

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки
Кафедра ЕОМ

**МЕТОДИ ПРОЄКТУВАННЯ ТА РОЗРОБКИ
ПРОГРАМНИХ КОМПОНЕНТІВ ДЛЯ СИСТЕМИ КЛАСИФІКАЦІЇ
КЛІЄНТІВ КОМПАНІЇ**

Кваліфікаційна робота
Другий (магістерський) рівень

Автор:
Ст.гр. СПзм-23-1
Малишенко Д.О.

Керівник:
Доцент кафедри ЕОМ
Шматко О.В.

Харків
2025

1

Актуальність теми

- ◆ **Зростаюча роль клієнтських даних**
 - У сучасному бізнес-середовищі ефективною робота з клієнтами базується на аналізі великих обсягів інформації.
- ◆ **Необхідність автоматизації**
 - ▶ Компанії потребують інструментів для швидкої та точної класифікації клієнтів з метою покращення обслуговування і збільшення прибутку.
- ◆ **Розвиток цифрових технологій**
 - ▶ Машинне навчання та аналітика даних відкривають нові можливості для персоналізованого підходу до клієнтів.
- ◆ **Потреба у програмних рішеннях**
 - ▶ Існує нагальна потреба у розробці гнучких та масштабованих програмних компонентів для інтеграції в IT-інфраструктуру компанії.



2

Об'єкта та предмет

Об'єктом дослідження є
методи та системи
класифікації клієнтів компанії.

Предметом дослідження є
методи та інструментальні
засоби розробки
програмного забезпечення
для побудови ефективних
моделей класифікації клієнтів
на основі аналізу даних.

3

Мета та задачі роботи

Метою дипломної роботи є підвищення точності класифікації клієнтів шляхом впровадження алгоритмів машинного навчання у процес проектування програмних компонентів.

Для досягнення поставленої мети необхідно виконати такі основні **задачі**:

1. Проаналізувати існуючі підходи до класифікації та сегментації клієнтів у бізнес-середовищі, зокрема методи, що базуються на інтелектуальному аналізі даних та алгоритмах машинного навчання.
2. Обґрунтувати вибір найбільш доцільних алгоритмів машинного навчання для задач класифікації клієнтів.
3. Розробити архітектуру програмної системи класифікації клієнтів, визначити основні компоненти та їх функціональну взаємодію.
4. Реалізувати програмний прототип системи, що здійснює обробку вхідних даних, навчання моделі та автоматичну класифікацію клієнтів за визначеними ознаками.
5. Провести тестування створеної системи, оцінити точність класифікації за допомогою відповідних метрик (точність, повнота, F1-оцінка, AUC тощо).

4

Порівняльна таблиця методів кластеризації

Метод	Необхідність задавати K	Виявлення викидів	Швидкість	Робота з нестандартними формами кластерів	Інтерпретованість	Підходить для візуалізації
K-means	Так	Ні	Висока	Ні	Висока	Частково
Ієрархічна	Ні	Частково	Середня	Так	Висока	Так
DBSCAN	Ні	Так	Середня	Так	Середня	Так
Fuzzy C-means	Так	Ні	Середня	Частково	Низька	Ні
SOM	Ні	Частково	Середня	Так	Середня	Так
PCA + кластеризація	Так (у кластеризації)	Залежить від алгоритму	Висока/середня	Так	Висока	Так

5

Функціональні вимоги

Актор	Сценарій використання	Функціональні вимоги
Маркетолог	Перегляд результатів класифікації	F4 – Візуалізація сегментів
	Формування таргетованих кампаній	F6 – Експорт класифікованих даних
Бізнес-аналітик	Завантаження набору даних	F1 – Імпорт вхідних даних
	Аналіз активності клієнтів у сегментах	F5 – Виведення метрик класифікації
	Побудова нової моделі сегментації	F2 – Навчання моделі класифікації
Системний адміністратор	Налаштування алгоритмів класифікації	F3 – Вибір параметрів моделі
	Моніторинг продуктивності системи	F7 – Журналізація дій користувачів

