

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Інформаційних управляючих систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження та використання методів прогнозування при розробці ІС
для планування роботи підприємства електронної торгівлі

(тема)

Виконав:

здобувач 2 року навчання, групи ІУСТм-23-1
Тимошицький Богдан Олексійович
(прізвище, ім'я, по батькові)

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

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

Освітня програма Інформаційні управляючі системи та технології
(повна назва освітньої програми)

Керівник зав. каф. ІУС Петров К.Е.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ІУС



(підпис)

Петров К.Е.

(посада, прізвище, ініціали)

2025 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Інформаційних управляючих систем


Рівень вищої освіти другий (магістерський)

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

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

Освітня програма Інформаційні управляючі системи та технології
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри 
(підпис)

« 09 » грудня 2024 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Тимошицькому Богдану Олексійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження та використання методів прогнозування при розробці ІС для планування роботи підприємства електронної торгівлі

затверджена наказом по університету від « 27 » листопада 2024 р. № 1249Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії _____

3. Вхідні дані до роботи матеріали передатестаційної практики, аналітичні матеріали про методи прогнозування, інформація про тенденції ринку електронної торгівлі, дані продажів підприємства електронної торгівлі, додаткові джерела інформації.


4. Перелік питань, що потрібно опрацювати у роботі Аналіз предметної області, аналіз існуючих методів прогнозування, аналіз проблем використання методів прогнозування для планування роботи підприємства електронної торгівлі, постановка задачі дослідження, дослідження методів прогнозування, формування порівняльних критеріїв, вибір методів прогнозування для подальшого комбінування, розробка комбінованого методу прогнозування, експериментальна перевірка отриманих результатів на практичному прикладі.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз предметної області	09.12.2024 – 15.12.2024	Виконано
2	Аналіз існуючих методів прогнозування	15.12.2024 – 17.12.2024	Виконано
3	Аналіз проблем використання методів прогнозування при розробці ІС для планування роботи підприємства електронної торгівлі	17.12.2024 – 18.12.2024	Виконано
4	Формулювання постановки задачі	18.12.2024 – 19.12.2024	Виконано
5	Дослідження методів прогнозування	19.12.2024 – 25.12.2024	Виконано
6	Оцінка можливостей використання обраних методів	25.12.2024 – 02.01.2025	Виконано
7	Розробка комбінованого методу прогнозування	02.12.2024 – 08.01.2025	Виконано
8	Експериментальна перевірка можливостей використання розробленого комбінованого методу прогнозування	08.01.2025 – 10.01.2025	Виконано
9	Оформлення пояснювальної записки	10.01.2025 – 12.01.2025	Виконано
10	Оформлення графічних матеріалів	12.01.2025 – 13.01.2025	Виконано
11	Захист кваліфікаційної роботи	22.01.2025	

Дата видачі завдання 09 грудня 2024 р.

Здобувач 
(підпис)

Керівник роботи 
(підпис)

зав. каф. ІУС Петров К.Е
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 69 с., 35 рис., 9 табл., 1 дод., 20 джерел.

АНАЛІЗ ПРОДАЖІВ, ЕЛЕКТРОННА ТОРГІВЛЯ, ІМОВІРНІСНЕ ПРОГНОЗУВАННЯ, КОМБІНАЦІЯ МЕТОДІВ, МЕТОДИ ПРОГНОЗУВАННЯ, SARIMA.

Об'єктом дослідження кваліфікаційної роботи є процес прогнозування соціально-економічних показників діяльності підприємств.

Предметом дослідження є методи прогнозування показників діяльності підприємств електронної комерції.

Метою дослідження є аналіз існуючих методів прогнозування параметрів діяльності підприємств електронної комерції, а також розробка комбінованого підходу до вирішення задачі прогнозування на основі інтеграції відомих та нових методів для підвищення якості та точності прогнозування показників діяльності підприємств.

В процесі виконання роботи було проведено порівняння методів прогнозування за обраними критеріями. Детально розглянуто процес прогнозування з використанням різних методів. Запропоновано комбінований метод прогнозування, для покращення результатів. Наведено опис експериментальної перевірки методу та проведено оцінку точності прогнозування показників.

Запропонований метод є корисним інструментом для керівників підприємств електронної торгівлі. Він дозволяє отримувати більш точні прогнозні значення показників, що дає змогу значно підвищити ефективність роботи підприємства.

ABSTRACT

Master's thesis: 69 pages, 35 figures, 9 tables, 1 appendices, 20 sources.

COMBINATION OF METHODS, E-COMMERCE, FORECASTING METHODS, PROBABILISTIC FORECASTING, SALES ANALYSIS, SARIMA.

The object of research of the qualification work is the process of forecasting socio-economic indicators of enterprises.

The subject of research is methods of forecasting the performance of e-commerce enterprises.

The purpose of the study is to analyze existing methods for forecasting the parameters of e-commerce enterprises, as well as to develop a combined approach to solving the forecasting problem based on the integration of known and new methods to improve the quality and accuracy of forecasting the performance of enterprises.

In the course of the work, a comparison of forecasting methods was made according to the selected criteria. The forecasting process using different methods is considered in detail. A combined forecasting method is proposed to improve the results. A description of the experimental verification of the method is given and the accuracy of forecasting indicators is evaluated.

The proposed method is a useful tool for managers of e-commerce enterprises. It allows obtaining more accurate forecast values of indicators, which can significantly improve the efficiency of the enterprise.

ЗМІСТ

	С.
Скорочення та умовні позначки.....	8
Вступ.....	9
1 Аналіз предметної області та постановка задачі дослідження.....	11
1.1 Ринок електронної комерції в Україні.....	11
1.2 Аналіз існуючих методів прогнозування.....	13
1.2.1 Класичне прогнозування.....	13
1.2.2 Квантильне прогнозування.....	14
1.2.3 Імовірнісне прогнозування.....	15
1.2.4 Нейромержеві методи прогнозування.....	16
1.2.5 Порівняння методів прогнозування.....	16
1.3 Аналіз проблем використання методів прогнозування при розробці ІС для планування роботи підприємства електронної торгівлі.....	18
1.4 Постановка задачі дослідження.....	19
2 Дослідження можливостей використання методів прогнозування.....	21
2.1 Оцінка можливостей використання імовірнісних методів прогнозування.....	23
2.2 Оцінка можливостей використання класичних методів прогнозування.....	26
3 Розробка комбінованого методу прогнозування.....	30
3.1 Метод SARIMA.....	34
3.2 Метод часткового імовірнісного прогнозування.....	37
3.3 Оцінка точності прогнозування.....	41
3.4 Комбінований метод прогнозування.....	44
4 Експериментальна перевірка можливостей використання розробленого комбінованого методу прогнозування.....	45

Висновки.....	52
Перелік джерел посилання.....	54
Додаток А Графічний матеріал кваліфікаційної роботи.....	57

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

Бутстрап – практичний метод дослідження розподілення статистик імовірнісного розподілення

ІС – інформаційна система

AR – autoregressive – авторегресія

ARIMA – autoregressive integrated moving-average – авторегресійна інтегрована ковзна середня

BSTS – Bayesian structural time series – Баєсів структурний часовий ряд

KPSS-тест – тест для перевірки стаціонарності часових рядів, що має назву на честь своїх засновників: Kwiatkowski, Phillips, Schmidt, Shin

MA – moving-average – ковзна середня

MAD – mean absolute deviation – середнє абсолютне відхилення

MAPE – mean percentage absolute error – середній відсоток абсолютної похибки

MSE – mean squared error – середня квадратична похибка

OCSB-тест – тест для перевірки стаціонарності часових рядів, що має назву на честь своїх засновників: Osborn, Chui, Smith, Birchenhall

RMSE – root mean square error – середньоквадратична похибка

SARIMA – seasonal ARIMA – сезонна версія ARIMA

ВСТУП

Електронна торгівля постійно стикається з численними проблемами, однією з яких є складність прогнозування соціально-економічних показників. Зміни в економічній ситуації, попит на товари, коливання валютних курсів і інші соціально-економічні фактори можуть значно впливати на ефективність бізнесу. Тому для підприємств, що працюють в онлайн-середовищі, важливо проводити точне і своєчасне прогнозування цих показників. Це дозволяє адаптувати стратегії розвитку, планувати витрати, оптимізувати асортимент та забезпечувати стабільне зростання в умовах мінливого ринку.

Сфера прогнозування існує вже багато років, і за цей час було розроблено безліч різних підходів та методів для прогнозування різноманітних показників, зокрема в економіці, фінансах та бізнесі. В останні роки цей розвиток особливо помітний, від традиційних математичних моделей до сучасних методів на основі штучного інтелекту. З кожним роком з'являються нові інструменти і техніки, які дозволяють більш точно оцінювати майбутні тенденції. Певні нові методи, наприклад методи імовірного прогнозування, є певною надбудовою над іншими більш відомими традиційними методами. Незважаючи на значний прогрес у цій галузі, дослідження та вдосконалення методів прогнозування продовжуються, адже реальний світ постійно змінюється, і важливо адаптувати підходи до нових умов.

Враховуючи все це, постає важливе питання щодо підвищення точності прогнозування параметрів діяльності підприємств. Сучасна сфера електронної торгівлі вимагає все більшої здатності прогнозувати зміни у попиті, поведінці споживачів, фінансових результатах та зовнішніх економічних умовах. Для цього необхідно вдосконалювати існуючі методи прогнозування, шляхом інтеграції відомих традиційних підходів з новими. Комбінація різних методів дозволяє забезпечити більш точне та комплексне

бачення майбутніх змін, що є критично важливим для прийняття стратегічних рішень.

Актуальність роботи обумовлена необхідністю покращення результатів прогнозування при плануванні роботи торговельного підприємства, що є важливим чинником для підвищення ефективності його діяльності в умовах динамічних змін ринку. Зараз це питання є важливішим ніж будь коли, бо через політичний стан, епідеміологічну ситуацію у світ, та інші зовнішні фактори, люди з кожним днем все більше віддають перевагу онлайн замовленням.

Метою дослідження є аналіз існуючих методів прогнозування параметрів діяльності підприємств електронної комерції, а також розробка комбінованого підходу до вирішення задачі прогнозування на основі інтеграції відомих та нових методів для підвищення якості та точності прогнозування показників діяльності підприємств.

Кваліфікаційна робота виконана у відповідності до методичних вказівок з організації виконання та захисту кваліфікаційних робіт [1].

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Ринок електронної комерції в Україні

На сьогоднішній день суспільство все більше розвивається та адаптується до нових інформаційних технологій. Цифровізація змінює умови ведення бізнесу в усьому світі, і Україна не є виключенням. Доступність інтернету та поява нових інформаційних технологій сприяли розвитку електронної торгівлі та значному підвищенню попиту на цю сферу. З кожним днем все більше компаній переходять на онлайн розповсюдження свого товару. В Україні, подібна торговельна діяльність охоплює різноманітні товари та послуги: від продуктів харчування і побутової техніки, до одягу, взуття та електроніки. Завдяки зручності, легкому доступу до асортименту та можливості заощадити час, цей формат користується великою популярністю.

Аналіз динаміки обсягів ринку електронної комерції в Україні показав, що до 2022 року Україна демонструвала значні успіхи на ринку електронної комерції. Найбільше зростання +48% відповідно до попереднього періоду спостерігалось у 2020 році, що пов'язано також із пандемією COVID-19, яка стала акселератором розвитку цифровізації. У 2021 році український ринок електронної комерції значно зріс – на 27 %, порівняно зі світовим темпом зростання на 15 %. Динаміку обсягів ринку електронної комерції в Україні за період 2017-2022 рр. [2] зображено на рис. 1.1.

Виходячи з усіх представлених фактів, перед компаніями, що вирішили займатися електронною торгівлею, виникає багато проблем та задач, що необхідно вирішити. Для цього створюють спеціалізовані ІС, у які закладають різноманітний життєво необхідний для ведення подібної діяльності функціонал. Окрім базових задач, по типу ведення фінансового та складського обліку, проведення транзакцій, оформлення звітності тощо, однією з найважливіших складових є процес прогнозування показників

діяльності підприємства електронної комерції.

