

Хмельницький національний університет
Факультет програмування
та комп'ютерних і телекомунікаційних систем
Кафедра телекомунікацій, медійних та інтелектуальних технологій

ДИПЛОМНА РОБОТА МАГІСТРА

Прогнозна модель споживчого попиту на підприємствах роздрібної торгівлі
на основі методів машинного навчання

Галузь знань 11 – Математика та статистика

Спеціальність 113 – Прикладна математика

Шифр ДРПМ.19/022.01.04.00

Виконав: студент 2 курсу, група ПМм19-1


Підпис

Н. А. Беляков

Керівник: канд. пед. наук, доцент


Підпис, дата

О. Я. Кучерук

До захисту допускаю:

Зав. кафедри ТМІТ, д-р.тех.наук, доцент


Підпис, дата

С. К. Підченко

7 12 2020 р.

Хмельницький 2020

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ПРОГРАМУВАННЯ ТА КОМП'ЮТЕРНИХ І ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ

Кафедра ТЕЛЕКОМУНІКАЦІЙ, МЕДІЙНИХ ТА ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ

Освітній рівень МАГІСТР

Галузь знань 11 МАТЕМАТИКА ТА СТАТИСТИКА

Спеціальність 113 ПРИКЛАДНА МАТЕМАТИКА

Освітня програма ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА ПІДГОТОВКИ МАГІСТРА

ЗАТВЕРДЖУЮ

Зав. кафедри _____

" 2 " 09. 2020р.

ЗАВДАННЯ

НА ДИПЛОМНУ РОБОТУ

Беляков Нікіта Андрійович

Прізвище, ім'я, по батькові студента

1. Тема роботи Прогнозна модель споживчого попиту на підприємстві роздрібної торгівлі на основі методів машинного навчання

Керівник роботи Кучерук Оксана Ярославівна, канд. пед. наук, доцент

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 01.09.2020 р. № 118

2. Строк подання студентом роботи на кафедру 01.12.2020 р.

3. Вихідні дані до роботи Статистичні дані продажів підприємства "Мамин хліб" за товарними групами за півроку (щоденна статистика).

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

1. Проаналізувати сучасні підходи до проблеми прогнозування загалом та прогнозування споживчого попиту, зокрема.

2. Вивчити основні методи прогнозування попиту.

3. Вивчити теоретичні основи технології OLAP-аналізу даних та машинного навчання.

4. Побудувати прогнозну модель споживчого попиту підприємства "Мамин хліб".

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень)

6. Консультанти розділів дипломного проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання «01» 09 2020 р.

КАЛЕНДАРНИЙ ПЛАН

Назва етапів (розділів) дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1. Затвердження теми	01.09.2020 – 02.09.2020	Виконано
2. Аналіз сучасних підходів до прогнозування. Підготовка 1 розділу роботи.	02.09.2020 – 28.09.2020	Виконано
3. Вибір математичних методів для побудови прогнозної моделі споживчого попиту	28.09.2020 – 06.10.2020	Виконано
4. Прогнозування попиту за допомогою штучних нейронних мереж	06.10.2020 – 24.10.2020	Виконано
5. Оформлення практичних результатів дослідження	24.10.2020 – 30.10.2020	Виконано
5. Написання вступу, висновків, формування переліку джерел посилання та додатків	30.10.2020 – 08.11.20	Виконано
6. Попередній захист дипломної роботи	09.11.2020 – 10.11.2020	Виконано
7. Внесення коректив в роботу, оформлення роботи згідно вимог	11.11.2020 – 18.11.2020	Виконано
8. Подача роботи на: кафедру, перевірка на плагіат, рецензування роботи	19.11.2020 – 03.12.2020	Виконано
7. Представлення роботи до захисту	04.12.2020 – 15.12.2020	Виконано

Студент



Н.А. Беляков

Ініціали, прізвище

Керівник проекту (роботи)



О.Я. Кучерук

Ініціали, прізвище

Анотація

Тема дипломної роботи: «Прогнозна модель споживчого попиту на підприємствах роздрібно́ї торгівлі на основі методів машинного навчання»

Автор роботи: Беляков Нікіта Андрійович

Керівник роботи: доцент кафедри ТМІТ Кучерук Оксана Ярославівна

Загальний обсяг роботи: 94 сторінок, 32 рисунка, 6 таблиці, 2 додатків, 53 посилань.

СПОЖИВЧИЙ ПОПИТ, ПРОГНОЗНА МОДЕЛЬ ПОПИТУ, МАШИННЕ НАВЧАННЯ, НЕЙРОННІ МЕРЕЖІ.

Метою дипломної роботи є прогноуюча модель попиту на підприємствах роздрібно́ї торгівлі. Результатом виконання дипломної роботи є створення прогнозної моделі попиту групи товарів у трьох регіонах. Отримані результати дозволили оцінити попит продукції та скорегувати роботу підприємства.

Annotation

Thesis topic: «Forecast model of consumer demand at retail enterprises based on machine learning methods»

Author of the work: Belyakov Nikita Andreevich

Supervisor: Associate Professor of TMIT Kucheruk Oksana Yaroslavivna

Total volume of work: 94 pages, 32 figures, 6 tables, 2 appendices, 53 references.

CONSUMER DEMAND, FORECAST DEMAND MODEL, MACHINE LEARNING, NEURAL NETWORKS.

The purpose of the thesis is a predictive model of demand in retail enterprises. The result of the thesis is the creation of a forecast model of demand for a group of goods in three regions. The obtained results allowed to estimate the demand for products and adjust the work of the enterprise.

2.12

Дата/Date

2020р.


Підпис/Signature

Зміст

Вступ	7
1. Теоретичні основи моделювання споживчого попиту	10
1.1. Прогнозування попиту споживачів як основне завдання автоматизації процесів роздрібного магазину.....	10
1.2 Сучасні підходи до прогнозування споживчого попиту	14
1.2.1 Моделі часових рядів	21
1.2.2 Класичний регресійний аналіз	22
1.2.3 Регресія на опорних векторах	24
1.2.4 Регресія на основі методу «випадковий ліс»	26
1.2.5 Регресія на основі нейромережевого підходу	28
1.2.6 Регресія на основі композиції (ансамблю)	29
1.2.7. Регресія на основі градієнтного бустінга	31
2. Математичні підходи до побудови прогновної моделі споживчого попиту	34
2.1. Технологія OLAP	34
2.2. Основні поняття технологій машинного навчання	41
2.3 Прогнозування часових рядів за допомогою штучних нейронних мереж	49
3. Практичні результати дослідження	55
3.1. Аналіз попиту засобами OLAP	55
3.2. Прогнозування попиту за допомогою штучних нейронних мереж	62
Висновки	74
Перелік джерел посилання	76
Додатки	81

Перелік скорочень

Скорочення, термін, позначення	Пояснення
ГН	Глибоке навчання
МН	Машине навчання
ЧР	Часовий ряд

ВСТУП

Сучасні підприємства-виробники здійснюють свою діяльність у мінливому ринковому середовищі та зустрічаються з великою кількістю проблем. Джерелами підвищеної складності управління є висока ступінь невизначеності ринкової ситуації, сезонна нестабільність попиту на послуги, жорстока конкуренція в галузі торгівлі, нестача фінансових ресурсів та інше. Моделювання і прогнозування процесів, які протікають в таких складних системах, як екологічні, соціальні та економічні, до сих пір залишається однією з найактуальніших проблем.

Аналіз споживчого попиту є одним із основних завдань товаровиробників, а для ефективної роботи організації використовують різні методи його прогнозування. В даний час компаніями здійснюється накопичення історичних значень економічних і фізичних показників в базах даних, що істотно збільшує обсяги вхідної інформації для задачі прогнозування. Крім того, сучасні підходи до економічного та технічного управління пред'являють все більш жорсткі вимоги до точності прогнозування. Нині задача прогнозування актуальна і є невід'ємною частиною щоденної роботи багатьох компаній.

Високу актуальність та широке поширення інструменти прогнозування набувають у роздрібній торгівлі продовольчими товарами. В умовах швидкоплинності бізнес-процесів замовлення, розміщення, зберігання та продаж товарів, що швидко псуються, край важливо забезпечити баланс між задоволеністю клієнта та залишками товару, який буде утилізовано по закінченню терміну придатності. Тому аналіз попиту на продукцію має першорядне значення і є нині однією з важливих задач бізнес-аналітики. Стохастичність попиту накладає значні обмеження на роботу підприємств роздрібної торгівлі. Зі збільшенням кількості товарних позицій, збільшенням торгових площ та масштабів торгівлі забезпечити якісне прогнозування

попиту за допомогою прямих калькуляцій неможливо на практиці. Отже, важливим є удосконалення існуючих технологій прогнозування споживчого попиту [35].

Актуальним є активне впровадження новітніх інтелектуальних модельних технологій, що використовують принцип навчання і є реальним інструментом розв'язання складних задач моделювання та прогнозування розвитку процесів на товарних ринках. Науковцями визначається, що ознакою сьогодення є стрімка цифровізація всіх сфер соціально-економічної діяльності та адміністрування, зокрема, впровадження в процеси управління математичного та модельного інструментарію Business Intelligence, Data Mining, Data Science, Machine Learning, Artificial Intelligence [22].

Виходячи із вищенаведеного, метою роботи є розробка концепції прогнозування попиту на продукцію підприємства з використанням методів машинного навчання.

Об'єктом дослідження є підприємство роздрібної торгівлі ТМ «Мамин хліб».

Предмет дослідження – методи прогнозування споживчого попиту на продукцію підприємства.

Завдання дослідження:

1. Вивчити теоретичну основу організації та управління процесом прогнозування споживчого попиту.
2. Дослідити особливості прогнозування попиту у магазинах роздрібної торгівлі.
3. Здійснити аналіз існуючих підходів до прогнозування попиту споживачів.
4. Побудувати прогнозну модель та спрогнозувати обсяги попиту на продукцію підприємства ТМ «Мамин хліб».

Наукова новизна дослідження: запропоновано підхід до побудови моделі прогнозування попиту конкретної товарної групи на основі часової

вибірки даних із застосуванням сучасних методів машинного навчання, зокрема нейронних мереж.

Робота складається із вступу, трьох розділів, висновків, переліку посилань та додатків. У першому розділі проаналізовано сучасні підходи до проблеми прогнозування загалом та прогнозування споживчого попиту, зокрема. Висвітлено основні методи прогнозування попиту. У другому розділі описано теоретичні основи технології OLAP-аналізу даних, машинного навчання та застосування штучних нейронних мереж до прогнозування часових рядів. У третьому розділі представлено практичні результати дослідження.

За темою роботи опубліковано тези : Беяков Н. А., Кучерук О. Я. Застосування технології OLAP для аналізу даних споживчого попиту / Збірник наукових праць за матеріалами XII всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2020». Хмельницький – 2020. – С. 29-32.

1 ТЕОРЕТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ СПОЖИВЧОГО ПОПИТУ

1.1 Прогнозування попиту споживачів як основне завдання автоматизації процесів роздрібного магазину.

Під час управління компанією виникають труднощі, які завжди супроводжуються невизначеністю, пов'язаною з прийнятими рішеннями. Стан компанії та її середовища визначає процес прийняття рішень, і їх точність повинна базуватися на достовірній інформації. Прогнозування значно зменшує ризик прийняття неправильного рішення, а значить, виключає можливість збитків у багатьох сферах підприємницької діяльності, дозволяє оптимізувати та планувати бізнес.

В останні роки дедалі більше усвідомлюється необхідність використання методів прогнозування у багатьох сферах бізнесу. Зміни, що відбуваються в економіці, викликали необхідність "бачити" майбутнє і використовувати ці прогнози для підтримки неперервної діяльності компанії. Прогнозування є надзвичайно важливим процесом у сучасній економічній реальності [8].

Високу актуальність і широке поширення інструменти прогнозування набувають в роздрібній торгівлі продовольчими товарами. У сучасних умовах жорсткої конкуренції успішна діяльність виробничого або торгівельного підприємства неможлива без прогнозування попиту. Зміна доходів споживачів, а відповідно, їх купівельної спроможності, зміна цін на товари, вплив цілої низки інших факторів на попит, вимагає від підприємств мати чіткий план прогнозу попиту на товари, які випускаються або продаються. Дослідження та прогнозування споживчого попиту є однією із найзатребуваніших, але при цьому і найскладніших задач аналізу ринку [8].

Попит є складною економічною категорією, відбиває, з одного боку, потреба покупця в продукції, бажання придбати товари або послуги в певній кількості і, з іншого боку, можливість сплатити покупку за ціною, що знаходиться в межах «доступного» діапазону [4].

Питання формування й прогнозування споживчого попиту на товари та послуги розглянуто в роботах таких науковців як: Дж. Армстронг, Л. Брейман, С. Брю, Г. Джеймс, Ф. Котлера, У. Маккінлі, А. Мюллер, Д. Уіттон, О. Бозуленко, Б. Грабовецький, Л. Ліпич, А. Нікульченко, В. Пучкова, Л. Строць, Е. Федусенко та ін.

Прогноз – це кількісне, ймовірне, науково-обгрунтоване судження про можливий майбутній стан системи або явища і (або) про можливі альтернативи і терміни їх реалізації. Таким чином, прогнозування є способом наукового передбачення, в рамках якого застосовується як сформований раніше досвід, так і поточні припущення для визначення майбутніх подій [22].

Б. Грабовецький визначає прогноз як «імовірне, аргументоване (тобто на підставі системи фактів і доказів) судження про стан об'єктів (процесів) в майбутньому або альтернативних шляхів і строків досягнення певних результатів» [13]. Прогнозування – це процес формування прогнозів.

Визначень прогнозу існує безліч, вони впливають із різних факторів, прогностичних ситуацій, методів дослідження або їх цілей.

В економіко-математичної термінології прогнозування – це система наукових досліджень якісного та кількісного характеру, спрямованих на з'ясування тенденцій розвитку народного господарства або його частин (галузей, регіонів, підприємств і т.д.) і пошук оптимальних шляхів досягнення цілей цього розвитку. Подібне поняття відображає конкретні аспекти прогнозування наявність якісних і кількісних висновків, на підставі яких здійснюється прогноз, але при цьому включає в себе також процедуру управління – в цьому випадку пошук способів досягнення значень отриманого прогнозу при здійсненні завдань [35].

Процес прогнозування складається з двох основних фаз: діагностування минулого та визначення майбутнього. Мета діагностики минулого – дізнатись про природу прогнозованого явища, механізми його функціонування, а також оцінити фактори, що його формують. Ця фаза відповідає за збір даних про минуле, які обробляються для побудови формальної моделі або ментальної моделі.

Визначення майбутнього – це перехід від оброблених даних до прогнозу. Цей перехід називається правилом прогнозування. Більшість прогностичних процесів виділяють кілька універсальних етапів [1]:

- визначення прогностичної проблеми,
- вибір змінних,
- отримання та попередній аналіз даних,
- вибір методу дослідження,
- побудова прогнозу та його допустимості,
- використання прогнозу та його перевірка.

Методи прогнозування – це спеціально визначені методи поведінки, що використовуються для вирішення завдань прогнозування. Методи обробки даних супроводжуються різними правилами прогнозування. Таким чином, метод прогнозування складається з двох елементів: моделі та правила прогнозування [1].

Вибір методів прогнозування в першу чергу залежить від історичних спостережень (даних), економічного стану даного підприємства та інформації, яку ми хочемо отримати.

Існує багато класифікацій методів прогнозування, і вони залежать від прийнятого критерію. Класифікація методів прогнозування детально досліджена, зокрема, в роботах О. Кривда, Ю. Сидоренко, Т. Кравченко.

За часом (або періодом), на який розроблюються прогнози, вони поділяються на коротко-, середньо- та довгострокові. Характерною особливістю короткострокових прогнозів є невеликий діапазон екстраполяції порівняно з кількістю наявних даних, більше того, це прогноз лише на кілька

проміжків часу вперед або на конкретний момент, для якого можуть відбутися кількісні зміни досліджуваного явища. Середньостроковий прогноз відрізняється невеликим діапазоном екстраполяції середньої кількості даних. Це дозволяє прогнозувати середню кількість періодів часу та прогноз моменту, до якого можуть відбутися кількісні зміни та незначні якісні зміни досліджуваного явища. З іншого боку, довгостроковий прогноз має найбільший діапазон, це прогноз на велику кількість одиниць часу, і аж до прогнозованого часу можуть відбуватися як кількісні зміни, так і суттєві якісні зміни досліджуваного явища [1].

За ступенем формалізації всі методи прогнозування можна умовно поділити на інтуїтивні (експертні методи прогнозування), формалізовані (методи прогнозування екстраполяції, системно-структурні методи та моделі, асоціативні методи, математичні методи та моделі) і комплексні [39].

Л. Строчень в своєму дослідженні запропонувала класифікацію методів прогнозування споживчого попиту (рис. 1.1)

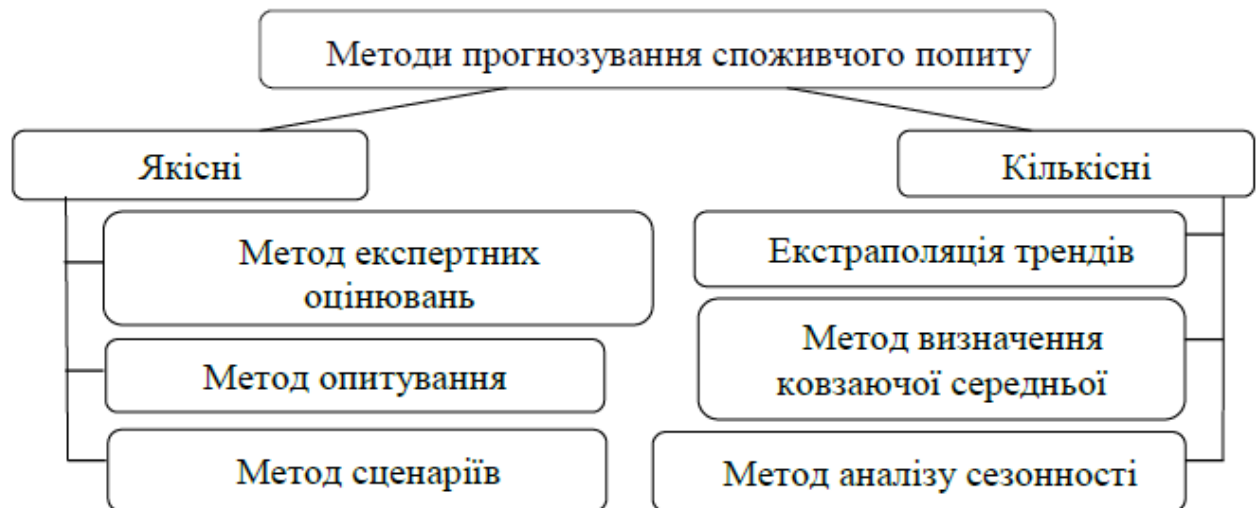


Рисунок 1.1 – Методи прогнозування попиту (за [40])

А. Нікульченко поділяє існуючі методи прогнозування споживчого попиту на три основні групи: методи експертних оцінок; методи аналізу та прогнозування часових рядів; казуальні (причинно-наслідкові) методи.

Методи експертні оцінки ґрунтуються на суб'єктивній оцінці (індивідуальній або груповій) поточного моменту і перспектив розвитку. Ці методи доцільно використовувати для кон'юнктурних оцінок, особливо в ситуаціях, коли статистична інформація про обсяги минулих продажів відсутній [32].

Друга і третя групи методів засновані на аналізі кількісних показників, але вони істотно відрізняються один від одного. Методи аналізу та прогнозування часових рядів пов'язані з дослідженням ізольованих показників, кожен з яких складається з двох елементів: прогнозу детермінованою компоненти і прогнозу випадкової компоненти. Найбільш відомими методами прогнозування, заснованими на екстраполяції, є методи, які використовують модель авторегресії і ковзного середнього (ARIMA). Своєю популярністю ці методи, в першу чергу, зобов'язані роботі Дж. Боксу і Г. Дженкінса, які запропонували та розвинули узагальнену модель ARIMA [32].

В основі казуальних методів прогнозування лежить пошук факторів, які впливають на значення прогнозованого показника, що призводить до необхідності побудови моделі, що враховує взаємозв'язок факторів. У разі прогнозування споживчого попиту на товар побудова подібної моделі вимагає обліку дуже великої кількості різноманітних факторів і тому є досить трудомісткою [32].

1.2 Сучасні підходи до прогнозування споживчого попиту

У країнах з розвинутою економікою для вирішення різних завдань управління все частіше використовуються математичні методи прогнозування, причому їх розподіл тим ширше, чим більше підприємство. У джерелах, які досліджують цю тему, показано, що понад 50% великих

підприємств і приблизно 18% середніх і дрібних підприємств при аналізі і прогнозуванні різних економічних і фінансових показників використовують саме такі підходи.

Математичні методи прогнозування можна поділити на три групи:

- методи експертних оцінок;
- детерміновані методи;
- стохастичні методи.

Методи експертних оцінок передбачають багатоступінчасте опитування експертів за спеціальними схемах і обробку отриманих результатів за допомогою інструментарію економічної статистики. Це найбільш прості і популярні методи, історія яких налічує не одне тисячоліття. Їх застосування на практиці зазвичай полягає у використанні досвіду і знань фахівців [34].

Детерміновані методи припускають наявність функціональних чи жорстко детермінованих зв'язків, коли кожному значенню факторної ознаки відповідає цілком визначене не випадкове значення результативної ознаки.

Стохастичні методи припускають імовірнісний характер як прогнозу, так і самого зв'язку між досліджуваними показниками. Імовірність отримання прогнозу зростає з ростом числа емпіричних даних. Ці методи займають провідне місце з позиції формалізованого прогнозування й істотно варіюються по складності використовуваних алгоритмів [39].

Концепцію прогнозування попиту на продукцію підприємства економіко-математичними методами представимо у вигляді алгоритму, який складається з таких етапів [34].

На першому етапі необхідно визначити фактори, які будуть включені в модель, вони мають відображати найсуттєвіші особливості модельованого процесу, і їх кількісні характеристики мають бути приведені до зіставленого виду за одиницями виміру, цінами, часом реєстрації, методологією розрахунку. Оскільки попит на кожний товар залежить від багатьох факторів, то багатофакторні моделі будуються диференційовано, тобто для різних товарів у модель вводяться різні аргументи. У найзагальнішому вигляді

залежність між величиною попиту і визначальними його факторами записують таким чином:

$$y = f(p, p_1 \dots p_n, t, Z, W, T, F, N, X), \quad (1.1)$$

де y – величина попиту на товар;

p – ціна товару;

$p_1 \dots p_n$ – ціни на інші товари;

t – час;

Z – дохід споживача;

W – купівельна спроможність споживача;

T – смаки і вподобання споживачів;

F – очікування споживачів щодо майбутніх цін;

N – кількість покупців;

X – інші фактори.

Включення в модель великої кількості факторів наближує її до реальності, але збільшує трудомісткість її розробки та можливу наявність мультиколінеарності між факторами [39].

Найпростішими є однофакторні моделі, наприклад, трендові або адаптивні моделі (Брауна, Хольта, Тейла – Вейджа) з єдиним фактором впливу – часом, але такі моделі не пояснюють причини зміни попиту, тому сфера їх використання обмежується короткостроковим прогнозуванням. Перевагою адаптивних моделей є те, що вони дозволяють врахувати різну економічну цінність рівнів динамічного ряду попиту та враховують результат прогнозу, розробленого на попередньому кроці, але вони не здатні передбачити еволюцію зміни попиту [44].

У структурних моделях попит розглядається як функція, яка залежить лише від доходу, знаючи частоти розподілу споживачів за рівнем доходу, такі моделі дозволяють розраховувати загальну структуру попиту.

