

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

**Факультет прикладної математики  
Кафедра системного програмування і  
спеціалізованих комп'ютерних систем**

«На правах рукопису»  
УДК 004.4

До захисту допущено:  
Завідувач кафедри  
\_\_\_\_\_ Віталій РОМАНКЕВИЧ  
«\_\_» \_\_\_\_\_ 2024 р.

**Магістерська дисертація**

**на здобуття ступеня магістра**

**за освітньо-професійною програмою «Системне програмування та  
спеціалізовані комп'ютерні системим»**

**зі спеціальності 123 «Комп'ютерна інженерія»**

**на тему: «Спосіб оцінки відповідності кандидата до вакансії на базі засобів  
машинного навчання»**

Виконав:

студент II курсу, групи КВ-21 мп  
ГЕЙ Антон Віталійович \_\_\_\_\_

Науковий керівник:

Д.т.н., професор,  
РОМАНКЕВИЧ Олексій \_\_\_\_\_

Рецензент:

Старший викладач кафедри ОТ ФІОТ,  
ВИНОГРАДОВ Юрій \_\_\_\_\_

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

Студент \_\_\_\_\_

Київ – 2024 року

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**  
**Факультет прикладної математики**

**Кафедра системного програмування і спеціалізованих комп'ютерних систем**  
Рівень вищої освіти – другий (магістерський)  
Спеціальність – 123 «Комп'ютерна інженерія»  
Освітньо-професійна програма «Системне програмування та спеціалізовані комп'ютерні системи»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Віталій РОМАНКЕВИЧ

« \_\_\_\_ » \_\_\_\_\_ 2024 р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**  
**Гею Антону Віталійовичу**

1. Тема дисертації «Спосіб оцінки відповідності кандидата до вакансії на базі засобів машинного навчання», науковий керівник дисертації Романкевич, Олексій Михайлович, професор кафедри СПіСКС, докт. техн. наук, затверджені наказом по університету від «9» листопада 2023 р. №5217-с
2. Термін подання студентом дисертації 10 січня 2024
3. Об'єкт дослідження - процес оцінки кандидатів, , зокрема, механізми і методи автоматизованого визначення їх відповідності зазначеним вимогам роботодавця.
4. Вихідні дані
5. Перелік завдань, які потрібно розробити: дослідження та оцінка існуючих методів та інструментів, які можуть бути застосовані для аналізу текстів резюме та описів вакансій, з особливим акцентом на векторні представлення тексту та методи визначення схожості. Створення алгоритмів для оцінки ключових характеристик кандидатів, включаючи їхні навички, освіту та досвід, та адаптація цих методів для забезпечення високої точності та об'єктивності оцінок.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу - презентація
7. Орієнтовний перелік публікацій - 2 Тези конференцій
8. Дата видачі завдання 1 листопада 2022р.

#### Календарний план

| № з/п | Назва етапів виконання магістерської дисертації                                                                                                                | Термін виконання етапів магістерської дисертації | Примітка |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|----------|
| 1     | Формування мети та цілі роботи                                                                                                                                 | 24.10.2022                                       |          |
| 2     | Дослідження теоретичного матеріалу                                                                                                                             | 28.12.2022                                       |          |
| 3     | Дослідження існуючих рішень                                                                                                                                    | 19.01.2023                                       |          |
| 4     | Підготовка матеріалів першого розділу дипломного проекту                                                                                                       | 12.03.2023                                       |          |
| 5     | Проведення наукового дослідження; робота над другим розділом магістерської дисертації                                                                          | 03.05.2023                                       |          |
| 6     | Робота над третім розділом магістерської дисертації                                                                                                            | 14.09.2023                                       |          |
| 8     | Підготовка матеріалів четвертого розділу магістерської дисертації                                                                                              | 30.10.2023                                       |          |
| 9     | Завершення роботи над основною частиною магістерської дисертації; підготовка ілюстративного матеріалу; підготовка матеріалів доповіді на конференції ПМК '2023 | 27.11.2023                                       |          |
| 10    | Оформлення текстової і графічної частини магістерської дисертації                                                                                              | 15.12.2023                                       |          |
| 11    | Попередній розгляд МД на кафедрі                                                                                                                               | 20.12.2023                                       |          |

Студент \_\_\_\_\_

Антон ГЕЙ

Науковий керівник \_\_\_\_\_

Олексій РОМАНКЕВИЧ

## РЕФЕРАТ

**Актуальність теми.** Оцінка відповідності кандидата до вакансії є критичним етапом у процесі відбору кваліфікованих співробітників для організацій різних галузей та сфер діяльності. Вірне визначення підходящих кандидатів допомагає підвищити продуктивність роботи, покращити показники компанії і забезпечити ефективний розвиток. У цьому контексті, спосіб оцінки відповідності кандидата до вакансії визначається як один із ключових аспектів успішного підбору персоналу.

*Об'єктом дослідження* є процес оцінки кандидатів, , зокрема, механізми і методи автоматизованого визначення їх відповідності зазначеним вимогам роботодавця.

*Предметом дослідження* є алгоритми та методи обробки природної мови, а також інтеграція зі засобами машинного навчання, що використовуються для аналізу текстів резюме та описів вакансій з метою оцінки відповідності кандидатів.

*Методи дослідження* включають аналіз та застосування технологій обробки тексту, алгоритмів машинного навчання, векторних представлень тексту та методів оцінки схожості.

**Мета роботи** полягає у створенні системи для оцінки відповідності кандидатів до вакансій, яка базується на предметному аналізі природної мови та використанні засобів машинного навчання. Ця система покликана підвищити точність, швидкість та об'єктивність процесу відбору, враховуючи специфіку різноманітних вакансій та індивідуальні особливості кандидатів. Для досягнення цієї мети в роботі передбачено вирішення наступних задач:

1. Дослідження та оцінка існуючих методів та інструментів, які можуть бути застосовані для аналізу текстів резюме та описів вакансій, з особливим акцентом на векторні представлення тексту та методи визначення схожості.

2. Створення алгоритмів для оцінки ключових характеристик кандидатів, включаючи їхні навички, освіту та досвід, та адаптація цих методів для забезпечення високої точності та об'єктивності оцінок.

**Наукова новизна** полягає в наступному:

1. Формалізація задачі автоматизації аналізу резюме та описів вакансій на базі алгоритмів машинного навчання та обробки природної мови (NLP), що сприяє швидкому та точному визначенню відповідності кандидатів до вакансії.
2. Алгоритми оцінки відповідності ключових характеристик кандидатів до вакансії, включаючи їхні навички, освіту та досвід.

**Практична цінність** цієї роботи полягає в полегшенні та покращенні процесу відбору кандидатів для вакансій у підприємствах та організаціях. Розроблений модифікований підхід дозволяє автоматично аналізувати резюме кандидатів та адаптувати до вимог конкретної вакансії, що полегшує та раціоналізує процес відбору персоналу.

**Апробація роботи.** Основні положення і результати роботи були представлені та обговорювались на:

- XVI науковій конференції магістрантів та аспірантів «ПРИКЛАДНА МАТЕМАТИКА ТА КОМП'ЮТИНГ – ПМК '2023»
- СХХХV Міжнародній науково-практичній конференції "НАУКОВІ ПІДСУМКИ 2023"

**Структура та обсяг роботи.** Магістерська дисертація складається з вступу, чотирьох розділів та висновків.

У *вступі* подана загальна характеристика роботи, проведена оцінка сучасного стану проблеми, обґрунтована актуальність напрямку досліджень, сформульовані мета та завдання дослідження. Також надано відомості про апробацію результатів та їхнє впровадження.

У *першому розділі* проведено теоретичний аналіз проблеми оцінки відповідності кандидата до вакансії, розглянуто основні методи та технології

цього процесу, а також визначено роль машинного навчання у вирішенні даної задачі.

У *другому розділі* визначено цільові критерії та метрики, які використовуються для оцінки відповідності кандидатів до вакансій. Також обгрунтовано вибір конкретних алгоритмів та моделей для вирішення цієї задачі.

У *третьому розділі* описано реалізацію системи оцінки відповідності кандидатів до вакансій, включаючи вибір технологій та програмного забезпечення, підготовку та обробку даних та аналіз результатів.

У *четвертому розділі* проведено валідацію та тестування розробленого програмного забезпечення, а також оцінку його практичної вартості та ефективності в реальних умовах.

У *висновках* подані підсумки дослідження, висновки та рекомендації щодо подальших досліджень у даній області.

**Ключові слова:** метрики оцінки кандидатів, машинне навчання, демонстраційна плата, вагова оцінка характеристик кандидата

## ABSTRACT

**Relevance of the topic.** Assessing the suitability of a candidate for a vacancy is a critical step in the process of selecting qualified employees for organizations in various industries and fields of activity. The correct identification of suitable candidates helps to increase productivity, improve company performance and ensure effective development. In this context, the method of assessing the candidate's suitability for a vacancy is defined as one of the key aspects of successful recruitment.

*The object of the study* is the process of evaluating candidates, in particular, mechanisms and methods for automated determination of their compliance with the specified requirements of the employer.

*The subject of the study* is algorithms and methods of natural language processing, as well as integration with machine learning tools used to analyze resume texts and job descriptions to assess the suitability of candidates.

*Research methods* include the analysis and application of text processing technologies, machine learning algorithms, vector representations of text, and similarity assessment methods.

**The scientific novelty** is as follows:

1. Formalization of the task of automating the analysis of resumes and job descriptions based on machine learning and natural language processing (NLP) algorithms, which helps to quickly and accurately determine the suitability of candidates for a job.
2. Algorithms for assessing the relevance of key characteristics of candidates to a vacancy, including their skills, education, and experience.

**The practical value** of this work is to facilitate and improve the process of selecting candidates for vacancies in enterprises and organizations. The developed modified approach allows to automatically analyze candidates' resumes and adapt them to the requirements of a particular vacancy, which facilitates and rationalizes the personnel selection process.

**Approbation of work.** The main provisions and results of the work were presented and discussed at:

1. XVI Scientific Conference of Undergraduate and Postgraduate Students "APPLIED MATHEMATICS AND COMPUTING - PMC '2023"
2. CXXXV International Scientific and Practical Conference "SCIENTIFIC RESULTS 2023"

**Structure and scope of the work.** The master's thesis consists of an introduction, four chapters and conclusions.

*The introduction* provides a general description of the work, assesses the current state of the problem, substantiates the relevance of the research area, and formulates the purpose and objectives of the study. It also provides information on the testing of the results and their implementation.

*The first section* conducts a theoretical analysis of the problem of assessing the suitability of a candidate for a vacancy, considers the main methods and technologies of this process, and defines the role of machine learning in solving this problem.

*The second section* defines the target criteria and metrics used to assess the suitability of candidates for vacancies. It also justifies the choice of specific algorithms and models for solving this task.

*The third section* describes the implementation of the system for assessing the suitability of candidates for vacancies, including the selection of technologies and software, data preparation and processing, and analysis of the results.

*The fourth section* validates and tests the developed software, as well as evaluates its practical value and effectiveness in real conditions.

*The conclusion* presents the results of the study, conclusions and recommendations for further research in this area.

**Keywords:** candidate evaluation metrics, machine learning, demo board, weighting of candidate characteristics

## ЗМІСТ

|                                                                                 |    |
|---------------------------------------------------------------------------------|----|
| ЗМІСТ                                                                           | 9  |
| ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ                                  | 11 |
| ВСТУП                                                                           | 12 |
| 1. ТЕОРЕТИЧНИЙ ОГЛЯД ПРОБЛЕМИ ОЦІНКИ РЕЛЕВАНТНОСТІ КАНДИДАТА ДО ВАКАНСІЇ        | 13 |
| 1.1 Сутність оцінки відповідності кандидата до вакансії.                        | 13 |
| 1.2 Методи і технології оцінки відповідності кандидата до вакансії.             | 15 |
| 1.2.1 Традиційний скрінінг резюме (CV)                                          | 16 |
| 1.2.2 ATS (Applicant Tracking Systems)                                          | 16 |
| 1.3 Машинне навчання в оцінці відповідності кандидата до вакансії.              | 18 |
| 1.3.1 Огляд сучасних методів машинного навчання                                 | 18 |
| 1.3.2 Ефективність алгоритмів машинного навчання в рекрутингу                   | 21 |
| 1.4 NLP в оцінці відповідності кандидата до вакансії.                           | 22 |
| 1.4.1 Техніки NLP у контексті аналізу резюме                                    | 23 |
| 1.4.2 Використання NLP для розуміння семантики резюме                           | 24 |
| 1.4.3 Порівняльний аналіз основних моделей NLP                                  | 25 |
| 1.5 Постановка задачі                                                           | 26 |
| 1.6 Висновки до розділу 1                                                       | 28 |
| 2. МЕТОДИ ОЦІНКИ ВІДПОВІДНОСТІ КАНДИДАТА ДО ВАКАНСІЇ                            | 30 |
| 2.1. Обґрунтування вибору напряму досліджень                                    | 30 |
| 2.2 Формулювання цільової функції та метрик оцінки                              | 33 |
| 2.3 Вибір конкретних алгоритмів та моделей та його обґрунтування.               | 35 |
| 2.3.1. Семантичне векторне представлення тексту                                 | 35 |
| 2.3.2. Алгоритми оцінки                                                         | 36 |
| 2.3.2.1 Оцінка навичок                                                          | 37 |
| 2.3.2.2 Оцінка освіти                                                           | 37 |
| 2.3.2.3 Оцінка досвіду                                                          | 38 |
| 2.3.3. Вагові коефіцієнти                                                       | 38 |
| 2.3.4. Порівняльний аналіз з іншими методами                                    | 39 |
| 2.4 Висновки до розділу 2                                                       | 44 |
| 3. РЕАЛІЗАЦІЯ СПОСОБУ ОЦІНКИ ВІДПОВІДНОСТІ КАНДИДАТА ДО ВАКАНСІЇ                | 46 |
| 3.1 Вибір технологій та засобів реалізації системи оцінки відповідності.        | 46 |
| 3.2 Програмне забезпечення для оцінки відповідності кандидатів до вакансій.     | 50 |
| 3.3 Підготовка та обробка даних для оцінки ефективності системи.                | 54 |
| 3.4 Аналіз та порівняння результатів з традиційними методами відбору персоналу. | 58 |
| 3.5 Висновки до розділу 3                                                       | 60 |
| 4. ВАЛІДАЦІЯ ТА ОЦІНКА РОЗРОБЛЕНОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ                    | 62 |
| 4.1 Валідація та тестування розробленого програмного забезпечення.              | 62 |

|                                                                                                                     |    |
|---------------------------------------------------------------------------------------------------------------------|----|
| 4.2 Оцінка практичної вартості та ефективності в контексті реальних використань. (переваги запропонованого способу) | 71 |
| 4.3 Плани щодо покращення існуючого функціоналу                                                                     | 73 |
| 4.4 Висновки до розділу 4                                                                                           | 76 |
| ВИСНОВКИ                                                                                                            | 78 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ                                                                                          | 80 |
| ДОДАТКИ                                                                                                             | 81 |

## ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ

NLP - Обробка природної мови (Natural Language Processing)

API - Програмний інтерфейс застосування (Application Programming Interface)

ML - Машинне навчання (Machine Learning)

HR - Управління персоналом (Human Resources)

TF-IDF - Частота термінів - зворотна частота документів (Term Frequency-Inverse Document Frequency)

BERT - Бідерекційні трансформери з упередженнями кодера (Bidirectional Encoder Representations from Transformers)

Word2Vec - Модель векторного представлення слів (Word to Vector)

JSON - Об'єктна нотація JavaScript (JavaScript Object Notation)

UI - Інтерфейс користувача (User Interface)

cosine\_similarity - Косинусна схожість (Cosine Similarity)

Scikit-learn - Бібліотека для машинного навчання в Python

NumPy - Бібліотека для наукових обчислень в Python

Pandas - Бібліотека для аналізу даних в Python

HR - (Human Resource)

CV - Резюме кандидата (Curriculum Vitae)

JD - Опис вакансії (Job Description)

ATS - Система відстеження кандидатів (Applicant Tracking System)

CSV - формат файлу, який використовується для зберігання табличних даних (Comma-Separated Values)

SQL - мова програмування для управління даними у системах управління базами даних (Structured Query Language)

## ВСТУП

На сьогоднішньому конкурентному ринку праці процес визначення найбільш підходящих кандидатів на вакансії стає дедалі складнішим. Традиційні методи відбору персоналу часто покладаються на оцінку вручну, яка може займати багато часу та бути суб'єктивною. Проте технологічний прогрес, зокрема у сфері штучного інтелекту та машинного навчання, відкриває нові можливості для підвищення ефективності та результативності оцінювання кандидатів.

Актуальність дослідження полягає у вирішенні критичних потреб у сфері людських ресурсів і підбору персоналу. Оскільки організації ростуть і диверсифікуються, завдання виявлення ідеального кандидата з пулу претендентів стає складнішим. З ростом ринку праці та збільшенням кількості кандидатів, автоматизація процесу оцінювання резюме дозволяє компаніям швидше та ефективніше відбирати потенційних співробітників. З використанням технологій обробки великих даних та штучного інтелекту, таке програмного забезпечення може аналізувати резюме більш глибоко, виявляючи складні взаємозв'язки та компетенції кандидатів. Автоматизація процесу оцінювання резюме може значно знизити час та витрати, пов'язані з наймом нових співробітників, особливо в великих компаніях.

## 1. ТЕОРЕТИЧНИЙ ОГЛЯД ПРОБЛЕМИ ОЦІНКИ РЕЛЕВАНТНОСТІ КАНДИДАТА ДО ВАКАНСІЇ

### 1.1 Сутність оцінки відповідності кандидата до вакансії.

Сутність оцінки відповідності кандидата до вакансії полягає в систематичному та об'єктивному аналізі кожного кандидата на підставі різноманітних метрик, щоб визначити, наскільки він відповідає вимогам, поставленим для певної спеціальності або вакансії. Оцінка відповідності є важливою складовою процесу підбору персоналу і сприяє ефективному відбору найкращих кандидатів для вакансій, що забезпечує успішність підприємства та оптимізацію його діяльності.

Першою ключовою складовою оцінки є досвід роботи кандидата. Досвід визначається як кількістю років роботи у сфері веб-розробки, а також типом і рівнем попередніх робочих місць. Кандидати з більшим досвідом в цій галузі мають більше шансів відповідати вимогам, оскільки вони можуть приносити більше знань та експертизи.

Другою важливою метрикою є освіта. Хоча досвід є важливим, наявність вищої освіти в галузі комп'ютерних наук або веб-розробки також може бути важливим показником кваліфікації кандидата. Освіта може свідчити про те, що кандидат має теоретичну базу та розуміння основних концепцій, які важливі для вакансії.

Третьою ваговою складовою є стек технологій та рівень їх володіння. Фронт-енд розробники повинні бути знайомі з різними мовами програмування, фреймворками та іншими інструментами. Оцінка стеку технологій полягає в аналізі, чи має кандидат необхідні навички для виконання завдань на вакансії, і наскільки він володіє ними.

Додатковою складовою може бути тип попередньої компанії, де працював кандидат. Наприклад, досвід у великій технологічній компанії може бути більш цінним, оскільки він може свідчити про те, що кандидат мав можливість працювати над складними проектами та вчитися від досвідчених фахівців.

Додатковими важливими метриками для оцінки відповідності кандидата до вакансії можуть бути знання англійської мови та здатність працювати в команді.

*Наприклад:*

1. Знання англійської мови: у світі інформаційних технологій англійська мова часто є мовою комунікації та документації. Кандидати з відмінними знаннями англійської мови можуть бути більш корисними для компанії, яка веде міжнародний бізнес або співпрацює з клієнтами та партнерами з інших країн. Оцінка знань англійської мови може включати тестування навичок усного та писемного спілкування, а також здатність розуміти технічну лексику.
2. Робота в команді: у сучасному ІТ-середовищі робота в команді є надзвичайно важливою. Команди розробників спільно працюють над проектами, і здатність спілкуватися, співпрацювати та вирішувати проблеми в команді є ключовою. Оцінка цієї навички може включати інтерв'ю або тестування, які спрямовані на визначення здатності кандидата працювати у групі, його лідерські якості, вміння приймати рішення та вирішувати конфлікти.