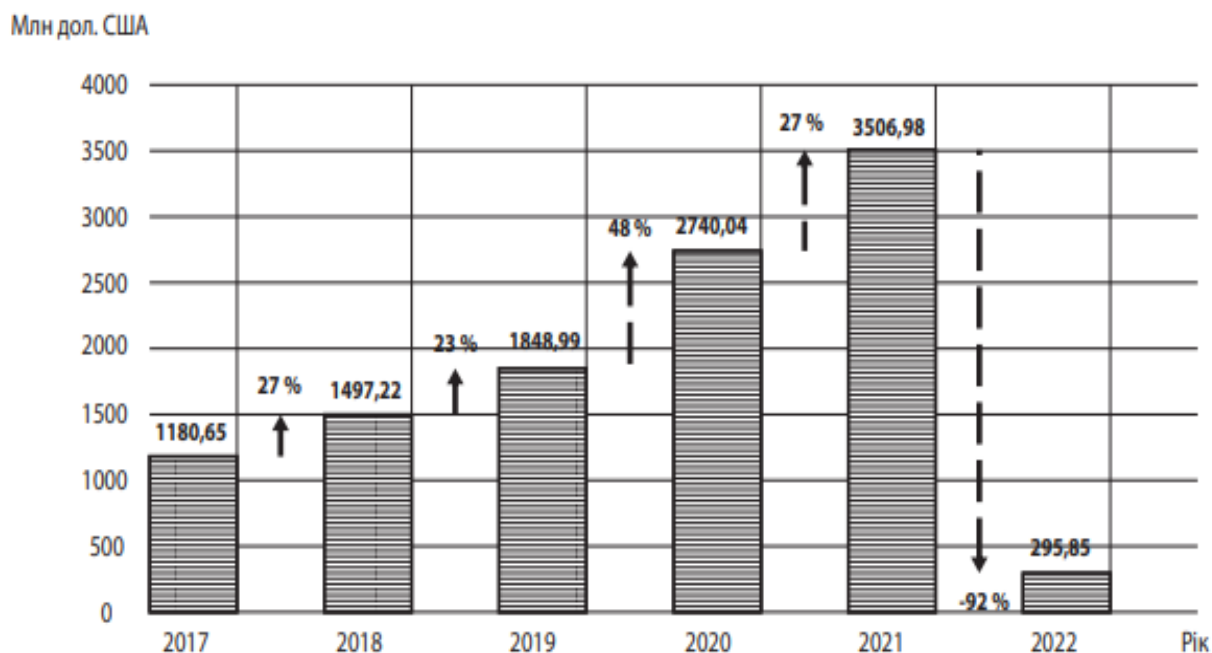


Рисунок 1.1 – Динаміка обсягів ринку електронної комерції в Україні

Задача прогнозування існує вже багато років, і науковці не перестають розробляти нові і вдосконалювати відомі методи прогнозування. Ця сфера постійно розвивається, особливо у останні роки, бо прискорення та покращення технічного обладнання відкриває можливості для більш трудомістких обчислень. Нові методи, наприклад методи імовірнісного прогнозування, є певною надбудовою над іншими більш відомими традиційними методами.

Враховуючи все це, постає питання щодо підвищення точності прогнозування параметрів діяльності підприємств та вдосконалення існуючих методів прогнозування на основі інтеграції відомих та нових методів, а також їх більш ефективного використання. Використання подібного підходу при розробці ІС допоможе збільшити прибуток та зменшити витрати підприємства.

1.2 Аналіз існуючих методів прогнозування

В світі налічується велика кількість різноманітних методів прогнозування. Кожен з них є унікальним, але є певні концепції зі своїми основоположними принципами, які закладені у відповідні методи. Для проведення аналізу буде доцільним поділити всю цю множину на чотири групи: класичне прогнозування, квантильне прогнозування, імовірнісне прогнозування, та нейромережеві методи прогнозування. Щоб виявити концептуальну різницю між ними, кожен групу методів розглянемо більш детально.

1.2.1 Класичне прогнозування

Класичне прогнозування об'єднує у собі методи, основним напрямком яких є отримання одного головного прогнозу, що представляє єдине значення цільового показника. Ці методи засновані на використанні лінійних або простих статистичних залежностей, і у припущенні про стаціонарність даних.

Подібне прогнозування прагне зрозуміти структуру даних, враховуючи тренди, сезонність і цикли. Ці методи створені для побудови базових моделей, аналізу часових рядів тощо. Вони не враховують імовірність розкиду даних і спрямовані на отримання єдиного, найімовірнішого точкового прогнозу. Вони підходять для роботи зі стабільними даними, де ступінь передбачуваності достатньо великий [3].

Серед представників класичного прогнозування можна виділити наступні:

- ковзне середнє;

- експоненціальне згладжування;
- регресійний аналіз;
- метод Хольта-Вінтерса;
- ARIMA;
- SARIMA.

1.2.2 Квантильне прогнозування

Квантильне прогнозування [4] є розширенням обчислень, які здійснюються за допомогою класичних методів, що передбачає додавання страхового запасу до прогнозованої оцінки. Подібні методи дають змогу оцінювати різні рівні, процентилю, розподілу даних, такі як медіана (50%), нижній (10%) і верхній (90%) квантилі, надаючи певний діапазон можливих значень замість єдиного результату.

Цей підхід може бути дуже корисним для виконання завдань, де важливо розуміти розподіл даних і передбачити не тільки середній результат, а й імовірні екстремальні значення. Це особливо актуально для управління ризиками, оцінки мінімальних і максимальних значень цільового показника [5].

Головне припущення такого прогнозування полягає в тому, що значення попиту на товари вважається нормально розподіленим. На жаль, у реальному світі, це твердження далеко не завжди є вірним. Це призводить до того, що для підтримування високого рівня сервісу підприємства утворюється надлишок товару, а при низькому рівні сервісу – його нестача.

Серед представників квантильного прогнозування можна виділити наступні:

- квантильна регресія;
- Quantile-ARIMA.

1.2.3 Імовірнісне прогнозування

Для пояснення імовірнісного підходу до прогнозування зручно провести аналогію з традиційними підходами до прогнозування попиту, такими як метод Хольта-Вінтерса, проста ковзна середня чи експоненціальне згладжування. Однак, на відміну від них, результат роботи алгоритму буде представлений не одним прогнозним значенням, а розподілом імовірностей, наприклад: 2 тисячі одиниць товару можуть бути продані з імовірністю 5%, 5 тисяч одиниць товару – з імовірністю 10%, а 10 тисяч одиниць товару – з імовірністю 85%. Таким чином, метод дозволяє отримати повну картину розподілу попиту.

Імовірнісне прогнозування [6] можна назвати новим кроком у своїй сфері. Люди усвідомили, що прогнозовані значення можна представляти не одним числом, а певним набором значень імовірностей. Така інновація дала можливість обчислювати імовірність утворення певного об'єму попиту, і вже виходячи з отриманого набору значень імовірностей приймати економічно обумовлені рішення.

Подібні методи будують прогноз у формі розподілу імовірностей, що дає повнішу картину майбутніх сценаріїв. Імовірнісне прогнозування використовується у поєднанні з іншими методами для проведення складних розрахунків, у яких присутній високий ступінь невизначеності, але їхнє застосування вимагає значних ресурсів. Вони найефективніші у великих і складних системах, де невизначеність відіграє ключову роль, але можуть бути надлишковими для простих або стабільних завдань. Часткове імовірнісне прогнозування та повне імовірнісне прогнозування [7] представляють два головні методи, з яких складаються ця група.

1.2.4 Нейромережеві методи прогнозування

Нейромережеві методи прогнозування [8] представляють собою методи, що базуються на використанні алгоритмів машинного навчання, які мають здатність знаходити складні залежності в даних, зокрема нелінійні, з використанням багатовимірних ознак і великих обсягів інформації. Ці методи адаптуються до особливостей даних та навчаючись на історичних прикладах.

Машинне навчання дає змогу будувати прогнози, використовуючи гнучкі моделі, які можуть містити як структурні залежності, так і зовнішні чинники, наприклад: економічні показники, погодні умови або поведінку споживачів, тощо. Ці методи застосовуються для роботи з великими наборами даних і складними часовими рядами, а також для завдань із численними змінними.

Подібні методи прогнозування мають високу гнучкість і потужність, але також існує низка недоліків, які можуть обмежувати їх застосування. Серед головних факторів, можна виділити наступні: висока складність розробки та впровадження, залежність від якості даних, потреба у неймовірно великій кількості даних для навчання моделей, необхідність регулярного оновлення моделей при змінах навколишнього середовища тощо [9].

1.2.5 Порівняння методів прогнозування

Для ефективного прогнозування показників підприємства важливо не лише розуміти основи кожної групи методів, але й усвідомлювати унікальність кожної з них. Вони мають свої специфічні особливості, результат прогнозування, переваги та обмеження, які роблять їх більш або

менш придатною для вирішення конкретних завдань. Наприклад, класичні методи забезпечують простоту і швидкість розрахунків, але не враховують невизначеність. Квантильні методи дозволяють зробити прогноз більш повним, але тримання стратегічного запасу може негативно впливати на підприємство. Методи імовірнісного прогнозування надають детальну картину розподілу майбутніх результатів, проте їх функціонал може бути надлишковим у простих системах. Методи машинного навчання, хоча й здатні працювати з великими даними та складними залежностями, є дорогими в реалізації та потребують дуже великого обсягу даних.

Саме через ці відмінності порівняння всіх груп методів стає необхідним для вибору найкращого підходу до конкретної задачі. Адже лише розуміння того, які інструменти найкраще підходять для врахування невизначеності, управління ризиками або швидкого реагування на зміни, дозволить розробити точні, адаптивні та практичні прогнози.

Підсумки проведеного аналізу, представлені у табл. 1.1.

Таблиця 1.1 – Аналіз методів прогнозування

Група методів	Результат прогнозування	Головна перевага	Недоліки
класичне прогнозування	єдине значення	простота реалізації	не враховує імовірність розкиду даних, ступінь передбачуваності повинна бути достатньо великою
квантильне прогнозування	квантиль	прогнозування діапазону з екстремальними значеннями	може утворитись надлишок товару, або його нестача
імовірнісне прогнозування	імовірнісний розподіл	дозволяє отримати повну картину розподілу попиту	може бути надлишковим для простих або стабільних завдань
нейромережеві методи	нелінійні залежності	висока гнучкість і потужність	висока складність розробки, залежність від якості даних, потреба у великій кількості даних для навчання, необхідність регулярного оновлення

1.3 Аналіз проблем використання методів прогнозування при розробці ІС для планування роботи підприємства електронної торгівлі

Розробка ІС для планування роботи підприємств електронної торгівлі вимагає використання ефективних методів прогнозування. Існує велика кількість проблем, які можуть суттєво вплинути на якість прогнозів і роботу системи в цілому.

Основна складність полягає у специфіці даних, що використовуються в електронній торгівлі. Дані про продажі, запаси, поведінку користувачів, сезонність, маркетингові активності тощо є не лише великими за обсягом, але й різноманітними. Ці фактори ускладнюють їхню обробку, а також підвищують ризик виникнення шуму або аномалій у даних.

Другою важливою проблемою є постійні зміни ринку електронної торгівлі. Поведінка споживачів, популярність продуктів і зовнішні фактори, такі як економічні коливання або зміни в законодавстві, можуть швидко змінюватися. Це створює складнощі для моделей прогнозування, особливо класичних і квантильних методів, які часто покладаються на історичні дані та не враховують швидких змін. У таких умовах навіть методи засновані на машинному навчанні, можуть зазнавати труднощів через необхідність регулярного оновлення моделей і перенавчання.

Також виникає проблема інтеграції прогнозів в управлінські рішення. Наприклад, класичні методи часто дають точкові прогнози, які легко інтерпретувати, але вони не враховують ризиків і невизначеності. У свою чергу, імовірнісні методи та методи машинного навчання генерують більш складні результати, які можуть бути незрозумілими для користувачів без глибоких аналітичних навичок. Це може ускладнити прийняття рішень на основі таких прогнозів.

Крім того, висока вартість і складність розробки моделей прогнозування часто стають нездоланими перешкодами для підприємств.

Методи машинного навчання, хоч і мають високу точність, потребують значних ресурсів для розробки, навчання моделей і підтримки їхньої актуальності. Для малих і середніх підприємств ці витрати можуть бути надмірними.

