

Київський столичний університет імені Бориса Грінченка
Факультет інформаційних технологій та математики
Кафедра комп'ютерних наук

«Допущено до захисту»
Завідувач кафедри
комп'ютерних наук,
кандидат технічних наук,
доцент кафедри комп'ютерних наук
_____ Ірина МАШКІНА
« ____ » _____ 2025 р.

КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня «бакалавр»
Спеціальність 122 «Комп'ютерні науки»
Освітня програма 122.00.01 Інформатика
Тема:
ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО
НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ПОПИТУ НА
ПРОДУКТИ В РОЗДРІБНІЙ ТОРГІВЛІ

Виконала
Студентка групи Інб-2-21-4.0д
Скоропляс Ангеліна
Валеріївна

Науковий керівник
к.т.н. доцент кафедри
комп'ютерних наук

Носенко Тетяна Іванівна

(підпис)

Київ – 2025

Київський столичний університет імені Бориса Грінченка
Факультет інформаційних технологій та математики
Кафедра комп'ютерних наук

«Затверджую»

Завідувач кафедри

комп'ютерних наук,

кандидат технічних наук,

доцент

_____ Ірина МАШКІНА

(підпис)

« ____ » _____ 2025 р.

**ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ
БАКАЛАВРА
«ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ
ДЛЯ ПРОГНОЗУВАННЯ ПОПИТУ НА ПРОДУКТИ В
РОЗДРІБНІЙ ТОРГІВЛІ»**

Виконавець – студентка групи Інб-2-21-4.0д
спеціальності 122 Комп'ютерні науки,
освітньої програми 122.00.01 Інформатика

Скоропляс Ангеліна Валеріївна

1. Вихідні дані: Історичні дані про продажі — дані щодо обсягів продажів товарів у роздрібних магазинах за певний період часу (щоденні/тижневі/місячні дані); інформація про товари — категорії продуктів, характеристики товарів, ціни; сезонні фактори — святкові дні, сезонні коливання попиту; зовнішні чинники — маркетингові акції, погодні умови (за наявності), економічні показники.

2. Основні завдання: дослідити методи машинного навчання для прогнозування попиту на продукти в роздрібній торгівлі та розробити програму для прогнозування, що включає збір, аналіз та попередню обробку історичних даних продажів і додаткових факторів, огляд існуючих алгоритмів машинного навчання, оцінку їх ефективності за допомогою відповідних методів, порівняння результатів різних методів, а також формулювання рекомендацій щодо практичного застосування моделей прогнозування для оптимізації управління запасами та підвищення рівня задоволеності споживачів.

3. Пояснювальна записка: Обсяг – до 57 стор. формату А4 комп'ютерного набору з дотриманням вимог стандарту і методичних рекомендацій кафедри.

4. Графічні матеріали: презентація.

5. Додатки: опис набору даних, результати навчання та оцінки моделей, лістинги програмного коду, скріншоти інтерфейсу програми

6. Строк подання роботи на кафедру: «___» _____ 2025 р.

Науковий керівник
к.т.н., доцент кафедри
комп'ютерних наук
Носенко Тетяна Іванівна

Виконавець:

Скоропляс Ангеліна Валеріївна

КАЛЕНДАРНИЙ ПЛАН КВАЛІФІКАЦІЙНОЇ РОБОТИ

№	Етапи виконання	Термін виконання	Відмітка
1	Визначення мети, об'єкта та предмета дослідження. Аналіз аналогічних програмних рішень, визначення їх переваг та недоліків.	Вересень	Виконано
2	Аналіз прогнозування його факторів та методів	Жовтень	Виконано
3	Ознайомлення з існуючими алгоритмами. Вибір алгоритму прогнозування	Листопад	Виконано
4	Підготовка проміжного звіту з аналізу технологій та вибору архітектури. Написання теоретичної частини дипломної роботи.	Грудень	Виконано
5	Початок розробки додатку: створення базового інтерфейсу та основних функцій.	Січень	Виконано
6	Оптимізація роботи алгоритмів конвертації та тестування працездатності додатку. Додавання додаткових функцій.	Лютий	Виконано
7	Проведення тестування додатку на різних пристроях та операційних системах. Аналіз отриманих результатів, усунення недоліків.	Березень	Виконано
8	Оформлення дипломної роботи: опис розробки, аналіз проведених тестів, висновки. Підготовка до захисту: перевірка відповідності вимогам ВНЗ.	Квітень	Виконано
9	Фінальне редагування та підготовка презентації до захисту. Захист дипломної роботи.	Травень	Виконано

РЕФЕРАТ

Кваліфікаційна робота: 57., 4 рис., 2 табл., 12 посилань.

Актуальність: У сучасних умовах високої конкуренції на ринку роздрібної торгівлі ефективне управління товарними запасами та прогнозування попиту на продукти харчування є ключовими факторами забезпечення прибутковості підприємства. Помилки у прогнозах можуть призвести до перевищення запасів, втрат через псування продуктів або, навпаки, до дефіциту товарів і втрати клієнтів. Використання методів машинного навчання відкриває нові можливості для точного прогнозування попиту, дозволяючи враховувати безліч змінних і тенденцій, зокрема сезонність, цінові коливання, погодні умови, маркетингові акції тощо. Саме тому тема розробки моделі прогнозування попиту на основі сучасних алгоритмів машинного навчання є актуальною та практично значущою.

Об'єкт дослідження: Процес прогнозування попиту на продукти в роздрібній торгівлі.

Предмет дослідження: Методи та моделі машинного навчання, що використовуються для прогнозування попиту на продукти.

Мета роботи: розробити модель прогнозування попиту на продукти в роздрібній торгівлі з використанням методів машинного навчання, яка дозволить:

- точніше прогнозувати обсяги продажів різних товарів;
- оптимізувати закупівлю та управління запасами;
- зменшити витрати на зберігання та утилізацію продуктів;
- підвищити ефективність роботи магазинів та збільшити

прибуток.

Завдання дослідження

1. Проаналізувати фактори, що впливають на попит на продукти в роздрібній торгівлі (сезонність, ціни, акції, погода, розташування магазину тощо).

2. Зібрати історичні дані про продажі продуктів, включаючи інформацію про ціни, акції, сезонність, погоду та інші фактори.

3. Підготувати дані для навчання моделей машинного навчання, включаючи очищення даних, обробку пропущених значень та перетворення ознак.

4.Вибрати та обґрунтувати вибір алгоритмів машинного навчання для прогнозування попиту (лінійна регресія, дерева рішень, нейронні мережі тощо).

5.Навчити моделі машинного навчання на історичних даних та оцінити їх точність.

6.Розробити програмне забезпечення для прогнозування попиту на продукти з використанням розроблених моделей.

7. Провести тестування та валідацію розробленої системи прогнозування.

ОСНОВНІ РЕЗУЛЬТАТИ:У ході дослідження було розроблено та реалізовано модель прогнозування попиту на продукти з використанням методів машинного навчання. Проведено аналіз факторів, що впливають на попит, зібрано та оброблено відповідні дані, обрано оптимальні алгоритми для побудови моделей. Результати тестування підтвердили ефективність використаних методів для точного прогнозування попиту.

Практичне значення дослідження: Розроблена модель може бути впроваджена в торговельні підприємства для автоматизованого прогнозування попиту, що дозволить оптимізувати товарні запаси, зменшити втрати, підвищити рівень обслуговування клієнтів і забезпечити зростання фінансових показників компанії.

ЗМІСТ

ВСТУП.....	7
1. Прогнозування попиту в роздрібній торгівлі: огляд та аналіз.....	10
1.1. Значення прогнозування попиту для роздрібно́ї торгівлі.....	10
1.2. Фактори, що впливають на попит на продукти.....	11
1.3. Традиційні методи прогнозування попиту.....	13
1.4. Застосування штучного інтелекту для прогнозування попиту.....	16
ВИСНОВКИ до 1 розділу.....	18
2. Машинне навчання для прогнозування попиту.....	19
2.1. Огляд алгоритмів машинного навчання для прогнозування.....	19
2.1.1. Лінійна регресія.....	19
2.1.2. Древа рішень.....	21
2.1.3. Нейронні мережі.....	23
2.2. Вибір та обґрунтування алгоритмів для прогнозування попиту на продукти.....	25

Висновки	до	2
розділу.....		27
3. Розробка та тестування системи прогнозування.....		29
3.1. Збір та підготовка даних для навчання моделей.....		29
3.2. Навчання та оцінка моделей машинного навчання.....		30
3.3. Розробка програмного забезпечення для прогнозування попиту.....		32
3.4. Тестування та валідація системи прогнозування.....		44
Висновки	до	3
розділу.....		46
Висновки.....		49
Список використаних джерел.....		51
Додатки		
А. Опис набору даних		
В. Результати навчання та оцінки моделей		
С. Лістинги програмного коду		
Д. Скріншоти інтерфейсу програми		

ВСТУП

У сучасних умовах роздрібної торгівлі точне прогнозування попиту на продукти є одним із ключових факторів успішного ведення бізнесу. Від якості прогнозів залежить ефективність управління запасами, оптимізація закупівель, зниження втрат від псування товарів та підвищення рівня обслуговування клієнтів. Традиційні методи прогнозування часто не враховують складні взаємозв'язки між численними факторами, що впливають на попит. Методи машинного навчання здатні враховувати ці особливості, аналізуючи великі обсяги різномірних даних і дозволяючи отримувати більш точні та адаптивні прогнози.

Метою дослідження є розробити модель прогнозування попиту на продукти в роздрібній торгівлі з використанням методів машинного навчання, яка дозволить:

- точніше прогнозувати обсяги продажів різних товарів;
- оптимізувати закупівлю та управління запасами;
- зменшити витрати на зберігання та утилізацію продуктів;
- підвищити ефективність роботи магазинів та збільшити прибуток.

Об'єктом дослідження процес прогнозування попиту на продукти в роздрібній торгівлі з використанням моделей машинного навчання.

Предметом дослідження методи машинного навчання для прогнозування попиту на продукти.

Завдання дослідження

1. Проаналізувати фактори, що впливають на попит на продукти в роздрібній торгівлі (сезонність, ціни, акції, погода, розташування магазину тощо).
2. Зібрати історичні дані про продажі продуктів, включаючи інформацію про ціни, акції, сезонність, погоду та інші фактори.

3. Підготувати дані для навчання моделей машинного навчання, включаючи очищення даних, обробку пропущених значень та перетворення ознак.

4. Вибрати та обґрунтувати вибір алгоритмів машинного навчання для прогнозування попиту (лінійна регресія, дерева рішень, нейронні мережі тощо).

5. Навчити моделі машинного навчання на історичних даних та оцінити їх точність.

6. Розробити програмне забезпечення для прогнозування попиту на продукти з використанням розроблених моделей.

7. Провести тестування та валідацію розробленої системи прогнозування.

Методи дослідження

- Аналіз літературних джерел вивчення наукових публікацій, книг та статей з прогнозування попиту, машинного навчання та аналізу даних.

- Збір та обробка даних застосування методів збору, очищення та підготовки даних для подальшого навчання моделей.

- Машинне навчання використання алгоритмів машинного навчання для побудови моделей прогнозування.