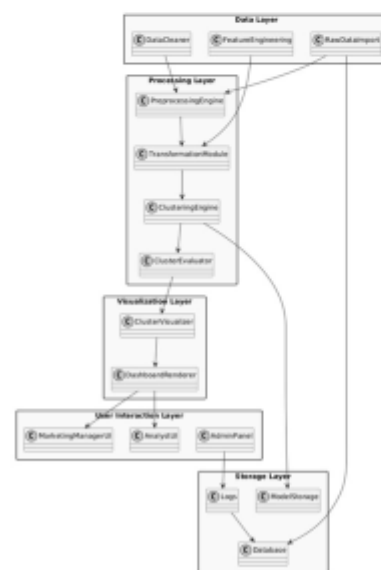
6

Нефункціональні вимоги

№	Нефункціональна вимога	Критерій вимірювання	Цільове значення
1	Масштабованість	Кількість клієнтів у базі	Підтримка понад 1 млн записів без зниження продуктивності
2	Продуктивність класифікації	Час виконання класифікації	Не більше 2 хвилин для 100 тис. записів
3	Безпека даних	Рівень доступу користувача	Розмежування доступу за ролями, логування дій
4	Інтеграція	Підтримка стандартів обміну	REST API, CSV, JSON
5	Надійність	Uptime системи	Не менше 99.5% на місяць
6	Візуалізація результатів	Наявність інтерфейсу аналітики	Дашборди, графіки, сегментні таблиці
7	Гнучкість налаштувань	Можливість зміни алгоритму класифікації	Підтримка змін без перезапуску системи