Варто відмітити, що недостатній рівень автоматизації процесів у деяких системах може призвести до затримок у використанні прогнозів. Тобто, якщо прогнози не інтегровані з іншими модулями, такими як управління запасами чи продажами, це знижує ефективність планування роботи підприємства.

Таким чином, успішне застосування методів прогнозування в ІС електронної торгівлі залежить від вирішення проблем якості даних, адаптивності моделей, складності інтеграції та високих витрат. Ефективне прогнозування потребує поєднання декількох методів, створення гнучкої архітектури ІС і регулярного оновлення моделей відповідно до змін ринкових умов.

1.4 Постановка задачі дослідження

Підприємство електронної торгівлі, прагне покрити найоптимальніший відсоток сервісу, не тримаючи при цьому страховий запас товарів, щоб зменшити кількість недоцільних витрат на закупівлю, зберігання тощо. Для задоволення цих умов і обмежень ідеально підходять імовірнісні методи прогнозування, що відходять від застарілої концепції стратегічного запасу. Постає потреба в дослідженні можливостей та визначені оптимального поєднання відомих та нових імовірнісних методів прогнозування.

Об'єктом дослідження є процес прогнозування соціально-економічних показників діяльності підприємств.

Предметом дослідження є методи прогнозування показників діяльності

підприємств електронної комерції.

Метою дослідження є аналіз існуючих методів прогнозування параметрів діяльності підприємств електронної комерції, а також розробка комбінованого підходу до вирішення задачі прогнозування на основі інтеграції відомих та нових методів для підвищення якості та точності прогнозування показників діяльності підприємств. Це, в свою чергу, може суттєво підвищити ефективність роботи компаній електронної торгівлі.

Для досягнення поставленої мети необхідно вирішити такі основні завдання:

- провести аналіз методів прогнозування та порівняти їх за визначеними показниками;
- знайти найкращу комбінацію відомих та нових методів, виходячи з існуючих умов та обмежень;
- розробити комбінований метод прогнозування показників діяльності підприємств електронної торгівлі;
- перевірити ефективність розробленого комбінованого методу на практиці.

2 ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ВИКОРИСТАННЯ МЕТОДІВ ПРОГНОЗУВАННЯ

Кількість різноманітних методів прогнозування у світі достатньо велика, розглядати і аналізувати кожний метод було б недоцільно. Тому, для вибору оптимальної комбінації відомих та нових методів прогнозування необхідно спочатку визначити, які з зазначених вище груп підпадають під умови і обмеження підприємства електронної торгівлі.

Нейромережеві методи прогнозування є безсумнівними лідерами з точки зору ефективності, але не варто забувати, як багато вони потребують для досягнення такого рівня: висока складність розробки, залежність від якості даних, потреба у великій кількості даних для навчання, необхідність регулярного оновлення. Середньостатистичне підприємство не має достатньої кількості грошей, для розробки, навчання, та впровадження методів цієї групи. Також, варто зазначити, що навіть при наявності коштів, все ще залишається потреба у кількості та якості даних, а більшість систем, що не були заздалегідь розраховані на додавання нейронних мереж, не мають навіть механізмів для збору необхідної інформації. Враховуючи всі вказані недоліки, ця група не підходить для вирішення поставленої задачі.

Квантильне прогнозування є ефективним методом для оцінки попиту, оскільки воно враховує імовірнісні аспекти розподілу даних. Використовуючи цей підхід, можна отримати більш точні прогнози щодо різних сценаріїв. Це допомагає підприємствам планувати закупівлі товарів та управляти складськими запасами більш усвідомлено. Однак, утримання страхового запасу, що базується на прогнозах квантильного методу, неминуче створює певні виклики для підприємства. З одного боку, надмірний обсяг страхового запасу призводить до збільшення витрат на зберігання та обслуговування товару, особливо якщо цей товар має обмежений термін або потребує спеціальних умов зберігання. З іншого боку, недостатній страховий

запас збільшує ризик дефіциту товару, що може призвести до втрати потенційного прибутку через незадоволений попит клієнтів і зниження рівня їхньої лояльності. Ці обставини створюють ситуацію, коли підприємство змушене балансувати між витратами на утримання надлишків товару та ризиками втрати продажів через нестачу. Як результат, існує великий ризик виникнення помилок через врахування сезонності, змін ринкових умов або непередбачуваного попиту. У нашому випадку, підприємство прагне позбутися подібних перешкод, тому квантильне прогнозування не підходить для вирішення задачі.

Імовірнісне прогнозування відходить від своїх попередників, роблячи акцент на розрахунках імовірностей, що можна охарактеризувати як відсотки покриття сервісу. Подібний ефект досягається за рахунок комплексного результату обчислень, який дозволяє отримати повну картину розподілу попиту [10]. Таким чином, можна знайти найоптимальніше значення, при якому відношення витрат та прибутків буде мати найбільшу вигоду для компанії. Методи цієї групи дають достатньо комплексний результат, у певних випадках, коли система проста, а дані легко передбачаються, додаткові обчислення створювали б лише зайве навантаження. Однак, у сфері електронної торгівлі, цей варіант є ідеальними для використання.

Класичне прогнозування дає простий результат, який легко інтерпретувати та використовувати. До цієї групи входить велика кількість методів та їх модифікацій, від самих базових і простих, до більш комплексних, з урахуванням трендів, сезонності, циклів, тощо. Багато з них використовуються і сьогодні, показуючи непогану точність розрахунків. Але, не рідко можна стикнутися з нестабільними даними, з низьким ступенем передбачуваності. У цьому випадку, точність прогнозу значно знижується. Однак, за рахунок своєї простоти, використання цих методів у поєднанні один з одним, або методами інших груп, дозволяє мінімізувати недоліки, при цьому не підвищуючи потреби до розробки занадто високо.

Отже, серед усіх груп ми можемо виділити дві, класичне та імовірнісне прогнозування. Кожна з них має власні особливості, але вони підпадають під умови і обмеження нашої задачі. А також, обидві групи мають гарні перспективи для поєднання між собою.

Наступним кроком необхідно розглянути конкретні методи прогнозування, скласти перелік показників, за якими їх можна порівняти, та знайти найкращий варіант для поставленої задачі. Для створення поєднання потрібно обрати по одному методу з кожної групи, тому доцільним буде розглядати обидві групи окремо.

2.1 Оцінка можливостей використання імовірнісних методів прогнозування

Головними представниками імовірнісного прогнозування можна назвати методи: часткового імовірнісного прогнозування, та повного імовірнісного прогнозування.

Часткове імовірнісне прогнозування базується на оцінці окремих точок у розподілі імовірностей. Цей метод дозволяє отримати ключові характеристики розподілу попиту або продажів, зосереджуючись на окремих важливих значеннях. Це робить його зручним для швидкої оцінки діапазону можливих значень, де потрібно врахувати ризики чи варіативність. Метод є відносно простим у реалізації, що робить його привабливим для малих і середніх підприємств, особливо в умовах обмежених ресурсів. Проте, його недоліком є обмежена інформація про проміжні значення між оціненими точками, що може знижувати точність прогнозів у більш складних ринкових умовах.

Повне імовірнісне прогнозування, на відміну від часткового, враховує весь спектр розподілу імовірностей, що дозволяє створювати прогнози з

урахуванням усієї варіативності можливих результатів. Це забезпечує глибший аналіз даних, зокрема імовірностей екстремальних сценаріїв, і робить метод незамінним для довгострокового планування. Однак, реалізація повного прогнозування є значно складнішою. Вона вимагає великих обсягів даних, потужних обчислювальних ресурсів і залучення висококваліфікованих фахівців. Для малих компаній це може бути недоступним, але для великих підприємств, які прагнуть максимізувати точність і врахувати всі можливі ризики, такий метод є оптимальним.

Для вибору конкретного методу, потрібно визначити чіткі критерії оцінювання. Вибір правильних критеріїв є значним етапом для подальшого отримання результатів, тому що вони дозволяють провести об'єктивне порівняння враховуючи специфічні умови задачі. У першу чергу, слід зазначити критерії, які впливають на успішність методу прогнозування, серед таких аспектів може бути точність прогнозу, гнучкість у застосуванні та інше. Не менш важливо, враховувати можливості та вимоги підприємства, доступні ресурси, складність реалізації, зручність інтерпретації результатів. Тому потрібно розглядувати кожен критерій, як окремий аспект, який сприяє вибору метода прогнозування. Перелік головних критеріїв оцінювання імовірнісних методів наведено у табл. 2.1

Таблиця 2.1 – Критерії оцінювання імовірнісних методів

Критерій	Опис	Значущість
Точність прогнозу	Точність визначає, наскільки добре прогноз відображає реальні результати. В електронній комерції це важливо для уникнення витрат.	Для підприємств, що працюють у конкурентному середовищі, точні прогнози мінімізують витрати на зберігання та підвищують рівень задоволеності клієнтів.
Складність реалізації	Включає технічну складність, потребу в спеціалізованих інструментах і кваліфікованих кадрах.	Вибір методу, який відповідає ресурсним можливостям компанії, особливо малих і середніх підприємств, впливає на ефективність його впровадження.

Продовження таблиці 2.1.

Критерій	Опис	Значення
Гнучкість у застосуванні	Відображає здатність методу адаптуватися до змін ринкових умов, сезонності чи інших зовнішніх факторів.	Гнучкі методи забезпечують стійкість бізнесу в умовах нестабільного попиту.
Ресурсні вимоги	Оцінює обсяг даних, необхідний для роботи методу, а також обчислювальні та фінансові витрати.	Методи з високими ресурсними вимогами можуть бути недоступними для підприємств з обмеженим бюджетом.
Зручність інтерпретації результатів	Ступінь зрозумілості отриманих прогнозів для менеджменту та співробітників.	Простота розуміння прогнозів допомагає швидко приймати обґрунтовані рішення.

Тепер, коли були встановлені чіткі критерії, ми можемо провести порівняння методів імовірнісного прогнозування. Результат порівняння наведено наведено у табл. 2.2.

Таблиця 2.2 – Порівняння імовірнісних методів

Критерій	Часткове імовірнісне прогнозування	Повне імовірнісне прогнозування
Точність прогнозу	Обмежена, враховує лише кілька точок розподілу	Висока, охоплює весь спектр можливих значень
Складність реалізації	Низька, легко реалізується з базовими інструментами	Висока, потребує складних алгоритмів і значних знань
Гнучкість у застосуванні	Помірна, добре працює зі стабільними ринками	Висока, адаптується до змінних умов і зовнішніх факторів
Ресурсні вимоги	Низькі, підходить для малих і середніх підприємств	Високі, вимагає великих обсягів даних і обчислювальних ресурсів
Зручність інтерпретації	Висока, результати легко зрозумілі	Складна, потребує інструментів для аналізу та візуалізації

За результатами порівняння, що наведені вище, можна зробити висновок стосовно цілеспрямованості використання представлених методів. Повне імовірнісне прогнозування здатне видавати більш чіткі результати, гарно адаптується до змін і зовнішніх факторів, але потребує зайвих ресурсів, що просто недоступні більшості підприємств. Часткове імовірнісне прогнозування спрощує загальну картину, що впливає на результативність,

але реалізація цього методу буде більш економічно-виправданим рішенням. Отже, саме цей метод краще використовувати для вирішення поставленої задачі [11].

2.2 Оцінка можливостей використання класичних методів прогнозування

Просте ковзне середнє є базовим методом аналізу, що використовується для згладжування коливань у часових рядах. Цей метод полягає в обчисленні середнього значення даних за фіксований період, що пересувається вперед із кожним кроком. Такий підхід дозволяє зменшити вплив короткострокових коливань, але не враховує змінні тренди чи сезонність.

Середньо-зважене середнє, як його запропонував Шрайбфедер, є модифікацією ковзного середнього, де кожному значенню у вибірці надається своя вага залежно від його значущості. Найбільш актуальні дані отримують вищу вагу, тоді як старі значення менш впливові. Це дозволяє враховувати зміни у даних із більшим акцентом на поточні тренди, хоча метод все ще обмежений у роботі зі складними часовими рядами.

Експоненціальне згладжування є більш складним методом, який використовує експоненційно зменшувані ваги для всіх попередніх значень даних. Цей підхід забезпечує згладжування даних із врахуванням як короткострокових, так і довгострокових компонентів. Він є гнучким, але також обмежений у випадках із явною сезонністю або значними структурними змінами.

