

МЕТОД АНАЛІЗУ ПРОЯВІВ ПОСТТРАВМАТИЧНОГО СТРЕСОВОГО РОЗЛАДУ ЗА НЕЙРОМЕРЕЖЕВИМ АНАЛІЗОМ ТЕКСТІВ

Овчарук О.М., off4aruk@gmail.com,

Мазурець О.В., exe.chong@gmail.com

Хмельницький національний університет

На тлі сучасних подій, які супроводжуються значним стресом, таких як військові конфлікти, стихійні лиха чи соціальні заворушення, зростає кількість випадків посттравматичного стресового розладу серед населення [1]. Традиційні методи діагностики виявляються малоефективними, оскільки вони вимагають прямої взаємодії з пацієнтом та значного часу для оцінки його стану.

З огляду на зростання доступності текстових даних, актуальною стає розробка методології створення навчальних наборів, які дозволять підвищити точність та чутливість моделей до текстових ознак ПТСР [2]. Такий підхід не лише покращить ефективність діагностики, але й сприятиме своєчасному наданню допомоги людям, які пережили травматичні події. В умовах швидких змін у соціально-психологічному середовищі використання нейронних мереж для виявлення симптомів ПТСР стає нагальною потребою, спрямованою на забезпечення цифрового психологічного добробуту суспільства.

Метою дослідження є розробка методу аналізу проявів посттравматичного стресового розладу за нейромережевим аналізом текстового контенту, що перетворює вхідні текстові дані у результат у вигляді оцінки ймовірності наявності ПТСР у користувацькому контенті.

Запропонований метод базується на використанні нейронної мережі, навченої на спеціально створеному наборі даних, який знижує ризик помилкового виявлення ПТСР у випадках інших психічних розладів [3]. Для текстового аналізу застосовувалася нейронна мережа з квантовим шаром. Архітектурна схема мережі для діагностики ПТСР за текстовими даними представлена на рис. 1.

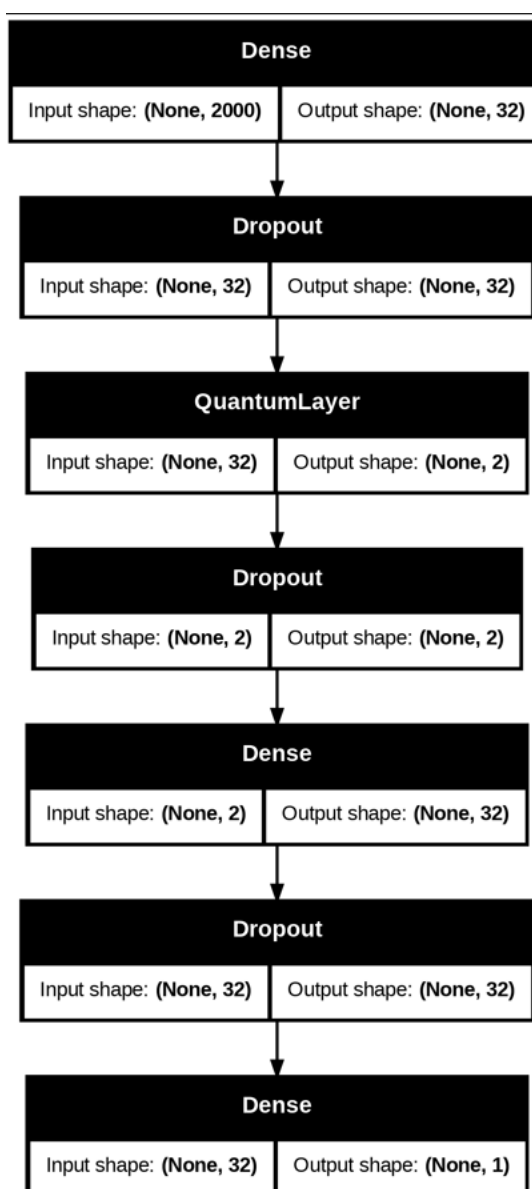


Рис. 1. Архітектура нейромережі для аналізу проявів посттравматичного стресового розладу за аналізом текстового контенту

Розроблений метод аналізу проявів посттравматичного стресового розладу за нейромережним аналізом текстового контенту дозволяє перетворювати вхідні текстові дані на оцінку ймовірності присутності ПТСР у контенті користувачів.

На першому етапі проводиться підготовка текстів, що включає їх токенізацію та перевірку на довжину й повноту даних. З текстів зберігаються розділові знаки, емодзі та інші елементи, які можуть бути важливими для діагностики ПТСР. Для токенізації використовується спеціалізований токенізатор, який відповідає моделі, застосованій у процесі навчання.