*Крім цього, інші можливі метрики можуть включати:*

1. Професійні сертифікати та курси: кандидати, які продемонструють наявність актуальних сертифікатів та успішно пройдені курси з веб-розробки, можуть вважатися більш кваліфікованими.
2. Портфоліо проектів: Аналіз робіт, які кандидат вже виконував у минулому, може допомогти визначити його рівень навичок та здатність до розробки веб-додатків.
3. Здатність до вирішення проблем: тестування кандидата на здатність розв'язувати складні завдання та аналізувати проблеми може бути корисним для оцінки його аналітичних навичок.

Ефективна оцінка враховує всі аспекти, які є важливими для успішної роботи кандидата на вакансії та сприяє вибору найкращих претендентів. Вага кожної з цих метрик може різнитися в залежності від конкретних вимог вакансії та

стратегії підбору персоналу компанії. Всі ці складові мають бути враховані при розробці системи оцінки відповідності кандидатів до вакансій, щоб забезпечити ефективний та об'єктивний процес відбору персоналу.

## 1.2 Методи і технології оцінки відповідності кандидата до вакансії.

У динамічному світі сучасного ринку праці, де кожна вакансія приваблює велику кількість кандидатів, а організації прагнуть до максимальної ефективності та якості найму, виникає гостра потреба в удосконаленні методів оцінки відповідності кандидатів. Це спонукає до пошуку нових підходів та розробки інноваційних технологій, здатних забезпечити точність, швидкість та об'єктивність в процесах відбору.

Сучасні технології оцінки відповідності кандидатів до вакансії базуються на широкому спектрі методів та інструментів, серед яких особливе місце займають алгоритми машинного навчання та семантичний аналіз тексту. Ці методи дозволяють обробляти великі обсяги даних, виділяючи значущі характеристики та шаблони, що сприяє підвищенню точності та ефективності оцінки кандидатів.

З одного боку, існують традиційні методи скрінінгу резюме, такі як ручний огляд та системи відстеження заявників (ATS). Ці методи, хоча й забезпечують певний рівень контролю та можливість для глибокого аналізу резюме, часто вимагають значних часових витрат та можуть бути суб'єктивними. Крім того, вони не завжди здатні ефективно обробляти великі обсяги даних, що є критично важливим в умовах великої кількості кандидатів.

З іншого боку, новітні підходи, засновані на машинному навчанні та комплексному аналізі даних, пропонують більш швидке та об'єктивне вирішення проблеми. Вони дозволяють автоматизувати процеси оцінки, знижуючи вплив людського фактора та підвищуючи швидкість обробки резюме. Застосування алгоритмів машинного навчання може допомогти ідентифікувати ключові навички та властивості кандидатів, оцінювати їх відповідність критеріям вакансії та ранжувати кандидатів за рівнем релевантності.

Проте, обрання між традиційними та сучасними методами вимагає зваженого підходу, оскільки кожен з них має свої переваги та недоліки. При виборі важливо враховувати специфіку вакансії, кількість кандидатів, доступні ресурси для найму, а також потребу в швидкості та точності процесу.

Надалі буде представлено детальний аналіз наявних методів та підходів до оцінки потенційних кандидатів, що використовуються у сфері рекрутингу.

### 1.2.1 Традиційний скрінінг резюме (CV)

Традиційний скрінінг резюме (CV) є одним із найпоширеніших методів відбору кандидатів у сфері рекрутингу. Цей процес включає ручний огляд резюме з боку рекрутерів або HR-менеджерів, які аналізують досвід роботи, освіту, навички та інші важливі критерії, щоб визначити потенційну відповідність кандидата вакансії.

Переваги традиційного скрінінгу резюме:

1. Глибоке розуміння контексту: Рекрутери можуть звернути увагу на специфічні деталі та контекст, який може бути неочевидний для автоматизованих систем. Це включає здатність оцінювати не тільки фактичні дані, а й потенціал, мотивацію, індивідуальні досягнення, які можуть бути представлені у нетиповій формі.
2. Виявлення унікальних аспектів: Ручний огляд дозволяє виявити унікальні особливості кандидата, такі як особливі досягнення, незвичайний кар'єрний шлях, або особливі проекти та досвід, які можуть бути важливими для певної ролі або компанії.

Недоліки традиційного скрінінгу резюме:

1. Часовитратність: Оцінка кожного резюме вручну є дуже часозатратною, особливо коли кількість кандидатів велика. Це обмежує кількість резюме, які можна ефективно переглянути, та збільшує час, необхідний для відбору кандидатів.
2. Підвищена ймовірність помилок: Людський фактор може призвести до помилок у процесі оцінки. Втома, упередженість, або просто помилки у судженні можуть вплинути на якість та об'єктивність відбору.
3. Суб'єктивність та вплив особистих уподобань: Рішення рекрутерів можуть бути під впливом їх особистих уподобань, стереотипів або упереджень, що може призводити до неконсистентності та недоцільності відбору.

Розуміння цих переваг і недоліків є важливим для визначення, коли і як найкраще застосовувати традиційний скрінінг резюме, а також для розробки стратегій його оптимізації та доповнення іншими методами. Особливо це стосується

### 1.2.2 ATS (Applicant Tracking Systems)

ATS (Applicant Tracking Systems) є широко використовуваними інструментами в сучасному рекрутингу, спрямованими на автоматизацію та оптимізацію процесу відбору персоналу. Ці системи надають можливість роботодавцям ефективно керувати великими обсягами заявок та резюме, спрощуючи пошук та відбір потенційних кандидатів.

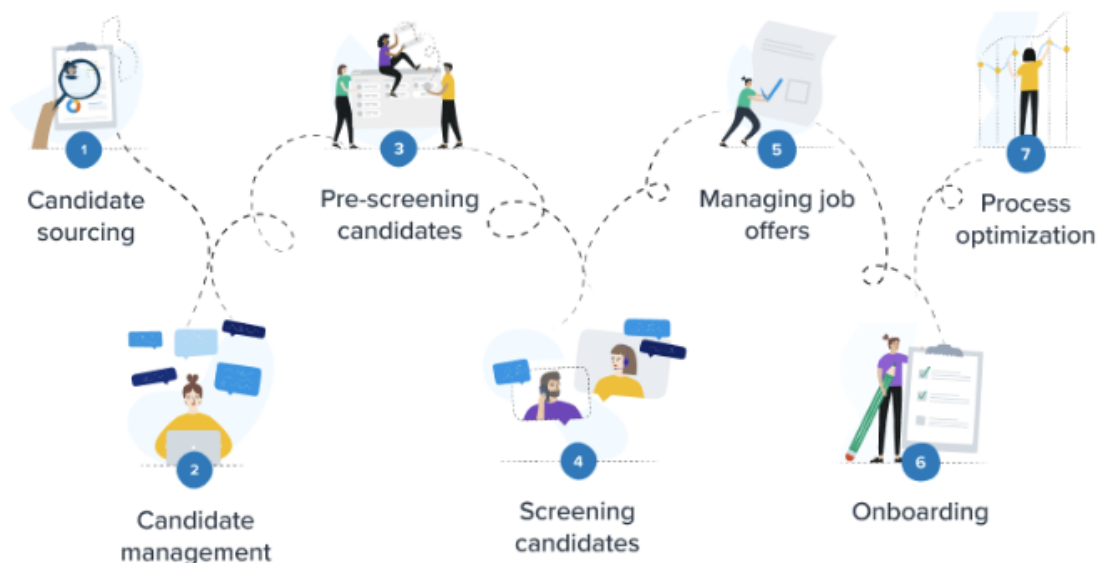


Рисунок 1 - Як працює ATS у процесі рекрутингу

#### Переваги ATS:

1. **Ефективність обробки**: ATS здатні швидко обробляти великі обсяги резюме, значно скорочуючи час, необхідний для первісного огляду кандидатів. Це особливо цінно у великих організаціях з високим об'ємом вакансій.
2. **Мінімізація людського фактора**: Автоматизація дозволяє знизити ймовірність помилок, пов'язаних з ручним оглядом резюме, а також забезпечує більшу консистентність та швидкість процесу скрінінгу.
3. **Структурування процесу відбору**: ATS надає інструменти для управління кандидатами, відстеження їх статусу на різних етапах відбору, а також аналізу ефективності рекрутингових кампаній.

#### Недоліки ATS:

1. **Відсутність глибини аналізу**: Більшість ATS зосереджені на пошуку ключових слів та фраз у резюме, що може не враховувати контекст та унікальні аспекти профілю кандидата. Це може призвести до втрати потенційно підходящих кандидатів, чий досвід або навички не відповідають строго заданим критеріям.
2. **Ризик відхилення кваліфікованих кандидатів**: Недосконалість алгоритмів та залежність від ключових слів можуть призвести до того, що кваліфіковані кандидати будуть помилково відфільтровані системою через неповну або неточну відповідність критеріям скрінінгу.
3. **Залежність від якості вхідних даних**: Ефективність ATS сильно залежить від якості та структури вхідних резюме. Нестандартні або погано структуровані резюме можуть неправильно оброблятися системою, що призводить до помилкових результатів.

З огляду на ці фактори, при виборі ATS для оцінки кандидатів необхідно звертати увагу на можливість налаштування критеріїв відбору, інтеграцію з іншими HR-системами, а також на гнучкість системи щодо обробки різноманітних типів даних. Крім того, важливою є наявність механізмів для перегляду та коригування результатів, щоб мінімізувати ризики пов'язані з автоматичним скрінінгом.

### 1.3 Машинне навчання в оцінці відповідності кандидата до вакансії.

У дослідженні сфери рекрутингу та оцінки кандидатів, застосування машинного навчання відкриває нові можливості для оптимізації та підвищення ефективності процесів відбору. Методи машинного навчання дозволяють автоматизувати та удосконалити аналіз великих обсягів даних, забезпечуючи більш точну та об'єктивну оцінку відповідності кандидатів до конкретних вакансій. Цей підрозділ зосереджується на огляді та аналізі різноманітних алгоритмів та моделей машинного навчання, їх застосуванні в контексті рекрутингу, а також на вивченні методів семантичного аналізу та обробки природної мови.

Використання сучасних технологій у цьому контексті дозволяє розширити можливості рекрутерів, зменшити час та витрати на пошук та оцінку кандидатів, а також підвищити якість процесу найму завдяки більш глибокому і всебічному розумінню даних. Оцінка ефективності методів машинного навчання, розгляд їх переваг та викликів, а також аналіз перспектив розвитку та можливих напрямків удосконалення забезпечують ґрунтовну основу для розуміння потенціалу цих технологій у сфері рекрутингу.

#### 1.3.1 Огляд сучасних методів машинного навчання

Сучасний світ машинного навчання пропонує широкий спектр алгоритмів та моделей, які здатні ефективно обробляти та аналізувати великі обсяги даних. У контексті оцінки відповідності кандидатів до вакансій, ці технології відіграють ключову роль у виявленні та прогнозуванні потенційно успішних працівників. Розглянемо декілька основних методів машинного навчання, які застосовуються в цій сфері.

Дерева рішень є одним із найпоширеніших інструментів у сфері машинного навчання для класифікації та регресії. Вони моделюють рішення і їх можливі наслідки, імітуючи логічне розділення, яке може бути інтуїтивно зрозумілим під час інтерпретації результатів. Окрім власне класифікації кандидатів, ці алгоритми

ефективні у визначенні найбільш важливих факторів для прогнозування успішності кандидата на певній роботі.

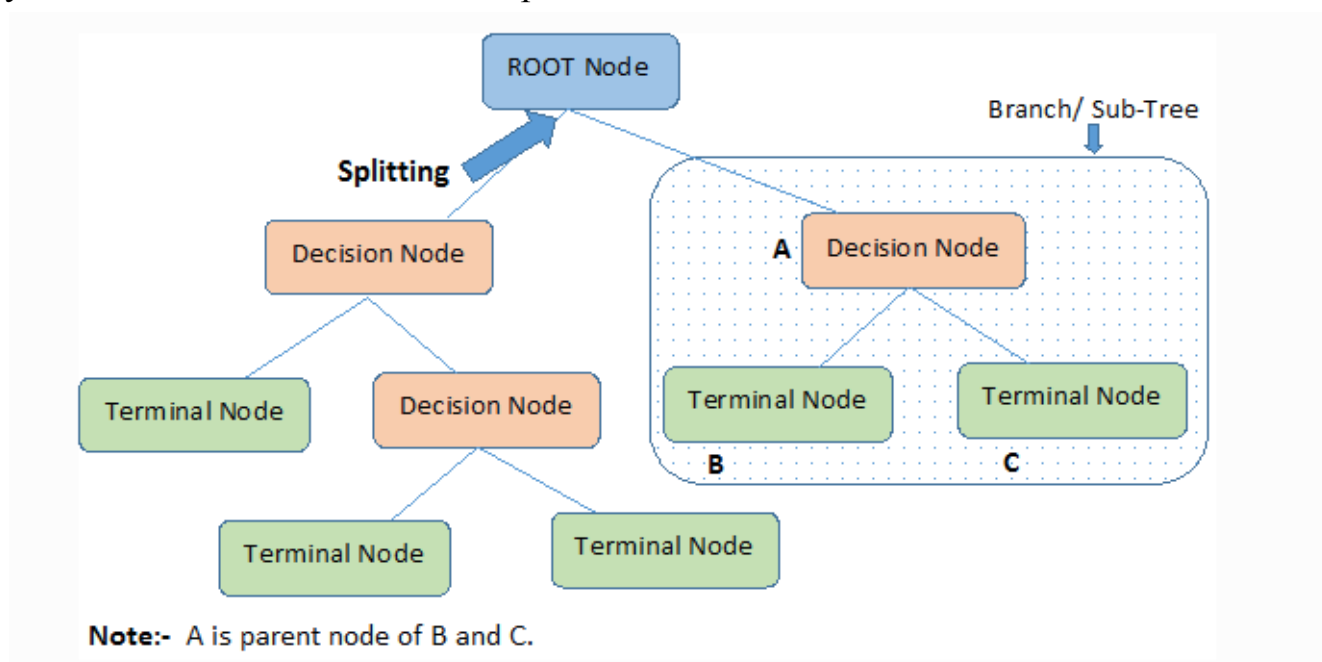


Рисунок 2 - Алгоритм дерева рішень

Випадкові ліси є ансамблевими алгоритмами, які працюють за принципом злісування багатьох дерев рішень для покращення загальної точності і запобігання перенавчанню. Кожне дерево в лісі прогнозує клас, і клас з найбільшою кількістю "голосів" стає передбаченням моделі. Ця техніка є особливо могутньою завдяки своїй здатності до класифікації складних датасетів.

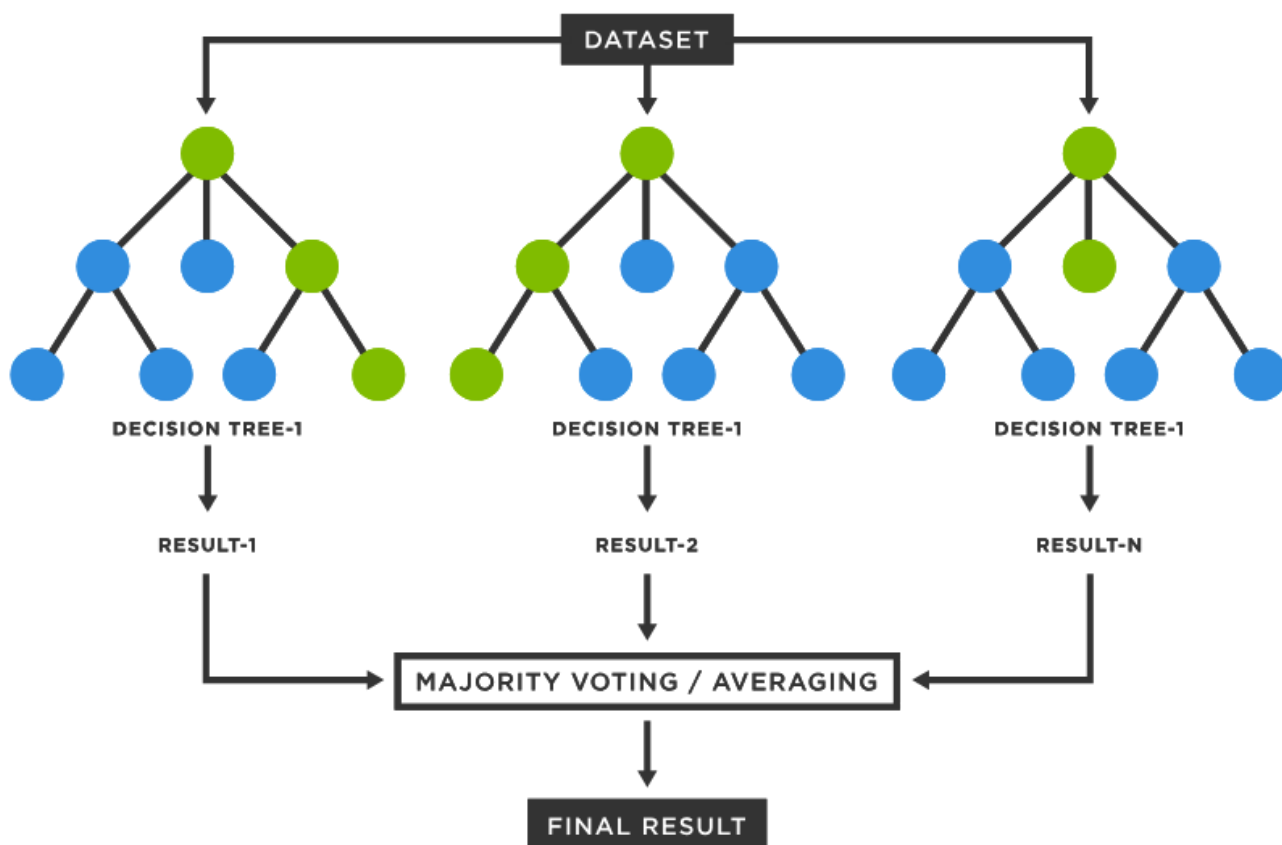


Рисунок 3 - Алгоритм випадкові ліси

Гرادієнтний бустинг використовує принцип покращення слабких передбачень шляхом побудови послідовності моделей, кожна з яких вносить свій вклад у загальне рішення. Такий підхід може збільшити точність передбачень, при цьому враховуючи, що кожна нова модель вчиться на помилках попередніх. На відміну від інших ансамблевих методів, градієнтний бустинг є чутливим до перенавчання, але водночас надає дуже точні прогнози, особливо у великих датасетах.