Метод Хольта-Вінтерса [12] розширює експоненціальне згладжування, включаючи можливість моделювання сезонності та тренду. У цьому методі виділяються три компоненти: рівень, тренд і сезонність, які оновлюються на

кожному кроці прогнозу. Цей підхід є ефективним для прогнозування даних, які мають регулярну сезонну структуру, але потребує налаштування параметрів для досягнення оптимальної точності.

ARIMA [13] – це одна з найпопулярніших моделей для аналізу та прогнозування часових рядів. Вона враховує залежності між спостереженнями в ряді, дозволяє працювати зі стаціонарними даними та інтегрує попередні значення для обліку довгострокових залежностей.

SARIMA [14] є певним розширенням свого попередника, яке додає моделювання сезонності, дозволяючи працювати з даними, що мають регулярні циклічні коливання.

BSTS [15] є сучасним методом, який поєднує підхід Баєса з структурними моделями часових рядів. Це дозволяє моделювати складні залежності, включаючи сезонність, тренди та вплив зовнішніх чинників. Цей підхід забезпечує прозорість моделювання, оскільки кожен компонент окремо визначається у рамках моделі, а також дозволяє враховувати невизначеність прогнозів. BSTS є потужним інструментом для прогнозування та аналізу часових рядів у складних бізнес-сценаріях.

Тепер, для вибору серед такої великої кількості методів, потрібно визначити чіткі критерії оцінювання. Як і раніше, слід зазначити критерії, які впливають на успішність методу прогнозування, серед таких аспектів може бути точність прогнозу, гнучкість у застосуванні та інше. Але, на відміну від критеріїв обраних для оцінки імовірнісних методів, ми не будемо розглядати можливості та вимоги підприємства. У цьому випадку, варто зосередити свою увагу на можливостях методів враховувати ті чи інші аспекти наявних даних. Серед таких важливих речей, можна виділити сезонність, тренди, та гнучкість. Перелік головних критеріїв оцінювання класичних методів наведено у табл. 2.3

Таблиця 2.3 – Критерії оцінювання класичних методів

Назва	Опис	Значущість
Точність прогнозу	Точність визначає, наскільки добре прогноз відображає реальні результати. В електронній комерції це важливо для уникнення втрат через дефіцит або надлишок товару.	Для підприємств, що працюють у конкурентному середовищі, точні прогнози мінімізують витрати на зберігання та підвищують рівень задоволеності клієнтів.
Складність реалізації	Включає технічну складність, потребу в спеціалізованих інструментах і кваліфікованих кадрах.	Вибір методу, який відповідає ресурсним можливостям компанії, особливо малих і середніх підприємств, впливає на ефективність його впровадження.
Гнучкість до змін даних	Відображає, наскільки метод здатен адаптуватися до змін у структурі даних або зовнішніх чинників.	Гнучкі методи забезпечують стійкість прогнозування в умовах нестабільного попиту.
Урахування трендів	Відображає, чи здатний метод враховувати загальні тенденції в даних, наприклад, зростання чи спад.	В умовах постійного розвитку підприємства, що займається безперервною торгівлею, так чи інакше виникнуть певні тренди. Їх врахування, дає можливість значно підвищити точність прогнозування.
Урахування сезонності	Показує, чи враховує метод сезонні коливання в даних, наприклад, щомісячні чи річні цикли.	Багато різноманітних товарів мають більший попит у конкретні моменти часу. Це може бути пов'язано зі святами, зміною пори року, або іншими зовнішніми факторами.

Встановивши чіткі критерії, стало можливим провести порівняння розглянутих класичних методів. Результат наведено у табл. 2.4.

Таблиця 2.4 – Порівняння класичних методів

Метод	Складність реалізації	Урахування трендів	Урахування сезонності	Гнучкість до змін даних	Точність прогнозу
Просте ковзне середнє	Низька	Ні	Ні	Низька	Низька
Середньо-взважене середнє	Низька	Частково	Ні	Середня	Низька-Середня
Експоненціальне згладжування	Середня	Так	Ні	Середня	Середня

Продовження таблиці 2.4.

Метод	Складність реалізації	Урахування трендів	Урахування сезонності	Гнучкість до змін даних	Точність прогнозу
Метод Хольта-Вінтерса	Висока	Так	Так	Середня	Висока
ARIMA	Висока	Так	Ні	Висока	Висока
SARIMA	Висока	Так	Так	Висока	Дуже висока
BSTS	Дуже висока	Так	Так	Дуже висока	Дуже висока

Підприємство електронної комерції повинно зосередитися на врахуванні трендів і сезонності, оскільки ці фактори є критично важливими для прогнозування попиту та планування запасів у цій галузі. Виходячи з результатів порівняння, вибір методів, які відповідають цим вимогам, звужується до SARIMA та BSTS.

SARIMA має низку переваг для електронної комерції. Цей метод хоч і вимагає складної роботи з параметрами, але має якісну документацію та є доступним у багатьох статистичних бібліотеках. Чітке уявлення про залежності в даних дозволяє менеджерам краще розуміти прогнози. BSTS, у свою чергу, є більш сучасним методом, заснованим на Баєсівському підході. Цей метод характеризується великою гнучкістю, що дозволяє дуже швидко адаптуватися до змін зовнішніх чинників. Враховуючи усі факти, для підприємства електронної комерції вибір SARIMA є більш доцільним.

3 РОЗРОБКА КОМБІНОВАНОГО МЕТОДУ ПРОГНОЗУВАННЯ

Тепер, коли методи для вирішення нашої задачі обрані, необхідно зрозуміти як вони працюють, які є обмеження на їх використання, яких обчислювальних ресурсів вони потребують, і як їх можна скомбінувати.

Перш за все, розберемося з вхідними даними. Для будь-якого прогнозування нам необхідно мати часовий ряд, щоб проаналізувати його і зробити якесь припущення про майбутнє. Виходячи з назви, можна зрозуміти, що це розподілення якоїсь величини в залежності від часу. Але, з точки зору математичної статистики, визначення часового ряду є більш комплексним. Часовий ряд – це послідовність випадкових величин, які можуть бути якось пов'язані між собою. Саме пов'язаність цих величин слугує передумовою для можливості прогнозування поведінки часового ряду, випадковість же варто розглядати лише як припущення.

Ряди можна поділити на стаціонарні та нестаціонарні. Якщо характер наших величини не змінюється з часом, то така поведінка відноситься до першої групи, а якщо навпаки – до другої. Стаціонарність часового ряду пов'язана з вимогою того, що він має стале середнє значення, і його рівні коливаються навколо цього середнього зі сталою дисперсією.

Ряд можна вважати строго стаціонарним, якщо для будь яких s і k розподілення випадкових величин $y_{t1}, y_{t2}, \dots, y_{tk}$ не відрізняється від $y_{t1+s}, y_{t2+s}, \dots, y_{tk+s}$. Фактично, це каже про те, що властивості нашого ряду не змінюються при зміні початкової точки у часі. На практиці, подібне визначення використовувати складно. Тому, як правило, використовується поняття слабкої стаціонарності. Ряд є слабо стаціонарним, якщо характеристики цього ряду не залежать від часу. Математичне очікування ряду – константа, дисперсія ряду – константа, і усі коваріації, не залежно від моменту часу, теж є константами. Ці твердження можна записати у наступному вигляді:

$$\begin{aligned} E[y_t] &= \text{cons} \\ \text{Var}(y_t) &= \sigma^2, \\ \text{Cov}(y_t, y_{t+s}) &= \gamma_s \end{aligned} \quad (3.1)$$

де y_t – значення ряду;

$E[y_t]$ – математичне очікування ряду;

$\text{Var}(y_t)$ – дисперсія ряду;

$\text{Cov}(y_t, y_{t+s})$ – коваріація між значеннями ряду на відстані s .

Перетворення нестационарного ряду на стаціонарний – це важливий крок у підготовці даних для багатьох методів аналізу часових рядів. Основні підходи включають видалення тренду, сезонності та згладжування змін. Одним з головних інструментів, який використовують у сучасному світі, є Диференціювання. Простими словами, це обчислення різниці між сусідніми значеннями часового ряду, щоб усунути тренд. Формула першого порядку:

$$y'_t = y_t - y_{t-1}, \quad (3.2)$$

де y'_t – диференційований ряд.

Формула другого порядку, може знадобитися, якщо після першого диференціювання ряд залишається нестационарним:

$$y''_t = y'_t - y'_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}). \quad (3.3)$$

Якщо ряд має чіткі сезонні коливання, їх можна усунути за допомогою сезонного диференціювання:

$$y'_t = y_t - y_{t-s}, \quad (3.4)$$

де s – довжина сезонного циклу.

Логарифмування даних допомагає стабілізувати дисперсію та

зменшити вплив великих значень. Якщо дисперсія ряду збільшується з часом, логарифмічне перетворення може зробити її стабільною:

$$y'_t = \log(y_t). \quad (3.5)$$

Звісно, не варто використовувати представлені перетворення для будь-яких часових рядів. Якщо початковий ряд не мав певної ознаки, або взагалі вже був стаціонарним, диференціювання та логарифмування зіпсують якість наших даних, що може значно вплинути на результат прогнозування. Існує популярний статистичний тест для оцінки стаціонарності ряду.

Тест Дікі-Фуллера [16] використовується для перевірки наявності одиничного кореня у часовому ряді, що вказує на його нестационарність. Тест перевіряє нульову гіпотезу (H_0) про те, що часовий ряд має одиничний корінь, тобто він нестационарний. Альтернативна гіпотеза (H_1) передбачає, що ряд стаціонарний. Існує три різні версії цього тесту: з константою, без константи, з константою та трендом. Перший варіант є найбільш розповсюдженим, бо інші два мають сенс лише у часних випадках.

Тест Дікі-Фуллера має наступний вигляд:

$$\Delta y_t = c + \beta y_{t-1} + d_1 \Delta y_{t-1} + \dots + d_p \Delta y_{t-p} + u_t, \quad (3.6)$$

де Δy_t – зміна часового ряду;

c – стала;

β – коефіцієнт, який тестується на значущість;

p – кількість лагів;

u_t – залишковий шум.

Гіпотеза H_0 полягає у тому, що $\beta=0$, а H_1 – що $\beta<0$. Для більш детального розуміння різниці рядів, що відносяться до цих двох гіпотез, наведено графіки на рис. 3.1.

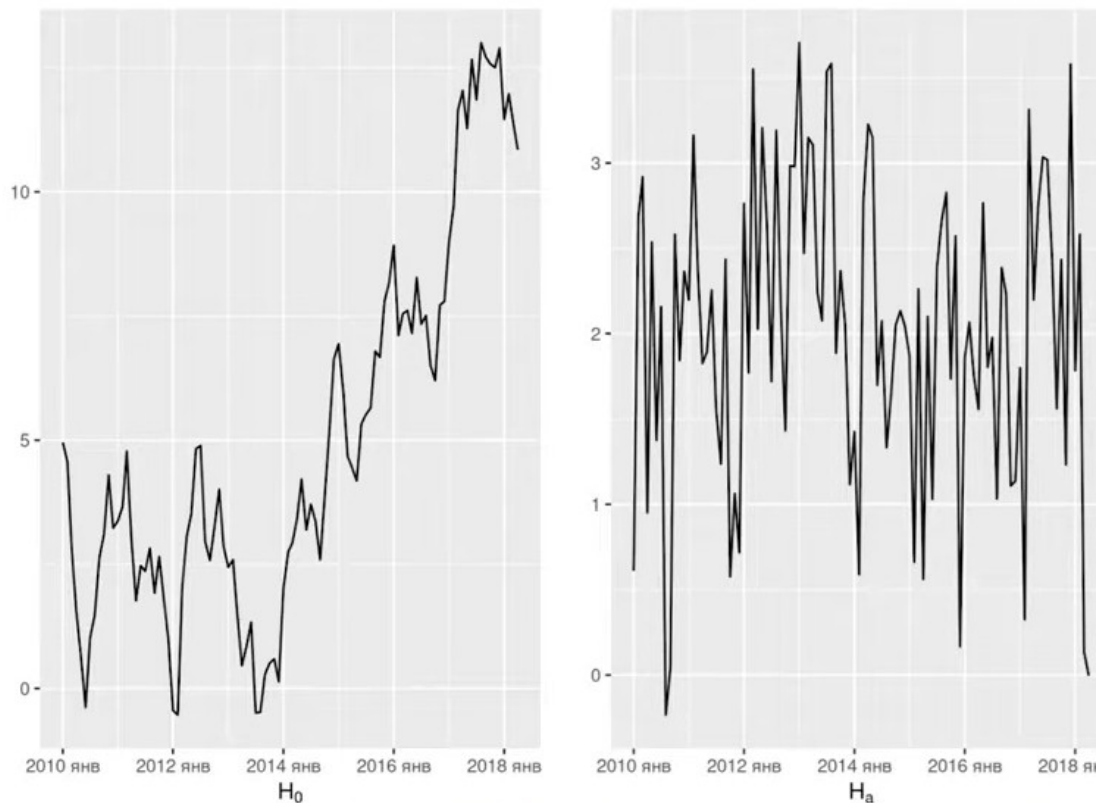


Рисунок 3.1 – Гіпотези тесту Дікі-Фуллера

Слід врахувати, що збільшення числа лагів призводить до зниження потужності тесту. Зазвичай обмежуються трьома-чотирма лагами. Підставляємо дані у модель і обчислюємо значення тестової статистики β .