Другий етап передбачає аналіз текстів за допомогою нейронної мережі, що базується на трансформерній архітектурі, орієнтованій на контекст, для визначення проявів ПТСР.

Вхідний шар моделі, який відповідає формату вхідних даних, має розмірність 2000, що відображає довжину вектора для аналізу слів. Цей шар на схемі не зображено. Далі йдуть щільні (Dense) шари, які забезпечують повне з'єднання нейронів між шарами. Dropout-шар, що деактивує частину нейронів у процесі навчання, допомагає запобігти перенавчанню; у цьому дослідженні застосовується значення Dropout на рівні 70%.

Вихідний шар класифікації представлений одним нейроном із функцією активації «Sigmoid», яка забезпечує результат у діапазоні від 0 до 1, що дозволяє оцінювати ймовірність проявів ПТСР.

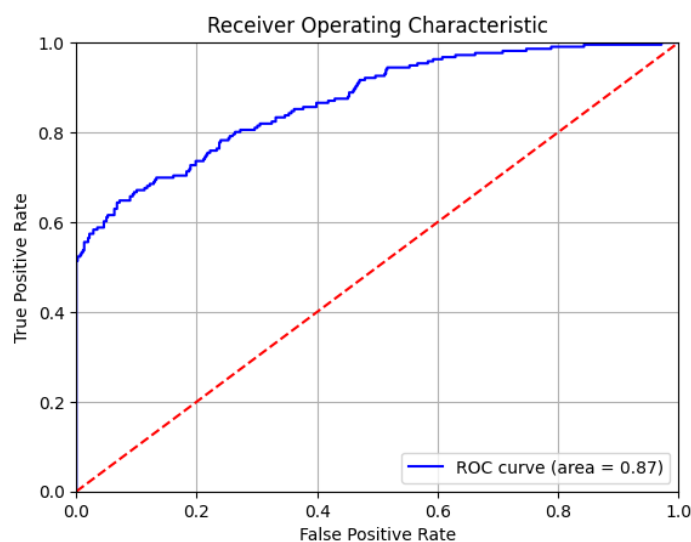


Рис. 2. ROC-криві нейронної мережі для аналізу проявів посттравматичного стресового розладу

Для підвищення точності діагностики й зменшення ризику помилкового ототожнення ПТСР із іншими психічними розладами створено спеціальний навчальний набір даних. Формування цього набору здійснювалось шляхом відбору даних із наявних ресурсів, таких як «AyaPTSD» [4] та *«HumanStressPrediction» [5], які доступні на платформі Kaggle.

Ефективність підходу була оцінена на хмарній платформі Google Colab. Результати представлені у вигляді ROC-кривої для на рисунку 2.

У результаті навчання моделі на створеному наборі даних було досягнуто наступних показників: точність склала 0.85, а площа під кривою (AUC) досягла 0.88. Навчання проводилося протягом 100 епох із використанням механізму ранньої зупинки, щоб запобігти перенавчанню та зберегти оптимальну версію моделі. Розмір навчальної партії (batch size) дорівнював 64, а словниковий запас охоплював 2000 токенів.

Отже, запропонована архітектура нейронної мережі продемонструвала значне покращення результатів у порівнянні з існуючими методами аналізу проявів посттравматичного стресового розладу за аналізом текстового контенту. Це має особливе значення для вирішення задачі виявлення проявів ПТСР у текстах. Розроблений підхід забезпечив вищу точність аналізу та зменшив ризик плутанини з іншими психічними розладами.

Список використаних джерел

1. Browning L., Rashid I., Javanbakht A. The Current State of Digital Technologies for the Treatment and Management of PTSD- A Look into the Future of Psychiatry. Preprints 2024, 2024091313. DOI: [10.20944/preprints202409.1313.v1](https://doi.org/10.20944/preprints202409.1313.v1)
2. Quillivic R., Gayraud F., Auxéméry Y. et al. Interdisciplinary approach to identify language markers for post-traumatic stress disorder using machine learning and deep learning. *Scientific Repository*. 2024. 14, 12468. DOI: [10.1038/s41598-024-61557-7](https://doi.org/10.1038/s41598-024-61557-7)
3. Мазурець О.В., Овчарук О.М. Підхід до збільшення розмірності вхідних даних для нейромережевого прогнозування значень показників епідеміологічної небезпеки. Матеріали VIII Міжнародної науково-практичної конференції «Перспективи сучасної науки: теорія і практика» (16-18 вересня 2024. Львів, Україна). 2024. С. 192-198.
4. Набір даних «AyaPTSD». URL: <https://www.kaggle.com/datasets/abdelrahmanahmed3/aya-ptsd>
5. Набір даних «HumanStressPrediction». URL: <https://www.kaggle.com/datasets/kreeshrajani/human-stress-prediction/data>