до аналізу поведінки, наприклад, як стадна поведінка впливає показники акцій, чи як відсоток внутрішніх директорів, які мають звичайними акціями підприємства, впливає показники певних акцій. Для розпізнавання таких моделей поведінки часто потрібна попередня обробка стандартних технічних індексів та інвестиційний досвід.

У відповідних роботах часто проводиться ретельний статистичний аналіз на основі спеціального набору даних та виводяться нові ознаки, а не проводиться відбір ознак. Деякі дані, такі як відсоток коливань певного індексу, довели свою ефективність щодо фондових показників. Ми вважаємо, що вилучення нових ознак з даних, а потім поєднання таких ознак із існуючими загальними технічними індексами значно покращить існуючі та добре перевірені моделі прогнозування.

3 ПОСТАНОВКА ЗАДАЧІ

3.1 Дослідницькі питання

Ми розглянемо три дослідницькі питання в кожному аспекті, відповідно: як інженерія характеристик може підвищити точність прогнозування моделі? Як висновки з фінансової галузі можуть допомогти в розробці моделі прогнозування? І який найкращий алгоритм для прогнозування короткострокових цінових тенденцій?

Перше Дослідницьке питання стосується розробки функцій. Ми хотіли б дізнатися, як метод відбору ознак впливає на продуктивність моделей прогнозування. З великої кількості попередніх робіт ми можемо зробити висновок, що дані про ціни на акції містять високий рівень шуму, а також існують кореляції між ознаками, що робить прогнозування цін досить складним. Це також є основною причиною того, що в більшості попередніх робіт в якості модуля оптимізації була введена частина, пов'язана з розробкою ознак.

Друге питання дослідження полягає в оцінці ефективності висновків, які ми витягли з фінансової сфери. На відміну від попередніх робіт, крім звичайної оцінки моделей даних, таких як вартість навчання і бали, наша оцінка буде підкреслювати ефективність нових доданих ознак, які ми витягли з фінансового домену. Ми вводимо деякі особливості з фінансової області. Хоча ми отримали тільки деякі конкретні висновки з попередніх робіт, пов'язані з ними необроблені дані необхідно переробити в придатні для використання характеристики. Після вилучення відповідних ознак з фінансової області ми об'єднуємо їх з іншими загальними технічними показниками, щоб вибрати ознаки з більш високим впливом. Існує безліч ознак, що вважаються ефективними у фінансовій сфері, і охопити їх все було б неможливо. Таким чином, питання про те, як правильно перетворити результати, отримані у фінансовій сфері, в модуль обробки даних нашої

системи, є прихованим дослідницьким питанням, на яке ми намагаємося відповісти.

Третє Дослідницьке питання полягає в тому, якими алгоритмами ми будемо моделювати наші дані? З попередніх робіт випливає, що дослідники докладали зусиль для точного прогнозування цін. Ми декомпозуємо проблему на передбачення тренда, а потім точного числа.

У даній роботі основна увага приділяється першому кроку. Таким чином, мета була перетворена в вирішення проблеми двійкової класифікації, при цьому був знайдений ефективний спосіб усунення негативного ефекту, викликаного високим рівнем шуму. Наш підхід полягає в тому, щоб розкласти складну проблему на підпроблеми, які мають менше залежностей, і вирішити їх одну за одною, а потім об'єднати рішення в ансамблеву модель в якості допоміжної системи для відстеження інвестиційної поведінки.

У попередніх роботах дослідники використовували різні моделі для прогнозування тенденцій зміни цін на акції. Хоча більшість моделей, які показали найкращі результати, засновані на методах машинного навчання, в даній роботі ми порівняємо наш підхід з моделями машинного навчання в частині оцінки і знайдемо рішення для даного дослідницького питання.

3.2 Високорівнева архітектура рішення

Високорівнева архітектура пропонованого нами рішення може бути розділена на три частини.

По-перше, це частина відбору ознак, щоб гарантувати високу ефективність обраних ознак.

По-друге, ми розглядаємо дані і виконуємо зниження розмірності. І остання частина, яка є основним внеском нашої роботи, полягає в побудові моделі прогнозування цільових запасів.

На рисунку 3.1 показана високорівнева Архітектура пропонованого рішення.

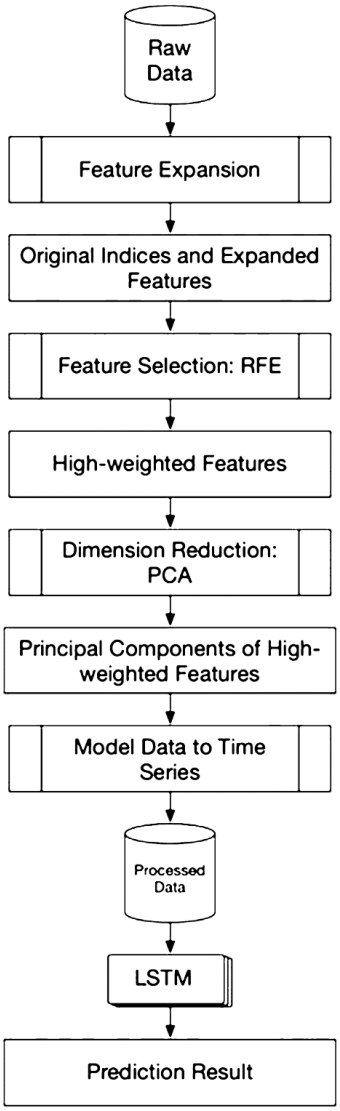


Рисунок 3.1 – Високорівнева архітектура пропонованого рішення

Існують способи класифікації різних категорій акцій. Деякі інвестори віддають перевагу довгостроковим інвестиціям, тоді як інші виявляють більший інтерес до короткострокових інвестицій. Часто можна бачити, що звіти, пов'язані з акціями, показують середні показники, в той час як ціна акцій різко зростає; це одне з явищ, що вказують на те, що прогнозування цін на акції не має фіксованих правил, тому необхідно знайти ефективні характеристики перед навчанням моделі на даних.

У даному дослідженні ми зосередилися на прогнозуванні короткострокових цінових тенденцій. В даний час у нас є тільки вихідні дані без міток. Тому найпершим кроком є маркування даних. Ми відзначаємо цінову тенденцію, порівнюючи поточну ціну закриття з ціною закриття N торгових днів тому, діапазон n – від 1 до 10, оскільки наше дослідження зосереджено на короткостроковому періоді.

Однак, щоб забезпечити найкращу продуктивність моделі прогнозування, ми спочатку вивчимо дані. У вихідних даних є велика кількість ознак; якщо ми будемо враховувати всі ознаки, це не тільки значно збільшить обчислювальну складність, але і викличе побічні ефекти, Якщо ми хочемо використовувати навчання без нагляду в подальших дослідженнях. Тому ми використовуємо рекурсивне виключення ознак (RFE) для забезпечення ефективності всіх обраних ознак.

Ми виявили, що більшість попередніх робіт в технічній галузі аналізували всі акції, в той час як у фінансовій галузі дослідники вважають за краще аналізувати конкретний сценарій інвестицій. Щоб заповнити прогалину між двома областями, ми вирішили застосувати розширення функцій на основі результатів, отриманих у фінансовій галузі, перш ніж почати процедуру RFE.

Оскільки ми плануємо моделювати дані у вигляді часових рядів, чим більше число ознак, тим складніше буде процедура навчання. Тому на початку запропонованої нами архітектури рішення ми скористаємося скороченням розмірності за допомогою рандомізованого PCA.

4 ПРОЕКТУВАННЯ ТЕХНІЧНОГО ПРОЕКТУ

4.1 Підготовка набору даних

У цьому розділі детально описані дані, які були витягнуті з відкритих джерел даних, і остаточний набір даних, який був підготовлений. Дані, пов'язані з фондовим ринком, різноманітні, тому ми спочатку порівняли відповідні роботи з огляду робіт з фінансових досліджень в області аналізу даних фондового ринку, щоб визначити напрямки збору даних. Після збору даних ми визначили структуру набору даних. Нижче ми детально описуємо набір даних, включаючи структуру даних і таблиці даних в кожній категорії даних з визначеннями сегментів.

Цей набір даних складається з 3600 акцій з китайського фондового ринку. Крім щоденних цінових даних, щоденних фундаментальних даних по кожному фондовому ідентифікатору, ми також зібрали історію призупинень і відновлень, 10 найбільших акціонерів і т. д.

Ми перерахували дві причини, чому ми вибрали 2 роки як часовий інтервал для цього набору даних:

1) більшість інвесторів проводять аналіз динаміки цін на фондовому ринку, використовуючи дані за останні 2 роки;

2) використання більш свіжих даних покращить результат аналізу.

Ми збирали дані через API з відкритим кодом, а саме Tusare [17], а також використовували техніку веб-скрапінгу для збору даних.

Веб-скрапінг – це процес автоматизованого збору структурованих веб-даних. Його також називають витяганням веб-даних. Деякі з основних випадків використання Веб-скрейпінг включають моніторинг цін, цінову розвідку, моніторинг новин, генерацію свинцю, маркетингові дослідження та багато іншого.

На рисунку 4.1 показані всі таблиці даних в наборі даних. У цьому наборі даних ми зібрали чотири категорії даних: (1) базові дані, (2) торгові дані, (3) фінансові дані та (4) інші довідкові дані.

