

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет інформатики та обчислювальної техніки**

**Кафедра інформаційних систем та технологій**

«На правах рукопису»  
УДК 004.891

До захисту допущено:

Завідувач кафедри

\_\_\_\_\_ Олександр РОЛІК

«\_\_» \_\_\_\_\_ 2023 р.

**Магістерська дисертація**  
**на здобуття ступеня магістра**  
**за освітньо-професійною програмою «Інформаційні управляючі системи**  
**та технології»**  
**зі спеціальності 126 «Інформаційні системи та технології»**  
**на тему: «Система підтримки роботи фахівця з маркетингу у соціальних**  
**мережах»**

Виконала:

студентка VI курсу, групи ІС-11мн  
Майборода Аріна Миколаївна \_\_\_\_\_

Керівник:

доцент, к.ф-м.н, доцент  
Гавриленко Олена Валеріївна \_\_\_\_\_

Рецензент:

ПШБ доцент, к.т.н, доцент  
Олійник Юрій Олександрович \_\_\_\_\_

Засвідчую, що у цій магістерській  
дисертації немає запозичень з праць  
інших авторів без відповідних  
посилань.

Студентка \_\_\_\_\_

Київ – 2023 року

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**  
**Факультет інформатики та обчислювальної техніки**  
**Кафедра інформаційних систем та технологій**

Рівень вищої освіти – другий (магістерський)

Спеціальність – 126 «Інформаційні системи та технології»

Освітньо-професійна програма «Інформаційні управляючі системи та технології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Олександр РОЛІК

« \_\_\_ » \_\_\_\_\_ 2023 р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**

**Майборода Аріна Миколаївна**

1. Тема дисертації «Система підтримки роботи фахівця з маркетингу у соціальних мережах», науковий керівник дисертації Гавриленко Олена Валеріївна, к.ф-м.н., доцент, затверджені наказом по університету від «20» березня 2023 р. № 1275-с.
2. Термін подання студентом дисертації.
3. Об'єкт дослідження: процес створення оптимального плану просування контенту за вхідними даними.
4. Предмет дослідження: модель сегментації факторів та користувачів соціальної мережі.
5. Перелік завдань, які потрібно розробити:
  - виконати огляд існуючих методів створення плану просування контенту;
  - сформулювати математичну постановку задачі;
  - здійснити порівняльний аналіз методів, які використовуються при побудові системи;
  - розв'язати задачу створення цільових аудиторій методом кластеризації;
  - розв'язати задачу визначення найвпливовіших факторів успішності методом головних компонент та алгоритмом прогнозування;
  - розв'язати задачу прогнозування очікуваної кількості приросту цільової аудиторії методом часових рядів;

- розв'язати задачу обрахунку КРІ;
  - розробити систему з просування контенту;
  - провести експерименти для дослідження методів;
  - виконати аналіз отриманих результатів.
6. Орієнтовний перелік графічного (ілюстративного) матеріалу:
- функціональна діаграма варіантів використання;
  - архітектура моделі системи підтримки фахівця з маркетингу у соціальних мережах;
  - алгоритм розв'язання задачі створення цільових аудиторій;
  - діаграма алгоритму порівняння методів кластеризації;
  - алгоритм розв'язання задачі визначення найвпливовіших факторів успішності;
  - архітектура методу прогнозування;
  - ER-діаграма бази даних;
  - графік активності розробки програмного забезпечення;
  - архітектура системи;
  - діаграма класів;
  - діаграма послідовностей дій користувача при створенні заявки та перегляду статистики за місяць;
  - діаграма послідовностей дій користувача при перегляді історії заявок;
  - діаграма послідовностей дій користувача при перегляді розрахунків КРІ;
  - креслення вигляду екранних форм.

7. Орієнтований перелік публікацій:

8. Дата видачі завдання: 01.01.2023

#### Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Систематизація результатів огляду літератури	15.02.2023	
2	Постановка та формалізація математичної моделі задачі	27.03.2023	
3	Розробка методів та програмного забезпечення	20.03.2023	
4	Проведення експериментальних досліджень розроблених алгоритмів	10.04.2023	
5	Оформлення документації	15.04.2023	
6	Подання роботи на попередній захист	19.04.2023	
7	Подання роботи на основний захист		

Студент

Аріна, МАЙБОРОДА

Науковий керівник

Олена, ГАВРИЛЕНКО

## РЕФЕРАТ

Магістерська дисертація: 105 с., 54 рис., 9 табл., 37 джерел, 2 додатків.

Актуальність. За останні декілька років, популярність та важливість задачі просування контенту в соціальних мережах значно зросла. Однак, з урахуванням поширення автоматизації процесів, виникає питання, чи можливо створити якісну рекламну кампанію, використовуючи стратегію, що базується на результаті прогностичних моделей та моделей кластеризації. Для того, щоб уникнути надмірного втручання фахівця з маркетингу в процес створення контенту, ми відповідаємо на це питання наступним чином. Однак, для успішного формування стратегії прогнозування потрібна якісна оцінка людиною, достатня кількість вхідних даних та чітко сформульована мета на етапі початку роботи.

До цього моменту ми не зустрічали системи з таким функціоналом, що робить дане завдання актуальним до цього часу.

Метою дослідження є вдосконалення процесу просування контенту, шляхом автоматизації цього процесу.

Для досягнення мети необхідно виконати наступні завдання:

- виконати огляд існуючих аналогів просування контенту;
- сформулювати математичну постановку задачі;
- здійснити порівняльний аналіз методів, які використовуються при побудові системи;
- розв'язати задачу створення цільових аудиторій методом кластеризації;
- розв'язати задачу визначення найвпливовіших факторів успішності методом головних компонент та алгоритмом прогнозування;
- розв'язати задачу прогнозування очікуваної кількості приросту цільової аудиторії методом часових рядів;
- розв'язати задачу обрахунку KPI;
- розробити систему з просування контенту;

- провести експерименти для дослідження методів;
- виконати аналіз отриманих результатів.

Об'єкт дослідження – процес просування контенту у соціальній мережі.

Предмет дослідження – модель сегментації факторів та користувачів соціальної мережі.

Методи дослідження, застосовані у даній роботі, базуються на методах кластеризації.

Науковою новизною є дослідження впливу обраних алгоритмів на коефіцієнт успішності.

Публікації. Результати досліджень були опубліковані у тезах доповідей на науково-технічній міжнародній конференції «Інженерія програмного забезпечення і передові інформаційні технології (SoftTech-2023)».

СЕГМЕНТАЦІЯ, КЛАСТЕРИЗАЦІЯ, МЕТОД ГОЛОВНИХ КОМПОНЕНТ,  
ПРОГНОЗУВАННЯ, ЧАСОВІ РЯДИ, КОНТЕНТ

## ABSTRACT

Master's thesis: 105 pp., 45 figures, 9 tables, 37 sources, 2 appendix.

Relevance. Over the past few years, the popularity and importance of the task of promoting content on social networks has grown significantly. However, given the proliferation of process automation, the question arises whether it is possible to create a quality advertising campaign using a strategy based on the results of predictive models and clustering models. In order to avoid excessive intervention of a marketing specialist in the content creation process, we answer this question as follows. However, the successful formation of a forecasting strategy requires a high-quality human assessment, a sufficient amount of input data and a clearly formulated goal at the stage of starting work.

Until now, we have not met a system with such functionality, which makes this task relevant until now.

The purpose of the study is to automate the content promotion process using input data.

To achieve the goal, it is necessary to complete the following tasks:

- perform a review of existing methods of creating a content promotion plan;
- to formulate a mathematical formulation of the problem;
- carry out a comparative analysis of the methods used in the construction of the system;
- solve the problem of creating target audiences using the clustering method;
- solve the problem of determining the most influential factors of success using the method of principal components and the forecasting algorithm;
- to solve the problem of forecasting the expected amount of growth of the target audience using the time series method;
- solve the problem of KPI calculation;
- develop a content promotion system;
- conduct experiments to research methods.

The object of research is the process of promoting content in a social network.

The subject of the study is a model of segmentation of factors and users of the social network.

Research methods used in this work are based on clustering methods.

A scientific novelty is the study of the influence of the selected algorithms on the success rate.

Publications. The research results were published in abstracts of reports at the scientific and technical international conference "Software Engineering and Advanced Information Technologies (SoftTech-2023)".

SEGMENTATION, CLUSTERIZATION, METHOD OF PRINCIPAL COMPONENTS, PREDICTION, CONTENT

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ .....	10
ВСТУП .....	11
1 ОГЛЯД СУЧАСНИХ ДОСЛІДЖЕНЬ В ОБЛАСТІ ТЕРГЕТОВАНОЇ РЕКЛАМИ У СОЦІАЛЬНИХ МЕРЕЖАХ .....	13
1.1 Опис предметного середовища .....	13
1.2 Огляд існуючих аналогів в області автоматизації таргетованої реклами в соціальних мережах .....	14
1.3 Огляд існуючих публікацій в області автоматизації таргетованої реклами в соціальних мережах .....	16
1.4 Опис функціональної моделі .....	17
Висновок до розділу .....	18
2 МОДЕЛЬ ТА МЕТОДИ ДЛЯ СТВОРЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ПІДТРИМКИ РОБОТИ ФАХІВЦЯ З МАРКЕТИНГУ У СОЦІАЛЬНИХ МЕРЕЖАХ .....	20
2.1 Задача створення цільових аудиторій .....	20
2.2 Задача визначення найвпливовіших факторів успішності .....	28
2.3 Задача прогнозування очікуваного приросту цільової аудиторії .....	32
2.4 Задача обрахунку КРІ .....	37
Висновок до розділу .....	38
3 РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ .....	41
3.1 Засоби розробки .....	41
3.2 Інформаційне забезпечення .....	49
3.4 Бібліотеки .....	51
3.5 Вимоги до технічного забезпечення .....	53
3.6 Архітектура програмного забезпечення .....	54
3.7 Опис функціональної моделі .....	54

	9
3.8 Специфікація функцій .....	56
3.9 Діаграма класів .....	58
3.10 Діаграма послідовностей .....	60
Висновок до розділу .....	62
4 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ .....	64
4.1 Планування експериментів .....	64
4.2 Параметризація .....	64
4.3 План експериментів .....	67
4.4 Проведення експериментів .....	71
Висновок до розділу .....	86
5 РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ .....	88
5.1 Опис ідеї проєкту .....	88
5.2 Технологічний аудит ідеї проєкту .....	92
5.3 Аналіз ринкових можливостей запуску стартап-проєкту .....	94
Висновок до розділу .....	101
ВИСНОВКИ .....	102
ПЕРЕЛІК ПОСИЛАНЬ .....	104

## ПЕРЕЛІК СКОРОЧЕНЬ

SMM (Social media marketing) – Соціальний медіа маркетинг.

eWOM (Electronic word of mouth) – це обмін інформацією про продукт чи послугу у вигляді рекомендацій у соціальних мережах.

ROI (Return on Investment) – це популярний показник прибутковості, який використовується для оцінки ефективності інвестиції.

МГК – Метод головних компонент.

KPI (Key Performance Indicators) – Ключові показники ефективності.

ER-діаграма (Entity Relationship) – Діаграма сутності та зв'язків.

## ВСТУП

Тема магістерської дисертації полягає у розробці сервісу для підтримки роботи фахівця з маркетингу в соціальних мережах. Цей сервіс призначений для мінімізації часу, необхідного для просування контенту, що значно спрощує процес розробки рекламної кампанії. Слід зазначити, що подібна система може стати дуже популярною серед різних соціальних груп населення та мати вагомую позицію в сфері соціального медіа маркетингу.

Онлайн бізнес сьогодні став справжньою промисловою революцією, що дозволяє підприємствам отримувати високі прибутки. Соціальні мережі надають різноманітні формати реклами, що можуть задовольнити будь-які потреби бізнесу. Сила соціального медіа маркетингу полягає в потужності соціальних мереж у трьох основних сферах маркетингу: зв'язок, взаємодія та збір даних про клієнтів. Таким чином, соціальний медіа маркетинг став дуже важливим напрямком рекламної діяльності, а фахівці у цій сфері, такі як таргетологи, стали вкрай популярними на ринку праці.

Як відомо, для проведення досліджень в маркетинговій сфері потрібно мати обширну експертну оцінку, значну кількість даних для ефективного просування контенту та затрати значних людських ресурсів. Тому, щоб зменшити потребу в ресурсах для розробки програмного продукту, було прийнято рішення розробити алгоритм кластеризації, прогнозування та аналізу часових рядів. Головна ідея полягає в використанні функції витрат і її мінімізації.

В даний час на Facebook є інструмент, що дозволяє автоматизувати запуск рекламних кампаній. Автоматизована реклама на Facebook дає можливість публікувати рекламу, користуючись перевіреним контентом. Facebook генерує індивідуальний план користувача з крок за кроком рекомендаціями для публікації рекламних оголошень. Цей підхід є ідеальним для новачків у сфері бізнесу або для підприємців з невеликим товарообігом. Однак для масштабування бізнесу на більш велику межу все ж потрібен фахівець з маркетингу, такий як таргетолог.

Оскільки соціальні медіа є дуже поширеними в сучасному світі, то методи, алгоритми та стратегії, які використовуються на цих платформах, є дуже важливими для бізнесу. У нашій системі ми детально опрацювали методи та алгоритми, щоб мінімізувати або повністю прибрати потребу в втручанні фахівця. За допомогою нашої системи можна просувати контент на основі моделі сегментації критеріїв та методів її розв'язання. Цей підхід дуже корисний для початківців у бізнесі або для підприємців з невеликим товарним оборотом, хоча для більш серйозного масштабування може бути необхідне втручання фахівця з маркетингу - таргетолога.

Практичні результати нашої роботи були застосовані для створення просувального застосунку, який працює на основі моделі сегментації критеріїв та методів її розв'язання.

За матеріалами досліджень була написана стаття на тему «Сегментування споживачів для якісного маркетингу у соціальних мережах», в якій розповідаємо про нашу систему та її переваги.

## 1 ОГЛЯД СУЧАСНИХ ДОСЛІДЖЕНЬ В ОБЛАСТІ ТЕРГЕТОВАНОЇ РЕКЛАМИ У СОЦІАЛЬНИХ МЕРЕЖАХ

У сучасний час наукові дослідження відводять значну увагу соціальному медіа маркетингу. Це область, де реклама і гроші, або реклама і час, тісно пов'язані, та яка дозволяє вирішувати такі завдання:

- вибір цільової аудиторії;
- просування контенту;
- запуск рекламної кампанії;
- розрахунок КРІ;

### 1.1 Опис предметного середовища

Маркетинг у соціальних мережах (англ. Social media marketing), також відомий як соціальний медіа-маркетинг, це форма інтернет-маркетингу, яка використовує можливості соціальних мереж як інструмент для просування різних контентів. Це дозволяє збільшити трафік на сторінки та вивчити реакцію користувачів на контент. [1]

SMM базується на принципі органічного залучення зацікавлених осіб з цільової аудиторії, що означає, що при активній взаємодії з контентом або сторінкою в соціальній мережі, позиція цього контенту в пошукових системах або рекомендаціях соціальних мереж, наприклад Instagram, зростає. [1]

Таргетована реклама в соціальних мережах створює динамічний характер взаємодії, де пряме або пасивне спілкування, таке як лайки, дозволяє компаніям скористатися безкоштовними рекламними можливостями, що базуються на рекомендаціях eWOM (електронне сарафанне радіо) між існуючими та потенційними клієнтами. Ці взаємодії можуть бути виміряні в соціальних мережах, що робить їх цінними для компаній, оскільки вони можуть вимірювати свою "соціальну справедливість" - термін, що відноситься

до рентабельності інвестицій (ROI) їх маркетингових кампаній у соціальних мережах (SMM).

Це стосується і пошукових систем, таких як Google. З урахуванням того факту, що 90% користувачів, які здійснюють пошук в пошукових системах, не переходять за межі першої сторінки результатів, а 70% з них клацають лише на перші три посилання, очевидно, що досягнення першої позиції на першій сторінці пошуку є головною метою при просуванні контенту.

Маркетинг в соціальних мережах є універсальним інструментом, що дозволяє досягати різноманітних маркетингових цілей, в залежності від загальної стратегії. Ось лише кілька прикладів таких цілей, що дуже популярні в SMM:

- залучення цільової аудиторії;
- збільшення продажів та генерація потенційних клієнтів;
- забезпечення лояльності клієнтів;
- збільшення трафіку на сайті або в особистому блозі;
- збільшення кількості підписників;
- отримання зворотного зв'язку від аудиторії;
- моніторинг конкурентів і трендів;
- підвищення впізнаваності бренду;
- дослідження споживчого попиту. [2]

## 1.2 Огляд існуючих аналогів в області автоматизації таргетованої реклами в соціальних мережах

Автоматизація маркетингу використовує програмне забезпечення для автоматизації повторюваних маркетингових завдань. Це може включати маркетинг шляхом розсилки електронною поштою, публікації в соціальних мережах, рекламні кампанії та інше. Маркетингові відділи використовують технологію автоматизації маркетингу, щоб збільшити ефективність своєї роботи та забезпечити більш персоналізований досвід для клієнтів. Ця

технологія відтворює повторювані завдання швидше та легше, що дає змогу маркетологам сконцентруватись на більш складних та креативних завданнях.

На сьогоднішній день існує багато розроблених рішень та застосунків, які являються помічниками таргетологів. Одним з найпопулярніших є Audiense — це програмне забезпечення для соціального маркетингу, створене для того, щоб допомогти користувачам максимально використати свої облікові записи в соціальних мережах, як-от Twitter, у маркетингових цілях. Це дозволяє їм ефективніше взаємодіяти зі своєю аудиторією, надаючи їм інструмент для аналізу даних, зібраних із цих сайтів соціальних мереж. [3]

На сьогоднішній день існує значна кількість розроблених рішень та програм, які допомагають таргетологам. Одним з найпопулярніших таких інструментів є Audiense — програмне забезпечення для соціального маркетингу, створене для того, щоб максимально використовувати облікові записи в соціальних мережах, таких як Twitter, у маркетингових цілях. Це дозволяє користувачам більш ефективно взаємодіяти зі своєю аудиторією, надаючи їм інструмент для аналізу даних, зібраних з цих соціальних мереж. Audiense - це програмне забезпечення, що надає ключові рішення для соціальної реклами, маркетингу та соціальної інформації. Його основна мета - допомогти компаніям зв'язатися зі своєю аудиторією та зрозуміти її поведінку в Інтернеті. Audiense надає необхідні інструменти для залучення більшої аудиторії та спонукання їх до дій. Маркетологи можуть використовувати це програмне забезпечення для швидкого визначення унікальних підходів до маркетингу для споживачів на основі багатих і різноманітних соціальних даних, що дозволяє забезпечувати конкурентну перевагу. [3]

Також, існує досить популярна платформа StackAdapt — це рекламна платформа, яка допомагає агентствам і брендам охоплювати та шукати своїх ідеальних клієнтів за допомогою привабливих блоків нативної та відеореклами. Є існуючий аналог цієї платформи — WOORISE. WOORISE — це інструмент інтерактивного вмісту, який дозволяє легко створювати потужні рекламні кампанії, такі як вірусні розіграші та конкурси, цільові сторінки,

цікаві форми, опитування, вікторини та багато іншого. Зрозуміло, що ці сервіси дуже важливі і є доречними при створенні рекламної кампанії.

### 1.3 Огляд існуючих публікацій в області автоматизації таргетованої реклами в соціальних мережах

Публікації у соціальних медіа маркетингу є надзвичайно важливим інструментом для будь-якої компанії, яка прагне залучити нових клієнтів та зберегти існуючих. Ці публікації допомагають підтримувати активну взаємодію з аудиторією та створювати позитивний імідж бренду. Вони також дають можливість компанії отримувати цінні дані про споживачів, їхні потреби та поведінку в Інтернеті. Необхідно розуміти, що ефективний соціальний медіа маркетинг потребує не тільки регулярних публікацій, але й аналізу результатів та відповідну стратегію для досягнення поставлених цілей.

У своєму дослідженні про ефективність моделі Кано, Девід Петровські та Жоао Педро Пестана Нето визначили модель Кано як основу для встановлення найважливіших факторів і досліджили важливість вибору цільової аудиторії для досягнення успіху в цій моделі. Ця дослідницька робота має на меті заповнити прогалину в знаннях та визначити, наскільки важливим є відбір цільової аудиторії для досягнення ефективності моделі Кано. [4]

У своєму дослідженні Христина Ліпяніна-Гончаренко, Тарас Лендюк, Анатолій Саченко та Яцек Волошин досліджували методи формування контекстної реклами та цільової аудиторії на основі навчання асоціативними правилами. Застосування запропонованої методики дозволило збільшити ефективність реклами в соціальних мережах на мінімум 23% та знизити ціну на 90%. Крім того, автори рекомендують збільшити кількість споживацьких опитувань для поліпшення вивчення асоціативних правил та отримання кращих ключових слів для формування контенту. Це призведе до підвищення ефективності реклами та зменшення її витрат. [5]

Дослідження, проведене Янхоу Луо, Хувей Пан, Шихион Ванг та Юніг Хуанг, зосередилося на визначенні значення цільової аудиторії для підприємства в соціальній мережі. Дослідники заявили про великий потенціал адаптивного та динамічного визначення цільової аудиторії, особливо в контексті швидкого розвитку інформаційних онлайн-служб та соціальних медіа. Однак, науковці також вказали на обмеження цього підходу, оскільки для визначення відповідної тематичної моделі може знадобитися допомога ручної оцінки. [6]

Мухаммед Усман провів дослідження ефективності реклами для різних культур та цільових аудиторій з метою переконання споживачів. Він зосередився на рекламних кампаніях, спрямованих на цільову аудиторію з різних культур. У своїх дослідженнях автор виявив, що ефективність реклами залежить від того, наскільки вона впливає на споживача. Результати дослідження дали автору імпульс для подальшої роботи в цій області, адже вони підтвердили важливість створення рекламних кампаній, що здатні впливати на цільову аудиторію. [7]

#### 1.4 Опис функціональної моделі

Дійовою особою (актором) застосунку для створення плану просування є Користувач. В межах застосунку Користувач може виконувати такі основні функції:

- створювання нової заявки для просування контенту;
- перегляд історії заявок цього користувача;
- перегляд показників КРІ для кожної із заявок цього користувача;
- перегляд статистики по всіх заявках: середній КРІ, статистика за мережами Facebook та Instagram, кількість заявок за останній місяць.

Функціональна діаграма варіантів використання користувачем застосунку для прогнозування курсу валют представлена на рисунку 1.1.