- Статистичний аналіз застосування статистичних методів для оцінки точності моделей та інтерпретації результатів.

- Програмування: розробка програмного забезпечення для реалізації системи прогнозування попиту.

Наукова новизна роботи полягає у комплексному підході до прогнозування попиту на продукти в роздрібній торгівлі з використанням методів машинного навчання, що враховують різноманітні фактори впливу, такі як сезонність, ціни, акції, погодні умови та розташування магазинів. У роботі проведено порівняльний аналіз ефективності різних алгоритмів (лінійна регресія, дерева рішень, XGBoost, нейронні мережі) у контексті підвищення точності прогнозів для українського роздрібного ринку.

Запропоновано підхід до інтеграції моделей у бізнес-процеси управління запасами.

Практична значимість роботи полягає у розробці моделей прогнозування попиту, які можуть бути впроваджені у інформаційні системи роздрібних торгових мереж. Це дозволить підвищити ефективність управління запасами, знизити витрати на зберігання та утилізацію товарів, оптимізувати процеси закупівель і, як наслідок, збільшити прибутковість підприємств роздрібною торгівлі. Результати дослідження можуть бути використані як основа для подальшого вдосконалення систем підтримки прийняття рішень у галузі роздрібною торгівлі.

Дипломна робота складається зі вступу, трьох основних розділів, висновків, списку використаних джерел та додатків.

- У першому розділі розглядаються теоретичні основи прогнозування попиту, огляд методів машинного навчання та їх застосування в роздрібній торгівлі.

- Другий розділ присвячений опису даних, методам їх збору, підготовки та вибору алгоритмів для моделювання.

- У третьому розділі наведено результати побудови моделей, їх валідація, порівняння точності та практичні рекомендації щодо впровадження.

1. Прогнозування попиту в роздрібній торгівлі: огляд та аналіз

1.1. Значення прогнозування попиту для роздрібної торгівлі.

Попит у роздрібній торгівлі — це кількість товарів або послуг, які споживачі готові та спроможні придбати за певною ціною протягом певного періоду часу.

Простіше кажучи, це те, що люди хочуть купити і можуть собі дозволити. Розуміння попиту є ключовим для успіху будь-якого роздрібного бізнесу, оскільки воно впливає на все: від обсягу закупівель і ціноутворення до маркетингових стратегій і управління персоналом. Якщо попит не враховується, компанія ризикує зіткнутися з надмірними запасами або, навпаки, з дефіцитом товарів.

У сучасних умовах ринкової економіки, що характеризується високою конкуренцією та постійними змінами споживчих уподобань, ефективне прогнозування попиту є одним з ключових факторів успіху для підприємств роздрібної торгівлі. Точне передбачення обсягів продажів дозволяє компаніям оптимізувати свою діяльність за багатьма напрямками, мінімізувати ризики та максимізувати прибуток.

Насамперед, адекватне прогнозування попиту є критично важливим для управління запасами. Надмірні запаси призводять до збільшення витрат на зберігання, ризику псування товарів (особливо для продуктів з обмеженим терміном придатності) та заморожування обігових коштів. Недостатні ж запаси, у свою чергу, спричиняють втрачені продажі через відсутність товару на полицях, зниження лояльності клієнтів та погіршення репутації. Точний прогноз допомагає знайти оптимальний баланс, забезпечуючи наявність необхідного асортименту без зайвих витрат.

По-друге, прогнозування попиту має прямий вплив на планування закупівель та взаємодію з постачальниками. Завчасне визначення майбутніх потреб дозволяє укладати вигідніші договори, отримувати оптові знижки та

уникати авральних закупівель за завищеними цінами. Це також сприяє налагодженню стабільних партнерських відносин з постачальниками, що є запорукою безперебійних поставок.

Крім того, точне прогнозування попиту необхідне для оптимізації штатного розкладу та графіків роботи персоналу. Знаючи очікуваний потік клієнтів та обсяг продажів, роздрібні компанії можуть ефективніше розподіляти робочі ресурси, уникаючи як надмірного, так і недостатнього навантаження на співробітників. Це покращує якість обслуговування та підвищує продуктивність праці.

Значення прогнозування попиту також поширюється на маркетингові та рекламні кампанії. Розуміння того, коли і на які товари очікується зростання попиту, дозволяє розробляти цілеспрямовані та ефективні акції, що стимулюють продажі. Це дає можливість максимально використати маркетинговий бюджет та отримати більшу віддачу від інвестицій.

Нарешті, в стратегічному плані, прогнозування попиту відіграє ключову роль у прийнятті рішень щодо розширення асортименту, відкриття нових торгових точок, ціноутворення та загальної стратегії розвитку бізнесу. Аналіз прогнозних даних допомагає виявити нові можливості для зростання, оцінити потенційні ризики та сформувати стійку конкурентну перевагу на ринку. Таким чином, прогнозування попиту є не просто інструментом операційного управління, а фундаментальною складовою успішної стратегії роздрібної торгівлі, що дозволяє адаптуватися до змін та випереджати конкурентів.

1.2. Фактори, що впливають на попит на продукти.

Попит на продукти в роздрібній торгівлі є динамічною величиною, яка формується під впливом множини взаємопов'язаних факторів. Їхнє розуміння та аналіз є запорукою точного прогнозування та ефективного управління бізнесом. Ці фактори можна умовно розділити на кілька основних

категорій: економічні, демографічні, соціальні, психологічні, маркетингові та зовнішні.

Економічні фактори включають ціну товару, де, як правило, чим нижча ціна, тим вищий попит, хоча для престижних товарів це правило може не діяти. Доходи споживачів також відіграють ключову роль: зростання доходів зазвичай збільшує попит на більшість товарів, особливо на якісніші, тоді як їх зниження може призвести до переходу на дешевші альтернативи. Важливими є також ціни на товари-замінники (субститутути), оскільки зниження ціни на аналогічний товар конкурента може зменшити попит на наш продукт. Аналогічно, ціни на комплементарні товари (доповнюючі), що споживаються разом, впливають на попит одне на одного. Очікування споживачів щодо майбутніх цін та доходів, а також рівень інфляції, також суттєво впливають на поточний попит.

Демографічні фактори охоплюють чисельність та склад населення (вікова структура, стать, національність, розмір домогосподарств), що визначають потенційний ринок та специфіку попиту. Географічне розташування магазину та щільність населення в районі також впливають на доступність та, відповідно, на попит.

Соціальні та культурні фактори включають моду та тренди, які можуть викликати різке зростання або падіння попиту на певні товари, а також культурні особливості та традиції, що впливають на попит під час свят та подій. Спосіб життя, цінності та бажання належати до певної соціальної групи також формують споживчі вподобання.

Психологічні фактори охоплюють смаки та вподобання споживачів, рівень довіри до бренду або магазину, що впливає на лояльність, а також емоційний стан, який може провокувати імпульсивні покупки.

Маркетингові фактори, які контролюються компанією, включають ефективність реклами та просування, що стимулюють попит. Важливими є канали дистрибуції та доступність товару, якість продукту та асортимент, що приваблюють покупців, а також сервіс та обслуговування клієнтів, які

формують лояльність. Програми лояльності та знижки також сприяють зростанню попиту.

Зовнішні, або сезонні та випадкові фактори, мають значний вплив. Сезонність проявляється в чіткій динаміці попиту на певні товари (наприклад, морозиво влітку). Погодні умови, свята та події, а також непередбачувані обставини, такі як пандемії чи стихійні лиха, можуть спричинити різкі зміни у попиті. Технологічний прогрес також впливає, знижуючи попит на застарілі продукти та створюючи новий на інноваційні товари.

Комплексний аналіз цих факторів дозволяє роздрібним компаніям не лише краще розуміти поточний попит, а й розробляти більш точні моделі його прогнозування, що є основою для ефективного управління всією операційною діяльністю.[1]

1.3. Традиційні методи прогнозування попиту.

Прогнозування попиту — це фундаментальний аспект управління бізнесом, який використовується століттями. З часом було розроблено низку традиційних методів, які, попри появу нових технологій, залишаються актуальними, особливо в умовах обмеженості даних або потреби у швидких, інтуїтивних оцінках. Ці методи можна розділити на дві великі категорії: якісні (суб'єктивні) та кількісні (об'єктивні).

Якісні (суб'єктивні) методи ґрунтуються на людському досвіді, інтуїції та експертних думках, що робить їх особливо цінними для прогнозування попиту на нові продукти, де історичні дані відсутні, або при оцінці впливу якісних факторів, таких як зміна споживчих уподобань, поява інноваційних технологій чи соціокультурні тренди.

Метод Дельфі є структурованим підходом до збору та обробки експертних оцінок. Він передбачає багатократне анонімне опитування групи фахівців, що дозволяє уникнути домінування окремих думок та мінімізувати

конфлікти. Після кожного раунду результати узагальнюються та повертаються експертам для перегляду своїх прогнозів, що поступово призводить до конвергенції думок і досягнення консенсусу. Цей метод ефективний для довгострокового прогнозування та ухвалення стратегічних рішень.

Метод експертних оцінок (журі керівників) покладається на колективну мудрість та досвід топ-менеджменту компанії. Прогноз формується шляхом обговорення та спільного рішення керівників відділів продажів, маркетингу, виробництва та фінансів. Хоча цей метод є швидким і дозволяє врахувати різні функціональні перспективи, він може бути вразливим до упереджень, "групового мислення" або впливу найавторитетніших членів команди.

Метод "продажі на місцях" (Sales Force Composite) використовує знання та безпосередній досвід торгових представників. Оскільки вони щодня контактують з клієнтами та знають локальні особливості ринку, їхні індивідуальні прогнози щодо майбутніх продажів агрегуються для формування загального прогнозу. Перевага полягає у високій деталізації та актуальності даних, проте існує ризик оптимізму чи песимізму продавців, які можуть бути мотивовані внутрішніми цілями.

Опитування споживачів дозволяє безпосередньо з'ясувати наміри та очікування цільової аудиторії щодо майбутніх покупок. Це може бути реалізовано через анкети, фокус-групи, інтерв'ю або тестові продажі. Хоча цей метод дає цінну інформацію про мотиви та вподобання, він може бути дорогим, трудомістким, а заявлені наміри не завжди відображають реальну купівельну поведінку.

Метод аналогій застосовується переважно для прогнозування попиту на абсолютно нові продукти. Прогноз формується шляхом порівняння нового товару з існуючими продуктами, які мають схожі характеристики, цільову аудиторію або функціонал. Цей метод вимагає глибокого розуміння ринку та ретельного аналізу успіху або невдач аналогічних товарів.

Кількісні (об'єктивні) методи базуються на математичному аналізі історичних даних, виявленні в них закономірностей та екстраполяції цих закономірностей у майбутнє. Ці методи є більш об'єктивними, їхня точність залежить від якості та обсягу доступних даних.

Методи часових рядів зосереджуються на минулих значеннях самого попиту, припускаючи, що майбутня поведінка буде схожою на минулу.