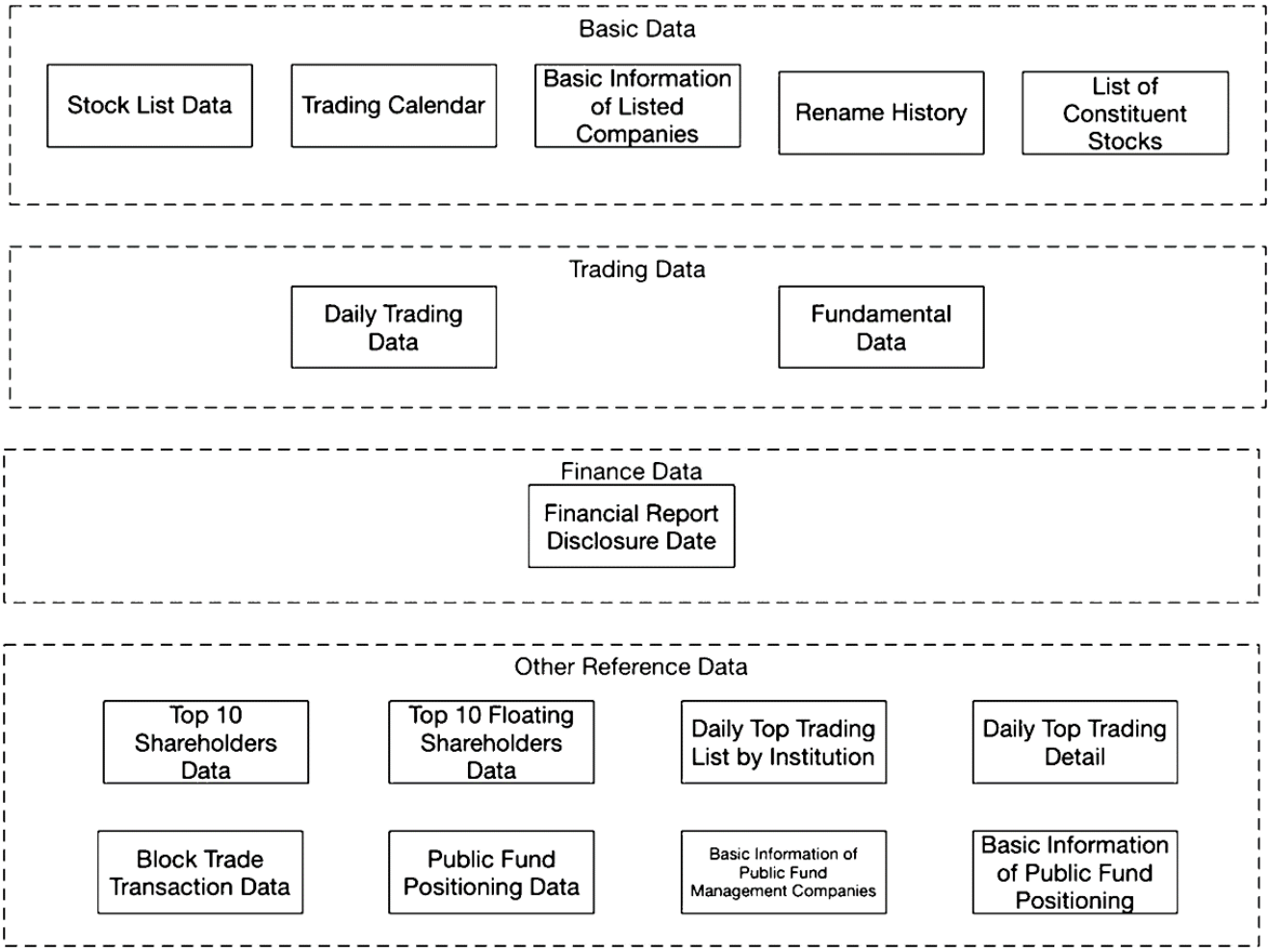


Рисунок 4.1 – Структура даних для витягнутого набору даних

Всі таблиці даних можуть бути пов'язані один з одним загальним полем під назвою «Stock ID».

«Stock ID» – Це унікальний ідентифікатор акцій, зареєстрований на китайському фондовому ринку.

На рисунку 4.2 представлен детальний огляд набору даних, інформація про поля кожної таблиці даних, а також до якої категорії відноситься таблиця даних.

Data table name	Category	Field
Stock list	Basic data	Stock ID, Stock name, Geographic info, Industry, Full name, English name, Market type, Stock exchange ID, Currency, List status, List date, Delist date, If the stock is HS constituent
Trading calendar	Basic data	Stock exchange ID, Calendar date, If the date is open for trading, Pervious trading date
Basic information of listed companies	Basic data	Stock ID, Stock exchange ID, Corporate representative, General manager, Secretary, Authorized capital, Registration date, Province, City, Introduction, Website, Email, Office address, Number of employees, Main business, Business scope
Renamed history	Basic data	Stock ID, Stock name, Start date, End date, Announcement date, Rename reason
Constituent stock information	Basic data	Stock ID, Constituent type, Included date, Excluded date, If the stock is new
Daily trading data	Trading data	Stock ID, Trading date, Opening price, Highest price, Lowest price, Closing price, Previous closing price, Price change, Price change percentage, Volume, Amount
Fundamental data	Trading data	Stock ID, Trading date, Closing price, Turnover rate, Free turnover rate, Volume ratio, Price-to-earning ratio, Price-to-earning ratio TTM, Price-to-book ratio, Price-to-sales ratio, Price-to-sales TTM, Total share capital, Circulating shares, Tradable circulating shares, Aggregate market value, Circulation market value
Financial report disclosure date	Finance data	Stock ID, Latest disclosure date, Reporting period, Scheduled disclosure date, Actual disclosure date, Disclosure modification date
Top 10 shareholders data	Other reference data	Stock ID, Announcement date, End date, Shareholder name, Holding amount, Holding ratio
Top 10 floating shareholders data	Other reference data	Stock ID, Announcement date, End date, Shareholder name, Holding amount
Daily top trading list by institution	Other reference data	Stock ID, Trading date, Institution name, Trading amount—buy, Trade ratio—buy, Trading amount—sell, Trade ratio—sell, Net turnover
Daily top transaction detail	Other reference data	Stock ID, Trading date, Stock name, Closing price, Price change percentage, Turnover rate, Amount—overall, On-list amount—sell, On-list amount—buy, On-list turnover, On-list net trading amount, On-list net trading ratio, On-list net turnover ratio, Circulation market value, Reason
Block trade transaction data	Other reference data	Stock ID, Trading date, Price, Volume, Amount, Buyer, Seller
Public fund positioning data	Other reference data	Fund ID, Announcement date, End date, Stock ID, Market value, Volume, Market value ratio, Circulation market value ratio
Basic information of public fund management companies	Other reference data	Company name, Short name, Province, City, Address, Phone, Office
Basic information of public fund positioning	Other reference data	Fund ID, Name, Management company, Custodian, Fund type, Founded date, Due date, List date, Issued date, Delist date, Issued amount, Fee, Duration, Value, Min Amount, Expecting return, Benchmark

Рисунок 4.2 – Таблиця огляду набору даних з різними категоріями і підмножинами полів

Як ми можемо бачити у нас є п'ять основних таблиць – акції, торговий календар, інформація о компаніях, історія переіменування, інформація про акції, що входять до складу компанії.

4.2 Опис технічного проекту

У цьому розділі наводиться докладний опис технічного проекту, який являє собою комплексне рішення, засноване на використанні, комбінуванні та адаптації декількох існуючих методів попередньої обробки даних, розробки ознак і глибокого навчання. На рисунку 4.3 представлений детальний технічний проект від обробки даних до прогнозування, включаючи дослідження даних.

Ми розділили зміст на основні процедури, і кожна процедура містить алгоритмічні кроки. Подробиці алгоритмів описані в наступному розділі. Зміст цього розділу буде зосереджено на ілюстрації процесу обробки даних.

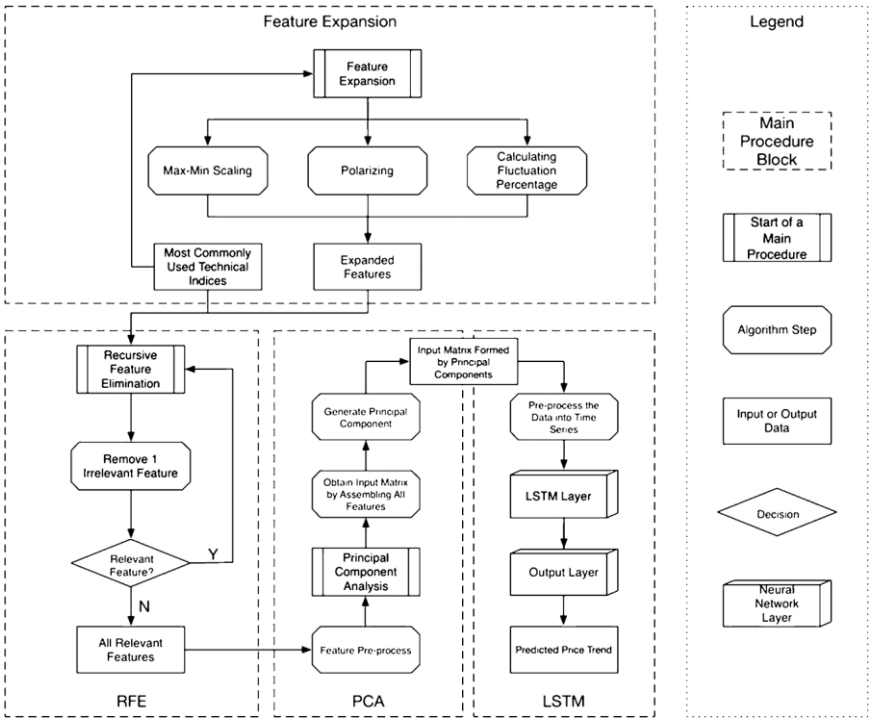


Рисунок 4.3 – Детальний технічний дизайн пропонованого рішення

На основі огляду літератури ми відбираємо найбільш часто використовувані технічні індекси і потім вводимо їх в процедуру розширення ознак для отримання розширеного набору ознак. З розширеного набору ознак ми виберемо найбільш ефективні ознаки. Потім ми подамо

дані з вибраними ознаками в алгоритм PCA, щоб зменшити розмірність на необхідну кількість ознак. Після отримання найкращої комбінації, ми обробляємо дані в остаточний набір ознак і подаємо їх в модель LSTM, щоб отримати результат прогнозування цінового тренда.

