

**Пилипчак Максим Володимирович**

Допускається до захисту:  
завідувач кафедри  
інформаційних технологій,  
д. т. н., доцент  
\_\_\_\_\_ Т. В. Нескородєва  
« \_\_\_\_\_ » \_\_\_\_\_ 2022р.

**ДОСЛІДЖЕННЯ СИСТЕМИ ІНТЕРНЕТ-МАГАЗИНУ НА ПІДСТАВІ  
МЕТОДІВ DATA SCIENCE**

Спеціальність 122 «Комп'ютерні науки»

**Кваліфікаційна (магістерська) робота**

Науковий керівник:  
Федоров Є. Є., професор кафедри  
інформаційних технологій  
д-р техн. наук, професор

\_\_\_\_\_  
(підпис)

Оцінка: \_\_\_\_\_ / \_\_\_\_\_ / \_\_\_\_\_  
(бали за шкалою ЄКТС/за національною шкалою)  
Голова ЕК: \_\_\_\_\_  
(підпис)

Вінниця – 2022

## АНОТАЦІЯ

Актуальність: зараз більшість людей роблять покупки в Інтернеті, тому магазин електроніки може стати дуже популярним серед людей.

Мета роботи: створити проект, який дозволить швидко та зручно здійснювати покупки в Інтернеті. Основна мета будь-якого інтернет-магазину – залучити більше потенційних клієнтів і збільшити продажі з найменшим рекламним бюджетом.

Методи дослідження: phpMyAdmin — це PHP веб-додаток із відкритим вихідним кодом із графічним веб-інтерфейсом для керування базами даних MySQL або MariaDB. За допомогою phpMyAdmin ви можете керувати сервером MySQL, запускати SQL-запити та переглядати та редагувати вміст таблиць бази даних із браузера. Ця програма дуже популярна серед веб-розробників, оскільки дозволяє керувати базою даних MySQL без необхідності вводити команди SQL через зручний інтерфейс і з будь-якого комп'ютера, підключеного до Інтернету, без встановлення додаткового програмного забезпечення.

Прогнозні припущення щодо розвитку об'єкта дослідження – додавання додаткових функцій: можливість додавати нові товари, робити прогнози щодо продажів у різні сезони.

## ABSTRACT

**Relevance:** Currently, most people shop online, so the electronics store can become very popular with people.

**Purpose:** To create a project that will make shopping online quickly and conveniently. The main purpose of any online store is to attract more leads and get more sales with the least budget for advertising.

**Research Methods:** phpMyAdmin is a PHP-based open source web application with graphical web interface to administer MySQL or MariaDB. phpMyAdmin allows you to administer MySQL server through a browser, run SQL queries, view and edit the contents of database tables. This program is very popular with web developers because it allows you to manage the MySQL database without typing SQL commands through a friendly interface and from any computer connected to the Internet without the need for additional software.

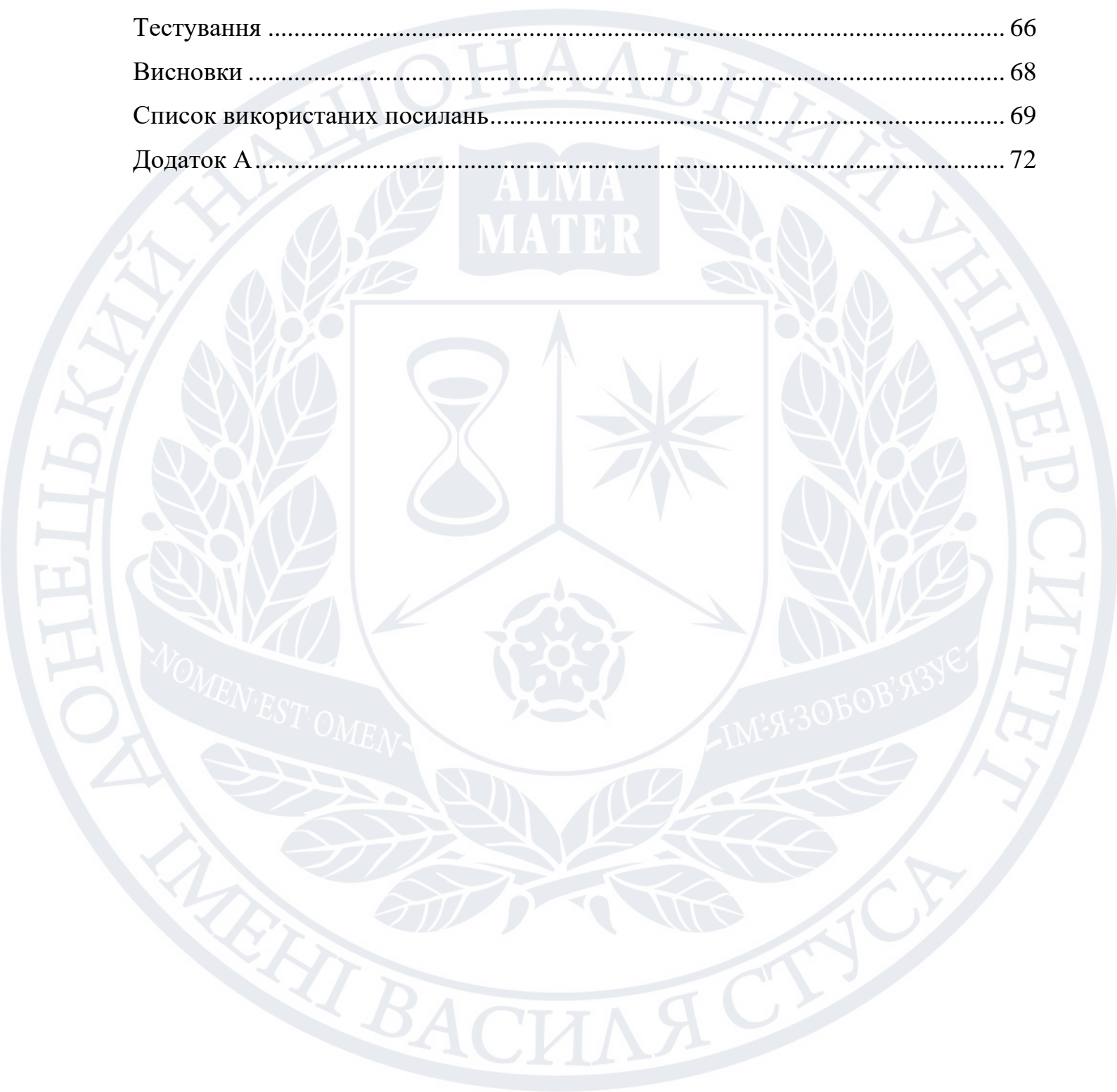
**Predictive assumptions about the development of a research object - adding more functionality:** the ability to add new products.

## ЗІМІСТ

РЕФЕРАТ .....	<b>Помилка! Закладку не визначено.</b>
ВСТУП.....	7
РОЗДІЛ 1 .....	9
Актуальність теми.....	9
Огляд аналогів .....	12
1.1 Zulu Longines.....	12
1.2 NIKE.....	13
1.3 Allbirds .....	14
Постановка задачі.....	16
1.1 Дослідити предметну область і створити модель, побудова ER-моделі; 16	
1.2 Логічне проектування. Побудова і перевірка реляційної моделі даних; 16	
1.3 Проектування і розробка користувальницьких уявлень; .....	16
1.4 Створення сайту для користувача; .....	17
1.5 Аналіз нейромережі для прогнозування кількості продажів;.....	17
1.6 Комунікація сайту з базою даних. ....	17
РОЗДІЛ 2 .....	18
Технічне завдання .....	18
2.1 Найменування програмного продукту. ....	18
2.2 Призначення і область застосування. ....	18
2.3 Основні характеристики. ....	18
2.4 Цільова аудиторія продукту – користувачі.....	19
2.5 Вимоги до функціональних характеристик .....	19
Умови експлуатації .....	20
2.1 Вимоги до складу і параметрів технічних засобів. ....	20
2.2 Вимоги до вихідного коду і мов програмування .....	20
2.3 Вимоги до інформаційної та програмної сумісності .....	20
2.4 Вимоги до вихідного коду і мов програмування .....	20
2.5 Вимоги до програмних засобів, які використовуються програмою .....	20
2.6 Вимоги до захисту інформації .....	21

Графічний інтерфейс.....	22
2.1 Вимоги до зовнішнього інтерфейсу .....	22
2.2 Зовнішній інтерфейс.....	22
РОЗДІЛ 3 .....	27
3.1 Структура продукту.....	27
3.2 Опис бізнес-логіки програми .....	27
3.3 Опис «ядра» програми. ....	28
3.4 Опис предметної області.....	30
3.5 Опис бази даних.....	30
РОЗДІЛ 4 .....	33
Опис засобів реалізації. ....	33
Опис стандартних класів, методів класів, використаних в реалізації. ....	34
Опис авторських класів, методів класів, використаних в реалізації .....	34
Опис методів Data science .....	34
4.1 Виявлення проблеми та підхід до вирішення проблеми .....	35
4.2 Визначте вимоги до даних і методи збору.....	36
4.3 Розуміти дані та підготовку даних .....	37
4.4 Створіть моделі та оцініть їх.....	37
4.5 Розгорніть модель і отримайте відгук.....	38
Проектування.....	39
4.1 Опис етапів вирішення поставленого завдання. ....	39
4.2 Опис архітектури додатку, що розроблюється.....	39
4.3 Опис функціональності додатку, що розроблюється. ....	39
4.4 Опис характеристик обраної ОС, необхідних служб і додатків.....	39
4.5 Обґрунтування вибору технології і мови програмування для розробки системи.....	40
4.6 Опис модулів, що забезпечують функціонування додатка / системи..	40
Огляд систем інтелектуального аналізу даних .....	41
Огляд методів інтелектуального аналізу даних .....	47
Нейромережі та принцип їхньої роботи .....	56
Історія.....	56

Як працюють нейронні мережі.....	57
Нейромережа нелінійної авторегресії.....	58
Нейромережа - ковзного середнього.....	61
Тестування .....	66
Висновки .....	68
Список використаних посилань.....	69
Додаток А.....	72



## ВСТУП

Сьогодні інтернет-магазини – індустрія, що активно розвивається. Перехід традиційної комерції в Інтернет робить її більш гнучкою, оскільки електронна комерція, яка працює з цифровою інформацією в комп'ютерних мережах, полегшує співпрацю людей. Розвиток телекомунікацій створив можливість здійснювати покупки, не виходячи з дому, і при цьому отримувати інформацію про товари в необхідній для прийняття рішення кількості, знайомитися з відгуками інших покупців, висловлювати свою думку тощо. Електронні магазини – це щороку стають найпопулярнішими інструментами для угод, продажу та покупок.

Крім того, кожен інтернет-магазин повинен десь зберігати інформацію про товар. І всю цю інформацію потрібно зберігати в базі даних. База даних є найважливішим компонентом будь-якої інформаційної системи. База даних дозволяє структуровано зберігати великі обсяги інформації конкретної компанії, що значно спрощує ведення звітності та створення архівів. Оптимізовані бази даних значно підвищують продуктивність програм, які будуються на їхньому використанні.

З розвитком інформаційних технологій і підприємництва актуальність використання баз даних значно зросла. Успішні та великі компанії вже не можуть обходитися без чітко структурованої інформаційної системи свого бізнесу. Бази даних, побудовані на SQL Server, відповідають високим вимогам до продуктивності та безпеки.

Тема: Інформаційна система магазину електроніки та побутової техніки.

Предметом є процеси обробки та аналізу інформації інформаційною системою торгівлі електротехнікою та побутовою технікою

Предмет – інтелектуальні моделі та методи обробки та аналізу інформації через інформаційну систему підприємства роздрібною торгівлі електротехнікою та побутовою технікою.

Мета – створення інформаційної системи «Магазин електроніки та побутової техніки» для автоматизації діяльності та забезпечення зберігання,

збору, доставки інформації та прогнозування продажів за допомогою нейронної мережі.





## РОЗДІЛ 1

### ПОСТАНОВКА ЗАДАЧІ, АКТУАЛЬНІСТЬ ТЕМИ ТА ОГЛЯД АНАЛОГІВ, МЕТОДІВ ТА СИСТЕМ

#### Актуальність теми

Простіше кажучи, термін «інтернет-магазин» - це магазин послуг або товарів у глобальній мережі.

Згідно з офіційним визначенням, інтернет-магазин є невід'ємною частиною процесу, відомого як онлайн-шопінг. Покупки в Інтернеті охоплюють процес, за допомогою якого відвідувачі можуть придбати певні продукти або послуги, які пропонуються в інтернет-магазині.

Існує три поширені способи ведення бізнесу в системі онлайн-покупок або інтернет-магазині: B2C (скорочено від Business to Consumer), що означає, що процес роботи інтернет-магазину між покупцем і підприємцем, тобто виробником або послугою, займає постачальників, потім B2B (скорочено від business to business), у якому процес онлайн-торгівлі відбувається між двома підприємцями, і метод B2B2C, який розроблено як свого роду комбінація цих двох методів онлайн-торгівлі, що є результатом скороченої англ. термін «бізнес-бізнес-клієнт», що означає, що онлайн-торгівля відбувається або між двома підприємцями, або між покупцем і підприємцем, за умови, що в цій формі торгівлі між ними також є посередник.

Цей третій спосіб також є найбільш часто використовуваним в сучасних інтернет-магазинах. Справа в тому, що онлайн-шопінг значно полегшує придбання певних товарів і послуг. Однак не варто забувати, що це не пряма торгівля, тут є багато недоліків, пов'язаних здебільшого з питанням якості продукції, що пропонується в інтернет-магазині. Це особливо актуально, наприклад, при купівлі товарів, пов'язаних з одягом, оскільки дуже часто зображення товару, яке користувач може побачити в Інтернеті, не повністю відповідає товару. З цієї причини багато посередників у сфері електронної комерції також вказують на можливість повернення товару на своїх сайтах.

Однак це лише одна з можливих незручностей, з якими можуть зіткнутися клієнти, які користуються послугами інтернет-магазину.

Усі платіжні картки можна використовувати для оплати в більшості інтернет-магазинів. Однак багато в чому це стосується інтернет-магазинів, що працюють на вітчизняному ринку. Для тих, хто є світовим лідером у цьому виді торгівлі, зазвичай потрібна спеціальна картка для онлайн-розрахунків, що значно ускладнює використання послуг подібних інтернет-магазинів вітчизняним користувачам.

Незважаючи на те, що онлайн-шопінг стає все більш поширеним способом купівлі різноманітних товарів і послуг, питання про безпеку такого виду покупки лунають досить часто. Звичайно, оскільки Інтернет відкритий майже для всіх, завжди є можливість для когось зловживати даними користувача. Саме тому майже в кожному інтернет-магазині повинна бути спеціальна програма, яка забезпечує захист інформації та гарантує безпеку даних користувача, а також запобігає несанкціонованим змінам збережених даних користувача, а також можливість перевірити особу користувача. Відповідно, майже всі інтернет-магазини змушені використовувати методи шифрування, щоб краще захистити інформацію, яку їм довіряє користувач.

Крім перерахованих вище типів інтернет-магазинів, існують також: С2В від англійських термінів *consumer to business*, в яких онлайн-торгівля здійснюється покупцем, який запитує у продавця певну послугу або товар, маючи при цьому власні побажання. вказує (наприклад, бюджет, який він має для придбання певного продукту чи послуги), С2С (*Consumer to Consumer*), який є онлайн-магазином, схожим на аукціон, який відбувається між двома користувачами, В2А від англійських слів *бізнес до адміністрування*, основною діяльністю якого є здійснення фінансових операцій між підприємцями, тобто компаніями та численними органами влади, С2А від англійських слів «*consumer*» до «*administration*», що позначає всі необхідні операції, які особа з різними здійснює В2Е (*Business to Employee*) насамперед для працівників державних установ конкретної компанії, і Р2Р (*Peer-to-Peer*), тобто одне й те саме, що є

певним видом спільного спілкування, а не онлайн-комерції ю в класичному розумінні цього слова.

Незважаючи на те, що більшість із цих типів інтернет-магазинів не є широко поширеними, принаймні, не в нашій країні, неодноразово робиться висновок, що запровадження згаданих моделей інтернет-магазинів полегшить бізнес як для окремих осіб, так і для цілих організацій зі значним погіршенням та фізичним станом. державна влада.



## Огляд аналогів

Веб-сайти електронної комерції стали невід’ємною частиною успіху будь-якого бізнесу. Зростання онлайн-продажів на 55% під час пандемії спонукало власників бізнесу перейти від фізичних вітрин до цифрових онлайн-рішень для продажів, допомагаючи розширити охоплення та збільшити продажі.

Незалежно від того, чи ви тільки починаєте роботу, чи маєте відомий бізнес, створення веб-сайту електронної комерції вимагає від вас продумати весь дизайн веб-сайту, від естетики до навігації та доступності. Ці аспекти мають вирішальне значення для того, щоб справити незабутнє враження та забезпечити безперешкодний досвід покупки.

### 1.1 Zulu Longines

Longines є виробником розкішних годинників із багатою історією, і цей сайт електронної комерції віддає шану піонерам авіації, які використовували годинники компанії у своїх кампаніях.

Інтерактивна домашня сторінка показує, як передові функції та довговічність продуктів допомагають авіаторам різних поколінь адаптуватися до часових поясів і подорожувати в екстремальних умовах.

Кращі елементи дизайну:

Анімація запускається прокручуванням. Коли відвідувачі прокручують домашню сторінку, вони бачать рухомі об’єкти, наприклад літаки та карту світу.

Шрифти із засічками. Використання цього шрифту в назвах і цитатах підтримує елегантний імідж бренду Longines.

Інтерактивний повзунок шкали часу. Цей унікальний елемент сторінки історії ілюструє шлях компанії за допомогою ефекту прокручування паралакса.

Ключові висновки:

Розкішним і класичним брендам слід розглянути можливість використання шрифтів із засічками, щоб передати вишуканість і елегантність. Крім того, інтегруйте інтерактивну часову шкалу для відображення історії компанії або дорожньої карти проекту (Рисунок 1.1).



Рисунок 1.1

## 1.2 NIKE

Ще один онлайн-магазин модного одягу, який може стати чудовим орієнтиром, — Nike. Від навігації до параметрів фільтрації, платформа електронної комерції цього бренду добре демонструє принцип «менше — це краще». Він використовує достатньо білого простору та блоків зображень із простим фоном, що дозволяє продуктам виділятися.

Найкращі елементи дизайну:

Мінімалістичний вигляд. Простота веб-сайту електронної комерції Nike усуває відволікання. Цей інтернет-магазин виділяє великі зображення продуктів і підтримує чорно-білу кольорну палітру (Рисунок 1.2).

Виписка гостей. Близько 23% споживачів покинуть свій кошик для покупок, якщо їм доведеться створити обліковий запис користувача. Щоб

уникнути цього, Nike пропонує опцію гостьової перевірки, що дозволяє покупцям робити покупки без входу в систему.