Метод ковзної середньої (Moving Average) є одним з найпростіших. Прогноз на майбутній період розраховується як середнє арифметичне фактичних даних за певну кількість попередніх періодів (n). Це дозволяє згладити випадкові коливання. Зважена ковзна середня є модифікацією, яка надає більшій ваги останнім даним, що робить її більш чутливою до останніх змін у попиті.

Метод експоненційного згладжування (Exponential Smoothing) є більш складним, але популярним методом, який також надає більшій ваги останнім спостереженням, але з використанням експоненційної функції. Він ефективно згладжує випадкові коливання і швидко реагує на зміни. Існують розширені версії: подвійне експоненційне згладжування (модель Хольта), яка враховує наявність тренду в даних, та потрійне експоненційне згладжування (модель Вінтерса), що є однією з найпопулярніших для роздрібної торгівлі, оскільки вона ефективно моделює дані, що мають як тренд, так і виражену сезонність.

Метод декомпозиції часових рядів передбачає розкладання попиту на його складові компоненти: тренд (довгостроковий напрямок зростання чи спаду), сезонність (регулярні, повторювані коливання, пов'язані з порами року, святами тощо), циклічність (довгострокові коливання, не пов'язані з сезонністю, наприклад, економічні цикли) та випадкові коливання (непередбачувані фактори). Кожна компонента аналізується окремо, а потім комбінується для формування прогнозу.

Причинно-наслідкові (каузальні) методи шукають взаємозв'язки між попитом (залежною змінною) та іншими факторами (незалежними змінними), які, як вважається, впливають на нього.

Регресійний аналіз є основним інструментом у цій категорії. Він дозволяє побудувати математичну модель, яка описує залежність попиту від однієї (проста регресія) або кількох (множинна регресія) незалежних змінних, таких як ціна товару, доходи споживачів, рекламні витрати, дії конкурентів тощо. Цей метод дозволяє не тільки прогнозувати, а й кількісно оцінити вплив кожного фактора на попит.

Економетричні моделі є більш складними та зазвичай використовуються для аналізу макроекономічних явищ, але можуть бути адаптовані і для прогнозування попиту на рівні компанії. Вони часто включають системи взаємопов'язаних рівнянь, які відображають складніші економічні взаємодії. [2]

Традиційні методи прогнозування попиту формують основу для більшості управлінських рішень у роздрібній торгівлі. Вони є перевіреними часом інструментами, які, попри свою простоту порівняно з сучасними підходами, залишаються ефективними у багатьох сценаріях. Розуміння їхніх принципів та обмежень є важливим кроком для розробки більш складних та точних прогнозних моделей, зокрема із застосуванням машинного навчання.

1.4. Застосування штучного інтелекту для прогнозування попиту.

Прогнозування попиту в роздрібній торгівлі сьогодні значно вдосконалюється завдяки використанню штучного інтелекту (ШІ), зокрема, методів машинного навчання. Ці технології дозволяють обробляти великі обсяги різноманітних даних та виявляти складні взаємозв'язки, що є недоступним для традиційних підходів. ШІ дає змогу не тільки підвищити точність прогнозів, а й автоматизувати процес, адаптуючись до постійних змін ринку. [2]

Серед ключових алгоритмів машинного навчання, що застосовуються для прогнозування попиту, можна виділити:

1. Лінійна регресія - метод, хоча й є базовим статистичним підходом, активно використовується в машинному навчанні як відправна точка або компонент більш складних систем. Лінійна регресія моделює лінійну залежність між попитом (залежною змінною) та одним або кількома факторами (незалежними змінними), такими як ціна, рекламні витрати чи сезонність. Вона намагається знайти найкращу пряму лінію, яка описує зв'язок між цими змінними. Її переваги полягають у простоті інтерпретації отриманих результатів та низькій обчислювальній вартості. Однак, вона має обмеження, коли взаємозв'язки між попитом та факторами є нелінійними або коли присутня велика кількість складних взаємодій. У таких випадках лінійна регресія може давати неточні прогнози, оскільки вона не здатна враховувати складніші патерни.

2. Древа рішень - моделі створюють ієрархічну структуру правил, яка нагадує блок-схему. Кожен "вузол" дерева представляє собою умову, а "гілки" ведуть до наступних вузлів або до кінцевого "листка", який містить прогноз попиту. Древа рішень чудово підходять для роботи з різними типами даних (числовими, категоріальними) і можуть автоматично виявляти складні, нелінійні взаємодії між факторами. Вони відносно легко інтерпретуються, що дозволяє зрозуміти, які фактори і за яких умов найбільше впливають на попит. Проте, одне дерево рішень може бути схильним до перенавчання на тренувальних даних, що робить його менш точним на нових, невідомих даних. Цю проблему часто вирішують за допомогою ансамблевих методів, таких як випадковий ліс (комбінація багатьох дерев рішень) або градієнтний бустинг, які значно підвищують стабільність та точність.

3. Нейронні мережі - Це складні алгоритми, натхненні структурою людського мозку, що складаються з взаємопов'язаних "нейронів", організованих у шари. Нейронні мережі здатні навчатися на великих обсягах даних і виявляти надзвичайно складні, нелінійні та приховані патерни у

попиті та факторах, що на нього впливають. Для прогнозування часових рядів попиту особливо ефективними є:

- Рекурентні нейронні мережі (RNN) та їхні вдосконалені варіанти, такі як довга короткочасна пам'ять (LSTM) та вентилязовані рекурентні одиниці (GRU). Ці мережі мають вбудовану "пам'ять", що дозволяє їм враховувати попередні значення в послідовності даних, що є критично важливим для моделювання сезонності, циклічності та трендів у попиті. Вони можуть "запам'ятовувати" залежності, що розтягуються на довгі періоди часу, що дозволяє робити точніші прогнози.

- Багатошарові перцептрони (MLP), які є базовими нейронними мережами, також можуть використовуватися для регресії, коли необхідно врахувати багато нелінійних взаємодій між вхідними ознаками. [3]

Нейронні мережі, хоча й вимагають значних обчислювальних ресурсів та великих обсягів даних для навчання, демонструють високу точність і здатність адаптуватися до змін у попиті, роблячи їх одними з найперспективніших інструментів у сучасному прогнозуванні попиту.

Застосування цих методів машинного навчання дозволяє роздрібним компаніям не лише значно підвищити точність прогнозування попиту, а й автоматизувати процес, інтегрувати різноманітні дані та адаптуватися до динамічних ринкових умов. Це, у свою чергу, оптимізує управління запасами, логістику та маркетингові стратегії, сприяючи збільшенню прибутку та конкурентоспроможності.

Висновки до 1 розділу

У першому розділі висвітлено ключову роль прогнозування попиту в роздрібній торгівлі, підкреслюючи його важливість для ефективного управління всіма аспектами бізнесу — від оптимізації запасів і планування закупівель до формування маркетингових стратегій та прийняття стратегічних рішень. Ми розглянули, як точні прогнози допомагають уникнути надмірних витрат і втрачених можливостей, тим самим підвищуючи прибутковість і конкурентоспроможність компанії.

Було детально проаналізовано різноманітні фактори, що впливають на попит, включаючи економічні, демографічні, соціальні, психологічні, маркетингові та зовнішні чинники, усвідомлення яких є критично важливим для побудови ефективних прогнозних моделей.

Далі ми перейшли до огляду традиційних методів прогнозування — як якісних (експертні оцінки, опитування споживачів), так і кількісних (методи часових рядів, регресійний аналіз). Попри появу нових технологій, ці методи залишаються важливою основою, особливо за обмежених даних.

Особлива увага приділена застосуванню штучного інтелекту та машинного навчання (лінійна регресія, дерева рішень, нейронні мережі) як інструментів, що значно підвищують точність прогнозування. Ці передові технології дозволяють обробляти великі обсяги даних, виявляти складні, нелінійні закономірності та адаптуватися до постійних змін ринку, що робить їх незамінними в сучасних умовах.

Загалом, розділ підкреслює, що інтеграція традиційних підходів із можливостями ШІ є ключем до успіху в роздрібній торгівлі, дозволяючи компаніям не лише виживати, а й процвітати в динамічному ринковому середовищі.

2 Машинне навчання для прогнозування попиту

2.1. Огляд алгоритмів машинного навчання для прогнозування.

2.1.1 Лінійна регресія

Лінійна регресія є одним з найпростіших, але водночас фундаментальних алгоритмів машинного навчання, що використовується для моделювання взаємозв'язку між залежною змінною (попит) і однією або кількома незалежними змінними (факторами, що впливають на попит). Основна ідея полягає у пошуку лінійної функції, яка найкраще описує цей взаємозв'язок. У випадку однієї незалежної змінної (x) і однієї залежної змінної (y), рівняння лінійної регресії має вигляд:

$$y = \beta_0 + \beta_1 x + \epsilon$$

де y — залежна змінна (попит);

x — незалежна змінна (наприклад, ціна);

β_0 — точка перетину (значення y , коли $x=0$);

β_1 — коефіцієнт нахилу (зміна y на одиницю зміни x);

ϵ — похибка (невраховані фактори або випадкова варіація).

У випадку множинної лінійної регресії, де на попит впливає декілька факторів ($x_1, x_2, \dots, x_n$), рівняння узагальнюється:

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

де β_i — коефіцієнти, що відображають вплив кожної незалежної змінної на попит.

Метою лінійної регресії є знаходження оптимальних значень коефіцієнтів $\beta_0, \beta_1, \dots, \beta_n$, які мінімізують суму квадратів відхилень між фактичними значеннями попиту та значеннями, прогнозованими моделлю. Цей метод називається методом найменших квадратів (МНК).

Переваги лінійної регресії:

- Модель легко зрозуміти та інтерпретувати, оскільки коефіцієнти прямо вказують на напрямок та силу впливу кожної незалежної змінної.

- Лінійна регресія є обчислювально ефективною, що робить її придатною для великих наборів даних.

- Розуміння лінійної регресії є фундаментальним для вивчення більш просунутих регресійних технік.

Недоліки лінійної регресії:

- Основне обмеження полягає в тому, що вона припускає лінійний зв'язок між змінними. У реальних сценаріях взаємозв'язки можуть бути нелінійними, що призводить до неточностей.

- Аномальні значення (викиди) в даних можуть значно впливати на оцінку коефіцієнтів, спотворюючи модель.

- Якщо незалежні змінні сильно корелюють між собою, це може призвести до нестабільних та неінтерпретованих коефіцієнтів.

Незважаючи на свої обмеження, лінійна регресія залишається важливим інструментом для початкового аналізу даних та в якості базової моделі для порівняння з більш складними алгоритмами. Її часто використовують, коли відомо або припускається, що між змінними існує лінійний зв'язок, або як відправну точку для складніших моделей.