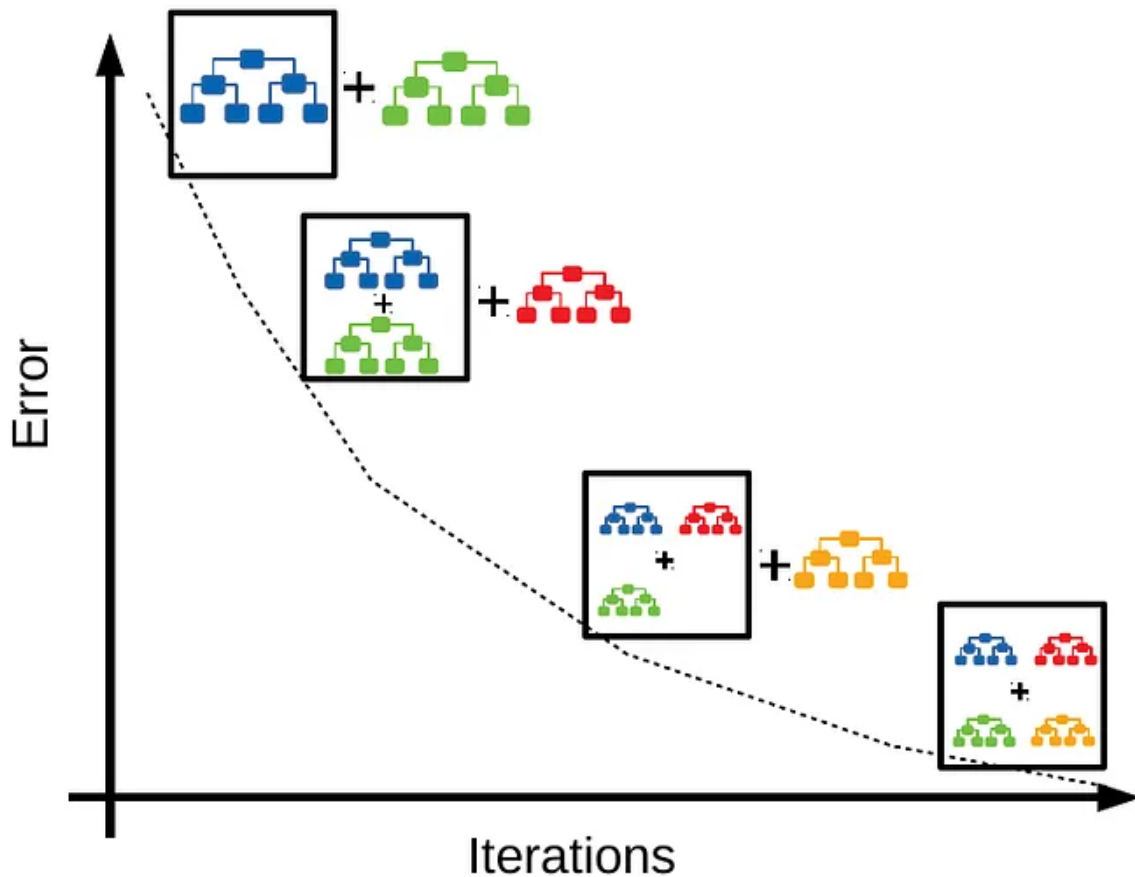


Рисунок 4 - Алгоритм градієнтного підсилення

### 1.3.2 Ефективність алгоритмів машинного навчання в рекрутингу

Огляд ефективності сучасних алгоритмів машинного навчання в прийнятті рішень про найм є винятково значущим для світової промисловості HR-технологій, особливо з огляду на сталий ріст запиту на автоматизацію та оптимізацію рекрутингових процесів. Впровадження таких технік, як дерева рішень, випадкові ліси, градієнтний бустинг та їх модифікації, пропонує радикально новий підхід до аналізу даних кандидатів, що дозволяє значною мірою підсилити якість та точність відбору потенційних працівників.

У рамках даного дослідження було вивчено та порівняно ефективність різних алгоритмів машинного навчання з метою оптимізації процесу відбору кандидатів. Використані алгоритми, такі як дерева рішень, випадкові ліси, та градієнтні дерева, були протестовані на наборі даних, поділеному на навчальний та тестовий сегменти. Метою було визначити, який з алгоритмів найбільш точно та ефективно дозволяє класифікувати кандидатів відповідно до вакансій.

Згідно з результатами експериментів, алгоритм випадкового лісу продемонстрував найвищу загальну точність, а також високі показники за іншими метриками, такими як час обробки, точність, пригадування, та оцінка F1. Водночас було зазначено, що хоча ентропійний критерій мав незначний вплив на ефективність деревовидних моделей, він забезпечував дещо кращі результати порівняно з індексом Джині. Це підкреслює здатність обох критеріїв ефективно використовуватися для задач класифікації в контексті оцінки кандидатів.

Також зазначимо, що час навчання моделей відрізнявся між алгоритмами, де дерева рішень виявилися найшвидшими, в той час як випадкові ліси та градієнтні дерева вимагали більше часу для обробки. Це підкреслює важливість вибору оптимального алгоритму з урахуванням конкретних потреб та ресурсів організації.

На закінчення, важливо зазначити, що машинне навчання слід розглядати як інструмент підтримки прийняття рішень, а не як самостійний механізм для вирішення всіх задач відбору персоналу. Його використання повинно бути інтегроване в ширший контекст HR-стратегій, забезпечуючи високу точність та ефективність при найманні, а також постійне удосконалення та адаптація до змінних умов ринку праці.

#### 1.4 NLP в оцінці відповідності кандидата до вакансії.

Сучасний процес найму вимагає від рекрутерів оперативного та точного аналізу великої кількості резюме кандидатів, щоб знайти найбільш відповідних для конкретних вакансій. Обробка природної мови (NLP) відіграє ключову роль у цьому процесі, надаючи інструменти для автоматизації та поліпшення якості скрінінгу кандидатів.

NLP є розділом штучного інтелекту, який займається розумінням, інтерпретацією та генерацією природної людської мови. Ці технології дозволяють комп'ютерам читати текст, розуміти його смисл, інтерпретувати мовлення та навіть генерувати відповіді. У контексті рекрутингу, NLP може бути використане для аналізу резюме та описів вакансій, виявлення ключових навичок, досвіду та інших важливих атрибутів, які необхідні для відповідності кандидата вимогам роботодавця.

Завдяки використанню NLP, процес скрінінгу резюме стає більш швидким, точним та об'єктивним. Рекрутери можуть сконцентруватися на більш кваліфікованих кандидатах, в той час як система здійснює попереднє сортування та аналіз резюме. Однак, важливо відзначити, що, незважаючи на значні переваги, NLP не може повністю замінити людське розуміння та інтуїцію, і краще використовувати його як допоміжний інструмент у процесі найму.

### 1.4.1 Техніки NLP у контексті аналізу резюме

Застосування NLP в рекрутингу включає такі основні аспекти:

#### 1. Токенізація

Цей процес передбачає розбиття тексту резюме на базові елементи - токени, що можуть бути словами, фразами або навіть символами. Токенізація є фундаментальним кроком у NLP, оскільки вона дозволяє подальше використання тексту для аналізу та вилучення корисної інформації. У контексті аналізу резюме, токенізація допомагає визначити ключові слова та фрази, які можуть бути індикаторами навичок, досвіду чи освіти кандидата.

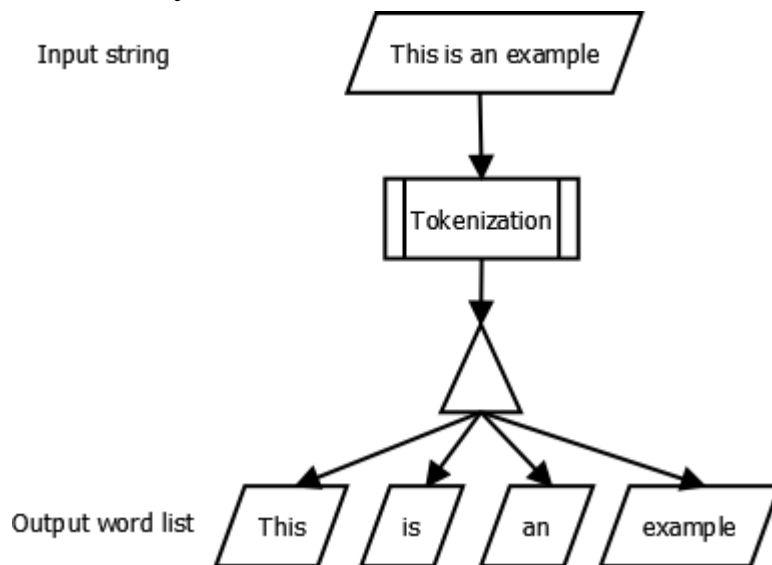


Рисунок 5 - Приклад процесу токенизації

#### 2. Лемматизація та стемінг:

Ці методи спрямовані на зведення слів до їх базової форми. Лемматизація розглядає слово в контексті та використовує словникову базу та морфологічний аналіз для точного визначення базової форми. Стемінг, з іншого боку, зрізає кінцеві частини слів з метою досягнення цієї ж мети, але робить це менш точно. Обидва ці методи важливі для уніфікації термінології в резюме і підвищення точності співставлення кандидатів із вакансіями.

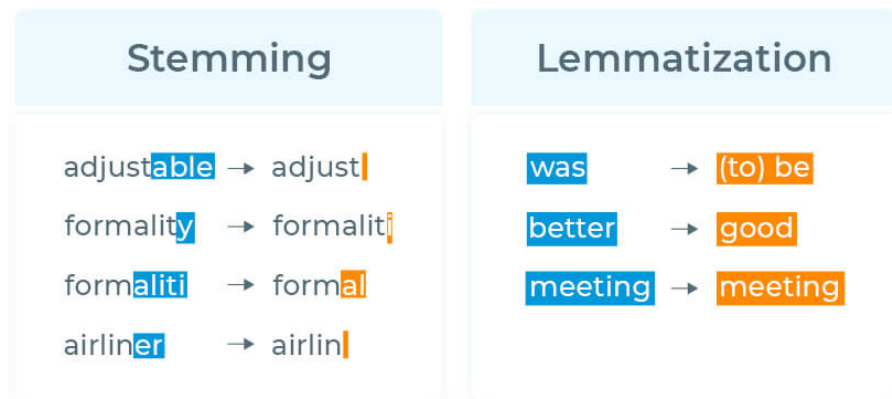


Рисунок 6 - Приклад лемматизації та стемінгу

Вилучення ключових слів та фраз: Ця техніка полягає в ідентифікації та вилученні важливих слова і фраз з резюме, які відображають кваліфікацію та професійні навички кандидатів. Це може включати спеціалізації, технічні терміни, назви посад та інші індикатори, які важливі для оцінки відповідності кандидата вимогам вакансії. Вилучення ключових слів та фраз є критично важливим для створення семантичних відображень та розуміння контенту резюме на глибокому рівні.

Разом ці техніки NLP допомагають трансформувати неструктуровані текстові дані резюме в структуровану інформацію, яка може бути ефективно використана для автоматизованого скрінінгу та оцінки кандидатів.

#### 1.4.2 Використання NLP для розуміння семантики резюме

Використання методів обробки природної мови (NLP) для аналізу резюме та описів вакансій значно розширює можливості оцінки відповідності кандидатів. Аналіз контексту та семантичної близькості дозволяє не просто виявляти наявність ключових слова в тексті, але й розуміти глибинний зміст та відносини між словами та фразами. Це особливо важливо при оцінці відповідності між компетенціями кандидата та вимогами вакансії.

Застосування моделей NLP як Word2Vec чи BERT, чи Embeddings від OpenAI API, дозволяє перетворити слова та фрази у векторні представлення, які відображають їхні семантичні та синтаксичні властивості. Ці вектори можуть бути порівняні за допомогою різних метрик схожості, таких як косинусна схожість, що

дозволяє обчислити ступінь відповідності між резюме кандидата та описом вакансії.

Такий підхід відкриває нові можливості для рекрутингу, дозволяючи не тільки автоматизувати процес скрінінгу резюме, але й забезпечити більшу точність та глибину аналізу кандидатур. Особливо ефективним він є при роботі з великими обсягами даних та в умовах, де потрібна швидка та об'єктивна оцінка великої кількості кандидатів.

### 1.4.3 Порівняльний аналіз основних моделей NLP

Word2Vec, BERT та Embeddings від OpenAI - це сучасні методи обробки природної мови, які використовуються для перетворення тексту у векторне представлення. Кожна з цих моделей має свої особливості та області застосування.

1. Word2Vec - це група моделей, які використовуються для векторизації слів. Вона може виробляти векторне представлення слів, так що слова зі схожими значеннями мають схожі вектори.

Переваги:

- Здатна ефективно вловлювати семантичні відношення між словами.
- Швидка та ефективна в роботі з великими корпусами тексту.

Недоліки:

- Ігнорування контексту слова (одне слово має один вектор незалежно від контексту вживання).
- Не враховує слова, які не зустрічалися в тренувальному наборі даних.

2. BERT - це передова модель для визначення контексту слів у реченні. Вона використовує механізм "трансформерів", що дозволяє моделі краще розуміти контекст та зв'язки між словами.

Переваги:

- Розуміння контексту слова в реченні, тому вектори більш точно відображають значення слова.
- Висока точність та ефективність у багатьох задачах обробки тексту.

Недоліки:

- Вимогливість до обчислювальних ресурсів.
- Складність та вартість тренування моделі.

3. Embeddings від OpenAI - це семантичні вектори, які генеруються за допомогою алгоритмів машинного навчання та нейромереж, розроблених OpenAI.

Вони забезпечують глибоке розуміння тексту та можуть використовуватися в різних задачах NLP.

Переваги:

- Висока точність та здатність розуміти складний текстовий контент.
- Підтримка широкого спектра мов та доменів.

Недоліки:

- Потреба в доступі до API OpenAI, що може вимагати додаткових витрат.
- Залежність від зовнішнього сервісу та його обмежень.

Коли вибирати яку модель:

1. Word2Vec ефективний для швидкого та простого аналізу семантики слів, коли контекст не є критично важливим.
2. BERT ідеально підходить для задач, де необхідно глибоко розуміти контекст та зв'язки між словами в тексті, наприклад, в задачах перекладу, класифікації тексту, відповідей на питання.
3. Embeddings від OpenAI можуть бути використані для широкого спектру задач, де потрібно висока точність та розуміння тексту на різних мовах та в різних доменах.

Кожен метод має свої особливості та найкраще підходить для певних типів задач. При виборі методу варто враховувати специфіку задачі, доступність ресурсів, потребу в швидкості та точності обробки, а також вимоги до контекстуального розуміння тексту.

## 1.5 Постановка задачі

Розробити та впровадити систему автоматичної оцінки відповідності кандидата до вакансії на основі аналізу резюме, використовуючи комплексні метрики, які враховують різні характеристики кандидата.

Метою дипломного проекту є створення програмного рішення, яке забезпечить ефективний та об'єктивний відбір кандидатів для певної спеціальності (наприклад, front-end розробника) за допомогою системи оцінки, яка враховує різноманітні аспекти кандидатського досвіду та навичок. Основна інновація полягає в розпаршиванні та ваговому визначенні характеристик, таких як досвід

роботи, освіта, стек технологій та рівень їх володіння, та інші, що дозволить більш точно визначати відповідність кандидата до конкретної вакансії.

*Ключові компоненти та вимоги:*

1. Аналіз резюме:

Розробити алгоритм аналізу резюме, який виявлятиме та екстрагуватиме ключові характеристики, такі як досвід, освіта, стек технологій тощо.

2. Визначення ваг характеристик:

Розробити методику для визначення ваги кожної характеристики в контексті конкретної спеціальності. Вага повинна відображати важливість кожної характеристики для вакансії.

3. Створення метрики оцінки:

Розробити систему метрик, яка враховуватиме вагу кожної характеристики та генеруватиме кінцеву оцінку відповідності кандидата до вакансії.

4. Впровадження в роботу:

Інтегрувати розроблену систему в процес відбору кандидатів, забезпечуючи швидку та ефективну фільтрацію резюме.

*Очікувані результати:*

Створення системи, яка дозволить ефективно та об'єктивно оцінювати відповідність кандидатів до вакансій, спираючись на вагово розподілені та збалансовані характеристики, тим самим оптимізуючи процес відбору та підвищуючи його об'єктивність.

У роботі буде зосереджено увагу на важливості точності, швидкості, адаптивності та об'єктивності у процесі оцінки кандидатів, а також на

потенційному зниженні суб'єктивного впливу та помилок, пов'язаних із людським фактором. Важливим аспектом стане також вивчення можливостей зменшення часу та ресурсів, необхідних для відбору персоналу, та підвищення якості та ефективності прийнятих рішень.

## 1.6 Висновки до розділу 1

У першому розділі магістерської дисертації було проаналізовано сучасну проблематику та актуальність оцінки відповідності кандидатів до вакансій, визначивши її як значущу потребу в умовах зростаючої конкуренції та вимог до якості найму на ринку праці. Це підкреслює необхідність розвитку та впровадження ефективних інструментів для вдосконалення процесів відбору та оцінки кандидатів.

Огляд існуючих методів, включаючи традиційний скрінінг резюме та системи відстеження заявників (ATS), виявив їх переваги, такі як здатність до глибокого аналізу та широкого охоплення, а також недоліки, що полягають у часових затратах, суб'єктивності оцінки та ризику пропускання кваліфікованих кандидатів. Це вказує на потребу розробки більш досконалих методів і інструментів оцінки.

Значний внесок у вирішення цієї проблеми може зробити застосування машинного навчання та семантичного аналізу. Ці технології дозволяють автоматизувати процеси оцінки, підвищуючи їх точність, швидкість та об'єктивність, а також забезпечують здатність системи до адаптації під специфічні вимоги роботодавців.

Водночас, напрямки подальших досліджень у цій області включають розвиток нових алгоритмів та моделей, здатних ефективно обробляти великі обсяги даних, вдосконалення методів врахування контекстуальної інформації, а також розробку рішень для зменшення упередженості у прийнятті рішень.

Особливу увагу в роботі було приділено порівнянню різних моделей векторного представлення тексту, таких як Word2Vec, BERT та Embeddings від

OpenAI. За результатами аналізу та порівняння, Embeddings від OpenAI виявилися найбільш ефективними для реалізації поставленої задачі. Ця технологія забезпечує високу точність у визначенні семантичної близькості між текстом резюме та описами вакансій, враховуючи контекст та нюанси мови. Водночас модель OpenAI є адаптивною та здатною до навчання, що дозволяє їй постійно вдосконалюватися та адаптуватися до змінних умов та вимог ринку праці.

Виходячи з проведеного аналізу, ми сформулювали задачу нашого дослідження як розробку інноваційної системи оцінки відповідності кандидатів, яка інтегрує передові методи машинного навчання та семантичного аналізу для підвищення ефективності та точності процесів відбору персоналу. Основними критеріями успіху для цієї системи є її здатність до швидкого адаптування під змінні умови та потреби ринку праці, а також надання об'єктивних та надійних результатів оцінки.

## 2. МЕТОДИ ОЦІНКИ ВІДПОВІДНОСТІ КАНДИДАТА ДО ВАКАНСІЇ

### 2.1. Обґрунтування вибору напряму досліджень

У першому розділі було чітко визначено актуальність теми. Сучасний ринок праці вимагає від систем оцінки кандидатів більшої точності, об'єктивності, та швидкості обробки інформації. Це вимагання є особливо актуальним у контексті зростаючої конкуренції та динаміки ринкових умов, що ставить перед дослідниками завдання розробки ефективних інструментів для оцінки кандидатів.

Було проаналізовано та визначено основні недоліки існуючих методів оцінки кандидатів, таких як скрінінг резюме та ATS системи, включаючи їх суб'єктивність, часові затрати, та високий ризик пропускання кваліфікованих кандидатів. Як було показано, машинне навчання та семантичний аналіз можуть істотно підвищити ефективність оцінки, забезпечуючи більшу точність та об'єктивність.

Як вже було викладено, технології машинного навчання та семантичного аналізу відіграють ключову роль у підвищенні ефективності процесів відбору. Особливу увагу варто приділити аналізу та використанню ембедінгів від OpenAI, які демонструють високу точність у семантичному зіставленні текстів резюме та описів вакансій. Ці технології не тільки дозволяють ефективно обробляти великі обсяги даних, але й враховувати нюанси та контекст мови, що є критично важливим для точної оцінки кандидатів.

На основі аналізу першого розділу магістерської дисертації, виявлено прогалини у сучасних методах оцінки відповідності кандидатів, що відкриває шлях для розвитку нових підходів і технологій. У контексті цього дослідження пропонується зосередитися на наступних аспектах:

#### 1. Використання Embeddings від OpenAI:

У рамках дослідження використання ембедінгів від OpenAI, особливу увагу буде приділено аналізу та застосуванню цих технологій для створення точної та адаптивної системи оцінки кандидатів. Ембедінги - це векторні представлення слів чи фраз, які зберігають семантичну близькість та контекстуальні зв'язки, дозволяючи машинам "розуміти" текст на рівні, подібному до людського сприйняття. Їх використання може кардинально покращити процес оцінки відповідності кандидата, забезпечивши наступні переваги:

- Визначення семантичної близькості: Ембедінги від OpenAI ефективні в визначенні семантичної близькості між резюме кандидата та описом вакансії. Це означає, що система може точно зіставляти вимоги до вакансії з кваліфікацією та досвідом кандидата, враховуючи не тільки ключові слова, а

й глибинний контекст та семантичні нюанси, які можуть бути втрачені при простому текстовому аналізі.

