ТЕХНОЛОГІЇ РОЗРОБКИ МУЛЬТИМОДАЛЬНОЇ СИСТЕМИ ПІДТРИМКИ HR-РІШЕНЬ ДЛЯ ОЦІНКИ КАНДИДАТІВ НА ОСНОВІ АНАЛІЗУ ОСОБИСТІСНИХ ХАРАКТЕРИСТИК

Павленко О.С., Чала Л.Е

Кафедра штучного інтелекту,
Харківський національний університет радіоелектроніки,
Україна

E-mail: oksana.pavlenko@nure.ua,
larysa.chala@nure.ua

Abstract

This article adressed the task of creating a system of HR-solution support for assessment of candidates based on the analysis of personal characteristics. The main task was to examine, analyze and compare existing tools and approaches to text analysis. Suitable models will help to obtain the most accurate personality analysis. In the course of the work, four main models for text analysis were examined: Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Transformers. The working principles, advantages and disadvantages of each model were investigated. As a result, it was determined which model is most suitable for the development of the specified system. At the end, a complete diagram of the system for personality analysis is presented, which shows the use of several transformer models that perform the tasks set.

Розвиток цифрових технологій суттєво вплинув на формування ринків праці. Зокрема зазнав змін обсяг задач працівника у сфері управління персоналом (HR). На сьогодні процес подачі заявки на вакансію став більш оперативним, а HR може аналізувати в реальному часі дані щодо великої кількості. При цьому необхідно оцінювати не лише професійні навички, але й аналізувати особистісні якості кандидатів, їх емоційну стабільність та комунікаційний стиль. Доцільним слід вважати створення автоматизованої HR-системи, яка здатна виконувати особистісний аналіз кандидата на основі текстів резюме та запису співбесід.

При розробці такої системи необхідно обрати технології та методи, при використанні яких буде отримано найбільш точний аналіз особистості, що передбачає насамперед порівняння існуючих підходів до аналізу текстів.

В сучасних нейромережевих технологіях аналізу текстів найчастіше використовуються:

- рекурентні нейронні мережі (RNN);
- довга короткочасна пам’ять (LSTM);
- вентильні рекурентні вузли (GRU);
- трансформери [1].

RNN (Recurrent Neural Networks) є мережами прямого зв'язку з внутрішньою пам'яттю, що допомагає передбачати наступний елемент аналізованої послідовності. Рекурентна природа RNN дозволяє реалізовувати таке передбачення на основі аналізу прихованих станів вхідної послідовності. RNN обробляють вхідні дані, використовуючи свою внутрішню пам'ять, на відміну від інших мереж прямого зв'язку, які безпосередньо застосовують операції перетворення заданих вхідних даних. Таким чином, на прогноз моделі впливають результати її корекції протягом попереднього кроку в часі. У даного методу є як переваги, так і недоліки. Завдяки своїй пам’яті про попередні стани, RNN можуть моделювати залежності між словами у реченнях, що важливо для таких завдань, як аналіз настроїв, машинний переклад або розпізнавання мовлення. Однак при цьому суттєвою проблемою може бути зникання та вибух градієнтів, через що мережі важко навчати на довгих послідовностях внаслідок забування моделлю інформації, яка була на початку ре-

чення чи документа. Крім того RNN працюють досить повільно, оскільки обробляють дані послідовно, крок за кроком, що унеможливлює ефективну паралельну обробку, на відміну від трансформерів. Через обмежену здатність до запам'ятовування довгострокових залежностей, RNN гірше працюють із великими текстами, що потребують глибшого розуміння.