Комплексна довідкова служба. Це оптимізує потік інформації, тому клієнти можуть знаходити відповідну інформацію, не чекаючи служби підтримки.

Ключові висновки:

Продавайте в Інтернеті широкій аудиторії, увімкнувши гостьову оплату для швидшої покупки та меншого відсотка залишення кошика.

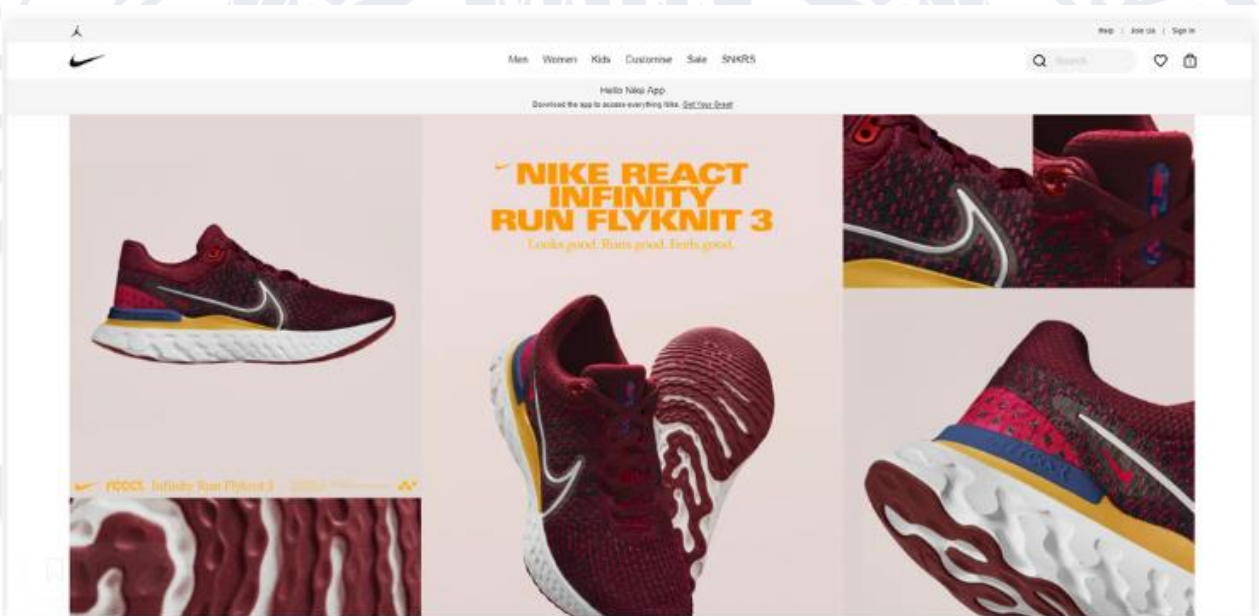


Рисунок 1.2

### 1.3 Allbirds

Онлайн-магазин одягу та взуття Allbirds ефективно відображає його цінності, простоту та екологічність – від заголовка та тексту до редакторських знімків у теплих тонах. Потрібно лише кілька прокручувань на домашній сторінці, щоб прочитати про екологічні практики компанії та основні категорії продуктів, як-от найновіші колекції та колаборації.

Найкращі елементи дизайну:

Інформаційні сторінки товарів. Allbirds надає вичерпну інформацію про продукт, описуючи основні функції та матеріали, а також включає фотографії та відео великим планом, щоб допомогти клієнтам візуалізувати кожен продукт.

Розширена система огляду. Під описом продукції Allbirds наводить відгуки користувачів. Це чудовий спосіб допомогти відвідувачам прийняти рішення про покупку та підвищити довіру клієнтів (Рисунок 1.3).

Рекомендації щодо продукції. Allbirds показує рекомендації щодо пов'язаних товарів на всьому сайті, включаючи сторінки оформлення замовлення, продукту та категорії продукту, щоб залучити клієнтів і збільшити конверсію.

Ключові висновки:

Зробіть ваші сторінки продуктів максимально інформативними та привабливими, щоб допомогти відвідувачам сайту електронної комерції перетворитися на клієнтів. Використання описової мови та цільових ключових слів також може покращити вашу пошукову оптимізацію на сайті (SEO).



Рисунок 1.3

## Постановка задачі

Сьогодні інтернет-магазини – індустрія, що активно розвивається. Перехід традиційної комерції в Інтернет робить її більш гнучкою, оскільки електронна комерція, яка працює з цифровою інформацією в комп'ютерних мережах, полегшує співпрацю людей. Розвиток телекомунікацій створив можливість здійснювати покупки, не виходячи з дому, і при цьому отримувати інформацію про товари в необхідній для прийняття рішення кількості, знайомитися з відгуками інших покупців, висловлювати свою думку тощо. Електронні магазини – це щороку стають найпопулярнішими інструментами для угод, продажу та покупок.

Крім того, кожен інтернет-магазин повинен десь зберігати інформацію про товар. І всю цю інформацію потрібно зберігати в базі даних. База даних є найважливішим компонентом будь-якої інформаційної системи. База даних дозволяє структуровано зберігати великі обсяги інформації конкретної компанії, що значно спрощує ведення звітності та створення архівів. Оптимізовані бази даних значно підвищують продуктивність програм, які будуються на їхньому використанні.

З розвитком інформаційних технологій і підприємництва актуальність використання баз даних значно зростає. Успішні та великі компанії вже не можуть обходитися без чітко структурованої інформаційної системи свого бізнесу. Бази даних, побудовані на SQL Server, відповідають високим вимогам до продуктивності та безпеки.

Тема: Інформаційна система магазину електроніки та побутової техніки.

Метою даного проекту є:

- 1.1 Дослідити предметну область і створити модель, побудова ER-моделі;
- 1.2 Логічне проектування. Побудова і перевірка реляційної моделі даних;
- 1.3 Проектування і розробка користувацьких уявлень;



- 1.4 Створення сайту для користувача;
- 1.5 Аналіз нейромережі для прогнозування кількості продажів;
- 1.6 Комунікація сайту з базою даних.



## РОЗДІЛ 2

### ТЕХНІЧНЕ ЗАВДАННЯ, УМОВИ ЕКСПЛУАТАЦІЇ, ГРАФІЧНИЙ ІНТЕРФЕЙС

#### Технічне завдання

##### 2.1 Найменування програмного продукту.

Інтернет-магазин електроніки та цифрової техніки

##### 2.2 Призначення і область застосування.

Інтернет-магазин призначений для здійснення покупок в Інтернеті. Користувач має можливість самостійно отримати необхідну інформацію про товар, переглянути фотографії, прочитати опис, порівняти ціни та вибрати найкращий для себе товар. Це дуже зручно.

##### 2.3 Основні характеристики.

Відмінними рисами інтернет-магазинів є те, що вони можуть пропонувати набагато більшу кількість товарів і послуг, ніж звичайні магазини, і надавати споживачам набагато більший обсяг інформації, необхідної для прийняття рішення про покупку. Завдяки використанню Інтернет-технологій також можлива персоналізація підходу до споживачів з урахуванням попередніх відвідувань магазину та здійснених у ньому покупок, а також використання Інтернет-магазину як ефективного методу дослідження ринку (опитування, конференції клієнтів тощо).

Інтернет-магазини потребують значно менших витрат на утримання та організацію роботи, оскільки мають значно обмеженішу матеріально-технічну базу (будівлі, споруди, приміщення) та кількість обслуговуючого персоналу.

#### 2.4 Цільова аудиторія продукту – користувачі

Сайт матиме практичний і простий інтерфейс, який дозволить усім користувачам швидко і легко здійснювати покупки. Якщо вам потрібно купити товар, ви можете його швидко замовити. Сайт призначений для людей різного віку та для користувачів з першою комп'ютерною освітою.

#### 2.5 Вимоги до функціональних характеристик

Цей інтернет-магазин дозволяє користувачеві оформити замовлення, а також покупець може дізнатися характеристики обраного товару.



## Умови експлуатації

### 2.1 Вимоги до складу і параметрів технічних засобів.

В склад технічних засобів повинен входити IBM-сумісний персональний комп'ютер, що включає в себе:

- персональний комп'ютер (ПЕОМ), що включає в себе:
- процесор, не менше intel core 2 duo;
- оперативну пам'ять, об'ємом не менше 2 Гігабайт;
- графічний адаптер nvidia geforce gt 630;
- маніпулятор типу "миша(touchpad)" та "клавіатура";

### 2.2 Вимоги до вихідного коду і мов програмування

Основною мовою програмування є PHP, Java script і використання бази даних MySQL для зберігання результатів програми.

### 2.3 Вимоги до інформаційної та програмної сумісності

Інтерфейс має бути інтуїтивно зрозумілим, щоб зробити замовлення швидко та легко. Для стабільної роботи інтернет-магазину також потрібен IBM-сумісний ПК або мобільний пристрій.

### 2.4 Вимоги до вихідного коду і мов програмування

Базова мова програмування PHP 5.5, MySQL 5.1, Java script. Потрібне використання компонентної технології та можливість динамічного компонування бібліотек.

### 2.5 Вимоги до програмних засобів, які використовуються програмою

Інтернет магазин повинен працювати в браузері Google Chrome, Mozilla Firefox, Opera, Safari.

## 2.6 Вимоги до захисту інформації

Вся інформація яку користувач залишає на сайті записується в базу даних, і ніхто крім власника сайту не зможе нею розпоряджатися.



## Графічний інтерфейс

### 2.1 Вимоги до зовнішнього інтерфейсу

Інтерфейс користувача повинен бути зрозумілим для користувачів. Можливість підтримки всіх розширень повинна адаптуватися до кожного розширення екрана.

### 2.2 Зовнішній інтерфейс

Графічний інтерфейс повинен бути інтуїтивно зрозумілий та легкий у використанні новими користувачами. Головна сторінка буде містити 6 головних кнопок: Каталог, Доставка, Оплата, Гарантія, Контакти, Вхід (Рисунок 2.3.1).

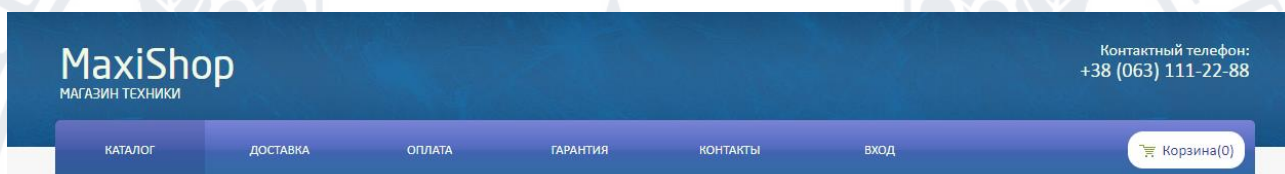


Рисунок 2.3.1

При натисканні на кнопку Каталог, користувачу буде показано список доступної продукції яку він зможе замовити (Рисунок 2.3.2).

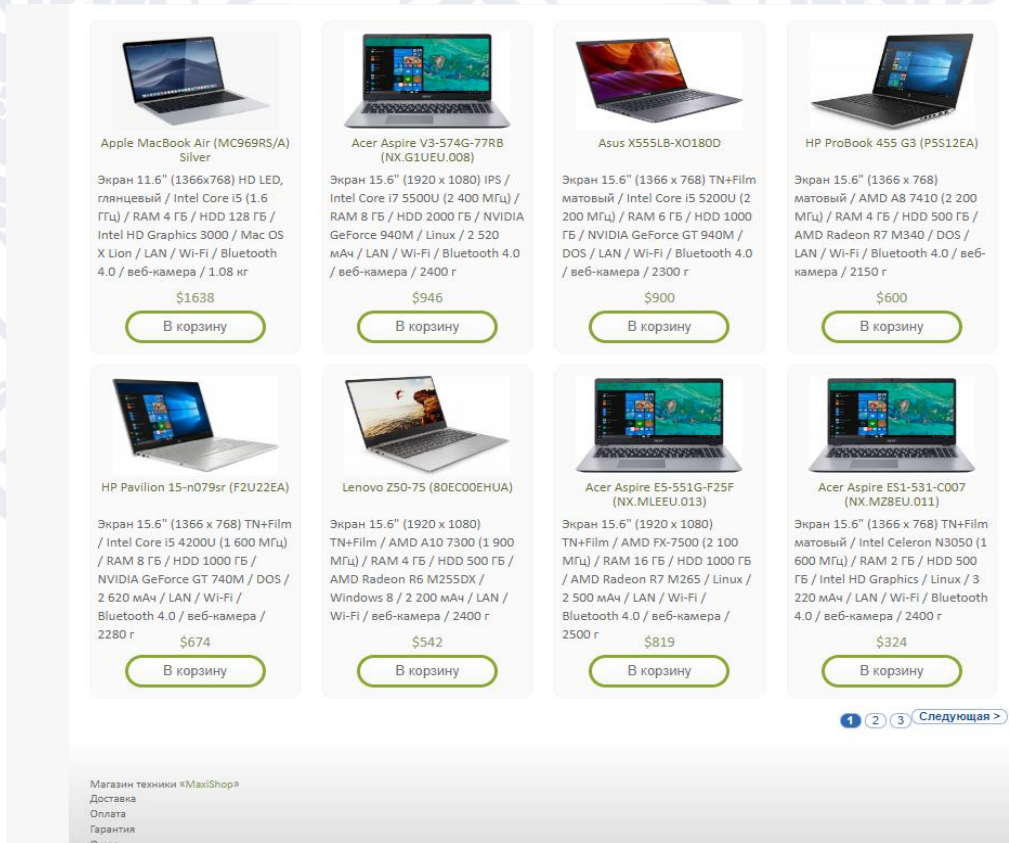


Рисунок 2.3.2

При натисканні на кнопку Доставка, користувачу буде показано інформацію про умови доставки (Рисунок 2.3.3).

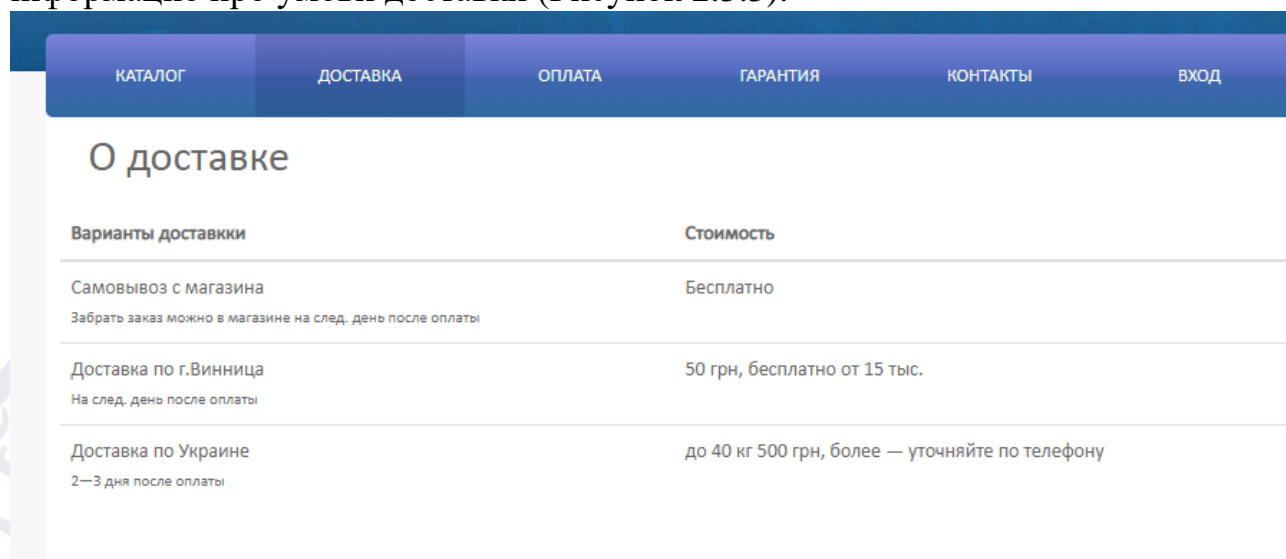


Рисунок 2.3.3

При натисканні на кнопку Оплата, користувачу буде показано інформацію про види оплати (Рисунок 2.3.4).

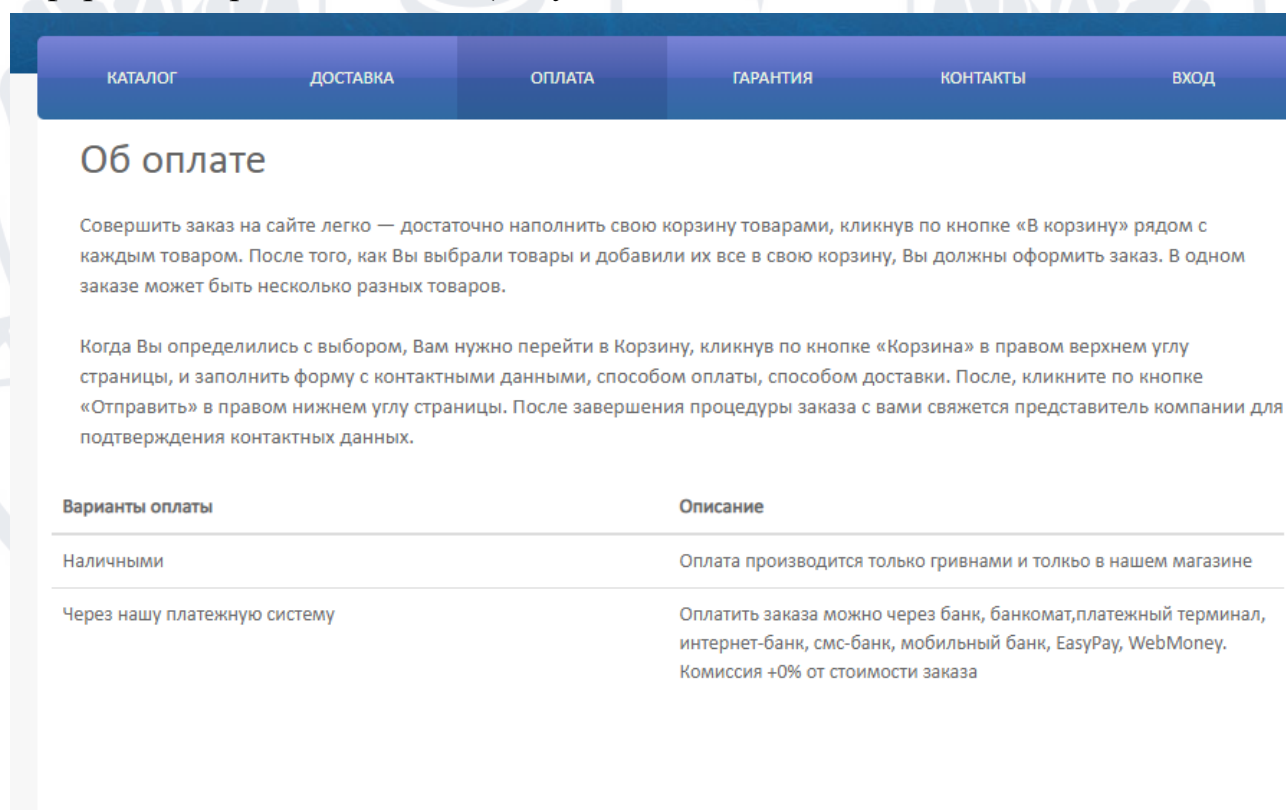


Рисунок 2.3.4

При натисканні на кнопку Гарантія, користувачу буде показано інформацію про тривалість гарантії на товар (Рисунок 2.3.5).