- Аналіз ключових навичок та кваліфікацій: Через здатність до глибокого семантичного аналізу, ембедінги можуть використовуватися для ідентифікації та оцінки ключових навичок та кваліфікацій, вказаних у резюме. Це включає не тільки визначення наявності певних технічних знань або м'яких навичок, але й оцінку глибини та релевантності цих навичок для конкретної роботи.
- Релевантність досвіду: Ембедінги також можуть бути використані для оцінки релевантності професійного досвіду кандидата. Завдяки аналізу контексту та семантичних відносин, система може розрізняти досвід, який точно відповідає вимогам вакансії, від менш релевантного або застарілого досвіду.

У процесі дослідження будуть вивчені різні типи ембедінгів, їхні особливості та можливості для застосування в контексті оцінки кандидатів. Окрему увагу буде приділено практичному застосуванню цих технологій: як вони можуть інтегруватися в існуючі HR-системи, які виклики можуть виникнути під час їх імплементації, та як забезпечити їх ефективність та точність. Це дозволить створити комплексний підхід до оцінки відповідності кандидатів, що базується на найновіших досягненнях у сфері машинного навчання та обробки природної мови.

## 2. Алгоритми оцінки:

Розробка та адаптація алгоритмів для оцінки кваліфікації кандидата є ключовим аспектом покращення процесу відбору та найму. Ці алгоритми мають на меті автоматизувати та оптимізувати процес оцінки, використовуючи сучасні методи машинного навчання та глибокого навчання. Детальніше, в рамках дослідження будуть розроблені наступні компоненти:

- Алгоритми аналізу навичок: Ці алгоритми будуть спрямовані на ідентифікацію та оцінку фахових навичок та знань, які вказані в резюме кандидата. Завдяки аналізу тексту резюме, система зможе визначити ключові слова та фрази, що відображають специфічні навички, та оцінити їх релевантність до вимог конкретної вакансії. Такий підхід дозволить не тільки виявляти наявні навички, але й оцінювати їх рівень, враховуючи контекст та специфіку роботи.
- Оцінка освіти та кваліфікацій: Алгоритми будуть аналізувати освітній та професійний шлях кандидата, включаючи рівень освіти, спеціалізацію, додаткові курси та сертифікати. Це допоможе роботодавцям отримати більш повне уявлення про кваліфікацію кандидата та його потенціал для навчання та розвитку.

- Аналіз досвіду роботи: Спеціально розроблені алгоритми будуть оцінювати професійний досвід кандидата, аналізуючи інформацію про попередні місця роботи, тривалість трудової діяльності, характер виконуваних завдань та досягнення. Це дасть змогу не тільки оцінити досвід, але й зрозуміти його актуальність та вплив на потенційну роботу кандидата на новій посаді.

Кожен з цих аспектів вимагає детального та глибокого аналізу, що становить суттєві виклики для дослідників. Зокрема, важливо забезпечити високу точність та об'єктивність оцінки, мінімізувати упередженість та надати системі здатність до адаптації під специфічні вимоги різних вакансій. Також важливим є забезпечення прозорості та зрозумілості результатів оцінки для роботодавців та кандидатів.

Використання методів машинного навчання та глибокого навчання дозволить створити алгоритми, здатні ефективно обробляти великі обсяги даних, враховувати різноманітність та складність інформації, представленої в резюме, та забезпечувати високу точність та об'єктивність оцінки. Таким чином, дослідження спрямоване на створення комплексної системи оцінки кандидатів, яка враховує всі ключові аспекти їхньої кваліфікації та забезпечує роботодавцям надійний інструмент для прийняття обґрунтованих кадрових рішень.

### 3. Використання вагових коефіцієнтів:

Використання вагових коефіцієнтів в системі оцінки відповідності кандидатів є важливим інструментом для налаштування процесу відбору під конкретні вимоги вакансії. Цей підхід дозволяє роботодавцям визначати важливість кожного аспекту кваліфікації кандидата, такого як освіта, досвід роботи, навички, та інші характеристики, згідно зі специфікою посади. Детальніше, в рамках дослідження буде розглянуто наступні аспекти:

- Визначення критеріїв вакансії: Розробка методики визначення ключових критеріїв для кожної вакансії, яка дозволить роботодавцям визначати, які аспекти кваліфікації є найбільш важливими. Наприклад, для технічної посади може бути високим коефіцієнт для технічних навичок, тоді як для менеджерської позиції - для навичок комунікації та лідерства.
- Розробка системи вагових коефіцієнтів: Створення гнучкої системи, яка дозволяє роботодавцям встановлювати вагові коефіцієнти для різних характеристик кандидата. Це може бути реалізовано через інтерфейс, де роботодавець може "налаштовувати" вагу кожного критерію, відповідно до специфіки вакансії.
- Адаптація оцінки: Система буде використовувати встановлені вагові коефіцієнти для адаптації процесу оцінки. Це означає, що результати оцінки

будуть більше зосереджені на тих аспектах кваліфікації, які мають вищу вагу, тим самим забезпечуючи більшу релевантність та точність оцінки.

- Тестування та оптимізація: Проведення тестування системи для забезпечення її ефективності та точності. Це включає аналіз результатів, збір зворотного зв'язку від роботодавців, та внесення коректив для оптимізації вагових коефіцієнтів та оцінки в цілому.

Вагові коефіцієнти дозволяють створити більш гнучку та цільову систему оцінки, яка може бути адаптована під унікальні потреби кожної вакансії та організації. Це не тільки підвищує ефективність процесу відбору, але й забезпечує більш об'єктивний та виправданий підхід до оцінки кандидатів, що в кінцевому результаті може сприяти вибору найкращих претендентів на посаду.

Ці напрямки досліджень спрямовані на закриття існуючих прогалин у методах оцінки відповідності кандидатів та на розвиток більш ефективних, точних та адаптивних систем. Результати цих досліджень не тільки покращать процеси найму, але й забезпечать роботодавцям потужний інструмент для вибору найкращих кандидатів на основі об'єктивних та надійних даних.

## 2.2 Формулювання цільової функції та метрик оцінки

Сучасний ринок праці характеризується високою конкуренцією та швидкими змінами, що змушує роботодавців шукати все більш вдосконалені та точні способи оцінки кандидатів.

Цільова функція у цьому контексті виступає як інструмент, що дозволяє зібрати та обробити інформацію про кандидатів, враховуючи такі параметри як їх навички, освіту та досвід. Вона повинна бути максимально об'єктивною та адаптивною, щоб забезпечити високу точність оцінки та ефективність подальшого відбору кандидатів.

Цільова функція, яка запропонована в даному дослідженні, використовує комплексний підхід до аналізу даних про кандидатів, об'єднуючи оцінки їх навичок, освіти та досвіду.

Основною задачею цільової функції є надання кількісної оцінки відповідності кандидатів до певної вакансії, враховуючи різноманітні аспекти їхнього профілю, такі як навички, освіта, та досвід. Ця функція визначає спосіб

інтеграції та взаємодії різних оцінок для формування загальної карти відповідності.

У контексті нашого дослідження, цільова функція включає в себе визначення та аналіз трьох основних параметрів: навички, освіта та досвід, кожен з яких відображає відповідність кандидата за певним критерієм. Ці критерії взаємодіють за допомогою вагових коефіцієнтів, які задаються користувачем, для адаптації системи під специфічні потреби вакансії.

*Приклад розрахунку:*

Оцінка відповідності кандидата може бути розрахована як:

$$\text{Оцінка} = \sum (\text{Вага кожного критерію} \times \text{Значення критерію})$$

Метрики оцінки у даній роботі включають:

- Точність відповідності: Метрика, яка вимірює частку правильно ідентифікованих кандидатів відповідно до вимог вакансії. Ця метрика критична для забезпечення якісного та ефективного відбору.
- Загальний бал оцінки: Агрегований показник, що враховує всі ключові параметри оцінювання, надаючи інтегрований підсумковий бал для кожного кандидата.

Метрики оцінки включають в себе не лише кінцеві чисельні показники, але й методи їх розрахунку та інтерпретації:

1. Навички: Висока відповідність навичок до вимог роботи є критичною для успішного виконання професійних обов'язків, тому точне їх оцінювання є важливим для кінцевого відбору кандидатів.
2. Освіта: Освіта часто відображає теоретичну підготовку та можливості кандидата у певній області. Точне оцінювання освіти дозволяє виявити потенціал кандидата для навчання та професійного розвитку.
3. Досвід: Досвід роботи дає уявлення про практичні здібності кандидата та його здатність адаптуватися до нових робочих умов. Він є важливим показником для оцінки як поточних здібностей, так і потенціалу для майбутнього розвитку.

Дослідження також включає детальний аналіз та оптимізацію цих метрик із застосуванням сучасних методів статистичного аналізу та машинного навчання. Проведено порівняльний аналіз із іншими існуючими методами та системами оцінки, що демонструє переваги запропонованого підходу у контексті точності, швидкості відбору та здатності до адаптації під різноманітні вимоги вакансій.

## 2.3 Вибір конкретних алгоритмів та моделей та його обґрунтування.

В розробці систем, що базуються на машинному навчанні для оцінки відповідності кандидата до вакансії, вибір алгоритмів та моделей відіграє ключову роль. Цей підрозділ присвячений аналізу та обґрунтуванню вибору певних технологічних рішень, які були використані у програмі для оцінки кандидатів. Мета полягає не лише у висвітленні того, які інструменти були вибрані, а й чому саме ці інструменти є найбільш підходящими для вирішення поставлених завдань.

Оцінка кандидатів у сучасному світі вимагає точності, швидкості, та адаптивності. Саме тому важливо обрати алгоритми та моделі, які максимально відповідають цим вимогам і дозволяють ефективно обробляти великі об'єми даних, виявляти нюанси та контекст резюме та описів вакансій. У цьому підрозділі розглянуті основні категорії технологічних рішень, які були вибрані для системи оцінки.

### 2.3.1. Семантичне векторне представлення тексту

У сучасному рекрутингу ключову роль відіграє здатність системи швидко та точно аналізувати великі обсяги текстових даних - резюме кандидатів та описи вакансій. Саме тут на сцену виходить семантичне векторне представлення тексту.

*Використання семантичних векторів:*

Семантичні вектори, як правило, генеруються за допомогою алгоритмів машинного навчання, які перетворюють слова, фрази чи цілі документи у вектори чисел. Ці вектори відображають семантичну близькість текстів та їх контекст, що робить їх особливо корисними для розуміння та порівняння смислового змісту текстів.

У контексті розробленої програми семантичні вектори використовуються для:

1. Аналізу текстів резюме та описів вакансій:

- а. Тексти резюме кандидатів та описи вакансій перетворюються на вектори, що дозволяє оцінювати ступінь їх відповідності.
  - б. Система аналізує семантичну близькість між навичками, які зазначені в резюме, та вимогами, які представлені в описі вакансії.
2. Визначенню відповідності кандидата до вакансії:
- а. Застосування косинусної міри схожості між векторами дозволяє визначити наскільки близько навички та досвід кандидата відповідають запитам роботодавця.
  - б. Використання цих векторних представлень сприяє більш точному та глибокому аналізу даних, що значно підвищує якість відбору кандидатів.

#### *Технологічне обґрунтування:*

Використання семантичних векторів у програмі базується на наступних технологіях та принципах:

- Моделі векторного представлення тексту: Застосування передових моделей, таких як Word2Vec, GloVe або BERT для створення високоякісних векторних представлень тексту.
- Інтеграція з сервісами зовнішніх API: Використання API, наприклад, від OpenAI для отримання доступу до передових моделей обробки природної мови та векторних представлень.

#### *Переваги та значення:*

Семантичне векторне представлення тексту відіграє важливу роль у підвищенні точності та ефективності процесу відбору кандидатів:

- Підвищення точності оцінки: Завдяки аналізу семантичної близькості можна більш точно оцінити відповідність кандидата до вакансії, знизивши при цьому кількість помилкових відборів.
- Ефективність та швидкість: Автоматизація процесу аналізу тексту значно збільшує швидкість обробки даних та відбору кандидатів, що є особливо важливим у умовах великої кількості вакансій та кандидатів.

Використання семантичного векторного представлення тексту в програмі для оцінки відповідності кандидатів є важливим кроком на шляху до створення більш точних, ефективних та адаптивних систем рекрутингу.

### 2.3.2. Алгоритми оцінки

Визначення точних та об'єктивних методів оцінювання є ключовим завданням для ефективної системи відбору. У цьому контексті, особлива увага приділяється трьом фундаментальним аспектам кандидатів, а саме: навичкам, освіті та досвіду. Оцінка цих параметрів вимагає застосування специфічних методологій та

алгоритмів, кожен з яких має бути обґрунтований та детально описаний у відповідності до наукових стандартів.

#### 2.3.2.1 Оцінка навичок

У дослідженні пропонується застосувати комплексний підхід до оцінки навичок кандидатів. Цей підхід включає в себе аналіз резюме та описів вакансій на наявність ключових технічних навичок, які важливі для позиції. Подальше ранжування та класифікація кандидатів здійснюється на основі відповідності цих навичок вимогам вакансії.

Використання семантичних векторів дозволяє глибоко аналізувати текстові дані та ідентифікувати відповідність між навичками кандидатів та вимогами вакансій. Цей підхід не тільки покращує точність оцінки, але й надає можливість обробляти великі об'єми даних швидко та ефективно. Використання косинусної міри схожості забезпечує кількісну оцінку близькості профілів кандидатів до вимог вакансії, що дозволяє ранжувати кандидатів за ступенем їх відповідності. Застосування такого методу в контексті відбору персоналу є обґрунтованим, оскільки воно спирається на об'єктивні, кількісно вимірювані показники та забезпечує високий рівень відтворюваності результатів.

Вибір цього підходу зумовлений потребою в об'єктивній та масштабованій оцінці кандидатів. Використання машинного навчання та аналітичних інструментів дозволить систематизувати великий обсяг даних та виділити найбільш релевантних кандидатів, спираючись на кваліфікований аналіз їхніх навичок. Такий підхід дозволяє не тільки виявляти очевидні відповідності, але й розуміти більш тонкі семантичні зв'язки в текстах, забезпечуючи таким чином більш глибокий та всебічний аналіз профілів кандидатів.

#### 2.3.2.2 Оцінка освіти

В реалізації програми передбачено алгоритм оцінки освітнього рівня кандидатів, який базується на визначеному наборі освітніх категорій. Кожна категорія отримує певний числовий коефіцієнт, що відображає її вагу в контексті загальної оцінки освіти. Оцінка здійснюється шляхом порівняння рівня освіти, зазначеного у вакансії, з рівнем освіти кандидата, при цьому враховується конкретний числовий коефіцієнт кожної освітньої категорії.

Така схема оцінки дозволяє кількісно оцінити освітній рівень кандидатів, надаючи можливість систематичного та об'єктивного порівняння з вимогами, встановленими для вакансії. Застосування категорій типу "diploma", "degree", "master", "phd" та відповідних їм числових коефіцієнтів дозволяє уніфікувати процес оцінки та забезпечити його відповідність загальноприйнятим стандартам кваліфікації.

Додатково, використання розрахункової логіки, яка враховує розходження між освітою кандидата та вимогами вакансії, дозволяє зробити оцінку більш гнучкою та адаптивною. Це сприяє точному визначенню ступеню відповідності освітнього рівня кандидата потребам роботодавця, враховуючи не лише наявність певної освіти, але й її відносну значущість.

Застосування числових коефіцієнтів для кожної освітньої категорії дозволяє кількісно оцінити відмінності між різними рівнями освіти, надаючи основу для зваженої та об'єктивної оцінки. Такий підхід враховує як загальноприйняті освітні стандарти, так і специфіку конкретної вакансії, що робить його важливим інструментом у процесі відбору кандидатів.

### 2.3.2.3 Оцінка досвіду

В реалізації програми передбачено алгоритм для оцінки досвіду роботи кандидатів, який заснований на кількісному аналізі тривалості їхньої зайнятості на різних місцях роботи. Алгоритм обраховує загальну кількість місяців роботи кандидата на основі даних, наданих у резюме, та порівнює цю тривалість із вимогами до досвіду роботи, вказаними у вакансії. Результат оцінки досвіду — це числовий показник, який вираховується з використанням функції, що відображає відносну відповідність досвіду кандидата вимогам вакансії.

Цей підхід було обрано через його здатність надавати кількісні та суб'єктивно важливі оцінки досвіду роботи. Використання кількісних показників, таких як загальна кількість місяців роботи, дозволяє зробити процес оцінки більш стандартизованим і порівняльним між різними кандидатами. Цей метод трансформації призначений для перетворення агрегованого досвіду у кількісний показник, який можна ефективно використовувати для оцінки ступеня відповідності кандидата вимогам вакансії. Це враховує не тільки загальну тривалість робочого досвіду кандидата, але й його релевантність, перетворюючи числові дані про досвід в більш нюансований та інформативний показник. Такий підхід дозволяє зробити оцінку більш гнучкою та точною, адаптуючи її до специфіки кожної конкретної вакансії.

### 2.3.3. Вагові коефіцієнти

У рамках розробки ефективної системи оцінки відповідності кандидатів до вакансій, значну увагу слід приділити механізму вагових коефіцієнтів. Ця частина дослідження спрямована на вивчення та реалізацію системи вагових коефіцієнтів, які дозволяють налаштувати оцінювання відповідно до специфічних потреб та пріоритетів роботодавця.

### *Опис та обґрунтування використання вагових коефіцієнтів*

Вагові коефіцієнти у системі оцінки відповідності кандидатів є критично важливими для забезпечення гнучкості та адаптивності процесу відбору. Вони надають можливість користувачам системи - роботодавцям та HR-фахівцям - модифікувати оцінку кандидатів з урахуванням особливостей конкретної вакансії чи стратегії підбору персоналу. Використання вагових коефіцієнтів дозволяє акцентувати увагу на тих параметрах оцінки, які є найбільш критичними для конкретної позиції.

Наприклад, для технічної позиції може бути підвищено значення коефіцієнта для навичок, тоді як для керівної позиції більш важливим може бути досвід та освіта. Такий підхід забезпечує більшу точність оцінки та знижує ризик пропуску ідеально підходящих кандидатів.

### *Практичне застосування вагових коефіцієнтів*

На практиці, вагові коефіцієнти можуть бути інтегровані у систему оцінки через інтерфейс користувача, що дозволяє HR-фахівцям та роботодавцям вказувати важливість кожного аспекту оцінювання для кожної вакансії. Це може бути реалізовано через шкалу значень або через вибір з фіксованого набору коефіцієнтів. Такий інструментарій дозволяє зробити процес відбору більш цілеспрямованим та ефективним.

Включення вагових коефіцієнтів у систему оцінки відповідності кандидатів до вакансій відіграє важливу роль у підвищенні адаптивності та ефективності процесу відбору. Це забезпечує додаткову гнучкість, дозволяючи роботодавцям кастомізувати оцінювання з урахуванням унікальних характеристик та вимог вакансій, а також їхніх власних стратегічних пріоритетів. Застосування вагових коефіцієнтів сприяє створенню більш точної та об'єктивної системи оцінки, здатної адекватно відображати різноманітність та специфіку робочих ролей у сучасному бізнесі.