Long Short-Term Memory – це особливий вид послідовних моделей на основі RNN, яка вирішує проблеми зникнення та вибуху градієнтів і використовується в таких випадках, як аналіз настроїв. LSTM схожа на RNN, але з введенням механізму стробування, який дозволяє їй зберігати пам'ять протягом тривалішого періоду. В основі архітектури LSTM лежить стан комірки, який передає інформацію, зберігаючи цілісність даних. Поведінка стану комірки визначається на основі трьох «воріт»: forget gate вирішує, яку інформацію забути, input gate вирішує, яку нову інформацію запам'ятати, output gate контролює, що передати далі. Таким чином, мережі LSTM експертно контролюють потік інформації через ворота та стани комірок, гарантуючи надійну продуктивність, зберігаючи при цьому ключовий зміст оригінального тексту. Хоча у даної моделі багато сильних сторін, але недоліки також присутні. Головним недоліком даної моделі є висока обчислювальна складність, адже через велику кількість параметрів та внутрішніх операцій навчання моделі LSTM є повільними і потребують значних ресурсів. Їх архітектура складніша, ніж у звичайних RNN, що робить налаштування моделі трудомістким процесом. Крім того, LSTM все ще працює послідовно, тобто не може обробляти елементи паралельно, як трансформерні моделі, що знижує швидкість обробки при роботі з великими обсягами тексту. Ще одне обмеження полягає в тому, що, попри покращену пам'ять, LSTM усе ж має межу у здатності зберігати дуже довготривалі залежності, особливо якщо текст надто великий або має складну структуру.

Gated Recurrent Unit – це механізм стробування, який використовується в рекурентних нейронних мережах. Такий підхід ефективно вирішує давню проблему зникнення градієнтів, поширених у звичайних рекурентних нейронних мережах. Структурна сутність GRU базується на механізмі подвійних воріт: воріт оновлення та воріт скидання, які дозволяють визначати, яку інформацію потрібно зберегти або забути. Модель GRU була розроблена як спрощена альтернатива LSTM і поєднує в собі більшість її переваг при меншій обчислювальній складності. Завдяки технології подвійних воріт, GRU ефективно усуває проблему зникання градієнтів і здатна запам'ятовувати довготривалі залежності у послідовностях, але при цьому її архітектура є простішою та швидшою, адже має менше параметрів і не містить окремого блоку пам'яті (стану комірки), тому навчання відбувається швидше, а результати часто є співставними з LSTM. Ще однією перевагою GRU є те, що вона краще узагальнює дані при обмеженій кількості навчальних прикладів, а також менше схильна до перенавчання. Водночас, GRU має й свої недоліки. Через спрощену структуру вона менш гнучка, і може гірше працювати на дуже довгих послідовностях, де необхідне більш точне керування процесом запам'ятовування. Відсутність окремого блоку пам'яті іноді призводить до того, що модель втрачає частину контекстної інформації на довгих інтервалах. Ще одним недоліком є те, що архітектура GRU менш інтерпретована, тому складніше зрозуміти, які саме дані модель запам'ятовує або відкидає на кожному кроці [2].

Трансформери – це тип архітектури нейронної мережі, призначений для обробки послідовного матеріалу, такого як речення або дані часових рядів. Архітектура тренсформера відрізняється від попередніх моделей своєю здатністю надавати різне значення різним частинам послідовності слів, які їй надано. Головним в архітектурі є механізм самоуваги, про який доведено, що він є корисним для довгострокових залежностей у текстах. Механізм самоуваги, які дозволяє трансформерним моделям обробляти дані паралельно, а не послідовно, як у рекурентних нейронних мережах або мережах з довгою короткочасною пам'яттю. Загальна архітектура трансформера складається з кодувальника та декодера. Кодувальник складається з кількох шарів. Кожен шар має два підшари: мультиголовний механізм самоуваги та повнозв'язана мережа прямого зв'язку. Результат кожного підшару проходить через залишкове з'єднання та нормалізацію шару, перш ніж потрапити до наступного підшару. Мультиголовна увага, тобто модуль уваги повторює обчислення паралельно, надаючи більше можливостей кодувати нюанси значень слів. Оцінка уваги обчислюється шляхом об'єднання подібних обчислень уваги. Декодер має подібну до кодувальника структуру, але з однією відмінністю. Тут використовується замаскована мультиголовна увага. Основні компоненти декодера включають в себе шар замаскованої самоуваги, що подібний до шару самоуваги кодувальника, але включає маскування, після цього йде звичайний шар самоуваги та нейронна мережа прямого зв'язку. Замаскована самоувага означає, що при використанні такої

техніки, майбутні токени будуть приховані від моделі під час навчання. Причина застосування маскування полягає в тому, що подання всієї послідовності в модель одночасно створює проблему, в якій модель одразу знає наступне слово та не навчається його прогнозувати. Тому маскування видаляє наступне слово з послідовності, що дозволяє моделі його прогнозувати. На рисунку 1 представлено базову модель трансформера.