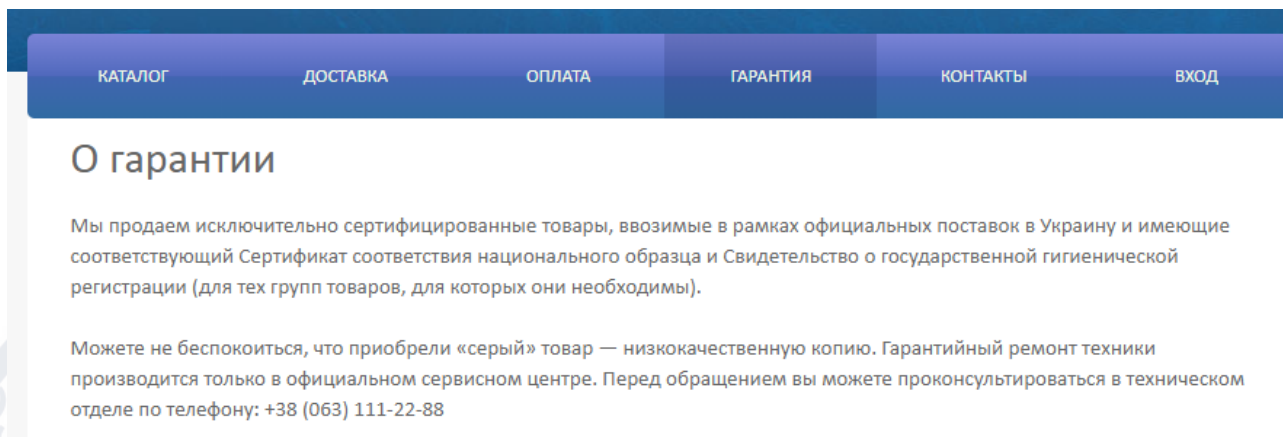


Рисунок 2.3.5

При натисканні на кнопку Контакти, користувачу буде показано інформацію про місцезнаходження магазину, та їхні номери телефонів (Рисунок 2.3.6).

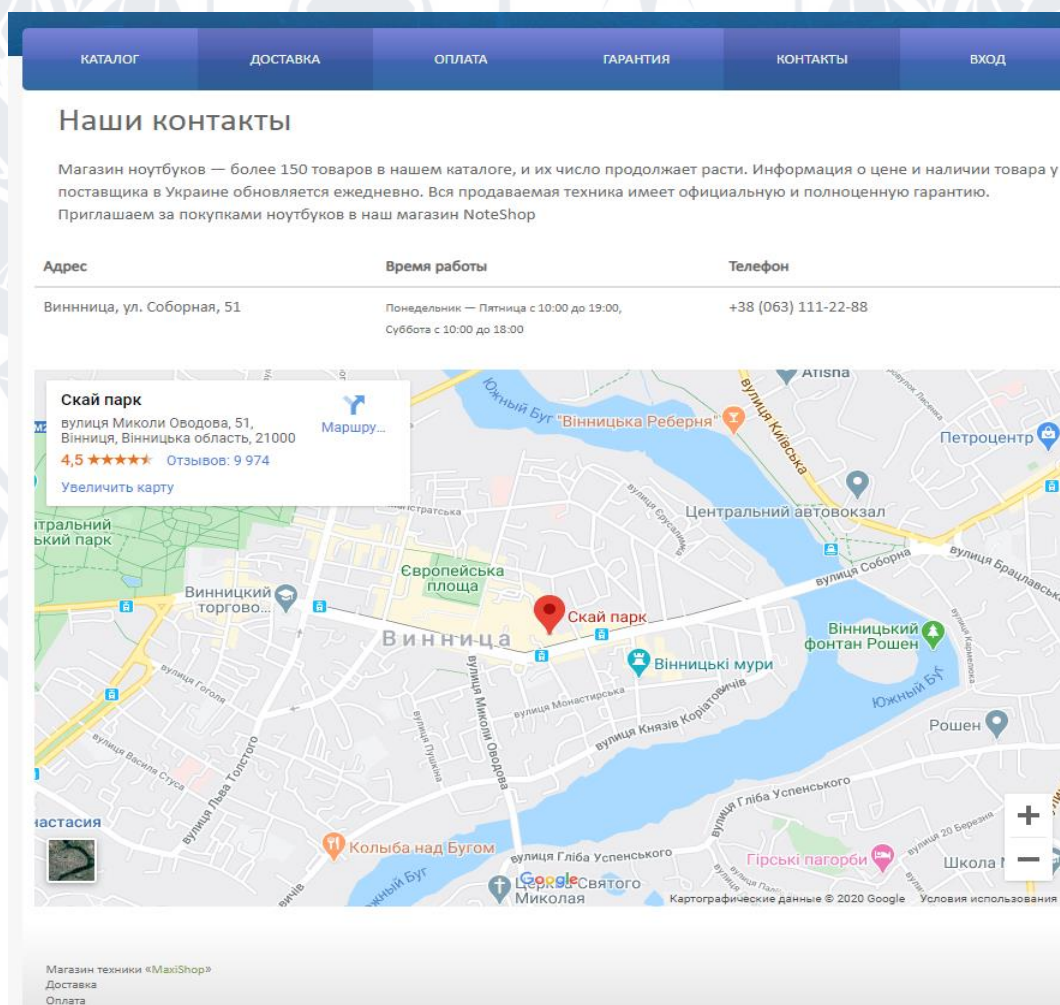


Рисунок 2.3.6



При натисканні на кнопку Вхід, власнику магазину можна буде побачити кількість заказів які зробили клієнти та добавляти нові товари (Рисунок 2.3.7).

Рисунок 2.3.7

The image displays two screenshots of a web application interface. The top screenshot shows a navigation bar with 'КАТАЛОГ', 'НОВЫЙ ТОВАР', 'ЗАКАЗЫ', and 'ВЫХОД (admin)'. Below it are two product listings. The first listing is for 'Макс, 1234567' with a price of \$946, listing 'Acer Aspire V3-574G-77RB (NX.G1UEU.008)'. The second listing is for 'Макс, 2211' with a price of \$540, listing 'Lenovo G50-70 (59433722)'. Both listings include payment and delivery options and a 'Удалить' button. The bottom screenshot shows the 'Добавление нового товара' form with fields for Brand (Acer), Model, Price, Image path, Screen diagonal (11.6"), Processor (Intel Celeron), RAM (1 GB), Hard disk, Video card, OS (Windows 7), and a 'Создать' button.

КАТАЛОГ    НОВЫЙ ТОВАР    ЗАКАЗЫ    ВЫХОД (admin)

Макс, 1234567  
Способ оплаты: Наличными  
Способ доставки: Самовывоз  
Удалить

Список товаров  
Acer Aspire V3-574G-77RB (NX.G1UEU.008) - \$946

Макс, 2211  
Способ оплаты: Наличными  
Способ доставки: Самовывоз  
Удалить

Список товаров  
Lenovo G50-70 (59433722) - \$540

КАТАЛОГ    НОВЫЙ ТОВАР    ЗАКАЗЫ    ВЫХОД (admin)

### Добавление нового товара

Бренд: Acer

Модель:

Цена:

Путь к картинке:  Файл не выбран

Диагональ экрана: 11.6"

Характеристики экрана:

Тип процессора: Intel Celeron

Процессор:

Оперативная память: 1 GB

Жесткий диск:

Видеокарта:

Тип ОС: Windows 7

Операционная система:

Дополнительное описание:

Создать

Рисунок 2.3.7

При натисканні на кнопку Кошик, користувачу можна буде побачити кількість товарів які він додавив у кошик (Рисунок 2.3.8).



Рисунок 2.3.8

## РОЗДІЛ 3

### СТРУКТУРА ПРОДУКТУ, БІЗНЕС-ЛОГІКА, ЯДРО ПРОГРАМИ, ПРЕДМЕТНА ОБЛАСТЬ

#### 3.1 Структура продукту

Програма складається з: графічного інтерфейсу користувача та описової частини програми: підключення модулів, оголошення змінних, описи процедур і функцій.

#### 3.2 Опис бізнес-логіки програми

Програма використовує функції, які з'єднують наші сторінки та ті, які мають спілкуватися з нашою базою даних. Головне вікно містить 6 кнопок: Каталог, Доставка, Оплата, Гарантія, Контакти, Вхід.

Після натискання кнопки «Каталог» користувачеві відкривається список доступних товарів для замовлення.

При натисканні кнопки доставки користувачеві відображається інформація про умови доставки.

При натисканні на кнопку Оплата користувачеві відображається інформація про способи оплати.

Натиснувши кнопку Гарантія, користувачеві буде представлена інформація про тривалість гарантії на товар.

При натисканні кнопки «Зв'язатися» користувач побачить інформацію про місцезнаходження магазину та його номери телефонів.

Натиснувши кнопку «Вхід», власник магазину може переглянути кількість замовлень від клієнтів і додати нові товари.

Натиснувши кнопку кошика для покупок, користувач може побачити кількість товарів, які він додав у кошик.

### 3.3 Опис «ядра» програми.

Функція використовується для запису введення користувача в базу даних MySQL і функції, які спілкуються з користувачем. Для реалізації основного функціоналу продукту використовується фреймворк Yii2.

Yii — це високопродуктивний компонентний PHP-фреймворк, призначений для швидкої розробки сучасних веб-додатків. Слово Yii (вимовляється Yi [ji:]) китайською мовою означає «простий і розвивається». Yii також можна використовувати для Yes It Is!

Крім того, це універсальний фреймворк, який можна використовувати в усіх типах веб-додатків. Завдяки своїй компонентній структурі та відмінній підтримці кешування фреймворк особливо підходить для розробки таких великих проєктів, як портали, форуми, CMS, магазини або додатки RESTful.

Програми Yii організовані відповідно до шаблону проєктування Model-View-Controller (MVC). Моделі представляють дані, бізнес-логіку та бізнес-правила; Представництва відповідають за відображення інформації, в тому числі на основі даних, отриманих від моделей; Контролери приймають вхідні дані від користувача і перетворюють їх у формат, зрозумілий моделям і командам, а також відповідають за відображення потрібного представлення.

Окрім MVC, додатки Yii також мають такі сутності:

Сценарії введення: це сценарії PHP, до яких кінцевий користувач програми має прямий доступ. Ви несете відповідальність за ініціювання та обробку вхідного запиту;

Програми: це глобально доступні об'єкти, які забезпечують коректну роботу різних компонентів програми та їх координацію для обробки запиту;

Компоненти програми: це об'єкти, зареєстровані в програмі, які забезпечують різні способи обробки поточного запиту;

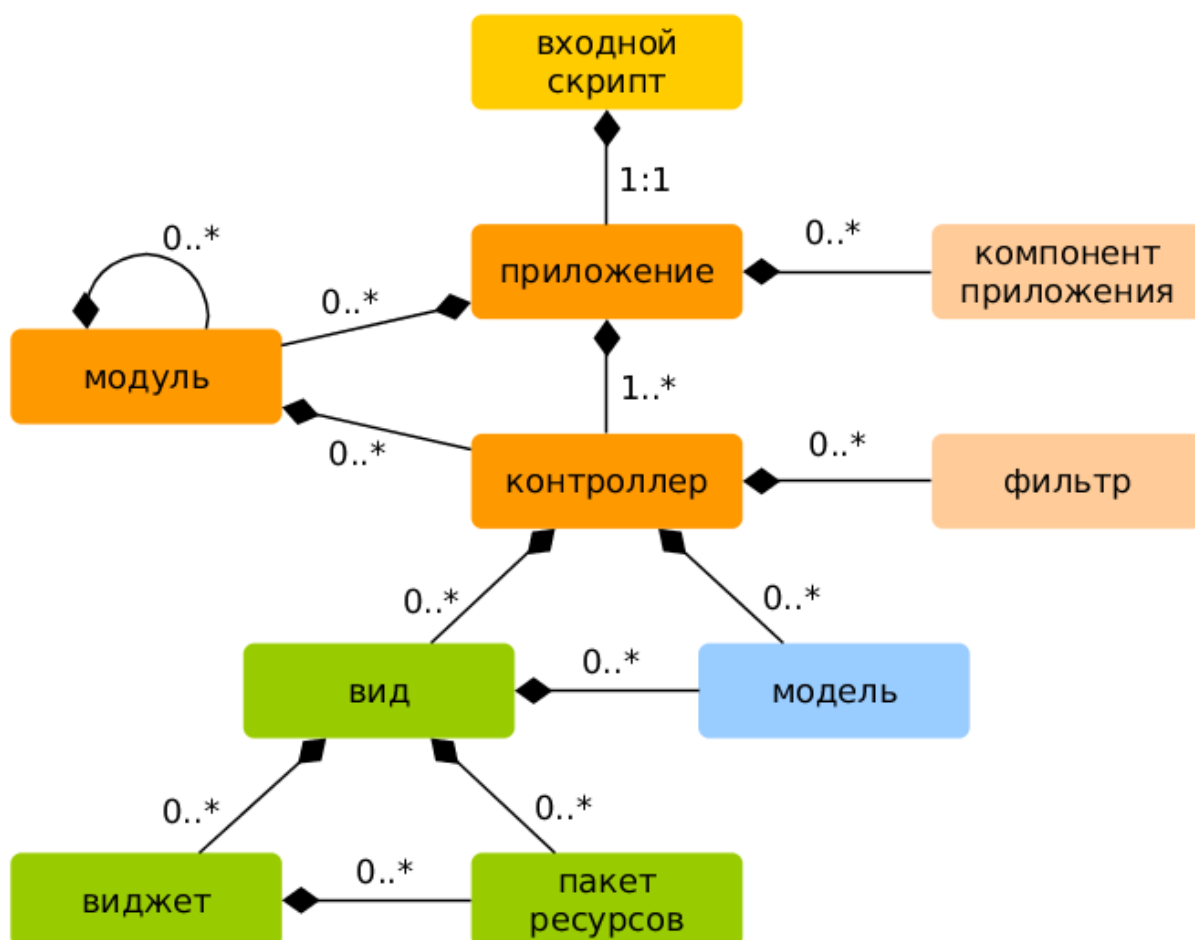
Модулі: це окремі пакети, які містять усі інструменти для MVC. Програма може бути організована за кількома модулями;

Фільтр: це код, який необхідно виконати до та після обробки запиту відповідальними особами;

Віджети: це об'єкти, які можуть містити перегляди. Вони можуть містити різну логіку і використовуватися в різних представленнях.

Нижче на діаграмі представлена структурна схема програми (Рисунок 3.3).

Рисунок 3.3



### 3.4 Опис предметної області

На сьогоднішній день електронна комерція є галуззю торгівлі, що активно розвивається. Перехід традиційної комерції в Інтернет робить її більш гнучкою, оскільки електронна комерція, яка працює з цифровою інформацією в комп'ютерних мережах, полегшує співпрацю людей. Розвиток телекомунікацій створив можливість здійснювати покупки, не виходячи з дому, і при цьому отримувати інформацію про товари в необхідній для прийняття рішення кількості, знайомитися з відгуками інших покупців, висловлювати свою думку тощо. Електронні магазини – це щороку стають найпопулярнішими інструментами для угод, продажу та покупок.

### 3.5 Опис бази даних

База даних — це сукупність інформації про об'єкти, процеси, події або явища реального світу, пов'язані з певною темою чи завданням, організована таким чином, щоб цю колекцію було зручно представити як цілком, так і частково. Реляційна база даних — це набір пов'язаних таблиць, кожна з яких містить інформацію про об'єкти певного типу. Кожен рядок таблиці містить дані про об'єкт (наприклад, клієнт, товар, документи), а стовпці таблиці містять різні характеристики цих об'єктів – атрибути (наприклад, імена та адреси клієнтів, марки та ціни товарів). Рядки таблиці називаються записами, всі записи мають однакову структуру - складаються з полів, які зберігають атрибути об'єкта. Кожне поле в записі містить характеристику об'єкта та має чітко визначений тип даних (наприклад, текстовий рядок, число, дата). Усі записи мають однакові поля, тільки вони містять різні значення атрибутів.

Будь-яка СУБД дозволяє виконувати чотири найпростіші операції з даними:

- додати в таблицю одну або кілька записів;
- видалити з таблиці одну або кілька записів;
- оновити значення деяких полів в одній або декількох записах;
- знайти одну або кілька записів, що задовольняють заданій умові.

Для виконання цих операцій використовується механізм запиту. Результатом виконання запиту є або набір записів, відібраних за певними критеріями, або зміна таблиць.

Для розробки програми також була обрана реляційна СУБД MySQL, один із найпоширеніших засобів створення додатків баз даних для бізнес-додатків. Програми такого типу на даний момент визнані стандартними, оскільки є найбільш поширеними, зручними для користувачів і не потребують тривалого навчання. Системі СУБД MySQL було віддано перевагу, оскільки вона дозволяє програмісту розробляти інтерфейс користувача дуже швидко, легко та комфортно.

Вдалий дизайн бази даних забезпечує легке обслуговування. Дані повинні зберігатися в таблицях, і кожна таблиця повинна містити інформацію одного типу, напр. В. Відомості про постачальника. Тоді достатньо ввести певні дані, напр. адресу лише в одному місці, щоб оновлена інформація відображалася в усій базі даних.

Однією з найскладніших фаз у процесі проектування бази даних є розробка таблиць, оскільки результати, які створена база даних (звіти, вихідні форми тощо), не завжди дають повне уявлення про структуру таблиці.

Предметною областю бази даних є діяльність у сфері виробництва електроніки та побутової техніки. Описується основний вид діяльності: продаж електроніки та побутової техніки.

Дана база даних передбачає врахування продажів товару, отже, повинні фіксуватися такі дані:

- Дата продажу;
- кількість;
- опис товару;
- знижка на товар;
- вартість продажу.





## РОЗДІЛ 4

### РЕАЛІЗАЦІЯ РІШЕННЯ, ПРОЕКТУВАННЯ

Опис засобів реалізації.

Мова програмування, використана при реалізації додатку, PHP версії 5.5, Java script, MySQL.

PHP — це сценарна мова програмування, яка використовується для розробки програм і веб-сайтів. В даний час він підтримується багатьма хостинговими компаніями і є однією з провідних мов програмування, які використовуються для написання динамічних веб-сторінок.

JavaScript — це мова програмування, яка додає інтерактивності вашому веб-сайту (наприклад, ігри, відгуки під час натискання кнопок або введення даних у форми, динамічні стилі, анімація). Ця стаття допоможе вам розпочати роботу з цією захоплюючою мовою та дасть уявлення про те, на що вона здатна.

