

Микола Губицький, Олександр Сторожук*, Ірина Борецька*

¹ НЛТУ України, Львів, Україна, ORCID: 0009-0001-4660-0613,
m.hubytskyi@gmail.com

² НЛТУ України, Львів, Україна. ORCID: 0000-0001-6566-5271,
storozhuk@nltu.edu.ua

³ НЛТУ України, Львів, Україна, ORCID: 0000-0002-6767-104X,
boretska@nltu.edu.ua

Математична модель інтелектуального аналізу резюме для систем автоматизованого рекрутингу

***Анотація.** У статті представлено математичну модель інтелектуального аналізу резюме, що забезпечує автоматизоване зіставлення кандидатів із вакансіями за допомогою технологій штучного інтелекту, обробки природної мови та хмарних сервісів аналізу документів. Розроблено архітектуру вебсистеми, яка поєднує модулі обробки PDF-документів, семантичного порівняння текстів і ранжування кандидатів. У реалізації використано стек NestJS, React, PostgreSQL, Drizzle ORM, OpenAI API та Azure Document Intelligence. Проведені експерименти підтверджують ефективність запропонованого підходу у підвищенні точності та швидкості рекрутингу.*

***Ключові слова** – інтелектуальний аналіз, рекрутинг, штучний інтелект, NLP, NestJS, Azure Document Intelligence.*

Вступ. Автоматизація процесу підбору персоналу є важливим напрямом цифрової трансформації бізнесу. Значний обсяг резюме, що надходять до HR-відділів, ускладнює попередній відбір, вимагаючи розробки систем, здатних інтелектуально аналізувати зміст документів і порівнювати їх із вимогами вакансій.

Сучасні підходи до автоматичного аналізу текстів базуються на використанні моделей штучного інтелекту (AI) і методів обробки природної мови (NLP) [5].

Запропоновано математичну модель, яка дозволяє формалізувати процес оцінювання відповідності кандидатів, а також вебсистему, що реалізує цю модель на практиці.

Постановка задачі. Необхідно побудувати модель M , яка для множини резюме R та множини вакансій V визначає ступінь відповідності. Модель має враховувати семантичну схожість описів навичок і досвіду, релевантність ключових компетенцій та вагові коефіцієнти важливості [1, 2].

Математична модель та алгоритм. Кожен документ представляється у векторному просторі ознак. Для оцінки подібності використовується косинусна метрика. Вектори формуються на основі embedding-моделей від OpenAI [3], що дозволяють відображати семантичні зв'язки між словами [1, 2].

Для обробки резюме застосовується Azure Document Intelligence [4], який автоматично витягує структуровані дані з PDF або зображень.

Ступінь відповідності між резюме та вакансією визначається косинусною мірою подібності між відповідними векторними представленнями.

$$M(r_i, v_j) = \frac{r_i \cdot v_j}{\|r_i\| \|v_j\|} \quad (1)$$

де: $r_i = [s_1, s_2, \dots, s_k]$ – вектор ознак резюме,

$v_j = [w_1, w_2, \dots, w_k]$ – вектор ознак вакансії,

$r_i \cdot v_j$ – скалярний добуток векторів,

$\|r_i\|$ та $\|v_j\|$ – їхні евклідові норми,

k – розмірність embedding-простору.

Для побудови векторних представлень використано сучасну модель text-embedding-3-small, що забезпечує компактні та високоточні семантичні embeddings.

Отримані документи автоматично нормалізуються: сервіс Azure Document Intelligence структурує дані за секціями, що спрощує подальшу обробку.

Архітектура інтелектуальної системи. Розроблена вебсистема реалізує модель у вигляді мікросервісної архітектури з компонентами: React (frontend), NestJS (backend), Azure Document Intelligence [4], OpenAI API [3], PostgreSQL + Drizzle ORM (база даних) та аналітичний модуль для оцінки ефективності.

- **Frontend (React)** – інтерфейс користувача для завантаження резюме, перегляду результатів та аналітики.
- **Backend (NestJS)** – REST API, що координує роботу всіх модулів, забезпечує авторизацію, логіку підбору та обробку запитів.
- **Parser Service (Azure Document Intelligence)** – автоматично розпізнає структуру резюме, витягує текстові дані, розділяє секції (досвід, освіта, навички).
- **Semantic Engine (OpenAI API)** – обчислює embeddings і семантичну схожість між профілями кандидатів та вакансіями.
- **Database Layer (PostgreSQL + Drizzle ORM)** – забезпечує гнучке зберігання даних, ведення історії запитів і аналітику.
- **Analytics Module** – агрегує статистику точності відбору, формує рекомендації для подальшого вдосконалення моделі.

Результати та обговорення. Система дозволяє автоматично обробляти та класифікувати резюме, будувати семантичні профілі та зіставляти їх з вакансіями [2].

Оцінювання проводилося на наборі з 1000 резюме, попередньо приведених до уніфікованої структури за допомогою Azure Document Intelligence. Тестування показало скорочення часу на обробку до менш ніж 3 хв та підвищення точності відбору на 30%.

Використання NestJS забезпечило модульність і швидкодію, а Drizzle ORM спростило роботу з даними [4].

Висновок. Запропонована математична модель інтелектуального аналізу резюме забезпечує ефективну автоматизацію початкового етапу рекрутингу. Поєднання лінгвістичного парсингу та глибоких мовних моделей дозволяє підвищити точність оцінки відповідності кандидатів, зменшити навантаження на HR-фахівців і прискорити процес підбору персоналу.

У подальших дослідженнях увагу буде зосереджено на вдосконаленні нейронних архітектур та розробленні методів інтеграції мультиагентних систем

у процес рекрутингу, що дозволить підвищити рівень автономності та адаптивності системи.

Список використаних літературних джерел

- [1] Mikolov T. et al. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781, 2013.
- [2] Devlin J. et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL, 2019.
- [3] OpenAI. Embeddings API Documentation. 2024.
- [4] Microsoft Azure. Document Intelligence Overview. 2024.
- [5] Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Jurafsky, D. & Martin, J. H. Third Edition draft edition, 2024.

A mathematical model of resume intelligence analysis for automated recruiting systems

Mykola Hubytskyi, Oleksandr Storozhuk, Iryna Boretska

The paper presents a mathematical model for intelligent resume analysis, enabling automated candidate–vacancy matching using artificial intelligence, natural language processing, and cloud-based document analysis. The developed web system architecture integrates modules for PDF parsing, semantic similarity evaluation, and candidate ranking. The implementation is based on NestJS, React, PostgreSQL, Drizzle ORM, OpenAI API, and Azure Document Intelligence. Experimental results demonstrate the effectiveness of the proposed approach in improving recruitment accuracy and efficiency.