Модель LSTM – це вдосконалена RNN, послідовна мережа, яка дозволяє зберігати інформацію. Вона здатна впоратися з проблемою зникаючого градієнта, з якою стикаються РНС. Рекурентна нейронна мережа, також відома як RNN, використовується для постійної пам'яті.

У нашому дослідженні у модель LSTM подається вхідна матриця, сформована основними компонентами.

LSTM складається с чотирьох компонентів:

1. попередня обробка даних у часові ряди;
2. LSTM шар;
3. вихідній шар;
4. прогнозована динаміка цін.

Новизна пропонованого нами рішення полягає в тому, що ми не тільки застосуємо технічний метод на необроблених даних, але і виконуємо розширення ознак, які використовуються інвесторами фондового ринку. Детально про розширення ознак розповідається в наступному підрозділі.

Досвід, отриманий при застосуванні та оптимізації рішень на основі глибокого навчання в [19], [20] був врахований при розробці та налаштуванні рішення по розширенню ознак і глибокому навчанню в даній роботі.

5 РЕАЛІЗАЦІЯ ТЕХНІЧНОГО ПРОЕКТУ

5.1 Застосування методу розширення ознак

Першою основною процедурою на рисунку 4.3 є розширення ознак. У цьому блоці вхідними даними є найбільш часто використовувані технічні показники, отримані з відповідних робіт. Три переваги методу розширення характеристик – це макс-мін масштабування, поляризація і обчислення відсотка коливань. Не всі технічні індекси застосовні для всіх трьох методів розширення характеристик; в даній процедурі застосовуються тільки значущі методи розширення технічних індексів.

Технічні індекси і відповідні методи розширення ознак показані на рисунку 5.1.

Feature	Polarize	Max–min scale	Fluctuation percentage
Price change			
Price change percentage			
Volume		√	
Amount		√	
SMA 10		√	√
MACD	√		
MACD SIGNAL	√		
MACD HIST	√		
CCI 24	√		
MTM 10	√		√
ROC 10	√		√
RSI 5		√	√
WNR 9	√	√	
SLOWK		√	√
SLOWD		√	√
ADOSC	√	√	
AR 26		√	
BR 26		√	
VR 26		√	√
BIAS 20	√		

Рисунок 5.1 – Вибір методу розширення характеристик

Ми вибираємо значущі методи розширення, дивлячись на те, як розраховуються індекси.

5.2 Застосування методу розширення характеристик

Прогнозування короткострокового тренду цін на фондовому ринку за допомогою комплексної системи глибокого навчанняпісля процедури розширення функцій, Розширені функції будуть об'єднані з найбільш часто використовуваними технічними індексами, тобто вхідні дані з вихідними даними, і подані в блок RFE в якості вхідних даних на наступному етапі.

5.3 Застосування методу рекурсивного усунення ознак

Після розширення функцій, описаного вище, ми досліджуємо найбільш ефективні функції і за допомогою алгоритму рекурсивного виключення функцій (RFE) [8]. Ми оцінюємо всі ознаки за двома атрибутами, коефіцієнтом і важливості ознаки. Ми також обмежуємо кількість ознак, що видаляються з пулу, що означає, що ми будемо видаляти за однією ознакою на кожному кроці і зберігати всі релевантні ознаки. Потім вихід блоку RFE стане входом наступного кроку, який відноситься до PCA.

5.4 Застосування методу аналізу головних компонент (PCA)

Найпершим кроком перед використанням PCA є попередня обробка ознак. Оскільки деякі ознаки після RFE являють собою процентні дані, а інші-дуже великі числа, тобто вихідні дані RFE представлені в різних одиницях виміру. Це вплине на результат вилучення головних компонент.

Таким чином, перед подачею даних в алгоритм PCA [11] необхідна попередня обробка ознак.

Після виконання попередньої обробки ознак наступним кроком є подача оброблених даних з обраними і ознаками в алгоритм PCA для зменшення масштабу матриці ознак до j ознак.

Цей крок спрямований на те, щоб зберегти якомога більше ефективних ознак і тим самим усунути обчислювальну складність навчання моделі. У даній дослідницькій роботі також оцінюється найкраща комбінація i і j, яка має відносно кращу точність передбачення і при цьому скорочує обчислювальні витрати. Після етапу PCA система отримає переформовану матрицю з j стовпцями.

5.5 Підбір моделі довготривалої короткочасної пам'яті (LSTM)

PCA зменшила розмірність вхідних даних, в той час як попередня обробка даних є обов'язковою перед подачею даних в шар LSTM. Причина додавання етапу попередньої обробки даних перед LSTM-моделлю полягає в тому, що вхідна матриця, сформована головними компонентами, не має тимчасових кроків. У той час як одним з найбільш важливих параметрів навчання LSTM є кількість тимчасових кроків. Отже, ми повинні моделювати матрицю з відповідними часовими кроками як для навчального, так і для тестового набору даних.

Після виконання частини попередньої обробки даних, останнім кроком є подача навчальних даних в LSTM і оцінка продуктивності за допомогою тестових даних. Як варіант нейронної мережі RNN, навіть з одним шаром LSTM, структура NN все ще є глибокою нейронною мережею, оскільки вона може обробляти послідовні дані і запам'ятовувати свої приховані стани в часі. Шар LSTM складається з одного або декількох блоків LSTM, а блок LSTM складається з комірок і затворів для виконання класифікації та прогнозування на основі даних часового ряду.

Структура LSTM складається з двох шарів. Вхідна розмірність визначається як матриця получена після алгоритму PCA. Перший шар є вхідним шаром LSTM, а другий шар – вихідним. Кінцевий вихідний сигнал 0 або 1 показує, чи буде результат прогнозування тренду ціни акцій падати

або рости, що є допоміжною пропозицією для інвесторів при прийнятті наступного інвестиційного рішення.

5.6 Розробка дизайну

Розширення функцій є одним з нововведень пропонованої нами системи прогнозування цінових тенденцій.

У процедурі розширення характеристик ми використовуємо технічні індекси для спільної роботи з евристичними методами обробки, отриманими від інвесторів, що заповнює прогалину між областю фінансових досліджень і областю технічних досліджень.

Оскільки ми запропонували систему прогнозування цінового тренду, розробка ознак надзвичайно важлива для кінцевого результату прогнозування. Не тільки метод розширення ознак допомагає гарантувати, що ми не пропустимо потенційно корелюючі ознаки, але і метод відбору ознак необхідний для об'єднання ефективних ознак. Чим більше нерелевантних ознак буде введено в модель, тим більше буде шуму.

Кожна основна процедура ретельно продумана і вносить свій внесок в розробку всієї системи.

Крім розробки функцій, ми також використовуємо LSTM, сучасний метод глибокого навчання для прогнозування часових рядів, що гарантує, що модель прогнозування зможе вловити як складний прихований патерн, так і патерн, пов'язаний з часовими рядами.

Відомо, що вартість навчання моделей глибокого навчання велика як в тимчасовому, так і в апаратному аспектах; ще однією перевагою нашої системи є процедура оптимізації-PCA. Вона може зберігати головні компоненти ознак при зменшенні масштабу матриці ознак, що допомагає системі заощадити витрати на навчання при обробці великої матриці ознак часових рядів.

5.7 Розробка алгоритмів

У цьому розділі представлені докладні відомості про алгоритми, які ми побудували, використовуючи і адаптуючи різні існуючі методи. Детально описані терміни, параметри, а також оптимізатори.

У правій частині рисунку 4.3, ми відзначаємо кроки алгоритму у вигляді восьмикутників.

Перш ніж заглибитися в кроки алгоритму, коротко розповімо про попередню обробку даних: оскільки ми будемо працювати з алгоритмами контрольованого навчання, нам також необхідно запрограмувати базову істину.

Базова істина в даному дослідженні програмується шляхом порівняння ціни закриття поточної торгової дати з ціною закриття попередньої торгової дати, з якої користувачі хочуть порівняти.

Зростання ціни позначається як 1, а «Базова істина» буде позначена як 0. Оскільки дана дослідницька робота спрямована не тільки на Прогнозування цінового тренда конкретного періоду часу, але і короткострокового в цілому, обробка «базової істини» проводиться відповідно до діапазону торгових днів. Хоча алгоритми не змінюються в залежності від довжини терміну прогнозування, ми можемо розглядати довжину терміну як параметр.

Алгоритми детально описані, відповідно, перший алгоритм – це гібридна частина розробки ознак для підготовки високоякісних навчальних і тестових даних. Він відповідає блокам розширення ознак, RFE і PCA на рисунку 4.3.

Другий алгоритм-блок процедури LSTM, що включає попередню обробку даних часових рядів, побудову NN, навчання і тестування.

5.7.1 Короткострокове прогнозування цінового тренду на фондовому ринку з використанням функції FE + RFE + PCA

Функція FE відповідає блоку розширення ознак. Для процедури розширення ознак ми застосовуємо три різних методи обробки, щоб перевести висновки з фінансової області в технічний модуль нашої системи. Хоча не всі показники застосовні для розширення, ми вибираємо відповідний метод для певних ознак, щоб виконати розширення ознак (FE), відповідно до рисунку 5.1.

Метод нормалізації зберігає відносні частоти термінів і перетворює технічні індекси в діапазон [0, 1]. Поляризація це добре відомий метод, часто використовуваний реальними інвесторами, іноді вони вважають за краще враховувати, вище або нижче нуля значення технічного індексу, ми програмуємо деякі ознаки за допомогою методу поляризації та готуємося до RFE.

