УДК 004.8

Овчарук О.М., Мазурець О.В.

Хмельницький національний університет

ПІДХІД ДО ВИЯВЛЕННЯ ОЗНАК ПСИХІЧНИХ РОЗЛАДІВ ЛЮДИНИ ЗА АНАЛІЗОМ КОРИСТУВАЦЬКИХ ДОПИСІВ АНСАМБЛЕМ НЕЙРОМЕРЕЖ-ТРАНСФОРМЕРІВ

Запропоновано підхід до аналізу користувацьких дописів ансамблем нейромереж-трансформерів для виявлення ознак психічних розладів людини. результатом є створений метод виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів, що відрізняється від існуючих можливістю ідентифікації одразу декількох психічних розладів без втрати точності. Метод дозволяє отримати в моделях покращену здатність розрізняти специфічні ознаки для кожного виду психічного розладу, за рахунок навчання на модифікованих множинах текстових даних. В результаті дослідження ефективності створеного методу було отримано точність виявлення ознак психічних розладів людини 85.4%.

An approach to analysis of user posts by ensemble of neural networks-transformers for identifying signs of human mental disorders is proposed. the result is a created method of identifying signs of mental disorders of a person based on the analysis of his posts by an ensemble of neural networks-transformers, which differs from the existing ones in the possibility of identifying several mental disorders at once without loss of accuracy. The method makes it possible to obtain in models an improved ability to distinguish specific signs for each type of mental disorder, due to training on modified sets of textual data. As a result of the study of the effectiveness of created method, accuracy of detecting signs of mental disorders of person was 85.4%.

На сучасному етапі обсяг користувацького контенту стрімко зростає у зв'язку із розвитком цифрових технологій і соціальних мереж [1, 2], аналіз цього контенту відкриває нові можливості щодо виявлення психічних розладів на ранніх стадіях, що дозволяє сприяти своєчасному наданню допомоги [3, 4].

Традиційні методи діагностики психічних розладів, такі як клінічні інтерв'ю та психометричні тести, мають вагомі обмеження, до прикладу вони потребують особистого контакту з пацієнтом й займають багато часу [5]. На противагу цьому, аналіз користувацького контенту в онлайн-середовищі дозволяє збирати дані в реальному часі, що підвищує точність та швидкість діагностики [6].

Метою роботи є створення методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів, що відрізняється від існуючих можливістю ідентифікації одразу кількох психічних розладів без втрати точності.

Запропонований метод виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів виконує автоматичну класифікацію текстового користувацького контенту щодо наявності в ньому психічних розладів за їх ознаками.

Схема виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів наведена на рисунку 1. Вхідними даними методу нейромережевого виявлення ознак психічних розладів людини є текст для аналізу, ансамбль моделей-трансформерів й множина відповідних токенізаторів.

На етапі 1 відбувається токенізація тексту для аналізу кожним токенізатором для переведення в векторне представлення. Токенізація включає в себе також додавання спеціальних токенів. Також доповнення або обрізання тексту до довжини заданої розмірності. В даному дослідженні максимальна довжина тексту становить 128 токенів.

На етапі 2 методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів здійснюється аналіз токенізованого кожним токенізатором тексту відповідною їм нейромережевою моделлю ансамбля. Результатом етапу є оцінки від 0 до 1 сили прояву кожного із досліджуваних психічних розладів.

Етап 3 відповідає за формування висновків про наявність кожного з п'яти досліджуваних видів психічних розладів з їх числовими мірами проявів. Психічний розлад вважається наявним, якщо його сила прояву вища за поріг. Оптимальний поріг для кожного психічного розладу визначається пошуком точки на ROC-кривій, що відповідає найкращому співвідношенню чутливості та специфічності.

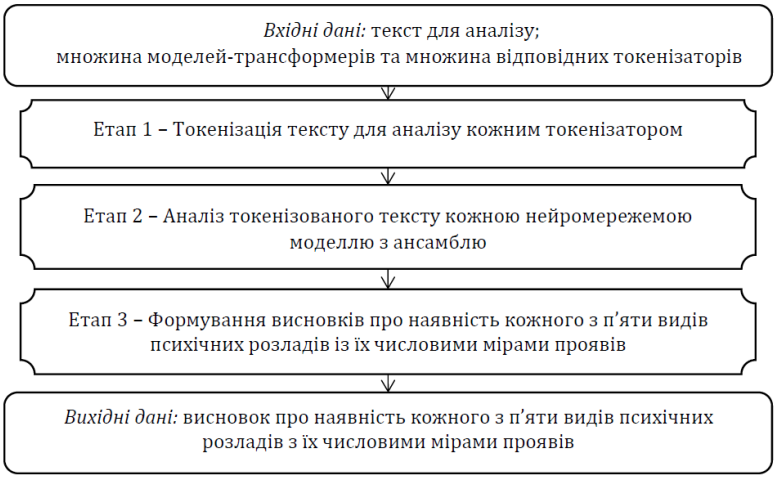

Рисунок 1 – Схема методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів

Вихідними даними методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів є висновок про наявність кожного з п'яти видів психічних розладів із їх числовими мірами проявів.