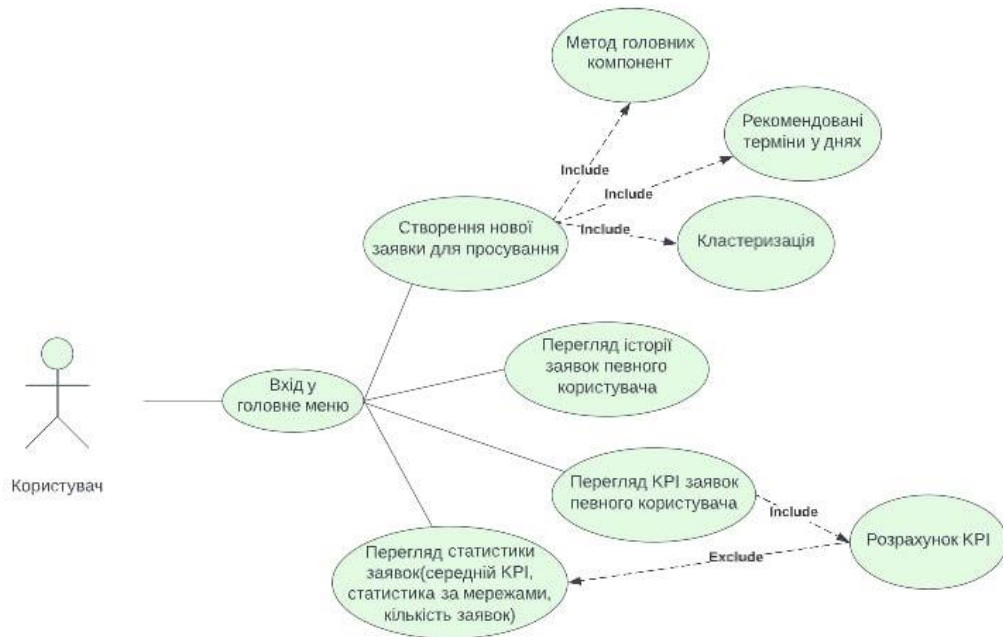


Рисунок 1.1 — Функціональна діаграма варіантів використання

### Висновок до розділу

Проведено аналіз наявного предметного середовища, включаючи аналоги та публікації з даної тематики, і після цього було зроблено висновок про необхідність розробки власної системи.

Наш застосунок відрізняється від інших аналогів тим, що забезпечує повну автоматизацію процесу. Іншими словами, для запуску рекламної кампанії необхідно лише ввести вхідні дані і розпочати процес. Це значно простіше, ніж купувати підписку на кілька застосунків та вручну поєднувати їх результати. Також, в ході аналізу було виявлено, що багато існуючих аналогів не мають повної функціональності, що можуть запропонувати наш застосунок. Зокрема, він може автоматично аналізувати результати рекламної кампанії — прорахунок KPI, що допоможе їм покращити ефективність своєї реклами та збільшити продажі.

Створений застосунок також має інтуїтивний і простий інтерфейс, що дозволяє користувачам швидко та легко використовувати його без додаткової

підготовки та навчання. Крім того, ми пропонуємо різноманітні інструменти та функції, що дозволяють налаштувати рекламну кампанію на потрібну аудиторію та підібрати оптимальні параметри для її запуску.

Отже, наш застосунок є інноваційним та ефективним рішенням, яке допоможе забезпечити повну автоматизацію процесу рекламної кампанії, що зекономить час та зусилля користувачів.

У рамках аналізу була створена функціональна діаграма варіантів використання та її опис, що детально описує функціональні можливості застосунку та дозволяє зрозуміти, як саме він працює.

## 2 МОДЕЛЬ ТА МЕТОДИ ДЛЯ СТВОРЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ПІДТРИМКИ РОБОТИ ФАХІВЦЯ З МАРКЕТИНГУ У СОЦІАЛЬНИХ МЕРЕЖАХ

Модель системи підтримки фахівця з маркетингу у соціальних мережах зображена на рисунку 2.1.



Рисунок 2.1 – Архітектура моделі

Архітектура моделі складається з наступних задач:

- задача створення цільових аудиторій;
- задача визначення найвпливовіших факторів успішності;
- задача прогнозування очікуваної кількості приросту цільової аудиторії;
- задача обрахунку КРІ.

Розглянемо більш детально кожен з цих задач та методи її розв’язання.

### 2.1 Задача створення цільових аудиторій

На даному етапі, сегментування аудиторії допоможе в розподілі користувачів на групи з спільними характеристиками, що дозволить легше визначити їх потреби та очікування. Це допоможе компанії більш ефективно налаштувати свої маркетингові та продажні стратегії для кожного сегменту та збільшити шанси на залучення та утримання клієнтів.

Сегменти споживачів використовуються для задачі прогнозування очікуваної кількості приросту цільової аудиторії як вхідна множина даних.

### 2.1.1 Математична постановка

Вхідні дані: люди, їх вік, їх інтереси.

Вихідні дані: цільові аудиторії — сегменти розподілені за віком та інтересами людей.

Алгоритм розв’язання задачі створення цільових аудиторій зображений на рисунку 2.2.



Рисунок 2.2 – Алгоритм розв’язання задачі створення цільових аудиторій

Для формування цільових аудиторій ми використовуємо метод кластеризації, що дозволяє об’єднати схожі об’єкти в групи. На вхід подається множина об’єктів, а на виході ми отримуємо групи об’єктів, які мають подібні характеристики. Для цього ми використовуємо метод кластеризації  $k$ -середніх.

Мета методу – розділити  $n$  спостережень на  $k$  кластерів, так щоб кожне спостереження належало до кластера з найближчим до нього середнім значенням. Метод базується на мінімізації суми квадратів відстаней між кожним спостереженням та центром його кластера, тобто на функції 2.1.

$$\sum_{i=1}^N d(x_i, m_j(x_i))^2, \quad (2.1)$$

де  $d$  – метрика;

$x_i$  –  $i$ -тий об’єкт даних;

$m_j(x_i)$  – центр кластера, на якому  $j$ -ій ітерації приписаний елемент  $x_i$ .

Головні переваги методу  $k$ -середніх — його простота та швидкість виконання. Метод  $k$ -середніх більш зручний для кластеризації великої кількості спостережень, ніж метод ієрархічного кластерного аналізу (у якому дендограми стають перевантаженими і втрачають наочність). [16]

Особливістю даного методу є те, що на вхід має подаватись кількість кластерів, які ми визначаємо наступними методами її розв'язання: методом градієнтного спуску та методом ліктя. Потім при порівнянні даних методів, ми маємо обрати найкращий та за ним буде створена цільова аудиторія для майбутнього просування контенту.

Основною задачею методу кластеризації у нашому випадку є сегментація цільової аудиторії на підмножини. Тобто, це категорії (кулінарія, медицина, IT, і так далі). Оскільки у людей може бути кілька інтересів, то сегментуючи дані, опір буде на найбільш впливовий інтерес для конкретної людини. [9]

### 2.1.2 Кластеризація методом $k$ -середніх з визначенням кількості кластерів методом градієнтного спуску

Градiєнтний спуск – ітераційний алгоритм оптимізації, який дозволяє мінімізувати функціонал помилки шляхом оновлення параметрів шуканої функції у зворотному напрямку градієнта. [17] Коефіцієнт навчання  $\alpha$  визначає розмір кроку, який ми використовуємо, аби досягти мінімуму функціоналу помилки.

Класичний алгоритм градієнтного спуску обчислює градієнт функціоналу помилки для параметрів усього набору точок матриці  $X$ . Оскільки необхідно обчислювати градієнти для всього набору даних, щоб виконати оновлення параметрів шуканої функції, класичний градієнтний спуск може бути дуже повільним. Однак, він гарантує знаходження глобального мінімуму для опуклих поверхонь і локального мінімуму для неопуклих поверхонь функціоналу помилки.

Для пошуку мінімуму множини кластерів функції  $k - \text{means } J(C)$  використовується метод градієнтного спуску.

Ідея градієнтного спуску заключається в тому, щоб на кожному кроці оновлювати параметри шуканої функції у зворотньому напрямку градієнта функціоналу помилки, тим самим здійснюючи оптимізацію у напрямку найшвидшого спуску.

За умови, якщо правильно підібрані гіперпараметри (максимальна кількість ітерацій, коефіцієнт навчання і так далі), тоді алгоритм градієнтного спуску приводить до допустимого розв'язку. Оскільки градієнт дозволяє знайти максимальний приріст функції, антиградієнт забезпечує пошук мінімуму функціоналу помилки. Також, цей алгоритм не гарантує досягнення глобального оптимуму, адже градієнт дорівнює нулю у будь-якому екстремумі, навіть якщо це локальний мінімум. [8]

#### 2.1.2.1. Математична постановка задачі

Для вибору числа кластерів використовується значення функції  $k - \text{means } J(C)$ . Обирають те число кластерів, починаючи з якого значення  $J(C)$  спадає не так швидко.

Введемо наступні позначення:

- $C = \{\vec{C}_1, \vec{C}_2, \dots, \vec{C}_k\}$  – множина центроїдів  $\vec{C}_i = \{C_{i,1}, C_{i,2}, \dots, C_{i,n}\}$ ;
- $X = \{\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n\}$  – множина даних, яку необхідно кластеризувати;
- $X_k \subset X$  –  $k$ -тий кластер, підмножина вихідної множини.

Розглянемо функцію класичного алгоритму  $k - \text{means}$ . Вона наведена у формулі 2.2.

$$J(C) = \sum_{k=1}^{|C|} \sum_{i=1}^{|X_k|} \sum_{j=1}^n (x_{i,j} - C_{k,j})^2 \quad (2.2)$$

Далі потрібно мінімізувати суму квадратів Евклідової відстані елементів кластера від відповідного центроїда. Представлення мінімізації наведено у формулі 2.3.

$$J(C) = \sum_{k=1}^{|C|} \sum_{i=1}^{|X_k|} f^2(\vec{X}_i, \vec{C}_k), \quad (2.3)$$

де  $f$  – евклідова відстань.

Знайдемо похідну нульової функції. Функція наведена у формулі 2.4.

$$\frac{\partial J(C)}{\partial C_{a,b}} = \sum_{i=1}^{|X|} \frac{\partial \min_{j=1}^{|C|} f^2\{\vec{x}_i, \vec{c}_j\}}{\partial C_{a,b}} \quad (2.4)$$

Фінальним етапом залишається визначити чому дорівнює похідна. Тобто, якщо  $a$ -тий центроїд це дійсно той, до якого відстань мінімальна від поточного елемента множини, то похідна у такому випадку дорівнює похідній квадрату відстані по елементу вектора, або в іншому випадку нуль. Це наведено у формулі 2.5. [8]

$$\frac{\partial \min_{j=1}^{|C|} f^2\{\vec{x}_i, \vec{c}_j\}}{\partial C_{a,b}} = \left\{ \begin{array}{l} 0, \min_{j=1}^{|C|} f^2\{\vec{x}_i, \vec{c}_j\} \neq f^2(\vec{x}_i, \vec{c}_a) \\ \frac{\partial f^2\{\vec{x}_i, \vec{c}_a\}}{\partial C_{a,b}}, \text{ in } \min_{j=1}^{|C|} f^2(\vec{x}_i, \vec{c}_j) = f^2(\vec{x}_i, \vec{c}_a) \end{array} \right\}, \quad (2.5)$$

де  $f$  – у цьому може бути будь-яким виміром метрики.

### 2.1.3 Кластеризація методом $k$ -середніх з визначенням кількості кластерів методом ліктя

При кластеризації методом  $k$  –середніх кількість кластерів найчастіше оцінюють за допомогою «методу ліктя». Він передбачає багаторазове циклічне виконання алгоритму зі збільшенням кількості кластерів, що вибираються, а також подальшим відкладанням на графіку балу кластеризації, обчисленого як функція від кількості кластерів. Основним показником методу ліктя є  $J(C)$  – сума квадратів помилок. [18]

Основна ідея методу ліктя полягає в тому, що в міру збільшення числа кластерів  $k$  розподіл вибірок буде більш точним, а ступінь агрегації кожного кластера поступово збільшуватиметься, тому квадрат помилок, природно, поступово стануть меншими. Більше того, коли  $k$  менше числа істинних кластерів, збільшення  $k$  значно збільшить ступінь агрегації кожного кластера,

тому зниження  $J(C)$  буде більшим, і коли  $k$  досягне числа істинних кластерів, значення, отримане шляхом збільшення  $k$ . Ступінь агрегації швидко стає меншим, тому зниження  $J(C)$  буде різко зменшуватися, а потім поступово згладжуватиметься, оскільки значення  $k$  продовжує збільшуватися, що означає, що відносини між  $J(C)$  і  $k$  мають форму коліна. Відповідне значення  $k$  є кількістю справжніх кластерів у даних. Звичайно, саме тому метод називається методом ліктя. [10] [18]

### 2.1.3.1 Математична постановка задачі

Математична постановка задачі полягає у мінімізації множини кластерів функції витрат  $k$  – середніх та наведена у формулі 2.6.

$$J(C) = \sum_{k=1}^K \sum_{i \in C_k} \|x_i - \mu_k\|^2 \rightarrow \min_C, \quad (2.6)$$

де  $C$  – множина кластерів потужності  $K$ ;

$\mu_k$  – центроїд кластера  $C_k$ ;

$x_i$  – множина даних ітерації.

Для вирішення питання вибору кількості кластерів потрібно знайти мінімум, але оскільки такий мінімум буде при тому, що кожна точка – це кластер одного елементу, то потрібно скористатись правилом, побудувавши графік, знайти точку, де  $J(C)$  падає не так швидко. Дане представлення наведене формулою 2.7.

$$D(k) = \frac{|J(C_k) - J(C_{k+1})|}{|J(C_{k-1}) - J(C_k)|} \rightarrow \min_C \quad (2.7)$$

Візьмемо частину нашої цільової аудиторії та побудуємо графік, аби визначити оптимальну кількість кластерів для такої кількості людей. Це показано на рисунку 2.3.

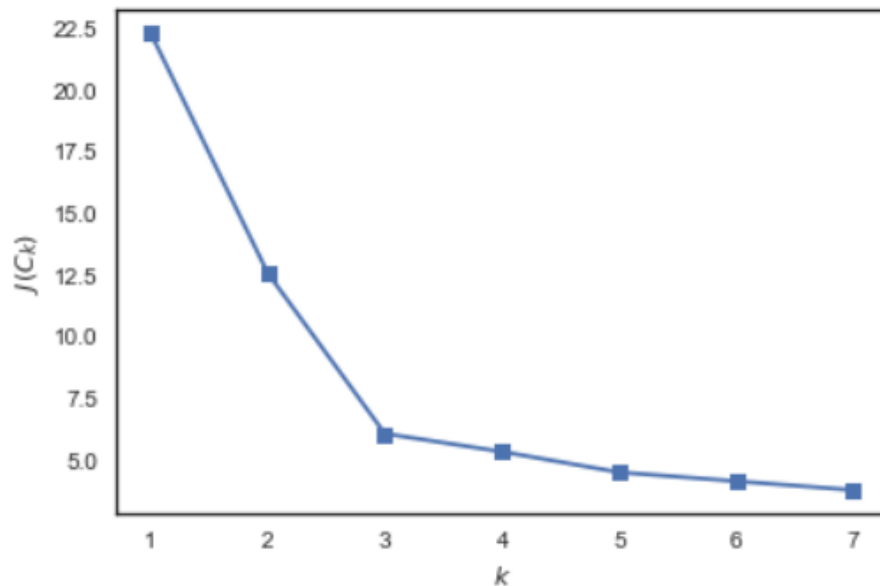


Рисунок 2.3 – Графік методу ліктя

На рисунку бачимо, що при зміні  $k$  від 3 до 4, значення  $J(C)$  падає не так швидко, тому для цього варіанту оптимальною кількістю кластерів буде 3 кластери.

#### 2.1.4 Порівняння методів визначення кількості кластерів та вибір кращого з них

Для початку побудуємо діаграму для нашого алгоритму порівняння методів кластеризації, вона зображена на рисунку 2.4.



Рисунок 2.4 – Діаграма алгоритму порівняння методів кластеризації

Розглянувши складності алгоритмів при кластеризації з використанням Євклідової відстані, дослідимо тільки одну ітерацію (аналіз кількості ітерацій є складною задачею).

Введемо наступні позначення:

- $k$  – кількість кластерів;
- $n$  – розмірність даних;
- $m$  – кількість даних ( $k \leq m$ ).

У крайньому випадку, коли  $k = m$ , складність алгоритму буде найвищою і складатиме  $O(m^2n)$ , що може стати проблемою для великих наборів даних і високих розмірностей.

Визначимо складність ітерації для EM-версії, вона позначена формулою 2.8.

$$O(2kmn + m) = O(m^2n + m) = O(m^2n) \quad (2.8)$$

Визначимо складність ітерації для GD-версії, вона позначена формулою 2.9.

$$O(2kmn + mk + mn) = O(m^2n + m^2 + nm) = O(m^2n) \quad (2.9)$$

На рисунку 2.5 відповідно показані графіки функції складності ітерації для методу ліктя та градієнтного спуску.

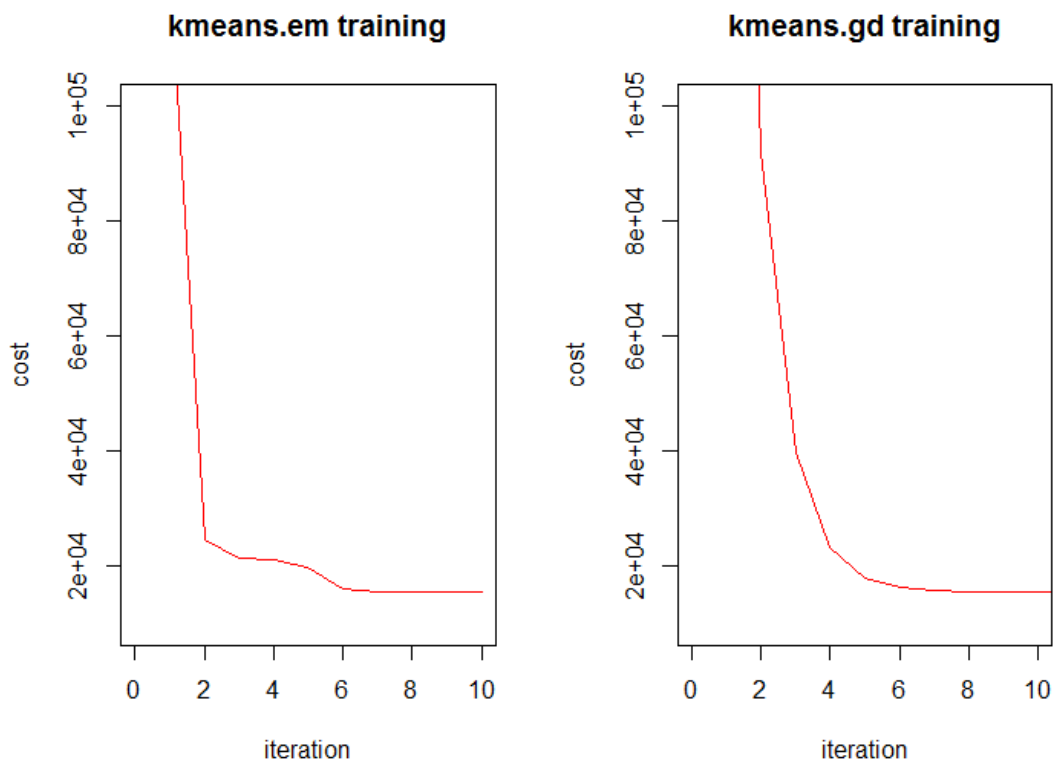


Рисунок 2.5 – Графіки складностей ітерації методів пошуку кількості кластерів

Додатково побудуємо лінійний графік з обрахунками складності, залежних від значення  $m$  – кількості даних. Він показаний на рисунку 2.6.

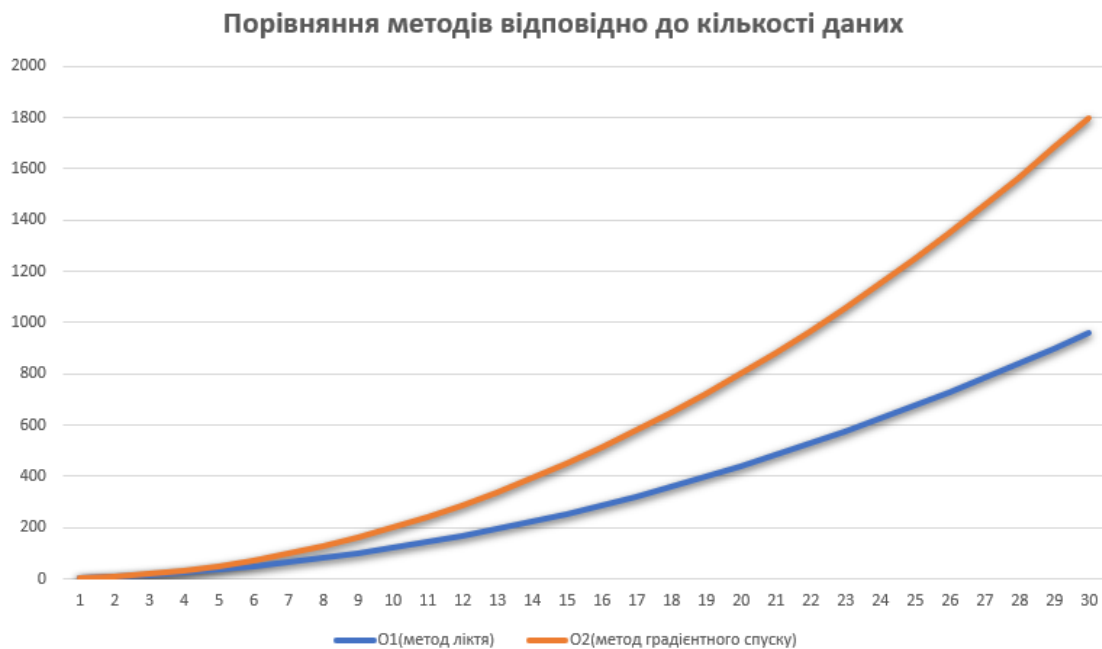


Рисунок 2.6 – Графік складностей обрахування методів

Отже, можна зробити висновок, що використовуючи метод ліктя, ітераційна складність з великою кількістю даних буде достатньо меншою, аніж використовуючи метод градієнтного спуску. На простих даних цієї переваги не можна побачити, але взявши великий набір даних буде видна різниця. Тому важливо використати метод ліктя для кластеризації споживачів.

## 2.2 Задача визначення найвпливовіших факторів успішності

Задля успішного просування контенту користувача, потрібно визначити найвпливовіші фактори, які можуть допомогти у просуванні. Таким чином, з множини даних потрібно вибрати лише ті, які дійсно якісно впливають на подальші результати.