Масштабування Max-min (або min-max) [16] – це метод перетворення, часто використовуваний в якості альтернативи масштабування нульового середнього і одиничної дисперсії.

Інший відомий метод – відсоток коливань, і ми перетворимо відсоток коливань технічних індексів в діапазон [- 1, 1].

Функція RFE () в першому алгоритмі відноситься до рекурсивного усунення ознак. Перш ніж виконувати скорочення масштабу навчальних даних, ми повинні переконатися, що вибрані нами ознаки ефективні. Неефективні ознаки не тільки знижують точність класифікації, але і збільшують складність обчислень.

Для відбору ознак ми вибрали рекурсивне виключення ознак (RFE). Як пояснюється в [15], процес рекурсивного виключення ознак можна розділити на алгоритм ранжирування, повторну вибірку і зовнішню перевірку.

Алгоритм ранжирування підганяє модель до ознак і ранжує їх за важливістю для моделі.

Ми задаємо параметр для збереження I числа ознак, і на кожній ітерації відбору ознак зберігаємо набір ознак, які отримали найвищий рейтинг, потім підганяємо модель і знову оцінюємо продуктивність, щоб почати нову ітерацію. У підсумку алгоритм ранжирування визначає набір кращих ознак.

Відомо, що алгоритм RFE страждає від проблеми надмірної підгонки. Щоб усунути проблему надмірної підгонки, ми запустимо алгоритм RFE кілька разів на випадково обраних акціях в якості навчального набору і переконаємося, що всі вибрані нами ознаки мають високу вагу.

Остання частина нашого гібридного алгоритму інженерії ознак призначена для оптимізації. Для зменшення масштабу матриці навчальних даних ми застосовуємо рандомізований аналіз головних компонент (PCA) [13], перш ніж визначитися з ознаками класифікаційної моделі.

Фінансові коефіцієнти зареєстрованої на біржі компанії використовуються для подання здатності до росту, здатності заробляти, платоспроможності і т. д.

Кожен фінансовий коефіцієнт складається з набору технічних показників, кожен раз, коли ми додаємо технічний показник (або ознаку), додається ще один стовпець даних в матрицю даних, що призводить до низької ефективності навчання і надмірності.

Якщо в навчальні дані будуть включені нерелевантні або менш релевантні ознаки, це також знизить точність класифікації.

Наведене нижче рівняння являє собою пояснювальну здатність головних компонент, витягнутих методом PCA для вихідних даних.

Algorithm 1

Algorithm 1: Short-term Stock Market Price Trend Prediction - Feature Engineering using FE + RFE + PCA

```
function FE(df)
  # Apply only the meaningful methods on data
  df_expandedfeatures  = Max-MinScaling(df)
  df_expandedfeatures  = Polarizing(df)
  df_expandedfeatures  = CalcFluctuationPercentage(df)
  return df_expandedfeatures
end function

function RFE(df)   # (Utilizing Recursive Feature Elimination function)
  Train the model on all the features of the training dataset in df
  Calculate performance of the model with samples from the test data
  Rank the weights of different features based on testing the model
  for each subset do
    Retain i most weighted features
    Train the model on all the features of the training dataset
    Calculate performance of the model with samples from the test data
  end for
  Calculate the overall performance profile for each feature over samples from the test data
  Rank and select top ranked features
  Train the model on the selected features using the training dataset in df
  return df_RFE    # (df_RFE is the processed data frame after RFE algorithm)
end function

function PCA (df)     # (Utilizing PCA to reduce dimension from i to j)
   df_PCA = applyPCA (n_components=j, whiten=False, copy=True, batchsize = 200)
   return df_PCA     # df_PCA is the optimized data frame after applying PCA algorithm)
end function

function MAIN()                    # (Main function)
  df_alldata = load data
  df_partition = DataPartition(df_alldata, method = resampling)
  df_FE = FE(df_partition)
  df_RFE = RFE(df_FE)
  df_PCA = PCA(df_RFE)
  return df_PCA
end function
```

Рисунок 5.2 – Опис першого алгориму

Якщо ACR нижче 85%, то метод PCA не підходить через втрату вихідної інформації. Оскільки коваріаційна матриця чутлива до порядку величин даних, перед проведенням PCA необхідно провести процедуру стандартизації даних. Зазвичай використовуються такі методи стандартизації, як стандартизація за середнім значенням і стандартизація за нормою.

5.7.2 Модель прогнозування цінового тренду з використанням LSTM

Після вилучення головних компонент ми отримаємо матрицю зі зменшеним масштабом, що означає, що І найбільш ефективних ознак

перетворюються в J головних компонент для навчання моделі прогнозування.

Ми використовували модель LSTM і додали процедуру перетворення для нашого набору даних про ціни на акції. Функція TimeSeriesConversion () перетворює матрицю головних компонент у часовий ряд шляхом зсуву вхідного кадру даних відповідно до кількості часових кроків, тобто довжини терміну в даному дослідженні.

Оброблюваний набір даних складається з вхідної послідовності та послідовності прогнозів. Алгоритм зображен на рисунку 5.3.

У даному дослідженні параметр LAG дорівнює 1, оскільки модель виявляє картину коливання ознак на щоденній основі. При цьому N_TIME_STEPS варіюється від 1 торгового дня до 10 торгових днів.

Algorithm 2

Algorithm 2: Price Trend Prediction Model using LSTM

```
function TimeSeriesConversion (df, term_length, lag)
   # Utilizing time series conversion technique to convert the training data matrix, after applying PCA
from Algorithm 1, to time series
   cols = list()
   for i in range (term_length, 0, -1) do                    # Input sequence
      shift df by i
      append shifted df to cols
   end for
   for i in range (0, lag) do                                # Forecast sequence
      shift df by -1
      append shifted df to cols
   end for
   df_TS = concat(cols, axis = 1)                            # Put all sequences together
   return df_TS
end function

function ModelCompile()                                      # Applying LSTM model with given structure
and compiling it
   Stack_method = Sequential()
   Layer_1 = LSTM(50, input _shape=(train_X.shape[1], train_X.shape[2]))
   Layer_2 = Dense(1)
   Loss_Function=mae
   Optimizer=adam
   Metrics=f1, metrics.binary_accuracy, metrics.mean_squared_error, metrics.mean_absolute_error
   return LSTMmodel
end function

function MAIN()                                              # Main Function
   df_TS = TimeSeriesConversion(df_PCA, N_TIME_STEPS, LAG)
   DataPartition(df_TS, method = resampling)
   ModelCompile(j)
   FitModel(X, y, epochs=50, batch_size=3000)                # Train and fit the model
   EvaluateModel(X_test, y_test)                             # Calculate evaluation metrics on the trained
model using test data
end function
```

Рисунок 5.3 – Опис другого алгоритму

Функції DataPartition(), FitModel(), EvaluateModel() є звичайними кроками без налаштування. Дизайн структури NN, рішення оптимізатора та інші параметри показані у функції ModelCompile().

6 РЕЗУЛЬТАТИ ТЕХНІЧНОГО ПРОЕКТУ

Деякі процедури впливають на ефективність, але не впливають на точність, і навпаки, в той час як інші процедури можуть впливати як на ефективність, так і на результат передбачення. Щоб повністю оцінити дизайн нашого алгоритму, ми структурували частину оцінки за основними процедурами і оцінили, як кожна процедура впливає на продуктивність алгоритму. По-перше, ми оцінили наше рішення на машині з процесором i7-10700K 3,8 ГГц та 16 ГБ оперативної пам'яті. Крім того, ми також оцінили наше рішення на екземплярі Amazon EC2, процесорі 2,2 ГГц з 12 vCPU та 12 ГБ оперативної пам'яті.

У частині реалізації ми розширили 20 ознак до 54, зберігши при цьому 30 з них, які є найбільш ефективними. У цьому розділі ми обговоримо оцінку вибору ознак. Набір даних був розділений на дві різні підмножини, тобто тренувальний і тестовий набори даних. Процедура тестування включала дві частини, одна тестова база даних призначена для відбору ознак, а інша – для тестування моделі. Ми позначаємо набір даних для відбору ознак і набір даних для тестування моделі як ds_test_f і DS_test_m, відповідно.

Для навчання RFE ми випадковим чином відібрали дві третини даних акцій за ідентифікатором акції і позначили набір даних як DS_train_f; всі дані складаються з повних технічних індексів і розширених характеристик за 2021 рік. Оцінювачем алгоритму RFE є SVR з лінійними ядрами. Ми ранжуємо 54 ознаки шляхом голосування і отримуємо 30 ефективних ознак, потім обробляємо їх за допомогою алгоритму PCA для зменшення розмірності і зведення ознак до 20 головних компонентів. Решта даних по акціях формує тестовий набір даних DS_test_f для перевірки ефективності головних компонент, витягнутих з обраних ознак. Всі дані за 2021 рік ми переформували в тренувальний набір даних моделі даних і відзначили як DS_train_m. набір даних для тестування моделі DS_test_m складається з

даних за перші 3 місяці 2022 року, які не перетинаються з набором даних, який ми використовували на попередніх етапах. Такий підхід дозволяє запобігти приховану проблему, викликану надмірною підгонкою.

6.1 Довжина терміна

Для побудови ефективної моделі прогнозування замість моделювання даних у часовий ряд ми вирішили використовувати дані за індексами на 1 день вперед для прогнозування цінової тенденції наступного дня. Ми протестували алгоритм RFE на діапазоні короткострокових періодів від 1 дня до 2 тижнів, щоб оцінити, як широко використовувані технічні індекси співвідносяться з ціновими тенденціями. Для оцінки довжини терміну прогнозування ми повністю розширили ознаки, як показано на рисунку 6.1, і подали їх в RFE.