### 2.3.4. Порівняльний аналіз з іншими методами

Порівняльний аналіз алгоритмів оцінки навичок, освіти та досвіду, що використовуються в нашій розробленій системі, здійснюється з метою виявлення та обґрунтування їхніх переваг порівняно з іншими доступними методами на ринку. Аналіз спрямований на демонстрацію ефективності обраних підходів, їхньої здатності задовольняти сучасні потреби у відборі персоналу та їхню перевагу перед традиційними та інноваційними методами оцінки.

### *Порівняння алгоритмів оцінки навичок*

Розглядаючи оцінку навичок, важливо відзначити, що в розробленій системі використовуються семантичні вектори для аналізу схожості між текстами резюме та описами вакансій. Цей підхід забезпечує глибокий і точний аналіз текстів, виділяючи релевантні навички кандидата. В порівнянні з традиційними методами, які можуть включати ручний пошук ключових слів або прості текстильні аналізи, семантичні вектори дозволяють досягти більшої точності та об'єктивності в оцінці.

Огляд методів оцінки навичок:

- Традиційний підхід (Keyword matching): Метод заснований на пошуку ключових слів у тексті резюме та вакансій. Часто включає ручний перегляд та оцінку HR-спеціалістами.
- Ручний аналіз тексту (Manual text review): Цей метод передбачає глибокий ручний аналіз текстів резюме та описів вакансій HR-спеціалістами для ідентифікації відповідності.
- Семантичне векторне представлення (Semantic Vector Analysis): Використання семантичних векторів для аналізу та порівняння текстів резюме та вакансій, як у запропонованій системі.

| Критерій / Метод                      | Традиційний підхід | Ручний аналіз тексту | Семантичне векторне представлення |
|---------------------------------------|--------------------|----------------------|-----------------------------------|
| Точність оцінки                       | Низька             | Середня              | Висока                            |
| Швидкість процесу                     | Швидкий            | Повільний            | Швидкий                           |
| Вимоги до кваліфікації оцінювачів     | Низькі             | Високі               | Середні                           |
| Адаптивність до різних вакансій       | Низька             | Середня              | Висока                            |
| Суб'єктивний вплив оцінювача          | Високий            | Високий              | Низький                           |
| Вартість впровадження та експлуатації | Низька             | Середня              | Середня/Висока                    |

Таблиця 1 – Порівняльна таблиця методів оцінки навичок кандидата

Таблиця вказує на загальні тенденції та може варіюватися залежно від специфіки вакансій та доступних ресурсів.

Порівняння та обґрунтування вибору методу:

1. Точність оцінки: Семантичне векторне представлення забезпечує високу точність оцінки, адже враховує не тільки наявність ключових слів, а й контекст їх використання, що дозволяє більш точно ідентифікувати відповідність кандидата вакансії.
2. Швидкість процесу: Хоча традиційний підхід може бути досить швидким, він часто не забезпечує необхідної глибини аналізу. Семантичні вектори забезпечують швидкий та в той же час глибокий аналіз.
3. Адаптивність до різних вакансій: Семантичні вектори можуть бути легко адаптовані до різних типів вакансій, забезпечуючи високу універсальність методу.
4. Вартість впровадження та експлуатації: Незважаючи на потенційно вищу вартість впровадження та експлуатації семантичних векторів, інвестиції окупаються через збільшення точності та ефективності відбору.

На основі порівняльного аналізу можна зробити висновок, що використання семантичних векторів для оцінки навичок є переважним методом у контексті сучасних вимог до точності, швидкості та адаптивності процесу відбору персоналу. Це дозволяє ефективно відповідати на потреби роботодавців, забезпечуючи високу якість відбору та оптимальне використання ресурсів.

#### *Порівняння алгоритмів оцінки освіти*

Щодо оцінки досвіду, метод, запропонований в системі, базується на кількісному аналізі загального стажу роботи кандидата. Він забезпечує можливість враховувати не лише загальну тривалість роботи, але й її релевантність до вимог вакансії. Відмінно від інших методів, які можуть не враховувати різноманітність досвіду або його актуальність, розроблена система надає більш комплексний погляд на досвід кандидата.

Для порівняльного аналізу обрано наступні методи оцінки освіти, які часто використовуються в процесі відбору персоналу:

1. Пряме порівняння дипломів (Direct Diploma Comparison): Метод базується на прямому порівнянні наявних дипломів кандидата з вимогами вакансії без додаткового аналізу спеціалізації або рівня навчального закладу.
2. Система ранжування університетів (University Ranking System): Оцінка базується на репутації навчального закладу, з якого кандидат отримав освіту, використовуючи різні рейтинги університетів.
3. Розроблена система числової оцінки (Developed Numerical Assessment): Метод, що застосовується в нашій системі, враховує різні рівні освіти та

надає кількісну оцінку, що відображає відповідність кандидата вимогам вакансії.

| Критерій / Метод                     | Пряме порівняння дипломів | Система ранжування університетів | Розроблена система числової оцінки |
|--------------------------------------|---------------------------|----------------------------------|------------------------------------|
| Точність оцінки                      | Низька                    | Середня                          | Висока                             |
| Врахування спеціалізації             | Ні                        | Частково                         | Так                                |
| Гнучкість у відборі кандидатів       | Низька                    | Середня                          | Висока                             |
| Врахування рівня навчального закладу | Ні                        | Так                              | Так                                |
| Вартість впровадження                | Низька                    | Середня                          | Середня                            |
| Швидкість процесу                    | Швидка                    | Середня                          | Швидка                             |

Таблиця 2 – Порівняльна таблиця методів оцінки освіти кандидата

Зазначені оцінки є приблизними та можуть варіюватися в залежності від специфіки вакансії та ринку праці.

Обраний метод числової оцінки освіти в нашій системі дозволяє з більшою точністю і гнучкістю оцінювати освітній рівень кандидатів, адаптуючись до різноманітних вимог роботодавців. Він забезпечує більш об'єктивне та всебічне розуміння освітнього фону кандидата, враховуючи не лише наявність диплому, але й його спеціалізацію, рівень навчального закладу та інші фактори. Це робить метод особливо ефективним для відбору кандидатів на позиції, де освіта відіграє ключову роль.

#### *Порівняння алгоритмів оцінки досвіду*

Щодо оцінки досвіду, метод, запропонований в системі, базується на кількісному аналізі загального стажу роботи кандидата. Він забезпечує можливість враховувати не лише загальну тривалість роботи, але й її релевантність до вимог вакансії. Відмінно від інших методів, які можуть не враховувати різноманітність досвіду або його актуальність, розроблена система надає більш комплексний погляд на досвід кандидата.

Для порівняльного аналізу розглянемо наступні методи оцінки досвіду, які часто застосовуються в рекрутингу:

1. Лічильник років (Years Counting): Простий підхід, де оцінка базується на загальній кількості років роботи, вказаних у резюме кандидата.
2. Категоризація за типами досвіду (Experience Categorization): Метод, що дозволяє класифікувати досвід кандидатів за типами (наприклад, менеджерський, технічний, творчий) та оцінювати його відповідність вимогам вакансії.
3. Розроблена система кількісного аналізу (Developed Quantitative Analysis): Метод, застосований у нашій системі, який включає оцінку не тільки загальної тривалості роботи, але й її специфіку, актуальність та релевантність до конкретної вакансії.

| Критерій /<br>Метод                         | Лічильник<br>років | Категоризація<br>за типами досвіду | Розроблена<br>система<br>кількісного<br>аналізу |
|---------------------------------------------|--------------------|------------------------------------|-------------------------------------------------|
| Точність<br>оцінки                          | Низька             | Середня                            | Висока                                          |
| Врахування<br>специфіки досвіду             | Ні                 | Частково                           | Так                                             |
| Адаптивність<br>до різних вакансій          | Низька             | Середня                            | Висока                                          |
| Врахування<br>актуальності<br>досвіду       | Ні                 | Частково                           | Так                                             |
| Швидкість<br>процесу                        | Швидка             | Середня                            | Швидка                                          |
| Вартість<br>впровадження та<br>експлуатації | Низька             | Середня                            | Середня                                         |

Таблиця 3 – Порівняльна таблиця методів оцінки досвіду кандидата

Розроблена система кількісного аналізу досвіду в нашій роботі вибрана через її здатність надавати глибокий і всебічний аналіз досвіду кандидата. Вона дозволяє не тільки оцінити загальну тривалість роботи, але й врахувати актуальність та релевантність досвіду до конкретної вакансії. Це забезпечує вищу

точність та об'єктивність відбору, а також дозволяє роботодавцям приймати більш виважені рішення щодо найму.

Завдяки глибокому аналізу та кількісній оцінці, метод, розроблений у нашій системі, перевершує альтернативні підходи з точки зору точності, адаптивності та швидкості. Це робить його важливим інструментом для сучасних роботодавців, які прагнуть оптимізувати процес відбору персоналу і знаходити найкращі таланти на ринку.

## 2.4 Висновки до розділу 2

У другому розділі магістерської дисертації було проведено детальний аналіз та обґрунтування методів оцінки відповідності кандидатів до вакансій, зокрема через формулювання цільової функції та вибір специфічних алгоритмів та моделей. Основні висновки з цього розділу представлені нижче:

1. Комплексний підхід до оцінки кандидатів: Цільова функція, яка була запропонована, об'єднує кілька ключових показників: навички, освіту, та досвід кандидатів. Це забезпечує багатовимірний аналіз та дозволяє оцінювати кандидатів з урахуванням різних аспектів їх професійного профілю.
2. Інтеграція сучасних технологій: Використання семантичних векторів та алгоритмів машинного навчання дозволяє системі ефективно обробляти великі об'єми текстових даних та забезпечує високу точність оцінки відповідності кандидатів до вакансій. Це підкреслює важливість інтеграції передових технологічних рішень у процесі рекрутингу.
3. Ефективність обраних алгоритмів та моделей: Вибрані алгоритми та моделі виявились ефективними для виконання задачі оцінки відповідності кандидатів. Їх обґрунтування базується на аналітичному порівнянні з іншими методами, демонструючи переваги у точності, швидкості та адаптивності.
4. Гнучкість та адаптивність системи: Система надає можливість користувачам кастомізувати процес оцінки через використання вагових коефіцієнтів, що робить її гнучкою та адаптивною до специфічних вимог кожної окремої вакансії.
5. Перспективність запропонованої системи: запропонована система відповідності кандидатів відкриває нові перспективи для підвищення ефективності процесів рекрутингу. Вона демонструє значні переваги перед

традиційними підходами, надаючи обґрунтовану основу для подальшого розвитку та оптимізації.

Ці висновки підкреслюють значення розробленої системи для сучасного рекрутингу та відзначають важливість продовження досліджень і розробок у цій сфері для досягнення ще більшої точності та ефективності відбору персоналу.

### 3. РЕАЛІЗАЦІЯ СПОСОБУ ОЦІНКИ ВІДПОВІДНОСТІ КАНДИДАТА ДО ВАКАНСІЇ

#### 3.1 Вибір технологій та засобів реалізації системи оцінки відповідності.

У розробці системи оцінки відповідності кандидата до вакансії важливим є вибір технологій, що забезпечує високу ефективність, надійність, та можливість адаптації під змінювані вимоги ринку. Розгляд та аналіз використаних технологічних рішень проведено на основі їх відповідності до заданих критеріїв та цілей дослідження.

Технології, які були задіяні у розробці системи:

1. Python - основна мова програмування для розробки.
2. Pandas - бібліотека, що забезпечує швидкі, гнучкі та виразні структури даних, призначені для роботи з "реляційними" чи "маркованими" даними, що є інтуїтивно зрозумілими та легкими в використанні.
3. Numpy - бібліотека для обробки та аналізу великих масивів даних.
4. OpenAI API - для використання семантичних векторів та аналізу текстових даних.
5. Scikit-learn - для використання алгоритмів машинного навчання та статистичного аналізу.
6. cosine\_similarity - функція з бібліотеки Scikit-learn для обчислення косинусної схожості між векторами.
7. NLTK (Natural Language Toolkit) - ведуча платформа для побудови програм Python для роботи з людською мовою, що надає інтерфейси до понад 50 корпусів та лексичних ресурсів, таких як WordNet, а також набір текстових обробних бібліотек для класифікації, токенизації, стемінгу, тегування, синтаксичного розбору та семантичного розуміння.
8. H2O Wave - фреймворк для швидкого створення інтерактивних веб-додатків з можливістю машинного навчання та передового аналізу даних. Особливо корисний для створення прогресивних інтерфейсів та інтеграції моделей машинного навчання.

Далі здійснено ґрунтовний аналіз та обґрунтування вибору кожної з технологій, задіяних у розробці системи.

#### *Основна мова програмування Python*

Python було обрано як основу для реалізації системи, оскільки він є широко відомим у сферах наукових досліджень, машинного навчання, а також розробки програмного забезпечення. Висока універсальність та читабельність коду, а також багата екосистема бібліотек, роблять Python надійним і гнучким інструментом для комплексного аналізу даних та розробки веб-додатків.

Переваги Python як основної мови програмування для реалізації системи оцінки відповідності кандидата до вакансії впливають із його широкої підтримки бібліотек і фреймворків, що спеціалізуються на машинному навчанні, обробці даних та розробці веб-додатків. Python відзначається своєю універсальністю та високим рівнем читабельності коду, що сприяє швидкій та зручній розробці програмного забезпечення. Він пропонує могутній інструментарій для аналітиків даних, розробників, та науковців, включаючи такі відомі бібліотеки, як NumPy для обробки великих масивів даних, Pandas для аналізу даних та Matplotlib для візуалізації.

Мова програмування Python володіє багатою екосистемою, яка включає в себе різноманітні інструменти та бібліотеки для розробки як простих, так і складних рішень в області обробки та аналізу даних. Це забезпечує високий рівень гнучкості та можливість адаптувати систему під конкретні потреби та вимоги. Широка підтримка спільноти та наявність готових рішень та модулів значно скорочують час на розробку та впровадження системи, забезпечуючи при цьому її надійність та ефективність.

Завдяки цим перевагам, Python став ключовим інструментом у багатьох інноваційних проектах, у тому числі і в розробці системи оцінки відповідності кандидатів до вакансій.

#### *Numpy & Pandas*

- Використання: Для ефективної роботи з числовими даними, статистичного аналізу та обробки великих масивів даних.
- Переваги: Оптимізація обчислень, великий набір інструментів для аналізу та маніпуляцій з даними.

Pandas використовується для ефективної обробки даних, зокрема для читання файлів CSV і застосування перетворень. Його структура DataFrame забезпечує зручний спосіб зберігання табличних даних і керування ними.

NumPy використовується для числових операцій, особливо під час роботи з масивами, що важливо для підготовки даних та подальшого їх аналізу.

NLTK (Natural Language Toolkit) є відкритим програмним пакетом, розробленим для обробки природної мови (Natural Language Processing, NLP) у Python. Цей інструментарій використовується у наукових та дослідницьких колах для аналізу, обробки та розуміння як письмових, так і усних форм людської мови.

Основні характеристики:

NLTK надає широкий спектр модулів та інструментів, включаючи токенізатори, тегери частин мови, синтаксичні аналізатори, класифікатори тексту та інше.

Основна функція NLTK - обробка тексту, яка включає токенізацію (розділення тексту на слова чи фрази), стемінг (видалення афіксів для знаходження основи слова), лематизацію (приведення слова до його базової форми) та видалення стоп-слів (виключення загальновживаних слів, які мають мало семантичного значення).

NLTK включає багатий набір корпусів та лексичних ресурсів, таких як WordNet, що є незамінними для лінгвістичних досліджень та розробки NLP-застосунків.

Бібліотека легко інтегрується з іншими популярними бібліотеками для обробки даних у Python, такими як NumPy та Pandas, та дозволяє налаштовувати та розширювати свої можливості для специфічних дослідницьких потреб.

В науковому контексті, NLTK використовується для проведення комплексних досліджень у галузі мовної обробки. Він дозволяє аналізувати мовні структури, вивчати лексичні зв'язки, визначати емоційне забарвлення тексту та розвивати системи, здатні виконувати складні мовні завдання, як-от автоматичний переклад, автоматизоване відповідання на запитання та семантичний аналіз.

У контексті наукової роботи, пов'язаної з оцінкою відповідності кандидатів до вакансій, Natural Language Toolkit (NLTK) використовується як інструмент для передової обробки текстових даних.

1. Токенізація та очищення тексту: Перед аналізом текстових даних, таких як резюме та описи вакансій, важливо перетворити великі блоки тексту на зручні для аналізу структури, такі як слова або фрази. NLTK дозволяє ефективно виконувати токенізацію, тобто поділ тексту на окремі слова або фрази, а також очищення тексту від зайвих символів, чисел чи іншого шуму.
2. Стемінг та лематизація: Для підвищення точності аналізу та порівняння текстів важливо звести слова до їх основної форми. NLTK включає алгоритми для стемінгу (зведення слова до кореня) та лематизації

(повернення слова до його базової, словникової форми), що дозволяє більш точно аналізувати та порівнювати текстові дані.

3. Тегування частин мови (POS tagging): Розпізнавання частин мови (іменник, дієслово, прикметник та ін.) може бути корисним для глибшого розуміння структури і смислу текстів. Така інформація може бути використана для вдосконалення алгоритмів оцінки відповідності, особливо при аналізі вимог до навичок у вакансіях.

*OpenAI API - для використання семантичних векторів та аналізу текстових даних.*

- Використання: Для інтеграції семантичних векторів та аналізу текстових даних.
- Переваги: Доступ до передових технологій обробки природної мови, швидка інтеграція та висока точність моделей.

У виборі технологічного рішення для аналізу текстових даних і використання семантичних векторів, вирішальним стало використання OpenAI API. Цей інструмент надає доступ до передових моделей обробки природної мови, зокрема, здатності генерувати семантичні вектори, які відображають глибоке розуміння текстового контенту. Вибір на користь OpenAI API обумовлений декількома ключовими аргументами, котрі відповідають вимогам наукової строгості та об'єктивності.

1. Перш за все, використання OpenAI API дозволяє скористатися останніми досягненнями у сфері машинного навчання без необхідності власної розробки складних моделей. Це забезпечує високу точність та надійність результатів, які постійно удосконалюються завдяки роботі спеціалістів OpenAI.
2. Другий аргумент полягає в економії часу та ресурсів. Розробка власних моделей глибокого навчання для обробки тексту є часозатратною та вимагає значних обчислювальних потужностей. Використання готового рішення від OpenAI дозволяє зосередитися на самій задачі оцінки відповідності, мінімізуючи необхідність в глибокому зануренні в деталі розробки та навчання моделей.
3. Третім аргументом є масштабованість та адаптивність. OpenAI API пропонує гнучкість у виборі моделей і можливість швидкої адаптації до змінних потреб користувача або специфіки задач. Це дозволяє системі оцінки залишатися актуальною та ефективною навіть у контексті постійно змінюваного ринку праці.

Таким чином, вибір OpenAI API як інструменту для інтеграції семантичних векторів та аналізу текстових даних у системі оцінки відповідності кандидата до вакансії є виправданим та обґрунтованим рішенням, яке спирається на переваги ефективності, надійності, та інноваційності, які надає це рішення.

Ці технології і бібліотеки вибрані з огляду на їх ефективність, спроможність до масштабування, велику підтримку спільноти та легкість інтеграції в різноманітні проекти, що забезпечує потужний і гнучкий інструментарій для створення сучасних систем оцінки відповідності кандидатів до вакансій.

### 3.2 Програмне забезпечення для оцінки відповідності кандидатів до вакансій.

У рамках розробки системи оцінки відповідності кандидатів до вакансій особлива увага приділяється програмному забезпеченню, яке є ключовим елементом у забезпеченні функціональності, ефективності та адаптивності процесу відбору. Програмне забезпечення розроблено з використанням сучасних технологій та передових практик, що забезпечують високу точність оцінки та оптимізацію робочих процесів.

Архітектура системи розроблена з урахуванням модульності, масштабованості та інтегрованості.