З числа методів, що дозволяють узагальнювати значення елементарних ознак, метод головних компонент виділяється простою логічною конструкцією. На його прикладі стають зрозумілими загальна ідея і цільові

настанови численних методів факторного аналізу. Для виявлення найбільш значущих факторів і їх структури, найбільш виправданим є метод головних компонент. Сутність методу полягає в заміні корельованих компонентів некорельованими факторами. Іншою важливою характеристикою методу є можливість виокремлення найбільш інформативних головних компонент і виключення інших з аналізу, що спрощує інтерпретацію результатів. Переваги методу також у тому, що він є математично обґрунтованим методом факторного аналізу. [21]

Метод головних компонент дає можливість за  $m$ -вхідними ознаками виділити  $m$ -головних компонент, або узагальнених ознак. Простір головних компонентів є ортогональним. Математична модель головних компонент базується на логічному допущенні, що значення безлічі взаємозалежних ознак породжують деякий загальний результат. [21]

Метод головних компонент використовується для вивчення взаємозв'язків між досліджуваними показниками. За його допомогою можна виявляти приховані показники (фактори), які відповідають за наявність лінійних статистичних зв'язків (кореляцій) між ними. Крім того, визначення більш впливових за умов проведення досліджень факторів серед первинно обраних показників, а також виявлення статистичного зв'язку визначають обґрунтованість висновків щодо ефективності тих чи інших впливів на досліджувану систему. [21]

### 2.2.1 Математична постановка задачі

Вхідні дані: потрібна початкова множина даних, яка включає в себе статистику впливовості кожного фактору:  $F = \{f_1, \dots, f_i, \dots, f_k\}$ ,

де  $f$  – фактор;

$i = \overline{1, k}$ ;  $k$  – кількість факторів.

На рисунку 2.7 показано фрагмент початкової множини даних.

№	Фактор	Початкове значення для майбутнього обрахування
1	Популярність теми	1200000
2	Правильне оформлення інформації на сторінці	1150000
3	Зовнішній вигляд сторінки	959000
4	Рівномірне планування постів	910000
5	Поточний стан сторінки	201000
6	Слідування за трендами	941000
7	Стильний дизайн	157000
8	Власний бренд	189000
9	Привертання уваги	1941000
10	Хештеги	121000

Рисунок 2.7 – Фрагмент початкової множини даних

Вихідні дані: найвпливовіші фактори з множини вхідних даних.

$$F^{imp} = \{f_1^{imp}, \dots, f_n^{imp}\}, \quad (2.18)$$

де  $n$  – кількість важливих факторів.

### 2.2.1.1 Метод розв'язання задачі

Нехай на вхід подається множина  $F = \{f_1, \dots, f_i, \dots, f_k\}$ ,

де  $i = \overline{1, k}$ ;  $k$  – кількість факторів.

Для кожного фактору  $i \in$  множина  $n$  значень за досліджувані періоди  $f_i = \{f_{i1}, \dots, f_{im}, \dots, f_{in}\}$ ,  $m = \overline{1, n}$ .

На початку необхідно знайти середні значення по кожному фактору. Це показано у формулі 2.10.

$$\bar{X}_i = \frac{1}{n} \sum_{m=1}^n f_{im}, \quad (2.10)$$

де  $i$  – індекс фактору;

$$i = \overline{1, k};$$

$$m – \text{індекс значення впливу } m = \overline{1, n}.$$

Потім для кожного з факторів  $i = \overline{1, k}$  необхідно порахувати відхилення від середніх значень (2.11) та суми квадратів відхилень (2.12).

$$d_{im} = f_{im} - \bar{X}_i, \quad (2.11)$$

$$S_i = \sum_{m=1}^n d_{im}^2, \quad (2.12)$$

де  $i = \overline{1, k}$ ;  $m = \overline{1, n}$ .

Далі рахується середнє значення відхилення кожного фактору. Це показано у формулі 2.13.

$$d_i = \frac{1}{n} \sum_{m=1}^n d_{im}, \quad (2.13)$$

де  $i$  – індекс фактору;  $i = \overline{1, k}$ .

Тепер можна обрахувати коефіцієнти кореляції Пірсона. Це показано у формулі 2.14.

$$r_{im} = \frac{\sum_{m=1}^n (d_i \times d_{im})}{\sqrt{\sum_{m=1}^n d_i^2 \times S_i}}, \quad (2.14)$$

де  $i, m = \overline{1, k}$ ;  $k$  – кількість кластерів.

Фінальним етапом є пошук взаємозалежних факторів. Якщо  $r_{im} < 0.7$ , то фактори  $f_i$  та  $f_m$  будемо вважати незалежними, а якщо  $r_{im} < 0.7$ , то залежними.

На наступному кроці переходимо до пошуку коефіцієнтів важливості методом головних компонент. На початковому етапі необхідно розв'язати рівняння. Це показано у формулі 2.15.

$$|R - \lambda E| = 0, \quad (2.15)$$

де  $R$  – матриця коефіцієнтів кореляції (формула 2.14);

$\lambda$  – діагональна матриця власних чисел розмірністю  $5 \times 5$ ;

$E$  – одинична матриця,  $i = \overline{1, k}$ ;

$k$  – кількість факторів.

Підставивши  $\lambda$  у систему рівнянь  $|R - \lambda E| U = 0i$ , знайдемо матрицю  $U$ , що є матрицею власних векторів і побудуємо пронормовану матрицю. Це показано у формулі 2.16.

$$V_j = \frac{U}{|U_j|} \quad (2.16)$$

З нормованої матриці знайдемо матрицю факторного відображення. Це показано у формулі 2.17.

$$A = V \lambda^{\frac{1}{2}} \quad (2.17)$$

За матрицею  $A$  оберемо ті фактори, у яких значення  $|a_{im}| > 0.1$ . Серед множини отриманих критеріїв знайдемо пари, які взаємозалежні, якщо такі

є, то будемо вважати значимим один фактор із такої пари. У результаті отримаємо множину важливих факторів. Це показано у формулі 2.18.

$$F^{imp} = \{f_1^{imp}, \dots, f_n^{imp}\}, \quad (2.18)$$

де  $n$  – кількість важливих факторів.

На рисунку 2.8 зображено схему дій задачі визначення найвпливовіших факторів успішності.

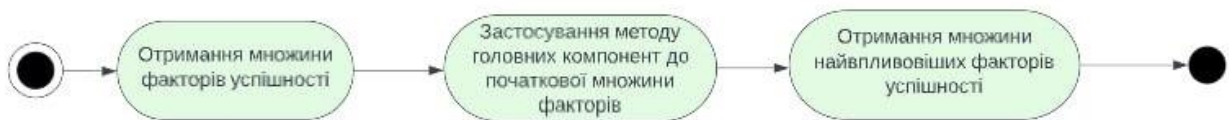


Рисунок 2.8 – Схема дій задачі визначення найвпливовіших факторів успішності

### 2.3 Задача прогнозування очікуваного приросту цільової аудиторії

Задача прогнозування очікуваного приросту цільової аудиторії являє в собі алгоритм, що обирає найвпливовіші фактори за допомогою методу головних компонент, обирає часовий проміжок із часового ряду, а також обирає кластер для майбутнього просування. Після чого прогнозує очікувану кількість підписників, вподобань контенту, а також переглядів. Задля роботи цього алгоритму, спочатку потрібно скористатись результатом задачі кластеризації і задачі знаходження найвпливовіших факторів успішності.

Оскільки метод часових рядів ARIMA упирається на результати минулих вимірів, а інші – ні, то у нас залишається тільки цей вибір. За допомогою часових рядів ми зможемо прогнозувати з опором на минулі дані і отримувати за допомогою інтерполяції похибок більш точну кількість приросту, аніж використовуючи будь-який інший метод.

Схема алгоритму показана на рисунку 2.9.



Рисунок 2.9 – Архітектура методу прогнозування

### 2.3.1 Визначення рекомендованих термінів просування контенту методом ARIMA

Часовий ряд – упорядкована в часі послідовність числових показників, котрі відповідають якісній характеристиці певного явища, об’єкту. Аналіз часових рядів – містить у собі методи обробки та аналізу набору даних часових рядів задля знаходження важливих статистик і характерних рис даних. Це потрібно для розуміння минулих трендів та для передбачення числових показників ряду у майбутньому. Основна ціль проведення аналізу часового ряду – сконструювати прогноз його значень на бажаний період часу у майбутньому. [22]

Важливими завданнями вивчення часових рядів є:

- знаходження оптимальної статистичної моделі для певного часового ряду;
- опис та формалізація основних ознак особливостей ряду;
- на основі попередніх спостережень мати можливість прогнозувати значення майбутніх показників. [15]

Аналіз часових рядів, зазвичай, складається з проведення основних етапів нижче:

- попередній аналіз досліджуваного часового ряду та його графічне відображення;
- отримання розуміння щодо закономірних складових часового ряду, а саме: тренду, циклічних та сезонних компонент;

- фільтрація високочастотних і низькочастотних складових;
- після видалення частотних складових описаних вище, проаналізувати випадкову складову ряду;
- перевірка побудови адекватності моделі залишеної складової;
- створення загальної моделі ряду, який досліджується у нашій роботі ;
- аналіз моделі, котру отримали і прогнозування поведінки об'єкта для майбутніх проміжків часу – формування висновку про взаємодію між різними досліджуваними часовими рядами, котрі характеризують певний процес або систему. [15]

Планування тісно пов'язане з прогнозуванням. Прогнозування можна використовувати як інструмент ефективного прийняття рішень. Завдяки прогнозуванню можна дізнатися, що з великою ймовірністю відбудеться з досліджуваним явищем, або процесом, чи які дії слід виконати для досягнення специфічного стану, котрий по тій чи іншій причині нам необхідний. [13] Саме тому ми використаємо часові ряди у програмі.

Тому це дозволить отримати нам такі результати:

- прогнозування терміну просування контенту;
- якісний розрахунок КРІ;
- статистичні дані як для розрахунку КРІ, так і для покращення алгоритму просування.

На рисунку 2.10 зображена схема використання часових рядів у програмі.



Рис. 2.10 – Схема використання методу ARIMA у програмі

### 2.3.1.1 Математична постановка задачі визначення термінів просування

Для якісного прогнозування очікуваної кількості приросту, потрібно використати метод часових рядів – ARIMA. Це допоможе обрати правильні

терміни для просування контенту, а також для покращення майбутнього значення KPI.

Вхідні дані: статистика минулих вимірів результатів користувачів (підписники, вподобання контенту, перегляди).

Вихідні дані: рекомендовані терміни просування, поточний приріст.

Авторегресійна модель – це модель часових рядів, в котрому фактичне значення ряду в теперішній момент лінійно залежний від попередніх значень досліджуваного нами ряду. Процес авторегресії з параметром  $p$  описується у формулі 2.19.

$$Y_t = a + \sum_{i=1}^p b_i Y_{t-i} + \varepsilon_t, \quad (2.19)$$

де  $a$  – константа;

$b_i$  – коефіцієнти авторегресії;

$\varepsilon_t$  – білий шум.

Так як автокореляційна функція є парною і використавши рекурентне співвідношення для перших параметрів  $p$  автокореляції, можна отримати систему рівнянь Юла – Уокера за формулою 2.20.

$$1 \leq m \leq p, \sum_{j=1}^p \varepsilon_j l(|m - j|) = l(m), \quad (2.20)$$

де  $m$  – кількість лагів;

$p$  – модель порядку;

$l$  – обсяг вибірки;

$j$  – константа.

При використанні вибірових кореляцій замість невідомих автокореляцій, можна отримати оцінки невизначених коефіцієнтів авторегресії. Метод оцінки авторегресії є синонімічним простому методу найменших квадратів. У випадку, якщо для наших випадкових помилок спостерігається нормальний розподіл, то даний метод є аналогічним умовному методу максимальної правдоподібності.

Таким чином, за допомогою авторегресійної моделі в результаті ми отримаємо набір рекомендованих термінів для майбутнього прогнозування

очікуваної кількості приросту цільової аудиторії, а також поточний приріст, який відбувається до початку просування.

### 2.3.2 Математична постановка задачі прогнозування очікуваного приросту цільової аудиторії

Вхідні дані: рекомендовані терміни просування, обраховані в попередній задачі методом ARIMA.

Вихідні дані: похибка, через яку йде прогноз очікуваного приросту.

У нашому випадку ми скористаємось авторегресійною моделлю, оскільки нам потрібно схилитись на показники минулих даних. А саме ця модель виконує таку функцію.

Оскільки алгоритм прогнозування є методом інтерполяції похибок, то метод ARIMA допоможе отримувати більш точні результати. Таким чином, чим більше заявок вже було створено, тим більшу точність ми отримаємо у кінцевому результаті.

Задля обрахунку очікуваної кількості, потрібно використати статистику приросту користувача до просування за допомогою часових рядів, а також знайти середнє значення похибки для минулих прогнозів. Якщо ж минулих прогнозів не було, то використаємо стандартну константу – 10%. Це показано у формулі 2.21.

$$\varepsilon_c = \frac{\sum P_r}{K_p}, \quad (2.21)$$

де  $\varepsilon_c$  – середня похибка;

$P_r$  – похибка кожного прогнозу;

$K_p$  – кількість прогнозів.

Обрахуємо загальну похибку для прогнозу. Це показано у формулі 2.22.

$$\varepsilon = \frac{P_m - \varepsilon_c * P_m + P_c}{P_m} * 100\%, \quad (2.22)$$

де  $\varepsilon$  – загальна похибка;

$P_m$  – максимальна кількість приросту цільової аудиторії;

$\varepsilon_c$  – середньо-статистична похибка;

$P_c$  – приріст цільової аудиторії до моменту просування.

Отже, знаючи загальну похибку ми зможемо отримати кількість приросту цільової аудиторії. Це показано у формулі 2.23.

$$A_r = P_m - P_m * \varepsilon, \quad (2.23)$$

де  $A_r$  – очікувана кількість приросту цільової аудиторії.

## 2.4 Задача обрахунку КРІ

КРІ — показники діяльності підрозділу (підприємства), які допомагають організації в досягненні стратегічних і тактичних (операційних) цілей. [23]

Використання ключових показників ефективності дає підприємству можливість оцінити свій поточний стан і допомогти в оцінці успішності реалізації обраної стратегії. [23]

КРІ дозволяють виробляти контроль ділової активності співробітників, підрозділів і компанії в цілому. Для терміну «key performance indicators (КРІ)» часто використовується український переклад «ключові показники ефективності» (КПЕ), однак правильним перекладом вважається варіант «ключові показники результатів діяльності». [11] [12] [23]

### 2.4.1 Математична постановка

Вхідні дані:

- $A_{r1}$  – очікувана кількість підписників;
- $A_{r2}$  – очікувана кількість вподобань контенту;
- $A_{r3}$  – очікувана кількість переглядів;
- $A_{f1}$  – фінальна кількість підписників;
- $A_{f2}$  – фінальна кількість вподобань контенту;

–  $A_{f3}$  – фінальна кількість переглядів.

Вихідні дані: показник КРІ.

#### 2.4.1.1 Метод розв'язання

Показник КРІ обраховується за основними трьома показниками, що збільшуються упродовж просування контенту, а саме:

- кількість підписників;
- кількість вподобань контенту;
- кількість переглядів.

На основі цього і проходить розрахунок показника. Це показано у формулі 2.24.

$$KPI = \frac{A_{r1} + A_{r2} + A_{r3}}{A_{f1} + A_{f2} + A_{f3}} * 100\%, \quad (2.24)$$

де  $A_{r1}$  – очікувана кількість підписників;

$A_{r2}$  – очікувана кількість вподобань контенту;

$A_{r3}$  – очікувана кількість переглядів;

$A_{f1}$  – фінальна кількість підписників;

$A_{f2}$  – фінальна кількість вподобань контенту;

$A_{f3}$  – фінальна кількість переглядів.

Отже, враховуючи результати минулих задач, ми зможемо знайти показник КРІ для певного просування користувача.

#### Висновок до розділу

Під час проектування моделі було обрано ряд завдань, які мали на меті проблему майбутнього просування контенту користувача.

Перша задача полягала в кластеризації користувачів за їх віковими категоріями та інтересами. Це дозволило розподілити користувачів по

сегментах і розробити персоналізовану стратегію пропозиції для кожного сегмента.

Друга задача полягала в обранні найбільш продуктивних компонентів успішності, які були використані для подальшого використання в алгоритмі прогнозування. Це дозволило точніше прогнозувати результати просування і використовувати ці дані для вирішення наступних завдань.

Третя задача полягала в розробці алгоритму прогнозування, який враховував вікові категорії та інтереси користувачів, а також часові ряди. Цей алгоритм дозволяє прогнозувати очікувану кількість підписників, вподобань контенту та переглядів, що допомогло розробити ефективну стратегію просування.

Четверта задача полягала в розрахунку KPI, що дає змогу якісно оцінити ефективність просування того чи іншого користувача. Цей обрахунок проводився з урахуванням таких параметрів, як цільова аудиторія, вікова категорія, термін просування та вибір мережі. Це дозволило зробити об'єктивну оцінку результатів просування.

Таким чином, розробка моделей, що включають у себе кластеризацію користувачів, вибрання найвищих компонентів ефективності, розробку алгоритму прогнозування та розрахунку KPI, дозволяє забезпечити ефективність просування контенту користувача та досягнення мети рекламної кампанії. Ця модель дозволяє отримувати більш точні результати, які враховують інтереси та потреби конкретної аудиторії, що сприяє збільшенню конверсії та покращенню відносин між брендом та користувачем.

Отже, поступове вирішення цих задач дозволяє досягти якісного результату в просуванні контенту користувача. Використання таких методів та алгоритмів допомагає залучати нових користувачів та зберігати старих, забезпечуючи їм якісний та цікавий контент. Додатково, це дозволяє знизити ризики невдалого просування та збільшити шанси на успіх. Застосування аналізу даних та машинного навчання стає все популярнішим в галузі

маркетингу та просування контенту, тому вивчення та використання таких методів є ключовим для ефективного просування у сучасному світі.

### 3 РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

При розробці програмного продукту було використано середовище Visual Studio 2019 для програмування мовою C#. Це середовище надає можливість розробляти як консольні, так і графічні застосунки, включаючи ті, що використовують технологію Windows Forms.

Visual Studio містить редактор вихідного коду, який підтримує технологію IntelliSense і забезпечує можливість швидкого рефакторингу коду. У складі Visual Studio є вбудовані інструменти для створення графічного інтерфейсу програми, такі як редактор форм, веб-редактор, дизайнер класів та дизайнер схеми бази даних. Крім того, Visual Studio дозволяє створювати та використовувати сторонні доповнення (плагіни), що допомагають розширювати функціональність середовища на різних рівнях, включаючи підтримку систем контролю версій вихідного коду та розширення доступних наборів інструментів. [24]

Для забезпечення контролю над версіями проекту використовується система GitHub.

Розглянемо більш детально основні засоби розробки.

#### 3.1 Засоби розробки

##### 3.1.1 Мова розробки

Мова програмування C# є сучасною технологією, яка має безліч переваг порівняно з іншими мовами, такими як C/C++ або Java. Нижче наведено лише деякі з переваг C#:

- C# створювався паралельно з платформою Microsoft .NET Framework і повною мірою враховує всі його можливості; [25]
- C# є повністю об'єктно-орієнтованою мовою, де навіть типи, вбудовані в мову, представлені класами; [25]

- мова програмування C# є потужною об'єктно-орієнтованою мовою з можливостями наслідування та універсальності; [25]
- мова програмування C# успадкувала кращі риси мов C/C++, зберігаючи їхню потужність та функціональність; загальний синтаксис та знайомі оператори мови допомагають програмістам, які мають досвід роботи з C++, швидше переходити до використання C#; [25]
- мова C# зберегла основні риси C++, але в той же час стала простіше і надійніше; простота і надійність, головним чином, пов'язані з тим, що на C# хоча і допускаються, але не заохочуються такі небезпечні властивості C++ як покажчики, адресація, розіменування, адресна арифметика; [25]
- бібліотека платформи надає потужні інструменти для створення різноманітних застосунків на C#; доступність таких інструментів дозволяє легко будувати веб-сервіси та компоненти, а також зручно зберігати та отримувати дані з баз даних; [25]
- завдяки платформі Microsoft .NET Framework, що є надбудовою над операційною системою, програмісти, що працюють з C#, можуть скористатися тими ж перевагами роботи з віртуальною машиною, що і програмісти, що працюють з Java; виконання коду стає ще більш ефективним, оскільки виконавча середовище CLR є компілятором проміжної мови, тоді як віртуальна Java-машина є інтерпретатором байт-коду; [25]

Мова програмування C# була розроблена Microsoft і вона стала однією з найпопулярніших мов програмування, оскільки має великий потенціал та можливості вирішення різноманітних завдань. Вона має простий та логічний синтаксис, який дозволяє легко створювати програми, в той же час є дуже потужною та масштабованою мовою. C# застосовується для веб-розробки на стороні сервера з використанням технології ASP.NET, а також для розробки десктопних додатків та ігор. Крім того, вона може бути використана для

створення мобільних додатків з використанням фреймворка Xamarin. C# також має добре розвинуту бібліотеку класів та підтримує різні платформи, включаючи Windows, Linux та macOS. Це дозволяє розробникам створювати програми для різних операційних систем та пристроїв. Окрім того, C# також може бути використана для розробки алгоритмів машинного навчання, завдяки фреймворку ML.NET, що дозволяє створювати моделі машинного навчання без великих знань в цій галузі.

### 3.1.2 База даних

Компанія Microsoft розробляє комерційну систему управління базами даних - Microsoft SQL Server. Для створення запитів у цій системі використовується мова Transact-SQL, що була розроблена спільними зусиллями Microsoft і Sybase. Transact-SQL відповідає стандартам ANSI/ISO щодо мови запитів SQL з розширеннями, і використовується як для невеликих та середніх, так і для великих баз даних, що задовольняють потреби підприємств. [26] [27]

Система успішно конкурує зі схожими системами управління базами даних протягом багатьох років. Базовий код MS SQL Server був заснований на коді Sybase SQL Server, що дало можливість Microsoft вийти на ринок баз даних для підприємств і змагатися з Oracle, IBM та Sybase. Початково Microsoft, Sybase і Ashton-Tate об'єднали зусилля для розробки та запуску першої версії програмного продукту, який був названий SQL Server 1.0 для OS/2, що фактично був еквівалентом Sybase SQL Server 3.0 для Unix та VMS. У 1992 році була випущена версія Microsoft SQL Server 4.2, яка входила до складу операційної системи Microsoft OS/2 версії 1.3. [26]

Випуск версії 4.21 Microsoft SQL Server для ОС Windows NT відбувся одночасно з випуском самої ОС Windows NT (версії 3.1). Першою версією SQL Server, розробленою для архітектури NT без участі Sybase в процесі розробки, був Microsoft SQL Server 6.0.