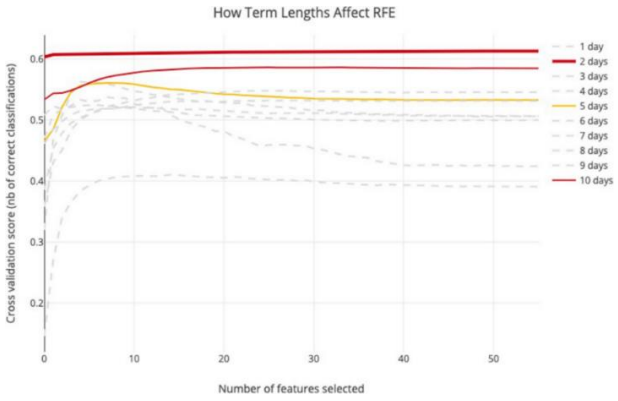


Рисунок 6.1 – Як довжина терміну впливає на крос-валідаційну оцінку RFE

Хоча ці криві мають різний характер, при довжині терміну прогнозування 2 тижні показник крос-валідації збільшується з ростом числа обраних ознак. Якщо довжина терміну прогнозування становить 1 тиждень, оцінка крос-валідації зменшується, якщо вибрано більше 8 ознак.

Для прогнозування цінового тренду на кожен другий день найкращий результат крос-валідації досягається при виборі 48 ознак. Для

прогнозування раз на два тижні для досягнення найкращого результату потрібно 29 ознак.

На рисунку 6.2 ми перерахували 15 найбільш ефективних ознак для цих трьох періодів. Якщо ми прогнозуємо цінову тенденцію на кожен другий день, то результат крос-валідації просто коливається в залежності від кількості обраних ознак. Тому на наступному етапі ми оцінимо результат RFE для цих трьох періодів, як показано на рисунку 6.1.

Relevant ranking	Every other day	Weekly	Bi-weekly
1st	Up_down	SLOWK_maxmin	MTM_10_plr
2nd	Change	SLOWK	ROC_10_plr
3rd	pct_chg	SLOWD_maxmin	WNR_9
4th	Low	RSI_5_maxmin	WNR_9_maxmin
5th	RSI_5_flc	SLOWD	SLOWK
6th	Open	RSI_5	SLOWK_maxmin
7th	Amount	SLOWK_flc	ROC_10
8th	Amount_maxmin	WNR_9_maxmin	SLOWD_flc
9th	Vol	WNR_9	WNR_9_flc
10th	BIAS_20_maxmin	CCI_24	RSI_5
11th	High	BIAS_20_maxmin	BIAS_20_maxmin
12th	Vol_maxmin	BIAS_20	RSI_5_maxmin
13th	ROC_10	ADOSC_maxmin	BIAS_20
14th	ADOSC_maxmin	ADOSC	SMA_10
15th	ADOSC	WNR_9_flc	SLOWD
...	...		...
Number of Features Selected	48 features selected	8 features selected	29 features selected

Рисунок 6.2 – Ефективні характеристики, відповідні тривалості терміну

Ми порівнюємо вихідний набір ознак RFE з вихідним набором ознак в якості базового, вихідний набір ознак складається з n ознак, і ми вибираємо N найбільш ефективних ознак з вихідних ознак RFE для оцінки результату за допомогою лінійного SVR. Ми використовували два різних підходи для оцінки ефективності ознак.

Перший метод – об'єднати всі дані в одну велику матрицю і оцінити їх шляхом одноразового запуску алгоритму RFE.

Інший метод – запустити RFE для кожної окремої акції і обчислити найбільш ефективні ознаки шляхом голосування.

6.2 Розширення ознак та RFE

З результатів попереднього підрозділу видно, що при прогнозуванні цінового тренда на кожен другий день або раз на два тижні найкращий результат досягається при виборі великої кількості ознак. Серед відібраних ознак деякі ознаки, оброблені за допомогою методів розширення, мають кращі ранги, ніж вихідні ознаки, що доводить корисність методу розширення ознак для оптимізації моделі.

Розширення ознак впливає як на точність, так і на ефективність, але в цій частині ми обговорюємо тільки аспект точності, а ефективність залишимо на наступний крок, оскільки PCA є найбільш ефективним методом оптимізації ефективності навчання в нашому проекті.

Ми провели оцінку того, як розширення функцій впливає на RFE, і використовували результати тестів для вимірювання поліпшення від розширення функцій.

Далі ми перевіряємо ефективність розширення ознак, тобто якщо поляризація, макс-хв масштаб і обчислення відсотка коливань працюють краще, ніж оригінальні технічні показники.

Найкращим випадком для використання цього тесту є щотижневий прогноз, оскільки для нього обрана найменш ефективна функція. З результатів, отриманих в попередньому розділі, ми знаємо, що найкращий результат крос-валідації з'являється при виборі 8 функцій.

Тест складається з двох етапів, на першому етапі тестується набір ознак, сформований тільки з вихідних ознак, в даному випадку тільки

SLOWK, SLOWD і RSI_5. Наступний крок-тестування набору ознак, що складається з усіх 8 ознак, які ми вибрали в попередньому підрозділі.

Ми використовували тест, визначивши найпростішу модель ДНК з трьома шарами.

Нормована матриця плутанини при тестуванні двох наборів ознак показана на рисунку 6.3.

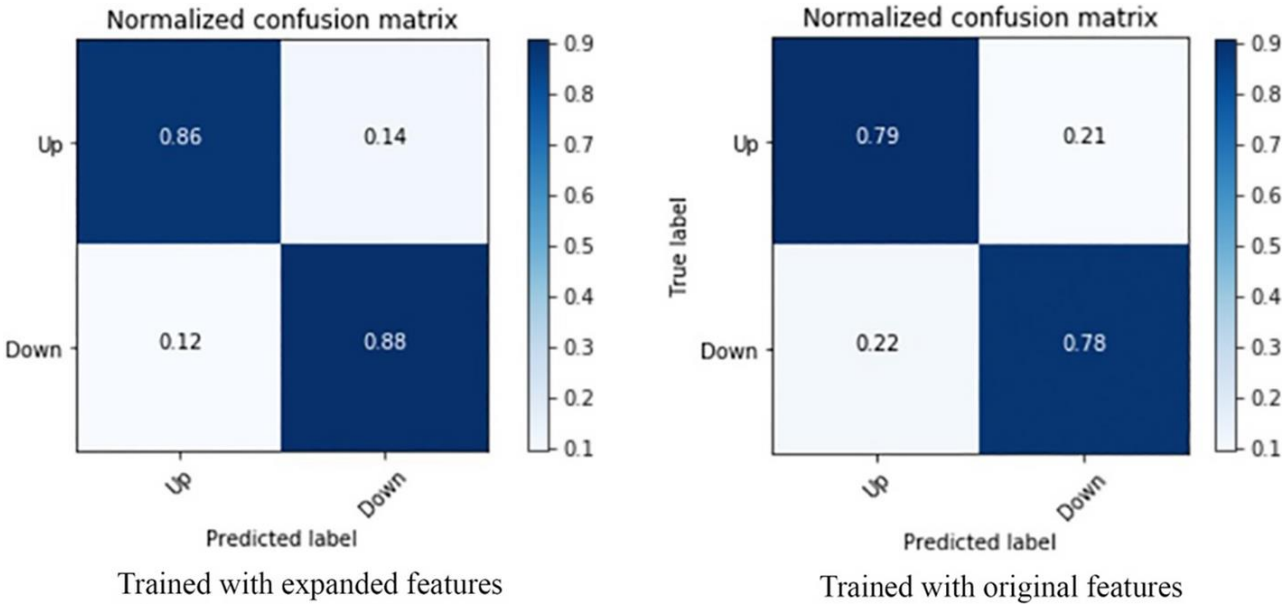


Рисунок 6.3 – Матриця плутанини для перевірки ефективності розширення функції

Зліва – матриця плутанини набору ознак з розширеними ознаками, а праворуч-результат тестування з використанням тільки оригінальних ознак.

Обидві точності істинно позитивних і істинно негативних результатів були покращені на 7% і 10%, відповідно, що доводить, що наша розробка методу розширення ознак досить ефективний вибір.

6.3 Скорочення ознак за допомогою аналізу головних компонент

PCA вплине на продуктивність алгоритму як на точність прогнозування, так і на ефективність навчання, в той час як ця частина повинна бути оцінена за допомогою моделі NN, тому ми також визначили найпростішу модель DNN з трьома шарами, як ми використовували в попередньому кроці для проведення оцінки.

У цій частині представлені метод оцінки та результат оптимізації моделі з точки зору обчислювальної ефективності та впливу на точність.

DNN – у найпростішому випадку нейронна мережа з певним рівнем складності, зазвичай не менше двох шарів, кваліфікується як глибока нейронна мережа (ГНМ), або скорочено глибока мережа. Глибокі мережі обробляють дані складним чином, використовуючи складне математичне моделювання. Вихідні дані отримуються шляхом контрольованого навчання з використанням наборів даних, що містять певну інформацію, засновану на «те, що ми хочемо», за допомогою зворотного поширення.

У цьому розділі ми виберемо двотижневе передбачення для аналізу прикладу, оскільки воно має плавно зростаючу криву крос-валідаційної оцінки, крім того, на відміну від всіх інших денних прогнозів, воно вже виключило більше 20 неефективних ознак.

На першому етапі ми вибираємо всі 29 ефективних ознак і навчаємо NN-модель без виконання PCA. Це створює базовий рівень точності та часу навчання для порівняння.

Щоб оцінити точність і ефективність, ми змінюємо кількість головних компонент на 5, 10, 15, 20, 25.

На рисунку 6.8 показаний аналіз точності та ефективності різних процедур попередньої обробки ознак.

