

**Мазурець О.В., Овчарук О.М.,** (Хмельницький національний університет, м. Хмельницький, Україна)

## ДІАГНОСТУВАННЯ ПРОЯВІВ ПОСТТРАВМАТИЧНОГО СТРЕСОВОГО РОЗЛАДУ ЗА НЕЙРОМЕРЕЖЕВИМ АНАЛІЗОМ ТЕКСТОВОГО КОНТЕНТУ

**Анотація:** Запропоновано підхід до діагностування проявів посттравматичного стресового розладу за нейромережєвим аналізом текстового контенту, який забезпечує трансформацію вхідних даних у вигляді тексту в результат у вигляді оцінки ймовірності проявів ПТСР у контенті користувачів. Розроблена архітектура нейронної мережі показала значне покращення результатів порівняно з існуючими аналогами. Після навчання моделі на запропонованому наборі даних було отримано точність понад 85%.

**Ключові слова:** посттравматичний стресовий розлад, нейронна мережа, обробка природної мови

**Abstract:** Approach to diagnosing manifestations of post-traumatic stress disorder using neural network analysis of textual content is proposed, which provides the transformation of input data in the form of textual data into a result in the form of an assessment of the probability of manifestations of PTSD in user content. The developed neural network architecture showed a significant improvement in results compared to existing analogues. After training the model on the proposed data set, an accuracy of over 85% was obtained.

**Keywords:** post-traumatic stress disorder, neural network, natural language processing

### Постановка проблеми

На тлі останніх подій, які викликають значний стрес, таких як війни, катастрофи чи соціальні конфлікти, спостерігається зростання випадків посттравматичного стресового розладу (ПТСР) серед населення [1, 2]. Традиційні підходи до діагностики часто виявляються недостатньо ефективними, оскільки потребують безпосереднього контакту з пацієнтом та тривалого часу для оцінки його стану.

У зв'язку з тим, що текстова інформація стає дедалі доступнішою, виникає необхідність у створенні методології формування навчальних наборів даних, яка забезпечить високу точність та чутливість моделей до текстових проявів ПТСР [3]. Це дозволить не лише підвищити ефективність діагностики, але й сприятиме своєчасному втручання та підтримці осіб, які пережили травматичні події. В умовах швидких змін у соціально-психологічному середовищі використання нейронних мереж для виявлення ознак ПТСР стає важливим і актуальним завданням, яке спрямоване на забезпечення цифрового психологічного добробуту суспільства.

Таким чином, актуальним є питання автоматизації діагностики ПТСР на основі текстового контенту, забезпечуючи при цьому стійкість до помилок у розрізненні ПТСР та інших психічних розладів. Цього можна досягти за допомогою нейронних мереж шляхом створення навчального набору даних, що включає тексти з яскраво вираженими проявами ПТСР як цільову категорію, а також тексти з іншими психічними розладами та без них – як контрольну групу.

### Аналіз останніх публікацій

Наразі широко застосовуються такі цифрові інструменти, як віртуальна реальність, нейрофідбек, носимі пристрої, мобільні додатки та машинне навчання для діагностики та лікування симптомів ПТСР [4]. Аналіз досліджень показав, що віртуальна реальність є ефективним методом лікування бойового ПТСР, тоді як результати використання мобільних додатків і нейрофідбеку залишаються обмеженими через невеликі вибірки та методологічну неоднорідність. Моделі машинного навчання демонструють потенціал у виявленні симптомів ПТСР за даними зі смартфонів і носимих пристроїв, однак для їх клінічного застосування потрібна стандартизація процедур збору даних і моделювання.

Дослідження [5] розглядає мову як можливий діагностичний біомаркер ПТСР, використовуючи дані від 148 осіб, які стали жертвами терактів у Парижі 13 листопада 2015

року. Інтерв'ю, проведені через 5-11 місяців після подій, охоплювали респондентів зі схожими соціально-економічними умовами. У дослідженні застосовувалася міждисциплінарна методологія, яка поєднувала психіатрію, лінгвістику та обробку природної мови для встановлення зв'язку між мовними особливостями та ПТСР.

Результати показали, що клінічний психіатр досяг AUC 0,72 при діагностиці ПТСР, що зіставно з результатами анкетування (AUC  $\approx$  0,80). Модель машинного навчання показала AUC 0,69, а модель глибокого навчання – AUC 0,64. У дослідженні також враховувалися змішувальні фактори, що впливають на результати, встановлювалися зв'язки між мовними особливостями та симптомами DSM-5, а також поєднувалися автоматизовані методи з якісним аналізом.

### Мета роботи та постановка завдань

Метою роботи є діагностування проявів посттравматичного стресового розладу за нейромережовим аналізом текстового контенту, що забезпечує перетворення вхідних даних у вигляді тексту в результат у вигляді оцінки ймовірності проявів ПТСР у контенті користувачів.