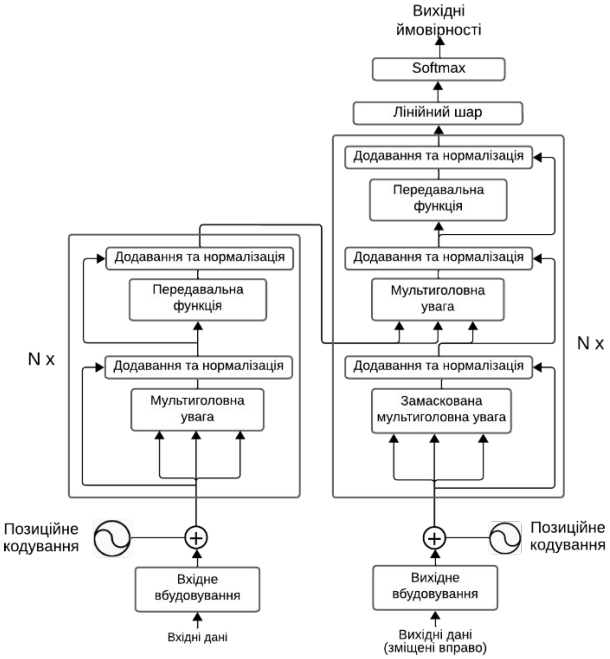

Рис. 1. Базова модель трансформера [3]

До переваг трансформерних моделей слід віднести їх незалежність від послідовної обробки, внаслідок чого вони можуть працювати паралельно, що суттєво прискорює навчання навіть на великих обсягах текстових даних. Трансформери також добре враховують контекст, тому що кожне слово розглядається у взаємозв'язку з усіма іншими словами в реченні, що забезпечує глибоке семантичне розуміння. Крім того, трансформери підтримують трансферне навчання, тобто можна використовувати вже попередньо натреновані моделі та донавчати їх на власних даних, що значно скорочує час і ресурси для досягнення високої якості результатів в особистих проектах. Недоліком таких моделей є висока обчислювальна вартість, тому що такі моделі потребують великої кількості пам'яті, потужних графічних процесорів і тривалого часу навчання. Незважаючи на це, інтеграція готових моделей є простішою за попередні підходи. Таким чином, використання трансформерів наразі є найефективнішим підходом до аналізу текстів. Такі моделі забезпечують високу точність, універсальність і здатність до глибокого розуміння мовлення. Завдяки паралельній обробці, здатності враховувати контекст на рівні всього документа й можливості використовувати попередньо навчені моделі, саме трансформери є найкращим вибором для виконання задач особистісного аналізу кандидатів на основі текстів резюме та запису співбесід.

Завдяки масштабованості, архітектура трансформерів легко адаптується для створення великих моделей. Одними з таких моделей є BERT (Bidirectional Encoder Representations from Transformers) та RoBERTa (Robustly Optimized BERT Pretraining Approach), які було використано для аналізу особистості.

BERT – це двонаправлений трансформер, що був попередньо навчений на немаркованому тексті для прогнозування замаскованих токенів у реченні та для прогнозування того, чи йде одне речення за іншим. Модель BERT використовує новий підхід, в якому одночасно враховуються обидва напрямки, що дозволяє моделі фіксувати контекстну інформацію як з попередніх, так і з наступних слів для кожного слова в послідовності [4]. Такий підхід дозволяє моделі глибше розуміти значення, відтінки та взаємозв'язки між словами, що особливо важливо для завдань, пов'язаних із аналізом текстів, класифікацією, емоційною оцінкою та розпізнаванням рис особи-

стості. В даному випадку модель на основі BERT була використана для виокремлення рис особистості за моделлю «Big Five». Використана модель «Minej/bert-base-personality» аналізує мовні патерни, частоту вживання певних слів, структуру речень та емоційне забарвлення тексту, щоб прогнозувати індивідуальні психологічні характеристики автора. Перевага цієї моделі полягає в тому, що вона поєднує потужність трансформерної архітектури BERT із психологічною теорією особистості, що робить її надзвичайно точною для задач профілювання особистості на основі тексту. Завдяки глибокому контекстному розумінню мови, модель здатна вловлювати тонкі відмінності у формулюваннях, які можуть свідчити про певні риси характеру. RoBERTa – це оптимізована похідна від моделі BERT, яка була спеціально розроблена для вирішення деяких проблем та обмежень. RoBERTa була навчена на дещо більшому датасеті протягом більшого часу, що в деяких випадках дозволяє цій модель видавати кращі результати ніж BERT. В даному випадку моделі, що засновані на RoBERTa використовувалася для оцінки емоційного забарвлення да виділення настрою.

