

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ВІННИЦЬКИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ ІНСТИТУТ**

Кафедра інноваційної економіки та цифрових технологій

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

**«ПРОЄКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ АНАЛІЗУ ТА
ВІДБОРУ РЕЗЮМЕ КАНДИДАТІВ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ
ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТЕКСТІВ»**

(за матеріалами ТОВ «ТРЕВЕЛ СТАДІ ВОРК», м. Київ)

Здобувача вищої освіти
4 курсу, групи ІСТ-41д,
спеціальності 126 «Інформаційні
системи та технології»
освітньої програми «Інформаційні
технології у бізнесі»

Вікторії БАЖУРИ

Науковий керівник
кандидат технічних наук

Руслан НОВИЦЬКИЙ

Гарант освітньо-професійної
програми
кандидат технічних наук, доцент

Світлана ЯРЕМКО

Вінниця 2026

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТЕКСТІВ ДЛЯ ВІДБОРУ ПЕРСОНАЛУ.....	6
1.1. Проблематика автоматизованого відбору кандидатів за резюме.....	6
1.2. Методи обробки текстів природною мовою (NLP) у системах рекрутингу.....	8
1.3. Огляд аналогів інформаційних систем аналізу та відбору резюме.....	12
РОЗДІЛ 2. ХАРАКТЕРИСТИКА ТА АНАЛІЗ ДІЯЛЬНОСТІ ТОВ «ТРЕВЕЛ СТАДІ ВОРК».....	15
2.1. Організаційно-економічна характеристика підприємства.....	15
2.2. Аналіз HR-процесів та поточного стану автоматизації рекрутингу на підприємстві.....	19
2.3. Обґрунтування необхідності розробки інформаційної системи аналізу резюме.....	22
РОЗДІЛ 3. ПРОЄКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ АНАЛІЗУ ТА ВІДБОРУ РЕЗЮМЕ КАНДИДАТІВ.....	24
3.1. Архітектура та UML-моделювання системи.....	24
3.2. Проектування бази даних та інтерфейсу системи.....	30
3.3. Алгоритм інтелектуального аналізу та ранжування резюме.....	33
ВИСНОВКИ ТА ПРОПОЗИЦІЇ.....	38
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	40
ДОДАТКИ.....	43

ВСТУП

В умовах сучасного динамічного ринку праці, де щодня публікуються тисячі вакансій і надходять сотні відгуків на кожну з них, ручний відбір резюме стає критично неефективним. Особливо гострою ця проблема є для компаній туристичного та рекрутингового сектору, які здійснюють сезонний масовий набір персоналу. ТОВ «ТРЕВЕЛ СТАДІ ВОРК» – провідна українська компанія з організації міжнародних стажувань та працевлаштування у сфері туризму – щороку обробляє сотні анкет і резюме кандидатів на програми стажування у Туреччині, Греції, Болгарії та інших країнах. Відсутність автоматизованого інструменту аналізу резюме суттєво уповільнює процес відбору, підвищує ймовірність людської помилки та знижує загальну ефективність HR-відділу. [16]

Актуальність теми полягає в необхідності впровадження сучасних інформаційних технологій у процес відбору персоналу. Методи інтелектуального аналізу текстів (Text Mining) і обробки природної мови (Natural Language Processing, NLP) відкривають широкі можливості для автоматизованого вилучення структурованих даних з неструктурованих документів (резюме), порівняння профілів кандидатів з вимогами вакансій та ранжування претендентів за ступенем відповідності. Зважаючи на стрімкий розвиток великих мовних моделей (Large Language Models, LLM) і трансформерних архітектур, що відбувся після 2022 року, завдання точного семантичного аналізу резюме стало доступним для впровадження навіть у малих підприємствах. [5, 6, 22]

Ступінь наукової розробленості проблеми. Теоретичні засади інтелектуального аналізу текстів та інформаційного пошуку закладено у фундаментальних працях Salton G. і McGill M. J. («Introduction to Modern Information Retrieval», 1983), які запропонували модель векторного простору та метрику TF-IDF, що й досі застосовуються в задачах аналізу документів.

Статистичні та лінгвістичні основи обробки природної мови систематизовано у роботах Manning C. D., Raghavan P. i Schütze H. («Introduction to Information Retrieval», 2008), а також Bird S., Klein E. i Loper E. («Natural Language Processing with Python», 2009). Практичні аспекти автоматизованого відбору персоналу за допомогою методів NLP досліджувались у роботах зарубіжних учених: Al-Otaibi S. T., Ykhlef M. розглядали інтелектуальні системи рекомендацій кандидатів; Faliagka E. та ін. пропонували системи автоматичного ранжування резюме на основі машинного навчання. Якісно новий етап у розвитку методів аналізу текстів пов'язаний із появою трансформерних архітектур: Vaswani A. та ін. («Attention Is All You Need», 2017) запропонували механізм уваги, на основі якого Devlin J. та ін. розробили модель BERT (2019), а Reimers N. i Gurevych I. — Sentence-BERT (2019), що дозволяє ефективно обчислювати семантичну подібність цілих речень і параграфів. Зазначені підходи відкривають принципово нові можливості для точного семантичного порівняння резюме з вимогами вакансій і складають теоретичну основу проєктованої системи. Метою кваліфікаційної роботи є проєктування інформаційної системи для автоматизованого аналізу та відбору резюме кандидатів з використанням методів інтелектуального аналізу текстів на матеріалах ТОВ «ТРЕВЕЛ СТАДІ ВОРК». [1-3, 5, 6, 22]

Для досягнення мети необхідно вирішити такі завдання:

- дослідити теоретичні основи інтелектуального аналізу текстів та методів NLP у контексті рекрутингу;
- проаналізувати існуючі аналоги програмних систем автоматизованого відбору резюме;
- охарактеризувати діяльність ТОВ «ТРЕВЕЛ СТАДІ ВОРК» та проаналізувати поточний стан HR-процесів підприємства;
- обґрунтувати необхідність та доцільність впровадження інформаційної системи аналізу резюме;
- розробити UML-моделі та архітектуру проєктованої системи;

- спроектувати базу даних, алгоритм інтелектуального аналізу та інтерфейс системи.

Об'єктом дослідження є процес аналізу та відбору резюме кандидатів у компанії ТОВ «ТРЕВЕЛ СТАДІ ВОРК».

Предметом дослідження є методи і засоби інтелектуального аналізу текстів для автоматизованого відбору резюме кандидатів в інформаційних системах.

Методи дослідження: системний аналіз, методи обробки природної мови (токенізація, лематизація, видалення стоп-слів, векторне представлення текстів), косинусна подібність, TF-IDF, методи UML-моделювання, аналіз фінансово-господарської діяльності підприємства.

Практична значущість роботи полягає в тому, що спроектована інформаційна система дозволить ТОВ «ТРЕВЕЛ СТАДІ ВОРК» скоротити час обробки резюме на 75-80%, зменшити кількість помилок при відборі кандидатів та підвищити якість підбору персоналу для програм міжнародних стажувань.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ОСНОВИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТЕКСТІВ ДЛЯ ВІДБОРУ ПЕРСОНАЛУ

1.1. Проблематика автоматизованого відбору кандидатів за резюме

Сучасний ринок праці характеризується постійно зростаючою кількістю пропозицій і попиту на робочі місця. За даними дослідження LinkedIn за 2024 рік, на одну вакансію в середньому надходить від 200 до 300 резюме. За умов ручного перегляду HR-менеджер витрачає від 6 до 8 секунд на первинне ознайомлення з одним резюме і від 10 до 23 годин на комплексний відбір по одній позиції. Отже, людський фактор, обмеженість ресурсів та суб'єктивність оцінки стають критичними вузькими місцями процесу рекрутингу. [13, 14]

Процес відбору резюме складається з кількох послідовних етапів. На першому етапі здійснюється первинна фільтрація за мінімальними вимогами: наявність дозволу на роботу, відповідна освіта, базові навички. На другому – оцінка за бажаними вимогами: релевантний досвід, додаткові компетенції, мовні навички. На третьому – формування короткого списку (shortlist) для запрошення на співбесіду. Середні коефіцієнти конверсії в рекрутингу, за даними SHRM (Society for Human Resource Management), становлять: 12% від відправлених заявок до співбесіди, 17% від співбесіди до отримання пропозиції, 89% від пропозиції до прийняття. [13]

Резюме є неструктурованими текстовими документами – вони не мають стандартизованого формату, можуть бути подані у форматах .pdf, .docx, .txt, мати різну структуру та обсяг. Це суттєво ускладнює автоматизовану обробку. Для сфери туризму та гостинності, де ТОВ «ТРЕВЕЛ СТАДІ ВОРК» здійснює набір персоналу, особливо важливими є такі характеристики кандидата, як

мовні навички, досвід у сервісі, освіта за фаховим напрямом, вік, наявність закордонного паспорта та готовність до переїзду.

Виділяють три основних проблеми сучасного процесу відбору резюме:

- Відокремлення найбільш підходящих кандидатів з великого масиву заявок – через людський фактор неможливо перевірити всі резюме та знайти того, хто найкраще відповідає вимогам вакансії;
- Виділення ключових аспектів з резюме кандидатів – оскільки резюме на ринку не стандартизовані, кожне має різну структуру, що потребує ручного пошуку релевантних даних;
- Зіставлення резюме з описом вимог вакансії – порівняння кваліфікації кандидата з вимогами роботодавця для визначення ступеня відповідності.

Ці проблеми призводять до того, що до 75-88% резюме відхиляються ще до реального розгляду, а компанії несуть суттєві витрати на HR-ресурси. За оцінками Glassdoor, вартість одного найму становить у середньому 4 000 доларів США, а середній час закриття вакансії – 24 дні. Для сезонних компаній туристичного сектору, таких як ТОВ «ТРЕВЕЛ СТАДІ ВОРК», де одночасно потрібно набрати сотні кандидатів на різні позиції, ця проблема набуває критичного значення. [15]

Автоматизовані системи відбору резюме (ATS – Applicant Tracking System) з'явилися ще у 1990-х роках як інструменти управління базами даних кандидатів. Сучасні ATS-системи інтегрують методи машинного навчання та NLP для підвищення якості відбору. Однак навіть найсучасніші комерційні рішення мають обмеження: вони не завжди враховують семантичні нюанси, використання синонімів, специфіку галузі, та часто пропускають сильних кандидатів через відсутність в їх резюме точних ключових слів.