7

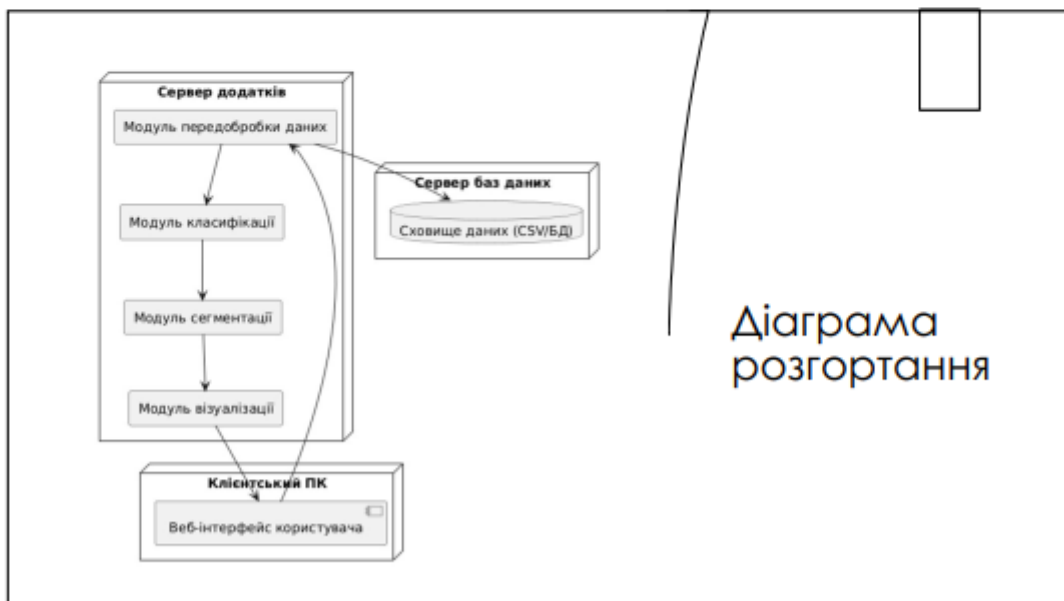
Архітектура системи



8



9



10



11

№	Назва змінної	Тип даних	Формат / Тип значень	Опис
1	InvoiceNo	Намітована	Ціле число (6 цифр), текст	Унікальний номер рахунку-фактури. Початок з "C" означає скасування.
2	StockCode	Намітована	Ціле число (5 цифр), текст	Ідентифікатор товарної позиції.
3	Description	Намітована	Текстовий рядок	Назва або опис товару.
4	Quantity	Кількісна	Ціле число	Кількість придбаних одиниць товару.
5	InvoiceDate	Часова	Дата і час (YYYY-MM-DD HH:mm:ss)	Час створення рахунку-фактури.
6	UnitPrice	Кількісна	Дісне число (float)	Ціна за одиницю товару (у фунтах стерлінгів).
7	CustomerID	Намітована	Ціле число (5 цифр), текст	Унікальний ідентифікатор клієнта.
8	Country	Намітована	Текстовий рядок	Назва країни проживання клієнта.

Експериментальні данні

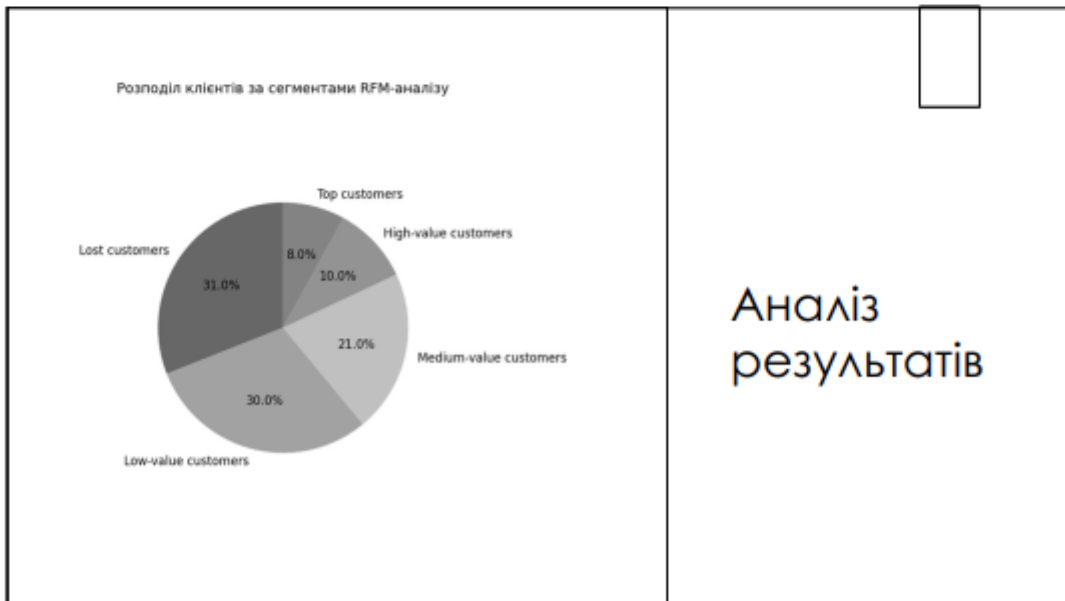
<https://archive.ics.uci.edu/dataset/352/online+retail>

12

Customer ID	Recency	Frequency	Monetary (£)
12346	325	1	77,183.60
12347	1	182	4,310.00
12348	74	31	1,797.24
12349	18	73	1,757.55
12350	309	17	334.40

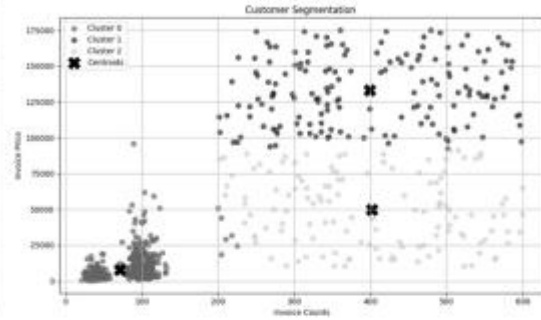
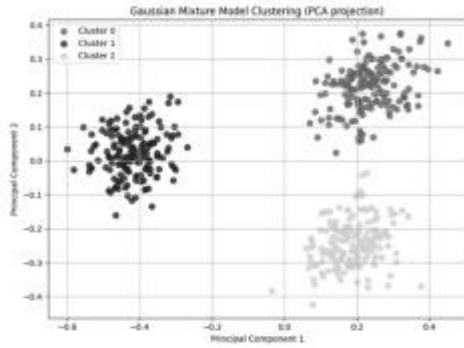
Аналіз
результатів