Ціна є одним з основних чинників зміни попиту, тому при прогнозуванні попиту поширений підхід, коли будують однофакторну модель залежності попиту від ціни (p) або двофакторну модель залежності попиту від ціни та її зміни у часі ($p(t)$).

На другому етапі встановлюється математична залежність між попитом і факторами (x_i), які його обумовлюють, і знаходяться параметри моделі. На практиці найчастіше використовуються такі залежності:

- дробово-лінійна:

$$y = \frac{a_1 Z}{Z + c_1}; \quad (1.2)$$

$$y = \frac{a_2(Z - b_2)}{Z + c_2}, (Z \geq b_2); \quad (1.3)$$

$$y = \frac{a_3 Z (Z - b_3)}{Z + c_3}, (Z \geq b_3). \quad (1.4)$$

- степенева:

$$y = a_0 + \sum_{i=1}^n a_i x_i; \quad (1.5)$$

- показникова:

$$y = A \prod_{i=1}^n x_i^{a_i}; \quad (1.6)$$

$$y = A \prod_{i=1}^n a_i^{x_i}. \quad (1.7)$$

Наприклад, Л. Торнквіст пропонує спеціальні види функцій попиту залежно від доходу (Z) для чотирьох груп товарів відповідно: товари першої необхідності, товари другої необхідності, предмети розкоші та нижчі товари [16].

Товари першої необхідності – найбільш потрібні споживачеві і часто вживані продукти. Величина еластичності попиту за доходом на такі товари

менше одиниці, оскільки з ростом доходів споживачі прагнуть купувати не товари першої необхідності, а більш рідкісні товари. Крім того, в таких товарів, як правило, коефіцієнт еластичності за доходом з ростом доходів знижується, оскільки в них існує межа насичення. Візьмемо, наприклад, такий товар, як зубна паста. Існує певний кількісний рівень, вище якого зубна паста людям просто не потрібна, так як вони не зможуть її використати. З наближенням до цього рівня коефіцієнт еластичності попиту за доходом прагне до нуля: як би ми не збільшували доходи, попит на зубну пасту не збільшиться [16]. Функція Торнквіста для товарів першої необхідності Y_1 виглядає так:

$$Y_1 = a_1 Z / (Z + C_1), \quad (1.8)$$

де a_1 — верхня межа попиту; C_1 — приріст доходу [28].

Товари другої необхідності –товари для яких крива ємності ринку має увігнутий вид і наближається зі зростанням доходів до верхньої межі споживання товарів даної групи, який має більше значення, ніж для товарів першої необхідності; при цьому попит на дану групу товарів з'являється після того, як дохід досягає певного розміру, після якого виникає можливість придбання товарів даної групи [16]. Функція Торнквіста для товарів другої необхідності Y_2 виглядає так:

$$Y_2 = a_2 (Z - b_2) / (Z + C_2), \quad (1.9)$$

де $Z \geq b_2$; a_2 –верхня межа попиту; b_2 –певний рівень доходу; C_2 –приріст доходу. Коли дохід досягає величину b_2 , з'являється попит на цю групу товарів [28].

Предмети розкоші –товари, без яких можна обійтися в житті, вони не пов'язані з задоволення базових потреб людини. Величина еластичності

попиту за доходом на такі товари більше одиниці, оскільки з ростом доходів частка предметів розкоші в споживанні зростає, а частка товарів першої необхідності падає. Крім того, слід зазначити, що попит на такі товари виникає лише при досягненні певного рівня доходів, в той час як попит на товари першої необхідності позитивний вже при найменшому рівні доходів. Це пов'язано з тим, що люди спочатку прагнуть задовольнити нагальні базові потреби, пов'язані з власним виживанням, і лише потім починають думати про предмети розкоші [16]. Функція Торнквіста для предметів розкоші Y_3 має такий вигляд:

$$Y_3 = a_3 Z (Z - b_3) / (z + C_3), \quad (1.10)$$

де $Z \geq b_3$; $a_3 > 1$; b_3 – певний рівень доходу; C_3 – приріст доходу. Функція Y_3 без меж, після того як дохід Z перевищить рівень b_3 , виникає попит на предмети розкоші [28].

Нижчі товари – товари низької якості, попит на які зі зростанням доходів зменшується, оскільки споживачі починають замінювати їх більш якісними товарами. До цієї категорії відносяться, наприклад, одяг «секонд-хенд», другосортні продукти харчування і т.д. Коефіцієнт еластичності попиту за доходом для таких товарів спершу позитивний, але потім з досягненням певного рівня доходів, опускається до нуля і слідом за тим досягає негативних значень [19].

Для знаходження параметрів регресійної моделі існує низка методів, найбільш часто використовуваним з яких є метод найменших квадратів.

На третьому етапі здійснюється перевірка адекватності та точності моделі. Модель вважається адекватною, якщо її залишкова компонента задовольняє властивостям випадкової компоненти:

- випадковість коливань рівнів залишкової послідовності (перевірка проводиться за допомогою критерію Кендела або критерію серій);

- відповідність розподілу випадкової компоненти нормальному закону розподілу (може бути проведена наближено за допомогою дослідження показників асиметрії та ексцесу);
- рівність математичного очікування випадкової компоненти нулю (перевірка проводиться на основі t-критерію Ст'юдента);
- відсутність автокореляції (перевіряється, наприклад, за критерієм Дарбіна – Уотсона).

Для адекватних моделей формулюється задача оцінки точності. Статистичними показниками точності виступають такі показники: середньоквадратичне відхилення, коефіцієнт детермінації, середня відносна помилка апроксимації. На підставі зазначених показників можна зробити вибір з декількох адекватних моделей найбільш точної [36].

На заключному, четвертому етапі, здійснюється прогноз попиту шляхом підстановки прогнозних значень факторів, врахованих у моделі, у розраховане рівняння регресії. Для оцінки прогнозуючих властивостей моделі доцільно використовувати ретроспективний прогноз. Оцінювання прогнозних властивостей моделі на ретроспективній ділянці особливо корисне при зіставленні різних моделей прогнозування з числа адекватних [27].

На цьому ж етапі проводиться верифікація прогнозової моделі, яка представляє собою сукупність критеріїв і способів, що дозволяють на основі багатобічного аналізу оцінити якість одержуваного прогнозу. Однак найчастіше на етапі верифікації здійснюється оцінка методу прогнозування, за допомогою якого був отриманий результат, ніж оцінка якості самого результату [26].

1.2.1 Моделі часових рядів

Найбільш часто використовується підхід прогнозування попиту є інструментарій аналізу часових рядів. Задача прогнозування майбутніх значень часового ряду на основі його історичних значень виникає в різних галузях природних і технічних наук, економічній сфері, соціологічних та інших дослідженнях.

Для аналізу часових рядів та прогнозування найчастіше використовуються методи математичної статистики та різні типи авторегресійних моделей [21].

На сьогоднішній день існує багато моделей прогнозування часових рядів: регресійні і авторегресійні моделі, нейромережеві моделі, моделі експоненціального згладжування, моделі на базі ланцюгів Маркова, класифікаційні моделі та ін. Найбільш популярними і таким, що широко використовуються, є класи авторегресійних і нейромережевих моделей.

Часовим рядом зазвичай називають послідовність вимірювань $y_{t1}, y_{t2}, \dots, y_{tn}$ випадкової величини ξ_t , які зроблені в послідовні моменти часу t_1, t_2, \dots, t_2 [4]. Часовий ряд включає в себе два обов'язкові елементи – позначку часу і значення показника ряду, отримане тим чи іншим способом і відповідне зазначеній позначці часу.

В рамках даної задачі мова йде про часовий ряд попиту на той чи інший товар. Відповідно, за допомогою методів аналізу часових рядів, можливо формалізоване уявлення математичної моделі попиту як змінюється в часі випадкової величини ξ_t з певним законом розподілу [7].

Найбільш популярними видами моделей часових рядів для прогнозування попиту є:

- тренд-сезонні моделі;
- моделі експоненціального згладжування;
- моделі авторегресії і ковзного середнього.

Це пояснюється тим, що дані моделі добре апроксимують специфічні елементи часового ряду попиту, наприклад, сезонність.

Тренд-сезонні моделі є фундаментальними моделями часових рядів, які визначають основні складові часового ряду: T – тренд часового ряду або так звану тенденцію (напрямок), в якій рухається тимчасової ряд, S – сезонна компонента тимчасового ряду, яка відображає амплітудні коливання навколо лінії тренду, і E – помилка моделі. Залежно від характеристик процесу елементи часового ряду можуть бути об'єднані двома способами:

- адитивно: $y_t = T_t + S_t + t$
- мультиплікативно: $y_t = T_t \times S_t \times t$

У даному випадку головною процедурою по створенню моделі часового ряду є декомпозиція або розкладання вихідного тимчасового ряду на складові T , S і E . Слід зазначити, що незважаючи на досить просту інтерпретацію процесу, методи, засновані на декомпозиції, в середньому відрізняються більш низькою якістю прогнозування [7].

1.2.2 Класичний регресійний аналіз

Більш сучасними методами побудови лінійної регресії є методи з регуляризацією. Під регуляризацією мається на увазі процедура звуження значення для певної групи коефіцієнтів b , які надають менший вплив на якісний результат прогнозування. Метод регуляризації зазвичай застосовується при великій кількості коефіцієнтів у моделі. Це актуально для завдання прогнозування попиту, так як вибірка може мати багато детермінант в залежності від проведеного попереднього аналізу. Найбільш поширеними методами, в основі яких лежить процедура регуляризації, є гребнева і ласо регресія [19]. При знаходженні коефіцієнтів для гребневої або ласо регресії

мінімізують наступний функціонал на основі МНК з деяким штрафними складовими:

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^m \beta_j x_{tj})^2 + \lambda \sum_{j=1}^m K = RSS + \lambda \sum_{j=1}^m K \rightarrow \min; \quad (1.11)$$

$$\begin{cases} k = \beta_j^2, \text{ якщо оцінюється гребнева регресія} \\ |k = \beta_j|, \text{ якщо оцінюється ласо регресія} \end{cases}$$

де β_j – це коефіцієнт моделі лінійної регресії (в матричній формі – b), β_0 – вільний коефіцієнт (зрушення) лінійної регресії, K – це тип оцінки коефіцієнтів, який вибирається в залежності від методу, λ – це гіперпараметр методу, за допомогою якого вибирається сила стиснення коефіцієнтів моделі.

У сучасному світі дані мають властивість накопичуватися у величезних масштабах, відповідно з'являються не тільки просторові або часові вибірки, а й просторово-часові вибірки або панельні дані. Зазвичай подібні вибірки формуються для вирішення економетричних проблем високого порядку [15]. Відповідно, в сфері торгівлі також формується великий обсяг даних. Це можуть бути дані по товарах і їх характеристикам за певний період часу, що цілком відповідає вимогам просторово-часової вибірки. Для панельних даних необхідний особливий підхід до аналізу та прогнозування на їх основі, тому були розроблені змішані регресійні моделі [32]:

$$y_{ij} = \beta_0 + \sum_{k=1}^p \beta_k x_{kji} + \sum_{k=1}^q d_{mj} z_{mji} + \varepsilon_{ji}; \quad (1.12)$$

де β_0 – константа, β_k – коефіцієнти для фіксованих ефектів, x_{kji} – значення фіксованих ефектів, d_{mj} – коефіцієнти для випадкових ефектів, z_{mji} – значення випадкових ефектів.

Поняття фіксованих і випадкових ефектів засновані на відношенні змінної до процесу дослідження. Фіксовані ефекти – це те, що є керованою змінною в системі, випадкові – це змінні, вплив яких виникає випадковим

чином. Змішані регресійні моделі є по суті своїй розвитком стандартного методу МНК. Безсумнівна перевага розглянутого підходу до задачі прогнозування попиту полягає в тому, що інтерпретація отриманих результатів дає уявлення про процес формування купівельного попиту в цілому. Проте, лінійна регресія не завжди дає якісні результати в точності прогнозування в порівнянні з іншими методами.

Область наукового знання – машинне навчання, яке активно уточнює і розвиває методи прогнозування, впливає в тому числі і на вирішення завдань у сфері роздрібно́ї торгівлі. Далі описуються кілька методів прогнозування, які є регресійними за своєю природою, але дозволяють врахувати можливу нелінійність у взаємозв'язках між змінними.

1.2.3 Регресія на опорних векторах

SVM або метод опорних векторів (від англ. Support Vector Machine) є одним з алгоритмів, що вивчаються у машинному навчанні, який характеризується своєю універсальністю – він може бути використаний як в задачах регресії, так і в задачах класифікації. Метод заснований на поділу класів за допомогою гіперплощини, що знаходиться в $N - 1$ просторі. При цьому для розширення простору використовуються так звані ядерні функції. Згідно з визначенням, даним Воронцовим [13], функція $K: X \times X \rightarrow R$ називається ядром (від англ. Kernel function), якщо задана функція була подана таким чином: $K(\mathbf{x}, \mathbf{x}') = \langle \psi(\mathbf{x}), \psi(\mathbf{x}') \rangle$ при деякому відображенні $\psi: X \rightarrow H$, де H - простір зі скалярним добутком.

Розширенням машини опорних векторів для задач регресії називається метод регресії на основі опорних векторів (від англ. Support Vector Regression). На відміну від лінійної регресії, заснованої на методі найменших квадратів,

метод регресії на основі опорних векторів використовує інший функціонал. Він визначається тільки тими залишками, які мають значення вище деякої позитивної константи. Далі формулюється завдання побудови SVR для вирішення задачі нелінійної регресії [19]. Множина $\{(x_i, y_i)\}_{i=1}^n$ визначається як множина прикладів навчання в завданні регресії [4]. Необхідно знайти множники Лагранжа $\{\alpha_i\}_{i=1}^n, \{\alpha_i'\}_{i=1}^n$, які максимізують цільову функцію.

$$j(a_i, a_i') = \sum_{i=1}^N y_i(a_i - a_i') - \varepsilon \sum_{i=1}^N (a_i + a_i') - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (a_i - a_i')(a_j - a_j')K(x_i, x_j); \quad (1.13)$$

При наявності наступних обмежень:

$$\sum_{i=1}^N (a_i - a_i') = 0; \quad (1.14)$$

$$\begin{aligned} 0 \leq a_i \leq C, i = 1, 2, \dots, N, \\ 0 \leq a_i' \leq C, i = 1, 2, \dots, N, \end{aligned} \quad (1.15)$$

де C – константа, яка задається дослідником, $K(x_i, x_j)$ – ядро скалярного твору, яке, відповідно до теореми Мерсера [4], визначається наступним чином:

$$K(x_i, x_j) = \varphi^T(x_i) \varphi(x_j). \quad (1.16)$$

Два параметра у вихідній постановці і C вибираються вільно з урахуванням структури і природи даних [7]. Зазначені параметри характеризують розмірність Вапніка-Червоненкіса (VC – розмірність) апроксимуючої функції:

$$F(x, w) = w^T x = \sum_{i=1}^N (a_i - a_i')K(x_i, x_j); \quad (1.17)$$

де w – вектор ваг, x – вектор ознак.

1.2.4 Регресія на основі методу «випадковий ліс»

Ще одним універсальним алгоритмом при вирішенні задач класифікації і прогнозування є метод "випадковий ліс" (від англ. Random forest). Алгоритм заснований на побудові великої кількості дерев рішень і по суті відображає просту ідею про те, що усереднений результат по великій кількості простих моделей є найкращим в порівнянні з результатом по одній складній моделі. Метод розроблений Брейманом і Катлером [42] і по суті є ансамблем дерев рішень. Опис індуктивного алгоритму випадкового лісу розгорнуто представлено в роботі Чистякова [43]:

1. Для кожного з B створюваних дерев рішень, що беруть участь в ансамблі, виконуються наступні процедури:
 - Сформувати випадкову вибірку з повторенням S розміру N по вихідній навчальній вибірці $D = \{(x_i, y_i)\};_{i=1}^n$
 - За вибіркою S побудувати дерево рішень T_i , не вдаючись до процедури відсікання гілок дерева, з мінімальною кількістю спостережень в термінальних вершинах рівним n_{\min} [39]. Для цього необхідно рекурсивно слідувати наступним алгоритмом:
 - i. з вихідного набору n ознак випадково вибрати m ознак;
 - ii. з m вибрати ознака, який забезпечує найкраще розщеплення дерева. Тут визначення оптимальної розщеплюєності вершини дерева будується на понятті забрудненості цієї вершини. Під забрудненням розуміють міру неоднорідності, яка максимальна якщо прецеденти, пов'язані з цією вершиною належать різним класам. Для оцінки забрудненості зазвичай використовують поняття ентропії $i(t) = -\sum_{j=1}^c P(\omega_j) \log_2 P(\omega_j)$, де t - вершина дерева рішень, $P(\omega_j)$ - частка прикладів класу ω_j в підвибірці $D(t)$. Звідси оптимальне розщеплення визначається як

$$\Delta i_B(t) = \frac{\Delta i(t)}{-\sum_{i=1}^B P_i \log_2 P_i} \quad (1.18)$$

де B - кількість нащадків вершини дерева t , P_k - частка прикладів підвибірки $D(t)$. Тут вибирається розщеплення максимізуючи величину $\Delta i_B(t)$.

- iii. розщепити вибірку після знаходження оптимального розбиття на дві підвибірки;
2. Після проведення всіх ітерацій на кроці 1 отримано $\{T_i\}_{i=1}^B$ - ансамбль дерев рішень [35];
 3. Розрахунок прогнозу для задач регресії за отриманим ансамблем проводиться наступним чином:

$$f_{rf}^B(x) = \frac{1}{B} \sum_{i=1}^B T_i(x); \quad (1.19)$$

Out-of-Bag – метод оцінки помилки класифікації при побудові випадкового лісу. По суті, помилка out-of-bag оцінюється на вибірці неправильно класифікованих прикладів без врахування голосів дерев на прикладах, які і так входять в їх навчальну вибірку. Метод випадкового лісу є невибагливим до форматів даних, що безумовно призводить до його досить широкого поширення. Одним з найголовніших недоліків методу є розширені вимоги до розрахунку і зберігання кількості дерев B , яка зазвичай досягає декількох сотень. Випадкові ліса досить часто застосовуються, як інструмент аналізу даних і в економічному аналізі: наприклад, в страхуванні і в аналізі ймовірності банкрутств [16].

1.2.5 Регресія на основі нейромережевого підходу

Провідним алгоритмом машинного навчання, який в сучасному світі використовується в якості теоретичної основи багатьох програм з прогнозування, є штучна нейронна мережа. Штучна нейронна мережа є адаптація біологічного підходу до організації систем високого порядку. Це означає, що в основі методу лежить взаємодія нейронів – нервових клітин в живому організмі. Нейрони реагують на подразники і передають інформацію наступним нейронам [20]. Організація процесу в штучній нейронній мережі аналогічна. Як показала практика, застосування цього підходу виправдано і успішно, особливо в задачах розпізнавання зображень, відеоряду та звуку. По суті це односпрямована мережа, в якій нейрони розташовані на різних рівнях, відповідних поняттю шару. При вирішенні простих завдань машинного навчання використовуються багатошарові нейронні мережі з одним прихованим шаром, як це представлено на малюнку 1.2, рідше з двома [8].

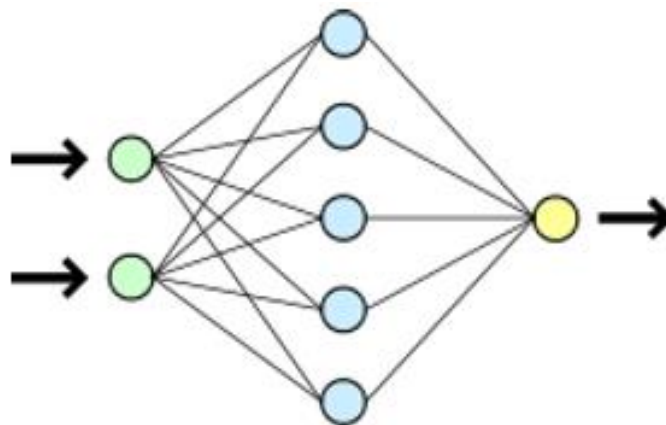


Рисунок 1.2 – Багатошарова нейронна мережа з одним прихованим шаром

Існує кілька типів функціонування нейронів, тобто базових одиниць нейронної мережі. Це може бути простий персептрон, крім багатошарового персептрона існує велика кількість інших видів нейромережевих технологій:

радіальні базисні мережі, рекурентні нейронні мережі, згорткові нейронні мережі, нейронні мережі з застосування нечіткої логіки. Вибір того чи іншого типу нейронної мережі, налаштування її топології, зазвичай визначається ресурсними можливостями і характеристикою самого завдання. В цілому застосування нейронних мереж в задачах економіки набуло великий характер. У тому числі нейронні мережі використовуються і для прогнозування рівня попиту і продажів [8]. Дуже часто нейронні мережі застосовуються для прогнозування економічних часових рядів. Все це робить нейронні мережі методом з дуже широкою сферою застосування. При цьому вони не позбавлені недоліків, пов'язаних зі спеціальною підготовкою даних і досить високими вимогами на обчислювальні ресурси. Це призводить до того, що на практиці, при вирішенні задач прогнозування на великому обсязі даних, використання нейромережевих технологій стає занадто витратним [5].

1.2.6 Регресія на основі композиції (ансамблю)

Велика кількість методів, які використовуються при побудові прогнозів, говорить про те, що немає унікального рішення для будь-якого завдання. Зазвичай, кожен з розглянутих методів машинного і статистичного навчання, має свої особливості і свою якість прогностичного рішення для того чи іншого завдання. Коли необхідно виконати високоточне рішення задачі прогнозування (або класифікації), використовується концепція композиції алгоритмів. Вводяться поняття множин X і Y як безліч ознак і безліч фактичних виходів (прогнозованої і класифікується величини), також позначають множину R як множина оцінок. Далі розглядаються алгоритми, які мають вигляд $a(x) = C(b(x))$, де функція $b: X \rightarrow R$ називається алгоритмічним

оператором, функція $C: R \rightarrow Y$ - вирішальним правилом. Звідси виводиться визначення композиції алгоритмів [10].

Композицією T алгоритмів $a_t(x) = C(b_t(x))$, $t = 1, \dots, T$ називається суперпозиція алгоритмічних операторів $b_t: X \rightarrow R$, коректує операції $F: R^T \rightarrow R$ і вирішального правила $C: R \rightarrow Y$ [15]:

$$a(x) = C(F(b_1(x), \dots, b_T(x))), x \in X. \quad (1.20)$$

Основною метою композиції є мінімізація помилок окремих базових моделей.

В якості методів побудови композицій або ансамблів алгоритмів машинного навчання можуть вважатися:

- Беггінг (від англ. Bagging). Сенс беггінга полягає в створенні великої кількості випадкових вибірок з вихідних даних простим вибором з заміщенням. Результатом методу вважається об'єднання прогнозів різних алгоритмів, зроблених на сформованих випадкових вибірках [33].
- Бустінг (від англ. Boosting, поліпшення). Для цього алгоритму характерно послідовне поліпшення результату, шляхом компенсування втрат при реалізації попереднього базового алгоритму машинного навчання. Бустінг і беггінг мають схожу природу з даним алгоритмом випадкового лісу, що говорить про генеральну ідею – безліч умовно поганих алгоритмів при створенні композиції можуть дати хороший алгоритм. Ефективність бустінга доведена на практиці, особливо бустінга, заснованого на деревах рішень. Це пояснюється тим, що при послідовному додаванні базових алгоритмів збільшуються відступи навчальних об'єктів [33].
- Стекінг (від англ. Stacking). В основі цього методу створення ансамблів моделей лежить поняття мета-алгоритму над існуючими базовими алгоритмами. У простій інтерпретації для реалізації ідеї стекінг

необхідно розбити вихідну вибірку на дві частини; на першій частині навчити кілька алгоритмів машинного навчання; на другій частині розрахувати результат. На фінальному етапі за прогнозами, отриманими на останньому етапі, навчається мета-алгоритм [33].