Для розробки бази даних використовувався Open Server (phpMyAdmin). phpMyAdmin допомагає вам керувати сервером MySQL, виконувати команди SQL і переглядати вміст таблиць і баз даних.

MySQL - це реляційна система управління базами даних, що містить понад шість тисяч об'єктів. Програма працює як сервер і пропонує «багатокористувацький доступ» до об'єктів бази даних.

Однією з найпоширеніших систем керування базами даних MySQL з 2010 року належить корпорації Oracle. Продукт розповсюджується за Загальною публічною ліцензією GNU (ліцензія на безкоштовне програмне забезпечення) або за власною комерційною ліцензією.

MySQL розроблена шведською компанією MySQL AB. У 2008 році MySQL AB купила американська компанія Sun Microsystems. Починаючи з квітня 2009 року корпорація Oracle почала придбання Sun Microsystems. Процес злиття двох компаній завершився в січні 2010 року.

MySQL часто використовується в безплатних програмних проектах, яким потрібна система управління базами даних з можливістю повнотекстового

пошуку, таких як програмне забезпечення WordPress, програмне забезпечення phpBB і багато інших.

Також використовується для крупних продуктів таких як Wikipedia Google Facebook.

Опис стандартних класів, методів класів, використаних в реалізації.

Бібліотека, якою я користуюся, не потребує жодних внесків у формі класів для написання цієї програми. Однак існують допоміжні функції, які полегшують написання коду.

Опис авторських класів, методів класів, використаних в реалізації

- Add – Записує дані в базу даних.
- Edit – Домагає редагувати дані в базі даних.
- SELECT - запити до БД.

Опис методів Data science

Наука про дані — це величезна галузь, і це не лише створення моделей машинного навчання чи прогнозування результатів для різних сценаріїв, які можуть виникнути під час роботи з даними. Спеціаліст із обробки даних має багато різних обов'язків і може нести відповідальність за одну чи декілька з наведених нижче дій:

- Ділове розуміння
- Розуміння даних
- Підготовка даних
- Моделювання
- Оцінка
- Розгортання

Кожне з цих завдань взаємопов'язане та допомагає виконувати інші ролі в методології науки про дані (Рисунок 4.4). Ось на чому я збираюся зосередитися в цьому блозі.

Що таке методологія науки про дані, яку ви можете запитати? Це методологія, яка використовується науковцями з обробки даних, щоб знайти рішення конкретної проблеми та використати залучені дані.

Методологія науки про дані складається з п'яти ітераційних кроків, і в межах цих кроків у вас є підкатегорії, які допомагають дослідникам даних.

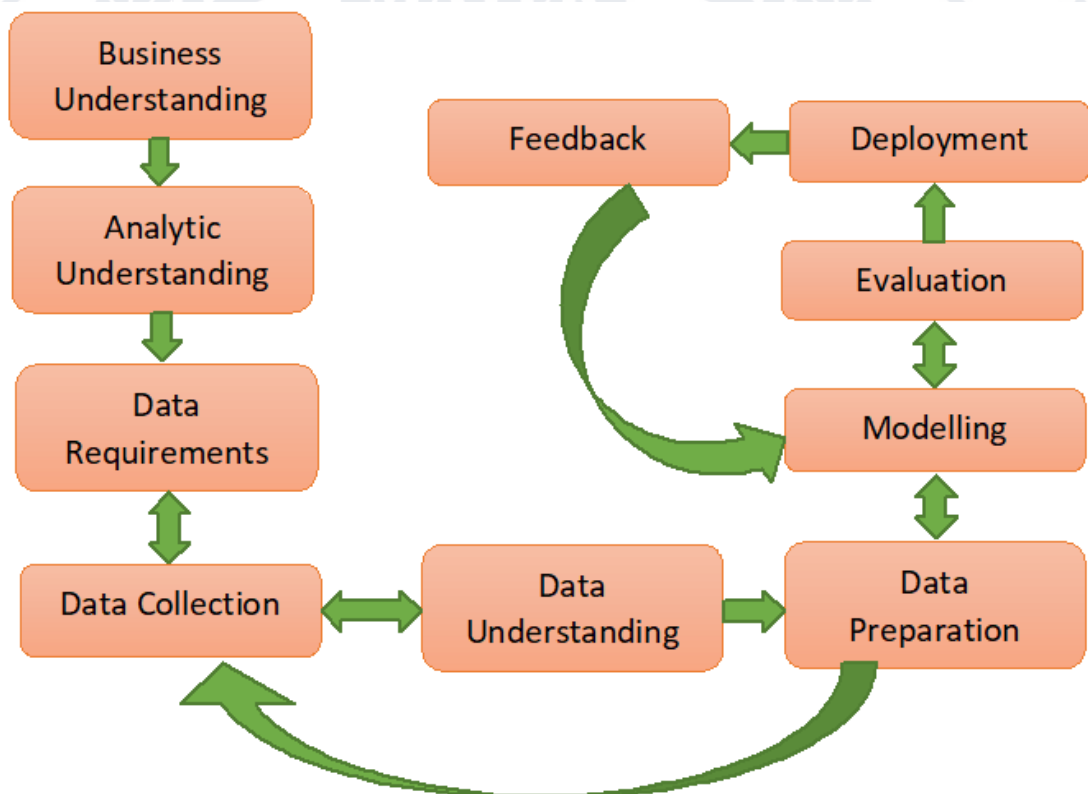


Рисунок 4.4

#### 4.1 Виявлення проблеми та підхід до вирішення проблеми

По-перше, подивіться, що ви намагаєтеся вирішити як компанія. Перший крок до цього — зрозуміти компанію: чим вона займається, який внесок вона робить, які прибутки компанії та які інші фактори ведуть до прибутку. З цією інформацією ви отримаєте чітке уявлення про бізнес.

Під час аудиту ви також висвітлюєте проблемні місця компанії. Тепер, коли ви маєте цю інформацію, переходьте до наступного кроку, який є

аналітичним підходом. На цьому етапі ви зазвичай обираєте відповідний аналітичний підхід для вирішення проблеми. Зазвичай це починається з описового аналізу, під час якого ви ставите запитання на зразок «що пішло не так». Потім ми переходимо до діагностичного аналізу, також відомого як статистичний аналіз, який більше стосується того, «чому це пішло не так». Далі ви перемкнетеся на підхід прогнозного аналізу, де ви враховуєте наслідки, якщо та сама тенденція збережеться. Це більше питання "що буде далі?" Нарешті, ви переходите до аналітичного підходу, де дивитесь, як вирішити проблему «Що мені робити?».

За допомогою цих двох кроків ви виконали перший крок у методології, дізналися більше про проблему в деталях і зрозуміли бізнес-конвеєр.

#### 4.2 Визначте вимоги до даних і методи збору

Вимоги до даних і збір даних знаходяться на другому етапі методології. Тут ви визначаєте дані, необхідні для вирішення проблеми, з якою стикається компанія. На цьому кроці ви повинні розглянути формат даних, які збираються, а також дані, які збираються. Наприклад, фахівець з обробки даних, який працює в банківському секторі та прогнозує кредитоспроможність, повинен зібрати дані, які включають місячний дохід, професію, вік і роки проживання в країні, щоб передбачити, чи має особа право на кредит.

Після узгодження вимог до даних наступним кроком є збір даних і те, як ці дані збиратимуться. Наприклад, це може бути збір даних прямих трансляцій із сервісів потокової аналітики, які надають вам необроблені дані для роботи. Сама компанія може надати набір даних або використовувати веб-колекцію для збору текстових даних. Більшість даних на цьому етапі є необробленими та потребують очищення та розуміння даних.

### 4.3 Розуміти дані та підготовку даних

Коли дані зібрані, почніть їх розуміти та готувати, що є третім кроком у методології. Коли справа доходить до розуміння даних, спочатку запитайте себе: «Чи допоможуть ці дані вирішити мою проблему?», а потім почніть пошуковий аналіз даних (EDA). В EDA ви створюєте візуалізації та діаграми, щоб знайти зв'язки між зібраними даними, зв'язки між стовпцями. Наприклад, деякі компанії можуть застосовувати зарплату окремої особи на основі її досвіду роботи в галузі.

Коли ви зрозумієте ці дані, наступним кроком буде їх підготовка самостійно, що займає найбільше часу в рамках методології. Тут ви виконуєте багато операцій з даними, напр. В. Зменшення відсутніх даних і стовпців, покращення читабельності даних, заміна та анонімізація даних, а також видалення непотрібних даних. Крім того, після маніпулювання даними ви також виконуєте розробку функцій, яка передбачає виконання операцій між стовпцями для отримання нових значущих даних, які генерують нове розуміння даних. Сама розробка функції є великою, оскільки ви виконуєте кілька операцій, які можуть включати синус, косинус, тан, а також додавання та порівняння даних у кожному стовпці набору даних.

Коли дані підготовлені та готові до використання, можна переходити до моделювання та перевірки моделі.

### 4.4 Створюйте моделі та оцінюйте їх

Побудова та оцінка моделей є наступним кроком у методології. Існує два типи моделювання: описове моделювання та прогнозне моделювання. Описове моделювання підкаже вам, яка послуга може бути цікавою для конкретної особи, і базується на системах рекомендацій і алгоритмах кластеризації. Прогнозне моделювання включає прогнозування майбутньої вартості на основі ключових вхідних даних і спирається на лінійну або логістичну регресію та алгоритми класифікації.

Оцінка моделювання стосується питання «чи робить модель те, для чого вона була розроблена», і ви також розглядаєте питання пояснення моделі. Коли ви оцінюєте модель на етапі діагностики, ви можете побачити, чи виконує модель те, для чого вона була розроблена. Якщо модель є прогностичною, ви можете використовувати дерева рішень, якщо результат збігається з оригінальним проектом. Якщо модель є описовою, ви можете використовувати тестовий набір даних, що складається з відомих результатів, щоб удосконалити модель відповідно.

Нарешті, на етапі оцінки ви також працюєте з пояснюваністю моделі, що дозволяє пояснити, як модель прийшла до цього рішення; які параметри враховувалися; і які показники відповідності моделі, точності та похибки були враховані для моделі для створення такого прогнозу.

Після створення та оцінки моделей ви переходите до розгортання моделі, збираєте відгуки та вдосконалюєте модель на основі відгуків.

#### 4.5 Розгорніть модель і отримайте відгук

Коли ви впевнені в моделі, переходьте до фази розгортання, де ви розгортаєте модель для обмеженої кількості людей, щоб отримати відгук щодо прогнозів і перевірити, чи підходить модель для загального використання.

Після розгортання моделі ви отримаєте відгук від користувачів моделі. Користувачі надають додаткові відгуки про модель, а потім ви можете покращити, оцінити та повторно розгорнути модель. Він повертається до фази зворотного зв'язку та повторює цей процес, доки ви не отримаєте остаточну модель. Тоді ви зможете розгорнути його для великої кількості користувачів.

Як бачите, ця методологія дозволяє отримати глибоке розуміння вашого конвеєра даних. Крім того, це гнучка методологія. Отже, витрачаючи більше часу на кожен крок, ви отримуєте глибше розуміння та впровадження моделі, що може допомогти вам досягти кращих результатів у науці про дані.

## Проектування

### 4.1 Опис етапів вирішення поставленого завдання.

Основним завданням даної роботи є створення інтернет сайту, а також БД для нього.

### 4.2 Опис архітектури додатку, що розроблюється.

Після запуску програми з'явиться сторінка з відповідними посиланнями та продуктами, включеними в нашу базу даних. Головна сторінка містить 6 кнопок: Каталог, Доставка, Оплата, Гарантія, Контакти, Вхід. Отже, якщо ви натиснете будь-яку кнопку, буде виконана відповідна функція, переміщаючись по сторінках.

Програма виглядає як веб-сторінка. Де розташовані відповідні поля для введення або вибору даних і кнопки, які дозволяють виконувати різні маніпуляції на сторінці.

### 4.3 Опис функціональності додатку, що розроблюється.

Головною особливістю цього продукту є те, що кожен товар користувач може вибрати і купити самостійно. Коли ви вибираєте товар, він автоматично додається в кошик і починається процес замовлення. Після успішного проходження всі дані заносяться в базу і власник сайту може легко обробляти замовлення.

### 4.4 Опис характеристик обраної ОС, необхідних служб і додатків.

Для розробки додатку використовувалась операційна система Windows

В склад технічних засобів повинен входити ІВМ-сумісний персональний комп'ютер, що включає в себе:

- Процесор або система на кристалі(SoC) з частотою 1 ГГц чи вищою
- Оперативну пам'ять, 4 ГБ (для 32-розрядної версії) або 8 ГБ (для 64розрядної версії)

- HDD, простір на жорсткому диску 128 ГБ (для 32-розрядної версії) або 256 ГБ (для 64-розрядної версії)
  - Відеокарта: DirectX10 або сучаснішої версії з драйвером WDDM 1.0
  - Дисплей: 1920x1080 і вище
  - Маніпулятори: миша, клавіатура.
- Необхідний додаток **Open Server**.

#### 4.5 Обґрунтування вибору технології і мови програмування для розробки системи.

Для написання коду була обрана мова PHP. PHP — популярна мова програмування загального призначення з відкритим кодом. PHP спеціально розроблений для веб-розробки, і його код можна вставляти безпосередньо в HTML.

Для розробки бази даних використовувався Open Server (phpMyAdmin). phpMyAdmin допомагає вам керувати сервером MySQL, виконувати команди SQL і переглядати вміст таблиць і баз даних.

#### 4.6 Опис модулів, що забезпечують функціонування додатка / системи.

При розробці використовувалися такі модулі: JavaScript (jQuery, Ajax), Bootstrap.

jQuery — популярна бібліотека JavaScript. Він був створений Джоном Резігом у 2006 році, щоб полегшити розробникам використання JavaScript на веб-сайтах. Це не окрема мова програмування і працює в поєднанні з JavaScript.

Написання коду може стати виснажливим, особливо якщо він охоплює багато рядків. jQuery конденсує кілька рядків коду в одну функцію, тож вам не потрібно переписувати цілі блоки коду, щоб виконати одне завдання.

AJAX — це асинхронний JavaScript і XML. Це набір методів веб-розробки, які дозволяють веб-додаткам працювати асинхронно та обробляти всі запити до сервера у фоновому режимі.



При використанні AJAX немає необхідності щоразу оновлювати всю сторінку, оскільки оновлюється лише певна її частина. Це набагато зручніше, оскільки не потрібно довго чекати, і економніше, оскільки не у всіх є безлімітний інтернет.

Bootstrap — це безкоштовний набір інструментів для створення веб-сайтів і веб-додатків. Містить шаблони дизайну HTML і CSS для типографіки, веб-форм, кнопок, міток, навігаційних блоків та інших компонентів веб-інтерфейсу, включаючи розширення JavaScript.

### Огляд систем інтелектуального аналізу даних

Системи IAD можна класифікувати за різними підходами. Давайте розглянемо деякі з найпоширеніших з них. Класифікацію можна здійснити за типами даних, що обробляються: просторові, мультимедійні, часові ряди, текстові, всесвітня мережа та інші. Класифікація систем IAD відповідно до моделей даних базується на задіяних моделях даних. Існує багато моделей аналізу даних (реляційна, об'єктно-орієнтована, об'єктно-орієнтована, сховище даних, транзакційна база даних, ієрархічна), кожна з яких використовує різні типи даних. Класифікація на основі виявлених типів знань базується на винайдених типах знань або функціях інтелектуального аналізу даних, таких як особливості, розрізнення, асоціації, класифікація, кластеризація та інші. Як правило, багато систем є досить складними системами і охоплюють декілька функцій IAD одночасно [13]. Систему IAD можна класифікувати відповідно до використовуваного методу: машинне навчання, нейронні мережі, генетичні алгоритми, статистика, візуалізація на базі даних або сховища даних тощо.

Системи інтелектуального аналізу даних також можна класифікувати за ступенем взаємодії з користувачами, які беруть участь у процесі. До них відносяться такі системи: запитова, інтерактивно-дослідницька, автономна. Велика інтегрована система надає можливість використовувати різні методи IAD, які підходять для конкретних завдань і середовищ, і забезпечує різні ступені взаємодії з користувачами.

Системи об'єктно-орієнтованого аналізу дуже різноманітні. Найширша підкатегорія цього типу системи стала домінувати в галузі дослідження фінансового ринку і стала відомою як «технічний аналіз». Це набір із десятків методів, які використовуються для прогнозування динаміки цін і вибору найкращої структури інвестиційного портфеля на основі набору емпіричних моделей ринкових тенденцій [53]. Як правило, ці методи використовують прості статистичні засоби, але враховують конкретні умови їх спеціальності (технічна мова, система різноманітних показників тощо). Окрім класичних статистичних методів, останні версії майже всіх відомих пакетів статистичних програм також містять елементи ІАД. Однак основна увага в них все ще зосереджена на класичних методах кореляції, регресії, факторного аналізу тощо. Недоліком цього типу системи є необхідність спеціального навчання користувача. Фахівці також зазначають, що потужні сучасні пакети статистичних програм занадто «важкі» для масштабного використання у фінансово-економічній сфері.

Крім того, ці системи зазвичай дуже дорогі. Статистичні пакети мають більш серйозну фундаментальну помилку, яка обмежує їх використання в ІАД. Більшість методів, включених до пакету, базуються на статистичній парадигмі, де основними учасниками є середні вибірки. Як було сказано вище, ці властивості часто є фіктивними значеннями при дослідженні реальних і складних життєвих явищ. Прикладами найпотужніших і розширених статистичних пакетів є SAS, SPSS, STATGRAPICS (Manugistics), STATISTICA, STADIA та інші. [25].

Нейронні мережі — це великий клас систем, архітектура яких нагадує генерацію нейронної тканини з нейронів. Однією з найпоширеніших архітектур є багат шаровий перцептрон із зворотним поширенням помилок, який імітує роботу нейронів в ієрархічній мережі, в якій кожен нейрон вищого рівня підключений до виходів нейронів нижчого рівня через власний вхід. Нейрони нижнього шару отримують значення вхідних параметрів для прийняття певних рішень, пов'язаних з прогнозуванням розвитку ситуації. Передані значення розглядаються як сигнали для передачі в наступну зміну. Ці сигнали

послаблюються або посилюються в залежності від значень (ваги), призначених зв'язкам між нейронами [54].

В результаті на виході верхнього нейрона виходить значення, яке вважається реакцією відповіді всієї мережі на вхідне значення вхідного параметра. Для подальшого використання мережі їй необхідно «вивчити» раніше отримані дані, для яких відомі значення вхідних параметрів і правильні відповіді на них. Навчання полягає у виборі ваг зв'язків між нейронами, тобто. Х. щоб надати мережі відповідь, яка є найближчою до відомої правильної відповіді. Основним недоліком парадигми нейронної мережі є те, що вона вимагає дуже великих вибірок. Ще одним серйозним недоліком є те, що навіть навчена нейронна мережа є «чорним ящиком». Знання, зафіксовані як вага сотень зв'язків між нейронами, просто не піддаються людському аналізу та інтерпретації. Відомі спроби пояснити конфігурацію структури нейронної мережі не видаються переконливими – PR-система KINOSuite. Прикладами нейромережевих систем є BrainMaker (CSS), NeuroShell (Ward Systems Group), OWL (HyperLogic) [39].