Порівнюємо обчислене значення зі значенням критичних точок для заданого рівня значущості. Якщо β менше критичного значення, ми відхиляємо H_0 і приймаємо H_1 , що вказує на стаціонарність ряду.

Існують також певні варіації тесту Дікі-Фуллера. В першу чергу, потреба створення нових тестів виникає через те, що потужність розглянутого достатньо мала. Одними з найпопулярніших тестів вважаються OCSB-тест та KPSS-тест. Перший використовується для перевірки наявності сезонності та одночасної стаціонарності часового ряду. Його головна мета – визначити, чи є в даних сезонні патерни та чи відповідає ряд умовам стаціонарності. Другий використовується для перевірки гіпотези стаціонарності часового ряду, зокрема наявності одиничного кореня.

3.1 Метод SARIMA

Підготувавши часовий ряд, можна переходити до побудови конкретних моделей прогнозування. SARIMA є дуже комплексним методом, який складається з багатьох складових. Виходячи з назви, ми можемо виділити такі ключові частини:

– S: сезонність свідчить про те, що під час вирішення задачі ми розглядаємо два різних патерна, в рамках цілого часового ряду та в рамках заданої сезонності;

– AR: авторегресійна модель має вигляд $AR(p)$. У цьому випадку, ми припускаємо, що прогнозовані величини можуть бути залежними від попередніх величин. Порядок моделі p характеризує кількість попередніх значень, що впливають на прогнозоване;

– I: інтегрованість свідчить про те, що наш часовий ряд повинен бути стаціонарним. Досягається цей результат шляхом проведення диференціювання, як було показано вище;

– MA: модель ковзного середнього має вигляд $MA(q)$. Тут ми будемо припущення, що прогнозовані величини можуть бути залежними від випадкових помилок. Порядок моделі q характеризує кількість попередніх значень, що впливають на прогнозоване.

У сукупності, наша модель має наступний вигляд:

$$SARIMA(p, d, q) - (P, D, Q) [12], \quad (3.7)$$

де p – порядок AR частини;

d – кількість разів проведення диференціювання;

q – порядок MA частини;

P – порядок SAR частини;

D – кількість разів проведення сезонного диференціювання;

Q – порядок SMA частини;

[12] – період сезонності.

Загальна модель SARIMA може бути записана наступним чином:

$$\Phi_p(B^s)\phi_p(B)\Delta^d\Delta_s^D Y_t = \Theta_q(B^s)\theta_q(B)\mu_t, \quad (3.8)$$

де B – оператор зсуву;

$\phi_p(B)$ – поліном авторегресії;

$\Phi_p(B^s)$ – сезонний поліном авторегресії;

Δ^d – звичайне диференціювання;

Δ_s^D – сезонне диференціювання;

$\theta_q(B)$ – поліном ковзного середнього;

$\Theta_q(B^s)$ – сезонний поліном ковзного середнього;

μ_t – випадкова помилка, яка має нормально розподілення з нульовими математичним очікуванням і дисперсією.

Щоб використати цю модель на практиці, треба підібрати достатньо велику кількість представлених параметрів та провести розрахунки. Зробити це можна за певним алгоритмом.

В першу чергу, перевіряємо гіпотезу о стаціонарності нашого часового ряду. Якщо ряд не стаціонарний, проводимо диференціювання, поки не досягнемо результату. Проводячи цю операцію, ми можемо знайти параметри d та D , що дорівнюють кількостям проведених диференціювань.

Як тільки ми отримали стаціонарний ряд, необхідно перебрати велику кількість значень інших параметрів. Для автоматизації цього процесу, необхідно провести велику кількість обчислень за нашим методом, і шляхом мінімізації інформаційного критерію, обрати параметри з найточнішим результатом. Мінімізація інформаційного критерію – це позначення процесу оцінки шляхом порівняння результату за певним методом.

Далі, коли усі параметри обрані, необхідно перевірити нашу модель на адекватність. Для цього, треба подивитись, чи відповідає вона усім

передумовам, що були в неї закладені, а саме: чи стабільна дисперсія, чи стабільне математичне очікування, чи правильно поведуть себе коваріації, чи дійсно ряд стаціонарний. Для цього можна провести графічний аналіз.

У разі, якщо модель не є адекватною, необхідно провести трансформацію початкового ряду. Одна з таких трансформацій має назву «трансформація Бокса-Кокса». Ідея цього методу достатньо проста, і може бути виражена наступним чином:

$$y_{new} = \frac{y^p - 1}{p}, \text{ якщо } p \neq 0, \quad (3.9)$$

$$y_{new} = \log(y), \text{ якщо } p = 0$$

Треба використовувати такий параметр p , при якому якість моделі буде максимальною. Для знаходження цього значення, використовується метод максимальної правдоподібності.

Якщо в наших даних є від'ємні значення, що може зламати логарифм, необхідно додати певний параметр A . Це може бути будь яке значення, наприклад, одиниця. Тоді наша трансформація буде мати наступний вигляд:

$$y_{new} = \frac{(y+A)^p - 1}{p}, \text{ якщо } p \neq 0, \quad (3.10)$$

$$y_{new} = \log(y+A), \text{ якщо } p = 0$$

Використання цього методу графічно представлено на рис. 3.2.

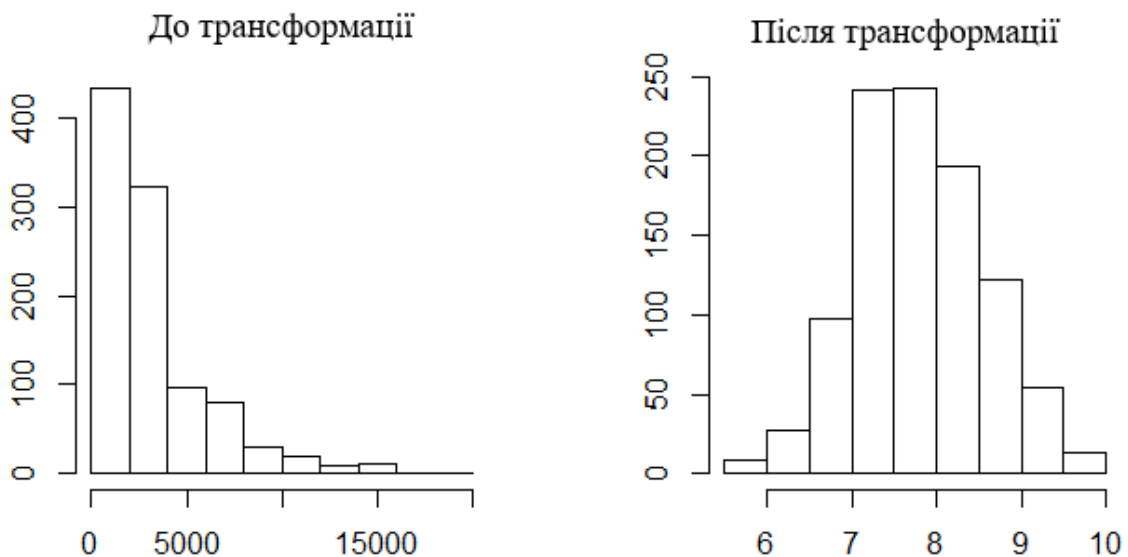


Рисунок 3.2 – Трансформація Бокса-Кокса

Коли поставленої мета була досягнена, і отримано бажану модель, можна розрахувати значення прогнозу. Кінцевим результатом буде певне число, або набір чисел, якщо прогноз будується на декілька часових точок наперед.

3.2 Метод часткового імовірнісного прогнозування

Метод часткового імовірнісного прогнозування достатньо легкий для розуміння, і виконується за 4 етапи:

- очищення даних;
- вибір періоду в минулому, який буде використовуватися;
- моделювання;
- розрахунок імовірності виникнення отриманих значень.

Очищення даних – це процес стабілізації нашого часового ряду. Якщо ми провели операції інтегрування і логарифмування, як було описано вище, то ряд вже повинен бути стаціонарним, тобто чистим.

Вибір періоду в минулому відіграє важливе значення, особливо у разі,

якщо спробувати побудувати прогноз використовуючи не стаціонарний часовий ряд. Цього варто уникати, але навіть з такої ситуації є вихід. Необхідно поділити часовий ряд на певні періоди. Не плутайте це зі значенням сезонної періодичності, хоча і його теж можна використовувати в певних випадках. На прикладі ярко вираженої сезонності, прогнозуючи значення на перші 4 місяці року ми могли б взяти не всі наявні дані, а лише дані з відповідною поведінкою, тобто дані за перші 4 місяці кожного року. Якщо провести якісь аналітичні дії стосовно наявних даних, можна обрати будь які періоди, що відповідають тим чи іншим потребам. Ще одним гарним прикладом, можна вважати відсіювання занадто старих даних.

Процес моделювання представляє собою витягування даних з обраного періоду. На основі цих даних вже можна проводити аналіз та будувати прогноз. Існують різні способи витягування, але у цій роботі ми розглянемо лише метод сформульований вченим, Бредлі Ефроном, у 1977 році, метод статистичного бутстрапу [17].

Сам англiцизм «бутстрап» використовується в багатьох галузях знань, де потрібно передати сенс того, що ви якимось чином з нічого отримуєте щось вартісне. Цей метод заснований на багаторазовій генерації підвiбірок методом Монте-Карло на базі наявної вибірки.

Метод Монте-Карло [18] – це чисельний метод, який використовує випадкові величини для моделювання та аналізу складних систем або процесів. Цей підхід отримав свою назву на честь знаменитого казино у Монте-Карло, оскільки метод ґрунтується на концепції випадковості та імовірностей.

Суть полягає в тому, щоб за наявною вибіркою побудувати емпіричний розподіл. Використовуючи його як теоретичний розподіл імовірностей, можна за допомогою генератору псевдовипадкових чисел отримати практично необмежену кількість підвiбірок довільного розміру. На їх основі можна оцінити не тільки аналізовані статистичні характеристики, а й вивчити їхні імовірнісні розподіли. Таким чином, наприклад, виявляється можливим

оцінити дисперсію або квантилі будь-якої статистики незалежно від її складності. Використовуючи такий підхід, нам важливо, щоб вибірка була репрезентативною, і наше емпіричне розподілення було схоже на реальне.

Отже, з часового ряду ми беремо вибірку. З неї випадковим чином беремо певну кількість даних з повтореннями, утворюючи псевдо-вибірки. Звісно, їх довжина може варіюватися в залежності від поставленої задачі, але не може бути менше 1 і більше розміру вибірки. У реальності, кількість псевдо-вбірок можна варіювати в залежності від обчислювальних потужностей, але не рекомендується брати менше 10000. На їх основі обчислюємо необхідну статистику.

Графічно, схему бутстрапу зображено на рис. 3.3.