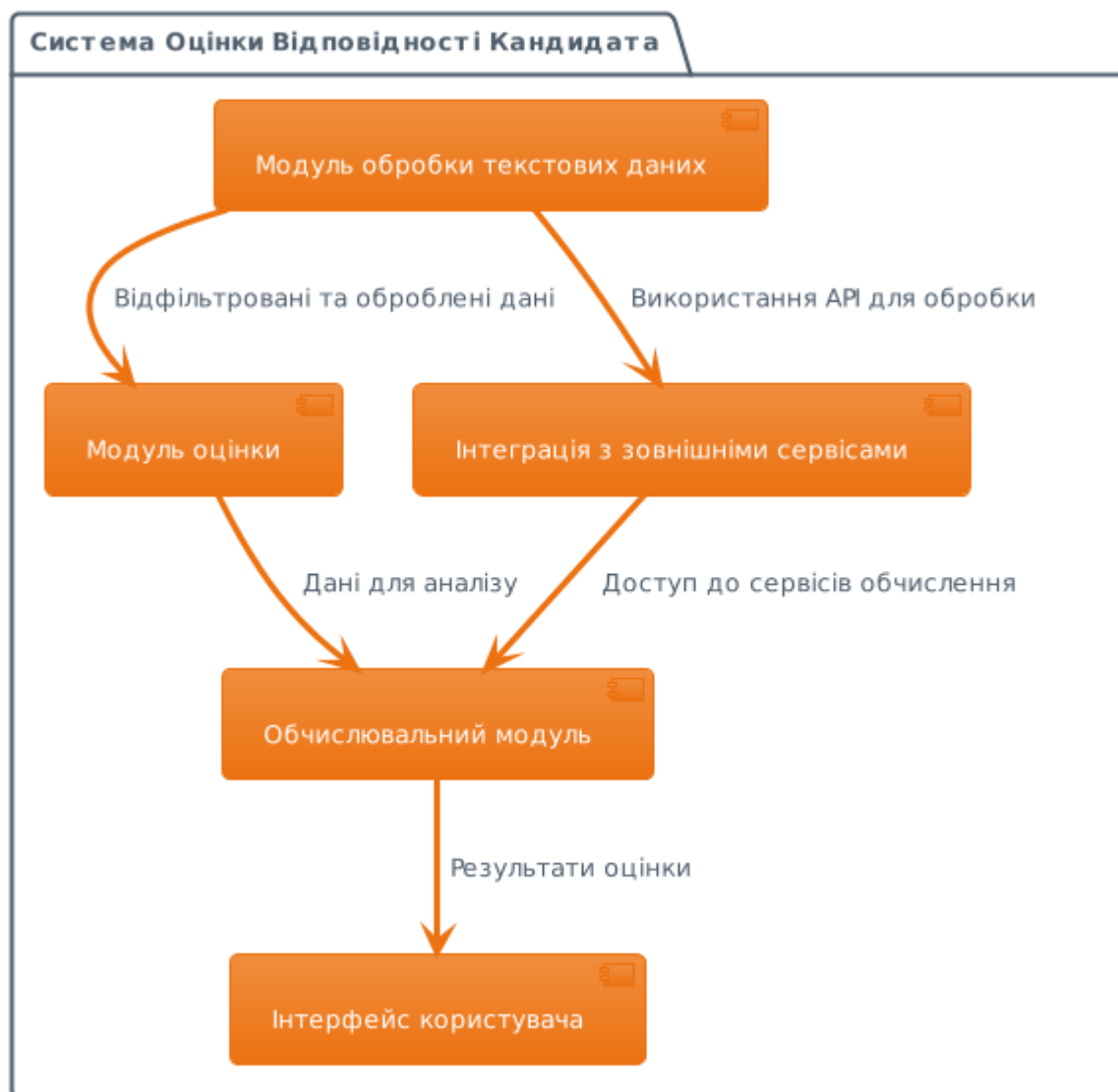


Рисунок 7 – Архітектура системи оцінки відповідності кандидата до вакансії

#### *Модуль обробки текстових даних*

Використовує OpenAI API для отримання семантичних векторів тексту.

Обробляє текстові дані кандидатів і вакансій, зокрема резюме та описи вакансій, для подальшого аналізу.

Завантаження та ініціалізація моделі (`load_model`) використовується модель `SentenceTransformer` з назвою `'all-MiniLM-L6-v2'`. Ця модель є варіантом трансформера, оптимізованим для згенерування векторних представлень речень (`embeddings`). Вона використовується для перетворення тексту вакансій та резюме в числові вектори, які потім можуть бути порівняні.

Обчислення схожості (`get_similarity_score`) функція використовує вищезазначену модель для перетворення тексту резюме та вакансій у вектори, а

потім розраховує косинусну схожість між ними. Це дозволяє оцінити, наскільки резюме відповідає вимогам вакансії.

Система використовує API від OpenAI для двох основних завдань: створення векторних представлень тексту та визначення схожості між текстовими даними.

Створення векторних представлень тексту:

- Система використовує OpenAI API для створення векторних представлень тексту з резюме кандидата та тексту робочого опису.
- Для цього викликається функція `openai.Embedding.create()`, передаючи текст, який потрібно векторизувати, та вказуючи модель, яка буде використана (у нашому випадку, "text-similarity-davinci-001").
- API повертає векторні представлення тексту у відповіді, які система потім використовує для обчислення схожості між текстами.

Визначення схожості між текстами:

- Після отримання векторних представлень тексту з резюме кандидата та тексту робочого опису система використовує косинусну схожість (`cosine_similarity`) для обчислення схожості між цими векторами.

Косинусна схожість - це метрика, яка визначає ступінь схожості між двома векторами у просторі. Вона вимірює кут між векторами та вказує на ступінь їхньої схожості (чим ближче значення до 1, тим схожіше текстові дані).

Загалом, OpenAI використовується для аналізу схожості тексту та отримання векторних представлень текстових даних, що допомагає в оцінці навичок та відповідності освіти та досвіду робочому опису.

### *Модуль оцінки*

Модуль оцінки є ключовою частиною системи оцінки відповідності кандидата до вакансії, що виконує комплексне оцінювання за трьома основними параметрами: навички, освіта та досвід. Кожен з цих параметрів аналізується за допомогою специфічних алгоритмів, орієнтованих на максимальну точність та об'єктивність оцінки.

1. Навички: Алгоритм оцінки навичок використовує семантичні вектори для порівняння текстів резюме кандидатів та описів вакансій. Він обчислює косинусну схожість, щоб визначити, наскільки набір навичок кандидата

відповідає вимогам роботи. Це дозволяє системі точно оцінювати відповідність кандидата на основі семантичного змісту текстів, забезпечуючи високу точність відбору.

2. Освіта: Модуль оцінки освіти зосереджений на аналізі рівня освіти кандидата. Він використовує просту числову шкалу для різних освітніх категорій, таких як диплом, ступінь бакалавра, магістра, аспірантури тощо. Кожен рівень освіти отримує певний числовий коефіцієнт, що дозволяє порівнювати освітній рівень кандидатів з вимогами, вказаними в описі вакансії.
3. Оцінка досвіду кандидатів проводиться через кількісний аналіз їхньої тривалості зайнятості. Методика оцінки враховує не лише загальну кількість місяців або років роботи, але й її релевантність до запропонованої вакансії.

#### *Модуль інтерфейсу користувача*

Інтерфейс користувача в реалізованій системі оцінки відповідності кандидатів до вакансій створений таким чином, щоб забезпечити зручність та інтуїтивність у використанні, дозволяючи користувачам ефективно управляти процесом оцінки та аналізувати результати. Основні характеристики та функції інтерфейсу користувача включають:

1. Введення даних: Інтерфейс дозволяє користувачам легко вводити або завантажувати дані про кандидатів та вакансії, включаючи резюме, описи вакансій та інші відповідні параметри.
2. Налаштування параметрів оцінки: Користувачі можуть налаштовувати різні параметри оцінки, такі як вагові коефіцієнти для навичок, освіти та досвіду, щоб краще адаптувати систему до специфічних вимог вакансії.
3. Показ результатів: Інтерфейс надає зрозуміле та зручне представлення результатів оцінки, включаючи загальні оцінки відповідності для кожного кандидата та детальніший аналіз їхніх навичок, освіти та досвіду.
4. Фільтрація та сортування: Користувачі можуть фільтрувати та сортувати кандидатів за різними критеріями, такими як загальна оцінка відповідності, навички, освіта або досвід, для полегшення процесу вибору.

Інтерфейс користувача розроблений таким чином, щоб спростити взаємодію з системою оцінки, зробити процес введення даних, налаштування параметрів та аналізу результатів максимально інтуїтивним та ефективним. Він служить ключовим зв'язком між користувачами та обчислювальною потужністю системи, забезпечуючи зручне та продуктивне використання всіх її можливостей.

### 3.3 Підготовка та обробка даних для оцінки ефективності системи.

У рамках розробки системи оцінки відповідності кандидата до вакансії важливо забезпечити ефективну підготовку та обробку даних. Цей процес включає збір, очищення, трансформацію, а також аналітичну обробку даних. Такі дії дозволяють визначити релевантність кандидатів з максимальною точністю, швидкістю та об'єктивністю.

#### *Збір та очищення даних*

**Збір даних:** Для аналізу використовуються датасети з резюме кандидатів та описами вакансій. Ці дані зазвичай збираються з різноманітних джерел, включаючи публічні бази даних, ресурси з пошуку роботи, а також внутрішні корпоративні бази.

**Очищення даних:** Видаляються неповні або некоректні записи, нормалізується форматування, виправляються помилки. Це забезпечує високу якість вхідних даних для аналізу.

#### *Трансформація та обробка даних для системи оцінки відповідності кандидата*

Трансформація даних – це процес перетворення зібраних текстових даних в структурований формат, який може бути ефективно аналізований. Для цього використовуються наступні методи обробки природної мови (NLP):

1. Токенізація: Розбиття тексту на окремі слова або фрази (токени), що є основою для подальшого аналізу.
2. Лемматизація: Приведення слів до їх базової форми (леми), що забезпечує узагальнення та зменшення варіативності тексту.
3. Вилучення ключових слів та фраз: Визначення та виділення найбільш значущих термінів, які можуть вказувати на специфічні навички, кваліфікації або вимоги до вакансій.
4. Векторне представлення тексту: Перетворення тексту в числові вектори. У контексті реалізованої системи оцінки відповідності кандидата до вакансії, векторне представлення тексту здійснюється шляхом використання семантичних векторів, які надаються через OpenAI API. Цей підхід включає:
  - a. Взаємодія з OpenAI API: Для кожного тексту резюме та опису вакансії здійснюється запит до API для отримання векторного представлення.
  - b. Семантичні вектори: Замість широкого спектра моделей як TF-IDF, Word2Vec або BERT, використовується конкретна технологія, що надається OpenAI, для отримання високоякісних векторних представлень тексту. Це дозволяє аналізувати та порівнювати текст на глибокому, семантичному рівні.

- с. Косинусна схожість: Для визначення ступеня відповідності між текстами резюме та описами вакансій використовується косинусна схожість векторів. Це дозволяє обчислити міру схожості та вибрати найбільш відповідних кандидатів.

Ці процеси є ключовими для створення ефективної та точної системи оцінки, що забезпечує об'єктивний та вимірюваний підхід до відбору кандидатів. Застосування цих методів дозволяє системі швидко та точно обробляти великі обсяги даних, виявляючи найбільш релевантних кандидатів для кожної вакансії.

| Функція                        | Опис                                                                                                                                                                                                                  |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>load_model()</code>      | Завантажує модель для векторизації тексту з використанням Sentence Transformers.                                                                                                                                      |
| <code>load_llm()</code>        | Ініціалізує модель мовленнєвого чатбота ChatOpenAI, яка використовує GPT-3.5 Turbo від OpenAI для чат-взаємодії.                                                                                                      |
| <code>load_js_data()</code>    | Завантажує дані з файлу <code>jds.csv</code> , які містять інформацію про вакансії (Job Descriptions).                                                                                                                |
| <code>load_cv_data()</code>    | Завантажує дані з файлу <code>cv_extractedV3.csv</code> , які містять інформацію про резюме кандидатів.                                                                                                               |
| <code>load_short_list()</code> | Завантажує дані з файлу <code>shortlist.csv</code> , які містять інформацію про кандидатів, які були відібрані для конкретних вакансій.                                                                               |
| <code>save_df(df, db)</code>   | Зберігає оновлені дані у файл CSV. <code>df</code> - це DataFrame, який необхідно зберегти, а <code>db</code> - це роздільник, що вказує, чи це дані про вакансії ( <code>jd</code> ) або резюме ( <code>cv</code> ). |

|                                                   |                                                                                                                            |
|---------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| <code>process_jd(q, jd_info)</code>               | Обробляє інформацію про вакансію, витягає ключові дані, такі як навички, освіта та досвід, та додає їх до списку вакансій. |
| <code>get_similarity_score(model, cvs, JD)</code> | Рахує косинусну подібність між вакансією (JD) та резюме кандидатів (cvs) з використанням моделі векторизації тексту.       |
| <code>shortlist_cvs(q, JD_dict, k=10)</code>      | Відбирає топ-резюме, які найбільше відповідають конкретній вакансії, на основі схожості та інших оцінок.                   |
| <code>update_short_list(q, df)</code>             | Оновлює список відібраних кандидатів для конкретної вакансії.                                                              |
| <code>download_jd(q)</code>                       | Завантажує вакансію з файлу та повертає інформацію про неї.                                                                |
| <code>get_jobs(q)</code>                          | Повертає список доступних вакансій та кількість вакансій для кожної з них.                                                 |
| <code>get_job_content(q, job_name)</code>         | Повертає вміст конкретної вакансії на основі її назви.                                                                     |
| <code>get_job_shortlist(q, job_name)</code>       | Повертає список відібраних кандидатів для конкретної вакансії та їх схожість з вакансією.                                  |
| <code>get_job_detail(q, job_name)</code>          | Повертає деталі конкретної вакансії на основі її назви.                                                                    |
| <code>get_cv_details(q, cv_id)</code>             | Повертає деталі конкретного резюме на основі його ідентифікатора.                                                          |

Таблиця 4 - Основні функції розробленої системи

В програмі реалізовані алгоритми оцінювання резюме кандидата (CV) на основі робочого опису (JD) за трьома основними категоріями: навичками (skills), освітою (education), та досвідом роботи (work experience).

#### Оцінка навичок (Skills):

- Внутрішня система видобутку інформації з резюме кандидата, виділяє навички (skills) з тексту.
- Потім використовується модель від OpenAI для обчислення векторних представлень тексту (embeddings) навичок з резюме кандидата та тексту робочого опису.
- Використовуючи косинусну подібність (cosine similarity) між векторами, обчислюється схожість між кандидатськими навичками та навичками, які вимагаються у робочому описі.
- Середнє значення схожості використовується як оцінка навичок.

#### Оцінка освіти (Education):

- Використовується внутрішня система видобутку інформації для визначення освіти з резюме кандидата.
- Перевіряється, чи відповідає освіта кандидата вимогам, зазначеним у робочому описі.
- Оцінюється освіта за власною шкалою, де різні рівні освіти мають ваги.

#### Оцінка досвіду роботи (Work Experience):

- Використовується внутрішня система видобутку інформації для визначення освіти з резюме кандидата.
- Переводяться різні формати досвіду роботи в кількість місяців.
- Далі враховується різниця між досвідом роботи кандидата та вимогами, зазначеними у робочому описі.
- Застосовується математичну функцію для трансформації цієї різниці в оцінку.

В кожній з цих категорій обчислюється окремий бал на основі схожості, відповідності вимогам та інших факторів, і потім обчислюється загальний бал як середнє значення цих балів, призначаючи вагу кожній категорії за необхідності.

### 3.4 Аналіз та порівняння результатів з традиційними методами відбору персоналу.

У сфері рекрутингу постійно відбувається пошук більш ефективних методів відбору персоналу, що зумовлює необхідність порівняльного аналізу традиційних підходів із сучасними автоматизованими системами. Розроблена система, яка базується на використанні алгоритмів машинного навчання та семантичного аналізу, пропонує інноваційний підхід до оцінки кандидатів. Основними перевагами цього методу є точність, швидкість та об'єктивність, що робить його важливим інструментом у пошуку найбільш підходящих працівників.

#### *Запропонований спосіб оцінки відповідності кандидата до вакансії*

##### Переваги:

1. Точність: Автоматизація процесу та використання алгоритмів машинного навчання забезпечують високу точність оцінки кандидатів.
2. Швидкість: Система може швидко обробляти великі обсяги даних, значно скорочуючи час відбору кандидатів.
3. Об'єктивність: Автоматизований підхід зменшує суб'єктивний вплив оцінювача, забезпечуючи більш об'єктивну оцінку кандидатів.
4. Адаптивність: Система легко адаптується під різні вимоги вакансій та може бути налаштована згідно специфіки організації.

##### Недоліки:

1. Вартість впровадження: Розробка та інтеграція комплексної системи може бути дорогавартісною, особливо на початкових етапах.
2. Залежність від даних: Ефективність системи значною мірою залежить від якості та повноти вхідних даних.
3. Потреба у технічному обслуговуванні: Система вимагає регулярного технічного обслуговування та оновлення для забезпечення її стабільної та ефективної роботи.

#### *Традиційні методи відбору персоналу*

##### Переваги:

1. Особистий контакт: Традиційні методи часто включають особисті співбесіди, що дозволяє оцінити не тільки професійні якості кандидата, але й його комунікативні навички та загальну поведінку.
2. Гнучкість: Людський фактор дає можливість гнучко реагувати на різноманітні ситуації, які можуть виникнути під час співбесіди або аналізу резюме.
3. Досвід HR-спеціалістів: Багаторічний досвід спеціалістів може допомагати у виявленні потенційно успішних кандидатів за допомогою інтуїції та неформального оцінювання.

Недоліки:

1. Суб'єктивність: Рішення, прийняті на основі особистих переконань чи вражень, можуть бути суб'єктивними та не завжди об'єктивно відображати кваліфікацію кандидата.
2. Масштабованість: Традиційні методи часто є часозатратними і не підходять для швидкої обробки великої кількості кандидатів.
3. Висока вартість: Ручний аналіз та співбесіди вимагають значних часових та фінансових витрат.

Для об'єктивної оцінки ефективності розробленої системи проведено порівняльний аналіз з такими традиційними методами відбору, як ручний аналіз резюме, інтерв'ювання та використання рекомендацій. Результати аналізу представлені у таблиці нижче, яка включає критерії порівняння, такі як точність відбору, швидкість процесу, об'єктивність оцінки та здатність до адаптації під різноманітні вимоги вакансій.

| Критерій              | Традиційні методи | Запропонований спосіб |
|-----------------------|-------------------|-----------------------|
| Точність відбору      | Середня           | Висока                |
| Швидкість процесу     | Повільна          | Швидка                |
| Об'єктивність оцінки  | Низька            | Висока                |
| Адаптивність          | Низька            | Висока                |
| Вартість впровадження | Змінна            | Середня               |

Таблиця 5 – Порівняльна таблиця традиційних методів оцінки відповідності кандидата та запропонованого способу

Традиційні методи часто залежать від суб'єктивного судження HR-спеціалістів та обмежені у швидкості обробки великої кількості даних. В той

час як розроблена система здатна швидко аналізувати великі масиви інформації, забезпечуючи високу точність та об'єктивність оцінок. Адаптивність системи до специфічних потреб вакансій дозволяє враховувати різноманіття вимог роботодавців, забезпечуючи більш точний підбір кандидатів. Вартість впровадження системи є середньою, але потенційна вигода від підвищення якості та швидкості відбору значно перевищує початкові витрати.

На основі проведеного аналізу можна зробити висновок, що розроблена система є більш ефективним інструментом порівняно з традиційними методами відбору персоналу. Вона забезпечує високу точність та швидкість процесу, що є особливо важливим у сучасних умовах динамічного ринку праці.

### 3.5 Висновки до розділу 3

У процесі аналізу та реалізації способу оцінки відповідності кандидата до вакансії було виявлено ключові аспекти, які впливають на ефективність та адекватність оцінювання. Основна увага зосереджена на виборі технологій, створенні програмного забезпечення, обробці даних та аналізі ефективності системи в порівнянні з традиційними методами. Розглянемо важливі висновки, до яких можна прийти на основі проведеного дослідження.