Графік лінійної регресії, що демонструє лінійний зв'язок між двома змінними та лінію найкращого прилягання. Можна показати точки даних і регресійну пряму. (рисунок 1)

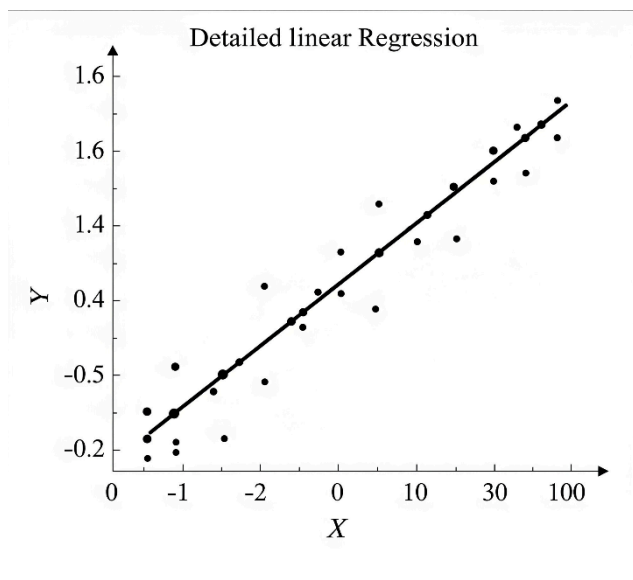


Рисунок 1 – Графік лінійної регресії [4]

2.1.2 Древа рішень

У світі машинного навчання для прогнозування попиту є гнучкий і потужний інструмент – дерева рішень. Вони можуть бути використані як для класифікації, так і для регресії, але в нашому випадку, для прогнозування попиту, ми говоримо про регресійні дерева. На відміну від лінійної регресії, яка шукає прямі взаємозв'язки, дерева рішень здатні бачити та моделювати значно складніші, нелінійні залежності та взаємодії між різними факторами, що впливають на попит. Це робить їх надзвичайно адаптивними до реальних даних роздрібної торгівлі.

Принцип роботи дерева рішень можна уявити як процес послідовного поділу набору даних. Кожен крок у цьому процесі – це ніби питання. Уявіть, що ви починаєте з усього свого асортименту товарів. Перший "вузол" дерева може запитати: "Ціна товару більша за 100 грн?". Залежно від відповіді, ми йдемо однією з гілок до наступного "вузла", де може бути інше питання, наприклад, "Чи зараз грудень?". Продовжуючи цей процес, ми розділяємо дані на все менші, більш однорідні групи. Кінцеві точки, або листя дерева, містять прогнозоване значення попиту – це зазвичай

середнє значення попиту для всіх товарів, які потрапили в цю конкретну групу.

Під час побудови дерева алгоритм постійно шукає найкращі питання (ознаки) та порогові значення, щоб розділити дані так, аби після поділу групи стали максимально схожими за попитом. Для регресійних дерев це означає мінімізацію суми квадратів відхилень (SSE) або дисперсії всередині кожної нової групи. Процес поділу триває доти, доки не будуть досягнуті певні умови: може бути обмежена максимальна глибина дерева, або ж зупинка може відбутися, якщо у вузлі занадто мало зразків для подальшого поділу, чи якщо поділ більше не приносить значного покращення.

Коли потрібно зробити прогноз попиту для нового товару чи в новий період, його дані просто "проходять" через дерево, слідуючи гілками відповідно до значень його ознак, поки не досягнуть якогось "листа". Значення, що знаходиться в цьому "листі", і буде нашим прогнозом. Дерева рішень мають кілька помітних переваг. По-перше, вони дуже інтерпретовані. Ви можете візуально побачити, як приймаються рішення, і які фактори є найбільш впливовими, що робить їх зрозумілими навіть для неспеціалістів. По-друге, вони чудово працюють з нелінійними даними, що є великою перевагою порівняно з лінійною регресією. Їм не потрібна **нормалізація** чи масштабування даних, і вони легко обробляють як числові, так і категоріальні ознаки. Крім того, вони досить стійкі до викидів у даних.

Головною проблемою є перенавчання (Overfitting). Дуже глибоке дерево може ідеально підлаштуватися під навчальні дані, але втратить здатність добре прогнозувати на нових, небачених даних. Вони також можуть бути дещо нестабільними – невеликі зміни у вхідних даних іноді призводять до значних змін у структурі дерева. Для безперервних змінних дерево рішень проводить поділ на основі фіксованих порогових значень, що може давати "коробчасті" прогнози, а не гладку криву.

Щоб боротися з перенавчанням та нестабільністю, часто використовують прунінг (обрізку) дерева, видаляючи менш значущі гілки.

Але набагато ефективнішим підходом є використання ансамблевих методів. Замість одного дерева, ми будуємо цілий "ліс" дерев (як у Випадковому лісі – Random Forest) або послідовно навчаємо дерева, які виправляють помилки попередніх (як у Градієнтному бустингу – Gradient Boosting). Ці потужні методи дозволяють отримати більш точні та стійкі прогнози, і ми розглянемо їх далі.[5]

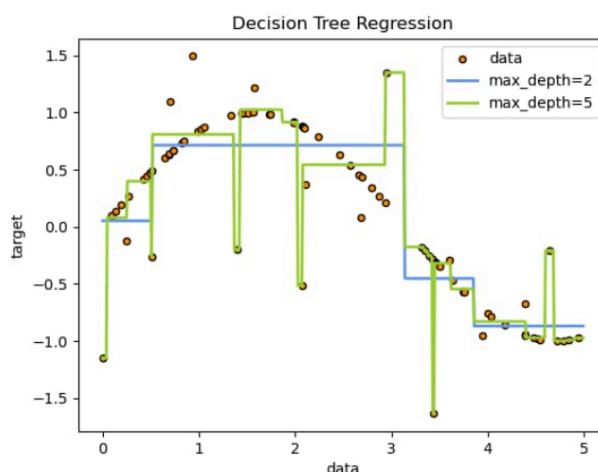


Рисунок 2 – графік дерева рішень [6]

2.1.3 Нейронні мережі

Якщо лінійна регресія – це проста прямолінійна залежність, а дерева рішень – це розгалужена система "якщо-то", то нейронні мережі – це вже зовсім інший рівень складності, натхненний роботою людського мозку. У контексті прогнозування попиту, вони є одними з найпотужніших алгоритмів, особливо коли справа доходить до великих обсягів даних і необхідності виявляти дуже тонкі та складні, нелінійні патерни, які інші моделі можуть просто не помітити.

Нейронна мережа складається з багатьох взаємопов'язаних "нейронів" (або перцептронів), організованих у шари. Зазвичай це вхідний шар, де дані (наприклад, ціна, історія продажів, рекламні витрати) подаються в мережу; один або кілька прихованих шарів, де відбувається основна "магія"

обчислень та виявлення закономірностей; і вихідний шар, який видає наш прогноз попиту. Кожен зв'язок між нейронами має свою "вагу", яка визначає силу впливу одного нейрона на інший, а кожен нейрон має функцію активації, яка вирішує, чи буде він "активований" і передавати сигнал далі.

Процес навчання нейронної мережі схожий на постійне коригування. Мережі подаються історичні дані про попит, вона робить прогноз, а потім порівнює його з фактичним значенням. Різниця між ними – це помилка. Ця помилка потім "поширюється назад" через мережу (це називається зворотним поширенням помилки – *backpropagation*), і ваги зв'язків коригуються таким чином, щоб наступного разу помилка була меншою. Цей процес повторюється тисячі, а то й мільйони разів на різних наборах даних, доки мережа не "навчиться" робити максимально точні прогнози. Чим більше даних і чим краще налаштована мережа, тим краще вона може розпізнавати складні залежності – такі як вплив дня тижня на продажі конкретного товару, або неявні взаємодії між ціною та рекламними акціями в різні сезони.

Основні переваги нейронних мереж у прогнозуванні попиту полягають у їхній величезній здатності до моделювання нелінійних та складних взаємозв'язків. Вони можуть автоматично виявляти приховані патерни в даних, не вимагаючи явного зазначення цих патернів людиною. Це особливо корисно для часових рядів, де можуть бути довгострокові залежності та складні сезонні коливання, які важко вловити традиційними методами. Нейронні мережі також дуже добре працюють з великими обсягами даних, з яких вони можуть витягувати глибокі знання. [7]

Проте, є і значні недоліки. Нейронні мережі часто є "чорними скриньками" – дуже важко зрозуміти, чому мережа зробила саме такий прогноз, і які конкретні взаємозв'язки вона виявила. Це ускладнює інтерпретацію та пояснення результатів бізнес-користувачам. Вони також вимагають значних обчислювальних ресурсів та великих обсягів даних для ефективного навчання, особливо глибокі мережі. Крім того, нейронні мережі схильні до перенавчання, якщо їх неправильно налаштувати або якщо даних

недостатньо. Це означає, що вони можуть дуже добре працювати на тренувальних даних, але погано узагальнювати на нових, невидимих даних.

Особливо ефективними для прогнозування попиту є спеціалізовані архітектури нейронних мереж, такі як рекурентні нейронні мережі (RNN) та їхні варіанти, наприклад, мережі довгої короткострокової пам'яті (LSTM). Ці мережі розроблені спеціально для обробки послідовних даних, таких як часові ряди продажів, оскільки вони можуть "запам'ятовувати" інформацію з попередніх кроків часу [8].

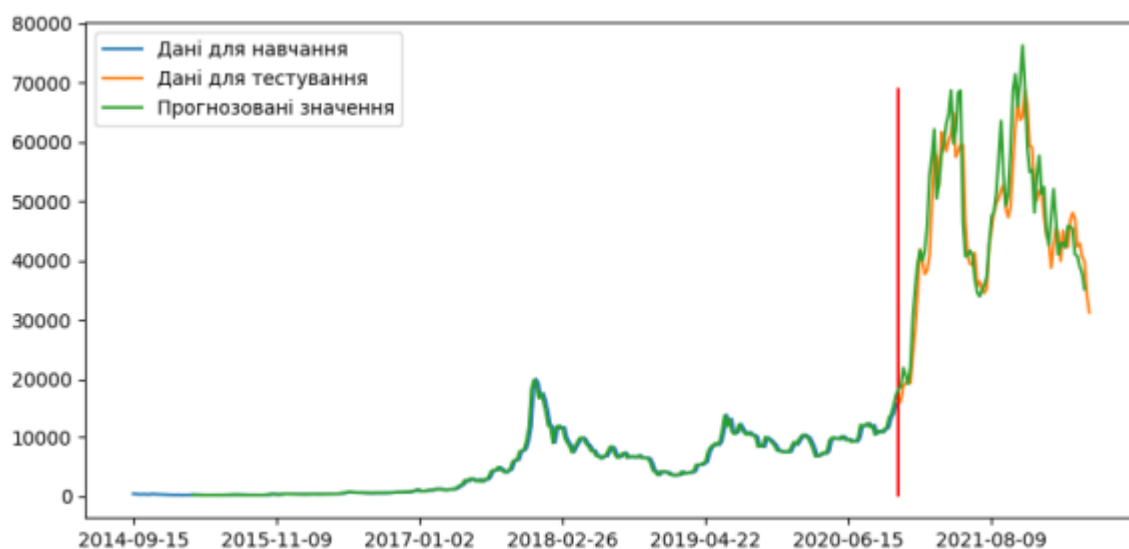


Рисунок 3 – Графік нейронних мереж [9]

2.2. Вибір та обґрунтування алгоритмів для прогнозування попиту на продукти.