Далі розглянемо іншу систему IAD – систему міркування на основі випадків (CBR). Його ідея полягає в тому, щоб передбачити майбутнє або прийняти правильне рішення, шукаючи подібності поточної ситуації в минулому та вибираючи ту саму відповідь, яка була правильною. Тому цей метод також відомий як «метод найближчого сусіда».

Нещодавно також отримав визнання термін Memory Based Reasoning (MBR), який підкреслює, що рішення приймаються на основі всієї інформації, що зберігається в пам'яті. Система CBR показала хороші результати в різних завданнях. Їх основний недолік полягає в тому, що під час використання не створюються моделі чи правила, які б узагальнювали попередній досвід – при виборі рішення система базується на повному спектрі наявних історичних даних, тому вони не змогли пояснити, на які конкретні фактори спирається система CBR про надання відповідей. Іншим недоліком є довільність, яку система ЦБР допускає у виборі ступеня «закритості». Діапазон прецедентів, які необхідно мати на увазі для досягнення задовільної класифікації або передбачення,

критично залежить від цього показника. Прикладами систем, що використовують CBR, є інструмент KATE (Acknosoft, Франція), Pattern Recognition Workbench (Unica, США) [23].

Найпопулярнішими серед систем IAD є ті, які покладаються на дерева рішень як основний метод вирішення проблем. Вони створили ієрархію правил, як-от «Якщо...» (якщо щось), у формі дерева. Щоб вирішити, який клас використовувати для класифікації об'єкта чи ситуації, ви повинні відповісти на запитання у вузлі цього дерева, починаючи з його кореня. Питання виглядає так: «Чи більше значення параметра  $a$  за  $x$ ?». Якщо відповідь «так», то перейдіть до правого вузла наступного шару, якщо відповідь «ні», то перейдіть до лівого вузла; тоді знову необхідно задавати питання про відповідні вузли. Метод дерева рішень зрозумілий і наочний, тому він широко використовується в інтелектуальному аналізі даних. Однак дерево рішень не може знайти «найкращі» (найбільш повні та точні) правила в даних. Вони базуються на нативному принципі поступового відбору ознак і фактично «зачіпають» лише фрагменти реальних закономірностей, створюючи ілюзію послідовних висновків. Крім того, більшість систем використовують цей метод. Найбільш відомі See5/C5.0 (RuleQuest, Австралія), Clementine (Integral Solutions, Великобританія), SIPINA (Ліонський університет, Франція), IDIS (Information Discovery, США), KnowledgeSeeker (ANGOSS, Канада) [25].

Також має сенс згадати про генетичні алгоритми, але IAD не є основною сферою їх застосування. Швидше, їх слід розглядати як потужні інструменти для вирішення різноманітних проблем комбінування та оптимізації. Однак генетичні алгоритми зараз є стандартними інструментами методу IAD, тому вони розглядаються. Першим кроком у побудові генетичного алгоритму є кодування початкового логічного шаблону в базі даних під назвою хромосома, а вся сукупність цього шаблону називається сукупністю хромосом. Далі для реалізації концепції відбору вводиться метод порівняння різних хромосом. Для роботи з популяціями, які імітують біологічні процеси, використовуються методи розмноження, мінливості (мутації) і генетичного складу. Найважливішими з них

є випадкові мутації одиночних хромосом, кросинговер і рекомбінація генетичного матеріалу в єдиних батьківських хромосомах, міграція генів.

При виконанні таких процедур на кожному етапі еволюції виникають популяції з дедалі більш досконалішими елементами. Зручність генетичних алгоритмів полягає в простоті їх розпаралелювання. Наприклад, можна розділити покоління на групи і розглядати кожну групу окремо, час від часу замінюючи кілька хромосом. Є й інші способи розпаралелення генетичних алгоритмів. Звичайно, генетичні алгоритми мають багато недоліків. Критерії відбору хромосом і процедури, які будуть використані, є евристичними, і немає гарантії, що буде знайдено «краще» рішення. Як і в реальному житті, еволюція може застрягти на непродуктивній гілці. Це особливо актуально при вирішенні масштабних завдань зі складними внутрішніми зв'язками. Прикладом може служити система GeneHunter компанії Ward Systems Group.

Приблизно в середині 1960-х років логічні закономірності в даних М.М. Алгоритм обмеженого перерахування був раніше запропонований Bongard [41]. З того часу довела свою ефективність у вирішенні багатьох проблем у різних сферах. Ці алгоритми обчислюють загальну частоту простих логічних подій у підмножинах даних. Приклади простих логічних подій:  $X = a$ ;  $X$ . На основі аналізу обчислених частот можна зробити висновки про корисність певних комбінацій для створення асоціацій даних, класифікацій і прогнозів.

Найбільш яскравим сучасним представником цього методу є система WizWhy компанії WizSoft. Хоча автор системи Авраам Майдан не розголошує конкретних подробиць алгоритму WizWhy, залишаючи результати ретельних тестів системи, результати її роботи, час отримання залежно від кількості аналізів. Параметри дозволяють зробити висновок, що це обмежене перевищення. Автор WizWhy стверджує, що його система розпізнає всі логічні правила «якщо-тоді» в даних, що є невірним. По-перше, максимальна довжина комбінації в правилі if-then у системі WizWhy дорівнює 6. По-друге, алгоритм є евристичним пошуком простих логічних подій із самого початку, на якому будуються всі наступні аналізи. Знаючи ці можливості WizWhy, не складно

поставити найпростішу тестову задачу, яку система взагалі не може вирішити [54].

Інший момент полягає в тому, що система надає рішення лише для відносно невеликої кількості даних протягом задовільного часу. Однак система WizWhy стала одним із лідерів на ринку продуктів інтелектуального аналізу даних, оскільки система завжди демонструвала вищу продуктивність, ніж будь-який інший алгоритм у вирішенні практичних завдань. Вартість системи трохи вище середнього. З огляду на перспективи ІАД можливі наступні напрямки розвитку:

- Виокремлення типів тематичних гілок з відповідними евристичними, формалізація яких допоможе вирішити проблеми, пов'язані з приналежністю ІАД до цих гілок;
- Винайдення формальних мов для формалізації засобів міркування та логіки, автоматизація яких стає інструментом для вирішення завдань ІАД у прикладних галузях;
- Розробка методів ІАД, які зможуть не тільки витягувати закономірності з даних, але й формувати деякі теорії на основі емпіричних даних;
- Подолання великого розриву між можливостями інструментів ІАД і теоретичними результатами в цій галузі.

Системи ІАД широко використовуються як комерційні програми та інструменти підтримки досліджень у генетиці, фармакології, медицині тощо. Фахівці в області ІАД пов'язують майбутнє цих систем з інтелектуальними додатками, вбудованими в сховища даних компаній. Хоча кількість методів ІАД досить велика, ці причинно-наслідкові правила поступово віддають перевагу логічним алгоритмам пошуку. Вони використовуються для вирішення таких проблем, як прогнозування, класифікація, розпізнавання образів, сегментація бази даних, вилучення «прихованих» знань із даних, інтерпретація даних та асоціація баз даних. Результати таких алгоритмів є ефективними та легкими для інтерпретації.

Крім того, основною проблемою методу логічного розпізнавання образів є проблема пошуку варіантів за прийнятний час. Відомі методи обмежують такі пошуки або штучно (алгоритми KORA, WizWhy), або будують дерева рішень (алгоритми CART, CHAID, ID3, See5, Sipina та ін.), які принципово обмежують ефективність пошуку за причинно-наслідковими правилами. Інші проблеми пов'язані з тим, що відомий метод пошуку логічного правила не підтримує функцію узагальнення наведених вище правил і функцію пошуку найкращої комбінації цих правил. Успішне вирішення цих проблем стає основою нових методик ІАД і відповідних розробок [39].

Системи та методи інтелектуального аналізу даних відіграють важливу роль у багатьох сферах: страхуванні, медицині, автомобілебудуванні, банківській справі тощо. Їх особливістю є складна системна організація, оскільки вони в основному належать до вищого рівня організації системи та їх закономірності в мові статистичних чи інших аналітичних математичних моделей неможливо описати точно. Дані в цих областях неоднорідні, різнорідні, тимчасові та часто багатовимірні. Виходячи з наведених вище матеріалів, можна сказати, що потенціал ІАД сприяє застосуванню та поширенню цієї технології в сучасному світі комп'ютерних технологій.

#### Огляд методів інтелектуального аналізу даних

Загалом, інтелектуальний аналіз даних – це процес пошуку значущих закономірностей у великих обсягах даних і встановлення зв'язків між ними [11, с. 400]. Термін інтелектуальний аналіз даних вперше використав М. Ловелл у 1983 році в критичному контексті поганих методів аналізу даних без попередніх припущень. З 1990-х років цей термін вже широко використовується в спільноті баз даних, штучного інтелекту та машинного навчання

позитивна конотація [43, с. 50]. Тому Г. Пятецький-Шапіро використав цей термін у 1989 р., щоб представити набір методів визначення властивостей інформації з великої кількості погано структурованих вихідних даних [8; 20].

При визначенні знань із бази даних інтелектуальний аналіз даних є важливим кроком. Він створює корисні шаблони або моделі з даних.

У сучасному середовищі, в якому багато даних, аналіз даних є потужним інструментом для вирішення різноманітних завдань від обробки даних до аналізу часових рядів.

Аналіз даних складається з трьох основних етапів: підготовка даних, скорочення даних і пошук важливої інформації. Однак конкретні методи відрізняються залежно від практики компанії або дослідників. Наприклад, Fayyad U. [8] Рекомендує наступні кроки:

- Отримати дані з великих баз даних. Вибрати важливі характеристики для конкретного дослідження.
- Вибрати відповідну підмножину для аналізу.
- Вибрати відповідну систему вибірки, очистити дані, усунути неточності, вирішити проблему з відсутніми полями та записами та видалити викиди.
- Застосувати відповідні перетворення, зменшити розмір і проєкцію та трансформувати масштаби, щоб застосувати вибраний метод.
- Пристосувати модель до попередньо оброблених даних та інтерпретувати результати в контексті значущих припущень.

Розроблено стандарти організації процесу аналізу даних, які містять покрокові рекомендації, цілі та завдання для всіх етапів процесу аналізу даних. Альянс, утворений NCR, SPSS і DaimlerChrysler, розробив міжгалузевий стандартний процес інтелектуального аналізу даних (CRISP-DM), який є продовженням Knowledge Discovery in Databases (KDD) і SEMMA (Sample — збір даних, Explore — дослідження зв'язку, Modify – модифікація, Model – моделювання, Assess – оцінка) [24; 39].

Створення та перевірка вихідного набору даних відбувається за один крок

«Підготовка даних». На цьому етапі важливо визначити придатність



вибраного набору даних для подальшої обробки. Якщо обсяг даних невеликий, можливо, нові дані потрібно буде зібрати відповідно до визначених стандартів. Крім того, на цьому етапі вам, ймовірно, потрібно буде переглянути цілі та завдання, поставлені організацією.

На наступному етапі «обробки даних» вихідні дані вже попередньо оброблені, щоб за допомогою машинного навчання можна було побудувати модель аналізу, яка відображає структурний опис інформації, яка міститься в навчальних даних. Тому на цьому етапі створюється робоча (внутрішня) модель даних для подальшої перевірки. Етапи підготовки даних і моделювання зазвичай виконуються паралельно. Зазвичай необхідна багаторазова ітерація, оскільки результати, отримані під час моделювання, дозволяють краще зрозуміти вимоги методу попередньої обробки.

На практиці для перевірки аналітичних моделей використовуються різноманітні методи. Якщо етап тестування показує низький рівень задоволеності моделлю, необхідно переглянути весь проект аналізу даних і повернутися до етапу «Зрозуміти бізнес-контекст», щоб узгодити цілі або підходи до збору даних. фаза «розгортання моделі» в програмній системі здійснюється при досягненні високого рівня точності моделі. Важливим етапом є верифікація побудованої моделі за допомогою тестових даних. Після того, як побудована модель аналізу досягає необхідного рівня точності та адекватності, її тестують на нових неструктурованих даних.

Методи інтелектуального аналізу даних також можна класифікувати за завданнями. Прогнозування та опис – категорії, на які можна розділити більшість проблем інтелектуального аналізу даних у найширшому сенсі [3; 25]. Проблема передбачення дозволяє передбачити змінну на основі іншої доступної інформації. Крім того, описові завдання узагальнюють дані. Відповідно до цієї класифікації ми виділяємо дві групи. Перший з них — це поділ методів IAD на ключові завдання сегментації (тобто завдання класифікації та кластеризації) і завдання прогнозування. Відповідно до другого поділу завдань, методи IAD спрямовані на отримання описових

результатів і прогнозів [23].

Описові методи служать для знаходження шаблонів або зразків, що описують дані, які піддаються інтерпретації з точки зору аналітика [53]:

- алгоритм k-середніх k-медіани;
- ієрархічні методи кластерного аналізу;
- самоорганізовані карти Кохонена;
- методи крос-табличній візуалізації;

Прогнозуючі методи використовують значення одних змінних для передбачення / прогнозування невідомих (пропущених) або майбутніх значень інших (цільових) змінних [54]:

- нейронні мережі;
- дерева рішень;
- лінійна регресія
- метод найближчого сусіда;
- метод опорних векторів.

Основним елементом процесу IAD є його методи, за допомогою яких в результаті інтерпретації винайдених закономірностей створюються нові зв'язки та закономірності та формуються закономірності та знання. Для отримання найкращих результатів обробки даних розроблено та використовується кілька основних методів IAD. До основних методів можна віднести використання статистичних методів, напр Б.: Кореляційно-регресійний аналіз, факторний та дискримінантний аналіз. Методи аналізу даних також включають: асоціацію, класифікацію, кластеризацію, прогнозування, розпізнавання образів, генетичні та еволюційні алгоритми та методи «м'якого обчислення» [11].

Відповідно до принципу використання необроблених навчальних даних, усі методи аналізу даних можна розділити на дві категорії. У цій класифікації верхній рівень визначається відповідно до того, зберігаються дані після

інтелектуального аналізу чи перетворюються для подальшого використання.

Перша група методів - це методи, які безпосередньо використовують або зберігають дані, наприклад Наприклад: кластерний аналіз, метод найближчого сусіда, метод k-найближчого сусіда тощо. У цьому випадку необроблені дані зберігаються в деталях і використовуються безпосередньо на етапі прогнозного моделювання та/або аналізу аномалій. Проблема цього набору методів полягає в складності аналізу дуже великих баз даних [46].

Друга група методів - це методи пошуку та використання формалізованих закономірностей або дистиляції закономірностей.

Принцип технології шаблонної дистиляції полягає в тому, що вибірка даних (шаблон) береться з вихідних даних і потім перетворюється на визначені формальні структури залежно від використовуваного методу. Описаний процес виконується на рівні вільного пошуку, на відміну від першої групи методів, яка не має такого рівня. Результати фази вільного пошуку використовуються в процесі прогнозного моделювання та аналізу винятків, оскільки, на відміну від самих баз даних, вони досить компактні і можуть інтерпретуватися аналітиком, а можуть і ні. Серед таких методів можна виділити: логічні (нечіткі запити та аналіз, символічні правила, дерева рішень, генетичні алгоритми), крос-табуляцію, методи візуалізації, методи на основі рівнянь тощо.

Логічні методи або методи логічної індукції зазвичай формулюють виявлені закономірності в досить зрозумілій і прозорій для користувача формі. Безперервні та дискретні змінні можуть бути включені в похідні правила. Що стосується дерев рішень, їх можна перетворити на набори символічних правил, створивши правило на шляху від кореня до кінцевого вузла дерева. Щоб вирішити ту саму проблему, можна застосувати дерева рішень або правила, відмінності яких стосуються лише можливостей. Крім того, реалізація правил відбувається за допомогою більш повільних алгоритмів порівняно з індукцією дерев рішень [54].

Методи крос-табуляції включають агентів, байєсівські (довірчі) мережі, візуалізацію крос-табуляції. Кростаблична візуалізація не зовсім відповідає особливості методів ІАД, а саме самостійному пошуку закономірностей за допомогою систем аналізу. Проте представлення інформації у вигляді перехресної таблиці представляє основне завдання ІАД – реалізацію пошуку за зразком, тому цей метод також можна розглядати як один із методів ІАД [49].

Метод, заснований на рівняннях, виражає визнану закономірність у формі математичного виразу - рівняння. Таким чином, вони можуть використовуватися лише з числовими змінними, а інші типи змінних повинні бути закодовані відповідним чином. Це певною мірою обмежує застосування цієї групи методів, але вони широко використовуються для вирішення різноманітних задач, особливо задач прогнозування. Ця класифікація поділяє різні методи аналізу даних на дві групи: статистичні методи та кібернетичні методи.

Варто зазначити, що є два способи пов'язати статистичні методи з аналізом даних. Перший порівнює статистичні методи з методами аналізу даних, і його прихильники вважають, що класичні статистичні методи є окремою областю аналізу даних. За другим методом до математичного інструментарію ІАД належать методи статистичного аналізу. Останнього дотримується більшість авторитетних джерел.

У цій класифікації є два набори методів [53]:

- на основі статистичних методів із застосуванням усередненого досвіду, відображеного в ретроспективних даних;
- кібернетичні методи, які містять багато різноманітних математичних підходів.

Недоліком цієї класифікації є те, що і статистичні, і кібернетичні алгоритми спираються на порівняння статистичного досвіду з результатами моніторингу поточної ситуації. Перевагою цієї класифікації є її легкість для пояснення, тому її використовують для опису сучасних методів вилучення знань із початкового поля спостереження (оперативного та ретроспективного), тобто в

задачах інтелектуального аналізу даних.

Статистичні методи ІАД мають чотири взаємопов'язані розділи [2]:

- попередній аналіз характеру статистичних даних (перевірка гіпотез стаціонарності, нормальності, незалежності, однорідності, оцінка характеру функції розподілу, її параметрів тощо);
- Визначення зв'язків і закономірностей (лінійний і нелінійний регресійний аналіз, кореляційний аналіз тощо);
- багатофакторний статистичний аналіз (лінійний та нелінійний дискримінантний аналіз, кластерний аналіз, компонентний аналіз, факторний аналіз тощо);
- динамічні моделі та прогнози на основі часових рядів.