Вибір технологій було здійснено на основі їх ефективності, доступності та підтримки спільноти. Використання Python, OpenAI API, Numpy, Pandas та інших інструментів дозволило створити гнучку та масштабовану систему, здатну ефективно обробляти та аналізувати великі масиви даних. Однак важливо підкреслити необхідність постійного оновлення та підтримки використовуваних технологій, аби забезпечити безперервну роботу та адаптацію системи до змінюваних умов ринку праці.

Створене програмне забезпечення відзначається модульною архітектурою та інтуїтивно зрозумілим інтерфейсом користувача. Особлива увага була приділена алгоритмам оцінки, що включають різні аспекти кандидата, такі як навички, освіта, та досвід. Програмне забезпечення забезпечує не лише кількісну оцінку відповідності, але й можливість гнучкого налаштування параметрів оцінки з урахуванням специфіки вакансій.

Ефективність системи значно залежить від якості та повноти даних, що використовуються для оцінки. Важливим аспектом стала розробка методів збору, очищення та трансформації даних, забезпечуючи їхню актуальність та відповідність. Інтеграція з зовнішніми джерелами дозволила розширити базу даних кандидатів та вакансій, збільшуючи таким чином шанси на успішний підбір.

Порівняння розробленої системи з традиційними методами відбору показало значні переваги в частині швидкості, точності та об'єктивності оцінки.

Автоматизація процесу дозволяє обробляти значно більшу кількість кандидатів за короткий час, зменшуючи витрати та підвищуючи ефективність відбору. Незважаючи на певні недоліки, такі як залежність від якості даних та потреба в технічному обслуговуванні, система демонструє високий потенціал для використання у сучасних HR-процесах.

#### 4. ВАЛІДАЦІЯ ТА ОЦІНКА РОЗРОБЛЕНОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

##### 4.1 Валідація та тестування розробленого програмного забезпечення.

Кожен тестовий приклад призначений для оцінки різних аспектів системи, від базової функціональності та обробки помилок до продуктивності та упередженості.

| Тест               | Мета                            | Вхідні дані | Очікуваний результат          | Фактичний результат           | Статус   |
|--------------------|---------------------------------|-------------|-------------------------------|-------------------------------|----------|
| Додавання вакансії | Задати нову вакансію і критерії | Критерії    | Нову вакансію створено        | Нову вакансію створено        | Пройдено |
| Оцінка резюме      | Отримання оцінки кандидата      | Резюме      | Програма дає оцінку кандидата | Програма дає оцінку кандидата | Пройдено |

Таблиця 6 - Тестування основних функцій

Ефективне тестування гарантує, що програма є не тільки технічно обґрунтованою, але й справедливою, надійною та відповідає реальним потребам підбору персоналу.

## Add a Job Description

×

**Job title**

**Vacancies count**

**Upload**  

Browse

**Skill Weight**  

1

**Education Weight**  

0.5

**Experience Weight**  

1

Submit

Рисунок 8 - Додавання опису вакансії

На рисунку 8 відображено модальне вікно з формою для додавання опису робочого місця в програму. Вікно містить наступні поля та елементи управління:

- **Job title:** Текстове поле для введення назви робочої позиції.
- **Vacancies count:** Текстове поле для вказівки кількості вакансій.
- **Upload:** Кнопка "Browse" для завантаження файлу з описом вакансії. Наразі підтримуються файли в форматі .txt.
- **Skill Weight:** Слайдер для налаштування ваги навичок, зі шкалою від 0 до 1.
- **Education Weight:** Слайдер для налаштування ваги освіти, також зі шкалою від 0 до 1.

- Experience Weight: Слайдер для налаштування ваги досвіду, аналогічно зі шкалою від 0 до 1.
- Submit: Кнопка для подання форми.

Користувач може використовувати цю форму для додавання нових вакансій до системи, вказавши назву позиції, кількість доступних місць, завантаживши відповідні документи та налаштувавши важливість різних аспектів кандидатури, таких як навички, освіта та досвід. Слайдери дозволяють встановити релевантність кожного з цих критеріїв, що може бути використано для відбору кандидатів або аналізу поданих резюме.

Після того, як користувач заповнить форму додавання вакансії та натисне кнопку "Submit", інформація з цієї форми розпаршується і записується в CSV файл.

| job                  | content                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | vacancies | skills                                                                                                                                                                                                                                                                                                                                                                                                       | education | experience |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------------|
| Full Stack Developer | <p>Joining our R&amp;D team might be one of the best ways to develop new skills as a software engineer. We work in small teams with simple yet efficient processes that include peer code review. You will be expected to take projects end-to-end, from the discovery and estimation phase to deployment. Our clients vary from small businesses needing a tool to get things done or develop a small application to large enterprises. LeanyLabs is a team of professionals, so we have near-zero bureaucracy and heavily in favor of automation.</p> <p><b>Requirements</b></p> <ul style="list-style-type: none"> <li>- 5+ years of experience with full-stack JavaScript/TypeScript (Node.js/React)</li> <li>- Willingness to take ownership of own work and the entire project</li> <li>- Experience with relational databases</li> <li>- Experience with AWS and Docker</li> <li>- Experience starting projects from scratch and deploying them to production</li> <li>- Upper-intermediate English level</li> </ul> <p>Will be a bonus</p> <ul style="list-style-type: none"> <li>- Leadership experience</li> <li>- An urge for direct communication with the clients</li> <li>- Experience with PostgreSQL and Sequelize or another ORM</li> <li>- AWS knowledge (we use ECS/ECR/ALB/S3/CloudFront/Lambda/SNS/SQS)</li> <li>- Understanding of TypeScript strengths and weaknesses</li> </ul> <p>Your day-to-day responsibilities will include</p> <ul style="list-style-type: none"> <li>- Engaging in all aspects of release cycles, including requirements review, work estimation, and deployment</li> <li>- Designing application architecture</li> <li>- Designing and developing REST APIs</li> <li>- Writing clean, testable, and maintainable code</li> <li>- Solving challenging tasks</li> <li>- Regular communication with the client</li> <li>- Mentoring other developers</li> <li>- Peer-to-peer code reviews</li> <li>- Occasionally participating in pre-sales</li> <li>- Leading the development process when needed</li> </ul> | 1         | <p>[</p> <ul style="list-style-type: none"> <li>"JavaScript/TypeScript",</li> <li>"Node.js",</li> <li>"React",</li> <li>"Relational databases",</li> <li>"AWS",</li> <li>"Docker",</li> <li>"PostgreSQL",</li> <li>"Sequelize",</li> <li>"ORM",</li> <li>"ECS",</li> <li>"ECR",</li> <li>"ALB",</li> <li>"S3",</li> <li>"CloudFront",</li> <li>"Lambda",</li> <li>"SNS",</li> <li>"SQS",</li> </ul> <p>]</p> | degree    | 5          |

Рисунок 9 - Структура файлу JDs.csv

На Рисунку 9 показано, як дані вакансії представлені у CSV форматі, який включає наступні колонки: "job", "content", "vacancies", "skills", "education", "experience". У колонці "job" вказана назва позиції - "Full Stack Developer", у "content" - опис вакансії, "vacancies" містить кількість вакансій (у цьому випадку 1), у "skills" перелічені необхідні технічні навички, "education" містить вимоги до освіти, а "experience" - необхідний досвід роботи. У цьому файлі кожна вакансія представлена рядком з відповідними даними, що дозволяє ефективно керувати вакансіями та проводити аналіз даних.

Для аналізу відповідності резюме кандидатів до вакансії, програма використовує введені дані з CSV файлів. Процес аналізу проходить наступним чином:

- Завантаження резюме: Користувач завантажує опис резюме кандидата у CSV файл, який містить різні поля, такі як ім'я, вміст (опис), навички, освіта, досвід, тощо.
- Розпарсинг даних: Після завантаження опису резюме, програма автоматично розпарсить текст на визначені критерії, як-от навички, освіта та досвід роботи, розділяючи інформацію на відповідні колонки у CSV файлі.
- Аналіз відповідності: На основі розпарсених даних програма може аналізувати, наскільки резюме кандидата відповідає вимогам вакансії, зважаючи на вказані вагові коефіцієнти для навичок, освіти та досвіду.
- Результати аналізу: Після аналізу програма надає оцінку щодо кожного кандидата, формуючи файл shortlist.csv допомагаючи рекрутеру або менеджеру по HR визначити найбільш підходящих кандидатів.

| id | name      | content                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | skills | education | experience | personal-content | skill-content | education-content | exp-content |
|----|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|-----------|------------|------------------|---------------|-------------------|-------------|
| 0  | Ivan      | <p>Experience</p> <p>Nothing is impossible for Me</p> <p>I am a highly skilled and experienced full-stack developer with a passion for creating efficient and user-friendly. With over 5 years of professional experience, I have worked on numerous projects, ranging from small startups to large enterprises. My expertise lies in both front-end and back-end development, allowing me to seamlessly integrate various technologies. I am a quick learner, adaptable to new technologies, and always strive to stay up-to-date with the latest industry trends.</p> <p>Programming Languages: JavaScript, Python, PHP, HTML5, CSS3, SQL</p> <p>Front-End Development: React.js, AngularJS, jQuery, CSS preprocessors (Sass, Less), Bootstrap</p> <p>Back-End Development: Node.js, Express.js, Django, Laravel</p> <p>Databases: MySQL, PostgreSQL, MongoDB</p> <p>Version Control: Git, GitHub, Bitbucket</p> <p>API Development: RESTful APIs, integration of third-party APIs</p> <p>Deployment and Infrastructure: AWS, Heroku, Docker</p> <p>Testing and Debugging: Jest, Mocha, Chai, Postman</p> <p>Agile Methodologies: Scrum, Kanban</p> <p>Skills</p> <p>JavaScript React REST API Node.js TypeScript Redux MongoDB Next.js Webpack HTML5 HTML Express.js</p> <p>Achievement</p> <p>Having 6+ years of Full-Stack Engineer experience I would like to apply for this position with 100% confidence</p> <p>Expectations from work</p> <p>I want to grow myself so that becomes the first in my workplace.</p> <p>Language of communication</p> <p>English</p> |        |           |            |                  |               |                   |             |
| 1  | Vladyslav | <p>Experience</p> <p>I'm Vladyslav, a full stack developer and software engineer with 10 years of development experience based in Ukraine.</p> <p>Skills</p> <p>HTML / CSS / JavaScript / TypeScript PHP/Laravel WordPress Themes WordPress Plugins WordPress API SASS</p> <p>Language of communication</p> <p>English</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |        |           |            |                  |               |                   |             |

Рисунок 10 - Структура файлу cv\_extracted.csv

На рисунку 10 показано приклад такого CSV файлу, де для кожного кандидата вказані його ім'я, текст резюме, перелік навичок, та інша інформація, яка може бути корисною для подальшої обробки та аналізу.

Спочатку користувач заповнює лише два поля у таблиці: "name", де вказує ім'я кандидата, та "content", де розміщує вміст резюме. Після цього програма



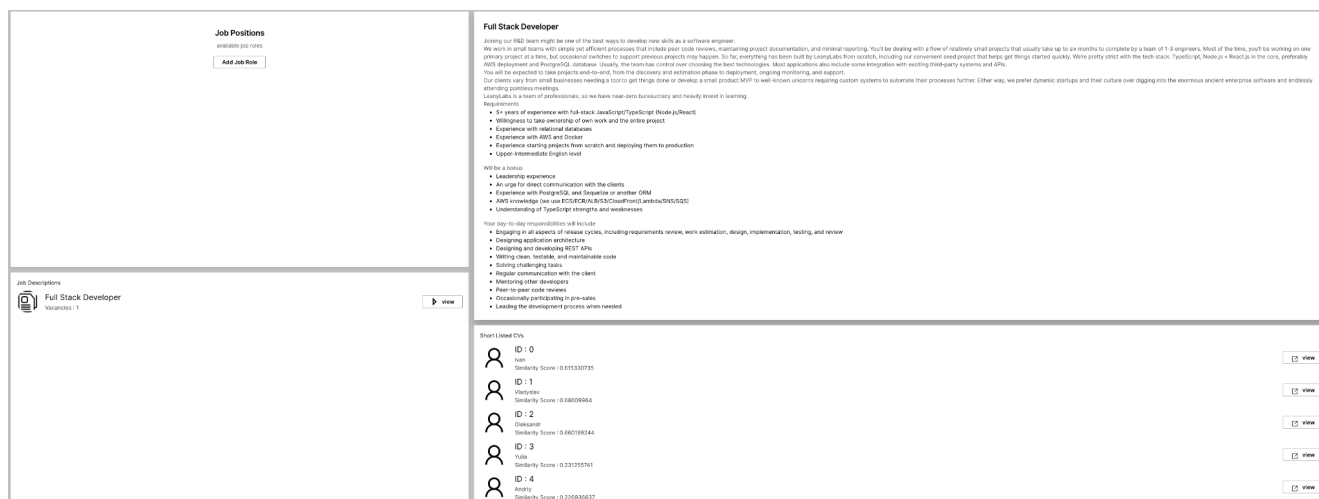


Рисунок 12 - Вигляд сторінки з вакансіями (JD\_Page)

У розділі "Job Positions" відображається наявність ролей, які наразі є відкритими для подання кандидатами, з опцією додати нову роль через кнопку "Add Job Role".

Детальний опис кожної вакансії включений у розділ "Job Descriptions", де користувач може знайти інформацію, таку як вимоги до кандидатів, бонуси за наявності певного досвіду, а також перелік очікувань від потенційних співробітників.

У нижній частині інтерфейсу, під назвою "Short Listed CVs", представлений список кандидатів з їх короткими резюме, ідентифікаційним номером (ID) та оцінкою схожості (Similarity Score). Ця оцінка може бути використана для визначення, наскільки профіль кандидата відповідає вимогам вакансії. Кнопка "view" поруч з кожним кандидатом дозволяє користувачу переглянути детальніше їх резюме та іншу інформацію.

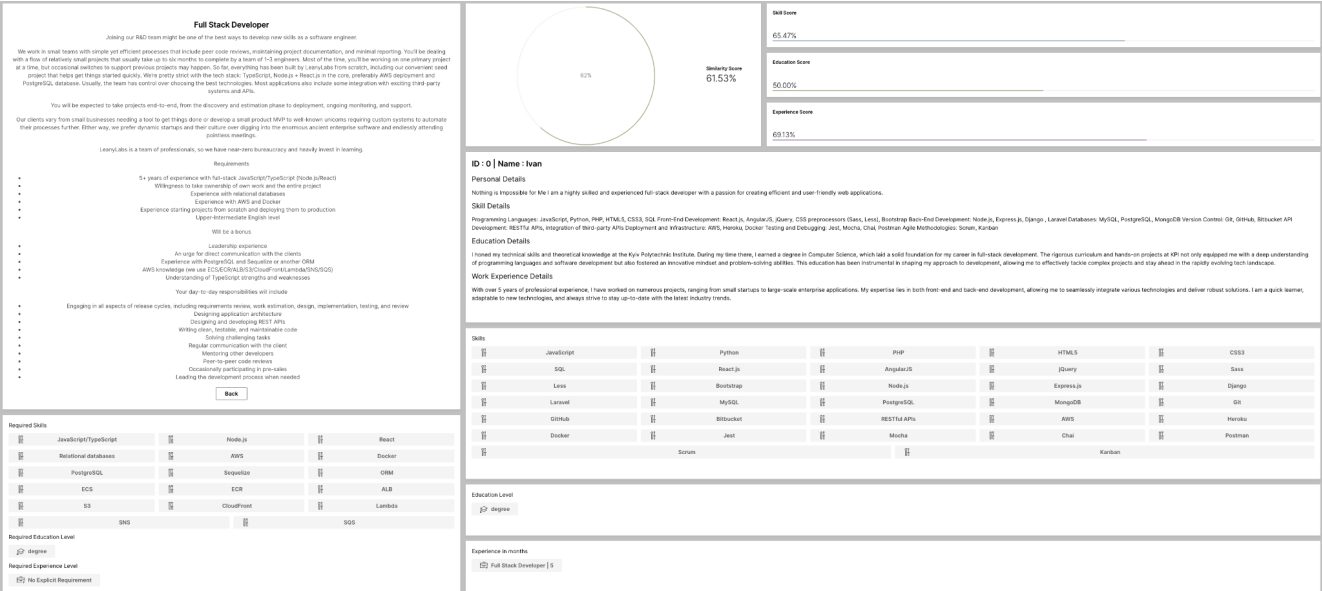


Рисунок 13 - Вигляд сторінки резюме кандидата (CV\_Page)

На зображенні 13 відображено сторінку cv\_page системи обробки резюме, яка інтегрує інформацію про вакансію та резюме кандидатів з оцінкою їх відповідності до зазначеної ролі.

В лівій верхній частині сторінки міститься детальний опис вакансії "Full Stack Developer", який включає зведену інформацію про робочі обов'язки, вимоги до кандидатів, а також додаткові умови, які будуть розглядатися як перевага. Також присутній перелік основних відповідальностей, які належать до цієї ролі.

В правій верхній частині відображено аналітичні діаграми, які представляють оцінку резюме кандидата на основі трьох критеріїв: навичок (Skill Score), освіти (Education Score) та досвіду (Experience Score). Кожен з цих параметрів представлений у вигляді відсоткового значення, що показує, наскільки профіль кандидата відповідає вимогам вакансії.

У нижній частині сторінки міститься інформація про кандидата з ідентифікаційним номером "0" та ім'ям "Ivan". Представлено розділи "Personal Details", "Skill Details", "Educational Details" та "Work Experience Details", які містять текстовий опис особистих даних, перелік навичок, інформацію про освіту та досвід роботи кандидата. Для кожного розділу вказано відповідні теги, які

допомагають системі в автоматизованому режимі аналізувати та визначати рівень відповідності між вимогами вакансії та резюме кандидата.

Відомості про резюме кандидата доповнюються списком "Required Skills", "Required Education Level" та "Required Experience Level", де вказано ключові навички, необхідний рівень освіти та досвід, який потрібен для вакансії. Ця сторінка є інструментом для оцінки та відбору кандидатів, надаючи користувачу зручний доступ до всієї необхідної інформації в єдиному інтерфейсі.

Для зручності є можливість перегляду оцінок відповідності по всім кандидатам через файл shortlist.csv

| id | name      | content    | skills                     | education | experience                         | target_job     | similarity_score | skill_score | edu_score      | exp_score      | total_score                                                          | personal-content | skill-content | education-content | exp-content                        |
|----|-----------|------------|----------------------------|-----------|------------------------------------|----------------|------------------|-------------|----------------|----------------|----------------------------------------------------------------------|------------------|---------------|-------------------|------------------------------------|
| 0  | Ivan      | Experience | ['JavaScript', 'Py degree  |           | ['Full Stack Deve Full Stack Devel | 0.547082305    | 0.065465049      | 0.5         | 0.691341716    | 0.293481804    | Nothing is Impos Programming La I honed my techi With over 5 year    |                  |               |                   |                                    |
| 1  | Vladyslav | Experience | ['HTML', 'CSS', ' degree   |           | ['Full Stack Deve Full Stack Devel | 0.530188203    | 0.070474553      | 0.5         | 0.853553391    | 0.338074118    | I'm Vladyslav, a Skills                                              |                  |               |                   | My journey as a full stack develop |
| 2  | Oleksandr | Oleksandr  | ['Python', 'Java', degree  |           | ['Full Stack Deve Full Stack Devel | 0.506143928    | 0.062704434      | 0.5         | 0.853553391    | 0.336002087    | Dedicated and e Programming La Master of Compi Senior Full Stadi     |                  |               |                   |                                    |
| 3  | Yulia     | Yulia      | ['Python', 'R', 'SC degree |           | ['Data Analyst   Full Stack Devel  | 0.056810759    | 0.023056546      | 0.5         | 0.1707106782   | 0.08443153525  | Resourceful data Programming: P Master of Scienc Data Analyst        |                  |               |                   |                                    |
| 4  | Andriy    | Andriy     | ['Python', 'R', 'SC degree |           | ['Data Analyst   Full Stack Devel  | 0.064926581714 | 0.025618384444   | 0.5         | 0.155191525636 | 0.076755941136 | Data Scientist wi Languages: Pyth Master's Degree pecialized in stal |                  |               |                   |                                    |

Рисунок 14 - Структура файлу shortlist.csv

На зображенні 14 представлено файл shortlist.csv, який використовується для перегляду та аналізу оцінок відповідності кандидатів до певної робочої позиції. Файл містить наступні колонки:

1. id: Унікальний ідентифікатор кандидата.
2. name: Ім'я кандидата.
3. content: Короткий опис досвіду або особисті дані кандидата.
4. skills: Список навичок кандидата, представлений у форматі масиву.
5. education: Освітній рівень кандидата.
6. experience: Специфікація досвіду роботи, зазначеного кандидатом.
7. target\_job: Позиція, на яку претендує кандидат.
8. similarity\_score: Загальна оцінка схожості резюме кандидата до опису вакансії.
9. skill\_score: Оцінка відповідності навичок кандидата вимогам вакансії.
10. edu\_score: Оцінка відповідності освіти кандидата вимогам вакансії.
11. exp\_score: Оцінка відповідності досвіду кандидата вимогам вакансії.
12. total\_score: Загальна оцінка кандидата.