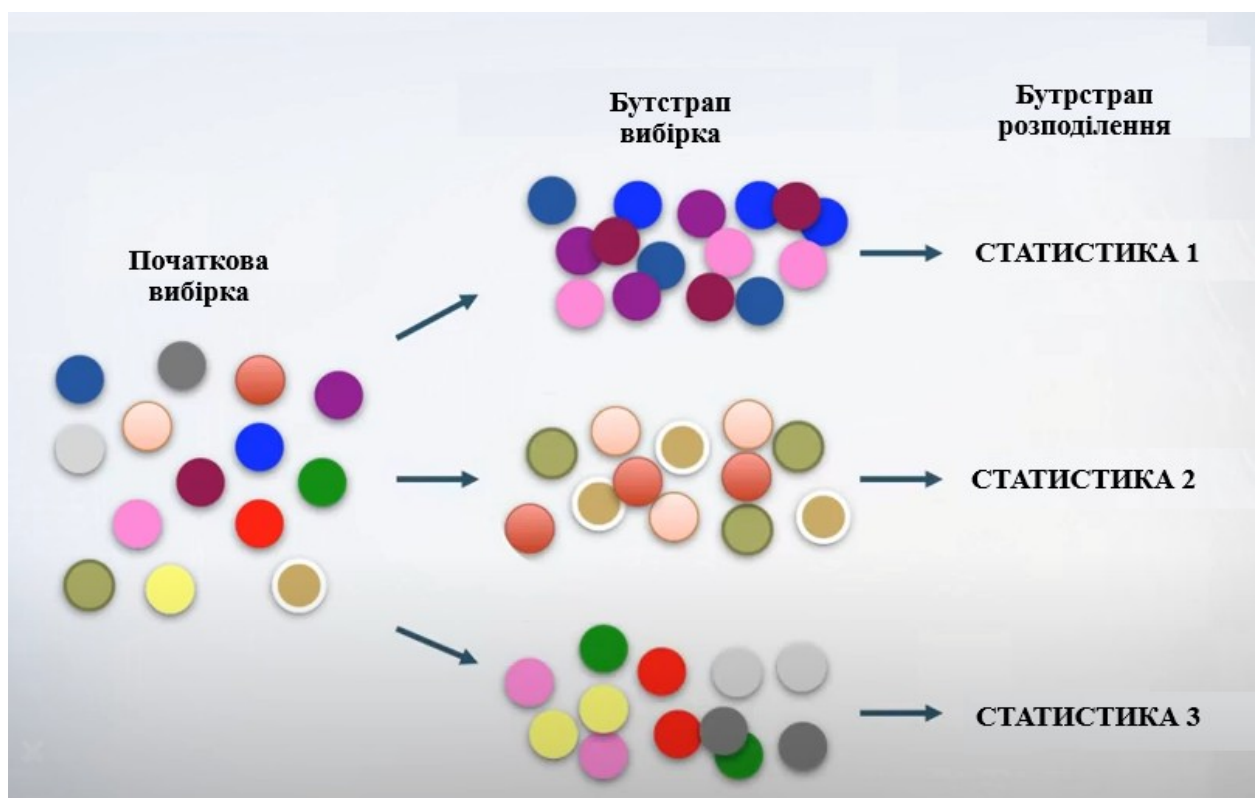


Рисунок 3.3 – Схема бутстрапу

Як результат, ми маємо велику множину результатів обчислення статистики. Щоб розрахувати імовірність виникнення певних значень, достатньо просто поділити кількість разів утворення цих значень на загальну

кількість розрахунків. Приклад такого розрахунку зображено у таблиці 3.1.

Таблиця 3.1 – Розподілення імовірностей

Результат	Кількість випадінь	Відносний відсоток	Загальний відсоток
5	1	0,10%	0,10%
6	1	0,10%	0,20%
7	3	0,30%	0,50%
8	4	0,40%	0,90%
9	13	1,30%	2,20%
10	23	2,30%	4,50%
11	43	4,30%	8,80%
12	55	5,50%	14,30%
13	64	6,40%	20,70%
14	92	9,20%	29,90%
15	92	9,20%	39,10%
16	121	12,10%	51,20%
17	119	11,90%	63,10%
18	97	9,70%	72,80%
19	80	8,00%	80,80%
20	81	8,10%	88,90%
21	47	4,70%	93,60%
22	32	3,20%	96,80%
23	19	1,90%	98,70%
24	9	0,90%	99,60%
25	3	0,30%	99,90%
26	1	0,10%	100,00%

Графічно ці дані можна зобразити у вигляді графіків розподілу імовірностей. Приклад цих графіків зображено на рис. 3.4.

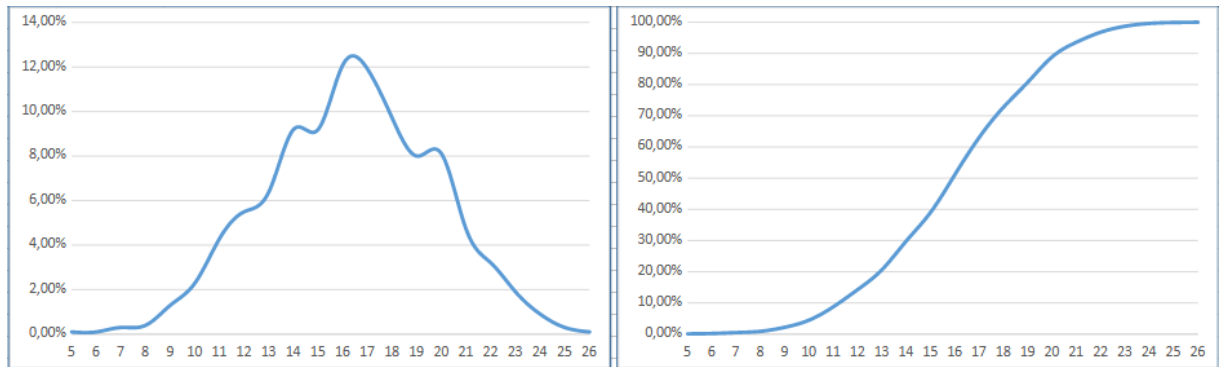


Рисунок 3.4 – Графіки розподілу імовірностей

3.3 Оцінка точності прогнозування

Отримавши якийсь результат, необхідно оцінити точність наших розрахунків. Основна величина, через яку оцінюється точність прогнозу - це похибка. Буквально, це різниця між прогнозованими значеннями і фактичними. Є очевидним, що чим більше похибка, тим більше ми помилилися. Для обчислення порівняльних коефіцієнтів похибка перетворюється, береться за модулем, або зводиться у квадрат. Без зазначених додаткових дій, результати оцінки втрачають будь-який сенс, бо від'ємні та позитивні значення можуть просто анулювати одне одного.

Загально прийнятими показниками, що використовують для оцінки результатів, можна вважати наступні: MSE, MAD, RMSE, MAPE.

MSE – середньоквадратична помилка. Коли нам потрібно підігнати криву під наші дані, то точність цієї підгонки буде оцінюватися саме за цією помилкою. Відповідно, ми прагне мінімізувати цей коефіцієнт. Обчислюється це все за наступною формулою:

$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n}, \quad (3.11)$$

де A_t – фактичне значення;

F_t – прогнозоване значення;

n – кількість спрогнозованих значень.

Чим ближче до нуля, тим модель краща, але Фізичного сенсу MSE не має. Тому науковці почали пошук альтернативних формул.

MAD – середнє абсолютне відхилення, визначається як частка від суми залишків за модулем до числа спостережень. Тобто, середній залишок за модулем. Виражений в абсолютних одиницях, MAD показує, наскільки одиниць у середньому помилятиметься прогноз. Обчислення відбувається за наступною формулою:

$$MAD = \frac{\sum_{t=1}^n |A_t - F_t|}{n}, \quad (3.12)$$

MAPE – середня абсолютна помилка у відсотках, що покликана надати моделі ще більш наочного сенсу. Фактично, це значення у відсотках, на яке в середньому може помилятися наша модель. Обчислення проводяться за наступною формулою:

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|}{n} \times 100. \quad (3.13)$$

RMSE – корінь середньоквадратичної помилки. Фактично має теж саме значення, що і MSE, але з більшою чутливістю до помилок. Чим значуща помилка в розрахунках, тим більший вплив вона буде мати на результат. Це ефективно, якщо ми особливо сильно боїмося великих помилок. RMSE рахується за наступною формулою:

$$RMSE = \sqrt{MSE}. \quad (3.14)$$

Щоб детальніше подивитись на роботу всіх цих показників, візьмемо дані за період в 12 одиниць часу, та якісь прогнозовані дані на цей самий період, і проведемо розрахунки. Обрані для перевірки дані з розрахунками представлені на рис. 3.5.

t	A_t	F_t	$A_t - F_t$	$ A_t - F_t $	$(A_t - F_t)^2$	$ A_t - F_t /A_t$
1	27,580	27,580	0,000	0,000	0,000	0,0000
2	25,950	26,765	-0,815	0,815	0,664	0,0314
3	26,080	26,015	0,065	0,065	0,004	0,0025
4	26,360	26,220	0,140	0,140	0,020	0,0053
5	27,990	27,175	0,815	0,815	0,664	0,0291
6	29,610	28,800	0,810	0,810	0,656	0,0274
7	28,850	29,230	-0,380	0,380	0,144	0,0132
8	29,430	29,140	0,290	0,290	0,084	0,0099
9	29,670	29,550	0,120	0,120	0,014	0,0040
10	30,190	29,930	0,260	0,260	0,068	0,0086
11	31,790	30,990	0,800	0,800	0,640	0,0252
12	31,980	31,885	0,095	0,095	0,009	0,0030
			2,200	4,590	2,968	0,160

Рисунок 3.5 – Обрані для перевірки дані з розрахунками

Розраховані показники оцінки точності представлені на рис. 3.6.

n	12
MAD	0,383
MSE	0,247
RMSE	0,497
MAPE	1,33

Рисунок 3.6 – Розраховані показники оцінки точності

3.4 Комбінований метод прогнозування

Метою дослідження є розробка комбінованого підходу до вирішення задачі прогнозування на основі інтеграції відомих та нових методів для підвищення якості та точності прогнозування показників діяльності підприємств електронної торгівлі.

Метод часткового імовірнісного прогнозування може базуватися на класичній моделі SARIMA, якщо він використовує імовірнісні властивості моделі для побудови прогнозів і оцінки невизначеності. Це досягається шляхом аналізу розподілів залишків, похибок прогнозу, моделі та параметрів часового ряду. SARIMA – це лінійна модель, яка передбачає, що залишки є випадковим процесом з нормальним розподілом і постійною дисперсією. Під час навчання моделі можна отримати оцінки розподілів параметрів і залишків, що дає змогу будувати імовірнісні прогнози.

Після побудови моделі можна спрогнозувати не одне значення, а множину можливих траєкторій часового ряду. Для цього із залишків виконується повторна вибірка, щоб врахувати їх статистичний розподіл. Потім до передбачених значень додаються випадкові шуми з вибірки залишків, що дає змогу отримати ансамбль прогнозів. Бутстрап генерує набір прогнозів, який наближається до справжнього розподілу майбутніх значень.

Отже, метод часткового імовірнісного прогнозування, заснований на статистичному бутстрапі, чудово поєднується з класичними моделями, такими як SARIMA. Бутстрап розширює можливості моделей, даючи змогу отримувати більш гнучкі та реалістичні імовірнісні прогнози.

4 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА МОЖЛИВОСТЕЙ ВИКОРИСТАННЯ РОЗРОБЛЕНОГО КОМБІНОВАНОГО МЕТОДУ ПРОГНОЗУВАННЯ

Перевіримо на практиці, як працює раніше описані методи прогнозування. Для початку, необхідно взяти певний часовий ряд з вираженою сезонністю і трендом, бо ці ознаки є характерними для підприємств електронної торгівлі. Для проведення розрахунків, у якості початкових даних, були взяті помісячні дані продажів магазину за останні 4 роки.

Графік продажів підприємства електронної торгівлі зображено на рис. 4.1.