13



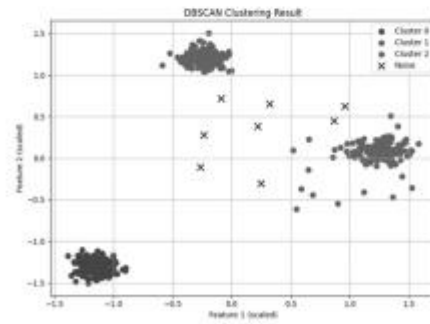
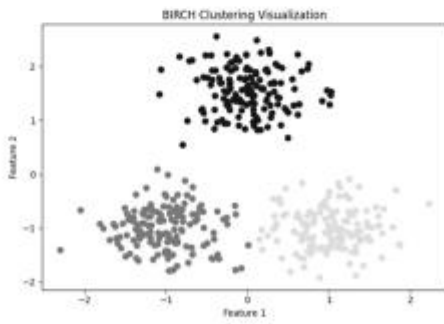
14

Аналіз результатів



15

Аналіз результатів



16

Висновки:

- ✓ **Комплексне дослідження.** Проведено аналіз та реалізацію методів кластеризації для системи класифікації клієнтів компанії з використанням підходів машинного навчання.
- ✓ **Теоретичне обґрунтування.** Оцінено сучасні алгоритми: K-Means, GMM, DBSCAN, BIRCH. Визначено їх придатність для задач сегментації.
- ✓ **Практична реалізація.** Розроблено прототип системи з повним циклом підготовки даних.
- ✓ **Оцінка ефективності.** Найвищий Silhouette Score = 0.80 продемонструвала модель GMM. DBSCAN був корисний для виявлення аномалій, BIRCH — ефективний при великих об'ємах даних.
- ✓ **Бізнес-застосування.** Профілювання клієнтів дозволило сформувати рекомендації для персоналізованих маркетингових стратегій та покращення взаємодії з клієнтами.
- ✓ **Практичне значення.** Розроблене рішення може бути інтегроване в IT-інфраструктуру компанії у сферах e-commerce, банківських послуг, телекомунікацій тощо.
- ✓ **Перспективи розвитку.** Можливе подальше розширення системи через інтеграцію з предиктивною аналітикою та рекомендаційними системами.

ДОДАТОК Б
Порівняння методів класифікації

Таблиця Б.1 - Порівняльна таблиця методів кластеризації

Метод	Необхідність задавати К	Виявлення викидів	Швидкість	Робота з нестандартними формами кластерів	Інтерпретованість	Підходить для візуалізації
K-means	Так	Ні	Висока	Ні	Висока	Частково
Ієрархічна	Ні	Частково	Середня	Так	Висока	Так
DBSCAN	Ні	Так	Середня	Так	Середня	Так
Fuzzy C-means	Так	Ні	Середня	Частково	Низька	Ні
SOM	Ні	Частково	Середня	Так	Середня	Так
PCA + кластеризація	Так (у кластеризації)	Залежить від алгоритму	Висока/середня	Так	Висока	Так