Бібліотека статистичних методів аналізу даних поділяється на такі види методів [27; 24]:

- описовий аналіз та опис вихідних даних;
- Аналіз кореляцій (кореляційно-регресійний аналіз, факторний аналіз, аналіз розкиду);
- багатофакторний статистичний аналіз (компонентний аналіз, дискримінантний аналіз, багатофакторний регресійний аналіз, канонічні кореляції тощо);
- Аналіз часових рядів (динамічні моделі та прогнози)

Другий напрямок ІАД – це набір методів, що поєднують ідеї обчислювальної математики та використання теорії штучного інтелекту або кібернетичних методів [13; 25]:

- штучні нейронні мережі (виявлення, кластеризація, прогнозування);
- Еволюційне програмування (включаючи алгоритми методу групового обчислення аргументів);
- генетичні алгоритми (оптимізація);
- асоціативна пам'ять (пошук аналогів, прототипів);

- незрозуміла логіка;
- дерева рішень;
- Системи обробки експертних знань.

Розробка проектів автоматизації на основі Super Big Data вимагає чіткої та глибокої оцінки існуючої методології IAD. Нижче наведено короткий огляд цих методів і їх порівняння.

Класифікація є найпростішим завданням в аналізі даних. Класифікація включає методи аналізу даних, які можна використовувати для вилучення шаблонів, які описують важливі категорії даних або прогнозують тенденції. Найчастіше методи класифікації використовуються для призначення елементів даних визначеним класам або групам. Іншими словами, завдання класифікації полягає у визначенні категорії об'єкта через властивості самого об'єкта. У цьому завданні багато класів, до яких можна віднести об'єкт, відомі заздалегідь. Різні типи методів класифікації включають [2]:

- найближчий сусід (найближчий сусід);
- k-найближчий сусід (k-найближчий сусід);
- Байєсовські мережі. Це ймовірнісна графічна модель, яка будує графіки набору випадкових величин та їхніх умовних залежностей за допомогою спрямованого ациклічного графа (DAG).
- індукція дерев рішень;
- нейронні мережі.

Задача регресії подібна до задачі класифікації тим, що дозволяє визначити значення певних параметрів об'єкта на основі відомих властивостей об'єкта. На відміну від задачі класифікації, значенням параметра є не набір кінцевих класів, а набір дійсних чисел. Тоді як у задачах класифікації цільова змінна зазвичай має невелику кількість дискретних значень (таких як «високе» та «низьке»), у задачах регресії цільова змінна є безперервною. Наприклад, виявлення шахрайства з кредитними картками є прикладом проблеми класифікації, тоді як

прогнозування майбутніх цін на акції є проблемою регресії.

Кластеризація є логічним продовженням ідей класифікації. Завданням кластеризації є пошук незалежних груп (кластерів) та їх характеристик у всьому наборі даних початкового аналізу. Вирішення цієї проблеми допомагає краще зрозуміти дані. Крім того, групування об'єктів одного типу може зменшити їх кількість, полегшуючи аналіз. Особливістю кластеризації є те, що клас об'єкта спочатку не визначений. Результатом кластеризації є поділ об'єктів на групи. Прикладом кластеризації є певний тип нейронної мережі (карта Кохонена), яка сама організовується без вчителя. Метод знаходження правил асоціації є найпоширенішим серед методів аналізу даних. Пошук асоціативних правил стосується виявлення повторюваних залежностей (або асоціацій) між об'єктами чи подіями. Виявлені зв'язки формулюються у вигляді правил, за допомогою яких можна краще зрозуміти природу аналізованих даних і передбачити настання подій. Цей метод широко використовується в роздрібній торгівлі для вивчення купівельної поведінки споживачів. Відмінність відображення від двох попередніх методів полягає в тому, що пошук шаблону ґрунтується не на властивостях аналізованого об'єкта, а на кількох подіях, які відбуваються разом. Априорний алгоритм є найвідомішим алгоритмом розв'язання задачі пошуку асоціативних правил [11].

Залежно від типу даних, етапу аналізу та завдання вибирається той чи інший метод відповідно до властивих йому властивостей: точність, масштабованість, інтерпретація, верифікованість, складність, гнучкість, швидкість і популярність.

Головною особливістю IAD є поєднання широкого спектру математичних інструментів (від класичного статистичного аналізу до нових кібернетичних методів) і останніх розробок інформаційних технологій. Технологія IAD гармонійно поєднує в собі чіткі формальні методи та неформальні методи аналізу, а саме кількісний та якісний аналіз даних.

Більшість методів аналізу, які використовуються в IAD, є добре відомими математичними алгоритмами та методами. З появою нового обладнання та

програмного забезпечення нові функції можна використовувати в їхніх програмах для вирішення певних конкретних проблем.

## Нейромережі та принцип їхньої роботи

### Історія

Перша нейронна мережа була розроблена Уорреном Маккалохом і Уолтером Піттсом у 1943 році. Вони написали основоположну роботу про те, як можуть працювати нейрони, і змоделивали свої ідеї, створивши просту нейронну мережу за допомогою електричних схем. Ця революційна модель проклала шлях до досліджень нейронних мереж у двох областях:

- Біологічні процеси в головному мозку.
- Застосування нейронних мереж до штучного інтелекту (ШІ).

Дослідження штучного інтелекту стрімко прискорилися, коли в 1975 році Куніхіко Фукусіма розробив першу справжню багатoshарову нейронну мережу. Початковою метою підходу нейронної мережі було створення комп'ютерної системи, яка могла б вирішувати проблеми, подібні до людського мозку. Однак з часом дослідники перемістили свою увагу на використання нейронних мереж для виконання конкретних завдань, що призвело до відходу від суто біологічного підходу. Відтоді нейронні мережі підтримують різноманітні завдання, включаючи комп'ютерний зір, розпізнавання мови, машинний переклад, фільтрацію соціальних мереж, настільні та відеоігри та медичну діагностику. Оскільки розмір структурованих і неструктурованих даних виріс до рівня великих даних, люди розробили системи глибокого навчання, які, по суті, є багаторівневими нейронними мережами. Глибоке навчання дає змогу збирати й отримувати більше даних, у тому числі неструктурованих.



## Як працюють нейронні мережі

Проста нейронна мережа складається з вхідного рівня, вихідного (або цільового) рівня та прихованого шару між ними. Шари з'єднані вузлами, і ці з'єднання утворюють «мережу» — нейронну мережу — взаємопов'язаних вузлів.

Вузол змодельований на нейроні в людському мозку. Подібно до поведінки нейронів, вузли активуються, коли є достатній стимул або введення. Ця активація поширюється мережею та створює відповідь на подразники (вихід). Зв'язки між цими штучними нейронами працюють як прості синапси, що дозволяє передавати сигнали від одного до іншого. Сигнали обмінюються між шарами на шляху від першого вхідного до останнього вихідного рівня та обробляються на цьому шляху.

Коли нейрони мають вирішити питання або проблему, вони виконують математичні обчислення, щоб визначити, чи достатньо інформації для передачі інформації наступному нейрону. Простіше кажучи, вони зчитують усі дані та з'ясовують, де найсильніші зв'язки. У найпростішому типі мережі отримані вхідні дані складаються, і коли сума перевищує певний поріг, нейрон «спрацьовує», активуючи нейрони, до яких він підключений.

У міру збільшення кількості прихованих шарів нейронної мережі утворюються глибокі нейронні мережі. Архітектури глибокого навчання виводять прості нейронні мережі на новий рівень. Використовуючи ці рівні, дослідники даних можуть створювати власні мережі глибокого навчання, які забезпечують машинне навчання, яке може навчити комп'ютер точно імітувати людські завдання, такі як розпізнавання мовлення, ідентифікація зображень або створення прогнозів. Не менш важливо, що комп'ютер може навчатися сам, розпізнаючи шаблони на багатьох рівнях обробки.

Застосуємо це визначення на практиці. Дані подаються в нейронну мережу через вхідний рівень, який передається на приховані рівні. Обробка виконується в прихованих шарах через систему зважених з'єднань. Потім вузли прихованого шару поєднують дані з вхідного рівня з набором коефіцієнтів і призначають відповідні ваги вхідним даним. Потім ці зважені продукти підсумовуються. Сума

передається через функцію активації вузла, яка визначає, як далеко сигнал має подолати мережу, щоб вплинути на кінцевий результат. Нарешті, приховані шари пов'язані з вихідним шаром, де витягуються результати.

Нейромережа нелінійної авторегресії

На рис. 11.1 приведена нейромережа нелінійної авторегресії (NARNN), яка є нерекурентною динамічною двошаровою мережею, причому прихований шар зображений прямокутником, за основу бралася модель нелінійної авторегресії NAR(5). NARNN є особливим класом багатошарового персептрона



Рис. 11.1. Нейросеть нелінійної авторегресії (NARNN)

Для NARNN використовується навчання на основі корекції помилок (навчання з учителем), при цьому найчастіше застосовується алгоритм зворотного поширення (BP) для ШНМ прямого поширення..

Навчання нелінійної авторегресії (метод зворотного поширення) в послідовному режимі

1. Номер ітерації навчання  $n = 1$ , ініціалізація зсувів (порогів)  $b^{(1)}(n)$ ,  $b^{(2)}(n)$ , і ваг  $w_{lj}^{(1)}(n)$ ,  $w_i^{(2)}(n)$ ,  $i, j \in \overline{1, N^{(1)}}$ ,  $l \in \overline{0, M^{(0)}}$ , де  $N^{(1)}$  - кількість нейронів в першому шарі,  $M^{(0)}$  - затримка у вхідному шарі.

2. Задається навчальна множина  $\{(x_\mu, d_\mu) \mid x_\mu \in R, d_\mu \in R\}$ ,  $\mu \in \overline{1, P}$ , де  $x_\mu$  -  $\mu$ -е навчальне вхідне значення,  $d_\mu$  -  $\mu$ -е навчальне вихідне значення,  $P$  - потужність навчальної множини. Номер поточної пари з навчальної множини  $\mu = 1$ .

3. Початкове обчислення вихідного сигналу для вхідного шару

$$y^{(0)}(n-v) = 0, \quad v \in \overline{1, M^{(0)}}.$$

4. Обчислення вихідного сигналу для кожного шару (прямий хід)

$$y^{(0)}(n) = x_{\mu},$$

$$y_j^{(1)}(n) = f^{(1)}(s_j^{(1)}(n)), \quad s_j^{(1)}(n) = b_j^{(1)} + \sum_{l=0}^{M^{(0)}} w_{lj}^{(1)}(n) y^{(0)}(n-l),$$

$$j \in \overline{1, N^{(1)}},$$

$$y^{(2)}(n) = f^{(2)}(s^{(2)}(n)), \quad s^{(2)}(n) = \sum_{i=0}^{N^{(1)}} w_i^{(2)}(n) y_i^{(1)}(n),$$

де  $N^{(1)}$  - число нейронів в першому шарі,  $w_{lj}^{(1)}(n)$  - вага зв'язку від вхідного нейрона в момент часу  $n-l$  до  $j$ -го нейрона в першому шарі в момент часу  $n$ ,  $w_i^{(2)}(n)$  - вага зв'язку від  $i$ -го нейрона до вихідного нейрона в момент часу  $n$ ,  $y_j^{(1)}(n)$  - вихід  $j$ -го нейрона в першому шарі,  $y^{(2)}(n)$  - вихід нейрона у другому шарі,  $f^{(k)}$  - функція активації нейронів  $k$ -го шару.

Вважається, що  $w_0^{(2)}(n) = b^{(2)}(n)$ ,  $y_0^{(1)}(n) = 1$ .

5. Обчислення енергії помилки ШНМ

$$E(n) = \frac{1}{2} e^2(n), \quad e(n) = y^{(2)}(n) - d_{\mu}.$$

6. Налаштування синаптичних ваг на основі узагальненого дельта правила (зворотний хід)

$$w_i^{(2)}(n+1) = w_i^{(2)}(n) - \eta \frac{\partial E(n)}{\partial w_i^{(2)}(n)},$$

$$w_{lj}^{(1)}(n+1) = w_{lj}^{(1)}(n) - \eta \frac{\partial E(n)}{\partial w_{lj}^{(1)}(n)},$$

$$\frac{\partial E(n)}{\partial w_i^{(2)}(n)} = y_i^{(1)}(n) g^{(2)}(n), \quad i \in \overline{0, N^{(1)}},$$

$$\frac{\partial E(n)}{\partial w_{lj}^{(1)}(n)} = y^{(0)}(n-l)g_j^{(1)}(n), \quad j \in \overline{1, N^{(1)}}, l \in \overline{0, M^{(0)}},$$

$$\frac{\partial E(n)}{\partial b_j^{(1)}(n)} = g_j^{(1)}(n), \quad j \in \overline{1, N^{(1)}},$$

$$g^{(2)}(n) = f'^{(2)}(s^{(2)}(n))(y^{(2)}(n) - d_\mu),$$

$$g_j^{(1)}(n) = f'^{(1)}(s_j^{(1)}(n))w_j^{(2)}(n)g^{(2)}(n)$$

#### 7. Перевірка умови завершення

Якщо  $n \bmod P > 0$ , то  $\mu = \mu + 1$ ,  $n = n + 1$ , перехід до 4.

Якщо  $n \bmod P = 0$  та  $\frac{1}{P} \sum_{s=1}^P (y(n-P+s) - d_s) > \varepsilon$ , то  $n = n + 1$ , перехід до 2.

Якщо  $n \bmod P = 0$  та  $\frac{1}{P} \sum_{s=1}^P (y(n-P+s) - d_s) < \varepsilon$ , то завершитися.

#### Переваги нелінійної авторегресії

1. Є універсальним апроксиматором. Забезпечує глобальну апроксимацію нелінійного відображення вхідного сигналу у вихідний.
2. Забезпечує хорошу якість узагальнення.
3. Автоматичне визначається кількість прихованих шарів (дорівнює одному).
4. На відміну від статичних мереж дозволяє здійснювати адаптивну фільтрацію, короткостроковий прогноз, адаптивне управління, параметричну ідентифікацію моделі, класифікацію нестационарних сигналів.
5. На відміну від NARMANN навчається швидше.
6. Можливо навчання в послідовному та пакетному режимі.

#### Недоліки нелінійної авторегресії

1. Навчання відбувається повільніше, ніж в разі MLP, RBFNN, PNN, мережі Хеммінга, SOM, CPNN.

2. Відсутня автоматичне визначення числа нейронів в прихованому шарі.
3. На відміну від методів навчання SVM, методи навчання NARNN вибирають положення гіперплощини, що розділяє, довільним чином.
4. На відміну від навчання SVM, навчання не зводиться до задачі квадратичного програмування в опуклою області, що має єдине рішення.
5. На відміну від ART, не вирішує проблему пластичності-стабільності.
6. Може бути тільки два класи, що не завжди підходить для задач класифікації.
7. На відміну від NARMANN гірше вирішує завдання адаптивної фільтрації, прогнозу, адаптивного управління, параметричної ідентифікації моделі, класифікації нестационарних сигналів.

Неймережа - ковзного середнього

На рис. 12.1 приведена неймережа нелінійної авторегресії - ковзного середнього (NARMANN), яка є рекурентною двошаровою мережею. причому прихований шар зображений прямокутником, за основу бралася модель нелінійної авторегресії-ковзного середнього NARMA(5,3). NARMANN є особливим класом багатшарового персеプトрона.

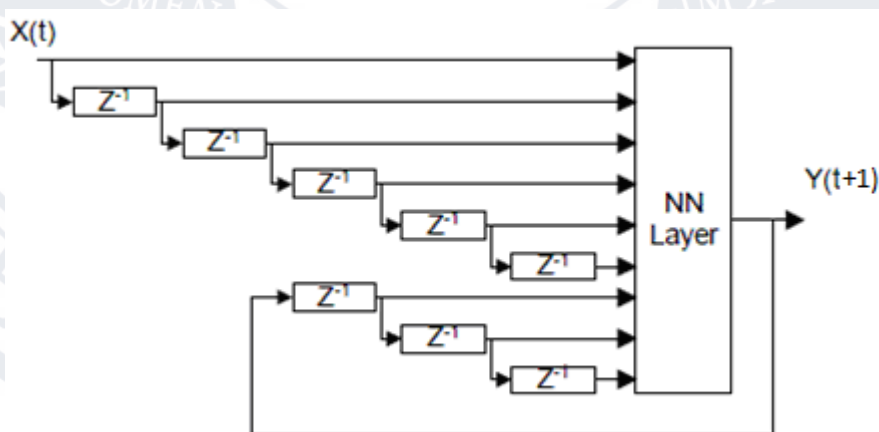


Рис. 12.1. Неймережа нелінійної авто регресії - ковзного середнього (NARMANN)

Для NARMANN використовується навчання на основі корекції помилок (навчання з учителем), при цьому найчастіше застосовується метод градієнтного спуску, на якому заснований метод зворотного поширення в часі (ВРГТ) для ШНМ зворотного поширення.

Навчання нелінійної авторегресії - ковзного середнього (метод зворотного поширення) в послідовному режимі

1. Номер ітерації навчання  $n = 1$ , ініціалізація за допомогою рівномірного розподілу на інтервалі  $(0,1)$  або  $[-0.5, 0.5]$  зсувів (порогів), і ваг  $w_{lj}^{(1)}(n)$ ,  $l \in \overline{0, M^{(0)} + M^{(2)}}$ ,  $j \in \overline{1, N^{(1)}}$ ,  $w_i^{(2)}(n)$ ,  $i \in \overline{1, N^{(1)}}$ , де  $N^{(1)}$  - кількість нейронів в першому шарі,  $M^{(k)}$  - затримка в  $k$ -ом шарі.

2. Задається навчальна множина  $\{(x_\mu, d_\mu) \mid x_\mu \in R, d_\mu \in R\}$ ,  $\mu \in \overline{1, P}$ , де  $x_\mu$  -  $\mu$ -е навчальне вхідне значення,  $d_\mu$  -  $\mu$ -е навчальне вихідне значення,  $P$  - потужність навчальної множини. Номер поточної пари з навчальної множини  $\mu = 1$

3. Початкове обчислення вихідного сигналу для вхідного та вихідного шару  $M = \max\{M^{(0)}, M^{(2)}\}$ .

$$y^{(0)}(n - v) = 0, v \in \overline{1, M}.$$

$$y^{(2)}(n - v) = 0, v \in \overline{1, M}.$$

4. Обчислення вихідного сигналу для кожного шару (прямий хід)

$$y^{(0)}(n) = x_\mu,$$

$$y_j^{(1)}(n) = f^{(1)}(s_j^{(1)}(n)), j \in \overline{1, N^{(1)}},$$

$$s_j^{(1)}(n) = b_j^{(1)}(n) + \sum_{l=0}^{M^{(0)}} w_{lj}^{(1)}(n) y^{(0)}(n - l) +$$

$$+ \sum_{l=M^{(0)}+1}^{M^{(0)}+M^{(2)}} w_{lj}^{(1)}(n) y^{(2)}(n - (l - M^{(0)})),$$

$$y^{(2)}(n) = f^{(2)}(s^{(2)}(n)), s^{(2)}(n) = \sum_{i=0}^{N^{(1)}} w_i^{(2)}(n) y_i^{(1)}(n),$$

де  $N^{(1)}$  - число нейронів в першому шарі,  $w_{lj}^{(1)}(n)$  - вага зв'язку від вхідного нейрона в момент часу  $n-l$  до  $j$ -го нейрона в першому шарі в момент часу  $n$  або вага зв'язку від вихідного нейрона в момент часу  $n-(l-M^{(0)})$  до  $j$ -го нейрона в першому шарі в момент часу  $n$ ,  $w_i^{(2)}(n)$  - вага зв'язку від  $i$ -го нейрона до вихідного нейрона в момент часу  $n$ ,  $y_j^{(1)}(n)$  - вихід  $j$ -го нейрона в першому шарі,  $y^{(2)}(n)$  - вихід нейрона у другому шарі,  $f^{(k)}$  - функція активація нейронів  $k$ -го шару.