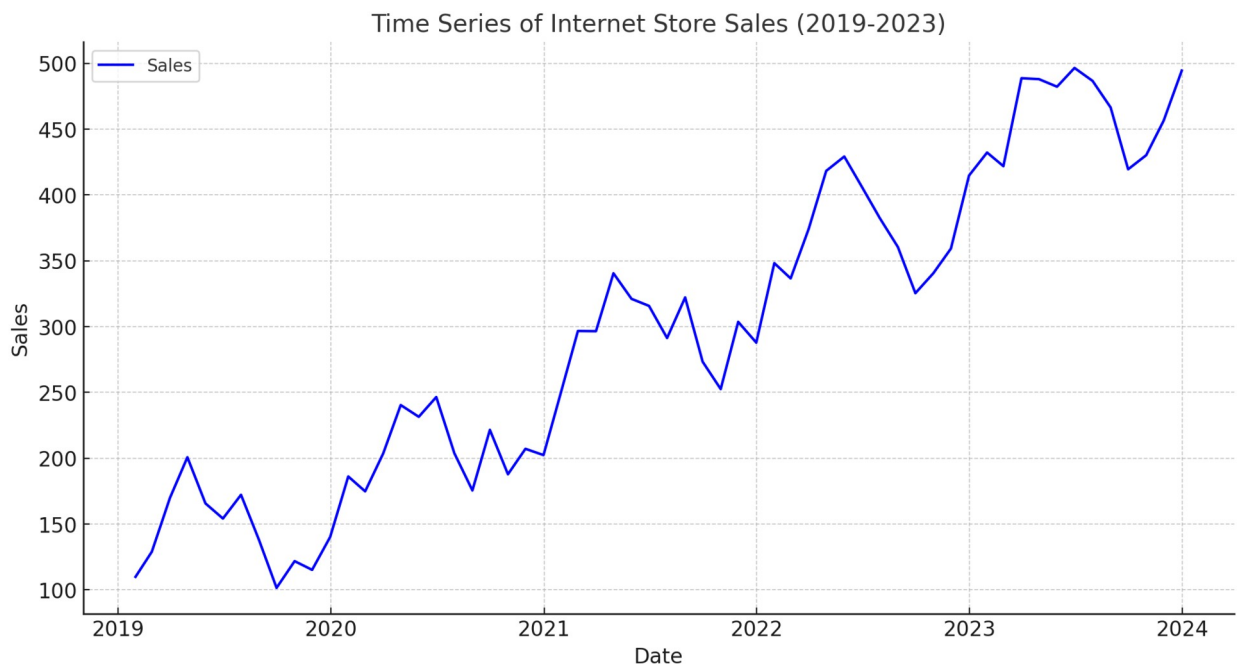


Рисунок 4.1 – Графік продажів підприємства електронної торгівлі

Як можна побачити, попит на товар має явні збільшення в першій половині кожного року, а у другій – зменшення. Це свідчить про наявність сезонності. Тим паче, з кожним роком продажі збільшуються, що є ознакою

тренду. Обидві ці характеристики є повною протилежністю стаціонарності часового ряду. Тому, щоб підготувати дані, та привести до необхідного виду, треба використати звичайне та сезонне диференціювання. Результати обох розрахунків зображено на рис. 4.2.

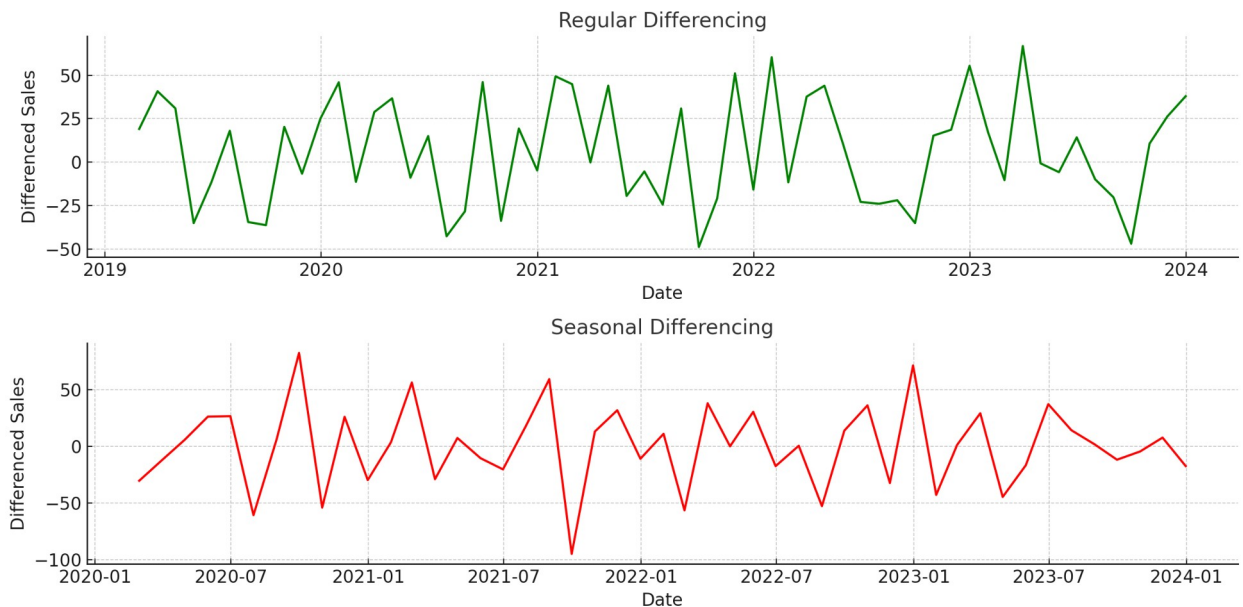


Рисунок 4.2 – Результати диференціювання

Використовуючи тест Дікі-Фуллера необхідно перевірити ряд на стаціонарність. За результатами цієї перевірки, підтверджується гіпотеза H_1 . Тобто, ряд все ще залишається не стаціонарним. Тому необхідно провести повторне диференціювання і повторити тест.

Результати другого тесту Дікі-Фуллера будуть наступні значення:

- статистика ADF: -6.9706;
- р-значення: $8.68e-10$;
- використано лагів: 9;
- спостережень: 36;
- критичні значення: -3.6267, -2.9460, -2.6117.

р-значення значно менше рівня значущості (0.05), що вказує на стаціонарність ряду. Графік двічі диференційованого ряду також побудовано та зображено на рис. 4.3.

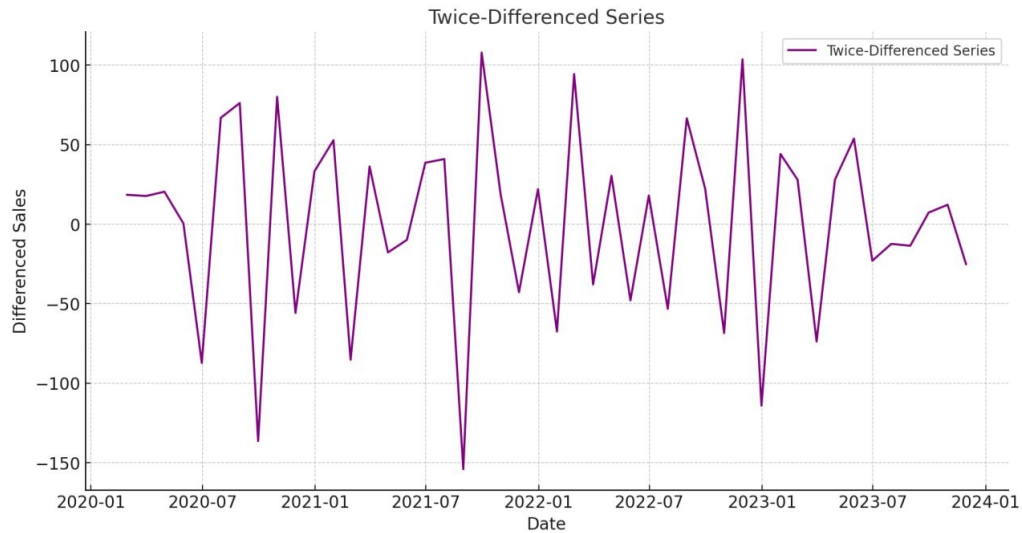


Рисунок 4.3 – Графік двічі диференційованого ряду

Коли дані готові для подальших обчислень, можна почати побудову прогнозу за методом SARIMA. Використовуючи сучасні бібліотеки, наприклад statsmodels у Python, можна досить швидко розрахувати прогноз. Нехай, останні 6 місяців будуть нашим тестовими даними, а усі інші – навчальними.

Результати прогнозування методом SARIMA у порівнянні з фактичними значеннями зображено на рис. 4.4.

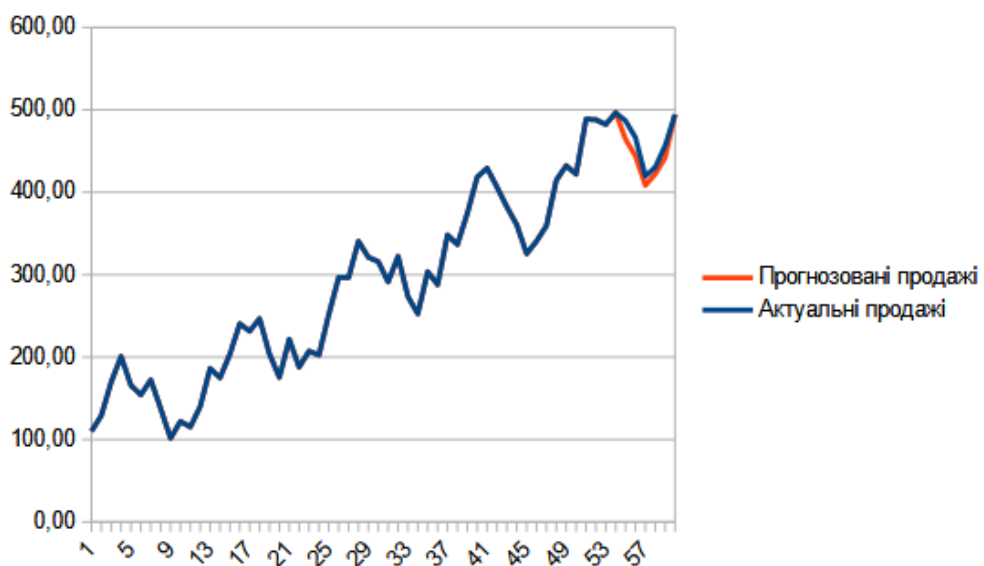


Рисунок 4.4 – Результати прогнозування методом SARIMA

Далі необхідно провести оцінку отриманого прогнозу. Для цього розраховуються наступні показники: MSE, RMSE, MAPE. Результати будуть наступними:

- MSE: 235,09;
- RMSE: 15,33;
- MAPE: 2,86.

Тепер, треба провести прогнозування з використанням розробленого комбінованого методу.

Першим кроком буде знаходження похибок отриманого раніше прогнозу за методом SARIMA.

Розраховані помилки відображено у табл. 4.1.

Таблиця 4.1 – Розраховані похибки

Актуальні продажі	Прогнозовані продажі	Похибка
486,72	464,30	22,42
466,51	444,04	22,47
419,58	408,33	11,24
430,26	422,04	8,22
456,54	442,09	14,45
494,51	494,47	0,04

Другим кроком, є використання методу статистичного бутстрапу. У якості початкової вибірки буде використовуватись множину отриманих залишків. Так як період прогнозування дорівнює кількості значень у вибірці, підвибірки будуть складатися з 6 чисел, тобто ідентичного розміру.

На цьому етапі, головним показником точності прогнозування є кількість створених підвибірок. У реальній практиці, рекомендується використовувати від 10000 до 50000. Але обчислювальні потужності для тестування достатньо невеликі, тому буде створено лише 500 підвибірок.

Після створення підвибірок, отриманні значення додаються до раніше спрогнозованих значень. Це створить велику множину варіацій, за якою

стане можливим оцінити імовірнісне розподілення. Графічно цей процес зображено на рис. 4.5.

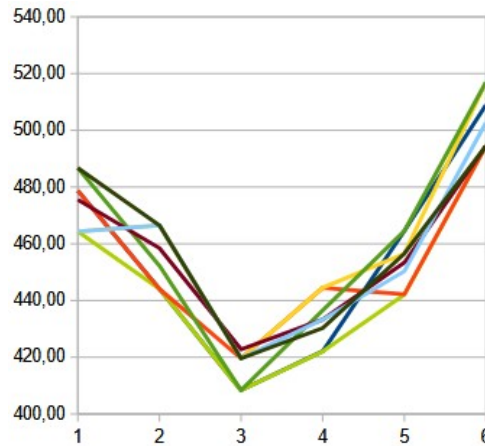


Рисунок 4.5 – Графічне представлення створення підвибірок

Тепер, маючи множину передбачень, можна оцінити імовірність виникнення тих чи інших результатів на кожному окремому значенні.