Система Microsoft SQL Server використовує мову запитів SQL з розширенням Transact-SQL, яка є програмною реалізацією стандарту SQL-92. T-SQL дозволяє використовувати додатковий синтаксис для збережених процедур та забезпечує підтримку транзакцій. Для взаємодії з мережею, Microsoft SQL Server та Sybase ASE використовують протокол Tabular Data Stream. Система також підтримує інтерфейс взаємодії ODBC та можливість підключення через веб-сервер-сервіси, що використовують протокол SOAP. Крім того, для застосунків під керування Java, Microsoft надає сертифікований драйвер JDBC. SQL Server надає можливість дзеркалювати та кластеризувати бази даних. [28]

Кластер серверу SQL являє собою сукупність серверів, що було однаково конфігуровано. Ця схема надає можливість розподілити робоче навантаження між декількома серверами. Кожний сервер має одне віртуальне ім'я, а всі дані розподіляються за IP-адресами машин кластеру протягом робочого циклу. У разі відмови або збою на одному з серверів кластеру буде надано автоматичне перенесення навантаження на інший сервер. Система SQL Server надає можливість надлишкового дублювання даних за трьома можливими сценаріями розробки:

- знімок виконує «знімок» бази даних, який сервер потім надсилає користувачам;
- історія змін розроблена для постійної безпомилкової передачі всіх змін користувачам;
- синхронізація з іншими серверами передбачає, що бази даних декількох серверів будуть синхронізуватися між один одним. Зміни в усіх базах даних будуть відбуватися незалежно на кожному сервері. А вже у той час коли синхронізація відбувається, системою буде проведена звірка даних. [29]

На рисунку 3.1 зображена діаграму сутності та зв'язків нашого застотуснку.

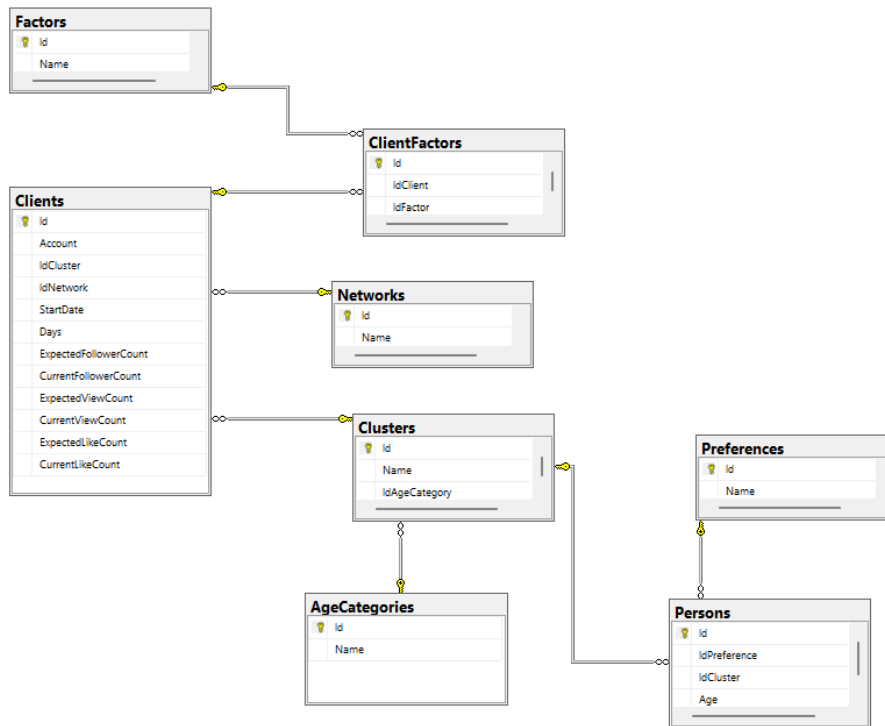


Рисунок 3.1 – ER-діаграма бази даних

### 3.1.3 Entity Framework

Entity Framework - це платформа Object-Relational Mapping (ORM) з відкритим вихідним кодом для додатків .NET, розроблена Microsoft. Вона дозволяє розробникам працювати з даними у вигляді об'єктів доменних класів, замість безпосередньої роботи з таблицями та стовпцями бази даних, де ці дані зберігаються.

За допомогою Entity Framework розробники можуть використовувати вищий рівень абстракції, коли мають справу з даними, що дозволяє їм більш ефективно та зручно працювати з даними. Entity Framework спрощує процес створення програм, що орієнтовані на дані, і зменшує кількість коду, що потрібно для розробки та підтримки таких програм.

Отже, Entity Framework дозволяє розробникам працювати з даними на більш високому рівні абстракції та забезпечує більш простий та зручний

спосіб роботи з базами даних, що робить розробку програм, орієнтованих на дані, більш ефективною та простою.

Entity Framework - це інструмент, який дозволяє розробникам .NET працювати з базами даних за допомогою об'єктів .NET, використовуючи концепцію Object-Relational Mapping (ORM). Це означає, що Entity Framework дозволяє замінити роботу з традиційними таблицями та стовпцями баз даних на роботу з об'єктами доменних класів, що представляють дані в програмі.

Це усуває необхідність в багатьох рядках коду доступу до даних, який зазвичай потрібно писати розробникам. За допомогою Entity Framework розробники можуть зосередитись на роботі з об'єктами та логікою програми, не на роботі з базою даних.

Entity Framework знаходиться в архітектурі програми програмного забезпечення як інструмент, який дозволяє розробникам .NET працювати з базою даних у вигляді об'єктів .NET, що спрощує розробку та підтримку програм, орієнтованих на дані.

Рисунок 3.2 може демонструвати розміщення Entity Framework у контексті архітектури програмного забезпечення.

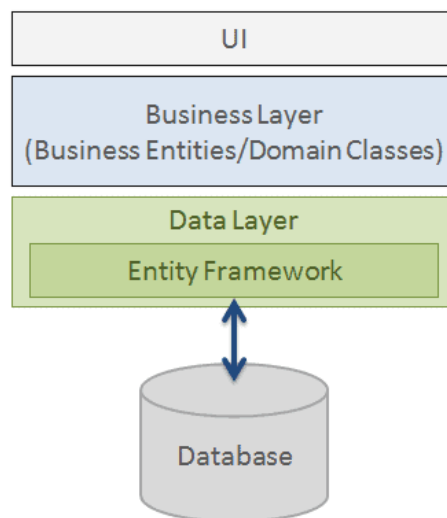


Рисунок 3.2 – Архітектура програмного забезпечення з Entity Framework

### 3.1.4 Visual Studio Code

Visual Studio Code (VS Code) є мінімалістичним, але потужним редактором вихідного коду, який можна запустити на комп'ютерах з операційними системами Windows, macOS та Linux. Основна мета VS Code полягає у забезпеченні зручного та ефективного досвіду редагування коду для розробників. Редактор підтримує широкий спектр мов програмування, включаючи JavaScript, Node.js та TypeScript, а також має розширення для інших мов, таких як C++, C#, Python, Go та Java. [35]

Загалом, VS Code є простим текстовим редактором, але завдяки розширенням він надає багато додаткових можливостей. Розширення для VS Code можуть додавати додаткові функції до наявного набору, наприклад, автоматичне завершення коду, що спрощує процес розробки. Це означає, що розробник може скористатися підказками для написання коду та автоматично завершувати певні блоки. Функція динамічного попереднього перегляду також може бути налаштована, що дозволяє автоматично переглядати веб-сторінку у режимі реального часу при виконанні певних змін. Ці можливості є дуже корисними для розробників та значно спрощують їх роботу. [35]

Висновуючи, VS Code є потужним та легким у використанні текстовим редактором для розробки веб-застосунків. Він працює ефективно на різних платформах та надає всі необхідні інструменти для швидкої та зручної розробки. Доступні розширення значно полегшують роботу розробників та дозволяють зосередитись на головних завданнях, що дозволяє їм прискорити процес розробки.

### 3.1.5 GitHub

Для керування версіями файлів було використано систему git, яка є розподіленою.

Порівняно з централізованими системами, використання розподіленої системи контролю версій має наступні переваги:

- розробник може періодично синхронізувати кілька своїх комп'ютерів (наприклад, робочий, домашній, ноутбук тощо) без потреби виділяти один із них як сервер. Використання розподіленої системи дозволяє здійснювати синхронізацію за потребою, що є перевагою порівняно з централізованими системами, де необхідно мати окремий сервер для цієї мети; [30][31]
- якщо невелика група розробників працює над спільним проектом з різних місць, вони можуть використовувати розподілену систему контролю версій, щоб працювати разом без необхідності виділяти загальні ресурси. При цьому вони можуть синхронізувати свої репозиторії періодично за схемою "кожен з кожним", що дозволяє уникнути потреби мати центральний сервер; [30][31]
- якщо розробники працюють над великим розподіленим проектом, що містить багато частин, і мають обмежене з'єднання з мережею, вони можуть використовувати централізований сервер для синхронізації своїх копій проекту з центральним репозиторієм. Якщо проект складається з декількох напрямків роботи, розробники можуть створити групи, і використовувати групові сервери для синхронізації роботи в межах кожної групи, а потім об'єднати зміни з різних груп на головному сервері. Якщо групові сервери не потрібні, розробники можуть синхронізувати свої зміни між собою і передати їх керівнику групи, який потім передає зміни на головний сервер. [30][31]

Git має такі переваги порівняно з іншими розподіленими системами контролю версій:

- висока продуктивність;
- продумана система команд;
- репозиторії git можуть поширюватися і оновлюватися загальносистемними файловими утилітами резервного копіювання та

оновлення, завдяки тому, що фіксації змін і синхронізації не змінюють існуючі файли з даними, а лише додають нові (за винятком деяких службових файлів, які можуть бути автоматично оновлені за допомогою наявних у складі системи утиліт);

- для поширення репозиторію по мережі досить будь-якого веб-сервера.

Згідно з метою магістерської дисертації було описано основні засоби розробки, які сприяють оптимізації процесу написання коду для програмного забезпечення. Далі необхідно провести аналіз і тестування даних моделей, підібрати вхідні параметри та забезпечити потрібну інформацію. Активність розробки програмного забезпечення показана на рисунку 3.3.



Рисунок 3.3 – Активність розробки програмного забезпечення

## 3.2 Інформаційне забезпечення

Розглянемо більш детально вхідні дані та параметризацію, які були застосовані при реалізації сервісу.

### 3.2.1 Вхідні дані

Система отримує на вхід наступні дані:

- логін користувача;
- вікова категорія для просування;
- соціальна мережа для просування;
- кількість днів просування;
- поточні фактори успішності (впливовості);
- метод ARIMA;
- інтереси людей.

### 3.2.2 Вихідні дані

Вихідними даними, які будуть отримані після розв'язання задачі створення цільових аудиторій методом кластеризації, є:

- кількість кластерів для розподілу;
- кластери, тобто цільові аудиторії, які розподілені за віковою категорією та інтересами людей.

Вихідними даними, які будуть отримані після розв'язання задачі визначення найвпливовіших факторів успішності методом головних компонент та алгоритмом прогнозування, є:

- список найвпливовіші факторів успішності просування контенту;
- статистика впливовості кожного фактору з цього списку.

Вихідними даними, які будуть отримані після розв'язання задачі прогнозування очікуваної кількості приросту цільової аудиторії методом ARIMA, є:

- певний проміжок часу;
- очікувана кількість підписників;
- очікувана кількість вподобань контенту;
- очікувана кількість переглядів контенту;
- прогноз результативності просування для цього часового проміжку.

### 3.2.3 Параметризація

Для реалізації можливого перегляду прогнозу за вхідними даними потрібно використати такі параметри:

- функція поточного стану просування – використовується для знаходження відсоткового співвідношення до очікуваного стану, що далі використовується у знаходженні показника КРІ;
- функція знаходження КРІ – використовується для вирахування показника КРІ на певний момент часу, за допомогою поточного та очікуваного станів просування контенту;
- кількість просувань за поточний місяць – використовується для побудови статистики у відповідному вікні;
- кількість просування у мережі Instagram – використовується для співвідношення просувань між мережами Instagram та Facebook;
- кількість просування у мережі Facebook - використовується для співвідношення просувань між мережами Instagram та Facebook;
- показник прогрес-бара – використовується для візуального показу співвідношень між мережами Instagram та Facebook.

### 3.4 Бібліотеки

Для успішної реалізації проєкту було використано різні бібліотеки мови програмування C#. Основні з них: EntityFramework та Material Design.

#### 3.4.1 Material Design

Material Design - це дизайн-бібліотека, яку винайшли та представили у компанії Google у 2014 році. Це не просто гайдлайн для однакового візуального оформлення - завдяки ньому були уніфіковані інтерфейси всіх продуктів та сервісів корпорації, зокрема ОС Android. У результаті їх загальна

сукупність сприймається як єдина цифрова система, створюючи таким чином новий користувацький досвід та забезпечуючи проникнення сервісів корпорації в усі сфери життя людини. [32]

До впровадження цієї концепції не то що в різних продуктах, а навіть в різних версіях одного сервісу (в мобільній та десктопній) дизайн значно відрізнявся. Через це користувачам було важко орієнтуватися при переході між ними. Навіть спроба розробити однаковий стиль для додатків Android Nolo не вирішила проблему - люди все одно плуталися в інтерфейсах. Лише з переходом до принципу Material Design користувацький досвід вдалося покращити. [32]

На основі концепції Material Design лежать 4 принципи:

Тактильні поверхні. Основою формування простору є "цифровий папір". Окремі його аркуші розташовуються на різній висоті і при цьому відбивають тіні один на одного. Крім того, вони також можуть розтягуватися, змінювати форму і кольори, а також з'єднуватися між собою. Завдяки цьому користувачі краще розуміють, як працює система та яка її ієрархія;

Поліграфічний дизайн. Це означає, що при створенні інтерфейсу для цифрових пристроїв використовуються традиційні засоби і підходи з графічного дизайну. Кажучи простіше, на "цифровому папері" елементи відображаються за допомогою "цифрових чорнил";

Грамотна анімація. Тут анімація не просто виринає з нікуди, а з'являється відповідно до логіки системи. Тобто один об'єкт, реагуючи на дії користувача, плавно переходить в інший;

Адаптивний дизайн. Тут все просто - інтерфейс повинен бути оптимізований на всіх пристроях і екранах, незалежно від того, який продукт ми беремо в якості прикладу. При цьому вигляд, функціональність та реакція повинні бути однаковими скрізь.

Google при створенні Material Design орієнтувався на користувацький досвід та функціональність деталей системи. Як наслідок, інтерфейс стає

більш простим і виразним. Саме анімація є основним елементом системи, завдяки якому робляться наступні акценти.

Демонстрація взаємозв'язків. Завдяки анімації стає простим і зрозумілим ієрархія, людина без труднощів зможе зрозуміти, що станеться при натисканні на конкретну деталь;

Фокусування уваги. Анімація ні в якому разі не відволікає від основної дії, що відповідає третьому принципу - її усвідомленості. Виразність. Кожен продукт компанії має свій особливий характер, стиль, і все це чітко відображено в анімованих деталях;

Навчання. Люди без особливих зусиль можуть зрозуміти, як виконувати певні дії;

Привабливість. Навіть якщо виключити практичну складову, то вся анімація сама по собі виглядає досить цікаво. У наслідок цього користувач виявляє інтерес та бажає взаємодіяти з продуктом.

Завдяки концепції Material Design багато компаній переглянули свій підхід до створення дизайнерських елементів. Зокрема, анімація з доповнення перетворилась на повноцінну складову інтерфейсу. При цьому на перший план було поставлено саме користувацький досвід, його оптимізацію та спрощення. Цей підхід став популярним і задав напрямок для розвитку всього цифрового дизайну. [33] [34]

### 3.5 Вимоги до технічного забезпечення

Технічні вимоги, які необхідні для запуску проєкту з програмного забезпечення:

- оперативна пам'ять – 4 ГБ або більше;
- жорсткий диск – 300МБ або більше;
- встановлений компілятор мови C#;
- встановлений фреймворк Entity Framework;
- встановлений SQL Server 2019.

### 3.6 Архітектура програмного забезпечення

На рисунку 3.4 зображена архітектура системи для просування контенту. Вона складається з основних двох компонент: графічна оболочка та основна логіка. Логіка включає в себе всі функції розрахунків, тобто показник КРІ, очікувана та поточна кількість підписників, вподобань, лайків, тощо. Графічна оболочка включає в собі бібліотеку Material Design та WPF, що розроблялась за допомогою XAML розмітки. Користувач взаємодіє з вікном, на якому за допомогою фрейм-менеджера показані сторінки. Таким чином архітектура складається в одне ціле та співпрацює з користувачем.

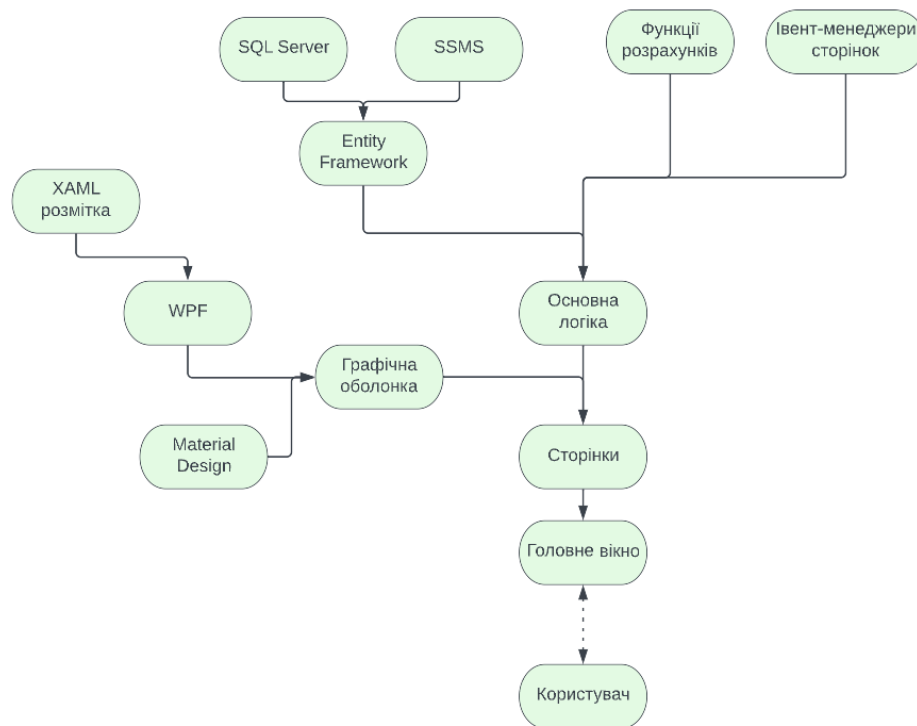


Рисунок 3.4 – Архітектура системи

### 3.7 Опис функціональної моделі

Основним актором розробленої системи є користувач. Користувач має змогу переходити на п'ять сторінок: «Головна сторінка», «Новий клієнт», «Переглянути історію клієнта», «Розрахунок КРІ», «Статистика за місяць».

За допомогою головної сторінки відбувається навігація між іншими. Тобто на інші сторінки можна перейти лише за допомогою головного меню, що якраз і є тільки на головній сторінці.

Сторінка «Новий користувач» відповідає за створення нової заявки для просування контенту, з вибором певної аудиторії (кластеру).

Сторінка «Переглянути історію клієнта» дає змогу переглянути виконані та невиконані заявки певного клієнта, відсортувати їх за будь-яким полем, або просто подивитись на їх кількість чи дату.

Сторінка «Розрахунок КРІ» відповідає за розрахунок КРІ для активних і завершених заявок клієнта.

Сторінка «Статистика за місяць» дає змогу отримати статистику по заявках, а саме:

- кількість просувань за місяць;
- порівняння кількості за мережами (Instagram, Facebook);
- середній КРІ за місяць.

На основі цих даних побудуємо таблицю 3.1 функціональних вимог.

Таблиця 3.1 – Функціональні вимоги

Актор	Варіант використання	Функціональна вимога	Пріоритет
Користувач	Перехід між сторінками	Застосування надає можливість вільно користуватись навігацією програми	Важливий
Користувач	Створення заявки клієнта	Основною функцією є заповнення даних про клієнта, його вибір аудиторії для просування, вибір часового проміжку та вибір вікової категорії	Важливий

## Продовження таблиці 3.1

		заявки, часовий проміжок, кластер	
Користувач	Перегляд історії	Система повинна містити можливість переглядати історію клієнта, тобто дату	Важливий
Користувач	Розрахунок КРІ	Система повинна містити можливість розрахунку показника КРІ для будь-якої активної або завершеної заявки	Важливий
Користувач	Статистика за місяць	Система повинна мати сторінку для показу статистики: кількості заявок, їх співвідношення між мережами, середній показник КРІ	Середній

## 3.8 Специфікація функцій

У таблиці 3.2 наведено функції, які використовуються в реалізованому програмному забезпеченні.

Таблиця 3.2 – Специфікація функцій

Сигнатура функції	Вхідні параметри	Опис функції
AddClient()	-	Отримує дані всередині класу та додає до бази

Продовження таблиці 3.2

Сигнатура функції	Вхідні параметри	Опис функції
		даних нову заявку
RadioButtonChecked()	Object sender, RoutedEventArgs e	Присвоювання змінній вибору мережі, чи Facebook, чи Instagram
SetKPI	-	Обрахунок КРІ для усіх активних і завершених заявок цього клієнта
SetHistory()	-	Зібрання інформації з бази даних про заявки клієнта та показ їх у таблиці
SelectAccount()	-	Вибір клієнта для подальших дій
Button_Click()	Object Sender, EventArgs e	Навігація відповідно до наступної сторінки

## Продовження таблиці 3.2

Сигнатура функції	Вхідні параметри	Опис функції
Stats()	-	Обрахування статистики за поточний місяць

## 3.9 Діаграма класів

При розробці програмного забезпечення використовувалось об'єктно-орієнтоване програмування, тому для збереження даних було створено класи. Це значно спростило реалізацію та зробило архітектуру проекту більш гнучкою до змін та структуризованою.

Діаграми класів наведені у графічних матеріалах (додаток-). Розглянемо детально реалізовані класи та їх призначення.

Клас `AgeCategory` містить атрибути, що використовуються для розподілу вікових категорій, а саме ідентифікатор та назва категорії.

Клас `Network` містить атрибути, що використовуються для розподілу між соціальними мережами, у нашому випадку це Facebook та Instagram. Атрибутами є ідентифікатор та назва мережі.

Клас `Factor` містить атрибути, що використовуються для створення плану для просування контенту. Це ідентифікатор та назва фактору впливовості.