Таблиця 1.1 – Порівняння ручного та автоматизованого відбору резюме

Критерій	Ручний відбір	Автоматизований відбір
Час обробки 100 резюме	10-23 год	1-5 хв
Суб'єктивність	Висока	Мінімальна
Точність відповідності вимогам	Залежить від HR	Алгоритмічна, стабільна
Врахування синонімів	Так (людський інтелект)	Потребує NLP-моделей
Масштабованість	Обмежена	Висока
Вартість	Висока (оплата праці HR)	Низька (після впровадження)
Ризик помилки	Втома, упередженість	Залежить від якості алгоритму

Таким чином, впровадження інформаційної системи з інтелектуальним аналізом текстів резюме є об'єктивно необхідним для підвищення ефективності рекрутингових процесів ТОВ «ТРЕВЕЛ СТАДІ ВОРК». Система повинна не лише фільтрувати резюме за ключовими словами, але й здійснювати семантичний аналіз на основі сучасних NLP-методів.

1.2. Методи обробки текстів природною мовою (NLP) у системах рекрутингу

Обробка текстів природною мовою (Natural Language Processing, NLP) є підгалуззю штучного інтелекту, що спеціалізується на наданні комп'ютерам здатності розуміти, інтерпретувати та генерувати людську мову. NLP поєднує обчислювальну лінгвістику з методами статистичного машинного навчання та

глибокого навчання. Саме NLP-методи лежать в основі сучасних систем автоматизованого відбору резюме. [2, 3]

Конвеєр обробки тексту в NLP для аналізу резюме включає такі основні кроки:

- Перетворення вхідних файлів (PDF, DOCX) у текстовий формат;
 - Нормалізація тексту: видалення зайвих символів, пунктуації, переведення до нижнього регістру;
 - Токенізація: розбиття тексту на окремі токени (слова або фрази);
 - Видалення стоп-слів (службових слів, що не несуть значущого змісту);
 - Лематизація або стемінг: приведення слів до базової форми;
 - Розпізнавання іменованих сутностей (Named Entity Recognition, NER): виявлення імен, організацій, дат, контактних даних;
 - Векторне представлення тексту та обчислення семантичної подібності.
- Основні задачі NLP в контексті аналізу резюме включають:
- Розпізнавання іменованих об'єктів (NER) – виявлення імені кандидата, навчального закладу, попередніх місць роботи, контактних даних;
 - Вилучення інформації (Information Extraction, IE) – структурований видобуток даних про навички, освіту, досвід роботи;
 - Класифікація тексту – категоризація резюме за галузями, рівнями посад;
 - Аналіз семантичної подібності – порівняння тексту резюме з описом вакансії.

Таблиця 1.2 – Основні NLP-методи та їх застосування в аналізі резюме

Метод NLP	Призначення	Застосування в аналізі резюме
Токенізація	Розбиття тексту на токени	Підготовка тексту резюме до аналізу
Видалення стоп-слів	Фільтрація службових слів	Зменшення шуму в даних
Лематизація	Приведення слів до базової форми	Уніфікація різних форм слів
NER	Виявлення іменованих сутностей	Вилучення імені, контактів, організацій
TF-IDF	Оцінка важливості термінів	Виявлення ключових навичок у тексті
Word2Vec / GloVe	Векторне представлення слів	Семантична близькість термінів
BERT / Sentence Transformers	Контекстуальні ембедінги	Точне семантичне порівняння резюме і вакансії
Косинусна подібність	Міра схожості векторів	Обчислення відсотку відповідності кандидата

Метод TF-IDF (Term Frequency-Inverse Document Frequency) є класичним підходом до оцінки важливості слів у документі відносно колекції документів. TF-IDF розраховується як добуток частоти терміна в документі (TF) на обернений логарифм частоти документів, що містять даний термін (IDF). Цей метод ефективний для пошуку ключових навичок і термінів у резюме, але не враховує семантику. [1, 10]

Косинусна подібність є однією з ключових метрик для порівняння текстових документів у NLP. Вона вимірює косинус кута між двома n -вимірними векторами в багатовимірному просторі. Математично косинусна подібність між векторами A і B визначається формулою:

$$\cos(A, B) = (A \cdot B) / (|A| \times |B|)$$

де $A \cdot B$ – скалярний добуток векторів, $|A|$ і $|B|$ – їхні євклідові норми. Значення косинусної подібності знаходиться в діапазоні від 0 до 1: значення 1 означає повну ідентичність векторів, 0 – повну відсутність схожості. У системі

аналізу резюме косинусна подібність дозволяє обчислити відсоток відповідності кожного резюме опису вакансії. [10]

Сучасним трендом у NLP є використання трансформерних архітектур, зокрема моделей типу BERT (Bidirectional Encoder Representations from Transformers) та Sentence Transformers. На відміну від TF-IDF та Word2Vec, ці моделі враховують контекст слова в реченні та здатні вловлювати тонкі семантичні відмінності. Sentence-BERT (SBERT) генерує ембедінги для цілих речень і параграфів, що дозволяє ефективно порівнювати фрагменти резюме з вимогами вакансії навіть при використанні різних слів і формулювань. [5, 6]

Для україномовних і мультимовних контекстів (що актуально для ТОВ «ТРЕВЕЛ СТАДІ ВОРК», яке працює з кандидатами, що подають резюме різними мовами) рекомендується використання мультилінгвальних моделей, таких як multilingual-BERT або LaBSE (Language-agnostic BERT Sentence Embeddings). Ці моделі забезпечують точну обробку текстів українською, англійською та іншими мовами в єдиному векторному просторі. [9]

Алгоритм вилучення інформації з резюме базується на таких техніках NLP:

- Токенізація на рівні речень і слів – бібліотека spaCy або NLTK; [8] [8]
- NER для виявлення контактних даних (електронна пошта – регулярні вирази; телефон – регулярні вирази; ім'я – NER-модель);
- Пошук навичок за переліком (skills matching) – порівняння токенів з еталонним словником навичок;
- Вилучення інформації про освіту – пошук за ключовими словами (бакалавр, магістр, університет);
- Вилучення даних про досвід роботи – аналіз дат і назв посад.

1.3. Огляд аналогів інформаційних систем аналізу та відбору резюме

Ринок програмного забезпечення для автоматизованого відбору резюме представлений різноманітними рішеннями – від хмарних ATS-платформ до спеціалізованих AI-інструментів аналізу резюме. Розглянемо найбільш відомі системи-аналоги.

Trakstar Hire (колишній RecruiterBox)

Trakstar Hire є онлайн-інструментом відстеження кандидатів з хмарними технологіями. Система надає: автоматизоване відстеження кандидатів, командну роботу з відгуками рекрутерів, профілювання кандидатів для фільтрації найкращих. Обмеження системи – відсутність глибокої NLP-інтеграції та повноцінного API для кастомізації.

CVVIZ

CVVIZ – AI-рішення для відбору резюме кандидатів з контекстним аналізом на основі штучного інтелекту. Переваги: використання еластичного пошуку (Elasticsearch) для швидкого знаходження результатів, знаходження резюме-дублікатів, відновлення процесу відбору. Недолік: незавершеність мобільного інтерфейсу.

Employa

Employa надає можливість HR-спеціалістам зосередитися на особистій взаємодії з кандидатами, автоматизуючи перевірку резюме. Система перевіряє резюме на наявність прогалин у стажі, пропонує необмежену кількість критеріїв пошуку та автоматизований пошук кандидатів.

HireVue

HireVue – сучасна AI-платформа для скринінгу кандидатів, що включає відеоінтерв'ю з аналізом мови та мімічних виразів, автоматичний скоринг кандидатів, інтеграцію з провідними ATS-системами. Недоліки: висока

вартість ліцензування, складність налаштування, питання етики AI-оцінювання.

Workable

Workable – комплексна ATS-платформа з вбудованим AI-пошуком кандидатів (AI Recruiter), рекомендаціями кандидатів на основі машинного навчання, автоматичною перевіркою резюме. Підходить для середніх і великих компаній, але має обмежені можливості кастомізації під специфічні вимоги галузі туризму.

Таблиця 1.3 – Порівняльна характеристика аналогів системи

Критерій	Trakstar	CVVIZ	Employa	HireVue	Проектована ІС
NLP-аналіз	Обмежений	Так	Так	Так	Так (TF-IDF + BERT)
% відповідності кандидата	Ні	Так	Частково	Так	Так
Збереження результатів у БД	Так	Так	Так	Так	Так
Підтримка PDF/DOCX	Так	Так	Так	Так	Так
Мультимовність	Обмежена	Обмежена	Ні	Обмежена	Так
Open Source / кастомізація	Ні	Ні	Ні	Ні	Так
Адаптація під туризм	Ні	Ні	Ні	Ні	Так

Проведений аналіз свідчить, що жодне з розглянутих рішень не забезпечує повноцінної адаптації до специфіки рекрутингу в туристичній галузі, не підтримує мультимовну обробку, включаючи українську мову, і не є налаштованим open-source рішенням. Проектована інформаційна система для ТОВ «ТРЕВЕЛ СТАДІ ВОРК» спрямована на усунення цих обмежень.

У першому розділі розглянуто проблематику автоматизованого відбору кандидатів за резюме, досліджено основні методи обробки текстів природною мовою (NLP) та проведено огляд програмних аналогів. Встановлено, що ручний відбір резюме є неефективним в умовах масового рекрутингу і

потребує автоматизації. Ключовими методами для аналізу резюме є: токенізація, лематизація, NER, TF-IDF, векторне представлення текстів та косинусна подібність. Аналіз аналогів підтвердив необхідність розробки власного рішення, адаптованого до потреб ТОВ «ТРЕВЕЛ СТАДІ ВОРК».