Часові витрати, наведені на рисунку 6.4 та 6.8, засновані на експериментах, проведених на стандартній машині користувача, щоб

показати життєздатність нашого рішення при обмеженій або середній кількості ресурсів.

Number of features	Training dataset preparation time (s)	Test dataset preparation time (s)	Training time (s)	Sum (s)
29 selected features	187.46	16.30	648.53	852.29
20 principal components	160.29	14.24	602.68	777.21
15 principal components	125.20	дек.18	591.93	729.31
10 principal components	96.54	окт.37	590.76	697.67
5 principal components	59.37	авг.22	572.88	640.47

Рисунок 6.4 – Взаємозв'язок між кількістю головних компонент і ефективністю навчання

На рисунку 6.4 записано, як кількість ознак впливає на ефективність навчання моделі. На рисунку 6.5 ілюструється, як PCA впливає на ефективність навчання.

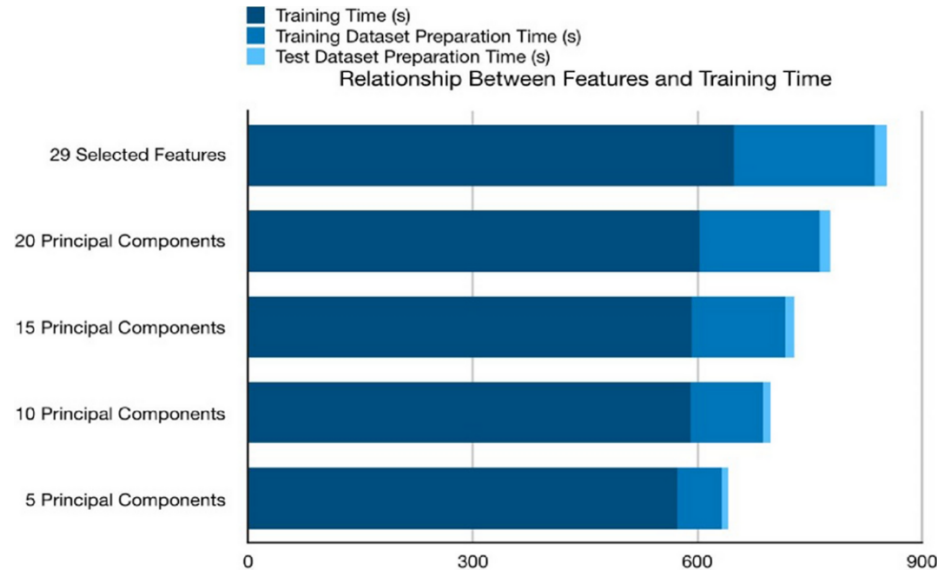


Рисунок 6.5 – Взаємозв'язок між кількістю ознак і часом навчання

Ми також привели матрицю заплутаності кожного тесту на рисунку 6.6. Діаграма показує, що загальний час, що витрачається на навчання

моделі, зменшується з ростом числа обраних ознак, а метод PCA значно ефективніше в оптимізації підготовки навчального набору даних.

Для часу, що витрачається на етап навчання, PCA не так ефективний, як на етапі підготовки даних.

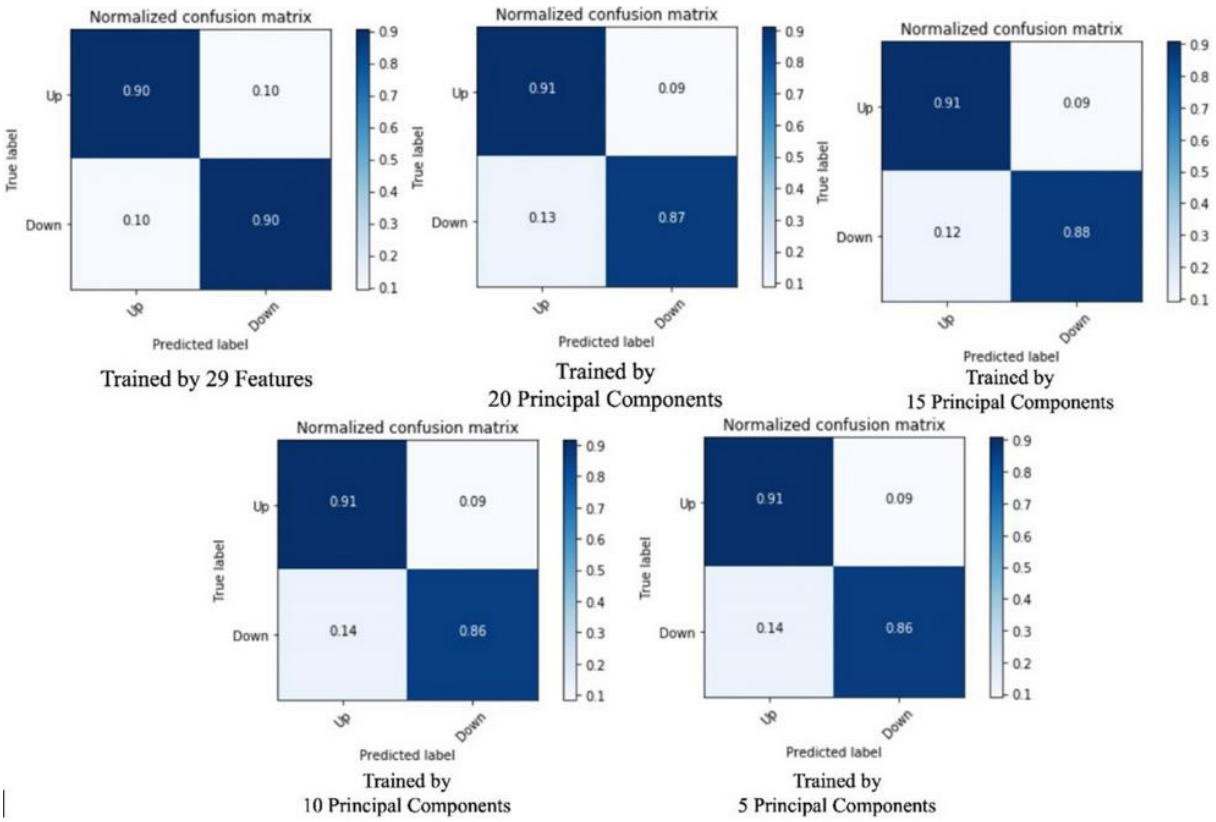


Рисунок 6.6 – Як кількість головних компонент впливає на результати оцінки

Тож можна зробити висновок, що PCA не має сильного негативного впливу на точність передбачення.

Рис 6.7 показує, що зниження розмірності не робить істотного впливу на загальну точність прогнозування.

Однак точність не може повністю підтвердити, що PCA не має побічного ефекту для прогнозування моделі, тому ми розглянули матриці плутанини результатів тестування.

Number of selected features	5 principal components	10 principal components	15 principal components	20 principal components	29 selected features
Accuracy	89.03%	89.35%	89.39%	89.30%	90.29%

Рисунок 6.7 – Як кількість вибраних ознак впливає на точність прогнозування

Показники істинно позитивних і хибнопозитивних результатів майже не змінилися, в той час як показники помилково негативних і істинно негативних результатів змінилися на 2-4%.

Окрім оцінки того, як кількість вибраних ознак впливає на ефективність навчання та продуктивність моделі, ми також використовували тест на те, як процедури попередньої обробки даних впливають на процедуру навчання та результат прогнозування.

Нормалізація і max-min масштабування – найбільш часто зустрічаються процедури попередньої обробки даних, що виконуються перед PCA, оскільки одиниці виміру ознак різні, і вважається, що це може підвищити ефективність навчання згодом.

Feature pre-processing	Overall accuracy (%)	Training dataset preparation time (s)	Testing dataset preparation time (s)	Training time (s)	Sum (s)
Max–min scaling	89.30	160.28	14.24	602.68	777.20
Normalization	78.17	157.63	14.73	596.22	768.58
N/A	78.88	142.17	13.00	595.52	750.69

Рисунок 6.8 – Як кількість вибраних ознак впливає на точність прогнозування

Ми використовували інший тест на додавання попередніх процедур перед витяганням 20 головних компонент з вихідного набору даних і

провели порівняння в аспектах часу, що минув з моменту навчання, і точності передбачення.

Однак результати тесту призводять до різних висновків. З рисунку 6.8 можна зробити висновок, що попередня обробка ознак не робить істотного впливу на ефективність навчання, але впливає на точність передбачення моделі.

Більш того, перша матриця плутанини на рисунку 6.9 показує, що без попередньої обробки ознак сильно страждають показники помилково негативних та істинно негативних результатів, в той час як показники істинно позитивних і помилково позитивних результатів не страждають.

Якщо виконати нормалізацію перед PCA, то показники справжніх позитивних і справжніх негативних результатів знижуються приблизно на 10%.