Таблиця Б.2 - Порівняльна таблиця методів машинного навчання

Характеристика	Decision Tree (DT)	Support Vector Machine (SVM)	K-Nearest Neighbors (KNN)	Random Forest (RF)
1	2	3	4	5
Тип навчання	Контрольоване	Контрольоване	Контрольоване	Контрольоване
Тип задачі	Класифікація, регресія	Класифікація, регресія	Класифікація, регресія	Класифікація, регресія
Необхідність підготовки даних	Мінімальна	Середня	Мінімальна	Мінімальна
Стійкість до перенавчання	Низька	Висока	Низька	Висока
Інтерпретованість результату	Висока	Низька	Середня	Середня
Чутливість до шумів/викидів	Висока	Низька	Висока	Низька
Обробка високовимірних даних	Посередня	Відмінна	Посередня	Добра
Необхідність налаштування гіперпараметрів	Середня	Висока	Низька	Висока

Продовження таблиці Б.2

1	2	3	4	5
Обчислювальна складність	Низька	Висока	Висока	Висока
Швидкість класифікації	Висока	Середня	Низька	Висока
Підтримка категоріальних змінних	Так	Лише після кодування	Так	Так
Точність у дослідженні (%)	83.24%	80.75%	79.78%	89.61%
Застосування	Персоналізовані пропозиції	Комплексна класифікація	Рекомендаційні системи	Масштабні системи прогнозу

ДОДАТОК В
Наукова публікація

ISSN 2073-7394

Національний університет
"Полтавська політехніка імені Юрія Кондратюка"
National University
"Yuri Kondratyuk Poltava Polytechnic"

**Системи
управління,
навігації
та зв'язку**

Випуск 3 (81)

**Control,
navigation and
communication
systems**

Issue 3 (81)

Щоквартальне видання

Засноване у 2007 році

У журналі відображені результати наукових досліджень з розробки та удосконалення систем управління, навігації та зв'язку у різних проблемних галузях.

Засновник і видавець:
Національний університет
"Полтавська політехніка імені Юрія Кондратюка"

Телефон:
+38 (050) 302-20-71

E-mail редакції:
kuchuk_nina@ukr.net

Інформаційний сайт:
<http://journals.nupp.edu.ua/sunz>

Quarterly

Founded in 2007

Journal represent the research results on the development and improvement of control, navigation and communication systems in various areas

Founder and publisher:
National University
"Yuri Kondratyuk Poltava Polytechnic"

Phone:
+38 (050) 302-20-71

E-mail of the editorial board:
kuchuk_nina@ukr.net

Information site:
<http://journals.nupp.edu.ua/sunz>

За достовірність викладених фактів, цитат та інших відомостей відповідальність несе автор

Журнал індексується міжнародними наукометричними базами: Index Copernicus (ICV = 82.05),
General Impact Factor, Google Scholar, Academic Resource Index, Scientific Indexed Service

Затверджений до друку Вченою Радою Національного університету
"Полтавська політехніка імені Юрія Кондратюка" (протокол від 12 березня 2025 року № 5).

Ідентифікатор медіа R30-04135 згідно з рішенням Національної ради України
з питань телебачення і радіомовлення від 25.04.2024 № 1416

Включений до "Переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук, кандидата наук та ступеня доктора філософії" до категорії Б – наказами МОН України від 17.03.2020 № 409 та від 09.02.2021 № 157

Полтава • 2025

© Національний університет "Полтавська політехніка імені Юрія Кондратюка"

З·М·І·С·Т·16¶

НАВИГАЦІЯ ТА ГЕОІНФОРМАЦІЙНІ СИСТЕМИ 1¶

<i>Пащенко Р. Е., Матроско М. В.</i> Вибір розмірів «вікна» під час розрахунку фрактальної розмірності космічних знімків сільськогосподарських земель	5¶
---	----

АВТОМОБІЛЬНИЙ, РІЧКОВИЙ, МОРСЬКИЙ ТА АВІАЦІЙНИЙ ТРАНСПОРТ 1¶

<i>Булгаков М. П., Логінов О. В., Мельник О. М.</i> Багатокритеріальна оцінка стратегій модернізації суден для зменшення викидів CO ₂ на базі методу TOPSIS	11¶
--	-----