13.personal-content: Додаткова інформація про кандидата.

14.skill-content: Детальний опис навичок кандидата.

15.education-content: Детальний опис освіти кандидата.

16.exp-content: Детальний опис досвіду роботи кандидата.

Файл shortlist.csv дозволяє HR-спеціалістам ефективно порівнювати кандидатів між собою, оцінюючи їх на основі зазначених метрик, та вибирати найбільш підходящих для вакансії. Оцінки відповідності допомагають швидко ідентифікувати кандидатів з найвищим потенціалом відповідно до визначених критеріїв вакансії.

Для верифікації функціональності та ефективності розробленого програмного забезпечення здійснено тестування на реальних даних. Вихідними матеріалами для тестування послужили опис вакансії, який було отримано з відомого ресурсу для пошуку роботи Djinni.co, та резюме потенційних кандидатів, відібрані у співпраці з рекрутерами. Для аналізу було вибрано п'ять резюме: три з них були визначені як відповідні вимогам вакансії, а два — як частково відповідні або невідповідні. Такий підхід дозволив забезпечити репрезентативність вибірки та перевірити здатність системи диференціювати кандидатів за ступенем їх придатності.

| Ім'я кандидата | Очікуваний результат від Рекрутера | Результат тестування | Similarity_Score |
|----------------|------------------------------------|----------------------|------------------|
| Ivan           | Задовольняє вимоги вакансії        | Пройдено             | 65,47%           |
| Vladyslav      | Задовольняє вимоги вакансії        | Пройдено             | 70,47%           |
| Oleksandr      | Задовольняє вимоги вакансії        | Пройдено             | 62,70%           |
| Yulia          | Не задовольняє вимоги вакансії     | Пройдено             | 23,06%           |
| Andriy         | Не задовольняє вимоги вакансії     | Пройдено             | 25,62%           |

Таблиця 7 - Результати тестування

Результати тестування підтвердили, що програмне забезпечення здатне ефективно обробляти резюме та класифікувати кандидатів відповідно до заданих параметрів. Отримані оцінки відповідності резюме критеріям вакансії відображали високий рівень точності, що у свою чергу засвідчило адекватність розробленої моделі оцінювання. Такі результати демонструють потенціал впровадження цього програмного забезпечення в реальні HR-процеси, забезпечуючи оптимізацію роботи рекрутерів та підвищення якості процесу відбору персоналу.

4.2 Оцінка практичної вартості та ефективності в контексті реальних використання. (переваги запропонованого способу)

Даний спосіб пропонує кілька практичних переваг і реальної цінності.

*Ефективність оцінювання кандидатів.* Традиційні методи відбору персоналу часто передбачають трудомісткі процеси, такі як перевірка резюме вручну та співбесіди. Запропонований метод автоматизує початковий відбір, значно скорочуючи час і зусилля, необхідні для оцінки кандидатів. Ця ефективність особливо цінна для великого набору персоналу.

*Об'єктивність і послідовність.* Метод застосовує конкретні ваги до атрибутів кандидатів на основі попередньо визначених критеріїв, забезпечуючи послідовні та справедливі оцінки. Це мінімізує потенційну упередженість у процесі перевірки, що призводить до більш справедливих рішень щодо найму.

*Настроюваність.* Організації можуть налаштовувати критерії та ваги відповідно до своїх конкретних потреб найму. Ця гнучкість дозволяє адаптувати процес перевірки до різних робочих ролей і галузей.

*Прийняття рішень на основі даних.* Використовуючи модель нейронної мережі, метод використовує інформацію, що керується даними, для оцінки придатності кандидата. Цей підхід враховує широкий спектр факторів, які можуть бути неочевидними під час традиційного скринінгу.

*Масштабованість.* Метод ефективно обробляє великі обсяги даних кандидатів, що робить його придатним для організацій із великою кількістю заявників. Це гарантує, що всі кандидати оцінюються на основі заздалегідь визначених критеріїв.

*Зменшення людських помилок.* Автоматизація зменшує ризик людських помилок під час перевірки вручну, зводячи до мінімуму шанси не помітити кваліфікованих кандидатів або просувати невідповідних.

*Постійне вдосконалення.* Метод можна точно налаштувати та вдосконалити з часом, коли стане доступним більше даних. Аналіз результатів найму дозволяє організаціям ще більше оптимізувати процес відбору.

*Інтеграція з існуючими системами.* Метод можна легко інтегрувати в існуючі системи управління персоналом, доповнюючи інші інструменти та процеси найму.

*Економія коштів.* Завдяки автоматизації початкового відбору та підвищенню точності відбору кандидатів організації зменшують витрати на підбір персоналу, пов'язані з перевіркою вручну та співбесідами.

*Оцінка придатності в режимі реального часу.* Метод надає оцінки придатності в режимі реального часу для кандидатів, дозволяючи командам

відділу кадрів визначати пріоритетність кандидатів для співбесід або оперативного подальшого оцінювання.

Загалом, запропонований метод пропонує практичність і ефективність, спрощує оцінку кандидатів, покращує якість найму та скорочує час і ресурси для найму. Його настроюваність, об'єктивність і масштабованість роблять його цінним для спеціалістів з управління персоналом у різних галузях і сценаріях найму.

#### 4.3 Плани щодо покращення існуючого функціоналу

У цьому підрозділі розглядаються потенційні шляхи покращення та розвитку системи оцінки відповідності кандидатів до вакансій на базі засобів машинного навчання. На основі аналізу отриманих результатів валідації та оцінки програмного забезпечення, було виявлено кілька ключових аспектів, які вимагають удосконалення для підвищення ефективності та корисності системи. Ці аспекти включають алгоритмічні покращення, розширення функціоналу, оптимізацію стабільності та масштабованості, інтеграцію з іншими системами, а також підвищення зручності користування.

Перш ніж перейти до обговорення конкретних планів покращення, слід зробити огляд поточного стану розробленої системи. Це включає оцінку сильних сторін, таких як висока точність класифікації та спроможність ефективно обробляти різноманітні дані кандидатів, а також слабких сторін, які можуть включати обмежену масштабованість або потребу в додаткових функціях для розширення можливостей системи.

На основі проведеного аналізу, одним із ключових напрямків для розвитку системи є покращення алгоритмічної бази. Це включає інтеграцію новітніх алгоритмів машинного навчання, які можуть підвищити точність та швидкість обробки даних. Особлива увага буде приділена алгоритмам глибокого навчання, які продемонстрували високу ефективність у сфері обробки природної мови та аналізу великих обсягів неструктурованих даних.

##### *Розширення функціоналу системи*

Для досягнення більш комплексного підходу в оцінці кандидатів, заплановане розширення функціоналу системи передбачає кілька ключових напрямків розвитку:

### 1. Автоматизація оновлення бази даних вакансій та резюме

- Ціль: Створення модулів, які автоматично оновлюють інформацію про вакансії та резюме, забезпечуючи актуальність даних для процесу оцінки.
- Реалізація: Інтеграція з веб-ресурсами та платформами пошуку роботи для збору оновлень про вакансії та кандидатів. Використання технологій веб-скрапінгу та API для збору та обробки даних.
- Переваги: Забезпечення оперативності та повноти інформації, необхідної для ефективної оцінки кандидатів.

### 2. Аналіз м'яких навичок та культурної відповідності

- Ціль: Врахування особистісних характеристик та культурної сумісності кандидатів з потенційним роботодавцем.
- Реалізація: Розробка алгоритмів NLP (Natural Language Processing) для витягування та аналізу відомостей з резюме, таких як міжособистісні навички, таких як комунікабельність, лідерські якості та здатність до командної роботи, на основі інформації з резюме та соціальних мереж.
- Переваги: Можливість більш глибокої оцінки кандидатів, не тільки за професійними якостями, але й за особистісними характеристиками.

### 3. Аналіз минулого досвіду роботи кандидатів в залежності від типу компанії

- Ціль: Визначити та аналізувати професійний шлях кандидатів, зокрема їхній досвід у різних типах компаній, для більш точної оцінки їх відповідності поточній вакансії.
- Реалізація: Розробка критеріїв для класифікації компаній. Спочатку необхідно визначити, які ключові характеристики розрізняють аутсорсингові та продуктові компанії (наприклад, тип проектів, ринок клієнтів, модель роботи).
- Переваги: Такий підхід дозволить більш всебічно оцінити професійні навички та потенціал кандидатів, забезпечуючи кращий підбір персоналу відповідно до специфіки роботи в компанії-роботодавця.

### *Забезпечення масштабованості та стабільності*

У відповідь на виклики, пов'язані з обробкою великих обсягів даних, запланована оптимізація архітектури програми ставить за мету підвищення масштабованості та стабільності системи. Особлива увага приділяється розробці ефективних механізмів для обробки великих даних, що надходять в реальному часі, та їх зберігання.

### Оптимізація архітектури для масштабованості

- Мікросервісна архітектура: Розробка системи на основі мікросервісної архітектури, яка дозволяє розподілити обробку даних та завантаження між

декількома серверами або сервісами. Це сприяє збільшенню гнучкості системи та її масштабованості.

- Використання розподілених баз даних: Застосування розподілених баз даних, таких як Cassandra або MongoDB, які оптимізовані для швидкого доступу та обробки великих обсягів даних. Це дозволить забезпечити високу швидкість читання та запису даних, а також їх надійне зберігання.
- Використання кешування: Імплементація кешування для часто запитуваних даних з метою зниження навантаження на базу даних та підвищення швидкості відповіді системи.
- Підвищення стабільності системи
- Автоматизоване масштабування: Впровадження автоматизованих механізмів масштабування, які забезпечують збільшення або зменшення ресурсів системи в залежності від поточного навантаження.
- Балансування навантаження: Реалізація балансувальників навантаження для рівномірного розподілу запитів між серверами або сервісами, що підвищить стабільність та доступність системи.
- Надійність та відмовостійкість: Запровадження стратегій відновлення після збоїв та резервного копіювання, щоб забезпечити високий рівень надійності та відмовостійкості системи.

Ці заходи дозволять створити масштабовану та стабільну систему, здатну ефективно обробляти великі потоки даних та забезпечувати високу якість обслуговування користувачів незалежно від обсягів запитів та даних.

### *Інтеграція з іншими системами*

З метою підвищення ефективності та гнучкості системи оцінки відповідності кандидатів до вакансій, планується розробка API (Application Programming Interface), який дозволить інтегрувати нашу систему з різноманітними зовнішніми HR-платформами та сервісами рекрутингу. Це важливий крок для забезпечення широкої сумісності та зручності використання, а також відкриває нові можливості для розвитку та розширення функціоналу.

#### *Реалізація API для інтеграції*

- Сумісність з Різними Платформами: API буде розроблено таким чином, щоб забезпечити легку інтеграцію з популярними HR-системами та платформами рекрутингу. Це включає можливість взаємодії з такими системами, як LinkedIn, Indeed, та іншими.
- Обмін Даними: API дозволить обмін даними між нашою системою та зовнішніми сервісами, включаючи передачу інформації про вакансії, кандидатів, результати оцінювання тощо.

- **Безпека та Конфіденційність:** При розробці API буде приділено особливу увагу питанням безпеки та конфіденційності даних. Всі взаємодії через API будуть відповідати стандартам захисту даних.
- **Гнучкість та Масштабованість:** API буде розроблено з урахуванням потреб масштабування, щоб він міг ефективно працювати як з невеликими компаніями, так і з великими корпораціями.

Переваги для Користувачів

- **Спрощення Процесу Рекрутингу:** Інтеграція з іншими HR-системами значно спростить процес залучення та оцінки кандидатів, дозволяючи HR-менеджерам використовувати одну уніфіковану платформу.
- **Автоматизація Робочих Процесів:** Автоматичний обмін даними між системами дозволить мінімізувати ручну роботу та підвищити продуктивність роботи рекрутерів.
- **Підвищення Ефективності Відбору:** Інтеграція з зовнішніми джерелами даних та інструментами оцінки забезпечить більш глибокий та всебічний аналіз кандидатів.

Розробка та впровадження API стане важливим кроком у напрямку розширення можливостей нашої системи та забезпечення її конкурентоспроможності на ринку HR-технологій.

#### 4.4 Висновки до розділу 4

У четвертому розділі ретельно розглянуто процес валідації та оцінки ефективності розробленого програмного забезпечення. Засновано на детальному тестуванні різних аспектів системи, цей розділ надає підтвердження її технічної надійності, об'єктивності, та придатності до реальних умов використання в контексті підбору персоналу. Основні висновки четвертого розділу можна сформулювати наступним чином:

1. **Ефективність оцінювання кандидатів:** Дослідження підтверджує, що автоматизований підхід до оцінювання кандидатів значно знижує час та зусилля, необхідні для первинного відбору, порівняно з традиційними методами. Це підтверджує ефективність та практичну значимість запропонованої системи.
2. **Об'єктивність та послідовність:** Система забезпечує об'єктивні та послідовні оцінки кандидатів, мінімізуючи упередженість та особистісні впливи, що є ключовою перевагою перед традиційними методами.
3. **Налаштовуваність та адаптивність:** Організації мають можливість налаштовувати критерії та ваги відповідно до своїх потреб, що робить

систему гнучкою та адаптивною до різноманітних умов та специфік вакансій.

4. Підтримка прийняття рішень на основі даних: Система використовує даних для глибокого аналізу та оцінки придатності кандидата, забезпечуючи більш обґрунтоване та ефективне прийняття рішень.
5. Масштабованість та інтеграція: Розроблена система ефективно обробляє великі обсяги даних, інтегрується з існуючими системами управління персоналом, підтримуючи різні сценарії найму.
6. Зниження людських помилок та постійне вдосконалення: Автоматизація значно знижує ризик помилок, а система може постійно вдосконалюватись на основі аналізу результатів найму та збору додаткових даних.
7. Економія ресурсів та забезпечення ефективності процесу найму: Використання системи дозволяє організаціям значно скоротити витрати, пов'язані з підбором персоналу, та підвищити загальну ефективність процесу найму.

У підсумку, дослідження демонструє не тільки технічну обґрунтованість та надійність розробленого програмного забезпечення, але й його практичну цінність для сучасного ринку праці, забезпечуючи ефективний, об'єктивний та масштабований підхід до відбору кандидатів.

## ВИСНОВКИ

У даній магістерській роботі був реалізований спосіб оцінки відповідності кандидата на основі засобів машинного навчання. Дослідження охоплювало всі ключові етапи розробки та оцінки такого додатку, включаючи аналіз потреб, визначення підходу до моделювання, реалізацію та валідацію. Створена система заснована на базі засобів машинного навчання з акцентом на об'єктивність, ефективність та точність при оцінці кандидатів. Вона інтегрується з існуючими даними та процесами, що забезпечує гнучкість та широкі можливості застосування в різних галузях та сценаріях найму. Визначення ефективних методів для аналізу та оцінки резюме, зокрема через застосування засобів машинного навчання, що дозволило глибше розуміти та обробляти складні шаблони даних.

Для оцінки кандидатів проводиться аналіз трьох основних категорій: навичок, освіти та досвіду роботи. Під час аналізу навичок модель OpenAI використовується для створення векторних представлень текстової інформації про вміння з резюме кандидата та необхідні навички з опису вакансії. Косинусна подібність використовується для визначення схожості між ними. Аналізуючи освітній рівень кандидата, перевіряємо, наскільки його здобутки відповідають критеріям, вказаним у описі вакансії. При оцінці професійного досвіду аналізується інформація з резюме, перетворюючи всі вказані формати стажу роботи на кількість місяців для зручності порівняння. Оцінка досвіду роботи здійснюється шляхом врахування різниці між досвідом кандидата та вимогами, зазначеними у робочому описі. В кожній категорії обчислюється окремий бал, а загальний бал обчислюється як середнє значення цих балів, призначаючи вагу кожній категорії за необхідності.

Виконано ретельну валідацію та оцінку системи, що включала ряд тестів для перевірки її надійності, точності та адаптивності. Результати підтвердили високу ефективність системи в різних умовах використання. Система демонструє значні переваги порівняно з традиційними методами відбору персоналу, такі як збільшення об'єктивності, зменшення витрат часу та ресурсів, а також забезпечення більш точної оцінки кандидатів.

Система показала свою здатність ефективно працювати з реальними даними, що включають різноманітні випадки та сценарії, забезпечуючи міцну основу для її використання в реальних умовах найму.

Центральним елементом системи є її гнучкість та масштабованість, які дозволяють адаптувати її до різних організаційних контекстів та величини даних. Це гарантує, що система залишається ефективною незалежно від кількості кандидатів або специфіки сектору.

Розроблення системи включало застосування новітніх технологій та підходів у галузі штучного інтелекту та машинного навчання. Це дозволило не лише забезпечити високу точність оцінок, але й відкрило шлях для подальших досліджень та розвитку в цій області.

Система демонструє велику практичну цінність для організацій, зменшуючи навантаження на HR-відділи та покращуючи якість процесу відбору. Її можливості роблять її ідеальним рішенням для сучасного динамічного ринку праці.

У підсумку, ця робота стала значущим внеском у галузі управління людськими ресурсами, пропонуючи новаторський підхід до автоматизації відбору персоналу з використанням засобів машинного навчання.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Alamelu, M., Kumar, D. S., Sanjana, R., Sree, J. S., Devi, A. S., & Kavitha, D. (2021). Resume validation and filtration using natural language processing.
2. Chou, Y.-C., & Yu, H.-Y. (2020). Based on the application of AI technology in resume analysis and job recommendation.
3. Daryani, C., Chhabra, G. S., Patel, H., Chhabra, I. K., & Patel, R. (2020). An automated resume screening system using natural language processing and similarity.
4. Garg, S., Sinha, S., Kar, A. K., & Mani, M. (2021b). A review of machine learning applications in human resource management.
5. Hemalatha, A., Kumari, P. B., Nawaz, N., & Gajenderan, V. (2021). Impact of artificial intelligence on recruitment and selection of information technology companies.
6. Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda.
7. Kadiwal, S. M., & Revanna, S. (2021). Design and development of machine learning based resume ranking system.
8. Abdel-Nasser Sharkawy (2020). Principle of Neural Network and Its Main Types: Review. Journal of Advances in Applied & Computational Mathematics, 7(1), 8-19. <https://doi.org/10.15377/2409-5761.2020.07.2>
9. Analysis of the Effectiveness of Using Machine Learning Algorithms to Make Hiring Decisions. Kirill Smelyakov, Yuliia Hurova and Serhii Osiievskyi, 6-15

## ДОДАТКИ