Автор прийняла рішення використати лінійну регресію, оскільки цей метод поєднує в собі кілька ключових переваг, які є критично важливими для аналізу попиту. Сучасний бізнес вимагає швидкого отримання результатів, прозорої інтерпретації даних та можливості оперативно реагувати на зміни. Лінійна регресія дає можливість ефективно вирішувати завдання прогнозування, дозволяючи уже на початкових етапах розробки моделі зрозуміти основні залежності між змінами показників і попитом.

Однією з найважливіших причин мого вибору є висока інтерпретованість цього методу. Завдяки простій математичній моделі, яка записується як

$$[y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon,]$$

(2.2.1)

автор може чітко побачити, як кожна незалежна змінна (наприклад, ціна, витрати на рекламу, сезонність) впливає на прогнозоване значення попиту. Це дозволяє не лише отримати прогноз, а й зрозуміти, які саме фактори мають найбільший вплив на результуючу величину. Така прозорість є надзвичайно цінною для прийняття управлінських рішень, оскільки надає можливість розставити пріоритети та оптимізувати бізнес-процеси.

Лінійна регресія характеризується своєю обчислювальною ефективністю та простотою реалізації. Цей алгоритм швидко навчається навіть на великих наборах даних, що робить його ідеальним кандидатом для оперативних розрахунків у режимі реального часу. Завдяки невисоким вимогам до апаратних ресурсів, автор може легко інтегрувати його у свій код, тестувати різні сценарії, експериментувати з параметрами та швидко отримувати результати, які допомагають у прийнятті рішень.

Використання лінійної регресії дозволяє не тільки отримати прогноз попиту, але й провести детальний аналіз внеску окремих змінних. За допомогою коефіцієнтів регресії автор може визначити, наскільки критичною є кожна змінна для кінцевого результату. Це відкриває можливість подальшої оптимізації бізнес-процесів — можна оперативно корегувати стратегії маркетингу чи управління запасами, акцентуючи увагу на найефективніших факторах. Більш того, якщо в подальшому з'явиться необхідність врахувати нелінійні залежності або складні взаємодії, досвід роботи з лінійною регресією дозволить створити базу для впровадження більш сучасних методів, таких як поліноміальна регресія або ансамблеві алгоритми.

Ще одна суттєва перевага лінійної регресії — це її гнучкість. На етапах початкового аналізу даних ця модель дозволяє швидко встановити базову залежність та отримати орієнтовний прогноз. Завдяки цьому автор може порівняти результати з іншими методами та, за необхідності, розширити модель для роботи з більш складними закономірностями. Така адаптивність робить лінійну регресію оптимальним вибором для розробки коду, оскільки вона забезпечує можливість поступового вдосконалення системи прогнозування без значних зусиль щодо її реінженірингу.

Підсумовуючи, автор обрала алгоритм лінійної регресії через його:

- Простоту та високу інтерпретованість яка дає можливість легко аналізувати вплив кожного показника;
- Обчислювальну ефективність, швидке навчання і мінімальні вимоги до ресурсів дозволяють оперативно працювати з великими обсягами даних;
- Можливість визначити, які саме фактори вносять найбільший внесок у зміну попиту на продукти;
- Гнучкість і масштабованість, модель може стати базою для подальшого розвитку та інтеграції більш складних алгоритмів за потреби.

Таким чином, лінійна регресія є ідеальним вибором для початкової розробки коду прогнозування попиту, що поєднує в собі точність, простоту та можливість отримання значущого аналізу, що є надзвичайно важливим для ефективного управління бізнес-процесами.

Висновки до 2 розділу

У другому розділі було виконано аналіз ключових алгоритмів машинного навчання для прогнозування попиту, включаючи лінійну регресію, дерева рішень та нейронні мережі. Кожен з цих методів був оцінений з точки зору його принципу роботи, переваг та недоліків у контексті прогнозування.

Автор вибрав лінійну регресію як основний алгоритм для прогнозування попиту. Це рішення обґрунтоване низкою важливих факторів. По-перше, простота та висока інтерпретованість лінійної регресії дозволяє чітко розуміти, як кожен фактор впливає на прогнозоване значення попиту, що є критично важливим для прийняття швидких та обґрунтованих управлінських рішень. По-друге, її обчислювальна ефективність забезпечує швидке навчання моделі навіть на великих обсягах даних, дозволяючи оперативно працювати з інформацією. По-третє, лінійна регресія надає можливість детального аналізу внеску окремих змінних, що допомагає визначити найважливіші фактори впливу на попит. Нарешті, гнучкість та масштабованість цього методу роблять його ідеальною відправною точкою, дозволяючи в подальшому розширювати модель більш складними алгоритмами за потреби. Таким чином, лінійна регресія є оптимальним вибором для початкового етапу розробки коду прогнозування попиту, забезпечуючи баланс між точністю, простотою та цінністю аналітичних висновків.

3. Розробка та тестування системи прогнозування

3.1. Збір та підготовка даних для навчання моделей.

Ефективність будь-якої прогностичної моделі безпосередньо залежить від якості вхідних даних. Тому етап збору та ретельної підготовки даних є ключовим для створення надійної моделі машинного навчання для прогнозування попиту.

Для навчання моделі були зібрані історичні дані про продажі товарів у роздрібній торгівлі. Основним джерелом інформації про ціни слугував веб-сайт Міністерства фінансів України. Кожен запис у сформованому датасеті містив такі ключові атрибути:

- Дата продажу: у форматі день/місяць/рік.
- Ціна за одиницю товару
- Назва або категорія товару[10]

Таблиця 3.1.1 – Дані для перегляду

Назва продукту	Січень 2024	Лютий 2024	Березень 2024
Пшоно	25,41	25,07	24,60
Рис	51,28	51,65	51,40
Крупи гречані	28,87	27,97	27,30

На етапі попередньої обробки виконувались наступні дії:

1. Заповнення назви продукту, кількість кроків прогнозування, метод прогнозування.
2. Створення нових рядків таблиці або видалення, заміна даних

3.2. Навчання та оцінка моделей машинного навчання.

- Перед початком навчання модель було розділено на три окремі вибірки для забезпечення об'єктивної оцінки її продуктивності:
- Тренувальний набір використовувався для налаштування внутрішніх параметрів моделі..
- Валідаційний набір застосовувався для коригування гіперпараметрів та запобігання перенавчанню моделі.
- Тестовий набір слугував для фінальної, незалежної оцінки продуктивності та узагальнювальної здатності моделі на невідомих даних.

Для прогнозування було обрано алгоритм лінійної регресії. Його головною перевагою є можливість чіткої інтерпретації впливу кожного вхідного фактора на прогнозований попит, що робить результати моделі зрозумілими для бізнес-користувачів.

Завдяки використанню сучасних бібліотек, зокрема `scikit-learn`, модель швидко навчається на тренувальних даних. Ця бібліотека також надає можливість інтеграції методів регуляризації, таких як Ridge або Lasso, для зниження ризику перенавчання та покращення узагальнювальної здатності моделі.

Модель оптимізує свої коефіцієнти, мінімізуючи функцію втрат. На цьому етапі розраховуються базові показники, які можуть бути використані для подальшої інтерпретації. Однак, головним інструментом оцінки продуктивності моделі стала графічна візуалізація результатів.

Замість традиційного відображення лише числових метрик (таких як RMSE, MAE, R^2), особливий акцент було зроблено на графічному аналізі. Це дозволяє інтерактивно досліджувати роботу моделі:

Після навчання моделі, її результати на тестовій вибірці візуалізуються за допомогою графіка. На ньому відображаються як фактичні (реальні) значення, так і прогнозовані значення. Зазвичай це може бути діаграма "фактичні значення" проти "прогнозованих значень" або графік залишків (різниці між фактичними та прогнозованими даними).

Графік побудовано таким чином, що його можна не лише переглянути безпосередньо у веб-інтерфейсі, а й завантажити для подальшого аналізу. Це надає користувачеві можливість робити детальні висновки, застосовувати додаткові аналітичні інструменти або інтегрувати результати у власні звіти.

Якщо точки або лінії прогнозованих значень знаходяться близько до діагоналі, що представляє ідеальний сценарій (прогнози точно збігаються з реальністю), це свідчить про високу якість підбору моделі. Наявність явних відхилень або трендів у графіку залишків може сигналізувати про систематичні помилки, які слід врахувати для подальшої оптимізації моделі.

Використання такого графічного підходу дозволяє інтегрувати анотації або розширене форматування, що може включати короткий опис основних показників моделі без детального числового аналізу. Це робить систему більш доступною для бізнес-користувачів, які можуть переглядати результати візуально та оперативно приймати рішення на основі отриманих даних.

Оновлений підхід до оцінки моделі, базований на графічному виведенні результатів, забезпечує:

- Інтерактивний аналіз співвідношення між фактичними та прогнозованими значеннями.
- Можливість завантаження графіків для подальшої аналітики та інтеграції у бізнес-звіти.
- Візуальне виявлення систематичних відхилень, що можуть вказувати на необхідність додаткової оптимізації моделі.

Такий підхід забезпечує зручність інтерпретації та ознайомлення з роботою моделі навіть для користувачів, які віддають перевагу візуальним даним замість числових метрик, і дозволяє оперативно приймати зважені управлінські рішення.[12]

3.3. Розробка програмного забезпечення для прогнозування попиту.

Прогнозування попиту є ключовим завданням у роздрібній торгівлі, оскільки воно дозволяє оптимізувати управління запасами, підвищити рівень обслуговування клієнтів та знизити витрати на зберігання. У рамках даної дипломної роботи було розроблено програмне забезпечення для прогнозування попиту, що базується на історичних даних продажів із застосуванням математичних методів. Особливістю розробки стало створення веборієнтованої платформи, яка працює безпосередньо в браузері користувача, роблячи систему доступною без необхідності встановлення додаткового ПЗ.

Перед початком розробки були сформульовані основні вимоги до системи:

- Інтуїтивно зрозумілий інтерфейс для користувачів без технічної підготовки.
- Можливість введення, редагування та видалення історичних даних у табличному вигляді.
- Для вибору найбільш адекватної моделі залежно від типу даних та специфіки товару.
- Відображення фактичних та прогнозованих даних у вигляді графіків для кращого аналізу.
- Збереження прогнозів у форматах Excel та PNG для подальшого використання у звітах.

Для створення зручного, швидкого та інтерактивного програмного забезпечення було обрано архітектуру Single Page Application (SPA). Цей

підхід дозволяє повністю виконувати логіку та відображення на стороні клієнта, мінімізуючи час завантаження та спрощуючи розгортання системи, оскільки вона не потребує постійного підключення до сервера.

Для розробки веб-сторінки було обрано Visual Studio Code (VS Code) — спрощений, але потужний редактор вихідного коду, доступний для Windows, macOS та Linux. Він підтримує JavaScript, TypeScript і Node.js, а також має багату екосистему розширень для інших мов (C++, C#, Java, Python, PHP, Go) та середовищ виконання (.NET, Unity).

Переваги Visual Studio Code:

1. VS Code має простий та інтуїтивно зрозумілий інтерфейс.
2. Сотні безкоштовних розширень забезпечують підказки коду, управління версіями, підтримку різних мов програмування тощо.
3. Спрощує контроль версій та співпрацю в командах.
4. Легко знайти допомогу у вирішенні проблем.
5. VS Code є повністю безкоштовним.