УПРАВЛІННЯ В СКЛАДНИХ СИСТЕМАХ 4¶

<i>Воронов А. О., Максимова Н. Г., Ухачков І. В., Безродний В. В., Ухачков В. І.</i> Інженерний кібернетичний компютер для моделювання та симуляції	27¶
<i>Євтушівко О. С., Захарович О. Ю.</i> Математична модель коливань рухомого складу на криволінійній ділянці шляху для бортової системи підтримки прийняття рішень	31¶
<i>Таранко Л. В., Таранко Б. Д., Ражнова Г. Г.</i> Моделі проектування біг-потоківого оцінювача ірраціональних функцій	36¶
<i>Ухачова Г. В., Ухачова І. В., Обрізан В. І., Шевченко О. Ю., Кулак Г. К.</i> Синтез дедуктивних формул для логічних примітивів	40¶
Якщо ...	44¶

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ 7¶

<i>Gordachuk V., Fedarov V.</i> Analysis of mathematical models of random graphs	76¶
<i>Dyachenko D., Korobeynikov M., Korobeynikov O., Kovalenko A., Kravchenko P.</i> Data processing methods in a corporate network	81¶
<i>Заїка В. В., Фікс О. В., Морозова О. І., Нікітіна Г. С.</i> Інтелектуальні веб-системи для автоматизації обробки та моніторингу інформаційних потоків	87¶
<i>Бірюк Г. А., Лисиця Д. О., Ігуд Д. М., Кириленко Д. В.</i> Інтелектуальний аналіз спеціалізованих комп'ютерних систем	97¶
<i>Сазаніков Д. В., Васильяков О. Г., Караман Д. Г.</i> Оптимізація множення квантованих одновісних матриць для великих моделей	101¶
<i>Shefer O., Yermilova N., Delychko O., Iermilov R., Parichko S.</i> Use of virtual measuring devices in metrology, electronics and electrical machines for the training of electrical engineering specialists	109¶
<i>Шпато О. В., Матвишук Д. О., Волощук О. Б.</i> Інформаційна система для інтелектуальної класифікації клієнтів: архітектура, реалізація та експериментальні дослідження	113¶

ЦИВІЛЬНА БЕЗПЕКА¶

Глава	121¶
-------	------

ЗВ'ЯЗОК, ТЕЛЕКОМУНІКАЦІЇ ТА РАДІОТЕХНІКА 3¶

<i>Аль-Амари А. Н., Іщенко Р. М., Дяченко П. В., Білик І. О.</i> Фізичні механізми створення технічних каналів витоку акустичної інформації	154¶
<i>Глоба Є. Ю., Сидінов В. Р., Натанський М. С.</i> Метод виявлення аномалій в корпоративній мережі	158¶
<i>Демченко О. І., Чумаченко С. В.</i> Технології прогресивного конструювання інтегральних мікросхем	165¶
<i>Васілішкін В., Татарова Г., Іашківська О.</i> Methods of data processing and analysis in a corporate network	171¶

АЛФАВІТНИЙ ПОКАЖЧИК	176¶
---------------------	------

Організації авторів¶

Федеральний інститут Парани, Парана, Бразилія¶
Вище професійне училище № 7, Кременчук, Україна¶
Державний торговельно-економічний університет, Київ, Україна¶
Державний університет «Київський авіаційний інститут», Київ, Україна¶
Київський національний університет будівництва і архітектури, Київ, Україна¶
Київський фаховий коледж зв'язу, Київ, Україна¶
Кременчуцький національний університет імені Михайла Остроградського, Кременчук, Україна¶

Authors affiliation¶

Instituto Federal do Parana, Parana, Brazil¶
Higher Vocational School No. 7, Kremenchuk¶
State University of Trade and Economics, Kyiv, Ukraine¶
State University "Kyiv Aviation Institute," Kyiv, Ukraine¶
Kyiv National University of Construction and Architecture, Kyiv, Ukraine¶
Kyiv Vocational College of Communication, Kyiv¶
Kremenchuk Mykhailo Ostrogradskyi National University, Kremenchuk, Ukraine¶