РОЗДІЛ 2

ХАРАКТЕРИСТИКА ТА АНАЛІЗ ДІЯЛЬНОСТІ ТОВ «ТРЕВЕЛ СТАДІ ВОРК»

2.1. Організаційно-економічна характеристика підприємства

ТОВ «ТРЕВЕЛ СТАДІ ВОРК» (код за ЄДРПОУ: 43671491) – приватне товариство з обмеженою відповідальністю, засноване у 2020 році в місті Київ. Юридична адреса підприємства: м. Київ, провулок Георгіївський, буд. 2, офіс 44. Компанія спеціалізується на організації міжнародних стажувань для студентів та працевлаштуванні у сфері туризму й готельного бізнесу.

Місія компанії: допомогти студентам та молодим фахівцям здобути міжнародний досвід у туристичній галузі, поєднуючи навчання, подорожі та роботу. Компанія успішно співпрацює з понад 55 державними та національними університетами України, організовує стажування на базі провідних туроператорів і в більш ніж 50 готелях у Туреччині, Греції, Болгарії, Кіпрі, Словаччині, Австрії, Чорногорії, Іспанії, Тунісі, Шрі-Ланці та Албанії.

Основні напрями діяльності ТОВ «ТРЕВЕЛ СТАДІ ВОРК»:

- Рекрутинг: пошук, відбір та найм персоналу для партнерів компанії – туроператорів і готелів. Включає проведення кастингів, проведення інтерв'ю, відкриття віз та страховок, логістичні послуги;
- Організація стажувань: програми практики для студентів у сферах туризму, анімації, спа та ресторанно-готельного сервісу з наданням міжнародного сертифіката ЄКТС;
- Працевлаштування за кордоном: конкурентні програми роботи для випускників університетів у міжнародних туристичних компаніях і готелях;

- Стажування освітян: програми для науково-педагогічних працівників університетів-партнерів.

Організаційна структура ТОВ «ТРЕВЕЛ СТАДІ ВОРК» є лінійно-функціональною. Очолює компанію директор-засновник. Структурні підрозділи включають відділ рекрутингу, відділ по роботі з університетами, відділ логістики та візового супроводу, маркетинговий відділ. Загальна кількість штатних працівників за даними фінансової звітності складає 1-2 особи, при цьому компанія активно залучає позаштатних фахівців (фрілансерів) у пікові сезони. [16, 17]

Таблиця 2.1 – Основні показники активу балансу ТОВ «ТРЕВЕЛ СТАДІ ВОРК» (тис. грн)

Показник	Кін. 2023	Кін. 2024	Кін. 2025	Відх. 2024/2023	Відх. 2025/2024
Основні засоби	22,30	168,40	131,90	+146,10	-36,50
Усього необоротних активів	22,30	168,40	131,90	+146,10	-36,50
Запаси	0,80	1,00	1,00	+0,20	0,00
Гроші та еквіваленти	1,20	9,30	9,00	+8,10	-0,30
Інші оборотні активи	0,00	88,10	0,00	+88,10	-88,10
Усього оборотних активів	2,00	98,40	26,70	+96,40	-71,70
Баланс	24,30	266,80	158,60	+242,50	-108,20

Аналіз активу балансу свідчить про активне нарощування основних засобів у 2024 році (з 22,30 до 168,40 тис. грн, тобто зростання в 7,5 рази), що може свідчити про придбання обладнання або оргтехніки. У 2025 році обсяг основних засобів дещо зменшився до 131,90 тис. грн внаслідок амортизації. Загальний баланс підприємства зростав у 2024 році (266,80 тис. грн), проте в 2025 скоротився до 158,60 тис. грн. [17, 18, 19]

Таблиця 2.2 – Основні показники пасиву балансу ТОВ «ТРЕВЕЛ СТАДІ ВОРК» (тис. грн)

Показник	Кін. 2023	Кін. 2024	Кін. 2025	Відх. 2024/2023	Відх. 2025/2024
Зареєстрований капітал	1,00	1,00	1,00	0,00	0,00
Нерозподілений прибуток (збиток)	-3,10	-17,50	-55,20	-14,40	-37,70
Усього власного капіталу	-2,10	-16,50	-54,20	-14,40	-37,70
Поточні зобов'язання, усього	26,40	283,30	212,80	+256,90	-70,50
Зобов'язання з оплати праці	13,70	31,30	28,70	+17,60	-2,60
Інші поточні зобов'язання	5,70	241,60	172,30	+235,90	-69,30
Баланс	24,30	266,80	158,60	+242,50	-108,20

Аналіз пасиву демонструє від'ємне значення власного капіталу протягом усіх трьох років (-2,10 тис. грн у 2023, -16,50 тис. грн у 2024 і -54,20 тис. грн у 2025), що пояснюється накопиченими збитками від операційної діяльності. Підприємство фінансується переважно за рахунок поточних зобов'язань, зокрема інших поточних зобов'язань, що в 2024 році різко зросли (до 241,60 тис. грн), а в 2025 скоротились до 172,30 тис. грн. [17, 18, 19]

Таблиця 2.3 – Аналіз фінансових результатів діяльності ТОВ «ТРЕВЕЛ СТАДІ ВОРК» (тис. грн)

Показник	2023	2024	2025	Темп зросту 2024/2023, %	Темп зросту 2025/2024, %
Чистий дохід від реалізації	126,80	245,60	486,90	193,7%	198,3%
Собівартість реалізованої продукції	157,40	286,10	492,90	181,8%	172,3%
Валовий прибуток (збиток)	-30,60	-40,50	-6,00	–	–

Продовження табл.2.3

Інші доходи	25,10	26,10	5,90	103,9%	22,6%
Інші витрати	0,00	0,00	37,60	–	–
Фін. результат до оподаткування	-5,50	-14,40	-37,70	–	–
Чистий прибуток (збиток)	-5,50	-14,40	-37,70	–	–

Фінансові результати підприємства за 2023-2025 роки демонструють стрімке зростання виручки: з 126,80 тис. грн у 2023 до 245,60 тис. грн у 2024 (+93,7%) і до 486,90 тис. грн у 2025 (+98,3%), що свідчить про активний розвиток бізнесу. Водночас підприємство є збитковим: чистий збиток зростає з 5,50 тис. грн у 2023 до 14,40 тис. грн у 2024 і 37,70 тис. грн у 2025. Перевищення собівартості над доходом від реалізації вказує на необхідність оптимізації операційних витрат, у тому числі витрат на HR-процеси. [17, 18, 19]

Динаміку виручки та збитків підприємства за 2023-2025 роки наведено на рисунку 2.1.

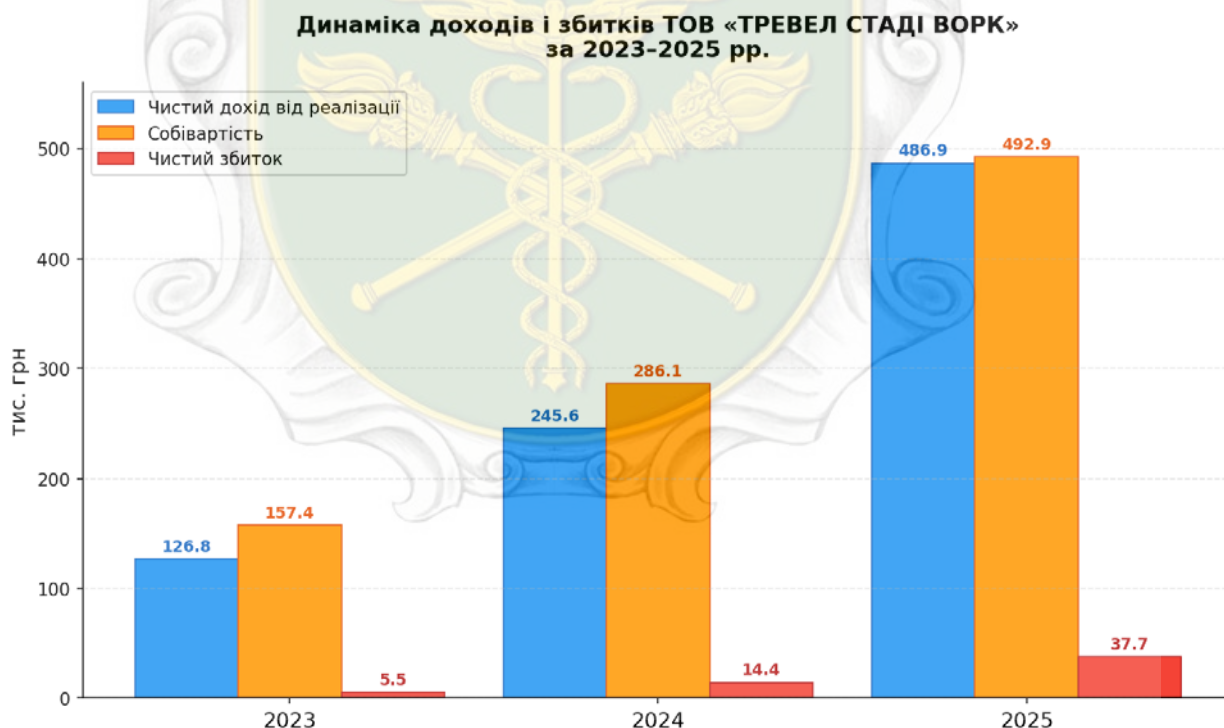


Рисунок 2.1 – Динаміка доходів і збитків ТОВ «ТРЕВЕЛ СТАДІ ВОРК» за 2023-2025 рр.

Примітка: діаграму побудовано на підставі даних фінансової звітності ТОВ «ТРЕВЕЛ СТАДІ ВОРК» за 2023, 2024 та 2025 рр.