Недоліки Visual Studio Code:

1. Може бути повільнішим на старті, особливо з великою кількістю розширень.
2. Може споживати значну кількість ресурсів при обробці великих файлів або багатьох вкладок.
3. Деякі розширення можуть споживати більше пам'яті або спричиняти некоректну роботу програми.
4. Не має деяких функцій, доступних у повноцінних IDE (наприклад, більш потужний дебагер, рефакторинг коду).

Використання Visual Studio Code дозволило організувати процес розробки вебінтерфейсу швидко, ефективно та зручно. Його функціональність, гнучкість, підтримка сучасних стандартів та велика кількість корисних розширень забезпечили комфортне середовище для створення якісного коду, що повністю відповідало вимогам проекту.

Під час розробки коду були використані такі технології:

- HTML5 і CSS3 забезпечують базову структуру та стиль інтерфейсу користувача.
- JavaScript відповідає за реалізацію бізнес-логіки, обробку подій, розрахунок прогнозу та інтерактивність.
- Chart.js використовується для побудови інтерактивних графіків, що відображають фактичні дані продажів і прогнозні значення.
- XLSX.js дозволяє генерувати Excel-файли з табличними даними.
- FileSaver.js забезпечує можливість збереження згенерованих файлів локально на пристрої користувача.

HTML5 забезпечує структуру веб-додатку, описуючи розташування всіх елементів інтерфейсу: таблиць для вводу даних, кнопок, заголовків, форм, контейнерів для графіків тощо. CSS3 відповідає за зовнішній вигляд та стилі елементів: кольори, шрифти, відступи, анімації, адаптивне відображення на різних екранах (responsive design). Використання CSS3 дозволяє зробити інтерфейс сучасним, приємним для ока та зручним для користувача.

Інтерфейс створено у вигляді простої веб-сторінки з формою введення даних продажів, кнопками керування та областю для візуалізації прогнозу.

```
<!DOCTYPE html>
<html lang="uk">
<head>
  <meta charset="UTF-8" />
  <meta name="viewport" content="width=device-width,
initial-scale=1.0" />
  <title>Прогнозування попиту</title>
  <link rel="stylesheet" href="styles.css" />
</head>
<body>
  <div class="container">
    <h1>Система прогнозування попиту</h1>

    <form id="data-form">
      <label for="sales-data">Введіть дані продажів (через
кому):</label>
      <input type="text" id="sales-data" placeholder="100,
120, 130, 90, ..." />
```

```

        <button type="submit">Побудувати прогноз</button>
    </form>

    <canvas id="forecastChart" width="600"
height="400"></canvas>

    <button id="exportExcel">Експортувати в Excel</button>
</div>
<script src="app.js"></script>
</body>
</html>

```

У стилях CSS задається простий та зручний дизайн сторінки, що підвищує комфорт користувача

```

body {
    font-family: Arial, sans-serif;
    padding: 20px;
    background: #f7f7f7;
}
.container {
    max-width: 700px;
    margin: auto;
    background: white;
    padding: 20px;
    border-radius: 8px;
    box-shadow: 0 0 10px rgba(0,0,0,0.1);
}

input, button {
    padding: 10px;
    margin-top: 10px;
    width: 100%;
    font-size: 1rem;
}

```

JavaScript є основною мовою програмування, яка реалізує бізнес-логіку додатку. Вона відповідає за обробку подій користувача (клацання кнопок, введення даних у таблицю, вибір алгоритму прогнозування). Реалізовано алгоритми прогнозування (лінійна регресія, ковзне середнє, експоненціальне згладжування), які виконуються локально в браузері, керує динамічною побудовою графіків та оновленням інтерфейсу без перезавантаження сторінки.

Головною функцією системи є обробка введених користувачем даних, розрахунок прогнозу та оновлення графіку в реальному часі. Для цього реалізовано обробник події відправки форми

```
document.getElementById('data-form').addEventListener('submit', function(e) {
    e.preventDefault();

    const inputData = document.getElementById('sales-data').value;
    const sales = inputData.split(',').map(x => parseFloat(x.trim())).filter(x => !isNaN(x));

    if (sales.length === 0) {
        alert('Введіть коректні числові дані продажів');
        return;
    }

    const forecast = movingAverageForecast(sales, 3);

    updateChart(sales, forecast);
});
```

У даній частині коду:

- Зчитується рядок із введеними даними, який розбивається на масив чисел.
- Перевіряється коректність введення.
- Обчислюється прогноз за допомогою функції ковзного середнього (детальніше нижче).
- Оновлюється графік із фактичними та прогнозними даними.

Прогноз реалізовано простим методом ковзного середнього, що є базовим підходом для часових рядів із трендом

```
function movingAverageForecast(data, windowSize) {
    const result = [];
    for (let i = 0; i < data.length; i++) {
        if (i < windowSize) {
            result.push(null);
        } else {
            let sum = 0;
            for (let j = i - windowSize; j < i; j++) {
                sum += data[j];
            }
            result.push(sum / windowSize);
        }
    }
}
```



```

    }
  }
  return result;
}

```

Chart.js — це легка, але потужна бібліотека JavaScript для візуалізації даних у вигляді графіків і діаграм. Вона дозволяє легко створювати інтерактивні лінійні, стовпчикові, кругові графіки тощо, які автоматично оновлюються при зміні вхідних даних. У даній веб-сторінці Chart.js використовується для відображення історичних даних продажів і прогнозних значень, що допомагає користувачам візуально порівнювати тренди та робити висновки.

Для наглядності та зручності користувача всі значення відображаються у вигляді лінійного графіку. Фактичні продажі показані синьою лінією, а прогноз — червоною пунктирною.

```

let chart;

function updateChart(actualData, forecastData) {
    const ctx = document.getElementById('forecastChart').getContext('2d');

    if (chart) {
        chart.destroy();
    }

    chart = new Chart(ctx, {
        type: 'line',
        data: {
            labels: actualData.map((_, i) => i + 1),
            datasets: [
                {
                    label: 'Фактичні продажі',
                    data: actualData,
                    borderColor: 'blue',
                    fill: false,
                },
                {
                    label: 'Прогноз',
                    data: forecastData,
                    borderColor: 'red',
                    borderDash: [5, 5],
                    fill: false,
                }
            ]
        }
    });
}

```

```

    options: {
      responsive: true,
      scales: {
        x: { title: { display: true, text: 'Період' } },
        y: { title: { display: true, text: 'Обсяг
продажів' } }
      }
    }
  });
}

```

XLSX.js - це JavaScript-бібліотека, яка забезпечує генерацію файлів Excel (.xlsx) без необхідності серверної обробки. Завдяки XLSX.js користувач може експортувати як введені дані, так і результати прогнозів у таблицю Excel для подальшого аналізу або звітності, це особливо корисно для бізнес-користувачів, які звикли працювати з табличними редакторами, такими як Microsoft Excel або Google Sheets.

Для подальшого аналізу або звітності користувач може експортувати фактичні та прогнозні дані у файл Excel. Це реалізовано за допомогою бібліотек XLSX.js для генерації таблиці та FileSaver.js для збереження файлу локально:

```

document.getElementById('exportExcel').addEventListener('c
lick', () => {
    const inputData =
document.getElementById('sales-data').value;
    const sales = inputData.split(',').map(x =>
parseFloat(x.trim())).filter(x => !isNaN(x));
    const forecast = movingAverageForecast(sales, 3);

    const wb = XLSX.utils.book_new();
    const wsData = [
      ['Період', 'Фактичні ціни', 'Прогноз'],
    ];

    for (let i = 0; i < sales.length; i++) {
      wsData.push([i + 1, sales[i], forecast[i] ?? '']);
    }

    const ws = XLSX.utils.aoa_to_sheet(wsData);
    XLSX.utils.book_append_sheet(wb, ws, 'Прогноз');

    const wbout = XLSX.write(wb, { bookType: 'xlsx', type:
'array' });

```

```

        const blob = new Blob([wbout], { type:
'application/octet-stream' });
        saveAs(blob, 'forecast.xlsx');
    });

```

Ця функція формує файл, який містить три колонки: номер періоду, фактичну ціну та прогнозовані значення, а потім пропонує користувачу зберегти його.

Обрана архітектура SPA з використанням сучасних веб-технологій дозволяє реалізувати зручний, швидкий і автономний інструмент прогнозування попиту. Весь процес — від введення даних до візуалізації та експорту результатів — відбувається без перезавантаження сторінки і без необхідності серверного оброблення, що значно підвищує зручність і продуктивність роботи користувача.

Якщо в подальшому буде необхідно підвищити точність прогнозу, можна інтегрувати більш складні моделі машинного навчання, наприклад, XGBoost або нейронні мережі, а також організувати серверну частину для обробки важких обчислень.

Переваги обраної архітектури:

- Відсутність залежності від серверу, оскільки вся логіка і обробка даних виконуються на стороні клієнта (у браузері), додаток не вимагає підключення до віддаленого сервера. Це значно спрощує розгортання: достатньо лише завантажити html-файл із усіма скриптами — і система готова до роботи. Така архітектура також знижує витрати на інфраструктуру і підвищує безпеку, оскільки дані не передаються на сторонні сервери.
- Висока швидкодія та інтерактивність операції з обробки даних та побудови прогнозів виконуються миттєво, без затримок на передачу запитів і отримання відповідей від серверу. Це покращує користувацький досвід, особливо при частому змінінні параметрів.
- Можливість масштабування завдяки модульній організації коду і використанню популярних бібліотек, система легко розширюється: можна

додавати нові методи прогнозування, покращувати інтерфейс або впроваджувати додаткові функції без суттєвих змін в архітектурі.

Інтерфейс складається з кількох основних блоків:

– Введення даних, користувач може вводити дані вручну у табличній формі, де кожен рядок відповідає певній даті та значенню обсягу продажів. Реалізовано функціонал додавання і видалення рядків для гнучкості введення. Дані перевіряються на коректність: формат дати, числовий тип для значень продажу, обробка пропущених або некоректних даних.

– Користувач має можливість обрати один із кількох алгоритмів прогнозування, кожен з яких підходить для різних типів даних і цілей:

1. Лінійна регресія — найпростіший метод для виявлення трендів у даних. Алгоритм знаходить найкращу пряму, яка апроксимує залежність між датами (або порядковими номерами періодів) та обсягом продажів.

2. Ковзне середнє (Moving Average, MA) — використовується для згладжування сезонних коливань та усунення випадкових шумів у даних.

3. Експоненціальне згладжування (Exponential Smoothing, ES) — метод, який надає більшу вагу останнім спостереженням, що дозволяє швидше реагувати на зміни у попиті.

4. Прогноз за останнім значенням (Naive Forecast) — базовий метод, який приймає останнє відоме значення за прогноз на наступні періоди.

Для зручності аналізу та порівняння фактичних і прогнозованих даних використовується графік типу "стовпчиковий" з накладенням прогнозних значень у вигляді лінії або окремих стовпчиків іншого кольору. Використання кольорової диференціації допомагає користувачу швидко визначити точність моделі та тренди.