Результати проведеного аналізу нейромережевих технологій обробки текстів дозволили запропонувати структуру системи прийняття HR-рішень. Для реалізації деяких додаткових задач цієї системи були використані трансформерні моделі від OpenAI. Зокрема для транскрибування відео співбесід використовувалась модель Whisper, а для остаточного мультимодального аналізу кандидата – модель gpt-4. Запропоновану схему роботи системи аналізу кандидатів представлено на рисунку 2.

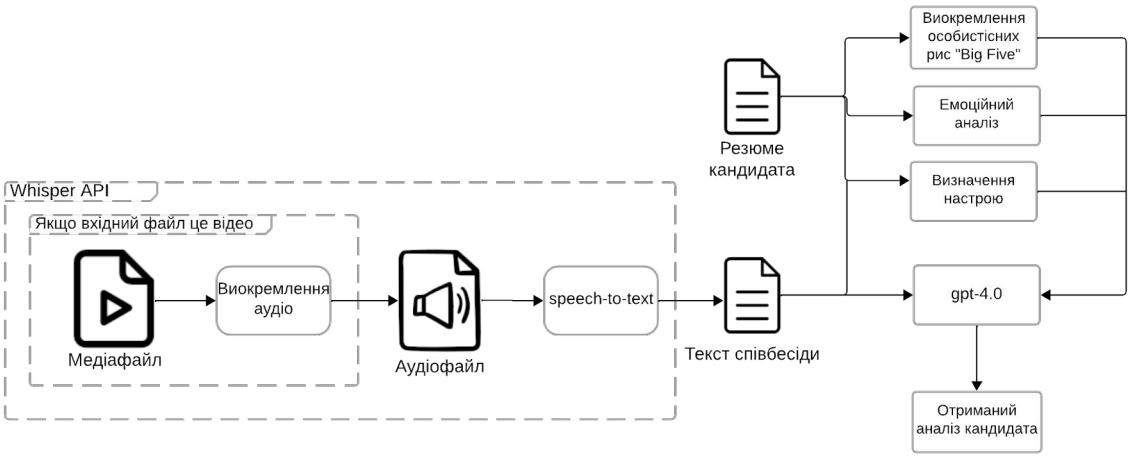

Рис. 2. Алгоритм роботи системи

Користувач спочатку завантажує два файли: резюме кандидата у форматі PDF або TXT та відео співбесіди. Відео співбесіди надсилається до моделі Whisper, де з нього виокремлюється аудіо, яке потім транскрибується у файл TXT. Застосунок обробляє обидва тексти та надсилає їх на серверну частину для аналізу. Модель HuggingFace «bert-base-personality» використовується для вилучення рис особистості «Big Five». Інші дві моделі «roberta-base-go_emotions» та «twitter-roberta-base-sentiment» використовуються для виявлення основної емоції та настрою. Після цього результати попереднього аналіуз та вхідні тексти передаються до моделі OpenAI «gpt-4» для глибшої інтерпретації загального профілю та створення узагальненого опису особистості кандидата.

Таким чином, отримані результати свідчать про доцільність використання трансформерних моделей в HR-системах аналізу особистості на основі текстових даних.

Література

1. Boesch G. Explore CNN-Based Sequence Models for Data Prediction [Електронний ресурс] // Режим доступу: https://viso.ai/deep-learning/sequential-models/.

2. Zhu J. Comparative study of sequence-to-sequence models: From RNNs to transformers // Applied and Computational Engineering. 2024. Vol. 42, No. 1. P. 67–75.

3. Vaswani A. et al. Attention is all you need // Proceedings of the 31st conference on neural information processing systems, Long Beach, CA. 2017.

4. Devlin J. et al.BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // 2018.