2.2. Аналіз HR-процесів та поточного стану автоматизації рекрутингу

Рекрутингова діяльність є ключовою для ТОВ «ТРЕВЕЛ СТАДІ ВОРК», оскільки компанія одночасно виступає і агентством з працевлаштування (підбирає кандидатів для готелів-партнерів), і організатором програм стажування (відбирає студентів для участі). Аналіз поточного стану HR-процесів підприємства виявив такі характеристики:

Джерела надходження резюме та анкет: сайт компанії (онлайн-анкета), платформи Work.ua і Rabota.ua, Instagram та LinkedIn, університетські партнери (прямі рекомендації). Відповідно, резюме надходять у різних форматах (PDF, DOCX, зображення), різними мовами (українська, англійська, рідше – інші) та з різним ступенем структурованості. [16]

Поточний процес обробки резюме є переважно ручним і включає такі кроки:

- Отримання резюме/анкет електронною поштою або через форму на сайті;
- Ручне ознайомлення HR-менеджера з кожним резюме;
- Складання переліку релевантних кандидатів у файлі Excel;
- Зв'язок з відібраними кандидатами для проведення онлайн-інтерв'ю;
- Фінальне рішення щодо кандидата і оформлення документів.

Основні проблеми поточного процесу HR ТОВ «ТРЕВЕЛ СТАДІ ВОРК»:

- Відсутність централізованої бази даних кандидатів – дані розпорошені між Excel-файлами і поштовим скринькою;
- Значні часові витрати – під час пікового сезону (лютий-квітень) надходить понад 200 резюме на тиждень, що перевищує можливості ручної обробки малою командою;
- Суб'єктивність відбору і ризик пропуску відповідних кандидатів;
- Відсутність стандартизованих критеріїв оцінки та системи ранжування;
- Неможливість повторного пошуку серед архіву кандидатів.

Таблиця 2.4 – SWOT-аналіз HR-процесів ТОВ «ТРЕВЕЛ СТАДІ ВОРК»

Сильні сторони (S)	Слабкі сторони (W)
Розвинена мережа університетів-партнерів (55+)	Відсутність автоматизованої системи обробки резюме
Власна база кандидатів, напрацьована за роки	Ручний відбір займає критично багато часу
Міжнародна мережа роботодавців (50+ готелів)	Розпорошеність даних кандидатів
Досвід у рекрутингу туристичного сектору	Обмежені HR-ресурси (1-2 штатних співробітники)
Можливості (O)	Загрози (T)
Впровадження NLP-системи для аналізу резюме	Втрата конкурентної позиції через повільний відбір
Зростання попиту на стажування (300+ учасників у 2025)	Ризик пропуску найкращих кандидатів
Інтеграція з платформами Work.ua / LinkedIn API	Погіршення якості підбору при збільшенні обсягу
AI-рекомендації для покращення відповідності	Збільшення операційних витрат на HR

Динаміку кількості учасників програм ТОВ «ТРЕВЕЛ СТАДІ ВОРК» за 2021-2025 роки наведено на рисунку 2.2.

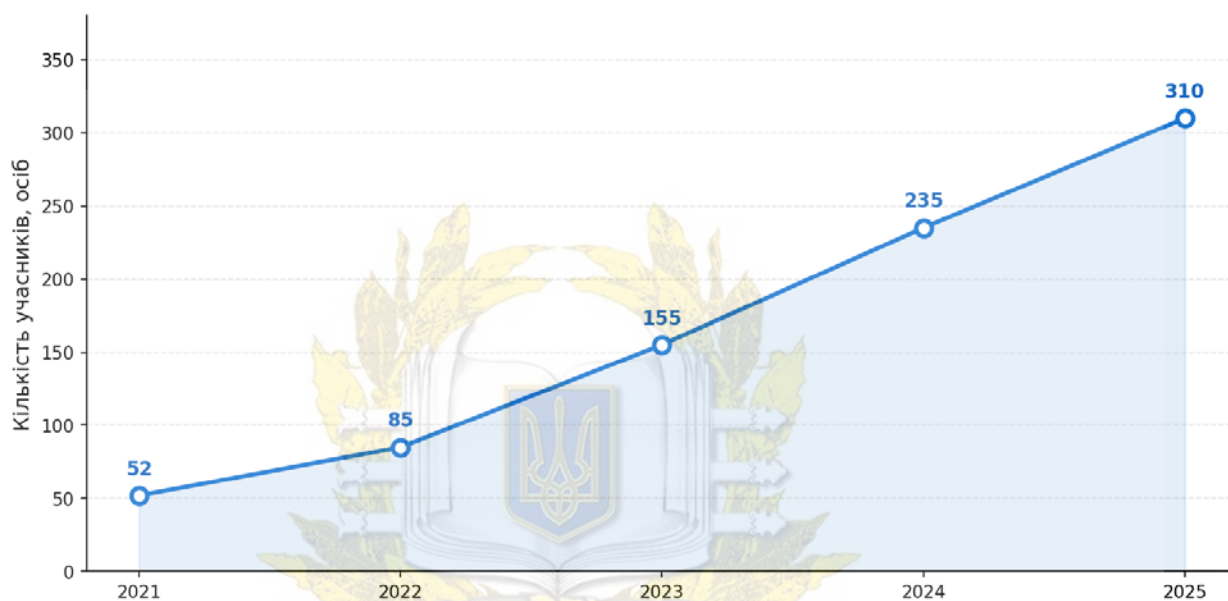


Рисунок 2.2 – Динаміка кількості учасників програм ТОВ «ТРЕВЕЛ СТАДІ ВОРК»

Примітка: дані за 2021-2023 роки є оціночними на підставі публічної інформації компанії; дані 2025 року – з офіційного сайту. [16]

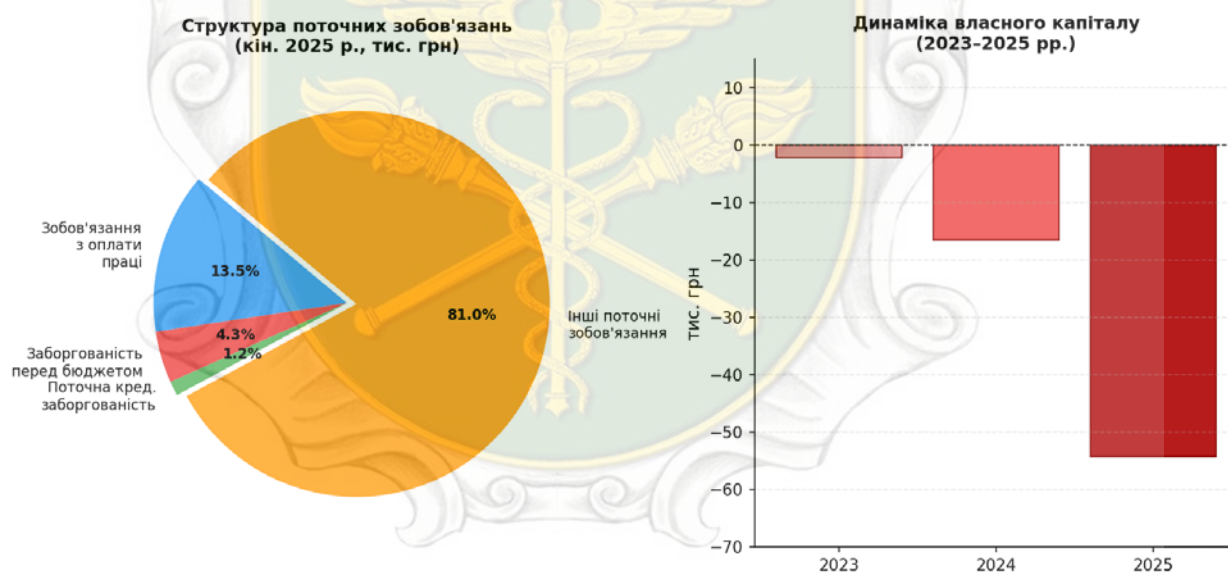


Рисунок 2.3 – Аналіз власного капіталу та структури зобов'язань ТОВ «ТРЕВЕЛ СТАДІ ВОРК»

2.3. Обґрунтування необхідності розробки інформаційної системи аналізу резюме

На підставі проведеного аналізу діяльності та HR-процесів ТОВ «ТРЕВЕЛ СТАДІ ВОРК» можна зробити висновок про нагальну необхідність впровадження спеціалізованої інформаційної системи для аналізу та відбору резюме. Ключові аргументи на користь розробки такої системи:

- Масштаб завдання. Кількість учасників програм компанії зросла з 50 у 2021 до 300+ у 2025 році. При коефіцієнті конверсії близько 12%, для набору 300 учасників необхідно обробити близько 2 500 резюме на рік, або понад 200 на місяць у пікові сезони. Ручна обробка такого обсягу є нереалістичною для малого штату. [16]
- Економічна ефективність. Виручка підприємства за 2 роки зросла майже в 4 рази (з 126,80 до 486,90 тис. грн), проте збиток також зростає. Оптимізація HR-процесів через автоматизацію дозволить скоротити операційні витрати на підбір персоналу і підвищити рентабельність. [17, 18, 19]
- Конкурентна перевага. Автоматизований відбір дозволить швидше реагувати на запити партнерів (готелів, туроператорів), пропонувати більш точно підібраних кандидатів і підвищити задоволеність клієнтів.
- Якість відбору. NLP-система здатна виявити відповідні кандидатури, що могли бути пропущені при ручному перелюді, на основі семантичного аналізу, а не лише точного збігу ключових слів.
- Збереження і використання накопиченої бази. Компанія має власну базу кандидатів, «напрацьовану роками». Інформаційна система дозволить систематизувати, зберігати та повторно використовувати цю базу.

Таблиця 2.5 – Оцінка ефекту від впровадження ІС аналізу резюме

Показник	До впровадження	Після впровадження	Ефект
Час обробки 100 резюме	10-23 год	15-30 хв	Скорочення в 40-50 разів
Частка релевантних кандидатів у shortlist	~60%	~85%	+25%
Потреба в HR-ресурсі (год/місяць)	~80 год	~20 год	Зменшення в 4 рази
Наявність централізованої БД кандидатів	Ні	Так	Якісна перевага
Можливість повторного пошуку	Обмежена	Повна	Нові можливості

Таким чином, впровадження інформаційної системи аналізу та відбору резюме є стратегічно виправданим і економічно обґрунтованим рішенням для ТОВ «ТРЕВЕЛ СТАДІ ВОРК». Система забезпечить суттєве підвищення ефективності рекрутингового процесу, якості відбору кандидатів та конкурентоспроможності компанії на ринку HR-послуг у туристичній галузі.