Клас `Preference` містить атрибути, що використовуються для організації цільової аудиторії. Тобто це ідентифікатор та назва кластеру.

Клас `Client` містить атрибути, що використовуються задля створення нової заявки, щодо просування контенту. Основними атрибутами є ідентифікатор, логін клієнта, цільова аудиторія, соціальна мережа, дата початку, час просування, очікувана кількість підписників, вподобань та переглядів, а також поточна кількість.

Клас ClientFactor містить атрибути, що використовуються задля створення зв'язку багато-багато у базі даних. Таким чином наявність ідентифікатора клієнта і фактору дозволяє при організації плану використовувати декілька факторів впливовості.

Клас Cluster містить атрибути, що відповідають за цільову аудиторію. Тобто це назва кластеру, вікова категорія, а також список людей, що туди входить.

Клас Person містить атрибути, що визначають людину за віковою категорією і інтересами. Таким чином вона може бути додана до відповідного кластеру.

Клас PRContent містить атрибути, які відповідають за базу даних, тобто таблиці, які знаходяться у ній, а також за зв'язки між ними. Метод OnConfiguring підключає базу даних та налаштовує її за певними налаштуваннями, а метод OnModelCreating дозволяє автоматично включати в себе ключі, які потрібні для виконання запитів.

Клас Singleton містить атрибути, задля глобального доступу до інформації. Це сама база даних, а також сторінки додатку.

Діаграма класів показана на рисунку 3.5.

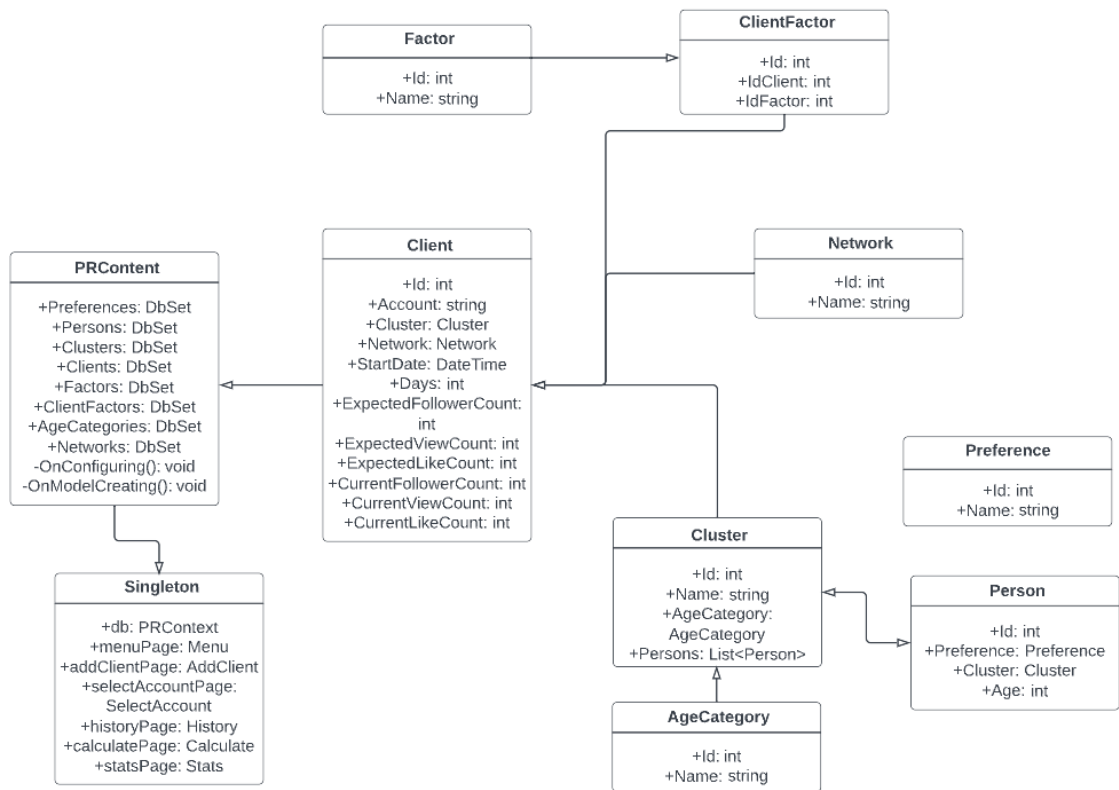


Рисунок 3.5 – Діаграма класів

### 3.10 Діаграма послідовностей

Діаграми послідовності відображають основні процеси взаємодії користувача з системою просування контенту. Основним завданням додатку є створення заявки для просування контенту, тому найголовніша діаграма послідовності пов'язана саме з цим. Також користувач має змогу подивитись історію заявок, або розрахувати КРІ для цих заявок. Це все робиться за допомогою діаграми послідовностей вибору в головному меню. Також користувач має змогу отримати інформацію щодо статистики заявок, тобто кількість заявок за останній місяць, відсоткове співвідношення між соціальними мережами та середній КРІ.

Отож, було створено три діаграми послідовності, які відображають ймовірні дії користувача при випробуванні системи. Усі діаграми послідовності наведені у графічних матеріалах у додатках.

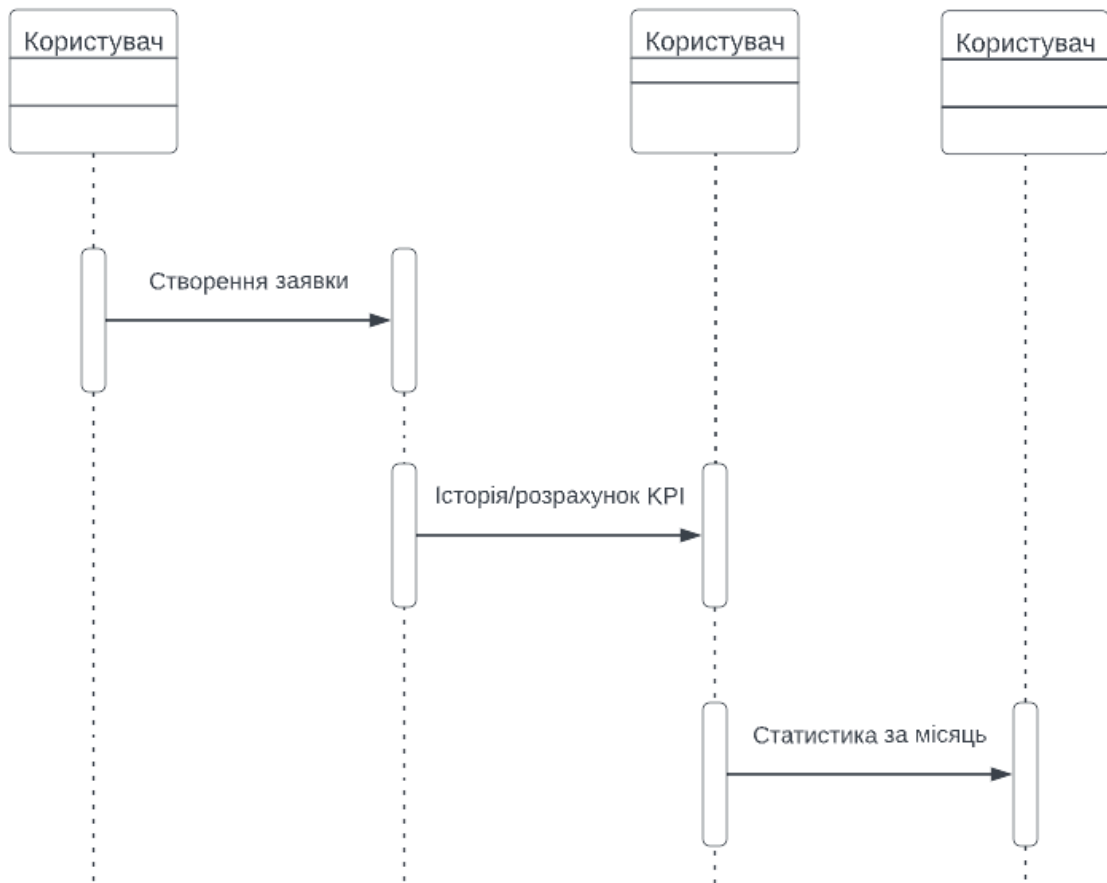


Рисунок 3.13 – Діаграма послідовностей дій користувача при створенні заявки та перегляду статистики за місяць

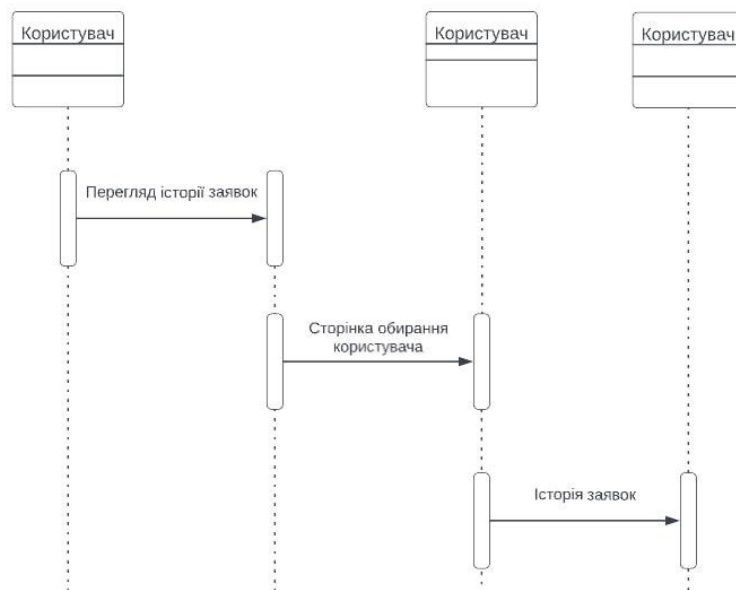


Рисунок 3.14 – Діаграма послідовностей дій користувача при перегляді історії заявок

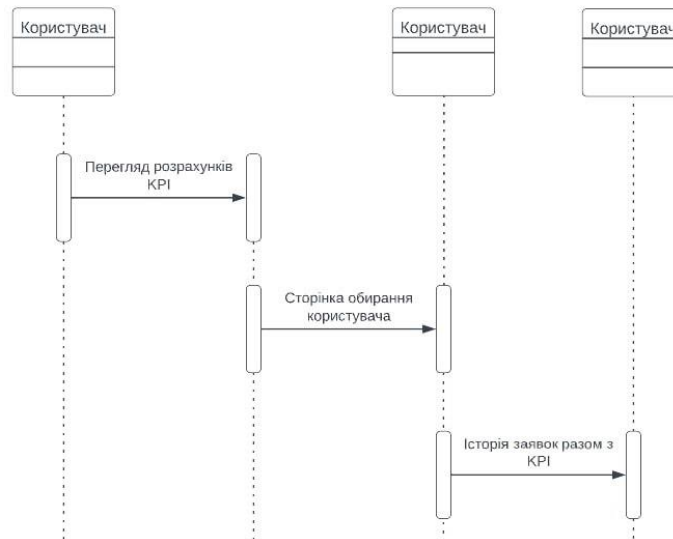


Рисунок 3.15 – Діаграма послідовностей дій користувача при перегляді розрахунків КРІ

#### Висновок до розділу

У даному розділі ми описали технічні деталі реалізації системи, включаючи мінімальні вимоги до програмного забезпечення, вибір засобів розробки та бібліотек. Ми також надали детальну інформацію про вимоги до програмного забезпечення та специфікацію функцій, а також створили діаграми класів, діаграму послідовності та схему варіантів використання.

Для реалізації сервісу була обрана мова програмування C#, оскільки ця мова має багато корисних фреймворків, які полегшують процес створення системи. Крім того, для графічної оболонки було використано WPF та Material Design, які є сучасними та вміють правильно поєднувати компоненти сторінки та відображати їх у сучасному стилі. Це дозволило створити зручну та привабливу для користувачів графічну інтерфейс.

Для зберігання та обробки даних у системі просування контенту було використано базу даних MS SQL Server з фреймворком Entity Framework. Цей фреймворк дозволяє легко реляційно моделювати дані та зручно використовувати їх у програмі.

Для написання коду розробники використовували текстовий редактор VS Code та систему контролю версій Git, GitHub - базові інструменти, які є стандартними для більшості розробників.

Інформаційне забезпечення системи просування контенту надає детальну інформацію про вхідні та вихідні дані, що обробляються в системі. Було проведено детальний опис вхідних та вихідних даних, які надані до системи просування контенту. Вхідні дані включають список контенту, що планується розмістити, інформацію про цільову аудиторію, параметри таргетування та параметри бюджету. Вихідні дані включають рекомендації щодо просування контенту, оцінку успішності кампанії та звіти про результати.

Функціональні вимоги та специфікація функцій значно спрощують процес підтримки проекту, адже документація є важливим етапом в розробці будь-якого програмного забезпечення.

Схема варіантів використання і діаграми послідовності в свою чергу доступно відображають основні дії користувача у взаємодії з системою. Всі діаграми та схеми можна знайти у графічних додатках.

Крім того, для забезпечення якості програмного забезпечення була проведена ручна система тестування. У разі виникнення проблем або помилок, розробники мають можливість звернутися до системи входу, що дозволяє детальніше досліджувати помилки та проблеми, які відображаються в системі.

Для забезпечення зручності використання програмного забезпечення, була створена документація користувача, яка містить інформацію про функціональні системи та інструкції з її використання.

Крім того, у додатку А можна переглянути програмний код реалізованого сервісу, що дозволяє більш детально ознайомитися з реалізаційною системою та її функціоналом.

## 4 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Цей розділ присвячений опису планування та проведенню експериментів над методами, які використовуються для розв'язання поставлених задач. Етап виконання експериментів є надзвичайно важливим, оскільки він дозволяє вибрати найбільш оптимальні параметри та метод для виконання поставленого завдання. Крім того, цей етап є своєрідною перевіркою здатності алгоритмів видавати задовільні результати для кінцевого користувача.

Для проведення експериментів буде використано дані з декількох рекламних кампаній. Для кожної кампанії будуть виконані експерименти з різними методами та параметрами для розв'язання задачі сегментації критерій.

Після проведення експериментів будуть оцінені результати.

### 4.1 Планування експериментів

Виконання експериментів буде виконуватись на наступних типах задач:

- задача створення цільових аудиторій;
- задача визначення найвпливовіших факторів успішності;
- задача прогнозування очікуваної кількості приросту цільової аудиторії;
- задача обрахунку KPI.

### 4.2 Параметризація

4.2.1 Параметри для методу кластеризації  $k$ -середніх з визначенням кількості кластерів методом градієнтного спуску

До параметрів даного методу відносяться:

- кількість ітерацій  $20 < k < 150$ , за замовчуванням  $k = 30$ ;

- складність ітерацій  $m^2n < O(n) < 2ktn + t$ , за замовчуванням  $O(n) = 2ktn + t$ ;
- швидкість навчання  $0.001 \leq lr \leq 0.1$ , за замовчуванням  $lr = 0.1$ .

При проведенні експерименту для пошуку кількості кластерів використовується класичний алгоритм методу градієнтного спуску.

Вхідні дані: люди, їх інтереси та вік.

Вихідні дані: кластери, їх кількість.

#### 4.2.2 Параметри для методу кластеризації $k$ -середніх з визначенням кількості кластерів методом ліктя

До параметрів даного методу відносяться:

- кількість ітерацій  $20 < k < 150$ , за замовчуванням  $k = 30$ ;
- складність ітерацій  $m^2n < O(n) < 2ktn + tk + tn$ , за замовчуванням  $O(n) = 2ktn + tk + tn$ ;
- швидкість навчання  $0.001 \leq lr \leq 0.1$ , за замовчуванням  $lr = 0.1$ .

При проведенні експерименту для пошуку кількості кластерів використовується класичний алгоритм методу ліктя.

Вхідні дані: люди, їх інтереси та вік.

Вихідні дані: кластери, їх кількість.

#### 4.2.3 Параметри для методу головних компонент

До параметрів даного методу відносяться:

- кількість факторів  $5 < k < 30$ , за замовчуванням  $k = 10$ ;
- впливовість кожного із факторів  $0 < f < 1$ .

При проведенні експерименту використовується класичний алгоритм методу головних компонент.

#### 4.2.4 Параметри для методу ARIMA

До параметрів даного методу відносяться:

- обсяг вибірки  $1 < l < 100$ , за замовчуванням  $l = 50$ ;
- кількість лагів  $1 < m < 5$ , за замовчуванням  $m = 2$ .

При проведенні експерименту використовується алгоритм методу часових рядів ARIMA.

#### 4.2.5 Параметри для алгоритму прогнозування

До параметрів даного методу відносяться:

- середня похибка  $0\% \leq \varepsilon_c < 30\%$ , за замовчуванням  $\varepsilon_c = 10\%$ ;
- кількість прогнозів  $1 < K_p < \infty$ ;
- загальна похибка  $0\% < \varepsilon < 10\%$ .

При проведенні експерименту використовується алгоритм методу інтерполяції похибок.

#### 4.2.6 Параметри для обрахунку KPI

До параметрів даного методу відносяться:

- очікувана кількість підписників;
- $A_{r2}$  – очікувана кількість вподобань контенту;
- $A_{r3}$  – очікувана кількість переглядів;
- $A_{f1}$  – фінальна кількість підписників;
- $A_{f2}$  – фінальна кількість вподобань контенту;
- $A_{f3}$  – фінальна кількість переглядів.

При проведенні експерименту використовується класичний алгоритм методу обрахунку KPI.

## 4.3 План експериментів

### 4.3.1 План експериментів для методу кластеризації методом градієнтного спуску

На початку ми виконаємо метод з кількістю ітерацій за замовчуванням, а потім порівняємо його результати з результатами методу з кількістю ітерацій, яка дорівнює 100 ( $k = 100$ ). Це дозволить нам оцінити, як кількість ітерацій впливає на результат методу градієнтного спуску.

Другим експериментом буде зміна швидкості навчання  $lr$ . Ми порівняємо результати методу градієнтного спуску зі значенням швидкості навчання за замовчуванням з результатами методу, який використовує значення  $lr = 0.5$ . Також ми порівняємо результати з середнім значенням  $lr$  вибірки, щоб оцінити вплив різних значень швидкості навчання на результат методу градієнтного спуску.

Обидва експерименти дозволять нам дослідити вплив різних параметрів методу градієнтного спуску на результат. Порівняння складності ітерацій та результатів покажуть, які параметри методу можуть бути налаштовані для покращення швидкості та якості збіжності.

Третім експериментом буде визначення кількості кластерів за різними вхідними даними. Перші вхідні дані будуть такими:

- підлітки (10-18);
- дорослі (18-60);
- люди похилого віку (від 60).

Другі вхідні дані будуть такими:

- підлітки (10-21);
- дорослі (21-65);
- люди похилого віку (від 65).

Після чого ми обрахуємо показник Recall Precision, що дасть визначити, наскільки точна кластеризація відбулась.

#### 4.3.2 План експериментів для методу кластеризації методом ліктя

Ми проведемо експерименти з використанням тих же значень параметрів, що і в методі градієнтного спуску. Це необхідно для порівняння результатів та вибору кращого підходу.

На початку ми виконаємо метод з кількістю ітерацій за замовчуванням, а потім порівняємо його результати з результатами методу з кількістю ітерацій, яка дорівнює 100 ( $k = 100$ ). Це дозволить нам оцінити, як кількість ітерацій впливає на результат методу градієнтного спуску.

Другим експериментом буде зміна швидкості навчання  $lr$ . Ми порівняємо результати методу градієнтного спуску зі значенням швидкості навчання за замовчуванням з результатами методу, який використовує значення  $lr = 0.5$ . Також ми порівняємо результати з середнім значенням  $lr$  вибірки, щоб оцінити вплив різних значень швидкості навчання на результат методу ліктя.

Обидва експерименти дозволять нам дослідити вплив різних параметрів методу градієнтного спуску на результат. Порівняння складності ітерацій та результатів покажуть, які параметри методу можуть бути налаштовані для покращення швидкості та якості збіжності.

Третім експериментом буде визначення кількості кластерів за різними вхідними даними. Перші вхідні дані будуть такими:

- підлітки (10-18);
- дорослі (18-60);
- люди похилого віку (від 60).

Другі вхідні дані будуть такими:

- підлітки (10-21);

- дорослі (21-65);
- люди похилого віку (від 65).

Після чого ми обрахуємо показник Recall Precision, що дасть визначити, наскільки точна кластеризація відбулась.

#### 4.3.3 План експериментів для методу головних компонент

Першим експериментом буде зміна кількості компонент. Ми порівняємо, які компоненти стали важливими, а які ні, у двох випадках:

- $k = 5$ ;
- $k = 10$ .

де  $k$  – кількість компонент.

Другим експериментом буде зміна впливовості кожного фактору. Ми порівняємо випадки, коли впливовість кожного фактору є за замовчуванням, та коли вони обернені.

Третім експериментом буде зміна впливовості кожного фактору, коли  $k = 10$ . Таким чином, ми можемо порівняти, як зміна впливовості кожного фактору впливає на результат, коли використовується більша кількість компонент.

#### 4.3.4 План експериментів для методу часових рядів ARIMA

Першим експериментом буде порівняння результатів для різних обсягів вибірки:

- $l = 10$ ;
- $l = 50$ ;
- $l = 100$ .

Другим експериментом буде дослідження впливу кількості лаг на результат при звичайній вибірці. Тобто дослідим часове відставання або випередження. Розглянемо такі варіанти, коли:

- $m = 2$ ;
- $m = 3$ ;
- $m = 4$ .

Третім експериментом буде дослідження впливу кількості лаг при вибірці розміром  $l = 100$ . Дані про кількість лаг беруться з другого експерименту.

#### 4.3.5 План експериментів для алгоритму прогнозування

Першим експериментом буде вивчення впливу середньої похибки на результати. Ми проведемо дослідження в трьох варіантах, де значення середньої похибки будуть:

- $\varepsilon = 0\%$ ;
- $\varepsilon = 10\%$ ;
- $\varepsilon = 20\%$ .

Другий експеримент буде полягати у вивченні впливу кількості прогнозів в базі на результат прогнозування. Ми розглянемо чотири можливі варіанти, де кількість прогнозів буде дорівнювати:

- $K_p = 1$ ;
- $K_p = 100$ ;
- $K_p = 1000$ ;
- $K_p = 10000$ .

Третім експериментом буде аналіз залежності від загальної похибки та кількості прогнозів, що є в базі. Ми розглянемо чотири можливі варіанти, де кількість прогнозів буде:

- $K_p = 1$ ;

- $K_p = 100$ ;
- $K_p = 1000$ ;
- $K_p = 10000$ .