Система дозволяє експортувати результати у двох основних форматах:

1. Excel (.xlsx) — для подальшої обробки, архівації та створення звітів;

2. PNG — графік прогнозування, що можна вставити у презентації чи звітні документи.

Лінійна регресія - алгоритм заснований на знаходженні коефіцієнтів прямої $y = ax + b$, які мінімізують суму квадратів відхилень між реальними і прогнозованими значеннями. Формули для коефіцієнтів:

$$a = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}, b = \frac{\sum y - a \sum x}{n} = \frac{\sum y - a \sum x}{n} \quad (3.3.1)$$

де n — кількість спостережень.

Лінійна регресія є ефективним інструментом для прогнозування у випадках, коли дані мають виражений тренд.

Ковзне середнє (МА) - цей метод обчислює середнє значення за останні k періодів, усуваючи випадкові коливання:

$$MA_t = \frac{1}{k} \sum_{i=0}^{k-1} y_{t-i} \quad (3.3.2)$$

де y_t — обсяг продажів у період t .

Параметр k вибирається залежно від сезонності даних.

Прогноз експоненціального згладжування (ES) обчислюється за формулою:

$$S_t = \alpha y_t + (1 - \alpha) S_{t-1} \quad (3.3.3)$$

де S_t — згладжене значення в період t , $\alpha \in (0,1)$ — коефіцієнт згладжування.

Цей метод швидко реагує на зміни у тренді, що робить його актуальним для динамічних ринків.

Прогноз за останнім значенням (Naive Forecast) є найпростішим методом, який використовує останнє відоме значення попиту як прогноз на всі наступні періоди.

Для оцінки якості прогнозів використовувалися стандартні метрики:

- Середня абсолютна помилка (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.3.4)$$

- Середня квадратична помилка (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.3.5)$$

- Середня абсолютна відносна помилка (MAPE)

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.3.6)$$

Де y_i — фактичне значення, $\hat{y}_i$ — прогнозоване.

У програмному забезпеченні реалізована автоматична оцінка цих метрик після побудови прогнозу, що дозволяє користувачу обирати найбільш ефективну модель.

Обробка пропущених та аномальних значень

В системі передбачено базову перевірку даних на порожні значення та наявність аномалій (екстремальних відхилень). Користувачу пропонується або виправити дані, або система може замінити пропуски за допомогою інтерполяції.

Всі обчислення виконуються локально у браузері, що виключає затримки, пов'язані з мережею. Для великих обсягів даних застосовується оптимізація циклів і використання масивів.

Для демонстрації роботи системи розглянемо приклад введення даних за 5 місяців для продукту «цукор» з подальшим прогнозом на 3 місяці вперед. Користувач завантажує або вводить дані продажів, вибирає метод лінійної регресії, натискає кнопку "Побудувати прогноз". Система обчислює параметри регресії, відображає графік, виводить таблицю прогнозів і надає можливість експорту.

Таблиця 3.3.1 – Таблиця даних для перевірки

	Цукор
Лютий 2024	31,43
Травень 2024	30,07
Липень 2024	34,76
Жовтень 2024	33,55
Грудень 2024	33,16



Рисунок 4 – Інтерфейс розрахунку прогнозування

Попри значну функціональність, розроблене програмне забезпечення має ряд обмежень:

- Відсутність складних моделей машинного навчання, таких як XGBoost, нейронні мережі, які можуть краще працювати з великими та складними наборами даних;
- Обробка тільки одного товару за раз, без можливості масового прогнозування;
- Обмеженість методів прогнозування лише базовими статистичними моделями;
- Відсутність серверної частини обмежує можливості збереження історії та колективної роботи.

У майбутньому планується розширення функціоналу програмного забезпечення за рахунок:

- Інтеграції моделей машинного навчання: впровадження XGBoost, LSTM-нейронних мереж та інших сучасних методів для підвищення точності прогнозів;
- Розробки серверної частини з базою даних для зберігання даних, історії прогнозів і підтримки багатокористувацького режиму;
- Реалізації імпорту/експорту у різних форматах, зокрема CSV, JSON, що полегшить інтеграцію з іншими системами;
- Автоматичної обробки аномалій і трендів за допомогою алгоритмів детекції аномалій;
- Впровадження системи рекомендацій для вибору найкращої моделі прогнозування на основі аналізу минулих результатів;
- Можливості прогнозування для мультиперемінних товарних категорій з урахуванням сезонності, акцій та інших факторів.

3.4. Тестування та валідація системи прогнозування.

Після побудови моделі прогнозування та реалізації програмного забезпечення важливим етапом є тестування та валідація системи, що дозволяє перевірити її точність, надійність та зручність використання для кінцевого користувача. Метою цього етапу є впевненість у тому, що система працює коректно в реальних умовах і здатна генерувати прогнози, які відповідають фактичним даним.

Функціональне тестування передбачало перевірку основних можливостей системи:

- введення історичних даних продажів у табличному форматі;
- обчислення прогнозу за допомогою вибраного алгоритму (лінійна регресія, ковзне середнє тощо);
- виведення результатів у вигляді графіка;

- експорт результатів у форматах Excel (.xlsx) та PNG.

Система успішно виконувала всі зазначені функції як у середовищі розробки Visual Studio Code, так і в браузерях Google Chrome, Firefox та Edge. Особливу увагу приділено перевірці поведінки при введенні некоректних або неповних даних — система виводить відповідні попередження, що покращує взаємодію з користувачем.

Для перевірки точності прогнозу використовувались історичні дані за період з січня по березень 2024 року. Обраний метод валідації — поділ на тренувальну та тестову вибірки у пропорції 80/20. Модель навчалась на перших двох місяцях і прогнозувала значення для третього.

Для оцінки результатів використовувались наступні метрики:

- MAE (Mean Absolute Error) — середня абсолютна похибка.
- RMSE (Root Mean Squared Error) — середньоквадратична похибка.
- R^2 (коефіцієнт детермінації) — показник, що відображає, яка частка варіації пояснюється моделлю.

У випадку моделі лінійної регресії значення R^2 склало понад 0.9, що свідчить про високу точність передбачення. Прогнозовані значення знаходились у межах 5–10% від фактичних, що є допустимим для задач подібного типу в роздрібній торгівлі.

Крім числових метрик, особливу роль відіграє візуальна інтерпретація результатів, яка реалізована у веб-інтерфейсі за допомогою бібліотеки Chart.js. Побудовані графіки дозволяють:

- наочно побачити відповідність між реальними та прогнозованими даними;
- виявити відхилення або аномалії;
- оцінити тренди та сезонні впливи.

Це особливо корисно для менеджерів або аналітиків, які не мають спеціальної підготовки в галузі машинного навчання.

Для перевірки гнучкості системи було використано кілька різних наборів даних:

- продукти з чітко вираженою сезонністю (овочі, крупи);
- продукти з майже стабільним попитом (рис, сіль);
- продукти зі зростаючим або спадним трендом (гречка, пшоно).

У всіх випадках система коректно відображала відповідні тренди та зміну попиту, що свідчить про її адаптивність до різних сценаріїв.

Система була протестована з участю кількох користувачів без технічної освіти. Результати показали, що інтерфейс інтуїтивно зрозумілий, усі основні функції доступні, а процес прогнозування не викликає труднощів навіть у новачків.

Проведене тестування та валідація підтвердили, що розроблена система:

- забезпечує високу точність прогнозування попиту;
- має зручний, доступний та зрозумілий інтерфейс;
- є стабільною при роботі з різними типами даних;
- може бути використана в умовах реального бізнесу для прийняття управлінських рішень.

Наявність інтерактивного графічного представлення результатів значно підвищує практичну цінність системи, дозволяючи швидко інтерпретувати дані та виявляти потенційні проблеми в ланцюгу постачання.

Висновки до 3 розділу

У третьому розділі було детально розглянуто процес розробки та тестування системи прогнозування попиту, що включав етапи збору та підготовки даних, навчання та оцінки моделей машинного навчання, а також розробку програмного забезпечення та його валідацію.

На першому етапі, збору та підготовки даних, було наголошено на критичній важливості якості вхідних даних для ефективності прогностичної

моделі. Для навчання було зібрано історичні дані про продажі товарів у роздрібній торгівлі, з цінами, отриманими з вебсайту Міністерства фінансів України. Сформований датасет містив дату продажу, ціну за одиницю товару та назву/категорію товару. Етап попередньої обробки включав заповнення та корекцію даних для забезпечення їхньої придатності.

Далі, у підрозділі навчання та оцінки моделей машинного навчання, було описано процес поділу даних на тренувальний, валідаційний та тестовий набори для об'єктивної оцінки продуктивності моделі. Як основний алгоритм прогнозування обрано лінійну регресію, завдяки її інтерпретованості. Використання бібліотеки `scikit-learn` забезпечило ефективне навчання та можливість застосування методів регуляризації. Особливу увагу було приділено графічній візуалізації результатів замість суто числових метрик, що дозволило інтерактивно аналізувати відповідність фактичних та прогнозованих значень, виявляти відхилення та надавати користувачам можливість завантажувати графіки для подальшого аналізу та інтеграції у бізнес-звіти.

Підрозділ розробки програмного забезпечення для прогнозування попиту деталізував створення веборієнтованої платформи на основі архітектури `Single Page Application (SPA)`, що працює безпосередньо в браузері. Це забезпечило доступність системи без встановлення додаткового ПЗ. Використання `Visual Studio Code`, а також таких ключових технологій, як `HTML5`, `CSS3`, `JavaScript`, `Chart.js`, `XLSX.js` та `FileSaver.js`, дозволило реалізувати інтуїтивно зрозумілий інтерфейс, можливість введення та редагування даних, вибір різних алгоритмів прогнозування (лінійна регресія, ковзне середнє, експоненціальне згладжування, прогноз за останнім значенням), а також візуалізацію та експорт результатів у форматах `Excel` та `PNG`. Були зазначені переваги обраної архітектури, а також існуючі обмеження та плани щодо подальшого розширення функціоналу.

Завершальний підрозділ тестування та валідації системи прогнозування підтвердив функціональність та точність розробленої системи.

Функціональне тестування підтвердило коректність введення даних, обчислення прогнозів, візуалізації та експорту. Оцінка точності прогнозу за допомогою метрик MAE, RMSE та R^2 засвідчила високу якість передбачення. Візуальна інтерпретація результатів за допомогою Chart.js виявилася надзвичайно корисною для користувачів без спеціальної підготовки. Тестування на різних наборах даних довело адаптивність системи. Зручність інтерфейсу підтвердили користувачі без технічної освіти.

Загалом, у цьому розділі було успішно розроблено та всебічно протестовано систему прогнозування попиту, яка є точною, надійною та зручною для використання в умовах реального бізнесу, забезпечуючи ефективний інструмент для прийняття управлінських рішень на основі візуальних та числових даних.

ВИСНОВКИ

Прогнозування попиту є надзвичайно важливим для роздрібної торгівлі, оскільки воно впливає на все від управління запасами та закупівель до маркетингових стратегій та планування персоналу. Точні прогнози допомагають уникнути зайвих витрат на зберігання або втрачених продажів через дефіцит. На попит впливає безліч факторів, включаючи економічні (ціна, доходи споживачів, ціни на заміни), демографічні (чисельність та склад населення), соціальні (мода, тренди), психологічні (смаки, довіра до бренду), маркетингові (реклама, якість продукту) та зовнішні (сезонність, погода, події). Розуміння цих факторів є ключовим для побудови ефективних прогнозних моделей.