Ключовим аспектом методу методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів є формування вхідних даних, а саме – ансамблю навчених моделей-трансформерів і їх токенізаторів. Для навчання нейромереж специфічним чином формуються навчальні вибірки: цільового й нецільового класів [7]. Цільовий клас складається із виключно текстових даних з проявами певного психічного розладу. Для запобіганню плутанини психічних розладів між собою та враховуючи, що в певному тексті можуть бути й прояви інших розладів, в нецільовій категорії записи формуються за певними правилами: кількість дописів у нецільовій категорії відповідає або наближається до цільової (похибка не більше 10 дописів); нецільова категорія складається з рівних пропорцій текстів з іншими видами психічних розладів та текстів що не містять таких проявів, або містять їх в дуже маленькій мірі (від 0.0 до 0.3 за шкалою від 0 до 1).

Для дослідження ефективності методу виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів було порівняно існуючий підхід до мультикласової класифікації з існуючим підходом мультилейблової класифікації та із запропонованим методом на основі використання ансамблю бінарних класифікаторів для кожного типу психічного розладу. Отож, при багатокласовій класифікації отримано такі результати за макро-метриками: Accuracy: 0.854, Precision: 0.867, Recall: 0.854, F1-score: 0.854. Відповідна матриця плутанини наведена на рисунку 2.

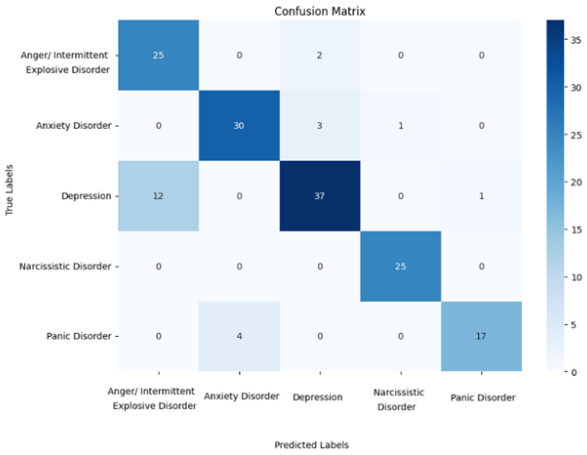

Рисунок 2 – Матриця сплутувань для мультикласифікації щодо методу виявлення ознак психічних розладів людини за аналізом її дописів

Отримані результати свідчать про високу ефективність використання ансамблю бінарних моделей для класифікації психічних розладів на основі текстових даних в методі виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів. Високі значення метрик (понад 84%) вказують на здатність моделі здебільшого правильно ідентифікувати різні розлади. Однак, існують певні труднощі з класифікацією деяких розладів, зокрема «Anxiety Disorder» та «Depression», які можуть бути пов'язані з їх клінічною подібністю. Це вимагає подальших досліджень для покращення моделі та залучення додаткових даних або інших характеристик тексту для підвищення точності класифікації.

Подальші дослідження будуть спрямовані на розширення наборів навчальних даних та дослідження інших видів психічних розладів, окрім 5-и розглянутих.

Загалом, створений метод виявлення ознак психічних розладів людини за аналізом її дописів ансамблем нейромереж-трансформерів дозволяє отримати в моделях покращену здатність розрізняти специфічні ознаки для кожного виду психічного розладу, за рахунок навчання на модифікованих множинах текстових даних, що зменшує ймовірність плутанини між розладами, оскільки модель вчиться розрізняти їхні характерні ознаки.

Перелік посилань

1. Mazurets O., Molchanova M., Klimenko V., Prosvitliuk M Practice Implementation of Neural Network Model BART-Large-CNN for Text Annotation. Prospects of Scientific Research in the Conditions of the Modern World. Proceedings of XXVII International scientific and practical conference. Rotterdam, Netherlands. 2024. Pp. 97-102.
2. Slobodzian V., Kovalchuk O., Molchanova M., Sobko O., Mazurets O., Barmak O., Krak I. Text Data Vectorization Model of Ukrainian-Language Internet Communication Content. CEUR Workshop Proceedings, 2022, vol. 3171, pp. 561–571.
3. Sobko O., Mazurets O., Didur V., Chervonchuk I. Recurrent Neural Network Model Architecture for Detecting a Tendency to Atypical Behavior Of Individuals by Text Posts. Theoretical and Practical Aspects of Modern Research. Proceedings of XXVI International scientific and practical conference. International Scientific Unity. Ottawa, Canada. 2024. Pp. 113-117.
4. Овчарук О.М., Мазурець О.В., Молчанова М.О., Дідур В.О. Метод прогнозування значень параметрів за їх часовими рядами рекурентною нейронною мережею із згортковим шаром. Матеріали IV Всеукраїнської науково-технічної конференції «Комп'ютерні ігри та мультимедіа як інноваційний підхід до комунікації - 2024». Одеса. 26-27 вересня 2024. С. 306-309.
5. Slobodzian V., Molchanova M., Kovalchuk O., Sobko O., Mazurets O., Barmak O., Krak I. An Approach Based on the Visualization Model for the Ukrainian Web Content Classification. 2022

12th International Conference on Advanced Computer Information Technologies, ACIT 2022. 2022. pp. 400-405.

6. Молчанова М.О., Мазурець О.В., Собко О.В., Кліменко В.І., Андрощук В.І. Метод нейромережевого виявлення кібербулінгу з використанням хмарних сервісів та об'єктно-орієнтованої моделі. Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2024. №2 (333). С. 200-206.
7. Овчарук О.М., Мазурець О.В. Прогнозування значень параметрів за їх часовими рядами рекурентною темпоральною нейронною мережею. Збірник наукових праць за матеріалами XV Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2023». Хмельницький, 2023. с. 227-232.