У другому розділі охарактеризовано діяльність ТОВ «ТРЕВЕЛ СТАДІ ВОРК» як провідного українського агентства з організації міжнародних стажувань у туризмі. Проведено аналіз фінансового стану підприємства за 2023-2025 роки, який показав стрімке зростання доходів (в 3,8 рази), але й наростання збитків, що обумовлює необхідність оптимізації операційних витрат. Аналіз HR-процесів підприємства виявив суттєві недоліки ручного відбору резюме: значні часові витрати, суб'єктивність оцінки, відсутність централізованої бази даних. На підставі SWOT-аналізу та кількісної оцінки ефекту обґрунтовано необхідність розробки інформаційної системи аналізу та відбору резюме.

РОЗДІЛ 3

ПРОЄКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ АНАЛІЗУ ТА ВІДБОРУ РЕЗЮМЕ КАНДИДАТІВ

3.1. Архітектура та UML-моделювання системи

Проектowana інформаційна система аналізу та відбору резюме кандидатів (далі – ІСАРК) ТОВ «ТРЕВЕЛ СТАДІ ВОРК» побудована за клієнт-серверною архітектурою. Такий підхід забезпечує централізоване зберігання і обробку даних на сервері, тоді як клієнтська частина відповідає за взаємодію з користувачем та відображення результатів. Система складається з трьох основних рівнів:

- Рівень презентації (Presentation Layer): веб-інтерфейс на базі React.js або Vue.js, що забезпечує зручну роботу HR-менеджера з системою;
- Рівень бізнес-логіки (Business Logic Layer): серверна частина на Python (Flask або FastAPI), що реалізує алгоритми NLP-аналізу, вилучення даних і ранжування резюме;
- Рівень даних (Data Layer): реляційна база даних PostgreSQL для збереження даних кандидатів, вакансій і результатів аналізу.

Для моделювання системи використано нотацію UML (Unified Modeling Language). Нижче наведено основні UML-діаграми проєктованої системи. [21, 26]

Діаграма варіантів використання (Use Case Diagram)

Діаграма варіантів використання описує функціональну поведінку системи з точки зору користувача. Основний актор системи – HR-менеджер (адміністратор). Виділяємо такі основні варіанти використання (Use Cases):

Таблиця 3.1 – Специфікація варіантів використання ICAPK

Актор	Варіант використання	Опис
HR-менеджер	Завантажити опис вакансії	Система приймає файл (.pdf або .docx) з описом вимог до вакансії і витягує ключові критерії відбору
HR-менеджер	Завантажити резюме кандидатів	Пакетне завантаження набору резюме у форматах .pdf, .docx для подальшого аналізу
HR-менеджер	Запустити аналіз резюме	Система автоматично обробляє всі завантажені резюме, вилучає ключові дані та обчислює відсоток відповідності кожного кандидата вимогам вакансії
HR-менеджер	Переглянути результати ранжування	Система відображає впорядкований список кандидатів з ключовими даними та відсотком відповідності у спадному порядку
HR-менеджер	Пошук по базі кандидатів	Пошук і фільтрація по збереженій базі кандидатів за ключовими параметрами
HR-менеджер	Переглянути профіль кандидата	Детальний перегляд вилучених даних з резюме конкретного кандидата
HR-менеджер	Управляти базою вакансій	Додавання, редагування та перегляд вакансій і їх вимог
Система (auto)	Зберегти результати в БД	Автоматичне збереження всіх отриманих результатів аналізу у базу даних

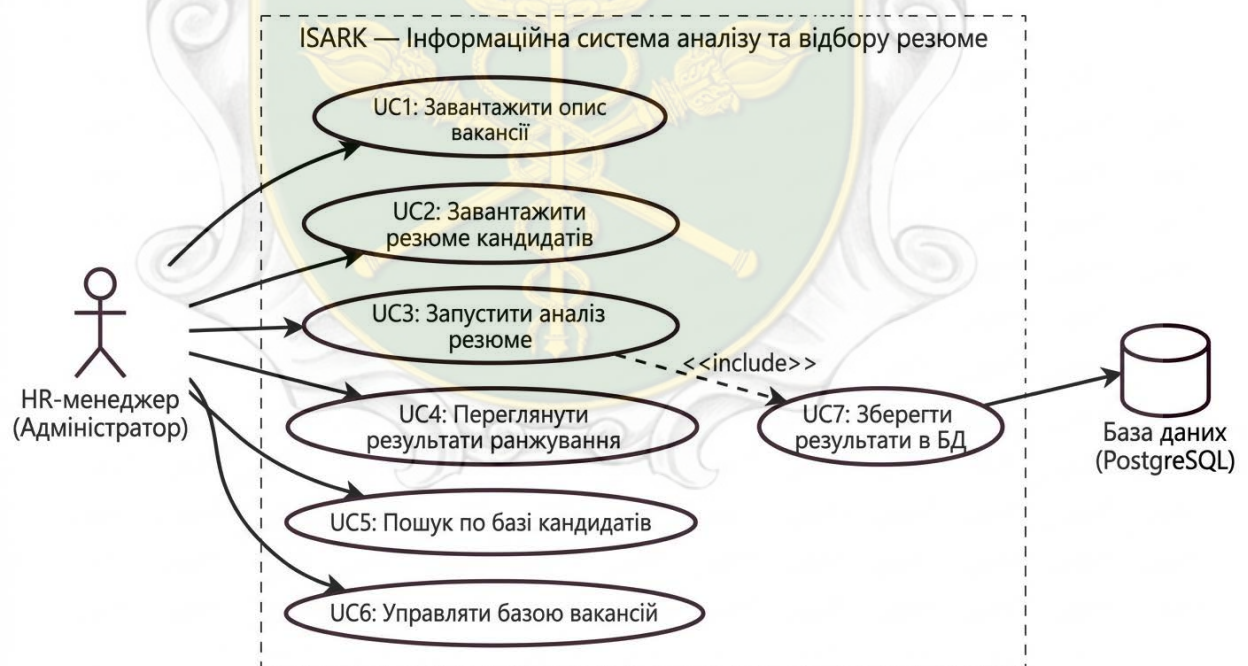


Рисунок 3.1 – Діаграма варіантів використання ICAPK (UML Use Case Diagram)

Діаграма послідовності (Sequence Diagram)

Діаграма послідовності описує порядок взаємодії між компонентами системи у часі для основного сценарію – аналізу та ранжування резюме.

Основна послідовність взаємодії:

- HR-менеджер завантажує файл з описом вакансії через веб-інтерфейс;
- Веб-інтерфейс надсилає запит на API-сервер;
- API-сервер передає файл на NLP-модуль для обробки;
- NLP-модуль конвертує файл у текст, токенізує, видаляє стоп-слова та лематизує;
- NLP-модуль вилучає ключові вимоги (навички, освіта) і повертає структурований об'єкт;
- HR-менеджер завантажує набір резюме через веб-інтерфейс;
- API-сервер передає файли в NLP-модуль для послідовної обробки;
- NLP-модуль для кожного резюме вилучає: ім'я, контакти, навички, освіту, досвід роботи;
- Модуль подібності обчислює TF-IDF вектори і косинусну подібність для кожного резюме та вакансії;
- Система ранжує кандидатів у порядку спадання % відповідності та зберігає результати в БД;
- Веб-інтерфейс відображає ранжований список кандидатів HR-менеджеру.

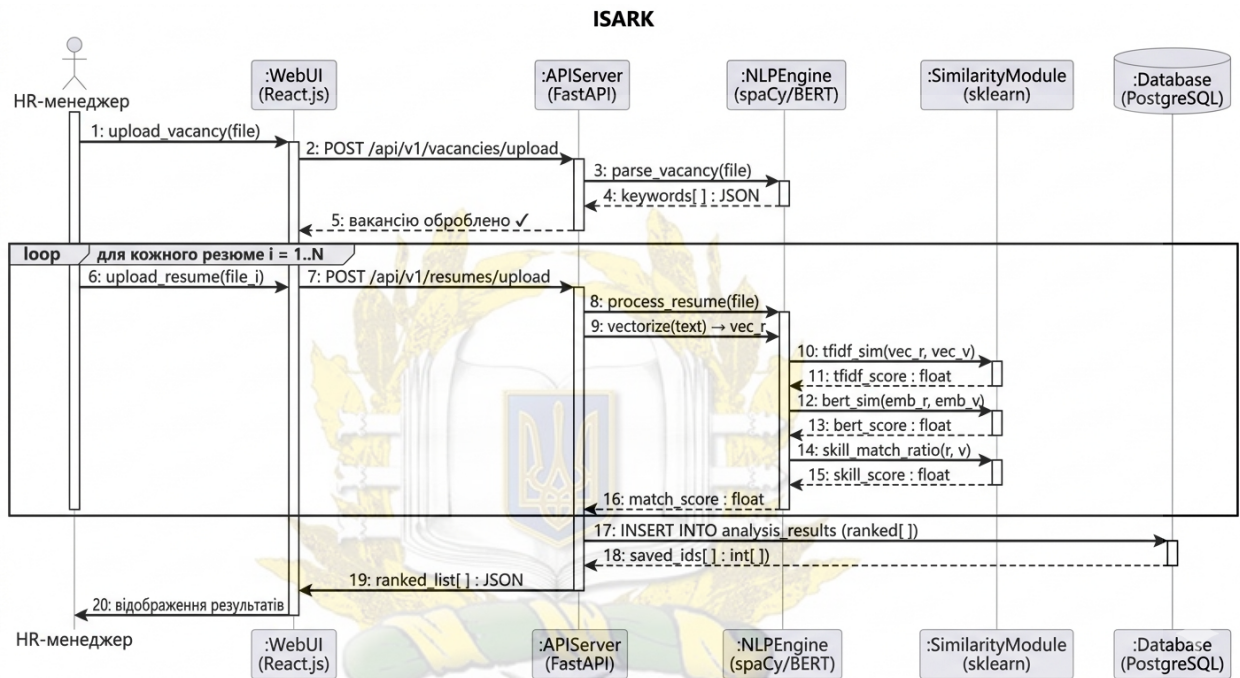


Рисунок 3.2 – Діаграма послідовності аналізу резюме (UML Sequence Diagram)