#### 4.3.6 План експериментів для методу розрахунку KPI

Першим експериментом буде проведено дослідження залежності між очікуваною кількістю підписників, вподобаннями контенту та переглядами. Для цього буде створена таблиця значень, яку буде використано для побудови відповідної діаграми, щоб зрозуміти, які фактори впливають на популярність контенту.

Другим експериментом буде проведено дослідження залежності між реальною кількістю підписників, вподобаннями контенту та переглядами. Для цього також буде створена таблиця значень, яку буде використано для побудови відповідної діаграми, щоб зрозуміти, які фактори насправді впливають на популярність контенту.

### 4.4 Проведення експериментів

4.4.1 Проведення експериментів з використанням методу кластеризації методом ліктя

#### Експеримент № 1

Ми розглянули складності ітерацій для різних показників  $k$ . Це показано на рисунку 4.1 та 4.2

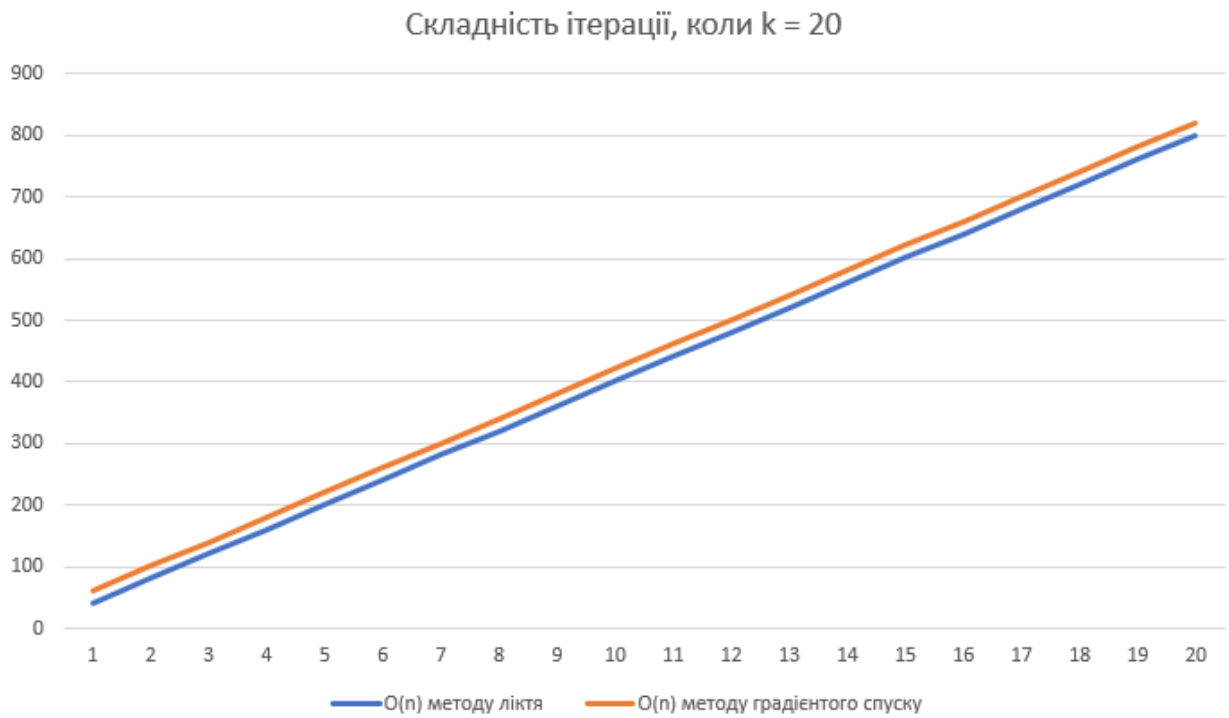


Рисунок 4.1 – Складність ітерації в залежності від  $k$



Рисунок 4.2 – Складність ітерації в залежності від  $k$

Отже, можна зробити висновок, що залежно від  $k$ , складність ітерації прямолінійно зростає.

Експеримент №2

В експерименті №2 ми розглянули зміну швидкості навчання, яка буде відповідати за часові результати. Це показано на рисунку 4.3

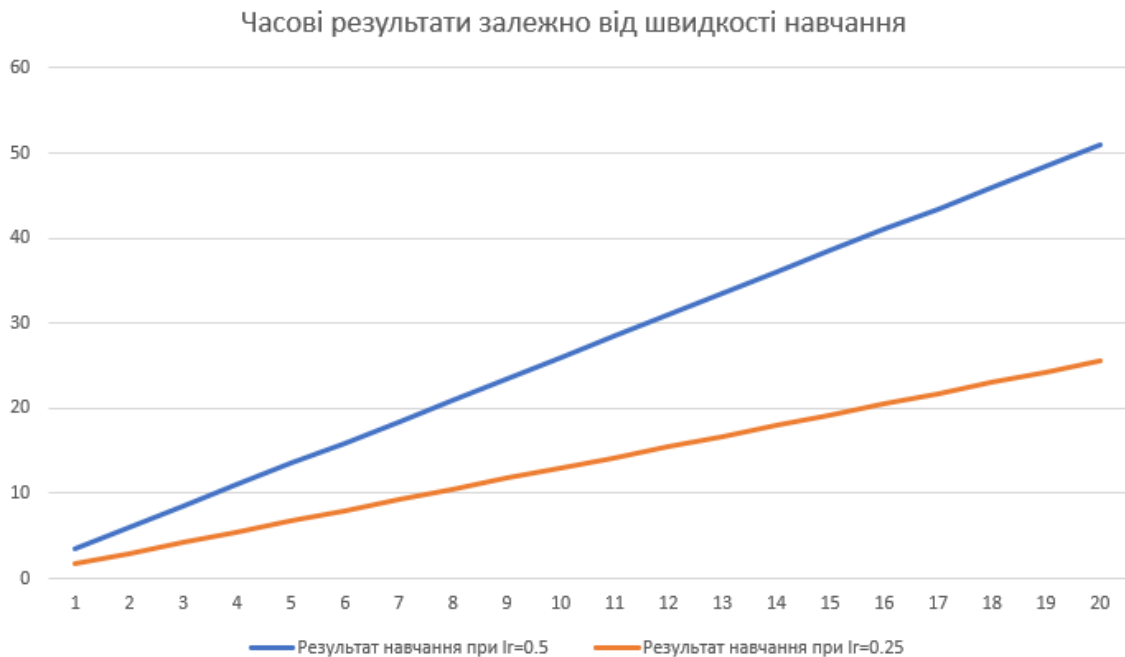


Рисунок 4.3 – Часові результати залежно від швидкості навчання

Отже, можна зробити висновок, що залежно від швидкості навчання, прямолінійно зростають і часові результати, тобто чим більша швидкість, тим менше часу потребується на кластеризацію.

### Експеримент №3

Отже, спочатку візьмемо таблицю даних можливих інтересів, для використання у подальшому обрахунку. Це показано на рисунку 4.4

Results		Messages
	Id	Name
1	1	Кулінарія
2	2	Медицина
3	3	ІТ
4	4	Ігрова сфера
5	5	Кіно
6	6	Театр
7	7	Фітнес та спортзал
8	8	Крипто-сфера
9	9	Бізнес
10	10	Туризм
11	11	Мистецтво

Рисунок 4.4 – Таблиця інтересів

Оскільки всього у нас вікових категорій три, то за максимального порогу ми отримаємо 33 кластери.

Візьмемо інформацію з бази даних та виконаємо кластеризацію. Спочатку знайдемо  $J(C)$  для кожної ітерації. Далі знайдемо квадратичне відхилення та обчислимо похідну, аби знайти кількість кластерів. Фрагмент обчислень показано на рисунку 4.5.

№ ітерації	Похідна	>0
28	0.15	TRUE
29	0.09	TRUE
30	0.07	TRUE
31	0.06	TRUE
32	0.02	TRUE
33	0.01	TRUE
34	0	FALSE
35	0	FALSE
36	0	FALSE
37	0	FALSE

Рисунок 4.5 – Фрагмент обчислень

Бачимо, що починаючи з 34-ої ітерації, похідна дорівнює нулю. Але дивлячись на 33-ю ітерацію, обчислюючи відстань вектору, вони співпадають між собою. Тому у цьому випадку є дві можливі дії: обрати кількість кластерів як 33, або обрати кількість кластерів як 34. Оскільки ми стараємось досягти якомога найкращого порогу, не втративши у цьому швидкість. То кількість кластерів виберемо як число 33.

За допомогою метрики Precision ми рахуємо точність кластеризації. Прорахунок точності представлений у формулі 4.1.

$$precision = \frac{tp}{tp+fp}, \quad (4.1)$$

де  $tp$  – сукупність людей, яка була кластеризована;

$fp$  – сукупність людей, яка не увійшла в кластер.

Якщо людина належить до певної вікової категорії та має свої інтереси, то вона може бути віднесена до конкретного кластеру з цими ж характеристиками, тому  $fp = 0$ .

Отже, оскільки метрика Precision становить 100%, ми можемо стверджувати, що точність дорівнює 100%. Це означає, що значення метрики Recall буде дорівнювати 0.

Отже, можна зробити висновок, що за допомогою методу ліктя ми отримаємо найкращу точність, що дасть користувачу обирати цільову аудиторію з найбільшим інтересом саме до цієї теми.

#### 4.4.2 Проведення експериментів з використанням методу кластеризації методом градієнтного спуску

##### Експеримент № 1

Ми розглянули складності ітерацій для різних показників  $k$ . Це показано на рисунку 4.6 та 4.7

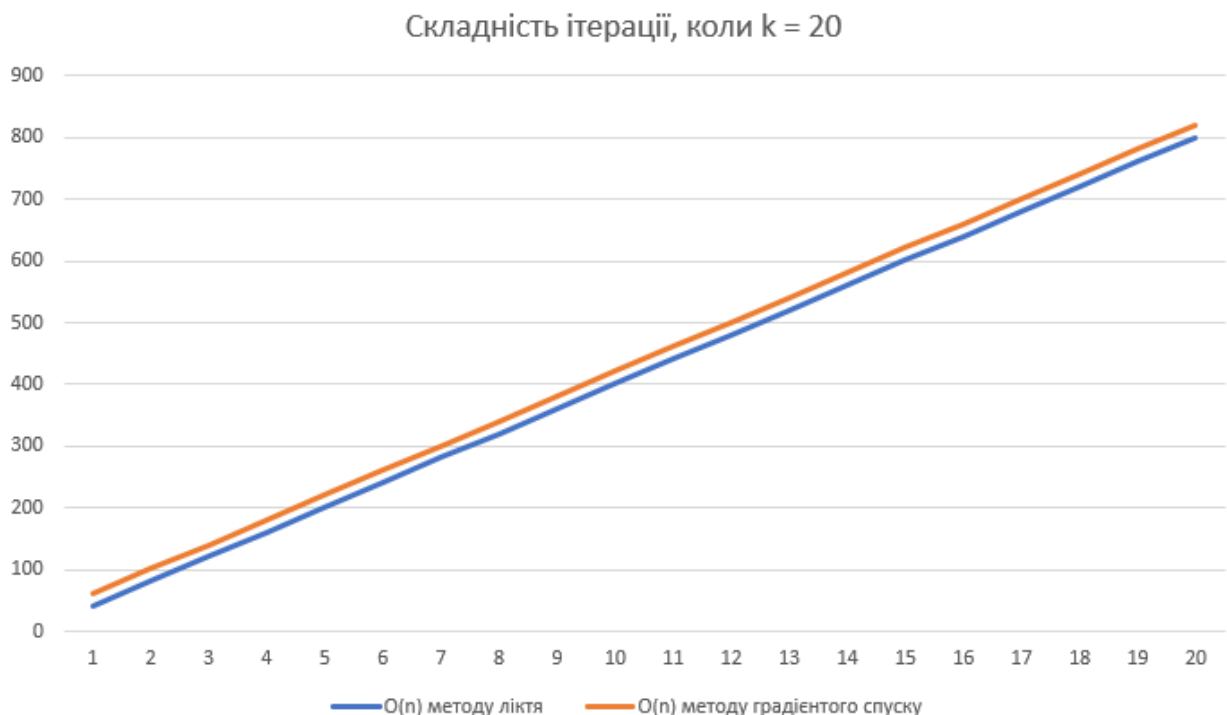


Рисунок 4.6 – Складність ітерації в залежності від  $k$



Рисунок 4.7 – Складність ітерації в залежності від  $k$

Отже, можна зробити висновок, що залежно від  $k$ , складність ітерації прямолінійно зростає.

#### Експеримент №2

В експерименті №2 ми розглянули зміну швидкості навчання, яка буде відповідати за часові результати. Це показано на рисунку 4.8



Рисунок 4.8 – Часові результати залежно від швидкості навчання

Отже, можна зробити висновок, що залежно від швидкості навчання, прямолінійно зростають і часові результати, тобто чим більша швидкість, тим менше часу потребується на кластеризацію.

### Експеримент №3

Отже, спочатку візьмемо таблицю даних можливих інтересів, для використання у подальшому обрахунку. Це показано на рисунку 4.9

	Id	Name
1	1	Кулінарія
2	2	Медицина
3	3	ІТ
4	4	Ігрова сфера
5	5	Кіно
6	6	Театр
7	7	Фітнес та спортзал
8	8	Крипто-сфера
9	9	Бізнес
10	10	Туризм
11	11	Мистецтво

Рисунок 4.9 – Таблиця інтересів

Оскільки всього у нас вікових категорій три, то за максимального порогу ми отримаємо 33 кластери.

Візьмемо інформацію з бази даних та виконаємо кластеризацію. Спочатку знайдемо  $J(C)$  для кожної ітерації. Далі знайдемо квадратичне відхилення і відсортувавши знаходимо ту ітерацію, де це відхилення майже не змінюється від наступного. Фрагмент обрахунків показано на рисунку 4.10.

№ ітерації	Квадратичне відхилення
21	0.94
22	0.46
23	0.31
24	0.25
25	0.24
26	0.23
27	0.2275
28	0.224
29	0.221375
30	0.22075

Рисунок 4.10 – Фрагмент обрахунків

За допомогою метрики Precision ми рахуємо точність кластеризації. Прорахунок точності представлений у формулі 4.1.

$$precision = \frac{tp}{tp+fp}, \quad (4.1)$$

де  $tp$  – сукупність людей, яка була кластеризована;

$fp$  – сукупність людей, яка не ввійшла в кластер.

Бачимо, що починаючи з 24-ої ітерації, квадратичне відхилення падає не так швидко, як раніше. Розглянемо два варіанти, де візьмемо кількість кластерів 25 та 30.

Для 25 кластерів  $fp = 25\%$ , для 30 кластерів  $fp = 10\%$ . Отже, чим більше кластерів – тим більше точність кластеризації.

Вони відрізняються порогом, у першому варіанті поріг складає 75%, а в другому 90%. Таким чином Recall дорівнює 25% та 15% відсотків відповідно.

Отже, можна зробити висновок, що за допомогою методу градієнтного спуску ми отримаємо більше повноту (в кластерах буде більше людей, але може бути втрачена пріоритетність інтересів), але втратимо в точності розрахунків. Оскільки точність задля кластеризації на цільові аудиторії є ваговим показником, тому обрати потрібно метод ліктя.

#### 4.4.3 Проведення експериментів з використанням методу головних компонент

##### Експеримент №1

У цьому експерименті ми визначили, які фактори припадуть до важливих, якщо змінити їх кількість, тобто ми розглянули 5 та 10 факторів, та виконавши метод головних компонент, обрали найвпливовіші. Результати показані на рисунках 4.11, 4.12, 4.13 та 4.14.

№	Фактор	Значення матриці факторного відображення
1	Популярність теми	0.11
2	Правильне оформлення інформації на сторінці	0.22
3	Зовнішній вигляд сторінки	0.11
4	Рівномірне планування постів	0.06
5	Поточний стан сторінки	0.04

Рисунок 4.11 – Набір даних з кількістю 5

№	Важливий фактор
1	Популярність теми
2	Правильне оформлення інформації на сторінці
3	Зовнішній вигляд сторінки
4	Слідування за трендами
5	Привертання уваги

Рисунок 4.12 – Результати методу головних компонент

№	Фактор	Значення матриці факторного відображення
1	Популярність теми	0.11
2	Правильне оформлення інформації на сторінці	0.21
3	Зовнішній вигляд сторінки	0.13
4	Рівномірне планування постів	0.09
5	Поточний стан сторінки	0.04
6	Слідування за трендами	0.1
7	Стильний дизайн	0.02
8	Власний бренд	0.03
9	Привертання уваги	0.19
10	Хештеги	0.01

Рисунок 4.13 – Набір факторів з кількістю 10

№	Важливий фактор
1	Популярність теми
2	Правильне оформлення інформації на сторінці
3	Зовнішній вигляд сторінки

Рисунок 4.14 – Результати методу головних компонент

Отже, змінюючи кількість факторів приходимо до висновку, що кількість важливих факторів також змінюється.

## Експеримент №2

В експерименті №2 були розглянуті фактори з інверсною статистикою, тому значення матриці відрізняються, після чого були обрані нові впливові фактори. Це показано на рисунку 4.15 та 4.16.

№	Фактор	Значення матриці факторного відображення
1	Популярність теми	0.06
2	Правильне оформлення інформації на сторінці	0.05
3	Зовнішній вигляд сторінки	0.03
4	Рівномірне планування постів	0.09
5	Поточний стан сторінки	0.1
6	Слідування за трендами	0.01
7	Стильний дизайн	0.12
8	Власний бренд	0.13
9	Привертання уваги	0.1
10	Хештеги	0.11

Рисунок 4.15 – Набір даних з інверсною статистикою

№	Важливий фактор
1	Хештеги
2	Привертання уваги
3	Власний бренд
4	Поточний стан сторінки

Рисунок 4.16 – Результати методу головних компонент

### Експеримент №3

В експерименті №3 були розглянуті фактори з іншими значенням матриці, аби перевірити які фактори на цей раз оберуться за допомогою метода головних компонент. Це показано на рисунку 4.17 та 4.18.

№	Фактор	Значення матриці факторного відображення
1	Популярність теми	0.05
2	Правильне оформлення інформації на сторінці	0.09
3	Зовнішній вигляд сторінки	0.07
4	Рівномірне планування постів	0.11
5	Поточний стан сторінки	0.2

Рисунок 4.17 – набір даних з різною статистикою для кожного фактору

№	Важливий фактор
1	Рівномірне планування постів
2	Поточний стан сторінки

Рисунок 4.18 – результати методу головних компонент

#### 4.4.4 Проведення експериментів з використанням методу часових рядів ARIMA

##### Експеримент №1

Цей експеримент заключається у знаходженні залежності від вибірки. Таким чином ми маємо три вибірки: 10, 50 та 100. Результати залежності показано на рисунку 4.19.

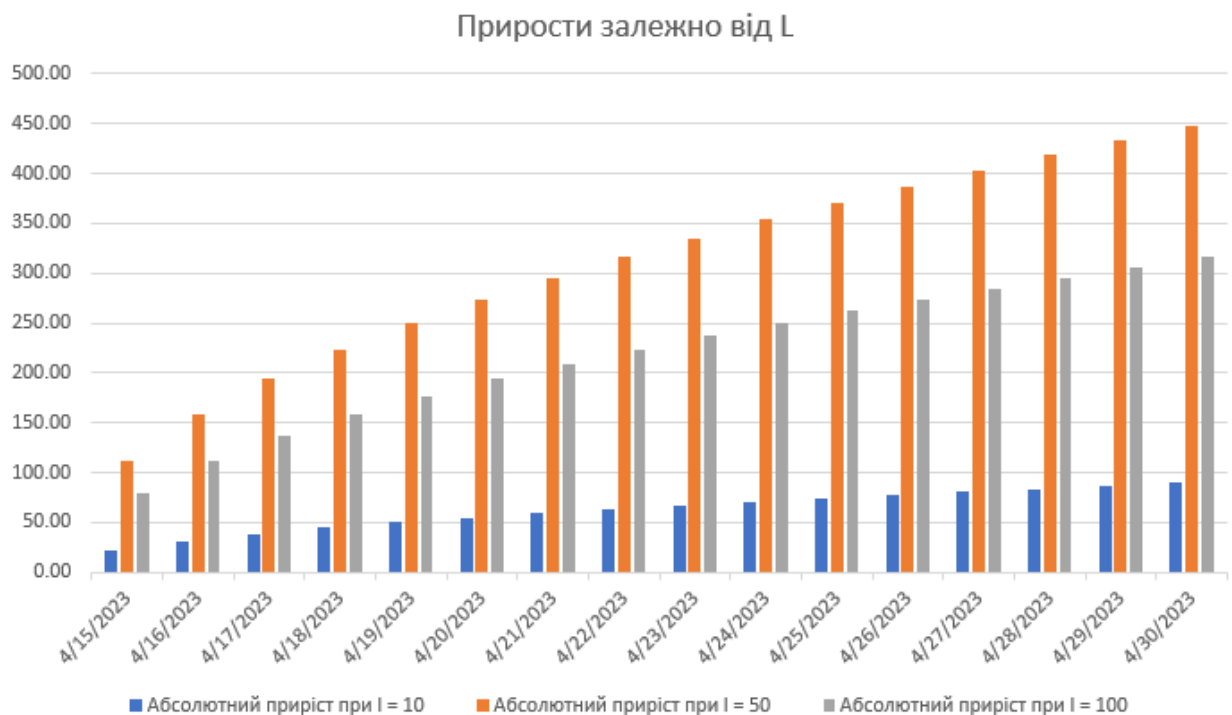


Рисунок 4.19 – Прирости залежно від значення вибірки

##### Експеримент №2

Цей експеримент заключається у знаходженні залежності від кількості лаг(часових відставань). У результаті отримаємо абсолютні значення відставань на кожен день, який передбачається у заявці. Це показано на рисунку 4.20.