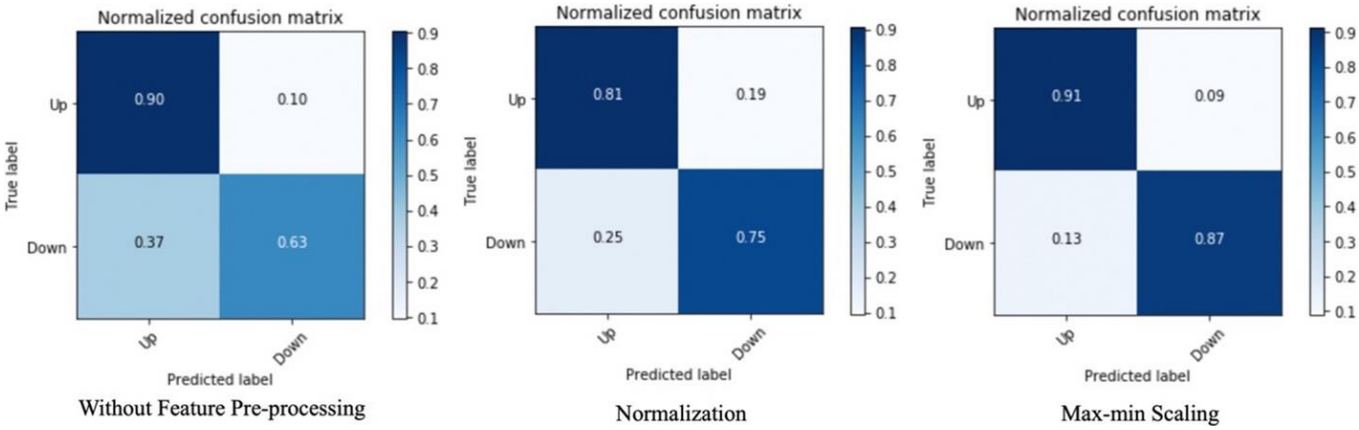


Рисунок 6.9 – Матриці змішування різних методів попередньої обробки ознак

Цей тест також довів, що найкращим методом попередньої обробки ознак для нашого набору ознак є використання шкали max-min.

7 АНАЛІЗ ТЕХЧНІЧНОГО ПРОЕКТУ

7.1 Порівняння з аналогічними роботами

З попередніх робіт ми з'ясували, що найбільш часто використовуваними моделями для короткострокового прогнозування тенденцій цін на фондовому ринку є машина опорних векторів(SVM), багатошарова перцептронна штучна нейронна мережа (MLP), Класифікатор Naive-Bayes(NB), Класифікатор випадкового лісу (RFC) і Класифікатор логістичної регресії (LR). Тестовим прикладом порівняння також є двотижневий прогноз цінового тренда, для оцінки найкращого результату всіх моделей ми зберігаємо всі 29 ознак, відібраних алгоритмом RFE. Для оцінки MLP, щоб перевірити, чи впливає кількість прихованих шарів на метричні оцінки, ми відзначили кількість шарів як n і протестували n = {1, 3, 5}, 150 епох навчання для всіх тестів, виявили незначні відмінності в продуктивності моделі, що вказує на те, що змінна кількості шарів MLP майже не впливає на метричні оцінки.

З матриць плутанини на рисунку 7.1 видно, що всі моделі машинного навчання показують хороші результати при навчанні на повному наборі ознак, відібраних за допомогою RFE. З точки зору часу навчання, найкращу ефективність показало навчання моделі NB. Алгоритм LR вимагає менше часу на навчання, ніж інші алгоритми, при цьому він може досягти аналогічного результату передбачення з іншими дорогими моделями, такими як SVM і MLP. Алгоритм RAF досяг відносно високого показника істинно-позитивних результатів, в той час як при передбаченні негативних міток він показав низьку ефективність. Запропонована нами модель LSTM досягає бінарної точності 93,25%, що є значно високою точністю передбачення двотижневого цінового тренда. Ми також попередньо обробили дані за допомогою PCA і отримали п'ять головних компонент, потім провели навчання протягом 150 епох. Крива навчання

запропонованого нами рішення, заснованого на інжинірингу ознак і моделі LSTM, показана на рисунку 7.2. Матриця плутанини показана праворуч на рисунку 7.3, а детальні оцінки метрик наведені на рисунку 7.4.

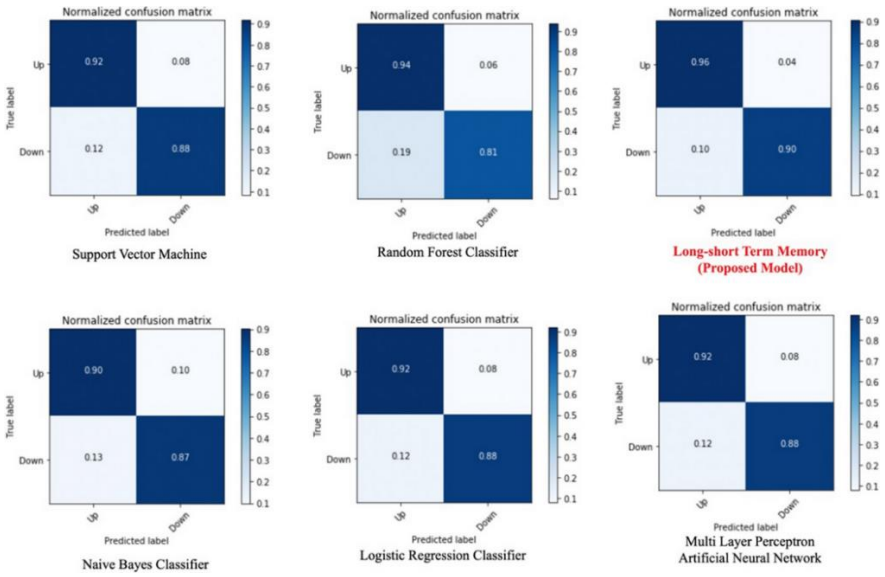


Рисунок 7.1 – Матриці порівняння-конфузії передбачень моделей

Нижче зображенна крива навчання пропонованого рішення

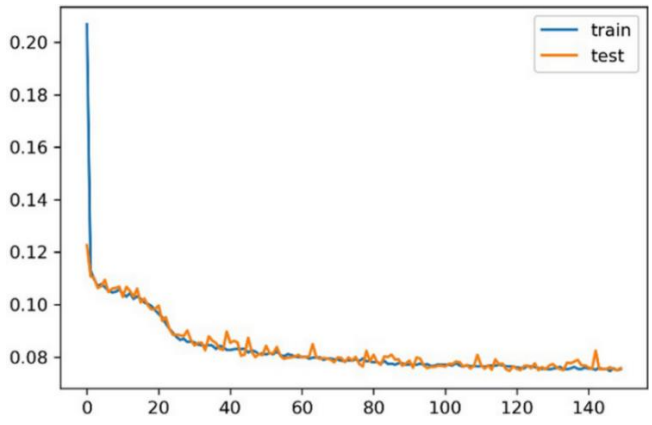
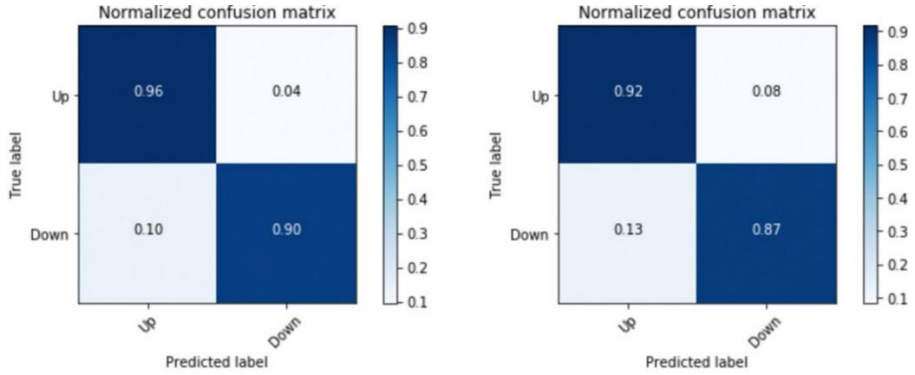


Рисунок 7.2 – Крива навчання пропонованого рішення

Нижче зображена модель передбачення точності порівняння-конфузії матриць.

Ця модель тренувалась за допомогою 29 ознак, а також за допомогою 5 головних компонентів.



LSTM Model
Trained by 29 Features

LSTM Model
Trained by 5 Principal Components

Рисунок 7.3 – Запропонована модель передбачення точності порівняння-конфузії матриць

Детальні результати оцінки ілюстровані на рисунку 7.4. У наступному розділі ми також почнемо обговорення результатів оцінки.

Model	F1 score	Binary accuracy	TPR (recall)	TNR (specificity)	FPR (fall-out)	FNR (miss rate)
LR	0.90	0.90	0.92	0.88	0.08	0.12
SVM	0.90	0.90	0.92	0.88	0.08	0.12
NB	0.89	0.89	0.90	0.87	0.10	0.13
MLP (Single hidden layer)	0.90	0.90	0.92	0.88	0.08	0.12
MLP (3 hidden layers)	0.90	0.90	0.92	0.87	0.08	0.13
MLP (5 hidden layers)	0.90	0.90	0.92	0.88	0.08	0.12
RAF	0.88	0.88	0.94	0.81	0.06	0.19
Propose d model	0.93	0.93	0.96	0.90	0.04	0.10

Рисунок 7.4 – Порівняння продуктивності моделі-метричні оцінки

Оскільки результуюча структура запропонованого нами рішення відрізняється від більшості відповідних робіт, було б важко провести наївне порівняння з попередніми роботами.

Наприклад, важко знайти точне число точності передбачення цінового тренда в більшості робіт, оскільки автори вважають за краще показувати коефіцієнт прибутку імітованих інвестицій.

Коефіцієнт прибутку-це оброблене число, засноване на імітаційних інвестиційних тестах, іноді одне правильне інвестиційне рішення з великим торговим обсягом може досягти високого коефіцієнта прибутку незалежно від точності прогнозування цінового тренда.

Крім того, унікальним і евристичним нововведенням в пропонованому нами рішенні є перетворення проблеми прогнозування точної ціни в дві послідовні проблеми, тобто. спочатку прогнозування цінового тренду, зосередження на побудові точної моделі бінарної класифікації, створення міцного фундаменту для прогнозування точної зміни ціни в майбутніх роботах.

Крім різної структури результатів, набори даних, на яких проводилися дослідження в попередніх роботах, також відрізняються від нашої роботи. Деякі з попередніх робіт використовують дані новин для аналізу настроїв і використовують частину SE як ще один компонент системи для підтримки своєї моделі прогнозування.