Діаграма класів (Class Diagram)

Діаграма класів описує статичну структуру системи – основні класи, їх атрибути, методи та зв'язки між ними. Основні класи ISARK:

Таблиця 3.2 – Специфікація основних класів ISARK

Клас	Атрибути	Методи
Resume	id, candidate_name, email, phone, skills[], education[], experience[], raw_text, created_at	parse(), extract_skills(), extract_education(), extract_contacts()
Vacancy	id, title, description, required_skills[], preferred_skills[], min_education, created_at	parse_requirements(), get_skill_vector()
NLPProcessor	model, stopwords[], lemmatizer	tokenize(text), remove_stopwords(tokens), lemmatize(tokens), extract_ner(text), vectorize(text)

Продовження табл. 3.2

SimilarityEngine	vectorizer (TF-IDF), model (BERT)	compute_tfidf(texts), compute_cosine_similarity(vec_a, vec_b), rank_candidates(resumes, vacancy)
AnalysisResult	id, resume_id, vacancy_id, match_score, matched_skills[], analysis_date	get_shortlist(n), export_to_excel()
DBManager	connection, session	save_resume(resume), save_result(result), search_candidates(query), get_vacancy(id)

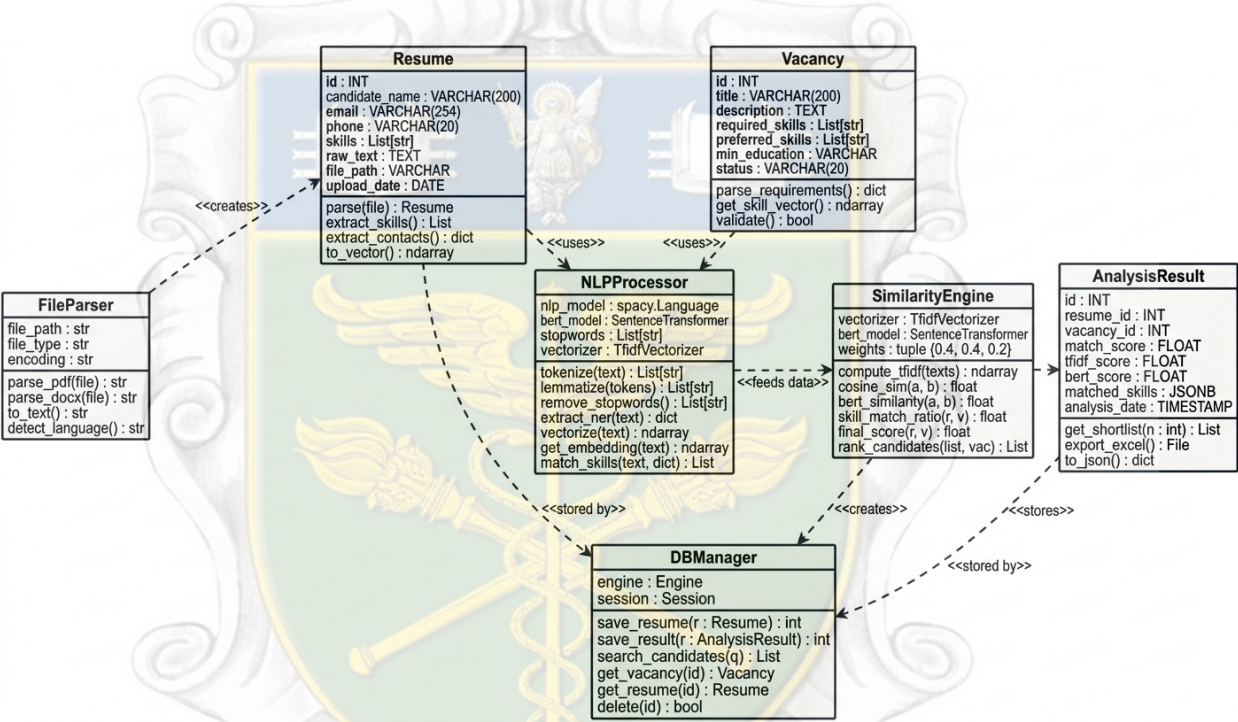


Рисунок 3.3 – Діаграма класів ICAPK (UML Class Diagram)

Діаграма компонентів (Component Diagram)

Діаграма компонентів описує фізичну архітектуру системи з точки зору програмних компонентів та їх взаємозв'язків. ICAPK складається з таких компонентів:

- Web UI Component – веб-інтерфейс (React.js), що надає функції завантаження файлів, перегляду результатів та управління вакансіями;
- API Gateway – REST API на базі FastAPI (Python), що обробляє HTTP-запити від клієнта та координує роботу серверних модулів;
- File Parser Module – модуль конвертації файлів PDF і DOCX у текстовий формат (бібліотеки pdfplumber, python-docx);
- NLP Engine – ядро системи: токенизація (spaCy), NER, TF-IDF (scikit-learn), BERT-ембедінги (sentence-transformers);
- Similarity Module – обчислення косинусної подібності та ранжування кандидатів;
- Database Component – PostgreSQL з ORM SQLAlchemy для управління даними.

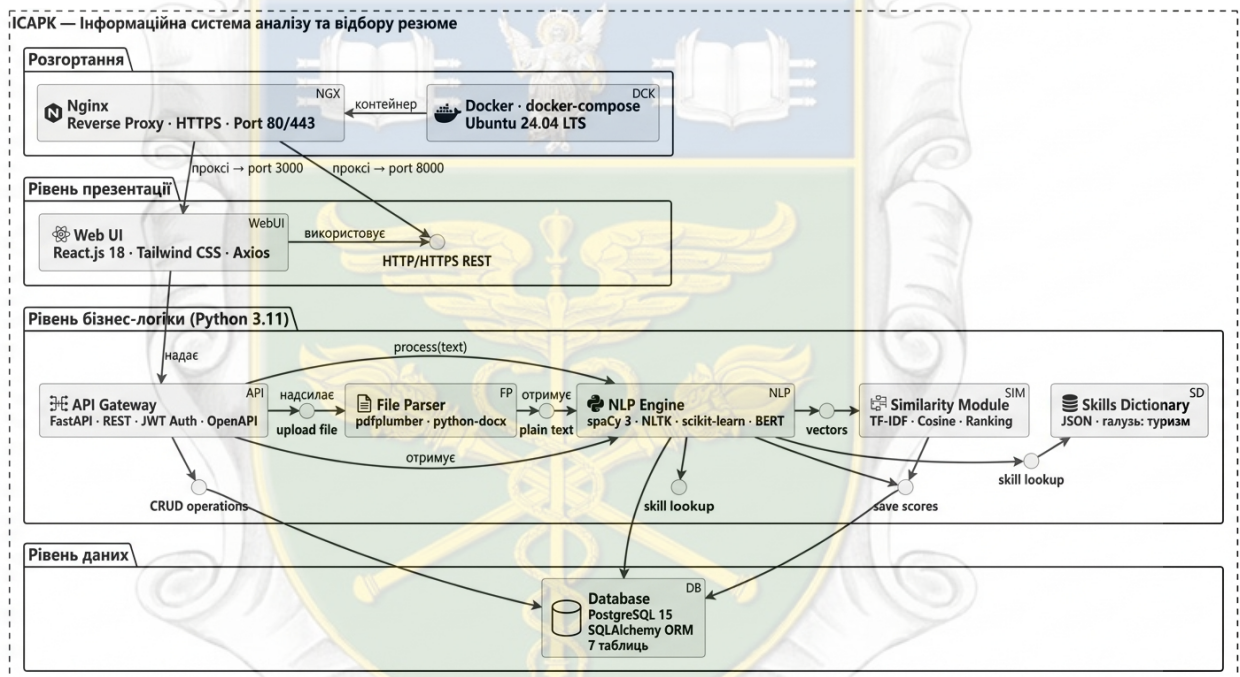


Рисунок 3.4 – Діаграма компонентів ICAPK (UML Component Diagram)

3.2. Проектування бази даних та інтерфейсу системи

Для зберігання даних ICAPK використовується реляційна база даних PostgreSQL. Структура БД розроблена з урахуванням принципів нормалізації (третья нормальна форма) та вимог до продуктивності. Схема бази даних включає такі основні таблиці: [28]

Таблиця 3.3 – Структура бази даних ICAPK

Таблиця	Поля	Призначення
candidates	id (PK), name, email, phone, created_at	Основна інформація про кандидатів
resumes	id (PK), candidate_id (FK), raw_text, file_path, upload_date, parsed BOOLEAN	Текст і мета-дані завантажених резюме
skills	id (PK), name, category	Довідник навичок (hard skills, soft skills, languages)
resume_skills	resume_id (FK), skill_id (FK), level	Зв'язок резюме і навичок кандидата
vacancies	id (PK), title, description, department, status, created_at	Опис вакансій компанії
vacancy_skills	vacancy_id (FK), skill_id (FK), is_required BOOLEAN	Вимоги вакансій до навичок
analysis_results	id (PK), resume_id (FK), vacancy_id (FK), match_score FLOAT, matched_skills[], analysis_date	Результати аналізу відповідності