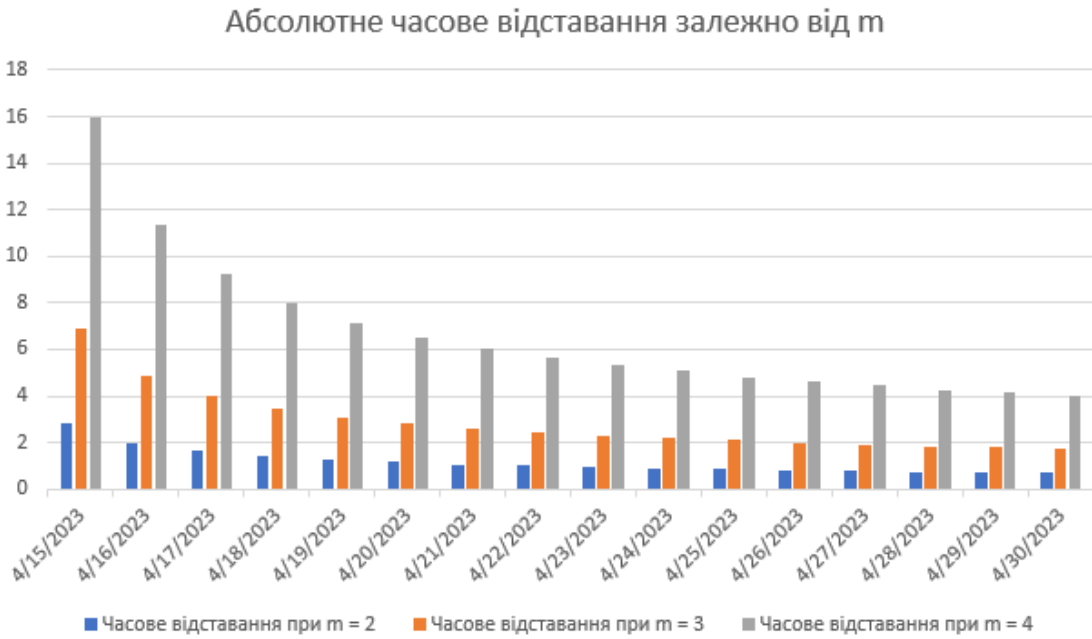


Рисунок 4.20 – Результати залежності від кількості лаг

### Експеримент №3

Цей експеримент заключається також у знаходженні залежності від кількості лаг, але вже при вибірці, що рівна 100. Оскільки вибірка є важливою у часових рядах, то і результати будуть інші. Це показано на рисунку 4.21.

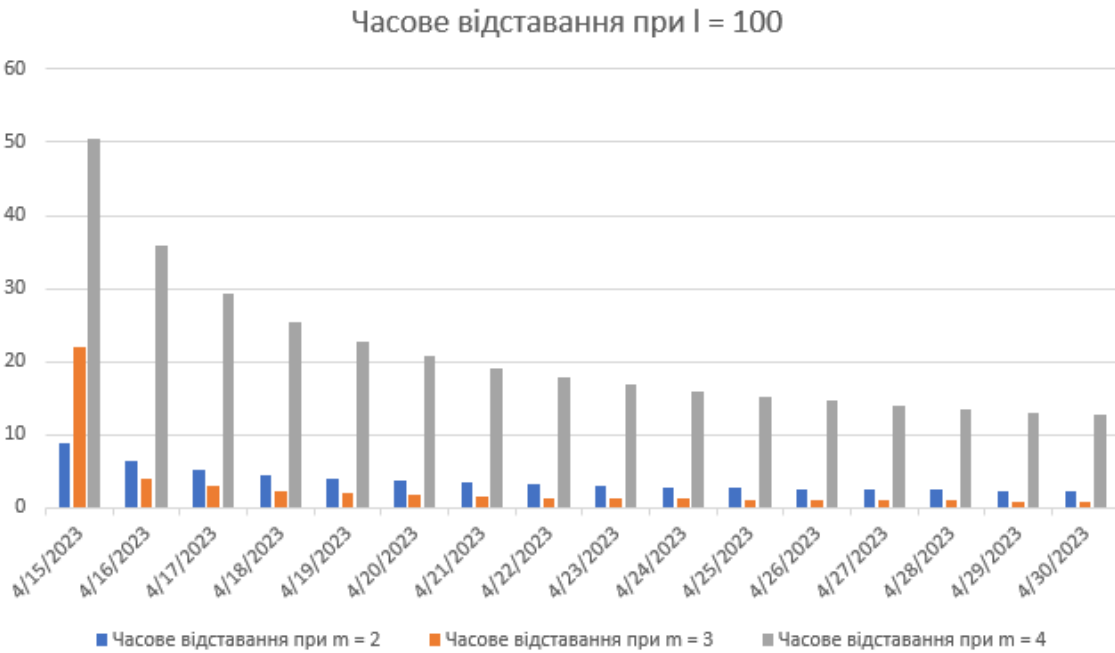


Рисунок 4.21 – Абсолютне часове відставання залежно від вибірки та кількості лаг

#### 4.4.5 Проведення експериментів з використанням алгоритму прогнозування

##### Експеримент №1

У цьому експерименті був проведений аналіз очікуваного та реального приросту залежних від середньої похибки. Було розглянуто похибки зі значеннями 0, 10 та 20. Результат показано на рисунку 4.22.

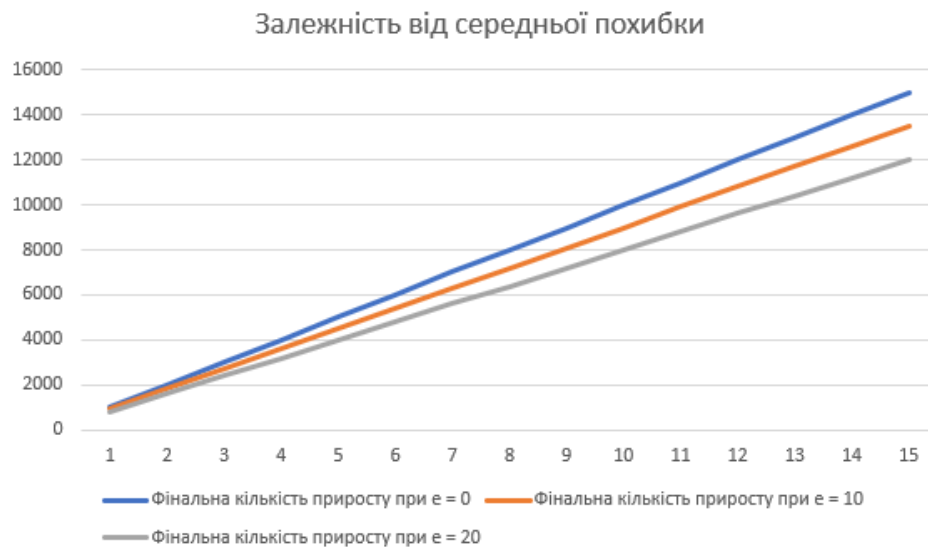


Рисунок 4.22 – Залежність від середньої похибки

##### Експеримент №2

Оскільки очікувана кількість приросту також залежить від кількості заявок, яка допомагає якомога точніше визначити цю кількість, то аналіз було проведено і з кількістю самих заявок. Результат показано на рисунку 4.23.

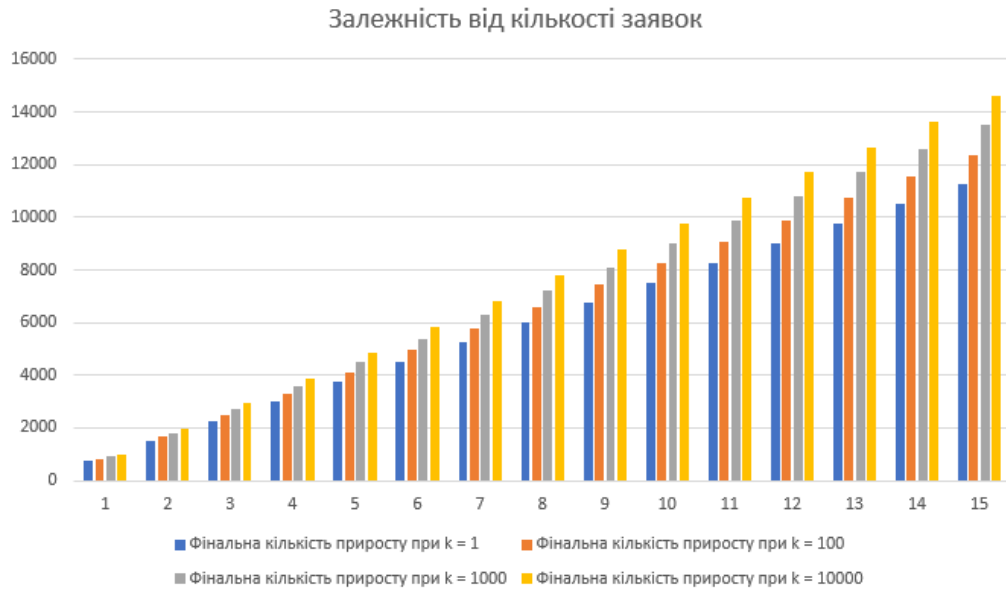


Рисунок 4.23 – Результати залежно від кількості заявок

### Експеримент №3

У цьому експерименті показано результати у залежності від кількості заявок та абсолютної похибки. Наприклад, якщо кількість заявок маленька, то і похибка буде ідентична з подібною кількістю. А ось коли заявок вже багато, то точність сягає до максимуму. Результат показано на рисунку 4.24.

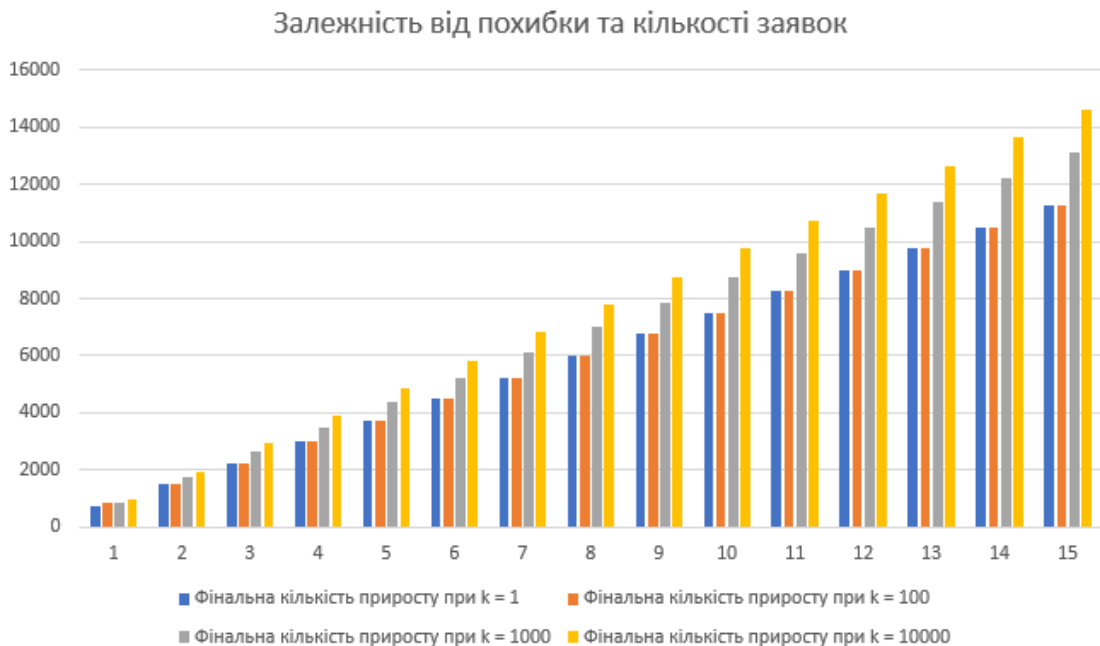


Рисунок 4.24 – Результати залежності від похибки та кількості заявок

#### 4.4.6 Проведення експериментів з методом розрахунку КРІ

##### Експеримент №1

У цьому експерименті був проведений аналіз, що допомагає визначити КРІ залежно від реальної кількості приросту. Таким чином, чим більш реальні кількість приросту збігається з очікуваною, тим більше КРІ доходить до 100%. Результат цього експерименту можна побачити на рисунку 4.25.



Рисунок 4.25 – Результати КРІ залежно від реальної кількості приросту

##### Експеримент №2

Цей експеримент показує залежність очікуваної кількості приросту від реальної. Тобто таким чином можна побачити, що чим більша очікувана кількість приросту, то тим більша і реальна кількість приросту. Тому можна зробити висновок, що якщо кластер буде більшим, то і приріст буде також більшим. А також, якщо похибка сягає мінімуму, то КРІ буде рости до максимуму. Результат цього показано на рисунку 4.26.

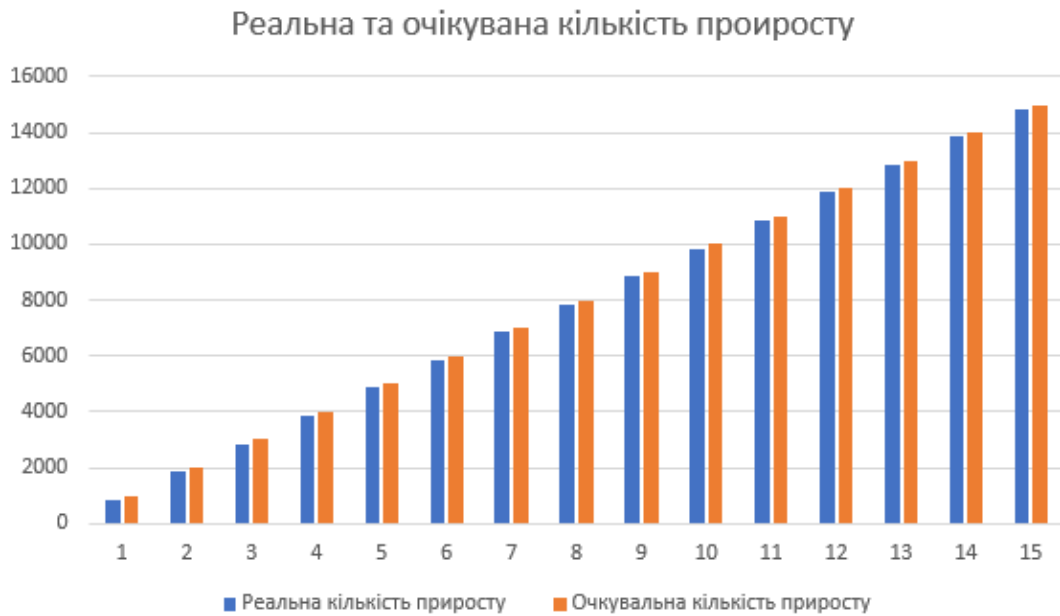


Рисунок 4.26 – Результати залежності від очікуваної кількості приросту

#### Висновок до розділу

У даному розділі було виконано планування і проведення експериментів для методів кластеризації градієнтного спуску, методу ліктя, методу головних компонент, методу часових рядів ARIMA, алгоритму прогнозування та методу обрахунку KPI. Планування було виконано ретельно з чітко зазначеною метою. Це в свою чергу дозволяє дослідити те, яким чином параметри впливають на вихідний результат та переконатись, що методи працюють правильно. Після проведення експериментів було зроблено детальний порівняльний аналіз.

За результатами експериментальних досліджень було зроблено декілька висновків:

- метод ліктя є більш ефективним для наших задач кластеризації, оскільки точність, яка була визначена за допомогою Recall Precision дорівнює ста відсоткам, а в методі градієнтного спуску лише від 70 до 90 відсотків, з урахуванням зміни повноти;

- метод головних компонент обирає лише ті компоненти, які є справді важливим, тобто такі, де значення матриці факторного відбору більше за одну десяту;
- метод головних компонент правильно працює з інверсними даними, що доказує його правильність з різними наборами даних;
- при суттєвій зміні вказаних параметрів у методі ARIMA створюються неправдоподібні ситуації для просування, тому великий обсяг вибірки брати не рекомендовано. Це відбувається, коли обсяг вибірки збільшується у два рази, аніж за замовчуванням, тому подальші прогнози мають неточності, що призведуть до порушення наступного методу – алгоритму прогнозування;
- алгоритм прогнозування найкраще працюватиме тоді, коли кількість заявок буде великою. Тобто збільшуючи кількість заявок, ми отримаємо більш точні дані за допомогою методу часових рядів ARIMA, тим самим обрахувавши більш точніше абсолютну похибку, отримаємо правдиві дані, які будуть схожими на фінальні;
- метод обрахунку KPI працює у зв'язку з методом прогнозування, тому від зміни параметрів алгоритму прогнозування, змінюється і обрахунок KPI.

## 5 РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ

Стартап - це компанія, зазвичай молода, заснована однією або кількома підприємливими особами з метою розробки унікальних продуктів чи послуг, що є новаторськими та конкурентоспроможними. Основна мета стартапу - швидке впровадження інновацій та здійснення креативних ідей. [36]. Головна мета розробки стартап-проєкту полягає в запуску нового та інноваційного продукту на ринок. На жаль, більшість стартапів не досягають успіху та закінчують своє існування на ранніх етапах, що не відповідає очікуванням інвесторів. Тому процес планування та розробки стартап-проєкту потребує відповідального підходу та ретельного аналізу ринку, потенційних конкурентів та споживачів. Це дозволить зробити обґрунтовану оцінку розробленого продукту та зменшити ризики.

### 5.1 Опис ідеї проєкту

Основна ідея проєкту полягає в створенні програмного застосунку, який зможе автоматизувати процес просування контенту у соціальних мережах на основі вхідних даних. Це дозволить значно спростити та прискорити процес рекламної кампанії для фахівців з маркетингу.

Більш детальний опис ідеї надано у таблиці 5.1.

Таблиця 5.1 – Опис ідеї проєкту

Програмний продукт	Напрямки застосування	Вигоди для користувача
Система з підтримки роботи фахівця по маркетингу у	Маркетинг	Є можливості користувача щодо просування контенту клієнта та розгортання рекламної кампанії. Він вказує на те, що користувач може з легкістю розмістити контент і залучити

Продовження таблиці 5.1.

Програмний продукт	Напрямки застосування	Вигоди для користувача
у соціальних мережах	Маркетинг	алгоритми та методи, використані при розробці застосунку, сприяють цьому. Крім того, таргетолог може запустити масштабну рекламну кампанію, що дозволяє розширити охоплення контенту та привернути нових споживачів та клієнтів. Загалом, функціонал застосунку вказує на важливість можливостей просування контенту та рекламних кампаній в соціальних мережах для успіху замовника в рекламному середовищі соціальних мереж.
	Приватний користувач	Цей застосунок не обмежений лише професіоналами у галузі маркетингу, а доступний для будь-якого користувача. Він дозволяє легко просунути власний контент, особистий блог та узагалі підтримувати бізнес в онлайн-середовищі, навіть якщо у користувача немає широких знань у сфері соціального медіа-маркетингу.

Перевага даної ідеї полягає в тому, що просування контенту у соціальних мережах має велику важливість для підприємств та брендів, оскільки це дозволяє залучати нових клієнтів, збільшувати свою аудиторію, зміцнювати відносини з існуючими клієнтами та підвищувати свою репутацію в Інтернеті. Присутність в соціальних мережах також дає змогу збільшити обсяг продажів та залучити увагу до нових продуктів та послуг. Тому,

просування контенту у соціальних мережах є важливою складовою успішної маркетингової стратегії.

### 5.1.1 Аналіз потенційних техніко-економічних переваг ідеї

Оскільки методи якісного просування контенту за вхідними даними є новинкою, то на даний момент немає багато аналогів для вирішення поставленої задачі. Це може стати перевагою для нашого проєкту, оскільки ми зможемо заповнити порожнє місце на ринку та стати інноваційним лідером у цій сфері. Однак, існує один конкурент на ринку – Audiense. [3]

Для того щоб провести якісний порівняльний аналіз необхідно порівняти характеристики систем. Слабкі, сильні та нейтральні характеристики ідеї проєкту визначаються на основі її потенціалу для досягнення поставленої мети, конкурентоздатності на ринку та здатності залучити фінансування. Сильні характеристики ідеї проєкту характеризуються високою потенційною дохідністю, високою конкурентоспроможністю на ринку та можливістю залучення достатнього фінансування. Слабкі характеристики ідеї проєкту, навпаки, можуть зменшити його потенціал успіху, наприклад, недостатній ринковій витривалості, відсутність ключових експертів у команді або недостатня потужність потрібної інфраструктури. Нейтральні характеристики не мають значення для успіху проєкту і можуть бути вирішені під час реалізації проєкту.

Порівняльний аналіз потенційних техніко-економічних переваг ідеї відображено у таблиці 5.2.

Таблиця 5.2 - Визначення слабких, сильних і нейтральних характеристик ідеї проєкту

№ п/п	Техніко-економічні характеристики	Продукти в		Оцінка
		Проект	Audiense	

Продовження таблиці 5.2

1	Зручний UI та функціональний дизайн	+	+	2
2	Контроль доступу до системи	+	-	1
3	Планування завдань до просування контенту	+	-	2
4	Створення цільових аудиторій	+	-	1
5	Робочий процес проекту	-	-	1
6	Швидкий результат	-	-	3

Позначення для оцінки характеристик:

- 1 – сильна характеристика;
- 2 – нейтральна характеристика;
- 3 – слабка характеристика.

З таблиці можна зробити висновок, що програмний продукт, що розроблений у рамках магістерської дисертації, має свої переваги та недоліки, які можна покращити. Треба продовжувати працювати над недоліками для покращення якості продукту.

Незважаючи на це, проєкт має потенціал для успіху за умови подальшого розвитку. Продукт може стати конкурентоспроможним на ринку, якщо будуть вжиті відповідні заходи щодо вдосконалення його функціональності та виправлення недоліків. Результатів таблиці помітно, що розроблений у рамках магістерської дисертації програмний продукт має як і певні переваги, так і недоліки, над якими варто працювати. Незважаючи на це, проєкт має певні шанси на успіх за умови подальшого розвитку.

## 5.2 Технологічний аудит ідеї проєкту

Сучасні технології програмування, які доступні та безкоштовні для всіх користувачів, забезпечують широкий вибір інструментів для вирішення різних завдань. Незважаючи на те, що такі технології є зручними та доступними, вони можуть створювати ілюзію вибору, яка не завжди відповідає потребам користувача.

Для того, щоб зробити обдумане рішення щодо вибору інструменту для вирішення певної задачі, важливо провести аналіз доступних інструментів та технологій розробки. Цей аналіз допоможе знайти оптимальний варіант для вирішення поставленої задачі. Крім того, при виборі технології слід враховувати не лише її властивості, а й майбутні перспективи розвитку та підтримки.

Існує багато підходів та технологій для реалізації застосування, якщо математичне забезпечення зрозуміле і зрозуміле. Спочатку важливо вирішити, який тип застосування буде створений: веб-застосування, мобільний додаток або настільний додаток. Настільний додаток не є оптимальним підходом для створення сучасних додатків у сфері розваг і не так популярний, як мобільні або веб-застосування, тому він не підходить. Хоча мобільний додаток може бути хорошим рішенням, проте мобільні пристрої зазвичай не мають достатньої пам'яті для тестування складних нейронних мереж та моделей. Тому варто обрати веб-застосування, які, якщо є доступ до Інтернету, завжди будуть доступні як для мобільних пристроїв, так і для комп'ютерів.