Ансамблі реалізуються для побудови надточних рішень, коли набагато важливіший результат прогнозування, ніж його інтерпретація. В рамках завдання прогнозування попиту цей підхід має місце бути виходячи з рівня помилки при побудові базових регресійних алгоритмів [33].

1.2.7 Регресія на основі градієнтного бустінга

Градiєнтний бустiнг є результатом розвитку тих iдей, якi закладенi в перших алгоритмах бустiнга. Алгоритм починає роботу з побудови початкової моделi i коригує її, крок за кроком створюючи послiдовнiсть дерев регресiї або використовуючи iншi базовi методи (лiнiйна регресiя, нейронна мережа i т.д.). Кожне дерево в послiдовностi створюється на пiдставi залишкiв моделi, яка зводиться на попередньому кроцi [26]. Залишки моделi по сутi використовуються в якостi цiльової змiнної. По сутi вирiшується завдання:

$$F = \sum_{i=1}^m L(y_i, a(x_i) + b_i) \rightarrow \min, \quad (1.21)$$

де $a(x_i)$ – спочатку побудований алгоритм, b_i – вибудовується наступний алгоритм, який здiйснює коригування вiдповiдей $a(x_i)$ до вiрних. Таким чином, полiпшується функцiонал $i = y_i - a(x_i)$. Розглядається вираз слiд визначати як мiнiмiзацiю функцiї $F(b_1, \dots, b_m)$, отже, необхідно вибрати ефективний метод

її мінімізації. В даному випадку функція від багатьох змінних максимально зменшується в напрямку свого антиградієнта:

$$-(L'(y_i, a(x_1)), \dots, L'(y_m, a(x_m))); \quad (1.22)$$

Відповідно відповіді алгоритму b_i можуть бути визначені наступним чином:

$$b_i = -L'(y_i, a(x_i)), i \in \{1, 2, \dots, m\}; \quad (1.23)$$

Налаштування алгоритму відбувається на навчальній вибірці:

$$(x_i, -L'(y_i, a(x_i)))_{i=1}^m; \quad (1.24)$$

Звідси визначається і назва методу – градієнтний бустінг [26]. Далі, після побудови відповідей алгоритму b_i , будується третій алгоритм, який коригує зазначену суму. А потім, як уже було сказано вище, процес триває. Це призводить до одного з основних параметрів градієнтного бустінга – кількості ітерацій бустінга N . Параметр впливає на точність одержуваних результатів. Наступним важливим параметром є η або швидкість навчання (learning rate). Суть його полягає в тому, що зміщення аргументу в F відбувається тільки на частину вектора для того щоб зберегти баланс між точністю і швидкістю збіжності алгоритму:

$$a_{t+1}(x) = a_t(x) + \eta \cdot b_t; \quad (1.25)$$

Параметр визначено в межах $\eta \in (0, 1]$.

Популярною реалізацією алгоритму градієнтного бустінга є градієнтний бустінг над вирішальними деревами [40].

Таким чином можна зробити наступні висновки:

1) Опис системи руху товару розкриває системну складність і економічну проблематику на торгових підприємствах. Для поліпшення роботи подібних систем необхідно чітко розуміти структуру руху товару і володіти інструментами економіко-математичного аналізу.

2) Розглянуті теоретичні аспекти управління рухом товарів і сутність купівельного попиту показують ключову важливість якісного прогнозування потреби для поліпшення ефективності прийнятих рішень. Тому вдосконалення функціонування організації роздрібної торгівлі в рамках дипломної роботи базується на побудові моделі прогнозування попиту на основі просунутого математичного і статистичного аналізу [40].

3) Окремо виділяються і розглядаються методи математичного прогнозування попиту для подальшої побудови методології прогнозування.

2 МАТЕМАТИЧНІ ПІДХОДИ ДО ПОБУДОВИ ПРОГНОЗНОЇ МОДЕЛІ СПОЖИВЧОГО ПОПИТУ

2.1 Технологія OLAP

Розробка рішень по управлінню підприємством потрапляє в розряд областей які найбільш складно піддаються автоматизації. На сьогоднішній момент проблему розуміння та встановлення взаємозв'язків між агрегованими даними найкращим чином вирішують програмні продукти, що використовують багатовимірний оперативний аналіз даних OLAP (On-line analytical processing) [5].

Технологія OLAP є сучасною концепцією аналізу даних, описану сукупністю вимог до програмних продуктів, які забезпечують оперативно аналітичну обробку та подання даних. Технологія OLAP застосовується, щоб спростити роботу з багатоцільовими накопиченими даними про діяльність підприємства в минулому і не загрузнути в їх великому обсязі, а також перетворити набір кількісних показників у якісні, дозволяє аналітикам, менеджерам і керуючим сформулювати своє власне бачення даних, використовуючи оперативний доступ до різноманітних форм подання інформації [6].

Технологія OLAP дає можливість швидко змінювати погляд на дані в залежності від обраних параметрів і забезпечує особі, яка приймає рішення, повний огляд ситуації з його власної стратегічної точки зору.

Вперше принципи OLAP були сформульовані основоположником теорії реляційних баз даних Е. Коддом (E.F.Kodd) в Книзі "OLAP для Користувачів-аналітиків: яким він повинен бути". Технологія інтерактивного багатовимірного аналізу даних (OLAP) була розроблена спеціально для вирішення завдань аналізу великих масивів накопичених даних, орієнтована на обробку нерегламентованих, несподіваних запитів аналітиків до даних [6].

Кодд сформулював концепцію комплексного багатовимірного аналізу даних, накопичених у сховищі, у вигляді 12 основних правил, які повинні задовольняти OLAP-системи, як продукти, що надають можливість виконання оперативної аналітичної обробки. У 1995 році до них були додані ще шість (які відомі в значно менше) [45].

У 1995 році на основі вимог, викладених Коддом, було сформульовано так званий тест FASMI (Fast Analysis of Shared Multidimensional Information – швидкий аналіз розподілу багатовимірної інформації), що включає такі вимоги до додатків для багатовимірного аналізу [37]:

- надання користувачеві результатів аналізу за прийнятний час (зазвичай не більше 5 с), нехай навіть ціною менш детального аналізу;
- можливість здійснення будь-якого логічного і статистичного аналізу, характерного для цього додатку, і його збереження в доступному для кінцевого користувача вигляді;
- розрахований на багато користувачів доступ до даних з підтримкою відповідних механізмів блокування і засобів авторизованого доступу;
- багатовимірне концептуальне представлення даних, включаючи повну підтримку для ієрархій і множинних ієрархій (це ключова вимога OLAP);
- можливість звертатися до будь-якої потрібної інформації, незалежно від її обсягу і місця зберігання [45].

OLAP забезпечує користувача природною, інтуїтивно зрозумілою моделлю даних, організовуючи їх у вигляді багатовимірних кубів. Багатовимірний підхід до аналізу надає аналітику широкі можливості моделювання даних у відповідності зі сформованим у нього уявленням про проблему, знижуючи тим самим ймовірність помилкової інтерпретації отриманих результатів. Багатовимірне уявлення, на думку відомих аналітиків, є найбільш природним поглядом керуючого персоналу на об'єкт управління [26].

Застосування OLAP-технологій дає можливість побудувати прогнози на основі узагальнених агрегованих даних і таким чином отримати загальну

картину або дозволяє провести більш детальне опрацювання питання, побудувавши моделі прогнозування і обчисливши значення прогнозів для елементів нижнього рівня.

Концепція OLAP-технології заснована на наступних поняттях [23]:

1) Багатомірний куб (n – число вимірювань багатомірному кубу) – це багатомірна структура, що складається з множини комірок і зберігає взаємопов'язані дані, що описують предметну область або її складову. Комірка є атомарною структурою куба і при відображенні розташовується всередині нього. Багатомірний куб може створюватися і зберігатися як на фізичному рівні представлення даних (багатомірні OLAP системи), так і на концептуальному рівні (віртуально) за допомогою схем (реляційні OLAP системи) [23].

2) Показник q_j ($j = 1 \dots p$, p – кількість показників) – це типізована величина, яка є предметом аналізу (наприклад, показниками є величина закупівель, величина продажів і т.д.). Один багатомірний куб може містити кілька показників, причому кожному конкретному значенню показника відповідає єдина комірка багатомірного куба [23].

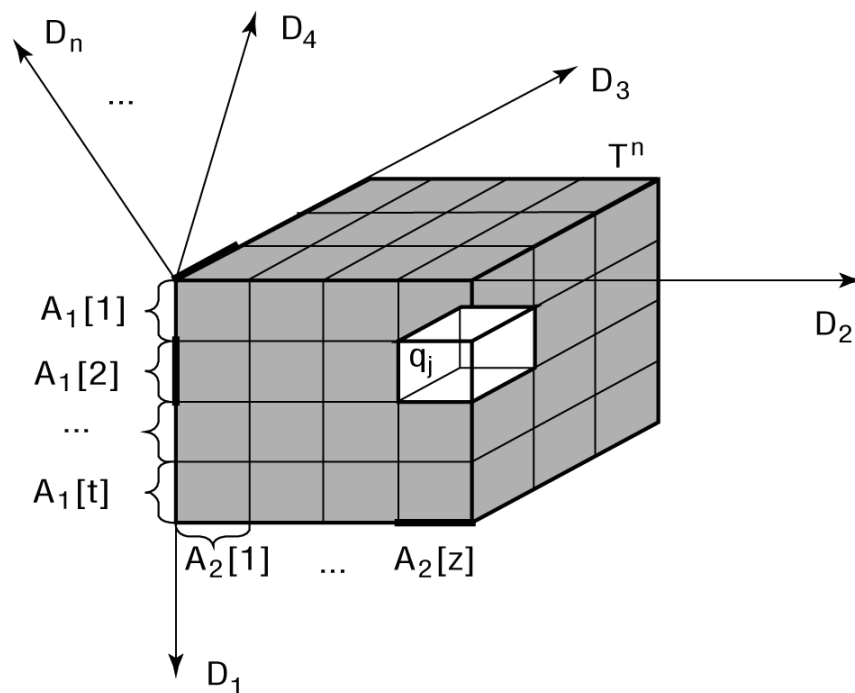


Рисунок 2.1 – Багатомірна модель даних (за [23])

3) Вимірювання D_i ($i = 1 \dots n$) – це множина об'єктів одного або декількох типів, організованих у вигляді ієрархічної структури і забезпечують інформаційний контекст типізованого показника. Наприклад, вимір «товар» може складатися з наступних значень: «побутова техніка», «комп'ютери», «мобільні телефони» і т.д. Сукупність вимірів утворює множину вимірів $D = \{D_1, D_2, \dots, D_n\}$ [23].

Значення, пов'язані з вимірюванням, характеризують будь-яку класифікаційну властивість сутностей предметної області. На рис. 2.1 $\{A_1 [1], A_1 [2], \dots, A_1 [t]\}$ – множина значень вимірювання A_1 (t – кількість елементів вимірювання A_1), $\{A_2 [1], \dots, A_2 [z]\}$ – множина значень вимірювання A_2 (z – кількість елементів вимірювання A_2). Вимірювання прийнято відобразити у вигляді ребра багатовимірного куба, а значення вимірювання – у вигляді точок або ділянок, що відкладаються на ребрі.

Множина всіх вимірювань куба утворює систему координат яка представляється у вигляді простору даних. Значення вимірювань відіграють роль індексів, що знаходяться в комірках куба.

Особливістю вимірювань є їх ієрархічна структура, яка використовується для агрегації і деталізації значень показників. Найбільш часто використовуються наступні два види ієрархій [23]:

– Збалансовані – ієрархії, в яких число рівнів визначено її структурою незмінно, і всі гілки мають однакову глибину. Для формування збалансованої ієрархії необхідна наявність зв'язку «один до одного» між об'єктами менш детального рівня по відношенню до об'єктів більш детального рівня [23].

– Незбалансовані – ієрархії, в яких число рівнів може бути змінено, деякі рівні логічно не однакові, і глибина гілок ієрархії може бути різною. Всі об'єкти незбалансованої ієрархії належать одному типу.

4) Елемент вимірювання – рівень в ієрархії значень координат вимірювання багатовимірного куба.

5) Атрибут вимірювання. Атрибути вимірювання описують його елементи, але не утворюють ієрархії.

При роботі з гіперкубом користувач, який аналізує інформацію, може «розрізати» куб за різними напрямками, отримати зведені (наприклад, по роках) або, навпаки, детальні (по місяцях) відомості та здійснювати інші маніпуляції, які необхідні в процесі аналізу [23].

Візуалізація OLAP-куба проводиться за допомогою спеціального виду таблиць, які будуються на основі зрізів OLAP-куба, що містять необхідну користувачу інформацію. Зрізи є результатом виконання відповідного запиту до бази даних.

Загальну схему роботи настільної OLAP системи можна представити таким чином:

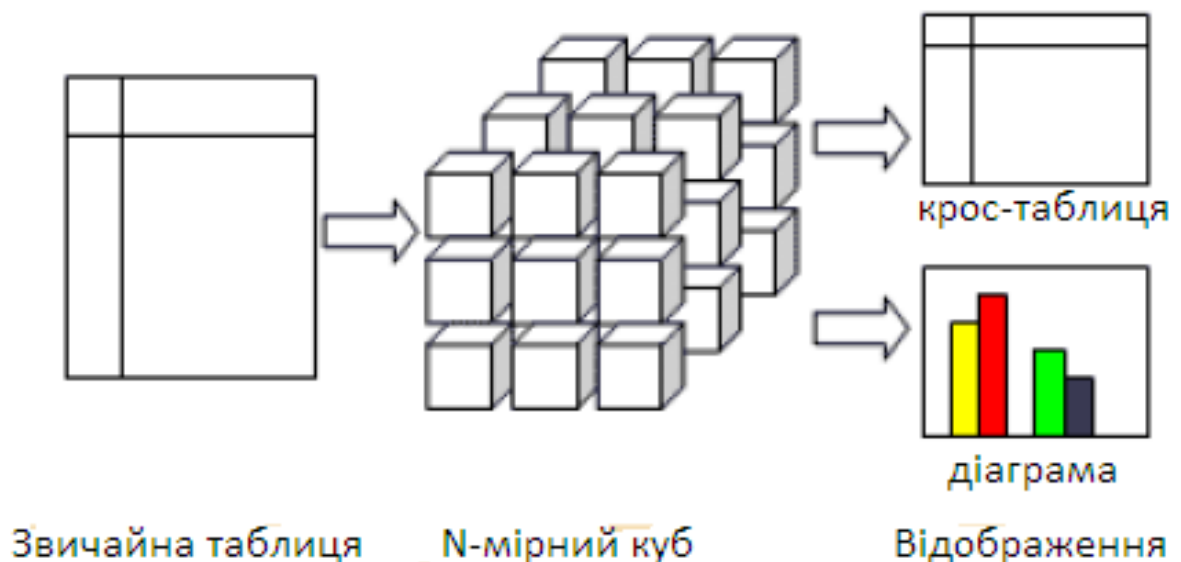


Рисунок 2.2 – Технологія OLAP

Відображення, використані в OLAP-системах, найчастіше бувають двох видів – крос-таблиці і крос-діаграми. Крос-таблиця є основним і найбільш поширеним способом відображення куба. Вона відрізняється від звичайної плоскої таблиці наявністю декількох рівнів вкладеності (наприклад, вона допускає розбиття рядків на підрядки, а стовпців – на підстовпці) [20]. Крос-діаграма являє собою діаграму заданого типу (гістограму, лінійну діаграму

і т.д.), побудовану на основі крос-таблиці. Основна відмінність крос-діаграми від звичайної діаграми в тому, що вона однозначно відповідає поточному стану куба і при будь-яких його змінах (транспонування, фільтрація за вимірюваннями і т.д.) також синхронно змінюється [37].

Над гіперкубом можуть виконуватися такі операції.

Розріз (Slice) (рис. 2.3) – формування підмножини багатовимірного масиву даних, відповідаючи єдиному значенню одного або декількох елементів виміру, що не входять в цю підмножину [9].

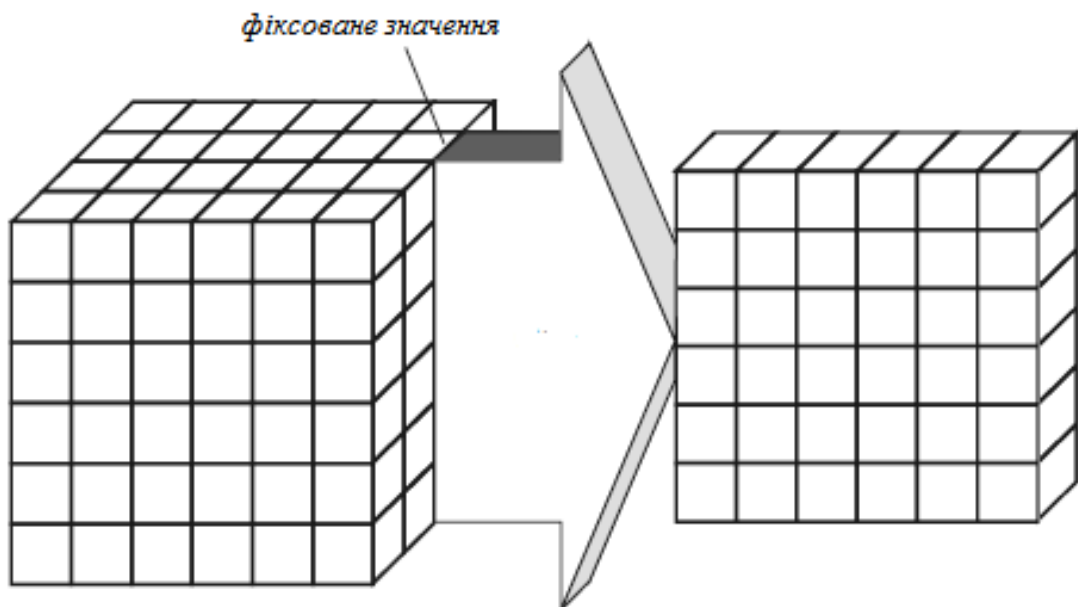


Рисунок 2.3 – Операція «розріз»

Обертання (Rotate) – зміна розташування вимірювань, представлених у звіті або на відтвореній сторінці. Наприклад, операція обертання може полягати в перестановці місцями рядків і стовпців таблиці або переміщенні цікавлять вимірювань в стовпці або рядки сформованого звіту, що дозволяє надавати йому бажаний вигляд [9].

Консолідація (Drill Up) і деталізація (Drill Down) (рис. 2.4) – операції, які визначають перехід вгору у напрямку від детального (down) представлення даних до агрегованого (up) і навпаки, відповідно. Напрямок деталізації

(узагальнення) може бути задано як по ієрархії окремих вимірювань, так і згідно іншим відносинам, встановленими в рамках вимірювань або між вимірами [9].

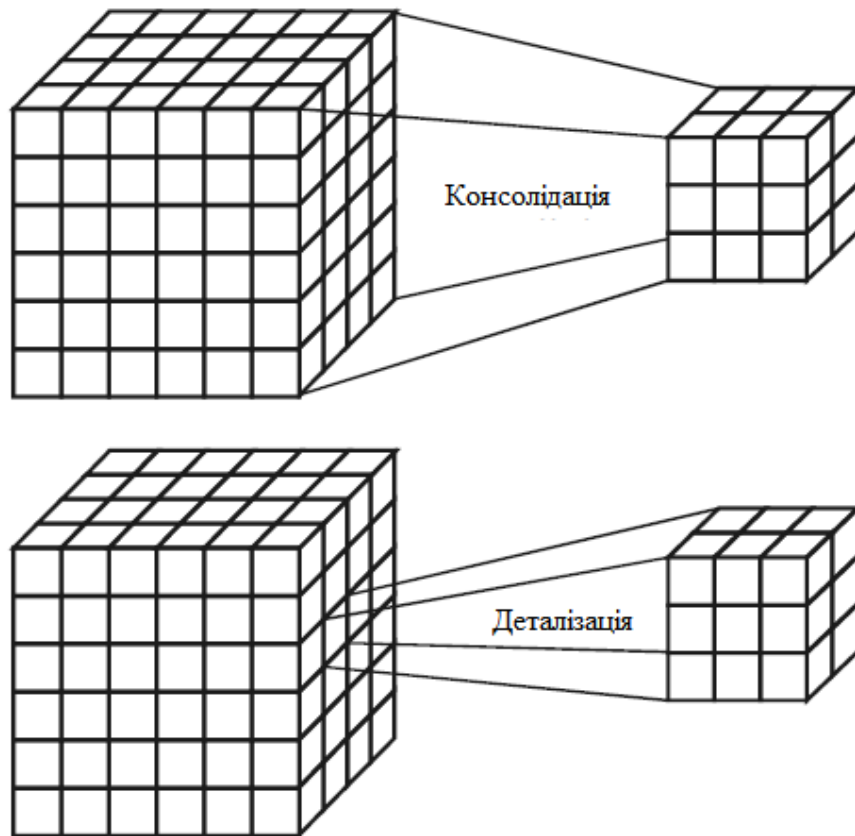


Рис. 2.4 – Операції «консолідація» та «деталізація»

Застосування OLAP-технологій дає можливість побудувати прогнози на основі узагальнених агрегованих даних і таким чином отримати загальну картину або дозволяє провести більш детальне опрацювання питання, побудувавши моделі прогнозування і обчисливши значення прогнозів для елементів нижнього рівня.

2.2 Основні поняття технологій машинного навчання

Поява і розвиток нейронних мереж в багатому сприяли досягнення медицини і біології в області знань про функціонування нервової системи живих істот. В наш час з нейронними мережами зв'язана ціла галузь знань – нейроінформатика. Це міждисциплінарна сфера яка об'єднує в собі елементи біокібернетики, електроніки, статистики, прикладної математики і навіть біомедицини [11].

Штучна нейронна мережа – це обчислювальна модель людського мозку. Кожен нейрон окремо виконує тільки прості перетворення, но паралельна робота великої кількості нейронів і величезна кількість зв'язків між ними приводить до того, що нейронна мережа в цілому може виконувати дуже складні перетворення сигналів у реальному часі. Крім того, велика кількість зв'язків робить нейронну мережу стійкою до помилок, які виникають в окремих зв'язках – робота всієї мережі не зазнає істотних змін.

Візуально нейронна мережа зазвичай представляється у вигляді графів, в яких нейрони – вузли, а зв'язки – лінії. Структурна схема типової нейронної мережі зображена на рис. 2.5 [20].



Рисунок 2.5 – Приклад нейронної мережі

З рисунка видно, що штучний нейрон, так само, як і живий, складається з синапсів, що пов'язують входи нейрона з ядром; ядра нейрона, яке здійснює обробку вхідних сигналів і аксона, який пов'язує нейрон з нейронами наступного шару. Кожен синапс має вагу, який визначає, наскільки відповідний вхід нейрона впливає на його стан. Стан нейрона визначається за формулою:

$$S = \sum_{i=1}^n x_i w_i; \quad (2.1)$$

Очевидно, що функціонування нейронної мережі, тобто дії, які вона здатна виконувати, залежить від величин синаптичних зв'язків. Тому, задавшись структурою нейронної мережі, що відповідає певній задачі, розробник повинен знайти оптимальні значення для всіх вагових коефіцієнтів.

Цей етап називається навчанням нейронної мережі, і від того, наскільки якісно він буде виконаний, залежить здатність мережі вирішувати під час експлуатації поставлені перед нею проблеми. Найважливішими параметрами навчання є: якість підбору вагових коефіцієнтів і час, який необхідно витратити на навчання. Як правило, два цих параметра пов'язані між собою зворотною залежністю і їх доводиться вибирати на основі компромісу [11].

В даний час всі алгоритми навчання нейронних мереж можна розділити на два великі класи: з учителем і без вчителя.