Остання робота, яку можна порівняти, це Броунлі [3], автор використовує множинний R-квадрат для вимірювання точності моделі. Множинний R-квадрат також називається коефіцієнтом детермінації, і він показує силу передбачувальних змінних, що пояснюють варіації прибутковості акцій [28]. Авторп використовував три набори даних (Індекс KSE 100, акції Lucky Cement, Engro Fertilizer Limited)для оцінки запропонованої моделі множинної регресії і досягли 95%, 89% і 97% відповідно. За винятком індексу KSE 100, вибір набору даних в даній роботі-це окремі акції; таким чином, ми вибрали результат оцінки першого набору даних запропонованої ними моделі.

На рисунку 7.5 ми перерахували провідні моделі прогнозування тренду цін на акції, з порівнянних метрик, метричні оцінки запропонованого нами рішення в цілому краще, ніж в інших суміжних роботах.

Замість того щоб робити довільний висновок про те, що запропонована нами модель перевершила інші моделі у відповідних роботах, ми спочатку подивимося на стовпець набору даних на рисунку 7.5.

Related work	Dataset	Model	Accuracy	Precision	Recall
Atsalakis [1]	Stock price data of AAPL, GE and Samsung Electronics Co. Ltd.	Random forest	0.83	0.82	0.81
Bharat [2]	Close price of stock data from New York Stock Exchange (NYSE)	ARIMA	0.90	0.91	0.92
Brounly [3]	KSE 100 Index Lucky Cement Stock Engro Fertilizer Limited	Multiple regression	0.94	0.95	0.93
(Proposed solution)	*Price data of 3558 stock ID from 2017 to 2018 collected from Chinese stock market*	*Proposed Model —FE + RFE + PCA + LSTM*	0.93	0.96	0.96

Рисунок 7.5 – Порівняння пропонованого рішення з аналогічними роботами

Якщо подивитися на набір даних, який використовується в кожній роботі [1], то він навчав і тестував запропоноване ним рішення тільки на трьох окремих акціях, що важко довести узагальненість запропонованої ними моделі. Бхарат [2] використовував аналіз даних по акціях з Нью-Йоркської фондової біржі (NYSE), але недоліком є те, що вони проводили аналіз тільки за ціною закриття, яка є характеристикою з високим рівнем шуму. Броунли [3] навчав запропоновану ними модель як на окремих акціях, так і на ціні індексу, але, як ми вже згадували в попередньому розділі, ціна індексу складається тільки з обмеженого числа характеристик і ідентифікаторів акцій, що ще більше погіршує якість навчання моделі.

Для запропонованого нами рішення ми зібрали достатньо даних з китайського фондового ринку і застосували алгоритм FE + RFE на оригінальних індексах для отримання більш ефективних ознак, результат комплексної оцінки 3558 ідентифікаторів акцій може розумно пояснити

узагальнення та ефективність пропонованого нами рішення на китайському фондовому ринку.

7.2 Оцінка ефективності запропонованої моделі-PCA

Крім порівняння ефективності популярних моделей машинного навчання, ми також оцінили, як алгоритм PCA оптимізує процедуру навчання запропонованої моделі LSTM. На рисунку 7.3 ми записали порівняння матриць плутанини між навчанням моделі за 29 ознаками і за п'ятьма головними компонентами.

Metrics name	LSTM trained on 29 features	LSTM trained on 5 principal components
Loss	0.0702	0.0848
F1 score	0.9323	0.9194
Binary accuracy	0.9325	0.9193
MSE	0.0669	0.0772
MAE	0.0702	0.0848
TPR	0.96	0.92
TNR	0.90	0.91
FPR	0.04	0.08
FNR	0.10	0.09

Рисунок 7.6 – Порівняння продуктивності пропонованої моделі – з PCA і без нього

Навчання моделі за повними 29 ознаками займає в середньому 28,5 с на епоху. У той час як на навчання по набору з п'яти головних компонент йде в середньому 18 с на епоху. PCA значно підвищила ефективність навчання моделі LTM на 36,8%.

ВИСНОВКИ

Дана робота складається з трьох частин: Витяг даних і попередня обробка набору даних китайського фондового ринку, проведення інженерії ознак і прогнозування тренда цін на акції на основі довготривалої короткочасної пам'яті (LSTM).

Було зібрано, очищено та структуровано дані китайського фондового ринку за 2 роки.

Було розглянуто різні методи, часто використовувані реальними інвесторами, розроблено новий компонент алгоритму – розширення ознак, який довів свою ефективність.

Були застосовані підходи розширення ознак (FE) та рекурсивного усунення ознак (RFE) з подальшим аналізом головних компонент (PCA), щоб створити процедуру інженерії ознак, яка є одночасно ефективною і дієвою.

Система, налаштована шляхом об'єднання процедури інжинірингу ознак з моделлю прогнозування LSTM, досягла високої точності прогнозування, що перевершує провідні моделі в більшості відповідних робіт.

Також була проведена комплексна оцінка цієї роботи. Порівнюючи найбільш часто використовувані моделі машинного навчання з запропонованою нами LSTM-моделлю в рамках функції інженерії запропонованої нами системи, ми зробили безліч евристичних висновків, які можуть стати питаннями майбутніх досліджень як в технічних, так і у фінансових областях.

Запропоноване рішення є унікальною розробкою в порівнянні з попередніми роботами, оскільки замість того, щоб просто запропонувати ще одну сучасну модель LSTM, було запропонована тонко налаштована та адаптована система прогнозування на основі глибинного навчання, а також

використанна комплексна інженерія ознак та об'єднання її з LSTM для виконання прогнозування.

Досліджуючи спостереження з попередніх робіт, були заповнені прогалини між інвесторами та дослідниками, запропонувавши алгоритм розширення ознак перед рекурсивним виключенням ознак і домігшись помітного поліпшення продуктивності моделі. Хоча вдалося досягти гідного результату від запропонованого рішення, дане дослідження має більший потенціал для вивчення в майбутньому.

В ході процедури оцінки було виявлено, що алгоритм RFE не чутливий до довжини термінів, відмінних від 2-денних, щотижневих і двотижневих. Більш глибоке дослідження того, які технічні індекси можуть впливати на нестандартну довжину термінів, було б можливим напрямком майбутніх досліджень.

Більш того, об'єднавши новітні методи аналізу настроїв з функціоналом і моделлю глибокого навчання, можна розробити більш комплексну систему прогнозування, яка буде навчатися на різних типах інформації, таких як твіти, новини та інші текстові дані.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Атсалакіс Г.С., Валаваніс К.П. Прогнозування короткострокових тенденцій фондового ринку за допомогою нейро-нечіткої методології. . 2017. Vol. 3(1). P.411–512.

2. Бхарат К.М. Прогнозування цін на акції з використанням моделі ARIMA. URL: https://www.businessperspectives.org (дата звернення: 15.04.2022)

3. Броунли Дж. Глибоке навчання для прогнозування часових рядів: передбачення майбутнє за допомогою MLP, CNN та LSTM в Python. URL: https://machinelearningmastery.com (дата звернення: 15.04.2022)

4. Фішер Т., Краус С. Глибоке навчання з використанням мереж з довготривалою короткочасною пам'яттю для прогнозування фінансових ринків. URL: https://deep-learning-nn.com/21.8456/k.erot (дата звернення: 15.04.2022)

5. Гійон І., Вестон Дж. , Вапнік В. Відбір генів для класифікації раку з використанням машин з опорними векторами. URL: https://gene-machine-learning.com/index.html (дата звернення: 15.04.2022)

7. Crash Course in Recurrent Neural Networks for Deep Learning. URL: https://machinelearningmastery.com/recurrent-neural-networks-deeplearning (дата звернення: 15.04.2022)

8. Рябова Н.В., Пахомов І.Ю. Нейромережевий підхід до прогнозування фінансового ринку та побудови інвестиційного портфелю. Матеріали XII Міжнародної науково-технічної конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління». Баку-Харків-Жиліна, 27-28 квітня 2022

9. Graves A., Mohamed A. and Hinton G., «Speech recognition with deep recurrent neural networks,» in Acoustics, Speech and Signal Processing. 2020. Vol. 1(1). P.502–589.

10. Maknickiene N., Investigation of financial market prediction by recurrent neural network. 2011. Vol. 2(2). P.3-8.

11. Iqbal Z., Mahmood Z.and Anjum J., Efficient Machine Learning Techniques for Stock Market Prediction. 2013. Vol. 3. P.855-867.

12. Box F., Jenkins G., and Reinsel I., Time Series Analysis: Forecasting and Control. 1994. Vol. 3. P.565-768.

13. Patton A. J., Volatility forecast comparison using volatility proxies. 2011. Vol. 160. P.246–256.

14. Kodogiannis V. and Lolis A., Forecasting financial time series using neural network and fuzzy system-based techniques. 2002. Vol. 11. P.90–100.

15. Refenes A.N, Constructive learning and its application to currency exchange rate forecasting. 1993. Vol. 4. P.465–493.

16. Colland, F.E., Advances in Neural Information Processing System. 2021. Vol. 43. P.551–556.

17. Fletcher, D., Goss, E. Forecasting with neural networks: An application using bankruptcy data. 2016. Vol. 160. P. 159–168.

18. Bontempi G, Taieb S B, Yann-Aël Le Borgne. Machine learning strategies for time series forecasting. 2013. Vol. 89. P. 20–47.

19. Jiang Q, Tang C, Chen C, et al. Stock Price Forecast Based on LSTM Neural Network. 2018. Vol. 101. P. 100–111.

20. Kim J, El-Khamy M, Lee J. Residual LSTM: Design of a Deep Recurrent Architecture for Distant Speech Recognition. 2020. Vol. 3. P. 89–97.