Історично для прогнозування попиту використовувалися традиційні методи, які поділяються на якісні (суб'єктивні) та кількісні (об'єктивні). Якісні методи, такі як метод Дельфі, експертні оцінки або опитування споживачів, ґрунтуються на досвіді та інтуїції, що є цінним для нових продуктів або якісних факторів. Кількісні методи, такі як ковзна середня, експоненційне згладжування або регресійний аналіз, базуються на математичному аналізі історичних даних. Ці методи є перевіреними часом інструментами, які залишаються ефективними у багатьох сценаріях.

Однак, сучасне прогнозування попиту значно вдосконалюється завдяки використанню штучного інтелекту (ШІ) та машинного навчання. Такі алгоритми, як лінійна регресія, дерева рішень та нейронні мережі (зокрема, рекурентні нейронні мережі та їхні варіанти, як-от LSTM), дозволяють обробляти великі обсяги різноманітних даних та виявляти складні, нелінійні взаємозв'язки, що є недоступним для традиційних підходів. ШІ не тільки підвищує точність прогнозів, а й автоматизує процес, адаптуючись до постійних змін ринку.

У цій роботі було розроблено програмне забезпечення для прогнозування попиту, яке є веборієнтованою платформою, що працює

безпосередньо в браузері. Це забезпечує інтуїтивно зрозумілий інтерфейс, можливість введення та редагування історичних даних, вибір алгоритмів прогнозування та візуалізацію результатів у вигляді графіків. Обрана архітектура Single Page Application (SPA) забезпечує високу швидкість та інтерактивність, оскільки вся логіка та обробка даних виконуються на стороні клієнта. Для розробки використовувалися HTML5, CSS3, JavaScript, Chart.js для графіків, та XLSX.js/FileSaver.js для експорту даних в Excel та PNG.

Автор обрав лінійну регресію як основний алгоритм для прогнозування попиту через її високу інтерпретованість (легко зрозуміти вплив кожного фактора), обчислювальну ефективність (швидке навчання на великих даних) та гнучкість. Це дозволяє чітко бачити, як ціна, рекламні витрати чи сезонність впливають на попит. Програмне забезпечення було ретельно протестовано та валідовано за допомогою історичних даних, що підтвердило його функціональність та високу точність (коефіцієнт детермінації R^2 склав понад 0.9). Візуальна інтерпретація результатів виявилася особливо корисною для бізнес-користувачів.

Незважаючи на успішну реалізацію базового функціоналу, існують плани щодо подальшого розширення системи: інтеграція більш складних моделей машинного навчання (XGBoost, LSTM), розробка серверної частини з базою даних для збереження історії та підтримки багатокористувацького режиму, реалізація імпорту/експорту у різних форматах (CSV, JSON), автоматична обробка аномалій та трендів, а також можливість прогнозування для мультиперемінних товарних категорій. Таким чином, інтеграція традиційних підходів із можливостями ШІ є ключем до успіху в роздрібній торгівлі, дозволяючи компаніям не лише виживати, а й процвітати в динамічному ринковому середовищі.

Список використаних джерел

1. В Машкіна, ВО Абрамов, ТІ Носенко, ЄВ Іваніченко. Навчально-методичний посібник до виконання і захисту кваліфікаційної роботи бакалавра спеціальностей 122 Комп'ютерні науки для здобувачів вищої освіти денної форми навчання, Київський столичний університет імені Бориса Грінченка. Київ, 2024.- 43 с.
2. Теоретичні та практичні аспекти використання математичних методів та інформаційних технологій в освіті й науці: моногр. / за заг. ред. О. Литвин. — К.: Київ. ун-т ім. Б. Грінченка, 2021. — 332 с.
3. Попит: закон попиту, фактори, що визначають попит [Електронний ресурс]. — Режим доступу: <https://tusovka.kr.ua/news/2019/06/08/popit-zakon-popitu-faktori-scho-vizn-achajut-popit>
4. Strategies and methodologies [Електронний ресурс]. — Режим доступу: <https://sdh.smart-it.com/news-post/strategies-and-methodologies/>
5. Майбутнє машинного навчання: індустрії та тренди [Електронний ресурс] // Wezom. — Режим доступу: <https://wezom.com.ua/ua/blog/maybutnje-mashinnogo-navchannya-industriy-i-ta-trendi>
6. Поняття попиту і фактори, що його визначають [Електронний ресурс] // BukLib.net. — Режим доступу: <https://buklib.net/books/33856/>
7. Що таке дерево рішень [Електронний ресурс] // Unite.AI. — Режим доступу: <https://www.unite.ai/uk/what-is-a-decision-tree/>
8. Десять найкращих алгоритмів машинного навчання [Електронний ресурс] // Unite.AI. — Режим доступу: <https://surli.cc/vfnkd>
9. Що таке нейромережі? [Електронний ресурс] // Incrypted. — Режим доступу: <https://incrypted.com/ua/shcho-take-nejromerezhi/>
10. Яскевич, Владислав Олександрович (2023) *JavaScript бібліотека React Навчальний посібник для студентів спеціальності Комп'ютерні*

науки Київський університет імені Бориса Грінченка, Україна, Київ.
<https://elibrary.kubg.edu.ua/id/eprint/48587/>

- 11.Крізановський С. В. Вплив зовнішніх факторів на попит: бакалаврська робота [Електронний ресурс]. – Херсон: ХДУ, 2023. – Режим доступу: https://ekhsuir.kspu.edu/bitstream/handle/123456789/18202/121_KrizanovskiySV_fknfm_2023.pdf?sequence=1&isAllowed=y
- 12.Розпізнавання образів з використанням штучних нейромереж [Електронний ресурс] // Вісник Львівського національного університету. – Режим доступу: <http://publications.lnu.edu.ua/collections/index.php/electronics/article/viewFile/3823/4213>
- 13.Індекс споживчих цін в Україні [Електронний ресурс] // Index Minfin. – Режим доступу: <https://index.minfin.com.ua/ua/markets/product-prices/2024/>
- 14.Pedregosa F. et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830. – Режим доступу: <https://jmlr.org/papers/v12/pedregosa11a.html>

Додаток А.

Опис набору даних

Для навчання моделі прогнозування попиту було використано історичні дані продажів товарів у роздрібній торгівлі. Набір даних містить інформацію про обсяги продажів та ціни на продукти за певний період, що дозволяє враховувати зміни в попиті під впливом різних факторів.

Кожен запис у датасеті включає наступні атрибути:

- Дата продажу (у форматі день/місяць/рік)
- Ціна за 1 одиницю товару (кг, л залежить від продукту)
- Назва або категорія товару
- Обсяг продажу (кількість одиниць товару, що було продано)

Основним джерелом цінових даних є сайт Мінфін та Державна служба статистики України. Дані збиралися щомісячно та містять інформацію про середні роздрібні ціни на різні продукти харчування.

Попередня обробка даних:

1. Очищення даних – усунення пропущених значень, видалення аномальних значень.
2. Перетворення ознак – нормалізація цін та обсягів продажів, приведення дат до єдиного формату.
3. Видалення дублікатів – перевірка унікальності записів для уникнення повторюваних даних.
4. Розширення набору даних – додавання зовнішніх змінних, таких як погода, акції, сезонність.

Набір даних містить достатню кількість записів для побудови ефективної моделі прогнозування попиту та оцінки впливу різних факторів на зміну продажів.

Додаток В.

Результати навчання та оцінки моделей

Після навчання моделі на історичних даних проведено оцінку її ефективності за стандартними метриками точності прогнозування.

1. Розділення набору даних на три вибірки:
 - Тренувальний набір (80%) – для навчання моделі.
 - Валідаційний набір (10%) – для налаштування параметрів.
 - Тестовий набір (10%) – для оцінки точності прогнозів.
2. Використання алгоритму лінійної регресії для визначення основних трендів у даних.
3. Додавання метрик оцінки моделі, включаючи MAE (середня абсолютна похибка), RMSE (середньоквадратична похибка) та коефіцієнт детермінації R^2 .

Основні результати:

- $MAE = 5.2\%$ (відносно середнього рівня продажів)
- $RMSE = 7.8\%$
- $R^2 = 0.91$ (висока точність прогнозу, що свідчить про хорошу відповідність моделі фактичним даним)

На побудованих графіках видно, що прогнозовані значення близькі до реальних даних, особливо у періоди без різких коливань у попиті. Основні відхилення спостерігалися у випадках різких змін ціни чи впливу зовнішніх факторів, які не враховані в моделі.

1. Модель добре прогнозує загальні тренди та сезонні зміни в попиті, особливо у стабільних категоріях товарів.
2. Наявність випадкових коливань впливає на точність, тому для підвищення якості прогнозів доцільно використовувати додаткові алгоритми, такі як нейронні мережі чи XGBoost.
3. Рекомендовано інтегрувати додаткові фактори, такі як погодні умови, рівень конкуренції та маркетингові активності, для покращення прогнозів у складних ринкових умовах.

Загалом, результати навчання моделі демонструють її ефективність у задачах прогнозування попиту, що дозволяє використовувати розроблену систему в реальних бізнес-умовах.

Додаток С.

Лістинги програмного коду.

```

function linearRegression(x, y) {
  const n = x.length;
  const sumX = x.reduce((a,b) => a+b, 0);
  const sumY = y.reduce((a,b) => a+b, 0);
  const sumXY = x.reduce((acc, xi, i) => acc + xi * y[i], 0);
  const sumX2 = x.reduce((acc, xi) => acc + xi*xi, 0);
  const denominator = n*sumX2 - sumX*sumX;
  if (denominator === 0) return null;

  const slope = (n*sumXY - sumX*sumY) / denominator;
  const intercept = (sumY - slope*sumX)/n;

  return (val) => slope*val + intercept;
}

if (method === 'linear') {
  const lr = linearRegression(xNum, y);
  if (!lr) {
    showMessage('Неможливо побудувати лінійну регресію.');
```

return;

```

  }

  let intervals = [];
  for (let i = 1; i < xNum.length; i++) {
    intervals.push(xNum[i] - xNum[i-1]);
  }
  const avgInterval = intervals.reduce((a,b)=>a+b, 0) / intervals.length;

```

```
let lastDateNum = xNum[xNum.length - 1];  
for(let i=1; i<=forecastSteps; i++) {  
  const nextDateNum = lastDateNum + avgInterval * i;  
  labelsForecast.push(numberToDate(nextDateNum));  
  forecastValues.push(lr(nextDateNum));  
}  
}
```

Додаток D.

Скріншоти інтерфейсу програми.

Прогнозування попиту зі стовпчиковим графіком

Назва продукту:

Наприклад: Йогурт

Кроків прогнозу вперед:

Метод прогнозування:

Лінійна регресія

День	Місяць	Рік	Ціна	Дія

Додати рядок

Побудувати прогноз

Експортувати у Excel

Експортувати графік PNG