Навчання з вчителем нейронної мережі пред'являються значення як вхідних, так і вихідних параметрів, і вона по деякому внутрішньому алгоритму підлаштовує ваги своїх синаптичних зв'язків [2].

Навчання з учителем передбачає, що для кожного вхідного вектору існує цільовий вектор, що представляє собою необхідний вихід. Разом вони називаються представницькою або навчальною вибіркою. Зазвичай нейронна мережа навчається на деякому числі таких вибірок. Пред'являється вихідний вектор, обчислюється вихід нейронної мережі і порівнюється з відповідним цільовим вектором, різницю (помилка) з допомогою зворотного зв'язку

подається в нейронну мережу, і ваги змінюються відповідно до алгоритму, які прагнуть мінімізувати помилку. Вектори навчальної множини пред'являються послідовно, обчислюються помилки і ваги підлаштовуються для кожного вектору до тих пір, поки помилка по всьому навчальному масиву не досягне прийняттого рівня.

Навчання без вчителя нейронної мережі пред'являються тільки вхідні сигнали, а виходи мережі формуються самостійно з урахуванням тільки вхідних і похідних від них сигналів [2].

Незважаючи на численні прикладні досягнення, навчання з учителем критикувалося за свою біологічну неправдоподібність. Важко уявити навчальний механізм в природному людському інтелекті, який би порівнював бажані і дійсні значення виходів, виконуючи корекцію за допомогою зворотного зв'язку. Якщо допустити подібний механізм в людському мозку, то звідки тоді виникають бажані виходи? Навчання без вчителя є більш правдоподібною моделлю навчання в біологічній системі. Розвинене Кохоненом і багатьма іншими, воно не потребує цільового вектору для виходів і, отже, не вимагає порівняння із зумовленими ідеальними відповідями [11].

Навчальна множина складається лише з вхідних векторів. Навчальний алгоритм підлаштовує ваги нейронної мережі так, щоб виходили узгоджені вихідні вектори, щоб значення досить близьких вхідних векторів давало однакові виходи. Процес навчання виділяє статистичні властивості навчальної множини і групує подібні вектори в класи. Пред'явлення на вхід вектору з даного класу дасть певний вихідний вектор, але до навчання неможливо передбачити, який вихід буде проводитися даним класом вхідних векторів [11].

Отже, виходи подібної мережі повинні трансформуватися в деяку зрозумілу форму, зумовлену процесом навчання. Це не є серйозною проблемою. Зазвичай не складно ідентифікувати зв'язок між входом і виходом, встановленою мережею.

Основні властивості нейронних мереж, які роблять їх привабливими для вирішення багатьох задач обробки інформації і сигналів:

- Нелінійність. Нейронна мережа в цілому є нелінійною системою. Дана властивість особлива важлива для моделювання процесів, які нелінійні по своїй природі.
- Навчання на прикладах. Нейронна мережа здатна змінювати вагу зв'язків за допомогою набору навчальних прикладів. Результатом навчання є налаштування параметрів мережі таким чином, що вони в неявному вигляді зберігають інформацію про об'єкти, рішення задачі і предметної області.
- Паралельна обробка даних. Поступаючи сигнали розмежовуються, що значно пришвидшує обробку інформації.
- Адаптивність. Нейронна мережа здатна адаптувати вагу зв'язків між нейронами до змін у навколишньому середовищі. Частково мережа, раніше навчена для вирішення одної задачі, може бути адаптована (перенавчена) для вирішення іншої задачі.
- Відмова стійкість. При великій кількості зв'язків між нейронами втрата чи спотворення даних в окремих зв'язках не призведе до істотного погіршення якості роботи мережі в цілому [34].

Машинне навчання (англ. machine learning, ML) – це підгалузь інформатики, яка еволюціонувала з дослідження розпізнавання образів та теорії обчислювального навчання в галузі штучного інтелекту.

Машинне навчання досліджує вивчення та побудову алгоритмів, які можуть навчатися з даних, і виконувати передбачувальний аналіз на них. Такі алгоритми діють шляхом побудови моделі зі зразкового тренувального набору вхідних спостережень, щоб створювати керовані даними прогнози або ухвалювати рішення, виражені як виходи [4], замість того, щоб суворо дотримуватись статичних програмних інструкцій [2].

Машинне навчання знаходиться на стику математичної статистики, методів оптимізації та класичних математичних дисциплін, але має також і власну специфіку, пов'язану з проблемами обчислювальної ефективності та перенавчання.

Бізнес-задачі, які можуть бути розв'язані засобами машинного навчання і нейронних мереж:

- Прогнозування попиту, обсягу продажів, наповнення складу, завантаження устаткування і інших ресурсів, подальшого розвитку підприємства і т.д.

- Виявлення тенденцій, прихованих взаємозв'язків, аномалій, повторюваних елементів і т.д.

- Автоматизація роботи операторів в онлайн-чатах, телефонних операторів і т.д.

- Класифікація: аналіз складу покупців, клієнтів, замовників і сегментація їх за різними параметрами [3].

Переваги та недоліки.

Існує багато прикладів успішного використання машинного навчання, починаючи від прогнозування поведінки клієнтів, до встановлення операційних систем для самокерованих автомобілів. Але те, що деякі галузі отримали переваги, не означає, що машинне навчання не має своїх мінусів.

Що стосується переваг, машинне навчання може допомогти підприємствам глибше зрозуміти своїх клієнтів. Збираючи дані про клієнтів та співвідносячи їх поведінку з часом, алгоритми машинного навчання можуть навчитися асоціаціям та допомогти командам адаптувати розробку продуктів та маркетингові ініціативи до попиту споживача [14].

Деякі інтернет-компанії використовують машинне навчання як основний двигун у своїх бізнес-моделях. Наприклад, Uber використовує алгоритми для узгодження водіїв між собою. Google використовує машинне навчання, щоб відображати потрібну рекламу під час пошуку [3].

Але машинне навчання має і недоліки. Перш за все, це може бути дорого. Проекти машинного навчання потребують дорогої програмної інфраструктури.

Існує також проблема упередженості машинного навчання. Алгоритми, навчені на наборах даних, що виключають певні сукупності або містять помилки, можуть призвести до неточних кінцевих моделей, які в кращому випадку дають збій, а в гіршому – дискримінаційні. Коли підприємство базує основні бізнес-процеси на упереджених моделях, це може нанести шкоду репутації [3].

Методи машинного навчання.

Нейронні мережі імітують структуру головного мозку: кожен штучний нейрон з'єднується з декількома іншими нейронами. Нейромережі мають багатошарову структуру: нейрони на одному шарі передають дані декільком нейронам на наступному і т. д. У кінцевому рахунку дані досягають вихідного шару, де мережа видає припущення про те, як вирішити задачу, класифікувати об'єкт і т. д. Нейромережі застосовуються в цілому ряді галузей. В охороні здоров'я їх використовують при аналізі медичних знімків з метою прискорення діагностичних процедур і пошуку ліків. В телекомунікаційної галузі і медіа індустрії нейромережі можна застосовувати для машинного перекладу, розпізнавання шахрайства та надання послуг віртуальних асистентів. У фінансовій галузі їх використовують для розпізнавання шахрайства, управління портфелями і аналізу ризику. У роздрібній торгівлі – для позбавлення від черг в касу і для персоналізації обслуговування покупців [30].

Дерево рішень. Алгоритм дерева рішень класифікує об'єкти, відповідаючи на «питання» про їх атрибути, розташовані у вузлових точках. Залежно від відповіді вибирається одна з гілок, і так до тих пір, поки не буде досягнуто «лист» – остаточну відповідь. Серед застосувань дерева рішень – платформи управління знаннями для клієнтського обслуговування, прогнозного призначення цін і планування випуску продукції [10].

«Випадковий ліс». Це універсальний механізм, який швидко навчається, для виявлення зв'язків всередині набору даних. У приклад можна привести небажані масові розсилки, що створюють проблеми не тільки користувачам, але і провайдером Інтернету, яким через спам доводиться мати справу з підвищеним навантаженням на сервери. Для боротьби з проблемою були розроблені автоматизовані методи фільтрації спаму, які за допомогою ансамблю вирішальних дерев швидко і ефективно визначають небажані листи [14].

Кластеризація – це групування елементів даних, які мають подібні характеристики, за допомогою статистичних алгоритмів. Це метод навчання без вчителя, який можна використовувати для вирішення завдань класифікації. Кластеризація також дієва, коли в складних наборах даних потрібно виявити групи, які важко помітити без спеціальних засобів. Приклади – від групування схожих документів в базі даних до виявлення по кримінальним новин територій з підвищеним рівнем злочинності [10].

Пошук асоціативних правил – це метод навчання без вчителя, що дозволяє знаходити відносини між змінними. Використовується в двигунах видачі рекомендацій – саме цей метод застосовується в багатьох інтернет-магазинах для складання фрази «Разом з цим товаром зазвичай купують ...».

Конкретний приклад – підвищення продажів в магазині делікатесів. Вивчивши купівельну поведінку шляхом пошуку асоціативних правил, можна пропонувати спеціальну упаковку і набори для свят і інших особливих випадків. Асоціативні правила дозволяють з'ясувати, коли і за яких обставин покупці купують ті чи інші поєднання товарів. Використовуючи відомості про минулі покупки і часу їх скоєння, можна скласти програму знижок і сформувані індивідуальні пропозиції з розрахунком на підвищення продажів [30].

Глибоке вивчення (ГН) – це такий тип машинного навчання, в якому модель навчається виконувати завдання класифікації безпосередньо з зображень, тексту або звуку. Глибоке вивчення, зазвичай, здійснюється за

допомогою архітектури глибокої штучної нейронної мережі. Термін «глибоко» означає кількість шарів в мережі – чим більше шарів, тим глибша мережа. Традиційні нейронні мережі містять тільки 2 або 3 шари, в той час як глибока мережа може мати їх сотні [14].

Глибоке навчання швидко зростаюче застосування машинного навчання. Використання цих алгоритмів в різних областях показує його універсальність [12].

Крім того, важливо відзначити, що ієрархія рівнів і контроль в навчанні є ключовими факторами для розробки успішного застосування щодо глибокого навчання. Ієрархія важлива для відповідної класифікації даних, в той час як контроль враховує важливість самої бази даних як частини процесу. Основна цінність глибокого навчання полягає в оптимізації існуючих додатків в машинному навчанні завдяки інноваційності ієрархічної обробки. Глибоке навчання може забезпечити ефективні результати при цифровій обробці зображень і розпізнаванні мови. Зниження відсотка помилок (від 10 до 20%) явно підтверджує поліпшення в порівнянні з існуючими та перевіреними методами.

У нинішню епоху і в майбутньому глибоке навчання може стати корисним інструментом безпеки завдяки поєднанню розпізнавання осіб і мови. Крім цього, цифрова обробка зображень є областю досліджень, яка може застосовуватися в безлічі інших областей. З цієї причини і показавши справжню оптимізацію, глибоке навчання є сучасним і цікавим предметом розвитку штучного інтелекту [13].

2.3 Прогнозування часових рядів за допомогою штучних нейронних мереж

Для моделювання ЧР (часових рядів) зараз існує багато методів, що само по собі ставить проблему вибору моделі ЧР для комп'ютерів. Чим більше методів вирішення задачі, тим складніше вибрати один з них, оскільки кожен з них має свої переваги і недоліки [16].

В задачах прогнозування ЧР, на відміну від аналізу випадкових вибірок дані розглядаються як послідовність вимірів, впорядковані у невідповідні моменти часу, тобто ЧР включає два обов'язкових елемента – час і конкретне значення показника.

Для побудови часового ряду мають бути виконані наступні умови [26]:

- Періодичність розвитку утворюється розчленуванням в часі на однорідні етапи в межах яких показники підпорядковуються загальному закону розвитку.
- Значення ЧР мають бути співставленні по всім ознакам за якими здійснюється його формування, наприклад по території, коло охоплених явищ, ціні і т.д.
- Періоди мають відповідати інтенсивності процесу.
- ЧР має бути повним, тобто не допускаються пропуски, якщо вони неминучі ряд доповнюють умовно-розрахунковими значеннями.

Тільки у випадку виконання перерахованих умов можна говорити, що послідовність яку розглядаємо є часовим рядом і для неї використані методи аналізу і прогнозування часових рядів [26].

У ситуації прогресуючої глобалізації та швидких змін на ринку багато економічних процесів характеризуються великою динамікою та мінливістю в часі. Часові ряди, що описують дані явища, у багатьох випадках можуть здаватися хаотичними. Відповідно, існує потреба у розробці рішень для

прогнозування, зокрема споживчого попиту, що забезпечить необхідну точність моделей.

Для моделювання попиту, прогнозування параметрів місткості ринку продукції конкретного товаровиробника та відтворення моделей поведінки споживачів найперспективнішими є нейромережеві моделі. Перевагами цих моделей є можливість моделювання нелінійних процесів, масштабованість, висока адаптивність, різноманітність сфер застосування [22, 16].

Часові ряди, що показують еволюцію економічних змінних, як правило, є результатом поєднання багатьох компонентів різної природи. При розв'язанні задачі прогнозування аналітику доводиться приймати рішення щодо таких характеристик часового ряду як тренд, сезонна і циклічна компоненти, робити припущення про модель часового ряду – адитивна, мультиплікативна та ін. Автоматичного способу виявлення трендів у часових рядах не існує. У той же час при вивченні кривої, що відображає результати спостережень, аналітику важко робити припущення щодо повторюваності форми кривої через рівні проміжки часу [33].

Загальним недоліком статистичних моделей є складність вибору типу моделі і підбору її параметрів. Все це істотно збільшує суб'єктивність при аналізі і прогнозуванні часових рядів. Таким чином, результат аналізу і прогнозування часового ряду залежить як від кваліфікації аналітика в предметній галузі, так і від його кваліфікації в методах аналізу [33].

Апарат нейронних мереж передбачає мінімальну участь аналітика в формуванні моделі часового ряду, оскільки нейронні мережі дозволяють будувати моделі, що відображають складні взаємозв'язки між вхідними та вихідними даними для явищ, структура, закони дії чи причинно-наслідкові зв'язки яких були недостатньо відомими для побудови ефективних математичних моделей. Мережі здатні виявляти взаємозв'язки в наборі даних, які неможливо виявити традиційними статистичними методами [34].

Як зазначає Заверач М.М.: «Важливою перевагою нейронних мереж є їх гнучка структура. Для зміни структури у рамках визначеної архітектури

нейронної мережі достатньо регулювати кількість шарів та нейронів, додаткові переваги надає також можливість зміни активаційної функції. Такі перетворення надають можливість повністю змінити структуру мережі, що дозволяє максимально пристосувати обрану архітектуру і мінімізувати похибку навчання мережі та підвищити точність прогнозування» [17].

Часовий ряд це послідовність упорядкованих у часі числових показників, що характеризують розвиток досліджуваного явища або процесу.

Моделювання часових рядів з використанням апарату нейронних мереж полягає в формуванні нейронної мережі певної структури, яка описує поведінку досліджуваного процесу в моменти часу, а прогнозування полягає в передбаченні майбутньої поведінки процесу за передісторією. Під навчанням нейромережевої моделі розуміють цілеспрямований процес зміни значень вагових коефіцієнтів, повторюваний до тих пір, поки мережа не придбає необхідні властивості. Зміна вагових коефіцієнтів відбувається на підставі зміни фактичної похибки прогнозування на ітераціях.

Процес прогнозування за допомогою нейронних мереж складається з наступних етапів [43]:

1. Підготовка вхідних даних. Готуються дані з продажу за попередній період.
2. Створення навчальної, контрольної та тестової вибірок, необхідних для навчання та перевірки мережі.
3. Формування архітектури нейромережі (вибір кількості шарів і кількості нейронів в кожному шарі) (варто використати кілька типів мережі, а потім відібрати найкращу).
4. Вибір алгоритму навчання (зворотне поширення, Левенберга-Маркара, швидке поширення, delta-bar-delta).
5. Навчання декількох нейронних мереж з різною архітектурою.
6. Розрахунок помилки.

7. Вибір найкращої мережі (з декількох мереж вибирається мережа з найменшою помилкою). Крос-перевірка – незалежна перевірка якості прогнозу є одним із важливих завдань в прогнозуванні [43].

Певною складністю в побудові нейромережевої моделі є формування навчальної вибірки, яка повинна задовольняти вимогам повноти (вибірка не повинна містити пропуски, повинна містити всі допустимі приклади досліджуваного діапазону) і несуперечності (вибірка не повинна містити суперечливих прикладів).

Множина вхідних даних розбивається на дві підмножини, що не перетинаються та хронологічно йдуть одна за одною. Одна з них є навчальною вибіркою, на якій буде виконуватися навчання нейронної мережі. Інша підмножина є контрольною вибіркою, яка не пред'являється нейронній мережі в процесі навчання і використовується для перевірки якості прогнозу. Навчальна та контрольна вибірки співвідносяться як 2:1 [8].

В задачі прогнозування часового ряду необхідно визначити, скільки попередніх значень вхідної змінної взяти та на скільки вперед прогнозувати значення вихідної змінної. Класичний і часто високоефективний підхід у процесі моделювання нейронних часових рядів базується на використанні так званого «ковзного вікна». Метод «ковзного вікна» полягає у визначенні двох вікон з фіксованими розмірами n і m відповідно. Ці вікна переміщуються з деяким кроком по часовій послідовності історичних даних, починаючи з першого елемента, і призначені для доступу до даних часового ряду, причому перше вікно, отримавши такі дані, передає їх на вхід нейронної мережі, а друге – на вихід. Отже, першим кроком є визначення ширини вікна вхідних даних (це також є кількістю вхідних змінних для нейронної мережі), а другий крок – визначення ширини вікна вихідних даних (кількість вихідних змінних), яка може бути більше 1, але на практиці передбачається рівною $m = 1$. Це означає, що одна вихідна змінна вважається прогнозом наступного, найближчого значення часового ряду [8].

Задача прогнозування є окремим випадком побудови регресії, тому вона може бути вирішена наступними типами нейронних мереж: багат шаровим перцептроном (MLP), радіально-базисною мережею (RBF), узагальнено-регресійною мережею (GRNN), мережею Вольтеррі і мережею Ельмана [8].

Архітектура нейронної мережі залежить від поставленого завдання, в більшості випадків найбільш оптимальною архітектурою для прогнозування часових рядів буде багат шаровий перцептрон зі зворотним поширенням помилки.

Вибір найбільш оптимальної моделі прогнозування визначається за допомогою порівнянь розрахункових помилок прогнозу:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|y(t) - \hat{y}(t)|}{y(t)} * 100\% ; \quad (2.2)$$

$$MAE = \frac{1}{N} \sum_t |y(t) - \hat{y}(t)|; \quad (2.3)$$

де MAPE (*Mean Absolute Percentage Error*) – середня абсолютна помилка у відсотках і MAE (*Mean Absolute Error*) – середня абсолютна помилка. Критерієм вибору моделі є мінімальне значення розрахункової помилки прогнозу [10].

Для якісного прогнозування попиту необхідно обробляти величезні масиви даних. Тому для зменшення людського фактора і отримання більш точного прогнозу попиту необхідно використовувати спеціальні програмні рішення для прогнозування.

А. Долгіх та О. Байбуз зазначають, що програмні продукти для прогнозування часових рядів «можна поділити на такі групи: статистичні пакети, які мають велику кількість реалізованих функцій і методів з аналізу даних, такі пакети містять модуль з прогнозування часових рядів (наприклад, R, MATLAB, STATISTICA, NCSS, Stata, SAS) і пакети, спеціалізовані на

прогнозуванні часових рядів, такі, як Zaitun Time Series, Forecast Pro, Sales Forecast» [16].

На сьогоднішній день існує ряд програмних продуктів, призначених для моделювання роботи нейронних мереж. Існує кілька десятків фірм виробників нейропакетів, а асортимент, який вони пропонують, становить кілька сотень найменувань. Повністю універсальних нейропакетів не існує, оскільки неможливо заздалегідь передбачити всі нейронні структури, які можуть знадобитися для рішення тих або інших прикладних задач [20]. Найбільш відомі та розповсюджені [20]:

1. NeuroSolutions – фірми NeuroDimension, Inc.;
2. Statistica з модулем Neural Networks – фірми StatSoft.;
3. Deductor – фірми BaseGroup.;
4. NeuroShell2 – фірми Ward Systems Group;
5. BrainMaker Pro – фірми California Scientific Software.

3 ПРАКТИЧНІ РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

3.1. Аналіз попиту засобами OLAP

На виробництвах з різноманітним асортиментом для якісного прогнозування попиту необхідно обробляти величезні масиви даних. Тому для отримання більш точного прогнозу попиту, як зазначалось в попередньому розділі, необхідно використовувати спеціальні програмні рішення для прогнозування. Нами було використано аналітичну платформу Deductor Academic Studio 5.3. Deductor – це платформа, що дозволяє проводити всебічний аналіз даних підприємства, прогнозувати показники його розвитку, проводити сегментацію і пошук закономірностей, тобто дозволяє виконати всі етапи бізнес-аналізу від збору даних і побудови моделей до візуалізації та інтеграції в бізнес-процес [6].

Дослідження проводилось у співпраці з компанією ТМ «Мамин хліб». Здійснювався аналіз та прогнозування попиту на продукцію за товарними групами по трьом регіонам Хмельницької області. Вхідними даними для процесу прогнозування споживчого попиту є дані про продажі за минулі періоди (історичні дані), а саме дані про щоденні продажі за пів року по товарним групам. Продукція, вироблена на підприємстві ТМ «Мамин хліб», поділяється на такі товарні групи:

- булочні вироби,
- кондитерські вироби,
- листові вироби,
- напівфабрикати,
- хлібні вироби.

Першим кроком до побудови прогнозної моделі споживчого попиту є попередній аналіз вхідних даних. Зручним інструментом швидкого аналізу

великих обсягів даних і наочного відображення результатів у вигляді малюнків, графіків, діаграм і таблиць є засоби OLAP.

Аналітична платформа Deductor дає можливість проведення OLAP-аналізу. Багатовимірний куб даних по продажам можна розглядати як систему координат, осями якої є виміри, в нашому випадку дата, регіон, товарна група (рис. 3.1).

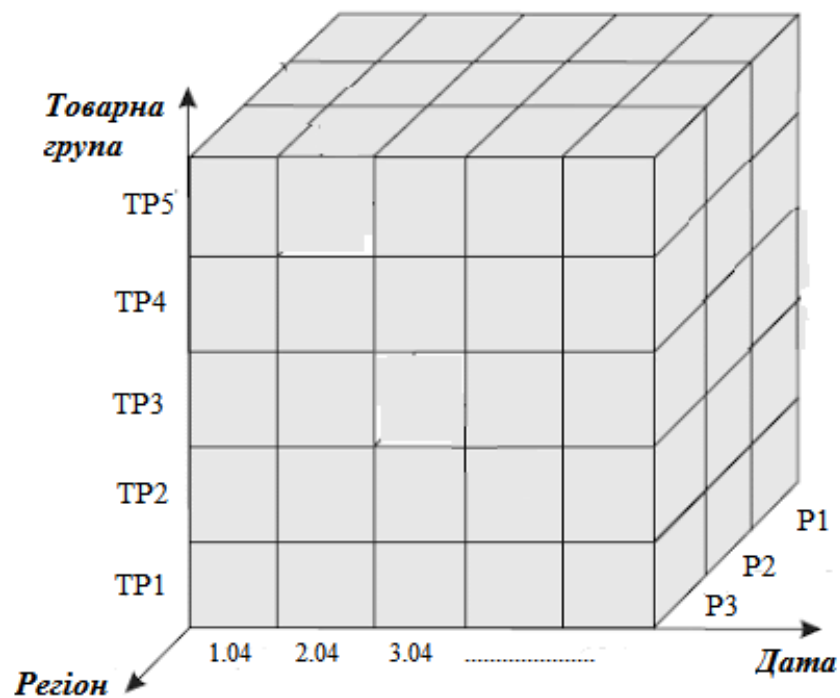


Рисунок 3.1 – OLAP-куб

Візуалізація OLAP-куба проводиться за допомогою спеціального виду таблиць – крос-таблиці, які будуються на основі зрізів OLAP-куба, що містять необхідну користувачу інформацію. Зрізи є результатом виконання відповідного запиту до бази даних. Крос-таблиці відображають багатомірні дані у вигляді двомірної таблиці. Дані таблиці зручні тим, що дають можливість швидко довільним чином групувати, фільтрувати дані, переставляти стовпці та рядки таблиці з метою побудови необхідних аналітичних звітів.