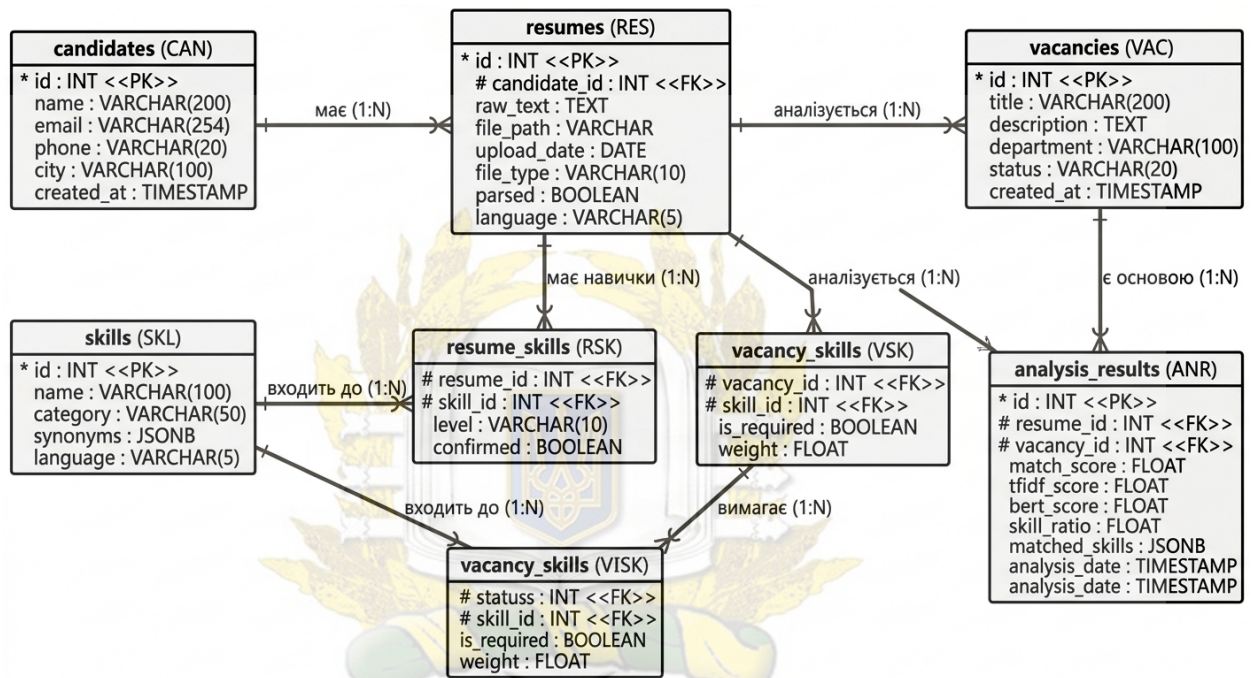


Рисунок 3.5 – Схема структури бази даних ICAPK (ER-діаграма)

Інтерфейс системи розроблено з урахуванням принципів UX-дизайну та вимог HR-менеджера. Система має такі основні екрани (сторінки):

- Головна сторінка (Dashboard) – відображає зведену статистику: загальна кількість кандидатів у базі, кількість активних вакансій, останні результати аналізу, швидкий доступ до основних функцій;
- Сторінка завантаження резюме – форма для вибору файлів резюме (підтримка множинного вибору, drag & drop), вибору вакансії для порівняння, запуску аналізу;
- Сторінка результатів аналізу – впорядкований список кандидатів з відсотком відповідності, вилюченими ключовими даними (ім'я, контакти, навички), можливість фільтрації та сортування, експорту в Excel;
- Профіль кандидата – повна інформація з вилученого резюме, порівняння навичок з вимогами вакансії у вигляді діаграми;
- База вакансій – управління описами вакансій і вимогами;
- Пошук кандидатів – пошук і фільтрація по базі з можливістю задати критерії пошуку.

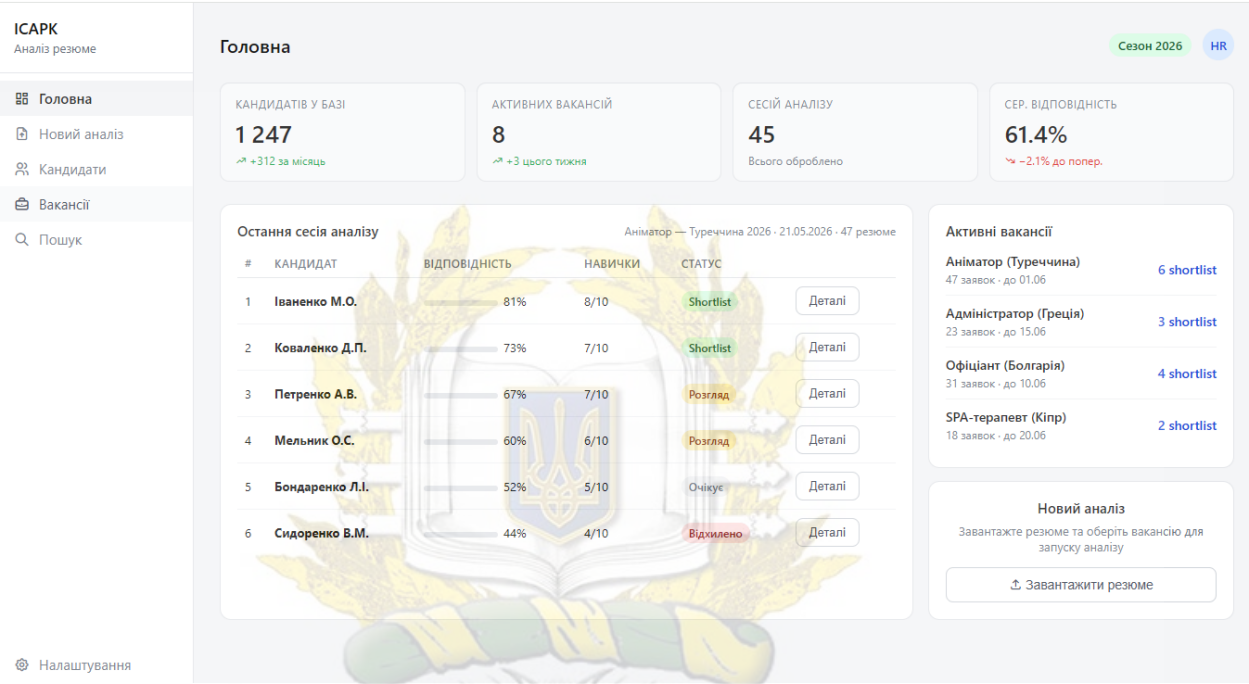


Рисунок 3.6 – Макет головної сторінки ICAPK (Dashboard)

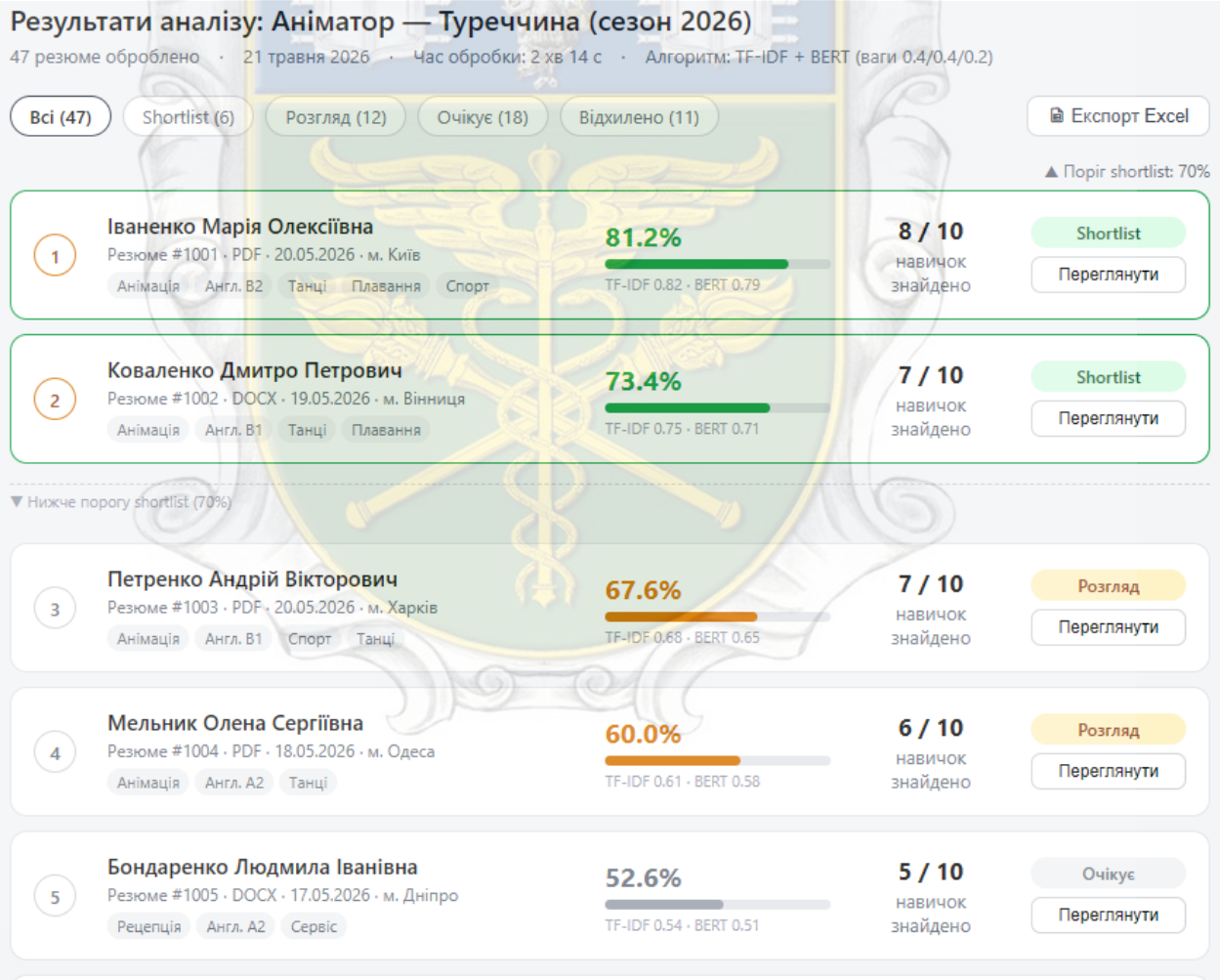


Рисунок 3.7 – Макет сторінки результатів аналізу резюме

3.3. Алгоритм інтелектуального аналізу та ранжування резюме

Алгоритм ІСАРК складається з двох основних частин: алгоритму вилучення інформації з резюме та алгоритму порівняння й ранжування кандидатів.