Вважається, що  $w_0^{(2)}(n) = b^{(2)}(n)$ ,  $y_0^{(1)}(n) = 1$ .

#### 5. Обчислення енергії помилки ШНМ

$$E(n) = \frac{1}{2} e^2(n), \quad e(n) = y^{(2)}(n) - d_\mu.$$

6. Налаштування синаптичних ваг на основі узагальненого дельта правила (зворотний хід)

$$w_i^{(2)}(n+1) = w_i^{(2)}(n) - \eta \frac{\partial E(n)}{\partial w_i^{(2)}(n)},$$

$$w_{lj}^{(1)}(n+1) = w_{lj}^{(1)}(n) - \eta \frac{\partial E(n)}{\partial w_{lj}^{(1)}(n)},$$

де  $\eta$  - параметр, що визначає швидкість навчання (при великому навчання відбувається швидше, але збільшується небезпека одержати невірне рішення),  $0 < \eta < 1$

$$\frac{\partial E(n)}{\partial w_i^{(2)}(n)} = y_i^{(1)}(n) \overline{g^{(2)}(n)}, \quad i \in \overline{0, N^{(1)}},$$

$$\frac{\partial E(n)}{\partial w_{lj}^{(1)}(n)} = \begin{cases} y^{(0)}(n-l) g_j^{(1)}(n), & l \leq M \\ y^{(2)}(n-(l-M^{(0)})) g_j^{(1)}(n), & l > M \end{cases},$$

$$l \in \overline{0, M^{(0)} + M^{(2)}}, \quad j \in \overline{1, N^{(1)}},$$

$$\frac{\partial E(n)}{\partial b_j^{(1)}(n)} = g_j^{(1)}(n), \quad j \in \overline{1, N^{(1)}},$$

$$g^{(2)}(n) = f'^{(2)}(s^{(2)}(n))(y^{(2)}(n) - d_\mu),$$

$$g_j^{(1)}(n) = f'^{(1)}(s_j^{(1)}(n))w_j^{(2)}(n)g^{(2)}(n).$$

7. Перевірка умови завершення

Якщо  $n \bmod P > 0$ , то  $\mu = \mu + 1$ ,  $n = n + 1$ , перехід до 4.

Якщо  $n \bmod P = 0$  та  $\frac{1}{P} \sum_{s=1}^P E(n - P + s) > \varepsilon$ , то  $n = n + 1$ , перехід до 2.

Якщо  $n \bmod P = 0$  та  $\frac{1}{P} \sum_{s=1}^P E(n - P + s) < \varepsilon$ , то завершитися.

Переваги нелінійної авторегресії - ковзного середнього

1. Є універсальним апроксиматором. Забезпечує глобальну апроксимацію нелінійного відображення вхідного сигналу у вихідний.
2. Забезпечує хорошу якість узагальнення.
3. Автоматичне визначається кількість прихованих шарів (дорівнює одному).
4. На відміну від статичних ШНМ дозволяє здійснювати адаптивну фільтрацію, прогноз, адаптивне управління, параметричну ідентифікацію моделі, класифікацію нестационарних сигналів.
5. На відміну від NARNN краще вирішує завдання адаптивної фільтрації, прогнозу, адаптивного управління, параметричної ідентифікації моделі, класифікації нестационарних сигналів.

Недоліки нелінійної авторегресії - ковзного середнього

1. Навчання відбувається повільніше, ніж в разі MLP, RBFNN, PNN, мережі Хеммінга, SOM, CPNN.
2. Відсутня автоматичне визначення числа нейронів в прихованому шарі.
3. На відміну від методів навчання SVM, методи навчання NARMANN вибирають положення гіперплощини, що розділяє, довільним чином.
4. На відміну від навчання SVM, навчання не зводиться до задачі квадратичного програмування в опуклою області, що має єдине рішення.
5. На відміну від ART, не вирішує проблему пластичності-стабільності.

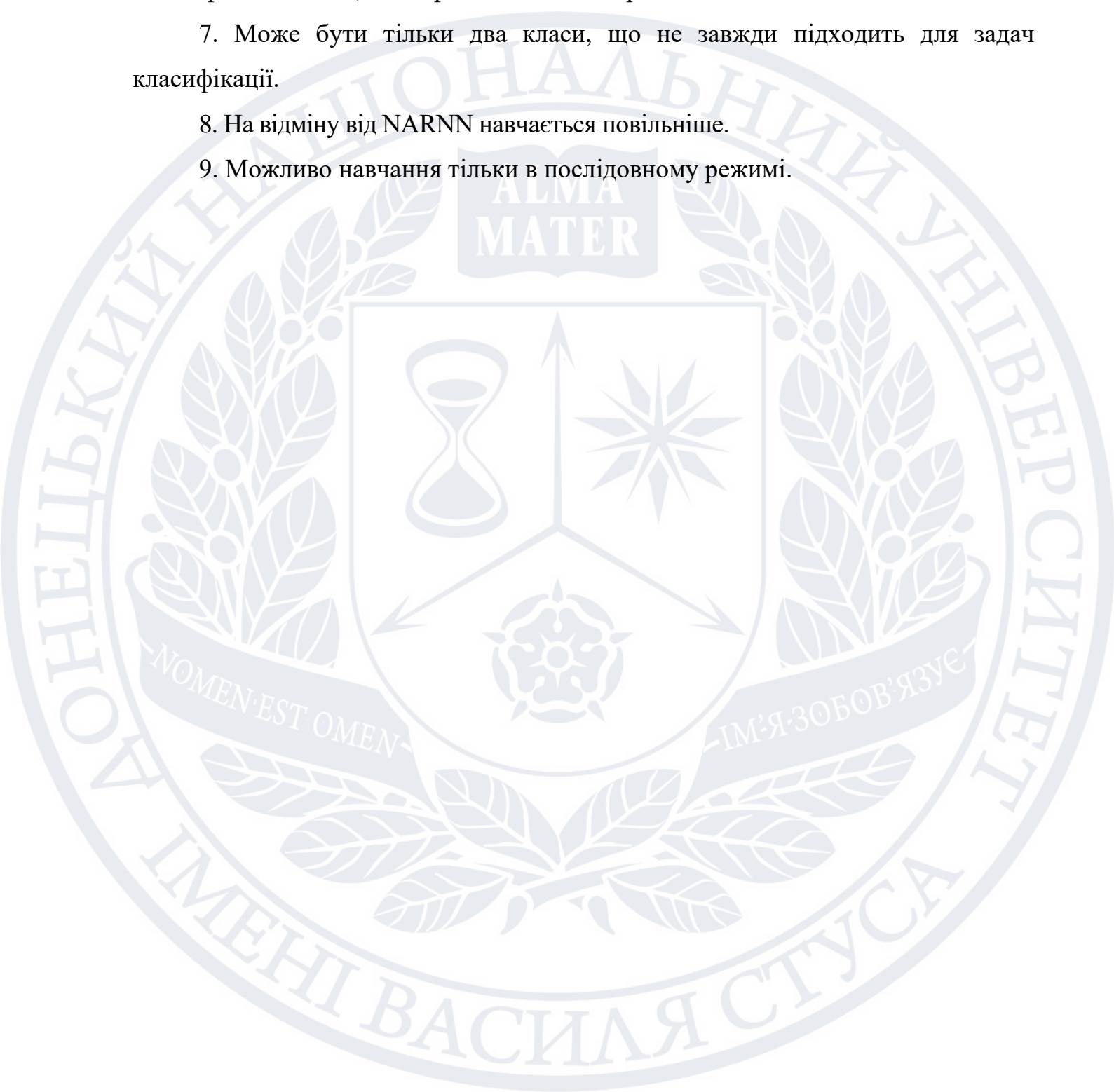


6. На відміну від LSTM, GRU, ESN, LSM, BRNN, NTM гірше вирішує завдання адаптивної фільтрації, прогнозу, адаптивного управління, параметричної ідентифікації моделі, класифікації нестационарних сигналів.

7. Може бути тільки два класи, що не завжди підходить для задач класифікації.

8. На відміну від NARNN навчається повільніше.

9. Можливо навчання тільки в послідовному режимі.



## Тестування

Для тестування виберемо пункт «Каталог» в якому відображаються усі товари. Виберемо довільний товар і натиснемо кнопку «Добавити в кошик». Як ми бачимо, товар добавився (Рисунок 10.1)

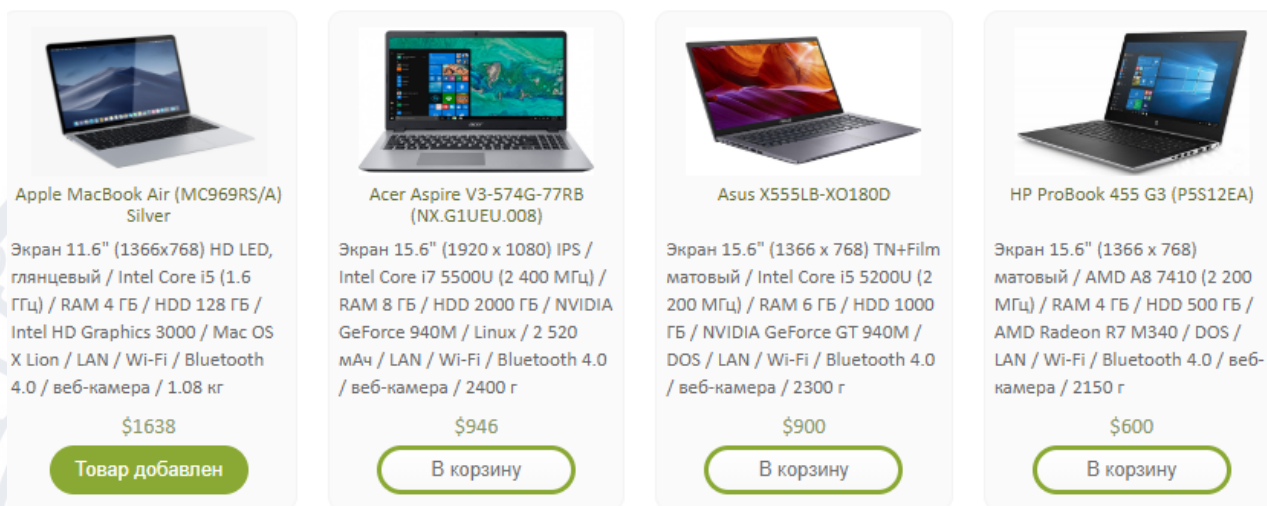


Рисунок 10.1

Перейдемо в «Кошик», та зробимо замовлення. (Рисунок 10.2).

**Способ оплаты:**

Наличными  
Оплата производится только в магазине.

Через нашу платежную систему  
Комиссия +0% от стоимости заказа. После оформления заказа вы получите его уникальный номер, который нужно указать при оплате.

**Доставка:**

Самовывоз  
Бесплатно. Забрать заказ можно в нашем магазине в Виннице.

Доставка по Виннице  
Стоимость — 50 грн. Бесплатно при заказе от 15000 грн. Доставка по Виннице в течение 24-х часов.

Доставка по Украине  
Стоимость — 500 грн. Доставка до 40 кг по Украине в течение 3-х дней.

**Контактные данные:**

Имя

Номер

Рисунок 10.2

Тепер ми можемо зайти до нашої бази даних і подивитися хто здійснив покупки (Рисунок 10.3).

Показати все | Число рядків: 50 ▼ | Фільтрувати рядки: Шукати в таблиці

+ Параметри

				id	pay_type	delivery_type	name	tel
<input type="checkbox"/>				11	1	1	Макс	2211
<input type="checkbox"/>				13	1	1	Макс	1234567
<input type="checkbox"/>				14	1	2	Макс	380631923333

Перевірити все | Вибрані:

Рисунок 10.3

## Висновки

Під час магістерської роботи були отримані навички роботи з бібліотеками JavaScript (jQuery, Ajax), Bootstrap, було розроблено «Базу даних Інтернет магазину» при використанні PHP+MySQL (phpMyAdmin).

Створено інформаційну систему «Магазин електроніки та побутової техніки» для автоматизації діяльності та забезпечення зберігання, збору, доставки інформації та прогнозування продажів за допомогою нейронної мережі.

Було зроблено БД, в якій можна легко подивитись необхідні дані, додати нові, редагувати їх і по необхідності видалити їх. Дані представлені в зручному для користувача вигляді. Інтерфейс програми був побудований у вигляді веб-сторінки. Розроблена база даних дозволяє виконати всі дії, які були необхідні.

База даних є найважливішим компонентом будь-якої інформаційної системи. База даних дозволяє структуровано зберігати великі обсяги інформації конкретної компанії, що значно спрощує ведення звітності та створення архівів. Оптимізовані бази даних значно підвищують продуктивність програм, які будуються на їхньому використанні.

Успішні та великі компанії вже не можуть обходитися без чітко структурованої інформаційної системи свого бізнесу. Бази даних, побудовані на SQL Server, відповідають високим вимогам до продуктивності та безпеки.

В майбутньому планується виконати такі завдання для модернізації додатку:

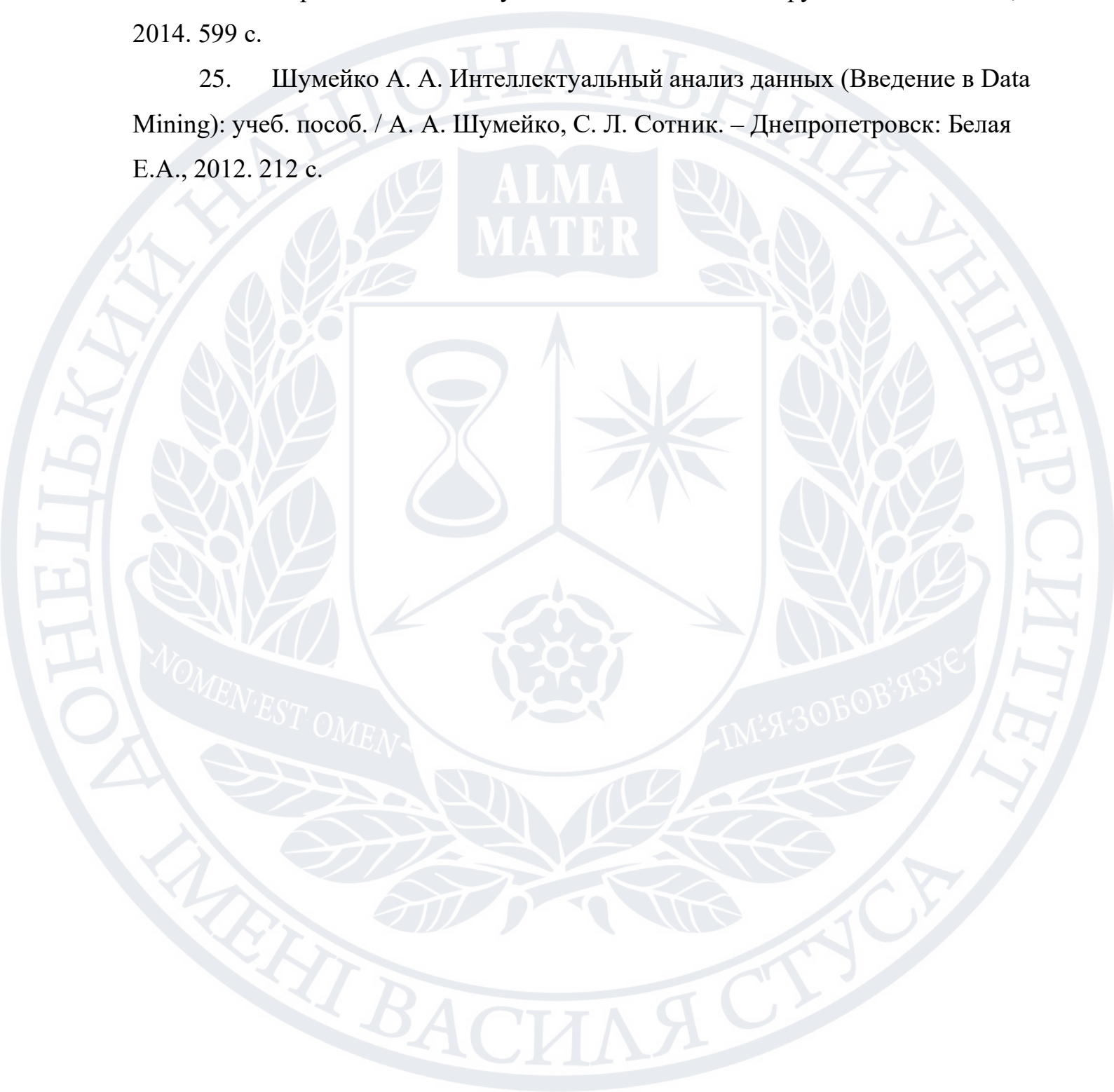
- ✓ Покращити інтерфейс
- ✓ Додати нові дані в БД

## Список використаних посилань

1. PHP Tutorial [Електронний ресурс]. Режим доступу:  
<https://www.w3schools.com/PHP/DEfaULT.asp> - Дата доступу: 01.05.2020
2. Java Script Tutorial [Електронний ресурс]. Режим доступу:  
<https://www.w3schools.com/js/default.asp> - Дата доступу: 01.05.2020
3. SQL Tutorial [Електронний ресурс]. Режим доступу:  
<https://www.w3schools.com/sql/default.asp> - Дата доступу: 01.05.2020
4. bootstrap Tutorial [Електронний ресурс]. Режим доступу:  
[https://www.w3schools.com/bootstrap/bootstrap\\_ver.asp](https://www.w3schools.com/bootstrap/bootstrap_ver.asp) - Дата доступу:  
01.05.2020
5. jQueryTutorial [Електронний ресурс]. Режим доступу:  
<https://www.w3schools.com/jquery/default.asp> - Дата доступу: 01.05.2020
6. Anushree R. Survey on Big Data Mining Algorithms. // International Journal for Research in Applied Science and Engineering Technology, 2019. N7. P. 1363-1367.
7. Bathla G., Aggarwal H., Rani R. Evolution of Data Analytics Techniques: From Data Mining to Big Data Mining // International Journal of Advanced Science and Technology, 2019. T 28. N19. P. 779-795.
8. Becerra-Bonache L., Bel-Enguix G., Jimenez-Lopez M.D., Martín-Vide C.: Formal Grammars and Languages. In R. Mitkov (Ed): The Oxford Handbook of Computational Linguistics (2nd Edition), Oxford University Press, 2018. P. 1-31.
9. Bevilacqua F. Finite-State Machines: Theory and Implementation / Game Design & Development Tutorials. 2013. URL:  
<https://gamedevelopment.tutsplus.com/tutorials/finite-state-machines-theory-and-implementation--gamedev-11867>.
10. Fayyad U., Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases. // AI Magazine, 1996. T17. N3. P. 37-54.
11. Fernau H. Algorithms for learning regular expressions from positive data // Information and Computation, 2009. N 207. P. 521-541.