На рисунку 3.2 представлено табличне зображення OLAP-куба.

		Регион			
- + Дата	Товарна група	P1	P2	P3	Итого:
01.04.2020		54 494,00	55 291,00	38 796,00	148 581,00
02.04.2020		57 749,00	46 848,00	45 170,00	149 767,00
03.04.2020		48 125,00	43 013,00	59 350,00	150 488,00
04.04.2020		52 893,00	47 730,00	48 626,00	149 249,00
05.04.2020		46 372,00	43 981,00	60 261,00	150 614,00
06.04.2020		46 920,00	52 247,00	49 937,00	149 104,00
07.04.2020		43 440,00	43 997,00	64 582,00	152 019,00
08.04.2020		52 972,00	53 232,00	44 553,00	150 757,00
09.04.2020		45 752,00	53 600,00	48 697,00	148 049,00
10.04.2020		42 093,00	47 618,00	59 483,00	149 194,00
11.04.2020		54 141,00	51 920,00	43 554,00	149 615,00
12.04.2020		52 505,00	56 805,00	40 817,00	150 127,00
13.04.2020		50 335,00	46 421,00	55 267,00	152 023,00
14.04.2020		54 644,00	51 180,00	43 946,00	149 770,00
15.04.2020		48 973,00	49 492,00	48 945,00	147 410,00
16.04.2020		40 506,00	50 018,00	59 512,00	150 036,00
17.04.2020		45 118,00	55 261,00	48 001,00	148 380,00
18.04.2020		44 222,00	51 345,00	52 898,00	148 465,00
19.04.2020		40 022,00	47 136,00	60 995,00	148 153,00
20.04.2020		51 809,00	50 423,00	47 154,00	149 386,00
21.04.2020		56 071,00	56 811,00	35 270,00	148 152,00
22.04.2020		47 272,00	49 453,00	52 109,00	148 834,00
23.04.2020		45 656,00	51 196,00	50 483,00	147 335,00
24.04.2020		61 720,00	56 447,00	34 186,00	152 353,00
25.04.2020		52 321,00	53 159,00	45 808,00	151 288,00
26.04.2020		47 680,00	48 445,00	54 461,00	150 586,00

Рисунок 3.2 – Фрагмент OLAP-куба

Виконуючи розрізи куба за значеннями різних вимірів, було проведено аналіз продажів компанії за різними товарними групами та окремо для кожного регіону.

Так на рисунку 3.3 представлено крос-таблиці продажів за товарними групами «булочні вироби» та «лишкові вироби».

Таблиця X Куб X					Таблиця X Куб X						
Регион					Регион						
Дата	Товарна група	P1	P2	P3	Итого:	Дата	Товарна група	P1	P2	P3	Итого:
01.04.2020	булочні вироби	1 688,00	1 642,00	1 231,00	4 561,00	01.04.2020	листкові вироби	8 343,00	10 589,00	13 156,00	32 088,00
02.04.2020	булочні вироби	1 934,00	1 844,00	720,00	4 498,00	02.04.2020	листкові вироби	10 935,00	9 327,00	11 900,00	32 162,00
03.04.2020	булочні вироби	1 747,00	1 425,00	1 425,00	4 597,00	03.04.2020	листкові вироби	14 595,00	9 843,00	9 504,00	33 942,00
04.04.2020	булочні вироби	1 870,00	1 309,00	1 496,00	4 675,00	04.04.2020	листкові вироби	9 880,00	11 924,00	12 265,00	34 069,00
05.04.2020	булочні вироби	1 157,00	1 024,00	2 270,00	4 451,00	05.04.2020	листкові вироби	14 303,00	14 643,00	5 108,00	34 054,00
06.04.2020	булочні вироби	1 937,00	1 153,00	1 522,00	4 612,00	06.04.2020	листкові вироби	12 005,00	9 337,00	12 005,00	33 347,00
07.04.2020	булочні вироби	1 543,00	1 350,00	1 929,00	4 822,00	07.04.2020	листкові вироби	10 636,00	7 645,00	14 958,00	33 239,00
08.04.2020	булочні вироби	1 144,00	1 244,00	2 587,00	4 975,00	08.04.2020	листкові вироби	14 442,00	12 762,00	6 381,00	33 585,00
09.04.2020	булочні вироби	1 407,00	1 117,00	1 613,00	4 137,00	09.04.2020	листкові вироби	8 626,00	13 418,00	9 904,00	31 948,00
10.04.2020	булочні вироби	1 240,00	1 796,00	1 240,00	4 276,00	10.04.2020	листкові вироби	7 528,00	13 420,00	11 783,00	32 731,00
11.04.2020	булочні вироби	1 048,00	1 135,00	2 183,00	4 366,00	11.04.2020	листкові вироби	12 693,00	13 344,00	6 509,00	32 546,00
12.04.2020	булочні вироби	1 123,00	1 167,00	2 200,00	4 490,00	12.04.2020	листкові вироби	12 868,00	13 198,00	6 929,00	32 995,00
13.04.2020	булочні вироби	1 400,00	1 400,00	1 716,00	4 516,00	13.04.2020	листкові вироби	10 858,00	8 143,00	14 930,00	33 931,00
14.04.2020	булочні вироби	1 644,00	1 550,00	1 503,00	4 697,00	14.04.2020	листкові вироби	10 091,00	10 416,00	12 044,00	32 551,00
15.04.2020	булочні вироби	1 325,00	2 035,00	1 372,00	4 732,00	15.04.2020	листкові вироби	8 156,00	13 050,00	11 418,00	32 624,00
16.04.2020	булочні вироби	1 112,00	2 079,00	1 644,00	4 835,00	16.04.2020	листкові вироби	8 214,00	8 871,00	15 770,00	32 855,00
17.04.2020	булочні вироби	1 673,00	1 673,00	1 574,00	4 920,00	17.04.2020	листкові вироби	7 775,00	9 071,00	15 550,00	32 396,00
18.04.2020	булочні вироби	1 780,00	1 684,00	1 347,00	4 811,00	18.04.2020	листкові вироби	9 066,00	11 979,00	11 332,00	32 377,00
19.04.2020	булочні вироби	1 424,00	1 708,00	1 613,00	4 745,00	19.04.2020	листкові вироби	9 034,00	9 357,00	13 874,00	32 265,00
20.04.2020	булочні вироби	1 251,00	1 204,00	2 177,00	4 632,00	20.04.2020	листкові вироби	8 843,00	13 944,00	11 223,00	34 010,00
21.04.2020	булочні вироби	1 598,00	1 686,00	1 154,00	4 438,00	21.04.2020	листкові вироби	13 182,00	12 860,00	6 109,00	32 151,00
22.04.2020	булочні вироби	1 091,00	1 679,00	1 427,00	4 197,00	22.04.2020	листкові вироби	9 572,00	13 719,00	8 614,00	31 905,00
23.04.2020	булочні вироби	1 825,00	1 477,00	1 043,00	4 345,00	23.04.2020	листкові вироби	11 123,00	12 077,00	8 581,00	31 781,00
24.04.2020	булочні вироби	1 640,00	1 265,00	1 781,00	4 686,00	24.04.2020	листкові вироби	14 369,00	12 364,00	6 683,00	33 416,00
25.04.2020	булочні вироби	1 387,00	1 430,00	1 517,00	4 334,00	25.04.2020	листкові вироби	13 191,00	13 521,00	6 266,00	32 978,00
26.04.2020	булочні вироби	1 253,00	1 253,00	2 134,00	4 640,00	26.04.2020	листкові вироби	14 669,00	12 964,00	6 482,00	34 115,00

Рисунок 3.3 – Розрізи OLAP-куба за значенням виміру «товарна група» (фрагменти)

Виконання розрізів OLAP-куба за значенням виміру «регіон» дає можливість проаналізувати продажі окремо для кожного регіону. Приклад такого розрізу для регіону P2 представлено на рисунку 3.4. Розрізаючи куб по виміру дата та виконуючи селективний відбір, можна окремо проаналізувати продажі за конкретний день або за декілька днів. Окрім крос-таблиць аналітична платформа одночасно будує крос-діаграми, що також допомагає візуалізувати результати аналізу даних. Так на рисунку 3.5 представлено крос-таблицю та крос-діаграму для аналізу продажів за 5 квітня 2020 року. А крос-таблиця та крос-діаграма на рисунку 3.6 візуалізують аналіз продажів в першому регіоні за 10 днів квітня.

		Регион	
Дата	Товарна група	P2	Итого:
01.04.2020		55 291,00	55 291,00
02.04.2020	булочні вироби	1 844,00	1 844,00
	кондитерські вироби	4 861,00	4 861,00
	листові вироби	9 327,00	9 327,00
	напівфабрикати	9 511,00	9 511,00
	хлібні вироби	21 305,00	21 305,00
	Итого:		46 848,00
03.04.2020	булочні вироби	1 425,00	1 425,00
	кондитерські вироби	6 539,00	6 539,00
	листові вироби	9 843,00	9 843,00
	напівфабрикати	8 034,00	8 034,00
	хлібні вироби	17 172,00	17 172,00
	Итого:		43 013,00
04.04.2020	булочні вироби	1 309,00	1 309,00
	кондитерські вироби	7 112,00	7 112,00
	листові вироби	11 924,00	11 924,00
	напівфабрикати	6 083,00	6 083,00
	хлібні вироби	21 302,00	21 302,00
	Итого:		47 730,00
05.04.2020	булочні вироби	1 024,00	1 024,00
	кондитерські вироби	5 309,00	5 309,00
	листові вироби	14 643,00	14 643,00
	напівфабрикати	5 842,00	5 842,00
	хлібні вироби	17 163,00	17 163,00
	Итого:		43 981,00

Рисунок 3.4 – Розріз OLAP-куба за значенням виміру «регіон» (фрагмент)

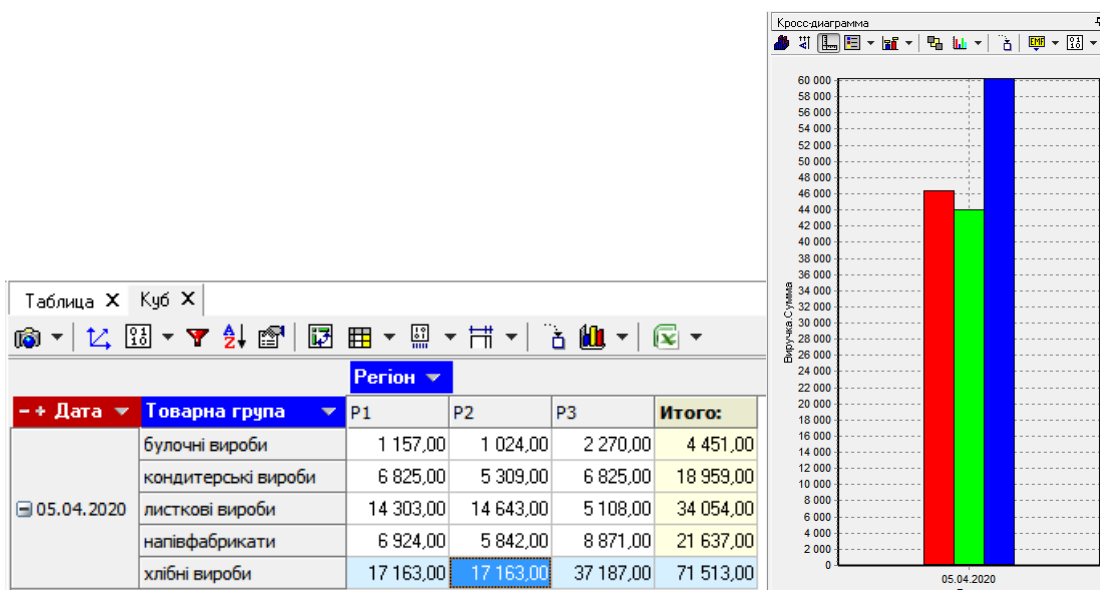


Рисунок 3.5 – Розрізи OLAP-куба за значенням виміру «дата»

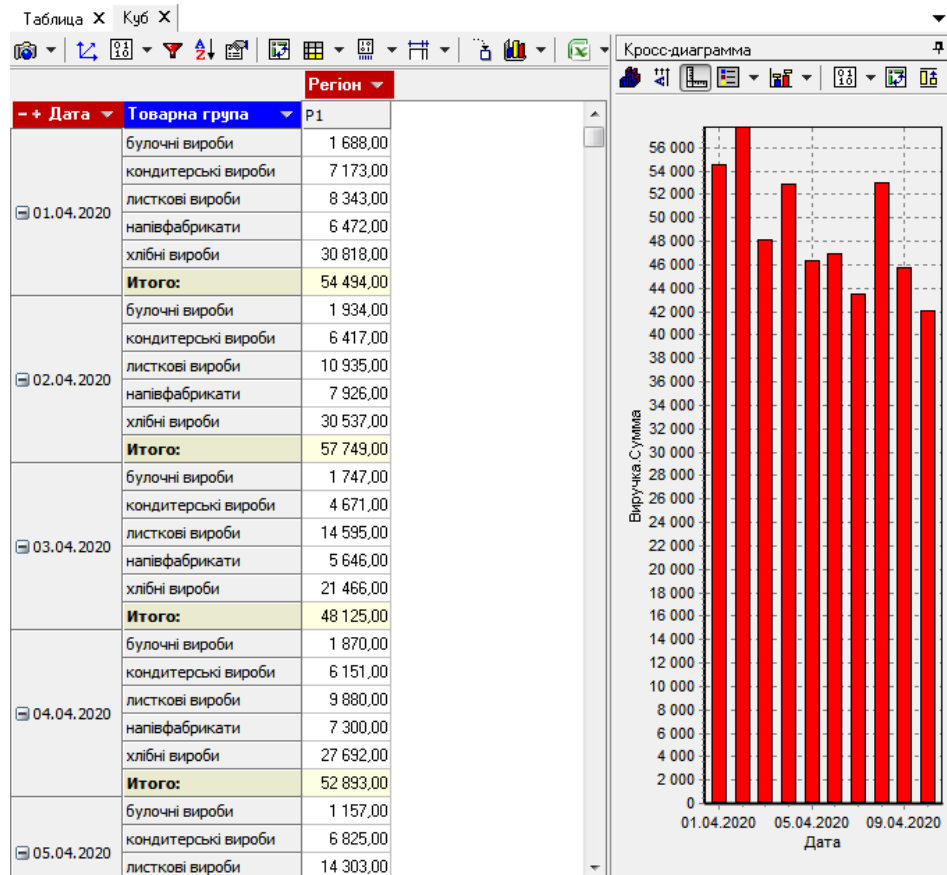
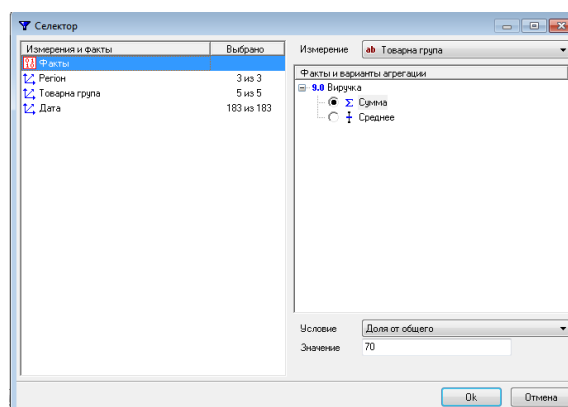


Рисунок 3.6 – Фрагмент звіту з продажів в першому регіоні за 10 днів квітня

Задавши до крос-таблиці селектор з умовою «доля від загального 70», було з'ясовано, що 70 % виручки від продажу забезпечують «хлібні вироби» та «листові вироби» (рис. 3.7).



		Регіон					
		P1		P2		P3	
		Виручка		Виручка		Виручка	
+ Товарна гр...	Дата	Σ Сумма	± Среднее	Σ Сумма	± Среднее	Σ Сумма	± Среднее
листові вироби		1 971 285,00	10 772,05	2 017 828,00	11 026,38	2 025 663,00	11 069,20
хлібні вироби		4 324 354,00	23 630,35	4 350 948,00	23 775,67	4 481 406,00	24 488,56

Рисунок 3.7 – Вироби, що забезпечують 70% виручки

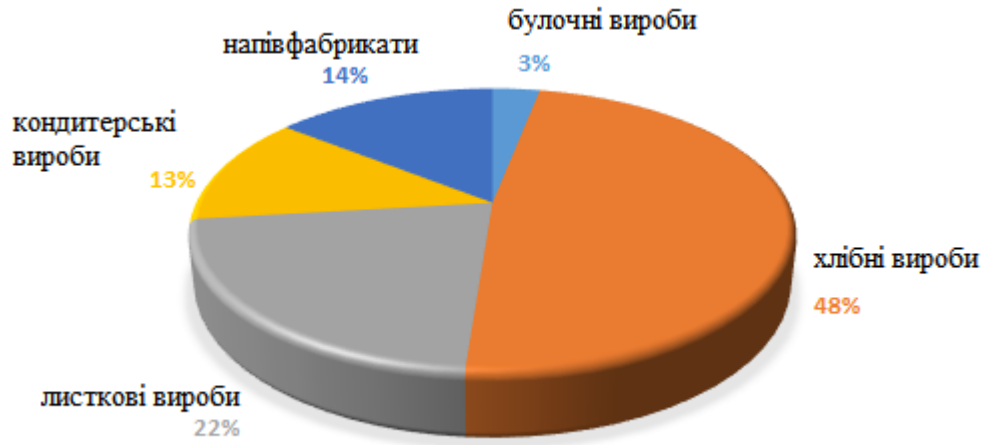


Рисунок 3.8 – Структура обсягів продажу

Як видно з діаграми (рис. 3.8) найбільшим попитом дійсно користуються вироби двох товарних груп: «хлібні вироби» та «лишкові вироби».

Топографічна діаграма представлена на рисунку 3.9 дозволяє проаналізувати рівень попиту за кожною товарною групою для кожного регіону, що досліджуються.

За результатами аналізу топографічної діаграми можемо зазначити, що дійсно «хлібні вироби» в загальному мають найвищий попит, проте самий високий попит на них у третьому регіоні, в той час як в двох інших попит середнього рівня. Достатньо низький попит у всіх регіонах на «булочні вироби». Попит на «кондитерські вироби» та «напівфабрикати» теж низький, проте не дуже відрізняється в різних регіонах. Щодо попиту на «лишкові вироби», то він змінюється від низького до середнього в різних регіонах.

Оскільки, «хлібні вироби» та «лишкові вироби» забезпечують підприємству 70 % виручки, то, на нашу думку, попит на них потребує більш детального аналізу. При тому варто провести аналіз попиту на дані вироби в кожному регіоні, оскільки, як видно з діаграми, він значно різниться в залежності від регіону.

Виконавши відповідні розрізи, було одержано статистичну інформацію для подальшого прогнозування.

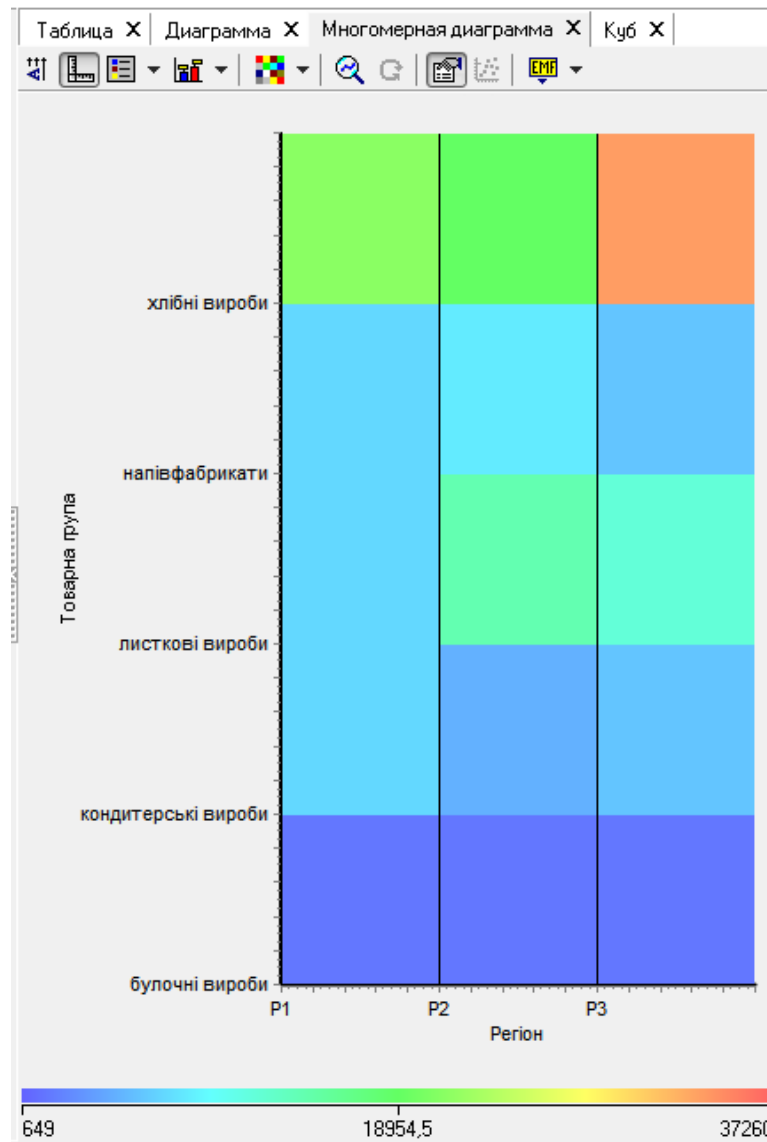


Рисунок 3.9 – Топографічна діаграма обсягів виручки

3.2 Прогнозування попиту за допомогою штучних нейронних мереж

Часові ряди одержані в результаті відповідних розрізів багатомірного кубу даних є вхідними даними для побудови нейронних мереж та, відповідно, прогнозів.

Зазначені ряди містять інформацію про щоденні обсяги продажу товарів за товарними групами та регіонами з 1.04.2020 по 19.09.2020. В якості вхідних

даних моделі прогнозування використовувались дані про обсяги продажу за період з 1.04.2020 по 09.09.2020. Прогноз здійснювався на період з 10.09.2020 по 19.09.2020. Статистичні дані за прогнозований період використовувались для перевірки достовірності прогнозу та розрахунку помилки прогнозування.

Перед побудовою нейронної мережі та прогнозуванням необхідно видалити аномалії, викиди та шуми у вхідних даних. Аналітична платформа Deductor дозволяє зробити це за допомогою, зокрема, спектральної обробки. Наступним кроком є трансформація вхідних даних за допомогою методу «ковзне вікно», після чого будується нейронна мережа. Повний алгоритм побудови прогнозу попиту за допомогою нейронної мережі з використанням аналітичної платформи Deductor представлено на рис. 3.10.