Алгоритм вилучення інформації (Information Extraction)

Алгоритм вилучення інформації з файлів резюме реалізується у такій послідовності:

КРОК 1. Конвертація вхідного файлу (PDF або DOCX) у текстовий формат. Для PDF використовується бібліотека pdfplumber, для DOCX – python-docx. [25]

КРОК 2. Нормалізація тексту: видалення зайвих пробілів, символів переводу рядка, спеціальних символів; приведення тексту до нижнього регістру.

КРОК 3. Токенізація: розбиття очищеного тексту на речення (sentence tokenization), а потім речень на окремі слова (word tokenization) за допомогою бібліотеки spaCy.

КРОК 4. Видалення стоп-слів. Стоп-слова – службові слова, що не несуть значущого змісту для аналізу (прийменники, сполучники, займенники тощо). Для мультимовної обробки застосовуються списки стоп-слів для англійської та української мов. [8, 3]

КРОК 5. Лематизація – приведення кожного токена до базової словникової форми (лема). Наприклад: "керував" → "керувати", "programming" → "program". Це дозволяє уніфікувати різні форми одного слова.

КРОК 6. Вилучення контактних даних за допомогою регулярних виразів: адреса електронної пошти (шаблон: [a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-

`]+\.[a-zA-Z]{2,})`, номер телефону (шаблон для UA: `\+?38?0?s?\d{2}s?\d{3}s?\d{2}s?\d{2}`).

КРОК 7. Вилучення імені кандидата за допомогою NER-моделі spaCy (розпізнавання іменованих сутностей типу PERSON).

КРОК 8. Пошук навичок (skill matching): порівняння токенів резюме з еталонним словником навичок (skills dictionary). Словник розроблено з урахуванням специфіки туристичної галузі та включає технічні навички, мовні компетенції та особистісні якості.

КРОК 9. Вилучення інформації про освіту: пошук за ключовими словами (бакалавр, магістр, університет, college, bachelor, master, спеціаліст тощо) з подальшим вилученням назви закладу та спеціальності.

КРОК 10. Структурування отриманих даних у вигляді JSON-об'єкта Resume та збереження в базі даних.

Алгоритм порівняння та ранжування кандидатів

Алгоритм порівняння резюме з вимогами вакансії реалізується у два етапи:

Етап 1 – TF-IDF векторизація та косинусна подібність:

- Всі оброблені тексти резюме та текст опису вакансії векторизуються за допомогою TF-IDF (TfidfVectorizer з бібліотеки scikit-learn);
- Для кожного резюме обчислюється косинусна подібність з вектором вакансії;
- Результат – базовий скор відповідності (base_score) від 0 до 1.

Етап 2 – Семантичне порівняння на базі BERT (Sentence Transformers):

- Використовується модель multilingual-e5-base або paraphrase-multilingual-MiniLM-L12-v2 для генерації ембедінгів речень; [9]
- Обчислюється семантична подібність між ключовими фрагментами резюме і вимогами вакансії;
- Результат – семантичний скор (semantic_score) від 0 до 1.

Фінальний скор відповідності обчислюється за формулою:

$$\text{final_score} = 0.4 \times \text{base_score} + 0.4 \times \text{semantic_score} + 0.2 \times \text{skill_match_ratio}$$

де *skill_match_ratio* – частка знайдених обов'язкових навичок з вакансії у резюме кандидата. Вагові коефіцієнти (0.4, 0.4, 0.2) визначені емпірично і можуть бути налаштовані HR-менеджером залежно від пріоритетів конкретної вакансії. [7]

Таблиця 3.4 – Приклад результатів ранжування кандидатів (демо-дані)

Рейтинг	Кандидат	TF-IDF скор	BERT скор	Навички	Фінальний скор
1	Іваненко М. О.	0.82	0.79	8/10 (80%)	81.2%
2	Коваленко Д. П.	0.75	0.71	7/10 (70%)	73.4%
3	Петренко А. В.	0.68	0.65	7/10 (70%)	67.6%
4	Мельник О. С.	0.61	0.58	6/10 (60%)	60.0%
5	Бондаренко Л. І.	0.54	0.51	5/10 (50%)	52.6%

← Назад до результатів



**Іваненко Марія
Олексіївна**
Кандидат #1001 · Резюме PDF

✉ m.ivanienko@gmail.com
☎ +380 97 123 45 67
📍 Київ, Україна
📅 Завантажено: 20.05.2026

Відповідність навичок 8 знайдено / 2 відсутніх

✓ Анімація / розваги	Обов'язк.	✓ Англійська B2	Обов'язк.
✓ Танці	Обов'язк.	✓ Плавання	Обов'язк.
✓ Спорт та фітнес	Бажана	✓ Перша мед. допомога	Бажана
✓ Водійське посвідчення	Бажана	✓ Командна робота	Soft skill
✗ Грецька мова A1	Бажана	✗ Досвід з дітьми	Бажана

Освіта

- Бакалавр з туризму та готельного бізнесу
КНТЕУ, Київ
2019 – 2023

81.2%
відповідність вакансії
Аніматор — Туреччина 2026

☆ Shortlist

⬇ Скачати резюме

📅 Запросити на інтерв'ю

✗ Відхилити

Досвід роботи 2 місяця

- Аніматор готельного клубу
Готель Ibis · Анталія, Туреччина
травень 2023 – жовтень 2023 · 6 місяців
- Менеджер рецепції
Готель «Академія» · Київ
вересень 2022 – травень 2023 · 8 місяців

Рисунок 3.8 – Скріншот профілю кандидата в ICAPK

Порівняння навичок кандидата з вимогами вакансії наочно візуалізується на сторінці профілю у вигляді кругової діаграми (pie chart) та переліку знайдених/відсутніх навичок, що значно спрощує роботу HR-менеджера.



Рисунок 3.9 – Діаграма відповідності навичок кандидата вимогам вакансії

Технічний стек проєктованої системи включає:

Таблиця 3.5 – Технічний стек ICAPK

Рівень	Технологія/Інструмент	Призначення
Backend	Python 3.11, FastAPI	Серверна логіка та REST API
NLP	spaCy 3.x, NLTK, scikit-learn, sentence-transformers	Обробка тексту, TF-IDF, BERT-ембедінги
File Parsing	pdfplumber, python-docx	Конвертація PDF/DOCX у текст
Database	PostgreSQL 15, SQLAlchemy (ORM)	Зберігання та управління даними
Frontend	React.js 18, Tailwind CSS	Веб-інтерфейс користувача
Deployment	Docker, Docker Compose	Контейнеризація та розгортання [29]
API Docs	Swagger / OpenAPI 3.0	Документування API [12]

ВИСНОВКИ

У кваліфікаційній роботі вирішено актуальну науково-практичну задачу проєктування інформаційної системи для аналізу та відбору резюме кандидатів із застосуванням методів інтелектуального аналізу текстів на матеріалах ТОВ «ТРЕВЕЛ СТАДІ ВОРК». За результатами виконаної роботи можна зробити такі висновки:

1. Досліджено теоретичні основи інтелектуального аналізу текстів та методів NLP у контексті рекрутингу. Встановлено, що ручний відбір резюме є неефективним при масовому рекрутингу: на одну вакансію надходить 200-300 заявок, на обробку яких HR-менеджер витрачає 10-23 годин. Ключовими методами аналізу резюме є: токенизація, лематизація, видалення стоп-слів, NER, TF-IDF, Word2Vec, BERT та косинусна подібність.

2. Проведено аналіз п'яти програмних аналогів (Trakstar Hire, CVVIZ, Employa, HireVue, Workable) і виявлено, що жоден з них не є повністю адаптованим до специфіки рекрутингу в туристичній галузі, не підтримує мультимовну обробку українською мовою і не є відкритим рішенням для кастомізації.

3. Охарактеризовано діяльність ТОВ «ТРЕВЕЛ СТАДІ ВОРК» – провідного українського агентства з організації міжнародних стажувань (300+ учасників у 2025 р., 55+ університетів-партнерів, 50+ готелів у 12 країнах). Фінансовий аналіз за 2023-2025 роки показав стрімке зростання виручки (з 126,80 до 486,90 тис. грн, або в 3,8 рази), але й наростання збитків (до 37,70 тис. грн у 2025 р.).

4. Аналіз HR-процесів підприємства виявив критичні недоліки: відсутність централізованої бази кандидатів, ручна обробка 200+ резюме на тиждень у пік-сезон малою командою (1-2 особи), суб'єктивність відбору. За

оцінкою, впровадження системи скоротить час обробки резюме в 40-50 разів і потребу в HR-ресурсах у 4 рази.

5. Спроектовано архітектуру ICAPK на базі клієнт-серверної моделі: серверна частина (Python FastAPI), NLP-ядро (spaCy, scikit-learn, sentence-transformers), база даних (PostgreSQL), веб-інтерфейс (React.js). Розроблено UML-діаграми: варіантів використання, послідовності, класів і компонентів.

6. Спроектовано реляційну базу даних з 7 основними таблицями (candidates, resumes, skills, resume_skills, vacancies, vacancy_skills, analysis_results). Розроблено макети 6 основних екранів інтерфейсу системи з орієнтацією на зручність роботи HR-менеджера.

7. Розроблено двоетапний алгоритм аналізу резюме: алгоритм вилучення інформації (10 кроків: конвертація → нормалізація → токенізація → лематизація → NER → регулярні вирази → skill matching) та алгоритм ранжування (TF-IDF + косинусна подібність + BERT-ембедінги). Фінальний скор відповідності розраховується за формулою: $0.4 \times \text{TF-IDF} + 0.4 \times \text{BERT} + 0.2 \times \text{skill_match}$.

Практична цінність результатів роботи полягає в тому, що спроектована система дозволяє ТОВ «ТРЕВЕЛ СТАДІ ВОРК» автоматизувати рекрутинговий процес, підвищити якість відбору кандидатів на програми міжнародних стажувань та суттєво скоротити операційні HR-витрати, що є критично важливим в умовах збиткової операційної діяльності підприємства.