Надалі треба зробити важливий крок – вибір технології для створення веб-застосування. Технологічна реалізація ідеї проєкту можна описати як можливість реалізації цієї ідеї з використанням наявних технологічних ресурсів та інструментів. Це включає в себе оцінку технічної складності проєкту, наявність необхідних ресурсів для його розробки та реалізації, а також потенційну можливість використання новітніх технологій та розробок

для поліпшення проекту. Технологічна здійсненність також враховує можливі ризики та виклики, пов'язані з розробкою та реалізацією проекту, і потребує оцінки доступності необхідних знань та навичок для розробки проекту. Оцінка технологічної здійсненності є важливою частиною процесу планування та реалізації проекту, яка допомагає забезпечити успішність та ефективність його виконання.

У таблиці 5.3 перераховано основні технології, які можуть бути використані для реалізації системи підтримки фахівця з маркетингу в соціальних мережах..

Таблиця 5.3 - Технологічна реалізація ідеї системи

№ п/п	Ідея реалізації проекту	Технології для реалізації	Наявність технологій	Доступність технологій
1	Реалізація за допомогою фреймворку WPF та Material Design.	C#, WPF, Entity	Усі технології наявні та можуть бути використані при розробці програмного забезпечення.	Технології є безкоштовними та доступними для всіх користувачів.
2	Реалізація за допомогою консольного додатку.	C#.	Усі технології наявні та можуть бути використані при розробці програмного забезпечення.	Технології є безкоштовними та доступними для всіх користувачів.
3	Реалізація за допомогою веб-застосунку	C#, ASP.NET	Усі технології наявні та можуть бути використані	Технології є безкоштовними та доступними

Продовження таблиці 5.3

№ п/п	Ідея реалізації проекту	Технології для реалізації	Наявність технологій	Доступність технологій
3	використання Windows Forms для створення десктоп-застосунку	C#, WinForms	при розробці програмного забезпечення.	для всіх користувачів.

Можна переформулювати так: Хоча існують різні технології для розробки додатків, ми вирішили використовувати фреймворк EntityFramework для нашого сервісу. Це дозволить спростити процес розробки та зосередити увагу на математичних методах, моделях та проведенні експериментів.

### 5.3 Аналіз ринкових можливостей запуску стартап-проекту

Попередньо було встановлено, що основні галузі, де можна використовувати систему підтримки роботи фахівців з маркетингу у соціальних мережах, - це маркетингові агенції та приватний бізнес. Ринок маркетингу є дуже прибутковим та популярним серед різних категорій споживачів. Отже, ми проведемо аналіз можливостей запуску стартап-проекту на цьому ринку.

#### 5.3.1 Аналіз попиту

Можна переформулювати так: На початку розгляду можна провести

аналіз попиту на ринку маркетингу у соціальних мережах у 2023 році. Відомо, що в зв'язку з пандемією попит на онлайн-маркетинг значно зріс.

Це пов'язано з тим, що під час карантину люди стали більше часу проводити в Інтернеті, зокрема в соціальних мережах, тому компанії змушені були перенести свої маркетингові кампанії в онлайн-середовище. Таким чином, ринок маркетингу у соціальних мережах набув ще більшої актуальності та значущості для бізнесу. Тому запуск стартап-проєкту на цьому ринку може бути дуже перспективним та успішним, якщо буде правильно розроблена та реалізована ідея.

У таблиці 5.4 зображено попередню характеристику потенційного ринку стартап-проєкту.

Таблиця 5.4 – Характеристика ринку розваг

№ п/п	Показники стану ринку (найменування)	Характеристика
1	Кількість головних гравців, од	необмежена
2	Загальний обсяг продаж, грн/ум.од	-
3	Динаміка ринку (якісна оцінка)	Зростає
4	Наявність обмежень для входу	-
5	Специфічні вимоги до стандартизації та сертифікації	-
6	Середня норма рентабельності в галузі (або по ринку), %	50-60%

Оцінити точний загальний обсяг продажів на ринку соціального медіа маркетингу є складним завданням через наявність багатьох гравців з різними доходами. Проте можна зазначити, що ринок не має жорстких обмежень для входу, а також не вимагає специфічної стандартизації або сертифікації. Це створює широкі можливості для підприємців та стартапів увійти на ринок зі своїми ідеями та продуктами, що призводить до значної конкуренції.

Однак, щоб виділитись серед інших гравців та привернути увагу потенційних клієнтів, важливо розробити унікальну та конкурентоспроможну стратегію маркетингу.

Необхідно враховувати, що ринок соціального медіа-маркетингу постійно змінюється, тому важливо вести моніторинг тенденцій, аналізувати конкуренцію та адаптувати стратегію маркетингу відповідно до нових умов.

### 5.3.2 Потенційні групи клієнтів

Давайте проаналізуємо потенційні групи клієнтів, які можуть скористатися розробленим сервісом. У таблиці 5.5 надано характеристики цих груп. Оскільки система з підтримки роботи фахівця з маркетингу у соціальних мережах надає можливості з автоматизації процесів роботи з соціальними мережами, то її можуть використовувати різні групи клієнтів, включаючи маркетингові агенції, власників бізнесів, менеджерів з продажу, а також фахівців з соціальних мереж.

Таблиця 5.5 – Характеристика потенційних клієнтів

№ п/п	Цільова аудиторія (цільові сегменти ринку)	Вимоги споживачів до товару
1	Маркетингові агенції	Вони пропонують свої послуги з реклами і продажу в соціальних мережах. Вони можуть скористатися розробленим сервісом для автоматизації роботи з різними соціальними мережами, щоб підвищити ефективність своєї роботи та скоротити час на її виконання.
2	Власники бізнесів та приватні підприємці	Вони можуть використовувати систему з підтримки роботи фахівця з маркетингу для розробки та запуску рекламних кампаній у соціальних мережах. Завдяки автоматизації процесів, вони можуть ефективніше контролювати свої рекламні кампанії та отримувати більше результатів.

## Продовження таблиці 5.5

№ п/п	Цільова аудиторія (цільові сегменти ринку)	Вимоги споживачів до товару
3	Менеджери з продажу послуг	Вони можуть використовувати систему з підтримки роботи фахівця з маркетингу для залучення нових клієнтів, підвищення продажів та контролю над репутацією компанії в соціальних мережах.
4	Фахівці з маркетингу у соціальних мереж.	Вони можуть використовувати розроблений сервіс для підвищення ефективності своєї роботи.

Отже, було виділено чотири основні групи потенційних клієнтів, які можуть скористатись системою з підтримки роботи фахівця з маркетингу у соціальних мережах. Перша група - це малі та середні підприємства, які шукають швидких та ефективних рішень у сфері соціального медіа маркетингу, а також не мають великих бюджетів на рекламу. Друга група - це великі корпорації, які потребують широкого та глибокого аналізу даних з соціальних мереж для вивчення поведінки своїх клієнтів та конкурентів. Третя група - це маркетингові агентства, які надають послуги з соціального медіа маркетингу своїм клієнтам. Четверта група - це незалежні фахівці з маркетингу, які працюють на фріланс-основі та потребують інструментів для ефективної роботи з клієнтами та управління проектами. Основні потреби цих груп клієнтів полягають у вивченні та аналізі даних з соціальних мереж, ефективному плануванні та виконанні маркетингових кампаній, управлінні та моніторингу соціальних мереж, а також взаємодії зі споживачами та конкурентами.

### 5.3.3 Аналіз ринкового середовища

Давайте проведемо аналіз можливих загроз та можливостей, що можуть виникнути на етапі впровадження проекту на ринок. Такий аналіз дозволить

нам зменшити ризики та належним чином використати можливості, які надає проект.

У таблиці 5.6 наведені фактори, які можуть бути загрозами для успішності проекту.

Таблиця 5.6 – Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1	Конкуренція з боку інших застосунків для створення таргетингової реклами.	<p>Такі застосунки можуть бути вже доведені до ринку та мати свою стабільну клієнтську базу. Для того, щоб наш сервіс міг здобути своє місце на ринку, нам потрібно розробити унікальну пропозицію, яка буде відрізнятися від конкурентів. Наприклад, ми можемо надати більше інструментів та функціоналу, який допоможе нашим клієнтам з легкістю створювати та</p>	<p>Можливі варіанти дій включають створення ефективної рекламної кампанії з використанням різних маркетингових каналів для залучення нових користувачів. Також можна розглянути можливість зниження ціни для залучення більшої кількості клієнтів та створення випробувального періоду для нових користувачів з метою залучення їх до використання застосунку та забезпечення позитивного досвіду користування. Для ефективного здійснення цих дій необхідно провести аналіз конкурентів та ринкової ситуації, а також</p>

Продовження таблиці 5.6

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
		<p>аналізувати рекламні кампанії, або можемо запропонувати низькі ціни для нових клієнтів, щоб залучити їх до співпраці з нами. Крім того, ми можемо розвивати нашу маркетингову стратегію, щоб залучати більше клієнтів, збільшувати свою клієнтську базу та збільшувати свою впізнаваність на ринку.</p>	<p>зрозуміти потреби та вимоги цільової аудиторії.</p>
2	<p>Застосунок видає некоректні результати.</p>	<p>Клієнти не задоволені результатами, які вони отримують від застосування системи. Можливо, проблема полягає у тому, що система</p>	<p>Необхідно провести додатковий аналіз та визначити причини некоректної роботи системи. Можливі причини можуть бути пов'язані з недостатньою кількістю даних, помилками в</p>

## Продовження таблиці 5.6

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
		недостатньо ефективно просуває контент або не враховує індивідуальні потреби кожного клієнта.	алгоритмах аналізу, неправильним внесенням даних користувачами або проблемами з інфраструктурою застосунку. Після виявлення причин несправності необхідно внести відповідні зміни та випустити оновлення застосунку для виправлення проблеми.

Фактори можливостей описані в таблиці 5.7.

Таблиця 5.7 – Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1	Конкурентоспроможність за ціною.	Продукт має конкурентну цінову перевагу.	Зберігати низьку ціну і високу якість обслуговування.
2	Проста функціональність.	Продукт має простий та доступний дизайн для будь-якого користувача.	Зберігати зручний підхід взаємодії з клієнтом і забезпечити зручність нового функціоналу.

Таким чином, було розроблено та обґрунтовано основний план дій компанії, який дозволить ефективно реагувати на можливості та загрози.

### Висновок до розділу

У ході роботи над цим розділом було представлено основну ідею нашого стартап-проєкту та описано наявні альтернативи розробки застосування. Також було обґрунтовано наш вибір реалізувати веб-застосування за допомогою фреймворку Entity.

Наш стартап-проєкт спрямований на таргетологів та маркетингові агенції, які мають потребу у такому продукті. Ми визначили основну цільову аудиторію, яка зацікавлена у нашому продукті, і саме вони становлять нашу основну мету.

З метою відповідності потребам цільової групи ми провели аналіз їхніх вимог та встановили функціонал продукту, що задовольняє їх базові потреби. Ми впевнені, що наш продукт надасть їм необхідні інструменти та можливості для успішного здійснення їхніх маркетингових стратегій та кампаній.

## ВИСНОВКИ

У результаті роботи над магістерською дисертацією та під час дослідження було розроблено систему, яка допомагає фахівцям з маркетингу ефективно створювати рекламні кампанії в соціальних мережах. Для досягнення цієї мети були використані різні методи аналізу та прогнозування, зокрема метод кластеризації, метод ARIMA та алгоритм прогнозування. Результатом роботи є інструмент, який може бути корисним для фахівців з маркетингу, що працюють у соціальних мережах.

Під час розробки пояснювальної записки було проведено дослідження різних методів та алгоритмів, які можуть допомогти забезпечити ефективне просування контенту. Було проведено аналіз наявних аналогів та публікацій, з метою визначення найкращих практик в цій галузі. На основі проведеного аналізу було зроблено висновки щодо створення власної системи та вибору оптимальних методів та алгоритмів для її реалізації.

У зв'язку з цим, в наступному розділі роботи було приділено більше уваги детальному дослідженню обраних методів та алгоритмів. Відповідно до визначених висновків, у цьому розділі було розглянуто різні аспекти застосування цих методів та алгоритмів в контексті просування контенту. Також було розглянуто можливості їх використання у власній системі, яка розроблялася в рамках дослідження. Результатом цієї роботи стало створення ефективної системи, яка може бути використана для якісного просування контенту в мережі.

У розділі з дослідження методів та моделей було описано використані метод кластеризації, метод ARIMA та алгоритм прогнозування, їх вміст та принцип роботи. Масштабне завдання було поділено на 5 задач і методи та алгоритми були адаптовані до цих задач.

У розділі, присвяченому реалізації програмного забезпечення, було обґрунтовано вибір засобів розробки, проведено детальний аналіз вимог до сервісу та створено діаграми послідовності, які демонструють процес взаємодії користувача з програмним забезпеченням. В результаті успішно

було реалізовано програмне забезпечення. Крім того, у додатку М були наведені креслення екранних форм веб-застосування.

Отож, було проведено експерименти за результатами яких було зроблено висновки про ефективність обраних методів.

В розділі розробки стартап-проєкту було сформульовано основну концепцію проєкту, надано детальний опис його складових та визначено цільову аудиторію. Крім того, був проведений аналіз ринку маркетингу з метою оцінки ймовірності успішного виходу на нього.

## ПЕРЕЛІК ПОСИЛАНЬ

1. What is Social Media Marketing. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://www.popwebdesign.net/what-is-smm.html>
2. Article about "Social Media Marketing". 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://sendpulse.com/support/glossary/social-media-marketing>
3. «Audiense» review. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://crozdesk.com/software/audiense>
4. Девід Петровскі та Жоао Педро Пестана Нето, Важливість цільової аудиторії. Вибір для ефективності моделі Кано // Хальмстадський університет, Швеція. [Електронний ресурс] – Режим доступу до ресурсу: <https://www.diva-portal.org/smash/get/diva2:1080906/FULLTEXT01.pdf>
5. Христина Ліпяніна-Гончаренко, Тарас Лендюк, Анатолій Саченко та Яцек Волошин, Методика формування контексту реклами та цільової аудиторії на основі навчання асоціативним правилам. [Електронний ресурс] – Режим доступу до ресурсу: <https://ceur-ws.org/Vol-3137/paper9.pdf>
6. Янхоу Луо, Хувей Пан, Шихион Ванг та Юніг Хуанг, Визначення цільової аудиторії в соціальній мережі підприємства, Zhejiang Sci-Tech University, Ханчжоу, Китай. [Електронний ресурс] – Режим доступу до ресурсу: [https://www.researchgate.net/publication/327191605\\_Identifying\\_target\\_audience\\_on\\_enterprise\\_social\\_network](https://www.researchgate.net/publication/327191605_Identifying_target_audience_on_enterprise_social_network)
7. Мухаммед Усман, Створення ефективної реклами для переконання цільової аудиторії. MS Scholar, Університет IQRA, Кампус Ісламабаду. [Електронний ресурс] – Режим доступу до ресурсу: [https://www.ejournalofbusiness.org/archive/vol2no1/vol2no1\\_5.pdf](https://www.ejournalofbusiness.org/archive/vol2no1/vol2no1_5.pdf)
8. Solution of the clustering problem by the method of gradient descent. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://habr.com/ru/articles/188638/>

9. Кластеризація методом k-середніх. 2022. [Електронний ресурс] – Режим доступу до ресурсу: [https://www.wiki-data.uk-ua.nina.az/Кластеризація\\_методом\\_k-середніх.html](https://www.wiki-data.uk-ua.nina.az/Кластеризація_методом_k-середніх.html)
10. Метод ліктя. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://www.helenkapatsa.ru/mietod-loktia/>
11. Що таке показники КРІ (ключові показники ефективності) і навіщо вони потрібні?. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://biznesua.com.ua/shho-take-pokazniki-kpi-klyuchovi-pokazniki-efektivnosti-i-navishho-voni-potribni/>
- Страшний сон менеджера. Що таке КРІ?. 2022. [Електронний ресурс] – Режим доступу до ресурсу: <https://www.dsnews.ua/ukr/economics/strashnyy-son-menedzhera-chtotakoe-kpi-18062021-428687>
- Часові-ряди. [Електронний ресурс] – Режим доступу до ресурсу: <https://kstat.pnu.edu.ua/wpcontent/uploads/sites/63/2018/04/Часові-ряди.pdf>
14. Юрченко М. Є. ПРОГНОЗУВАННЯ ТА АНАЛІЗ ЧАСОВИХ РЯДІВ
15. Марина Євгеніївна Юрченко. ЧЕРНІГІВ ЧНТУ. – 2018. – Електронний ресурс] – Режим доступу до ресурсу: <http://ir.stu.cn.ua/bitstream/handle/123456789/16992/Прогнозув.%20та%20аналіз%20часових%20рядів.pdf?sequence=1&isAllowed>
16. Кластеризація методом k-середніх. – 2021 – [Електронний ресурс] – Режим доступу до ресурсу: [https://www.wiki-data.uk-ua.nina.az/%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D1%96%D1%8F\\_%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D0%BE%D0%BC\\_%D0%BA%E2%80%93%D1%81%D0%B5%D1%80%D0%B5%D0%B4%D0%BD%D1%96%D1%85.html](https://www.wiki-data.uk-ua.nina.az/%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D1%96%D1%8F_%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D0%BE%D0%BC_%D0%BA%E2%80%93%D1%81%D0%B5%D1%80%D0%B5%D0%B4%D0%BD%D1%96%D1%85.html)
17. Градієнтний спуск. – [Електронний ресурс] – Режим доступу до ресурсу:

- [https://uk.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%B4%D1%96%D1%94%D0%BD%D1%82%D0%BD%D0%B8%D0%B9\\_%D1%81%D0%BF%D1%83%D1%81%D0%BA](https://uk.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%B4%D1%96%D1%94%D0%BD%D1%82%D0%BD%D0%B8%D0%B9_%D1%81%D0%BF%D1%83%D1%81%D0%BA)
18. Кваліфікаційна робота на тему “Дослідження методів кластеризації часових рядів за значеннями параметрів”. – Зінченко – 2021 – Харків: Харківський національний університет радіоелектроніки.
  19. Стаття “Планування перевезень транспортними засобами”/ О. Г. Жданова, А. М. Майборода// Матеріали VI всеукраїнської науково-практичної конференції молодих вчених та студентів “Інформаційні системи та технології управління ” (ІСТУ-2021) – Київ.: НТУУ «КПІ ім. Ігоря Сікорського», 22-23 квітня 2021 р.
  20. Дипломний проєкт “Система планування перевезень транспортними засобами”, О. Г. Жданова, А. М. Майборода – 2021 – Київ: НТУУ «КПІ ім. Ігоря Сікорського».
  21. Метод головних компонент. – [Електронний ресурс] – Режим доступу до ресурсу: <http://ebooks.git-elt.hneu.edu.ua/babap/8-4-id8-4.html>
  22. Часовий ряд. – [Електронний ресурс] – Режим доступу до ресурсу: [https://www.wiki-data.uk-ua.nina.az/%D0%A7%D0%B0%D1%81%D0%BE%D0%B2%D0%B8%D0%B9\\_%D1%80%D1%8F%D0%B4.html](https://www.wiki-data.uk-ua.nina.az/%D0%A7%D0%B0%D1%81%D0%BE%D0%B2%D0%B8%D0%B9_%D1%80%D1%8F%D0%B4.html)
  23. Key Performance Indicators, KPI. – [Електронний ресурс] – Режим доступу до ресурсу: <https://www.it.ua/knowledge-base/technology-innovation/key-performance-indicators-kpi>
  24. Visual Studio. – [Електронний ресурс] – Режим доступу до ресурсу: [https://znaimo.com.ua/Visual\\_Studiol](https://znaimo.com.ua/Visual_Studiol)
  25. Реферат “ Розробка автоматизованого робочого місця адміністратора кінотеатру ” [Електронний ресурс] – Режим доступу до ресурсу: <https://ukrbukva.net/page,3,100956-Razrobotka-avtomatizirovannogo-rabocheho-mesta-administratora-kinoteatra-Mir.html>

26. Дипломний проєкт “ Розробка програмного агента моніторингу та управління сонячної електричної станції”. Задачин С. С., Мірошніченко І.В. – 2019 – Київ: НТУУ «КПІ ім. Ігоря Сікорського».
27. Microsoft SQL Server. – [Електронний ресурс] – Режим доступу до ресурсу: [https://www.wiki-data.uk-ua.nina.az/Microsoft\\_SQL\\_Server.html](https://www.wiki-data.uk-ua.nina.az/Microsoft_SQL_Server.html)
28. Публічна діаграма “SQL Server”. – [Електронний ресурс] – Режим доступу до ресурсу: [https://coggle.it/diagram/W\\_6Kssk8l\\_6v51Qb/t/sql-server](https://coggle.it/diagram/W_6Kssk8l_6v51Qb/t/sql-server)
29. Microsoft SQL Server. – [Електронний ресурс] – Режим доступу до ресурсу: [https://uk.wikipedia.org/wiki/Microsoft\\_SQL\\_Server](https://uk.wikipedia.org/wiki/Microsoft_SQL_Server)
30. Стаття “Система керування версіями”. – [Електронний ресурс] – Режим доступу до ресурсу: [https://znaimo.com.ua/Система\\_керування\\_версіями](https://znaimo.com.ua/Система_керування_версіями)
31. Стаття “ Розподілені системи управління версіями”. – [Електронний ресурс] – Режим доступу до ресурсу: [https://studopedia.com.ua/1\\_174094\\_rozpodileni-sistemi-upravlinnya-versiyami.html](https://studopedia.com.ua/1_174094_rozpodileni-sistemi-upravlinnya-versiyami.html)
32. Що таке Material Design і чому він такий популярний. – 2022 – [Електронний ресурс] – Режим доступу до ресурсу: <https://timeweb.com/ru/community/articles/что-такое-material-design-i-pochemu-on-tak-populyaren>
33. Які реальні заслуги у Material Design? – 2015 – [Електронний ресурс] – Режим доступу до ресурсу: <https://prodesign.in.ua/2015/11/yaki-realni-zasluchy-u-material-design/>
34. Що таке Material Design від Google і чому про нього всі говорять. – 2021 – [Електронний ресурс] – Режим доступу до ресурсу: <https://androidinsider.ru/eto-interesno/что-такое-material-design-ot-google-i-pochemu-o-nem-vse-govoryat.html>
35. Використання розширення Visual Studio Code. – 2021 – [Електронний ресурс] – Режим доступу до ресурсу: <https://learn.microsoft.com/uk-ua/power-apps/maker/portals/vs-code-extension>

36. Startup. – 2022 – [Електронний ресурс] – Режим доступу до ресурсу:  
<https://sendpulse.com/support/glossary/startup>
37. Тези довідей “Сегментування споживачів для якісного маркетингу у соціальних мережах”. – М. О. Солдатова, А. М. Майборода// Матеріали ІV міжнародної конференції «Інженерія програмного забезпечення і передові інформаційні технології (SoftTech-2023)» – м. Київ, 9-11 травня 2023 р.