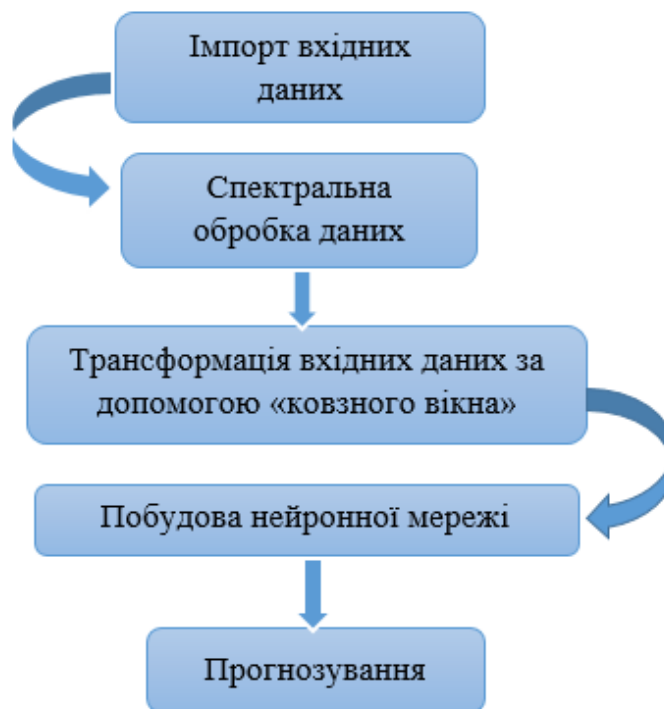


Рисунок 3.10 – Алгоритм побудови прогнозу попиту

Першою для побудови прогнозу попиту розглядалась товарна група «хлібні вироби» для першого регіону. Вхідні дані (фрагмент) для прогнозування представлено на рис. 3.11. Першим кроком є побудова графіка часового ряду на основі вихідних даних (рис.3.12). Аналіз графіка свідчить про

наявність викидів та шумів. Отже, наступний крок – спектральна обробка. Результат якої представлено на рис. 3.13.

	Дата	Регіон	Товарна група	Виручка
▶	01.04.2020	P1	хлібні вироби	30818
	02.04.2020	P1	хлібні вироби	30537
	03.04.2020	P1	хлібні вироби	21466
	04.04.2020	P1	хлібні вироби	27692
	05.04.2020	P1	хлібні вироби	17163
	06.04.2020	P1	хлібні вироби	21472
	07.04.2020	P1	хлібні вироби	16581
	08.04.2020	P1	хлібні вироби	24408
	09.04.2020	P1	хлібні вироби	25730
	10.04.2020	P1	хлібні вироби	17100
	11.04.2020	P1	хлібні вироби	26598
	12.04.2020	P1	хлібні вироби	23860
	13.04.2020	P1	хлібні вироби	23072
	14.04.2020	P1	хлібні вироби	31232
	15.04.2020	P1	хлібні вироби	28526
	16.04.2020	P1	хлібні вироби	20110
	17.04.2020	P1	хлібні вироби	24134

Рисунок 3.11 – Вхідні дані (фрагмент)

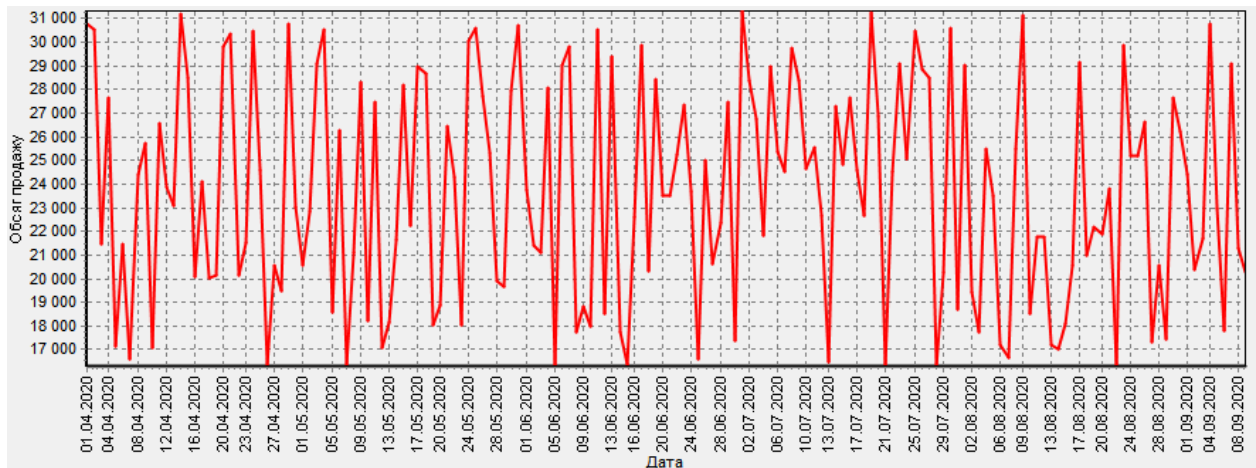


Рисунок 3.12 – Діаграма даних

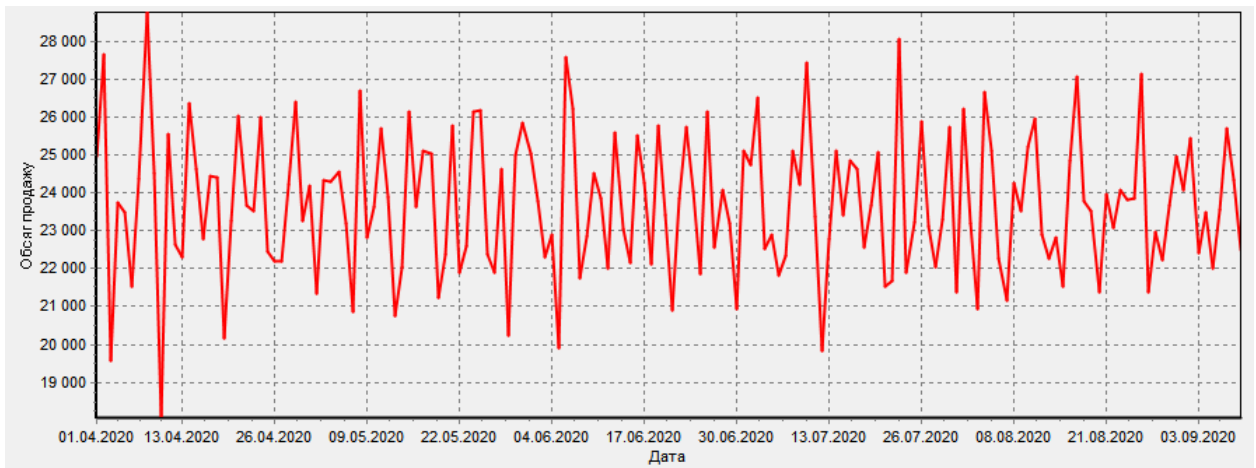


Рисунок 3.13 – Діаграма даних після спектральної обробки

Після спектральної обробки дані було трансформовано за допомогою «ковзного вікна». Основним параметром при побудові «ковзного вікна» є глибина занурення – кількість минулих значень, що потрапляють у вікно.

Для побудови прогностичної моделі необхідно знати три параметри:

- інтервал прогнозу – часовий інтервал, на якому буде здійснюватися прогнозування: день, тиждень, місяць, квартал або рік (в нашому випадку це день);
- горизонт прогнозу – інтервал (кількість днів, тижнів і т. д.), на який ми хочемо отримати прогноз;
- глибина занурення – кількість попередніх значень ряду, які ми будемо використовувати для передбачення майбутніх значень, тобто кількість минулих значень ряду, які потрапляють у вікно.

Для прогнозування обсягів продажів було використано вікно в сім днів, тобто глибина занурення була обрана рівною 7. Дані трансформовані за допомогою «ковзного вікна» представлено на рис. 3.14.

Дата	Реґіон	Товарна група	Обсяг продажу-7	Обсяг продажу-6	Обсяг продажу-5	Обсяг продажу-4	Обсяг продажу-3	Обсяг	
08.04.2020	P1	хлібні вироби	24960,349114349	27651,8507923205	19560,2023612834	23741,8901156693	23494,1820252558	21519,694241877	2436
09.04.2020	P1	хлібні вироби	27651,8507923205	19560,2023612834	23741,8901156693	23494,1820252558	21519,694241877	24369,6071489591	2877
10.04.2020	P1	хлібні вироби	19560,2023612834	23741,8901156693	23494,1820252558	21519,694241877	24369,6071489591	28770,0697072583	2451
11.04.2020	P1	хлібні вироби	23741,8901156693	23494,1820252558	21519,694241877	24369,6071489591	28770,0697072583	24518,7534026638	1806
12.04.2020	P1	хлібні вироби	23494,1820252558	21519,694241877	24369,6071489591	28770,0697072583	24518,7534026638	18062,3074660092	2556
13.04.2020	P1	хлібні вироби	21519,694241877	24369,6071489591	28770,0697072583	24518,7534026638	18062,3074660092	25562,7990628165	2262
14.04.2020	P1	хлібні вироби	24369,6071489591	28770,0697072583	24518,7534026638	18062,3074660092	25562,7990628165	22628,8265079252	2229
15.04.2020	P1	хлібні вироби	28770,0697072583	24518,7534026638	18062,3074660092	25562,7990628165	22628,8265079252	22298,7180382014	2638
16.04.2020	P1	хлібні вироби	24518,7534026638	18062,3074660092	25562,7990628165	22628,8265079252	22298,7180382014	26386,8633539665	2463
17.04.2020	P1	хлібні вироби	18062,3074660092	25562,7990628165	22628,8265079252	22298,7180382014	26386,8633539665	24638,5464108963	2443
18.04.2020	P1	хлібні вироби	25562,7990628165	22628,8265079252	22298,7180382014	26386,8633539665	24638,5464108963	24433,2713201405	2014
19.04.2020	P1	хлібні вироби	22628,8265079252	22298,7180382014	26386,8633539665	24638,5464108963	24433,2713201405	20148,9451154416	2328
20.04.2020	P1	хлібні вироби	22298,7180382014	26386,8633539665	24638,5464108963	24433,2713201405	20148,9451154416	23282,5733344773	2604
21.04.2020	P1	хлібні вироби	26386,8633539665	24638,5464108963	22782,6024162093	24433,2713201405	20148,9451154416	23282,5733344773	2365
22.04.2020	P1	хлібні вироби	24638,5464108963	22782,6024162093	24433,2713201405	24421,9808177434	20148,9451154416	23282,5733344773	2604
23.04.2020	P1	хлібні вироби	22782,6024162093	24433,2713201405	24421,9808177434	20148,9451154416	23282,5733344773	23656,8590986783	2354
24.04.2020	P1	хлібні вироби	24433,2713201405	24421,9808177434	20148,9451154416	23282,5733344773	23656,8590986783	23542,6771095854	2601
25.04.2020	P1	хлібні вироби	24421,9808177434	20148,9451154416	23282,5733344773	26042,8905618221	23656,8590986783	23542,6771095854	224
26.04.2020	P1	хлібні вироби	20148,9451154416	23282,5733344773	26042,8905618221	23656,8590986783	23542,6771095854	22441,516411723	2219
27.04.2020	P1	хлібні вироби	23282,5733344773	26042,8905618221	23656,8590986783	26014,7456513029	22441,516411723	22194,0222804456	2220
28.04.2020	P1	хлібні вироби	26042,8905618221	23656,8590986783	23542,6771095854	26014,7456513029	22441,516411723	22194,0222804456	2423
29.04.2020	P1	хлібні вироби	23656,8590986783	23542,6771095854	26014,7456513029	22441,516411723	22194,0222804456	22209,9813830381	263
30.04.2020	P1	хлібні вироби	23542,6771095854	26014,7456513029	22441,516411723	22194,0222804456	22209,9813830381	24232,7522244426	2326
01.05.2020	P1	хлібні вироби	26014,7456513029	22441,516411723	22194,0222804456	22209,9813830381	24232,7522244426	26391,719922031	2431
02.05.2020	P1	хлібні вироби	22441,516411723	22194,0222804456	22209,9813830381	24232,7522244426	26391,719922031	24310,2332275641	245
03.05.2020	P1	хлібні вироби	22194,0222804456	22209,9813830381	24232,7522244426	26391,719922031	24310,2332275641	24542,522387021	2085
04.05.2020	P1	хлібні вироби	22209,9813830381	24232,7522244426	26391,719922031	23268,1033849301	24186,7906709367	24310,2332275641	2669
05.05.2020	P1	хлібні вироби	24232,7522244426	26391,719922031	23268,1033849301	24186,7906709367	21333,9002169073	24310,2332275641	243
06.05.2020	P1	хлібні вироби	26391,719922031	23268,1033849301	24186,7906709367	21333,9002169073	24310,2332275641	24542,522387021	2431
07.05.2020	P1	хлібні вироби	23268,1033849301	24186,7906709367	21333,9002169073	24339,502942979	24310,2332275641	24542,522387021	245
08.05.2020	P1	хлібні вироби	24186,7906709367	21333,9002169073	24339,502942979	24310,2332275641	24542,522387021	23204,3873674434	2320
09.05.2020	P1	хлібні вироби	21333,9002169073	24339,502942979	24310,2332275641	24542,522387021	23204,3873674434	20856,6870430421	2085
10.05.2020	P1	хлібні вироби	24339,502942979	24310,2332275641	24542,522387021	23204,3873674434	20856,6870430421		2669

Рисунок 3.14 – Дані трансформовані методом «ковзного вікна»

Після трансформації відбувалась побудова та навчання нейронної мережі. Майстер обробки в аналітичній платформі Deductor дозволяє побудувати нейронну мережу із заданою структурою, визначити її параметри і навчити за допомогою одного з доступних в системі алгоритмів навчання.

Для навчання нейронної мережі вхідна множина значень розбивалась на дві підмножини: множина значень для навчання – 95% та для тестування – 5%. В ході дослідження розглядалось декілька типів мережі. Аналіз деяких досліджуваних типів мережі представлено у таблиці 3.1.

Найкращою виявилась мережа, що включає 7 нейронів на вхідному шарі, 20 нейронів на наступному прихованому шарі та 1 нейрон на вихідному шарі (рис. 3.15). Як функцію активації було обрано сигмоїду, а для навчання нейромережевої моделі обрано алгоритм Resilient Propagation (RPROP), що передбачає навчання в режимі «офлайн» (рис. 3.15, рис. 3.16).

Таблиця 3.1 Типи мережі

№	Архітектура мережі	Похибка
1	7 : 5: 1	4,7 %
2	7 : 9: 1	5,5 %
3	7: 11: 1	4,6 %
4	7: 15: 1	4,8 %
5	7: 18: 1	4,7 %
6	7: 20: 1	3,7 %

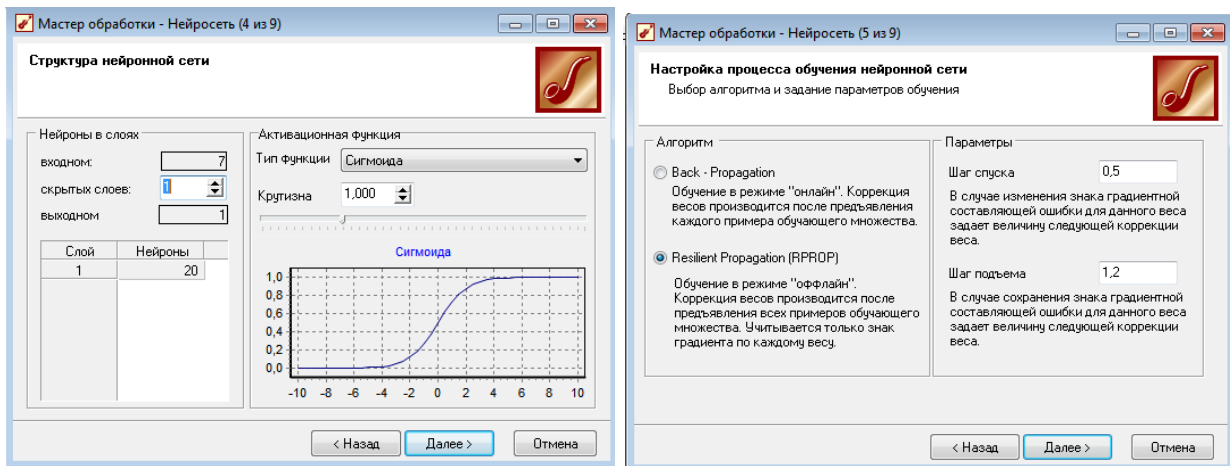


Рисунок 3.15 – Архітектура нейронної мережі

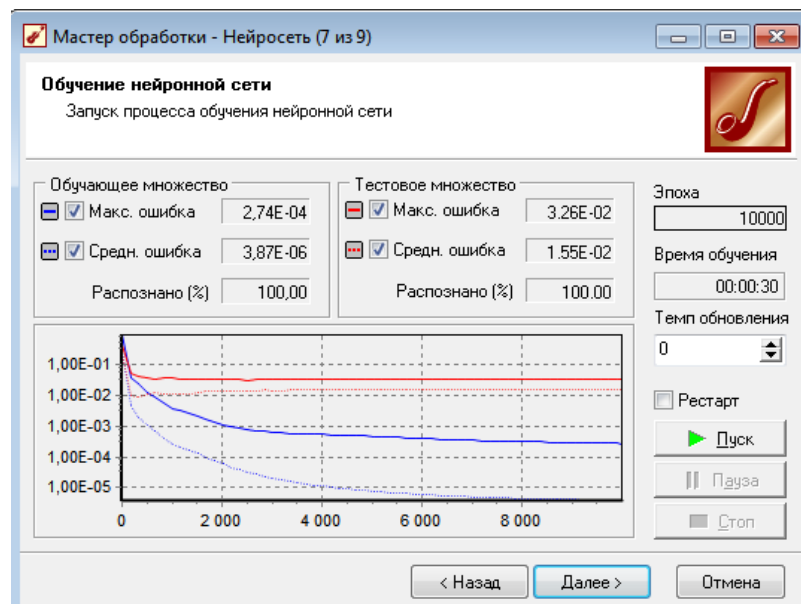


Рисунок 3.16 – Навчання нейронної мережі

Граф побудованої нейромережі, яку використовуємо для прогнозування попиту представлено на рисунку 3.17.

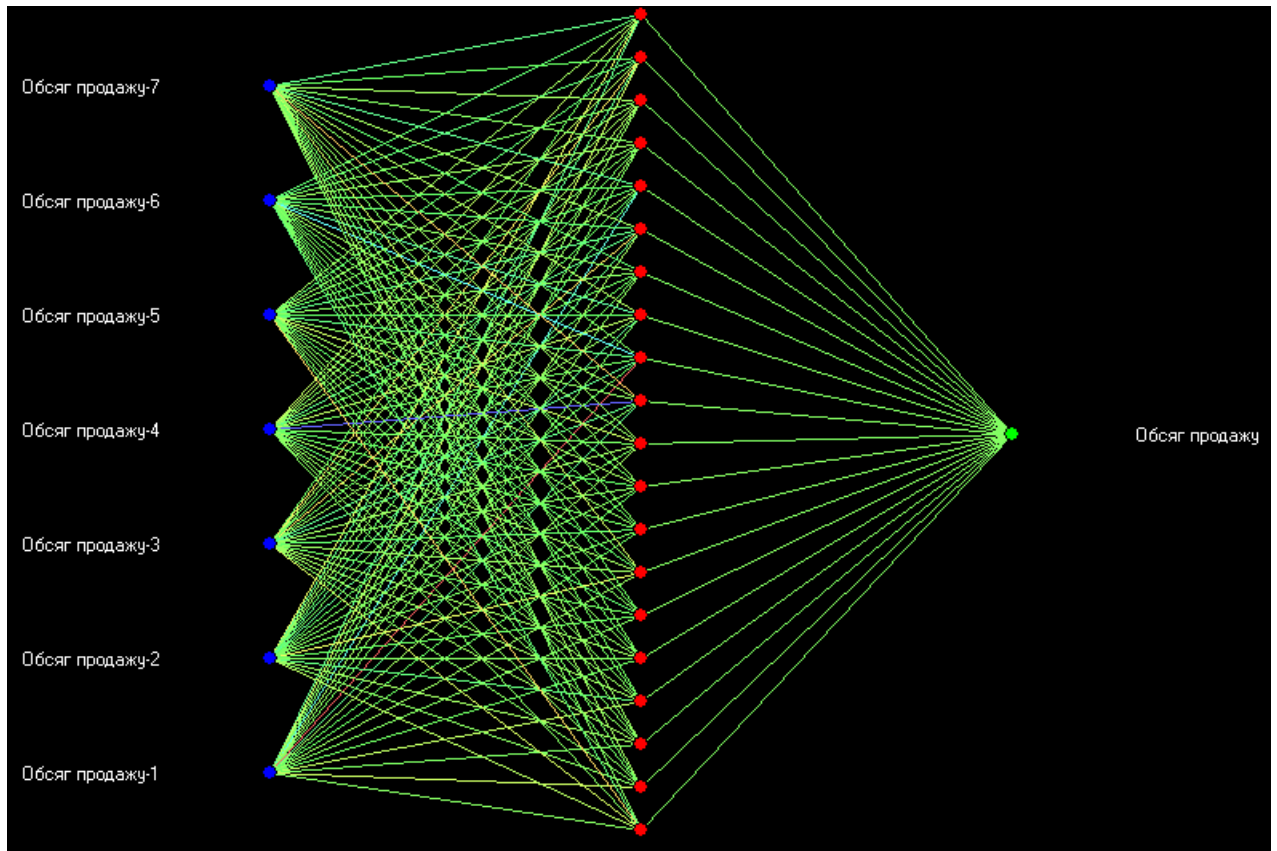


Рисунок 3.17 – Граф нейронної мережі

Для аналізу побудованої моделі зручно використовувати діаграму розсіювання (рис.3.18). Діагональна лінія на діаграмі – це лінія ідеальних значень, зеленими точками позначено реальні значення обсягів продажів, а червоними – вихідні значення моделі. Якщо більша частина точок знаходиться на невеликій відстані від лінії ідеальних значень та розташована в межах заданого «коридору» похибки, то модель добре описує реальний процес. У нашому випадку видно, що якість наближення хороша, всі прогнозовані значення знаходяться дуже близько до ідеальних.