12. Haryana R. Big Data Mining: a Comprehensive Analysis // International Journal of Creative Research Thoughts (IJCRT), 2018. Vol. 6(2). P. 398-404.
13. Hernandez I., Zhang Y. Using predictive analytics and big data to optimize pharmaceutical outcomes // American Journal of Health-System Pharmacy, 2017. Vol. 74. N 18. P. 1494-1500.
14. Ian H. Witten, Eibe Frank and Mark A. Hall. Data Mining: Practical Machine Learning Tools and Techniques. – 3rd Edition. – Morgan Kaufmann, 2011. 664 p.
15. Piatetsky-Shapiro G., Parker G. Lesson: Data Mining and Knowledge Discovery: An Introduction. Introduction to Data Mining. // KD Nuggets. 2011. 660 p.
16. Solanki H. Comparative Study of Data Mining Tools and Analysis with Unified Data Mining Theory // Int. J. Comput. Appl., 2013. N 75(16). P. 23-28.
17. Umair S., Haseeb Q. A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA) // International Journal of Innovation and Scientific Research, 2014. Vol.12. N1. P. 217-222.
18. Yashasree T., Kumar H., Hemantha K. A review on Data Mining & Big Data Analytics. // International Journal of Engineering & Technology. 2018. N7. P. 92- 94.
19. Багутдинов Р., Нарек С., Краснопахтыч М. Аналитика, инструменты и интеллектуальный анализ больших разнородных и разномасштабных данных // Экономика. Информатика, 2020. Т.47(4). С. 792-802.
20. Клейнберг Дж., Тардос Е. Алгоритмы: разработка и применение. Классика Computers Science / Пер. с англ. Е. Матвеева. СПб.: Питер, 2016. 800 с.
21. Козак Л. І. Основи програмування: навчальний посібник / Л. І. Козак, І. В. Костюк, С. П. Стасевич. – Л.: Новий Світ-2000, 2020. 328 с.
22. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. К., 2017. 150 с.

23. Матвієнко М. П. Теорія алгоритмів. Навчальний посібник. – К.: Видавництво Ліра-К, 2017. 340 с.
24. Черняк О.І. Інтелектуальний аналіз даних: підручник. – К: Знання, 2014. 599 с.
25. Шумейко А. А. Интеллектуальный анализ данных (Введение в Data Mining): учеб. пособ. / А. А. Шумейко, С. Л. Сотник. – Днепропетровск: Белая Е.А., 2012. 212 с.



## Додаток А

## SQL

```
-- phpMyAdmin SQL Dump
-- version 4.9.0.1
-- https://www.phpmyadmin.net/
--
-- Хост: 127.0.0.1:3306
-- Час створення: Трв 08 2020 р., 19:16
-- Версія сервера: 10.3.13-MariaDB-log
-- Версія PHP: 7.1.32

SET SQL_MODE = "NO_AUTO_VALUE_ON_ZERO";
SET AUTOCOMMIT = 0;
START TRANSACTION;
SET time_zone = "+00:00";

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!40101 SET NAMES utf8mb4 */;

--
-- База даних: `noteshop`
--
-----
--
-- Структура таблиці `admin`
--
CREATE TABLE `admin` (
  `username` varchar(128) NOT NULL,
  `password` varchar(128) NOT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8;

--
-- Дамп даних таблиці `admin`
--
INSERT INTO `admin` (`username`, `password`) VALUES
('admin', 'admin');
```



-----  
--  
-- Структура таблиці `lookup`  
--

```
CREATE TABLE `lookup` (  
  `id` int(11) NOT NULL,  
  `label` varchar(50) NOT NULL,  
  `value` varchar(50) NOT NULL,  
  `type` varchar(50) NOT NULL,  
  `position` int(11) NOT NULL  
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

--  
-- Дамп даних таблиці `lookup`  
--

```
INSERT INTO `lookup` (`id`, `label`, `value`, `type`, `position`) VALUES  
(1, 'Наличными', '1', 'PayType', 1),  
(2, 'Через нашу платежную систему', '2', 'PayType', 2),  
(3, 'Самовывоз', '1', 'DeliveryType', 1),  
(4, 'Доставка по Виннице', '2', 'DeliveryType', 2),  
(5, 'Доставка по Украине', '3', 'DeliveryType', 3),  
(6, '11.6"', '11.6', 'ScreenType', 1),  
(7, '12.5"', '12.5', 'ScreenType', 2),  
(8, '13.3"', '13.3', 'ScreenType', 3),  
(9, '15.6"', '15.6', 'ScreenType', 4),  
(10, '17.3"', '17.3', 'ScreenType', 5),  
(11, 'Intel Celeron', 'intel-celeron', 'CpuType', 1),  
(12, 'Intel Pentium', 'intel-pentium', 'CpuType', 2),  
(13, 'Intel i3', 'intel-i3', 'CpuType', 3),  
(14, 'Intel i5', 'intel-i5', 'CpuType', 4),  
(15, 'Intel i7', 'intel-i7', 'CpuType', 5),  
(16, 'AMD', 'amd', 'CpuType', 6),  
(17, '1 ГБ', '1', 'RamType', 1),  
(18, '2 ГБ', '2', 'RamType', 2),  
(19, '4 ГБ', '4', 'RamType', 3),  
(20, '6 ГБ', '6', 'RamType', 4),  
(21, '8 ГБ', '8', 'RamType', 5),  
(22, 'Windows 7', 'windows-7', 'OsType', 1),  
(23, 'Windows 8', 'windows-8', 'OsType', 2),  
(24, 'Windows 10', 'windows-10', 'OsType', 3),  
(25, 'Linux', 'linux', 'OsType', 4),  
(26, 'Mac OS', 'mac-os', 'OsType', 5),
```

```
(27, 'DOS', 'dos', 'OsType', 6),  
(28, 'Acer', 'Acer', 'BrandType', 1),  
(29, 'Apple', 'Apple', 'BrandType', 2),  
(30, 'Asus', 'Asus', 'BrandType', 3),  
(31, 'Dell', 'Dell', 'BrandType', 4),  
(32, 'HP', 'HP', 'BrandType', 5),  
(33, 'Lenovo', 'Lenovo', 'BrandType', 6),  
(34, 'Samsung', 'Samsung', 'BrandType', 7),  
(35, 'Sony', 'Sony', 'BrandType', 8);
```

```
-----  
--  
-- Структура таблиці `order`  
--
```

```
CREATE TABLE `order` (  
  `id` int(11) NOT NULL,  
  `pay_type` varchar(15) NOT NULL,  
  `delivery_type` varchar(15) NOT NULL,  
  `name` varchar(50) NOT NULL,  
  `tel` varchar(50) NOT NULL  
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

```
--  
-- Дамп даних таблиці `order`  
--
```

```
INSERT INTO `order` (`id`, `pay_type`, `delivery_type`, `name`, `tel`) VALUES  
(11, '1', '1', 'Макс', '2211'),  
(13, '1', '1', 'Макс', '1234567'),  
(14, '1', '2', 'Макс', '380631923333');
```

```
-----  
--  
-- Структура таблиці `product`  
--
```

```
CREATE TABLE `product` (  
  `id` int(11) NOT NULL,  
  `price` int(11) NOT NULL,  
  `image_source` varchar(100) NOT NULL,  
  `brand` varchar(30) NOT NULL,  
  `model` varchar(40) NOT NULL,  
  `screen_diagonal` varchar(5) NOT NULL,
```

```

`screen_desc` varchar(40) NOT NULL,
`cpu` varchar(50) NOT NULL,
`cpu_type` varchar(20) NOT NULL,
`ram` int(11) NOT NULL,
`hdd` int(11) NOT NULL,
`video_card` varchar(40) NOT NULL,
`os` varchar(30) NOT NULL,
`os_type` varchar(20) NOT NULL,
`additional_desc` varchar(100) NOT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8;

--
-- Дамп даних таблиці `product`
--

INSERT INTO `product` (`id`, `price`, `image_source`, `brand`, `model`, `screen_diagonal`,
`screen_desc`, `cpu`, `cpu_type`, `ram`, `hdd`, `video_card`, `os`, `os_type`, `additional_desc`)
VALUES
(1, 1638, 'images/notebook.jpg', 'Apple', 'MacBook Air (MC969RS/A) Silver', '11.6',
'(1366x768) HD LED, глянцева', 'Intel Core i5 (1.6 ГГц)', 'intel-i5', 4, 128, 'Intel HD
Graphics 3000', 'Mac OS X Lion', 'mac-os', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 1.08
кг'),
(2, 946, 'images/acer_aspire.jpg', 'Acer', 'Aspire V3-574G-77RB (NX.G1UEU.008)', '15.6',
'(1920 x 1080) IPS', 'Intel Core i7 5500U (2 400 МГц)', 'intel-i7', 8, 2000, 'NVIDIA GeForce
940M', 'Linux', 'linux', '2 520 мАч / LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2400 г'),
(3, 900, 'images/asus_x5.jpg', 'Asus', 'X555LB-XO180D', '15.6', '(1366 x 768) TN+Film
матовий', 'Intel Core i5 5200U (2 200 МГц)', 'intel-i5', 6, 1000, 'NVIDIA GeForce GT 940M',
'DOS', 'dos', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2300 г'),
(4, 600, 'images/hp_probook.jpg', 'HP', 'ProBook 455 G3 (P5S12EA)', '15.6', '(1366 x 768)
матовий', 'AMD A8 7410 (2 200 МГц)', 'amd', 4, 500, 'AMD Radeon R7 M340', 'DOS', 'dos',
'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2150 г'),
(5, 674, 'images/hp_pavilion.jpg', 'HP', 'Pavilion 15-n079sr (F2U22EA)', '15.6', '(1366 x 768)
TN+Film', 'Intel Core i5 4200U (1 600 МГц)', 'intel-i5', 8, 1000, 'NVIDIA GeForce GT 740M',
'DOS', 'dos', '2 620 мАч / LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2280 г'),
(6, 542, 'images/lenovo.jpg', 'Lenovo', 'Z50-75 (80EC00EHUA)', '15.6', '(1920 x 1080)
TN+Film', 'AMD A10 7300 (1 900 МГц)', 'amd', 4, 500, 'AMD Radeon R6 M255DX', 'Windows
8', 'windows-8', '2 200 мАч / LAN / Wi-Fi / веб-камера / 2400 г'),
(7, 819, 'images/acer_aspire.jpg', 'Acer', 'Aspire E5-551G-F25F (NX.MLEEU.013)', '15.6',
'(1920 x 1080) TN+Film', 'AMD FX-7500 (2 100 МГц)', 'amd', 16, 1000, 'AMD Radeon R7
M265', 'Linux', 'linux', '2 500 мАч / LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2500 г'),
(8, 324, 'images/acer_aspire.jpg', 'Acer', 'Aspire ES1-531-C007 (NX.MZ8EU.011)', '15.6',
'(1366 x 768) TN+Film матовий', 'Intel Celeron N3050 (1 600 МГц)', '', 2, 500, 'Intel HD
Graphics', 'Linux', 'linux', '3 220 мАч / LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2400 г'),
(9, 2140, 'images/notebook.jpg', 'Apple', 'MacBook Pro 13 Retina (MF840)', '13.3', '(2560 x
1600) IPS', 'Intel Core i5 (2 700 МГц)', 'intel-i5', 8, 256, 'Intel Iris Graphics 6100', 'OS X
(Yosemite)', 'mac-os', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 1580 г'),

```

(10, 424, 'images/asus\_x5.jpg', 'Asus', 'X200MA-KX622B', '11.6', '(1366 x 768) TN+Film', 'Intel Pentium N3540 (2 160 МГц)', 'intel-pentium', 2, 500, 'Intel HD Graphics', 'Windows 8', 'windows-8', '3 300 мАч / LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 1240 г'),

(11, 547, 'images/asus\_x5.jpg', 'Asus', 'F555LB-XO557D', '15.6', '(1366 x 768) TN+Film', 'Intel Core i3 4005U (1 700 МГц)', 'intel-i3', 4, 500, 'NVIDIA GeForce 940M', 'Windows 8', 'windows-8', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2300 г'),

(12, 760, 'images/hp\_probook.jpg', 'HP', 'ProBook 450 G1 (F7Y23ES)', '15.6', '(1366 x 768) TN+Film', 'Intel Pentium 3550M (2 300 МГц)', 'intel-pentium', 6, 750, 'AMD Radeon HD 8750M', 'Windows 8', 'windows-8', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2370 г'),

(13, 683, 'images/hp\_pavilion.jpg', 'HP', 'Pavilion 15-p079sr (J5C01EA)', '15.6', '(1366x768) глянец', 'Intel Core i5 4210U (1700 МГц)', 'intel-i5', 8, 1000, 'NVIDIA GeForce 840M', 'DOS', 'dos', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2.27 кг'),

(14, 552, 'images/hp\_pavilion.jpg', 'HP', 'Pavilion 15-p266ur (L2V61EA)', '15.6', '(1366 x 768) TN+Film', 'Intel Core i3 5010U (2 100 МГц)', 'intel-i3', 4, 500, 'Intel HD Graphics 5500', 'DOS', 'dos', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2270 г'),

(15, 576, 'images/hp\_pavilion.jpg', 'HP', 'Pavilion 15-p028sr (J6Z25EA)', '15.6', '(1366x768) глянец', 'AMD A8 6410 (2000 МГц)', 'amd', 6, 750, 'AMD Radeon R7 M260', 'Windows 8', 'windows-8', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2.27 кг'),

(16, 602, 'images/hp\_probook.jpg', 'HP', 'ProBook 455 G2 (G6W40EA)', '15.6', '(1366 x 768) TN+Film', 'AMD A6 Pro 7050B (2 200 МГц)', 'amd', 4, 500, 'AMD Radeon R5 M255', 'Windows 8', 'windows-8', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2200 г'),

(17, 605, 'images/lenovo.jpg', 'Lenovo', 'Z50-70 (59421900)', '15.6', '(1366 x 768) TN+Film', 'Intel Pentium 3558U (1 700 МГц)', 'intel-pentium', 4, 1000, 'NVIDIA GeForce 840M', 'DOS', 'dos', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2400 г'),

(18, 662, 'images/lenovo.jpg', 'Lenovo', 'G50-70 (59423447)', '15.6', '(1366 x 768) TN+Film', 'Intel Core i3 4030U (1 900 МГц)', 'intel-i3', 4, 500, 'Intel HD Graphics 4400', 'Windows 8.1', 'windows-8', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2500 г'),

(19, 540, 'images/lenovo.jpg', 'Lenovo', 'G50-70 (59433722)', '15.6', '(1366 x 768) TN+Film', 'Intel Core i3 4005U (1 700 МГц)', 'intel-i3', 4, 500, 'AMD Radeon R5 M230', 'DOS', 'dos', 'LAN / Wi-Fi / Bluetooth 4.0 / веб-камера / 2500 г');

```

-----
--
-- Структура таблиці `product_order`
--
CREATE TABLE `product_order` (
  `id` int(11) NOT NULL,
  `product_id` int(11) NOT NULL,
  `order_id` int(11) NOT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
--
-- Дамп даних таблиці `product_order`
--

```

```
INSERT INTO `product_order` (id, `product_id`, `order_id`) VALUES
(23, 19, 11),
(25, 2, 13),
(26, 1, 14);

--
-- Индекси збережених таблиць
--
--
-- Индекси таблиці `lookup`
--
ALTER TABLE `lookup`
  ADD PRIMARY KEY (id`);

--
-- Индекси таблиці `order`
--
ALTER TABLE `order`
  ADD PRIMARY KEY (id`);

--
-- Индекси таблиці `product`
--
ALTER TABLE `product`
  ADD PRIMARY KEY (id`);

--
-- Индекси таблиці `product_order`
--
ALTER TABLE `product_order`
  ADD PRIMARY KEY (id`),
  ADD KEY `product_id` (product_id`),
  ADD KEY `product_id_2` (product_id`),
  ADD KEY `order_id` (order_id`);

--
-- AUTO_INCREMENT для збережених таблиць
--
--
-- AUTO_INCREMENT для таблиці `lookup`
--
ALTER TABLE `lookup`
  MODIFY `id` int(11) NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=36;
```

```
--  
-- AUTO_INCREMENT для таблиці `order`  
--  
ALTER TABLE `order`  
  MODIFY `id` int(11) NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=15;  
--  
-- AUTO_INCREMENT для таблиці `product`  
--  
ALTER TABLE `product`  
  MODIFY `id` int(11) NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=24;  
--  
-- AUTO_INCREMENT для таблиці `product_order`  
--  
ALTER TABLE `product_order`  
  MODIFY `id` int(11) NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=27;  
--  
-- Обмеження зовнішнього ключа збережених таблиць  
--  
-- Обмеження зовнішнього ключа таблиці `product_order`  
--  
ALTER TABLE `product_order`  
  ADD CONSTRAINT `product_order_ibfk_1` FOREIGN KEY (`product_id`)  
REFERENCES `product` (`id`) ON DELETE CASCADE ON UPDATE CASCADE,  
  ADD CONSTRAINT `product_order_ibfk_2` FOREIGN KEY (`order_id`) REFERENCES  
`order` (`id`) ON DELETE CASCADE ON UPDATE CASCADE;  
COMMIT;  
  
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;  
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;  
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
```

Пилипчак Максим Володимирович

Прізвище, ім'я по батькові

Інформаційних і прикладних технологій

Факультет

122 Комп'ютерні науки

Шифр і назва спеціальності

Комп'ютерні технології обробки даних

Освітня програма

## ДЕКЛАРАЦІЯ АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ

Усвідомлюючи свою відповідальність за надання неправдивої інформації, стверджую, що подана кваліфікаційна (магістерська) робота на тему «ДОСЛІДЖЕННЯ СИСТЕМИ ІНТЕРНЕТ-МАГАЗИНУ НА ПІДСТАВІ МЕТОДІВ DATA SCIENCE» є написаною мною особисто.

Одночасно заявляю, що ця робота:

- не передавалась іншим особам і подається до захисту вперше;
- не порушує авторських та суміжних прав, закріплених статтями 21-25 Закону України «Про авторське право та суміжні права»;
- не отримувалась іншими особами, а також дані та інформація не отримувалась у недозволеній спосіб.

Я усвідомлюю, що у разі порушення цього порядку моя кваліфікаційна робота буде відхилена без права її захисту, або під час захисту за неї буде поставлена оцінка «незадовільно».

\_\_\_\_\_ (дата)

\_\_\_\_\_ (підпис здобувача освіти)