Усі наступні кроки залежать від конкретних потреб бізнесу. Можна побудувати довірчий інтервал, знайти для нього середнє значення, і це буде результатом прогнозу. Можна порахувати середнє значення помилки в усіх підвибірках, розрахувати імовірність виникнення цих даних, і оперувати вже їми. Всі ці варіанти мають право на існування і можуть бути використані в залежності від того, що саме прагне отримати підприємство.

В нашому прикладі, підприємство прагне знайти значення, що мають найбільшу імовірність не обмежуючи себе конкретним значенням рівня сервісу. Для цього, прогноз буде побудовано з використанням значень, що мають найбільшу власну імовірність.

Результати комбінованого прогнозу відображено у табл. 4.2.

Таблиця 4.2 – Результати прогнозу, що отриманий за допомогою комбінованого методу

Актуальні продажі	Комбінований прогноз
486,72	472,52
466,51	466,46
419,58	419,58
430,26	444,51
456,54	456,54
494,51	516,94

Графічне порівняння фактичних значень прогнозу за методом SARIMA та комбінованого методу, зображено на рис. 4.6.

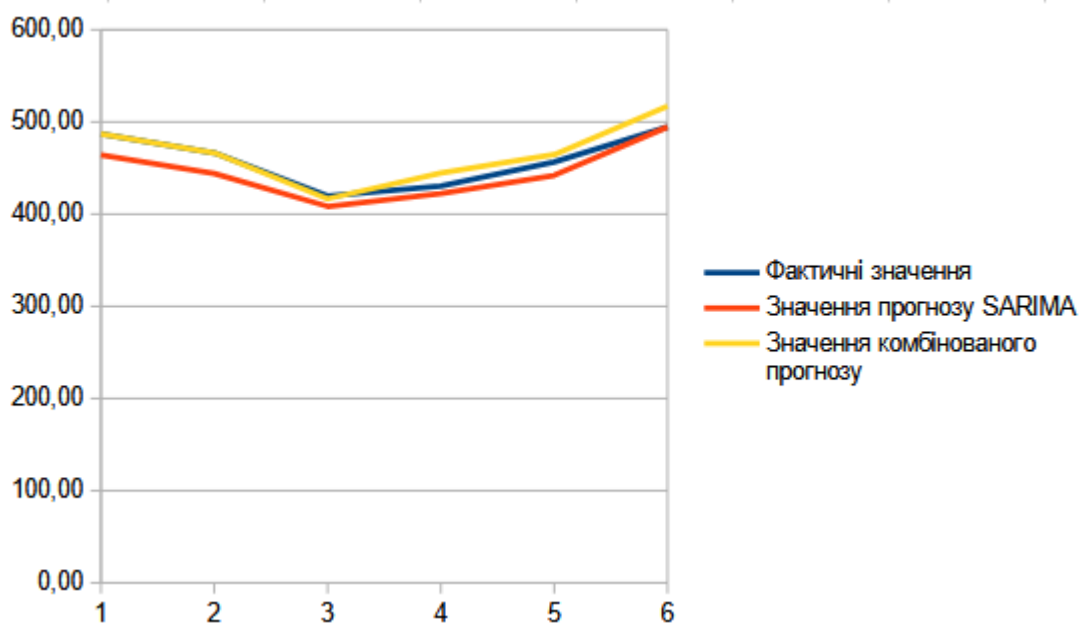


Рисунок 4.6 – Графічне порівняння отриманих прогнозів

Також, для впевненості у надійності прогнозу, що отриманий за допомогою комбінованого методу необхідно провести оцінку отриманих результатів і порівняти з оцінкою методу SARIMA.

Результати порівняння точності методів відображено у табл. 4.3.

Таблиця 4.3 – Результати порівняння точності методів

Метод	MSE	RMSE	MAPE
SARIMA	235,09	15,33	2,86
Комбінований метод	151,28	12,30	1,80

Можна побачити, що у розглянутому прикладі використання комбінованого методу прогнозування забезпечило збільшення точності приблизно на 1%. Це достатньо гарний результат, що слугує доказом підвищення ефективності прогнозування.

Звісно, багато залежить від того, чого потребує і якими інструментами збору даних володіє підприємство. У цій роботі було розглянуто базові дані продажів за часом. Але, якщо підприємство має більше інструментів збору статистики, це дає змогу провести більше різноманітних обчислень, отримуючи все більше і більше даних для прийняття рішень. Також, чим більше обчислювальних потужностей є в розпорядженні, тим більше вибірок можна створити. Що у свою чергу збільшить точність прогнозування.

Можна експериментувати з впливом довжини прогнозування на точність і зміну розподілення імовірностей. Теоретично, чим більша вибірка, тим більше розподілення можна побудувати. Але це питання залишиться для майбутніх досліджень.

ВИСНОВКИ

В ході виконання кваліфікаційної роботи проведено дослідження методів прогнозування соціально-економічних показників діяльності підприємства. Проаналізовано об'єкт дослідження та встановлено, що область електронної торгівлі має низку обмежень, що значно впливає на процес прогнозування. Тому, під час дослідження велика увага приділялась саме обмеженням і потребам відповідних підприємств.

Було визначено які методи прогнозування існують і на які групи поділяються. Проведено порівняння усіх груп за їх особливостям, та принципіальною різницею, що була в них закладена.

В ході проведення дослідження було встановлено які, з перелічених раніше груп методів, відповідають умовам і обмеженням підприємства електронної торгівлі. Було проведено детальний аналіз конкретних методів класичного і імовірнісного прогнозування, та сформовано перелік порівняльних характеристик. У результаті, для подальшого використання та комбінування, були обрані два метода, SARIMA та часткового імовірнісного прогнозування.

У третьому розділі роботи, було розглянуто технологію прогнозування обраними методами, та описано їх покрокове виконання. Також, для підвищення точності прогнозування, було розроблено комбінований метод.

На практичному прикладі, даних продажів підприємства електронної торгівлі, було проведено розрахунки класичним методом SARIMA та розробленою комбінацією. Порівняння точності прогнозування підтвердило ефективність розробленого комбінованого методу.

Розроблений під час дослідження метод може бути використаний при розробці ІС для планування роботи підприємства електронної торгівлі, що дозволить автоматизувати процеси прогнозування та прийняття рішень, зокрема, щодо управління запасами, ціноутворення, та маркетингових

стратегій. В залежності від ІС, можна провести прогнозування будь якого показника, якщо створено механізм для збору даних цього показника за часом.

Наприклад, використання розробленого методу для прогнозування попиту потребує від системи можливості автоматично отримувати дані продажів окремих товарів. В результаті, це дозволить знизити витрати на закупівлю товарів та зменшити ризики, пов'язані з надлишковими або дефіцитними запасами.

Окрім того, результати дослідження можуть бути корисними для дослідників у сфері прогнозування. У світі існує велика множина різноманітних галузей і кожна з них має власні потреби, обмеження, та властивості. А розробка комбінованих методів дозволить збільшити точність і гнучкість моделей, здатних враховувати численні змінні фактори. Властивості необхідної комбінації можуть змінюватись, відштовхуючись від потреб та обмежень конкретної галузі, для якої буде відбуватися прогнозування. Враховуючи розвиток сучасних технологій, навіть медична сфера може потребувати подібних рішень для прогнозування розвитку захворювань.

Робота виконана згідно з державними стандартами [19, 20].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Методичні вказівки щодо розробки та оформлення кваліфікаційної роботи другого (магістерського) рівня вищої освіти за освітньо-професійною програмою «Інформаційні управляючі системи та технології» спеціальності 122 Комп'ютерні науки / Упоряд.: Петров К.Е., Левикін В.М., Чалий С.Ф., Євланов М.В., Саєнко В.І., Міхнов Д.К., Міхнова А.В., Чала О.В. – Харків: ХНУРЕ, 2021. – 24 с.
2. Колупаєва І. В., Безсонов С. Стан і напрями розвитку електронної комерції в Україні. Проблеми економіки, 2 (56), 2023 с. 74 – 80.
3. Присенко Г. В. Прогнозування соціально-економічних процесів / Г. В. Присенко, Є. І. Равікович. – К.: КНЕУ, 2005. – 378 с.
4. Komunjer I. Quantile prediction. Handbook of economic forecasting. 2013. P. 961 – 994. URL: <https://doi.org/10.1016/b978-0-444-62731-5.00017-8> (дата звернення 02.01.2025).
5. Jianzhou Wang. Forecasting VaR and ES by using deep quantile regression, GANs-based scenario generation, and heterogeneous market hypothesis / Jianzhou Wang, Shuai Wang, Mengzheng Lv, He Jiang // Financial Innovation. – 2024. – №10. – Article number: 36 <https://doi.org/10.1186/s40854-023-00564-5>
6. Benth F. E., Kutrolli G., Stefani S. Dynamic probabilistic forecasting with uncertainty. SSRN electronic journal. 2019. URL: <https://doi.org/10.2139/ssrn.3405890> (date of access: 13.01.2025).
7. Probabilistic Forecasting [Електронний ресурс] // Forecast NOW. – Режим доступу: <https://fnow.pro/articles/veroyatnostnoe-prognozirovanie> (дата звернення 10.11.2024)
8. Sako K., Mpinda B. N., Rodrigues P. C. Neural networks for financial time series forecasting. Entropy. 2022. Vol. 24, no. 5. P. 657. URL: <https://doi.org/10.3390/e24050657> (date of access: 29.12.2024).

9. Калініна І. О. Дослідження нейромережових методів у задачах прогнозування / І. О. Калініна // Наукові праці [Чорноморського державного університету імені Петра Могили]. Сер.: Комп'ютерні технології. – 2009. – Т. 106, Вип. 93. – С. 132 – 138.
http://nbuv.gov.ua/UJRN/Npchduct_2009_106_93_19
10. Probabilistic Forecasting (Supply Chain) [Електронний ресурс] / Vermorel Joannes // lokad.com. – Режим доступу: <https://www.lokad.com/probabilistic-forecasting-definition/> (дата звернення 22.12.2024).
11. Тимошицький Б. О. Удосконалений метод прогнозування для планування роботи підприємства електронної торгівлі/ Тимошицький Б. О., Петров К.Е. // Collection of Scientific Papers with the Proceedings of the 1st International Scientific and Practical Conference «Scientific Innovation: Theoretical Insights and Practical Impacts» (January 13-15, 2025. Naples, Italy). European Open Science Space, 2025. – С. 234 – 235.
12. Canela M. Á., Alegre I., Ibarra A. Holt-Winters forecasting. Quantitative methods for management. Cham, 2019. P. 121 – 128. URL: https://doi.org/10.1007/978-3-030-17554-2_13 (date of access: 13.01.2025).
13. ARIMA. Прогнозування часових рядів [Електронний ресурс] // caseware.com. – Режим доступу: <https://caseware.com.ua/arima-3/> (дата звернення 10.11.2024)
14. Introduction to SARIMA Model [Електронний ресурс] / Ritu Santra // medium.com. – Режим доступу: <https://medium.com/@ritusantra/introduction-to-sarima-model-cbb885ceabe8> (дата звернення 27.12.24)
15. Almarashi A. M., Khan K. Bayesian structural time series. Nanoscience and nanotechnology letters. 2020. Vol. 12, no. 1. P. 54 – 61. URL: <https://doi.org/10.1166/nnl.2020.3083> (date of access: 13.01.2025).
16. Ritu Santra. Tests for Stationarity in Time Series – Dickey Fuller Test & Augmented Dickey Fuller(ADF) Test [Електронний ресурс] / Ritu Santra //

medium.com. – Режим доступу: <https://medium.com/@ritusantra/tests-for-stationarity-in-time-series-dickey-fuller-test-augmented-dickey-fuller-adf-test-d2e92e214360> (дата звернення 27.12.24).

17. An introduction to the bootstrap / ред. Т. Robert. New York : Chapman & Hall, 1993. 436 с.

18. Байда Є. І. Метод Монте-Карло [Електронний ресурс] : навч.-метод. посібник / Є. І. Байда ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2024. – 74 с. – URI: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/80846>

19. ДСТУ 3008:2015 Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання, 2016. – 16 с.

20. ДСТУ 8302:2015 Система стандартів з інформації, бібліотечної та видавничої справи. Бібліографічний запис. Бібліографічний опис, 2016. – 26 с.