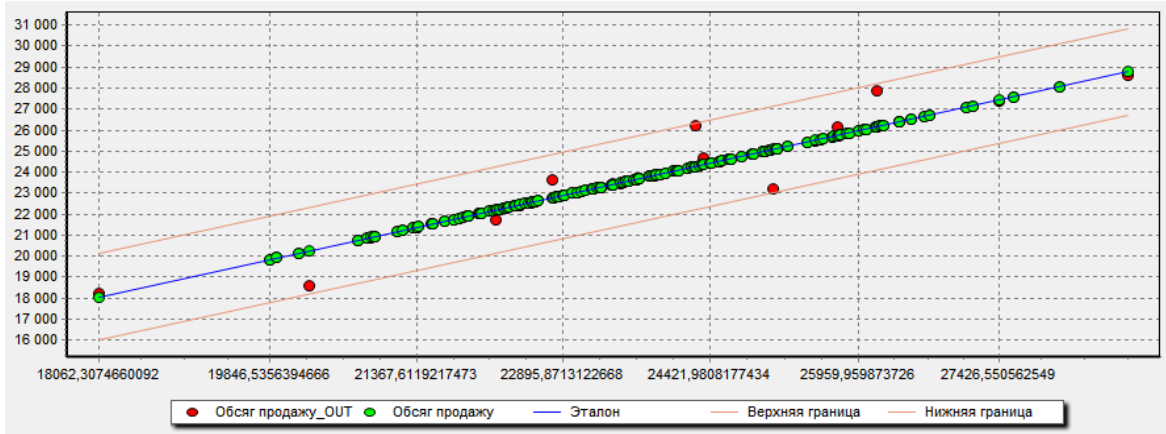


Рисунок 3.18 – Діаграма розсіювання

Далі, використовуючи інструмент «прогнозування» в аналітичній платформі, було побудовано прогноз обсягів продажів на 5 та на 10 днів (рис.3.19 та рис. 3.20)

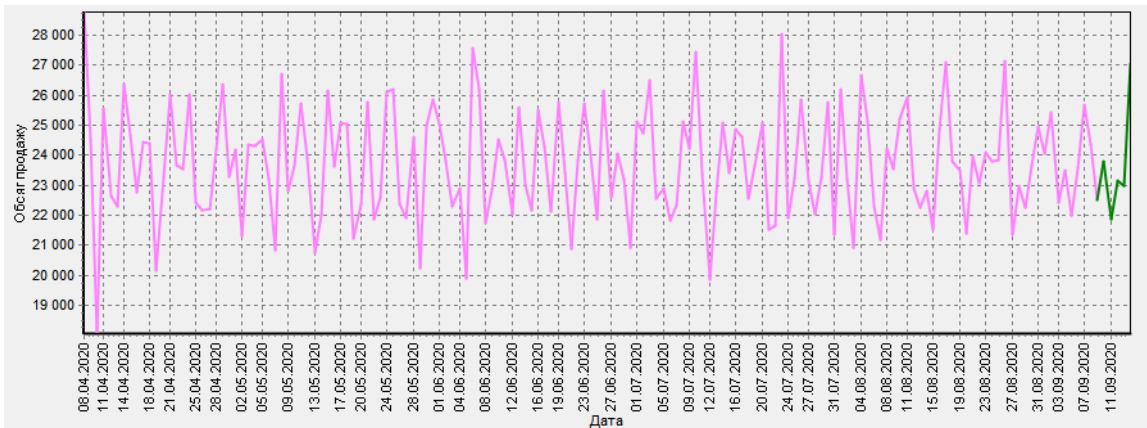


Рисунок 3.19 – Прогноз на 5 днів

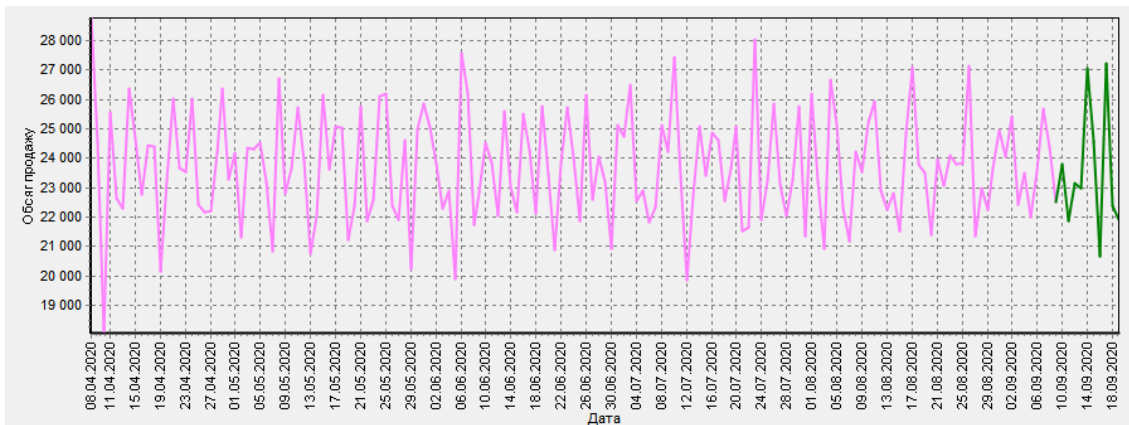


Рисунок 3.20 – Прогноз на 10 днів

Було проведено зіставлення прогнозованих значень з фактичними, що дозволило оцінити якість прогнозу (таблиця 3.2.).

Таблиця 3.2. Оцінка прогнозних значень обсягів продажів «хлібних виробів» у першому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	17211	17812,789	3,50
11.09.2020	26074	24884,991	4,56
12.09.2020	30753	28168,968	8,40
13.09.2020	21386	22974,796	7,43
14.09.2020	17867	18059,843	1,08
15.09.2020	29739	27589,183	7,23
16.09.2020	25185	23664,098	6,04
17.09.2020	26498	27231,199	2,77
18.09.2020	30384	28376,782	6,61
19.09.2020	25745	23936,883	7,02

За аналогічним алгоритмом було побудовано прогнози для товарної групи «хлібні вироби» для другого та третього регіонів, а також для товарної групи «лишкові вироби» для трьох регіонів.

Результати прогнозування та оцінки прогнозів для кожного випадку представлено в таблицях 3.3–3.7.

Таблиця 3.3. Оцінка прогнозних значень обсягів продажів «хлібних виробів» у другому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	23666	22812,879	3,61
11.09.2020	22453	21884,882	2,53
12.09.2020	19310	18168,977	5,91
13.09.2020	21386	22974,768	7,43
14.09.2020	17867	18059,753	1,08
15.09.2020	19584	18589,483	5,08
16.09.2020	17989	18664,087	3,75
17.09.2020	21485	21231,201	1,18
18.09.2020	22427	22376,673	0,22
19.09.2020	20739	21936,783	5,78

Таблиця 3.4. Оцінка прогнозних значень обсягів продажів «хлібних виробів» у третьому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	30837	29812,687	3,32
11.09.2020	23901	23884,891	0,07
12.09.2020	21455	21168,867	1,33
13.09.2020	28514	27974,733	1,89
14.09.2020	35733	34059,902	4,68
15.09.2020	23211	22589,271	2,68
16.09.2020	28782	28664,107	0,41
17.09.2020	23633	23231,200	1,70
18.09.2020	19533	20376,791	4,32
19.09.2020	25030	24936,863	0,37

Таблиця 3.5. Оцінка прогнозних значень обсягів продажів «лишкових виробів» у першому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	7292	7312,88	0,29
11.09.2020	10354	9994,82	3,47
12.09.2020	8526	8368,87	1,84
13.09.2020	9160	8974,96	2,02
14.09.2020	10333	10059,34	2,65
15.09.2020	8257	8589,28	4,02
16.09.2020	13057	12664,98	3,00
17.09.2020	13327	13531,91	1,54
18.09.2020	11199	10976,24	1,99
19.09.2020	14318	14636,33	2,22

Таблиця 3.6. Оцінка прогнозних значень обсягів продажів «лишкових виробів» у другому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	9194	9312,11	1,28
11.09.2020	12295	11994,18	2,45
12.09.2020	13772	13368,74	2,93
13.09.2020	12553	12974,65	3,36
14.09.2020	12666	12059,49	4,79
15.09.2020	7940	7579,58	4,54
16.09.2020	10119	10462,84	3,40
17.09.2020	12661	12532,51	1,01
18.09.2020	11877	11995,94	1,00
19.09.2020	8182	8456,56	3,36

Таблиця 3.7. Оцінка прогнозних значень обсягів продажів «листочкових виробів» у третьому регіоні

Дата	Фактичне значення	Прогнозне значення	Похибка (%)
10.09.2020	15218	14912,89	2,00
11.09.2020	9707	9994,84	2,97
12.09.2020	10493	10368,56	1,19
13.09.2020	12214	12674,78	3,77
14.09.2020	10333	10058,46	2,66
15.09.2020	15561	15779,88	1,41
16.09.2020	9466	9662,44	2,08
17.09.2020	7330	7538,59	2,85
18.09.2020	10859	10985,76	1,17
19.09.2020	11591	11456,36	1,16

Одержані результати прогнозів дозволили керівництву підприємства оцінити попит споживачів на товари досліджуваних товарних груп та скорегувати обсяги поставок даних товарів в торгівельні точки зазначених регіонів.

ВИСНОВКИ

Прогнозування часових рядів за допомогою стандартних методів прогнозування засноване лише на тій інформації, яка міститься в історичних даних ряду. Однак, існує багато сторонніх чинників, що впливають на споживчий попит. Моделі багатовимірної регресії і модифікації стандартних алгоритмів, які дозволяють враховувати зовнішні фактори, недостатньо гнучкі. Тому за метод прогнозування було вирішено взяти нейронні мережі. Використання методів прогнозування, заснованих на побудові багатосарової нейронної мережі, є доцільним при прогнозуванні різних бізнес-процесів та економічних явищ.

В ході дослідження було проаналізовано сучасні підходи до математичного прогнозування та виділено три основних: метод експертних оцінок, детерміновані методи, стохастичні методи. Наступним дуже важливим кроком дослідження було ознайомлення з видами моделей часових рядів для прогнозування попиту, було встановлено чотири найбільш популярних моделі це: тренд-сезонні моделі, моделі експоненціального згладжування, моделі авторегресії і ковзного середнього.

Всебічно було розглянуто проблему розуміння та встановлення взаємозв'язків між агрегованими даними, як виявилось, найкращим чином цю проблему вирішують продукти, що використовують багатовимірний оперативний аналіз даних OLAP (On-line analytical processing). Після чого, з ціллю майбутнього використання, дослідили будову нейрона, способи навчання нейронної мережі. Поетапно побудовано процес прогнозування за допомогою нейронних мереж який складається з семи етапів.

На практиці створено багатовимірний куб даних по продажам, зріз якого дає можливість побудувати крос-таблиці які відображають багатомірні дані у вигляді двомірної таблиці. Дані таблиці зручні тим, що дають можливість

швидко довільним чином групувати, фільтрувати дані, переставляти стовпці та рядки таблиці з метою побудови необхідних аналітичних звітів.

В процесі виконання дипломної роботи було розглянуто та створено прогнозну модель попиту групи товарів у трьох регіонах, виявили дві групи товарів які приносять найбільше прибутку – хлібні вироби і вироби з листкового тіста. Отримані результати дозволили оцінити попит продукції та скорегувати роботу підприємства.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Agnieszka Duda Prognozowanie w logistyce na przykładzie prognozy metodą średnich ruchomych, Zeszyty Naukowe ASzWoj. – 2017. – nr 3(108). – S. 206-227.
2. Samuel Arthur Some Studies in Machine Learning Using the Game of Checkers. IBM Journal. — 2011. – 3(3). – P. 210–229.
3. Machine learning [Електронний ресурс]. URL: <https://searchenterpriseai.techtarget.com/definition/machine-learning-ML> (дата звернення 29.09.2020).
4. Александров Ю.Л. Экономика товарного обращения: учебник / Ю.Л. Александров, Н.Н. Терещенко; Краснояр. гос. торг.экон. институт – 3-е изд., перераб. и доп. – Москва: МАКС Пресс, 2015. – 456с.
5. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ -Петербург, 2009. — 512 с.
6. Аналитическая платформа Deductor: консолидация данных, OLAP-анализ, ABC-XYZ-анализ, разработка аналитической отчетности : пособие для реализации содержания образовательных программ высшего образования I ступени и переподготовки руководящих работников и специалистов / авт.-сост. И. В. Трусевич. – Гомель: учреждение образования «Белорусский торгово-экономический университет потребительской кооперации», 2014. – 132 с.
7. Андрейшина Н. Б. Концептуальний підхід щодо прогнозування попиту Бізнес Інформ. – 2013. – № 6. – С. 120–124.
8. Арзамасцев А. А. Искусственный интеллект и распознавание образов : учеб. пособие / А. А. Арзамасцев, Н. А. Зенкова. Тамбов : Издательский дом ТГУ им. Г. Р. Державина, 2010. – 196 с.

9. Барсегян А.А. Методы и модели анализа данных: OLAP, Data Mining / А.А.Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – СПб.: БХВ-Петербург, 2011. – 336с.
10. Барсегян А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А.Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – СПб.: БХВ-Петербург, 2010. – 384 с.
11. Бродкевич В. М. Алгоритми машинного навчання та глибокого навчання і їх використання в прикладних додатках / В. М. Бродкевич, В. Я. Ремесло // Міжнародний науковий журнал «Інтернаука». – 2018. – № 11 (51), 1 т.– С.56-60.
12. Глубокое обучение (Deep Learning): обзор [Электронный ресурс]. URL: <https://habr.com/ru/company/otus/blog/459785/> (дата звернення 25.09.2020).
13. Грабовецький Б.Є. Планування та економічне прогнозування: навчальний посібник /Б.Є. Грабовецький. – Вінниця: ВНТУ, 2013. – 66с.
14. Гудфеллоу Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль; пер. с англ. А.А.Слинкина. – М.: ДМК Пресс, 2017. — 652 с.
15. Данилевич С.Б. Современные информационные технологии в экономике. Бизнес-анализ данных средствами аналитической платформы Deductor: учеб. пособие / С.Б. Данилевич, О.В. Дьячкова. – Харьков: Изд-во НУА, 2013. – 64с.
16. Долгіх А.О. Аналіз методів, моделей та програмних засобів прогнозування часових рядів / А.О. Долгіх, О.Г. Байбуз // Открытые информационные и компьютерные интегрированные технологии. – 2018. – № 79. – С. 74-87.
17. Заверач М.М. Застосування аналітичної платформи Deductor при вивченні навчальної дисципліни "системи штучного інтелекту та експертні системи в міжнародних відносинах". URL: <http://ena.lp.edu.ua:8080/bitstream/ntb/19145/1/8-Zaverach-44-51.pdf> (дата звернення 02.11.2020).

18. Заруба В. Я. Инвестиционная оценка конкурентных преимуществ рыночной конъюнктуры промышленных предприятий / В. Я. Заруба, П. А. Орлов, А. П. Косенко // Бизнес Информ. – 2012. – № 7. – С. 24 – 27.
19. Інформаційні системи та технології в управлінні. Методичні вказівки. Частина 1. Прогнозування часових рядів. / Укл.: Біла Н.І. – Запоріжжя: ЗНТУ, 2014. – 60 с.
20. Калініна І.О. Дослідження нейромережових методів у задачах прогнозування / І. О. Калініна // Наукові праці [Чорноморського державного університету імені Петра Могили]. Серія: Комп'ютерні технології. – 2009. – Т. 106, Вип. 93. – С. 132-138.
21. Капитанова О.В. Прогнозирование социально-экономических процессов: Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2016. – 74 с.
22. Катуніна О.С. Аналіз ринку продуктів харчування / О.С. Катуніна, О.Ю. Гузенко // Моделювання та інформаційні системи в економіці. – 2018. – № 95. – С. 146-169.
23. Каширин И.Ю. Интерактивная аналитическая обработка данных в современных OLAP-системах / И.Ю. Каширин, С.Ю. Семченков // Бизнес-Информатика. – 2009. – №2(08) – С.12-19.
24. Кравченко Т. В. Методи прогнозування регіонального економічного розвитку / Т. В. Кравченко// Економічний аналіз : зб. наук. праць / Тернопільський національний економічний університет; редкол.: С. І. Шкарабан (голов. ред.) та ін. –Тернопіль: Видавничо-поліграфічний центр Тернопільського національного економічного університету “Економічна думка”, 2013. –Том 13. – С. 88-94.
25. Круглов В.В. Нейронные сети: теория и практика / В.В. Круглов, В.В. Борисов. – М: Горячая линия – Телеком, 2010. – 382 с.
26. Крючин О.В. Прогнозирование временных рядов с помощью нейронных сетей и регрессионных моделей на примере прогнозирование котировок

- валютных пар / О.В. Крючин, А.С. Козадаев, В.П. Дудаков. URL: <http://zhurnal.apec.relarn.ru/articles/2010/030.pdf> (дата звернення 19.10.2020).
27. Кудрицька Н.В. Прогнозування розвитку транспортно-дорожнього комплексу України за допомогою нейронних мереж / Н.В. Кудрицька // Економіко-математичне моделювання соціально-економічних систем. Збірник наукових праць. Київ. – 2014. – Вип. 19. – С. 198-207.
28. Лароз Д.Т. Методы и модели интеллектуального анализа данных / Д.Т. Лароз. – Wiley Publishing, 2010. – 68 с.
29. Машинное обучение и анализ данных: методические указания / сост. : С.Н.Чуканов, С.Ю.Пестова. – (Серия внутривузовских методических указаний СибАДИ). – Омск : СибАДИ, 2018. – 60 с.
30. Модели оценки, анализа и прогнозирования социально-экономических систем : монография / Под ред. Т. С. Клебановой, Н. А. Кизима. – Х. : ФЛП Павленко А. Г.; ИД «ИНЖЭК», 2010. – 280 с.
31. Моделювання попиту і споживання в маркетингу [Електронний ресурс]. URL: <https://library.if.ua/book/97/6667.html> (дата звернення 11.09.2020).
32. Никульченко А. А. Прогнозирование потребительского спроса на сезонные товары с использованием вектора кривой продаж / А. А. Никульченко // Вісник Національного технічного університету «ХП». Серія: Системний аналіз, управління та інформаційні технології. – 2018. – № 21 (1297). С. 23-27.
33. Овакимян А.С. Нейросетевое прогнозирование временных рядов / А.С. Овакимян, С.Г. Саркисян, М.А. Зироян, В.И. Тинякова. URL: <http://ysu.am/files/Paper4.pdf> (дата звернення 19.10.2020).
34. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям: учеб. пособие / Н. Б. Паклин, В. И. Орешков. – СПб : Питер, 2013. – 704 с.
35. Пивкин К. С. Прогнозирование ключевых показателей розничной сети во времени // Вестник Пермского университета. Серия "Экономика". 2017. – Том 12. – №4. – С.592-608.

36. Писарева О.М. Методы прогнозирования развитие социально-экономических систем / О.М. Писарева.– М.: Высшая школа, 2009. – 592 с.
37. Прокопенко Н.Ю. Системы поддержки принятия решений: учеб. пособие / Н. Ю. Прокопенко; Нижегород. гос. архитектур.-строит. ун-т. – Н. Новгород: ННГАСУ, 2017. – 188 с.
38. Савченко Т. Г. Генезис теорій економічної рівноваги / Т. Г. Савченко // Економіка і регіон. – 2010. – № 1. – С. 198 – 206.
39. Соколов Е. В. Экономико-математическая модель и инструментарий прогнозирования и оптимизации расходов торгового предприятия по видам рекламы / Е. В. Соколов, Р. Н. Измайлов // Прикладная информатика. – 2011. – № 5 (35). – С. 5-15.
40. Строцень Л. Якісні методи прогнозування попиту / Л. Строцень // Галицький економічний вісник. – 2018. – № 1. – С. 113-118.
41. Уоллас Т., Сталь Р. Планирование продаж и операций. Практическое руководство. СПб.: Питер, 2010. – 272 с.
42. Федусенко Е. В. Моделирование и прогнозирование спроса методами многокритериальной оптимизации / Е. В. Федусенко, А. А. Федусенко // Управление развитием сложных систем. – 2010. – Вып. 3. – С. 66 – 70.
43. Фролова Т.А. Краткосрочное прогнозирование на фармацевтическую продукцию / Т.А. Фролова, Д.С. Туляков // Вопросы современной науки и практики. Ун-т им. В.И. Вернадского. – 2011. – № 2(33). – С. 78–82.
44. Хайкин С. Нейронные сети: полный курс 2е изд.; Пер. с англ. – М.: ООО «И.Д.Вильямс», 2016. – 1104 с.
45. Ярошенко Е.В. «Информационные технологии в менеджменте»: Учебное пособие. – М.: МЭСИ, 2014. – 114 с.
46. Эластичность спроса по доходу [Электронный ресурс]. URL: <http://files.school-collection.edu.ru/dlrstore/8fdc8dfe-c9e2-a730-9e63-43cbb5471870/00139084759399235.htm> (дата звернення 09.09.2020).

ДОДАТОК А
(обов'язковий)
Тези конференції

Актуальні проблеми комп'ютерних наук

УДК 004.6

Беляков Н. А., Кучерук О. Я.

Хмельницький національний університет

**ЗАСТОСУВАННЯ ТЕХНОЛОГІЇ OLAP ДЛЯ АНАЛІЗУ ДАНИХ
СПОЖИВЧОГО ПОПИТУ**

В останні роки дедалі більше усвідомлюється необхідність використання методів прогнозування у багатьох сферах бізнесу. Одним з способів вирішення такої задачі є використання сучасних математичних методів прогнозування. Для якісного прогнозування попиту необхідно аналізувати великі масиви даних. Застосування OLAP-технології підвищує ефективність оперативної аналітичної обробки даних

In recent years, there is a growing awareness of the need to use forecasting methods in many areas of business. One way to solve this problem is to use modern mathematical methods of forecasting. For high-quality demand forecasting, it is necessary to analyze large data sets. The use of OLAP-technology increases the efficiency of operational analytical data processing.

У країнах з розвинутою економікою для вирішення різних задач управління все частіше використовуються математичні методи прогнозування. У джерелах, що розглядають цю тему, показано, що більше 50% найбільших підприємств і приблизно 18% середніх та малих підприємств при аналізі та прогнозуванні різних економічних та фінансових показників використовують саме такі підходи [1]. В економіко-математичній термінології прогнозування – це система наукових досліджень якісного та кількісного характеру, створена для виявлення тенденцій розвитку господарства або його частин (галузі, регіони, підприємства та ін.) та пошук оптимального шляху досягнення цілей цього розвитку. Таке поняття відображає конкретні аспекти прогнозування – наявність якісних та кількісних висновків, на основі яких здійснюється прогноз [2].

Високу актуальність та широке поширення інструменти прогнозування набувають у роздрібній торгівлі продовольчими товарами. В умовах швидкоплинності бізнес-процесів замовлення, розміщення, зберігання та продаж товарів, що швидко псуються, край важливо забезпечити баланс між задоволеністю клієнта та залишками товару, який буде утилізовано по закінченню терміну придатності. Тому аналіз попиту на продукцію має першорядне значення і є нині однією з важливих задач бізнес-аналітики. Стохастичність попиту накладає значні обмеження на роботу підприємств роздрібною торгівлі. Зі збільшенням кількості товарних позицій, збільшенням торгових площ та масштабів торгівлі забезпечити якісне прогнозування попиту за допомогою прямих калькуляцій неможливо на

практиці. Отже, важливим є удосконалення існуючих технологій прогнозування споживчого попиту [2].

Однак, наскільки ця задача є актуальною, настільки непросто до сих пір іде пошук її більш перспективних рішень, втілення яких залежить від багатьох факторів і достатньо глибокого знання фахівців деталей предметної області [3]. Зокрема, для якісного прогнозування попиту на підприємствах нині необхідно аналізувати величезні масиви даних, на що витрачається багато часу, що в свою чергу, гальмує процес прийняття рішень. Тому виникає необхідність у використанні технологій комплексного багатовимірного аналізу даних. Такою технологією є OLAP (Online Analytical Processing).

Питання формування й прогнозування споживчого попиту на товари та послуги розглянуто в роботах таких науковців як: Дж. Армстронг, Л. Брейман, С. Брю, Г. Джеймс, Ф. Котлера, У. Маккінлі, А. Мюллер, Д. Уїттон, О. Бозуленко, Б. Грабовецький, Л. Ліпич, А. Нікульченко, В. Пучкова, Л. Строць, Е. Федусенко та ін.

Використання технології OLAP для підтримки прийняття рішень висвітлювали в своїх роботах Д. Акушко, А. Коробко та Т. Пенькова, І. Каширін, Ю. Кондрашов, Є. Ярошенко та ін.

Мета статті здійснення аналізу даних споживчого попиту з використанням OLAP технології.

Першим кроком до побудови прогнозної моделі споживчого попиту є збір даних та попередній їх аналіз. На сьогоднішній день найбільш зручною моделлю представлення інформації для аналізу є багатовимірна модель даних. Саме засоби OLAP є зручним інструментом швидкого аналізу великих обсягів даних і наочного відображення результатів у вигляді малюнків, графіків, діаграм і таблиць.

OLAP – технологія оперативної аналітичної обробки даних, яка використовує методи і засоби для збору, зберігання і аналізу багатовимірних даних в цілях підтримки процесів прийняття рішень [4]. Актуальність застосування OLAP систем полягає в тому, що вони дозволяють працювати з даними в термінах предметної області без знання архітектури зберігання інформації [5]. OLAP забезпечує користувача природною, інтуїтивно зрозумілою моделлю даних, організовуючи їх у вигляді багатовимірних кубів, дозволяє прогнозувати майбутнє та виявляти залежності на основі історичних даних, які неможливо помітити безпосередньо з необроблених даних. OLAP дозволяє витягувати, візуалізувати інформацію та отримувати дані в різних зрізах.

Дослідження проводилось у співпраці з компанією «Мамин хліб». Здійснювався аналіз та прогнозування попиту на продукцію за товарними групами по трьом регіонам Хмельницької області. База даних містить дані щоденної виручки за пів року по п'яти товарним групам. Багатовимірний куб даних по продажах можна розглядати як систему координат, осями якої є виміри, в нашому випадку дата, регіон, товарна група (рис. 1).

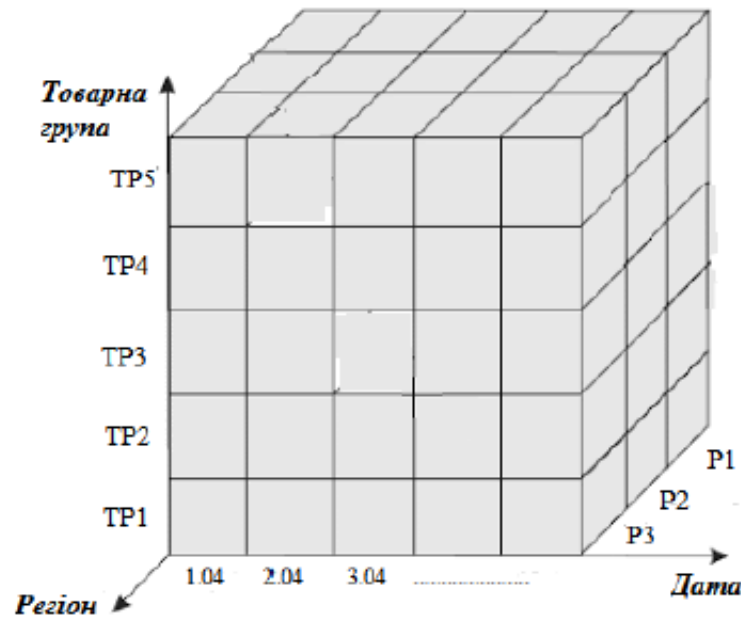


Рисунок 1 – OLAP-куб

Завдяки детальному структуруванню інформації OLAP-куб дозволяє оперативно здійснювати аналіз даних і формувати звіти в різних розрізах і з довільною глибиною деталізації. На рис.2 та рис.3 представлено розрізи OLAP-куба за вимірами «дата» та «товарна група» відповідно.

		Region			
Дата	Товарна група	P1	P2	P3	Итого:
05.04.20	булочні вироби	1 157,00	1 024,00	2 270,00	4 451,00
	кондитерські вироби	6 825,00	5 309,00	6 825,00	18 959,00
	листові вироби	14 303,00	14 643,00	5 108,00	34 054,00
	напівфабрикати	6 924,00	5 842,00	8 871,00	21 637,00
	хлібні вироби	17 163,00	17 163,00	37 187,00	71 513,00
	Итого:		46 372,00	43 981,00	60 261,00
Итого:		46 372,00	43 981,00	60 261,00	150 614,00

Рисунок 2 – Розріз OLAP-куба за виміром «дата»

Виконавши відповідний розріз, одержуємо статистичну інформацію для подальшого прогнозування.

		Регион			
Дата	Товарна група	P1	P2	P3	Итого:
01.04.20	кондитерські вироби	7 173,00	6 397,00	5 816,00	19 386,00
02.04.20	кондитерські вироби	6 417,00	4 861,00	8 166,00	19 444,00
03.04.20	кондитерські вироби	4 671,00	6 539,00	7 473,00	18 683,00
04.04.20	кондитерські вироби	6 151,00	7 112,00	5 959,00	19 222,00
05.04.20	кондитерські вироби	6 825,00	5 309,00	6 825,00	18 959,00
06.04.20	кондитерські вироби	5 138,00	7 992,00	5 899,00	19 029,00
07.04.20	кондитерські вироби	5 082,00	7 623,00	6 841,00	19 546,00
08.04.20	кондитерські вироби	6 476,00	7 555,00	3 958,00	17 989,00
09.04.20	кондитерські вироби	4 126,00	6 459,00	7 356,00	17 941,00
10.04.20	кондитерські вироби	7 166,00	6 778,00	5 423,00	19 367,00

Рисунок 3 – Фрагмент розрізу OLAP-куба за виміром «товарна група»

Під час управління компанією виникають труднощі, які завжди супроводжуються невизначеністю, пов'язаною з прийнятими рішеннями. Одним з інструментів мінімізації невизначеності є оперативний аналіз інформації та прогнозування. Застосування OLAP-технології підвищує ефективність оперативної аналітичної обробки даних, а використання методів сучасної економетрики, статистики та машинного навчання дозволяє здійснювати ефективні прогнози, зокрема споживчого попиту.

Перелік посилань

1. Соколов Е. В. Экономико-математическая модель и инструментарий прогнозирования и оптимизации расходов торгового предприятия по видам рекламы / Е. В. Соколов, Р. Н. Измайлов // Прикладная информатика. – 2011. – № 5 (35). – С. 5-15.
2. Пивкин К. С. Моделирование покупательского спроса на предприятиях розничной торговли на основе методов машинного обучения: дис. ... канд. тех. наук : 08.00.13 / К. С. Пивкин. – Ижевск, 2018. – 145 с.
3. Замятин А.В. Интеллектуальный анализ данных : учеб. пособие. – Томск : Издательский Дом Томского государственного университета, 2016. – 120 с.
4. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.
5. Keith Laker «OLAP Workshop 1: Basic OLAP Concepts» [Електронний ресурс] – URL: http://oracleolap.blogspot.com/2007/12/olap_workshop_1_basic_olap_concepts.html

ДОДАТОК Б

(обов'язковий)

Презентація захисту

ДИПЛОМНА РОБОТА МАГІСТРА

ПРОГНОЗНА МОДЕЛЬ СПОЖИВЧОГО ПОПИТУ НА ПІДПРИЄМСТВАХ РОЗЛІБНОЇ ТОРГІВЛІ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ

СТУДЕНТА ГРУПИ ПММ-19-1

БЕЛЯКОВА НІКІТИ АНДРІЙОВИЧА

НАУКОВОЇ КЕРІВНИК

ДОЦЕНТ КАФЕДРИ ТМІТ КУЧЕРУК О.Я.

ОБ'ЄКТ, ПРЕДМЕТ І МЕТА ДОСЛІДЖЕННЯ.

- Об'єктом дослідження було обрано підприємство роздрібної торгівлі «Мамин Хліб»
- Предмет дослідження – методи прогнозування споживчого попиту на продукцію підприємства.
- Метою дипломної роботи є розробка концепції прогнозування попиту на продукцію підприємства з використанням методів машинного навчання.

ЗАВДАННЯ ДОСЛІДЖЕННЯ:

- Вивчити теоретичну основу організації та управління процесом прогнозування споживчого попиту.
- Дослідити особливості прогнозування попиту у магазинах роздрібної торгівлі.
- Здійснити аналіз існуючих підходів до прогнозування попиту споживачів.
- Побудувати прогнозну модель та спрогнозувати обсяги попиту на продукцію підприємства ТМ «Мамин хліб».

OLAP-ТЕХНОЛОЇЯ

Таблиця X Куб X

Регіон

Дата	Товарна група	P1	P2	P3	Итого:
01.04.2020		54 634,00	55 291,00	38 756,00	148 581,00
02.04.2020		57 749,00	46 948,00	45 170,00	149 767,00
03.04.2020		48 125,00	43 013,00	59 350,00	150 488,00
04.04.2020		52 893,00	47 730,00	48 626,00	149 249,00
05.04.2020		46 372,00	43 801,00	60 251,00	150 614,00
06.04.2020		46 320,00	52 247,00	49 937,00	148 104,00
07.04.2020		43 440,00	43 597,00	64 592,00	152 019,00
08.04.2020		52 972,00	53 232,00	44 953,00	150 757,00
09.04.2020		45 752,00	53 600,00	48 697,00	148 049,00
10.04.2020		42 093,00	47 519,00	59 403,00	149 134,00
11.04.2020		54 141,00	51 320,00	43 954,00	149 615,00
12.04.2020		52 505,00	56 805,00	40 617,00	150 127,00
13.04.2020		50 335,00	46 421,00	55 267,00	152 023,00
14.04.2020		54 644,00	51 180,00	43 946,00	149 770,00
15.04.2020		48 973,00	49 492,00	48 945,00	147 410,00
16.04.2020		40 506,00	50 016,00	59 512,00	150 036,00
17.04.2020		45 118,00	55 251,00	48 001,00	148 380,00
18.04.2020		44 222,00	51 345,00	52 898,00	148 465,00
19.04.2020		40 022,00	47 136,00	60 995,00	148 153,00
20.04.2020		51 809,00	50 423,00	47 154,00	149 386,00
21.04.2020		56 071,00	56 811,00	35 270,00	148 152,00
22.04.2020		47 272,00	49 453,00	52 105,00	148 834,00
23.04.2020		46 695,00	51 196,00	50 483,00	147 376,00
24.04.2020		61 720,00	56 447,00	34 186,00	152 353,00
25.04.2020		52 321,00	53 159,00	45 806,00	151 286,00
26.04.2020		47 680,00	48 445,00	54 461,00	150 586,00

Фрагмент OLAP-куба

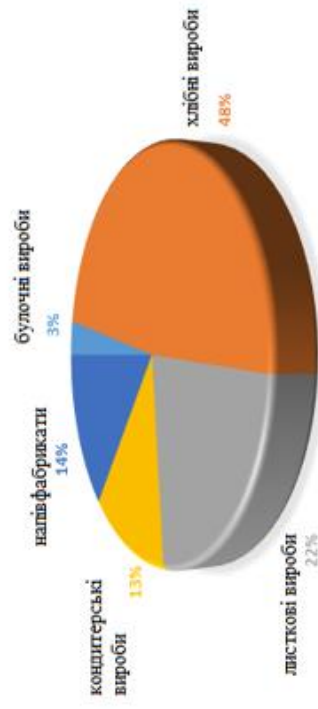
Таблиця X Куб X

Регіон

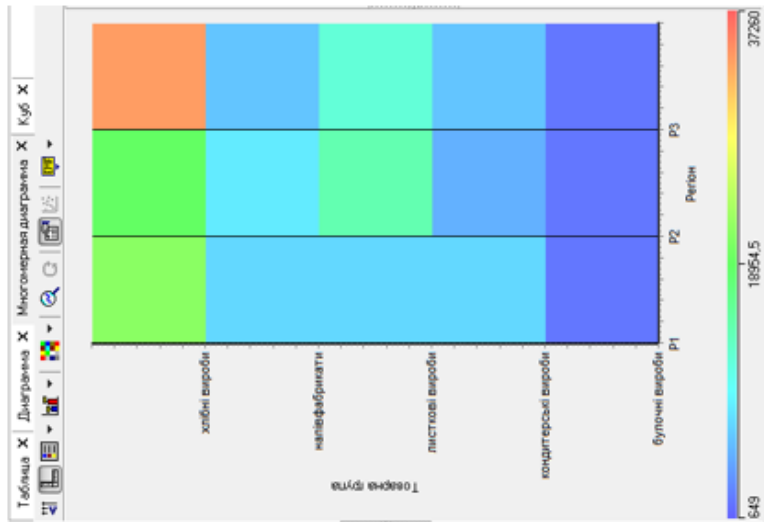
Дата	Товарна група	P2	Итого:
01.04.2020		55 291,00	55 291,00
	Булочні вироби	1 844,00	1 844,00
	кондитерські вироби	4 861,00	4 861,00
	ліктові вироби	9 327,00	9 327,00
	напівфабрикати	9 511,00	9 511,00
	хлібні вироби	21 305,00	21 305,00
	Итого:	46 848,00	46 848,00
	Булочні вироби	1 425,00	1 425,00
	кондитерські вироби	6 539,00	6 539,00
	ліктові вироби	9 843,00	9 843,00
	напівфабрикати	8 034,00	8 034,00
	хлібні вироби	17 172,00	17 172,00
	Итого:	43 013,00	43 013,00
	Булочні вироби	1 309,00	1 309,00
	кондитерські вироби	7 112,00	7 112,00
	ліктові вироби	11 524,00	11 524,00
	напівфабрикати	6 093,00	6 093,00
	хлібні вироби	21 302,00	21 302,00
	Итого:	47 730,00	47 730,00
	булочні вироби	1 024,00	1 024,00
	кондитерські вироби	5 309,00	5 309,00
	ліктові вироби	14 643,00	14 643,00
	напівфабрикати	5 842,00	5 842,00
	хлібні вироби	17 163,00	17 163,00
	Итого:	43 981,00	43 981,00

Розріз OLAP-куба за значенням виміру «регіон»

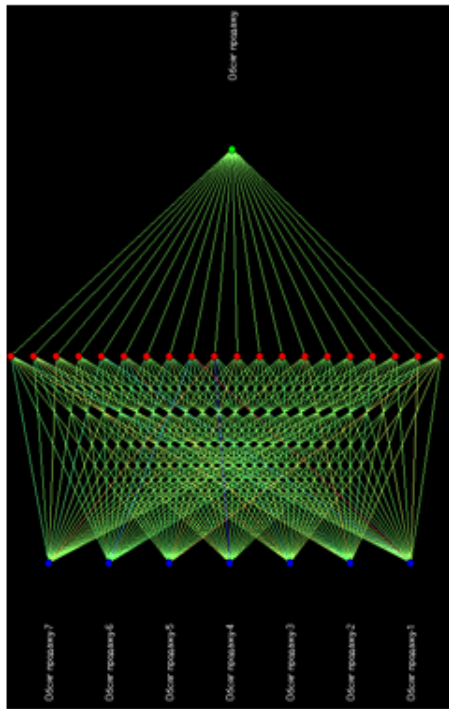
СТРУКТУРА ОБСЯГІВ ПРОДАЖУ



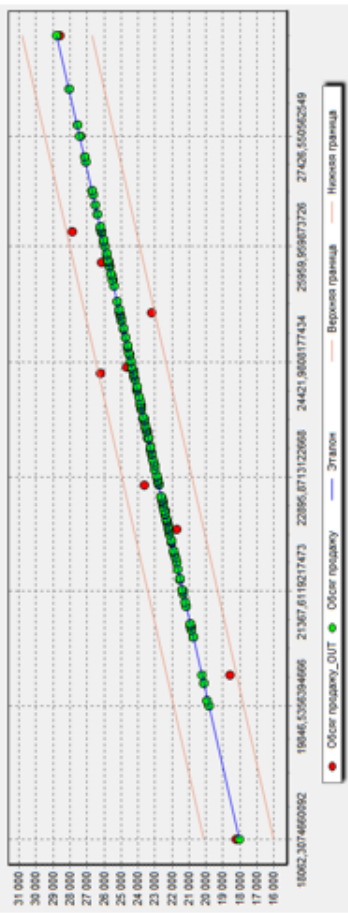
ТОПОГРАФІЧНА ДІАГРАМА ОБСЯГІВ ВИРУЧКИ



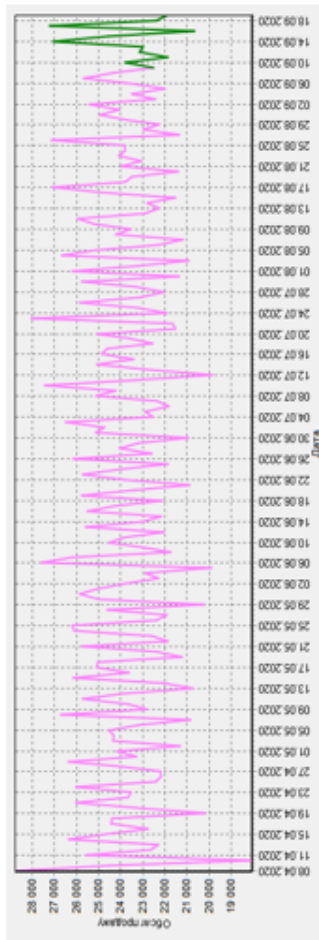
ПРОГНОЗУВАННЯ ЗА ДОПОМОГОЮ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ



Граф нейронної мережі



Діаграма розсіювання



Прогноз на 10 днів

ВИСНОВКИ:

- вивчено теоретичну основу організації та управління процесом прогнозування споживчого попиту;
- розглянуто особливості прогнозування попиту у магазинах роздрібної торгівлі та проведено аналіз існуючих підходів до прогнозування попиту споживачів;
- проведено OLAP-аналіз статистичних даних щоденних продажів за пів року;
- побудовано прогнози обсягів попиту на продукцію підприємства ТМ «Мамин хліб» за товарними групами для трьох регіонів.

Новизна дослідження: запропонований підхід до побудови моделі прогнозування попиту конкретної товарної групи на основі часової вибірки даних із застосуванням сучасних методів машинного навчання, зокрема нейронних мереж.

За темою роботи опубліковано тези: [Беляков Н. А., Кучерук О. Я.](#) Застосування технології OLAP для аналізу даних споживчого попиту / Збірник наукових праць за матеріалами XII всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2020». Хмельницький – 2020. – С. 29-32.

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА ДИПЛОМНУ РОБОТУ

Дипломник Беляков Нікіта АндрійовичТема Прогнозна модель споживчого попиту на підприємстві роздрібної торгівлі на основі методів машинного навчанняСпеціальність 113 – Прикладна математика**Обсяг дипломної роботи:**Кількість листів креслень _____; кількість сторінок записки 81

1. Короткий зміст ДР та прийнятих рішень Магістерська робота присвячена побудові прогнозів попиту споживачів на продукцію підприємства «Мамин хліб». Робота складається з вступу, трьох розділів, висновків, переліку джерел посилання і додатків. В роботі проаналізовано наявні підходи до прогнозування споживчого попиту; проведено аналіз даних попиту за допомогою сучасних засобів аналізу даних; побудовано прогнози з використанням нейронних мереж.

2. Висновок про відповідність ДР поставленому завданню Робота має чітко побудовану структуру та зміст, відповідає назві, поставленій меті та завданню.

3. Характеристика виконання кожного розділу роботи, ступінь використання останніх досягнень науки і техніки і передових методів роботи: _____
У першому розділі проаналізовано сучасні підходи до проблеми прогнозування загалом та прогнозування споживчого попиту, зокрема, описано основні методи прогнозування попиту. У другому розділі описано теоретичні основи технології OLAP-аналізу даних, машинного навчання та застосування штучних нейронних мереж до прогнозування часових рядів. У третьому розділі представлено практичні результати дослідження: проведено аналіз даних по продажах за допомогою технології OLAP; побудовано прогнози за товарними групами підприємства.

4. Позитивні сторони роботи _____
Проведені дослідження мають практичне значення для планування виробництва продукції на підприємстві. В дослідженні застосовано сучасні технології аналізу даних, зокрема, OLAP-аналізу, та прогнозування на основі методів машинного навчання

5. Негативні сторони роботи _____

В роботі детально описано результати прогнозування лише для двох товарних груп. Непогано було б виконати прогноз по всім товарним групам.

6. Оцінка графічного оформлення та пояснювальної записки роботи _____

Пояснювальна записка оформлена згідно чинних вимог

7. Відгук про роботу в цілому _____

Робота виконана у повному обсязі, відповідно до поставлених завдань. Робота відповідає вимогам вищої школи і вимогам, що пред'являються до освітньо-кваліфікаційного рівня «магістр», а її автор Беляков Н.А. заслуговує присвоєння кваліфікації магістра з прикладної математики.

8. Інші зауваження У роботі, по тим чи іншим причинам, не висвітлена інформація про назви регіонів Хмельницької області для яких проводилось дослідження попиту. Але відсутність цієї інформації не знижує загального враження від магістерської роботи.

9. Оцінка дипломної роботи Дипломна робота заслуговує на оцінку добре

РЕЦЕНЗЕНТ (прізвище, ім'я, по-батькові, посада, місце роботи) _____

Ярецька Наталія Олександрівна, кандидат фізико-математичних наук, доцент, доцент кафедри вищої математики та комп'ютерних застосувань

" 30 " жовтня 2020 р.


(підпис)

Завідувачу кафедри ТМІТ
д-р.техн.наук Підченку С.К.

Беляков М.А.

ПІБ здобувача вищої освіти

ФПКТС, 2 курсу, групи ПМм-19-1

ЗАЯВА

З правилами чинного Положення «Про дотримання академічної доброчесності в Хмельницькому національному університеті» від 26.09.2020 (зі змінами від 26.11.2020), згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на плагіатоповіщений (а) та надаю свою згоду на обробку та збереження університетом моєї роботи в інституційному репозитарії університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-технічних засобів(Unicheck та Anti-Plagiarism) та використання роботи для виявлення плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення плагіату в текстах робіт.

Робота для перевірки університетом надається в друкованому та електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

02.12.2020

дата

М.А. Беляков
підпис

Anti-Plagiarism v-15.257

Максимальное совпадение с одним документом 2.0%

Словари проверки: en_US, ru_RU, ua_UA. Ошибок в документах: 8%

ID: 81220 Название: Прогнозна модель споживчого попиту на підприємствах роздрібної торгівлі на основі методів машинного навчання Добавлено в БД: 2020-11-25 Авторы: Беляков Нікіта Андрійович Руководители: Кучерук Оксана Ярославівна Консультанты: Опоненты:	Документ		Суммарное совпадение по Базе Данных	
	Символы	Лексемы	Символы	Лексемы
	84330	671	4318 (5%)	47 (7%)

Источник плагиата

ID	Описание	Наличие плагиата в документе	
		Символы	Лексемы



Имя пользователя:
Kafedra TMIT KhNU

ID проверки:
1005363726

Дата проверки:
04.12.2020 13:30:01 EET

Тип проверки:
Doc vs Internet + Library

Дата отчета:
04.12.2020 14:03:37 EET

ID пользователя:
100005657

Название файла: **Беляков_ПМм-19-1**

Количество страниц: 75 Количество слов: 12816 Количество символов: 96060 Размер файла: 1.58 MB ID файла: 1005656329

1241 слово помечено как "исключенное" и не учитывается в подсчете слов

10.9% Совпадения

Наибольшее совпадение: 1.41% с Интернет-источником (<https://ukrbukva.net/page,2,97196-Metod-obucheniya-neiyron..>)

10.6% Источники из Интернета	181	Страница 77
0.57% Источники из Библиотеки	14	Страница 79

1.11% Цитат

Цитаты	3	Страница 80
--------	---	-------	-------------

Не найдено ни одной ссылки

0% Исключений

Нет исключенных источников

Модификации

Обнаружены модификации текста. Подробная информация доступна в онлайн-отчете.

Замененные символы	69
--------------------	----

РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ
КАФЕДРИ ТЕЛЕКОМУНІКАЦІЙ, МЕДІЙНИХ ТА ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Прогнозна модель споживчого попиту на підприємствах роздрібно́ї торгівлі на основі методів машинного навчання

Автор: Беляков Нікіта Андрійович

Спеціальність: 113 – прикладна математика

Освітня програма: освітньо-професійна

Науковий керівник: Кучерук Оксана Ярославівна, к.п.н доцент

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту.	+
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

Підтвердження:

Наявні текстові запозичення є законними і не є плагіатом, оскільки:

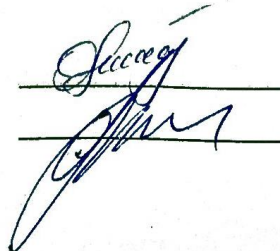
- запозичення є повторами відомих загальноприйнятих словосполучень і тому є допустимі;
- усі запозичення розміщені в розділах, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
- усі запозичення мають належним чином оформленні посилання.

Сумарний обсяг всіх запозичень складає 10.9% . Отже, критичні зауваження щодо плагіату відсутні.

Керівник роботи

Завідувач кафедри ТМІТ

Дата 10.12.2020 р.



О.Я. Кучерук

С.К. Підченко