### Виклад основного матеріалу

Метод діагностування проявів посттравматичного стресового розладу за нейромережовим аналізом текстового контенту базується на навчанні нейронної мережі з використанням спеціально створеного навчального набору даних. Цей набір забезпечує високу стійкість до помилкового виявлення ПТСР через змішування з іншими психічними розладами. Для аналізу текстів була застосована нейронна мережа, що включає квантовий шар. Схема архітектури мережі для діагностування проявів посттравматичного стресового розладу за аналізом текстового контенту представлена на рис. 1.

Запропонований підхід до діагностування проявів посттравматичного стресового розладу за нейромережовим аналізом текстового контенту забезпечує трансформацію вхідних даних у вигляді тексту в результат у вигляді оцінки ймовірності проявів ПТСР у контенті користувачів. На першому етапі здійснюється попередня обробка текстів, включаючи їх токенизацію. Тексти перевіряються на довжину і наявність даних. При цьому зберігаються розділові знаки, емодзі та інші елементи, які можуть містити значущу інформацію для виявлення ПТСР. Процес токенизації виконується із застосуванням токенизатора, що використовувався під час навчання моделі.

На другому етапі проводиться аналіз тексту за допомогою контекстно-орієнтованої нейромережі на основі архітектури трансформера, спрямованої на виявлення проявів ПТСР.

Вхідний шар (Input) моделі визначає формат даних і має розмірність 2000, що відповідає довжині вектора слів для аналізу. На рисунку цей шар не відображено. Подальші шари включають щільні (Dense) шари, що з'єднують усі нейрони попереднього шару з кожним нейроном поточного. Dropout-шар випадково деактивує певну кількість нейронів у процесі навчання, що зменшує ризик перенавчання моделі, зокрема для даного дослідження використовується значення 70%.

Вихідний шар Dense із одним нейроном використовується для задачі класифікації на два класи. Застосовується функція активації Sigmoid, що стискає результат у межах від 0 до 1, що ідеально підходить для моделювання ймовірності проявів ПТСР.

Для забезпечення точності діагностики та зменшення ймовірності змішування ПТСР із симптомами інших психічних захворювань було створено навчальний набір даних. У ньому цільова категорія містить тексти з високим рівнем проявів ПТСР, а контрольна – включає як тексти без ознак ПТСР, так і тексти, пов'язані з іншими психічними розладами.

Формування цього набору здійснювалося шляхом комбінування та відбору даних із наявних баз [6], таких як «AyaPTSD» [7] та «HumanStressPrediction» [8], розміщених на платформі Kaggle.

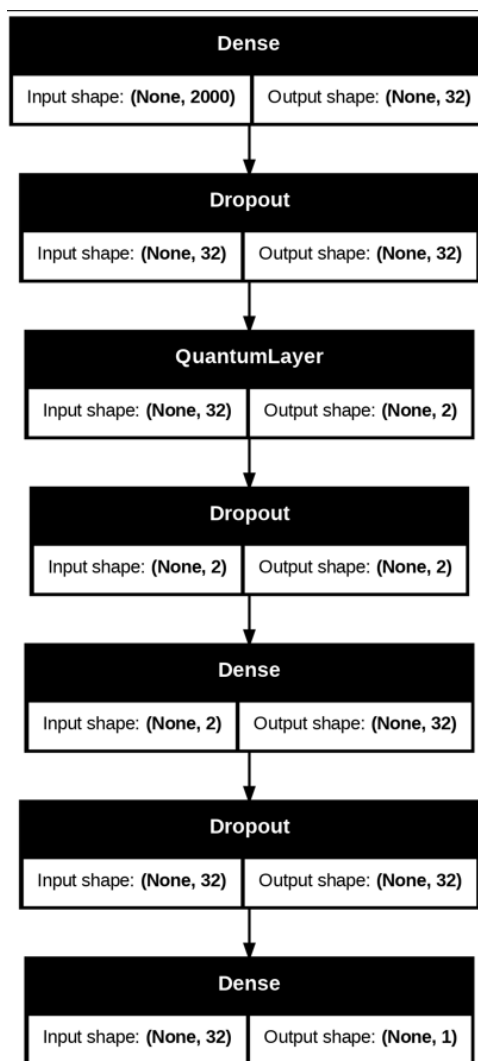


Рис. 1. Архітектура нейромережі для аналізу текстового контенту для діагностування проявів ПТСР

Ефективність методу діагностування проявів посттравматичного стресового розладу за нейромережним аналізом текстового контенту була оцінена за допомогою хмарної платформи Google Colab. На рисунках 2 і 3 продемонстровані ROC-криві для двох варіантів параметрів Dropout – 60% і 75%.

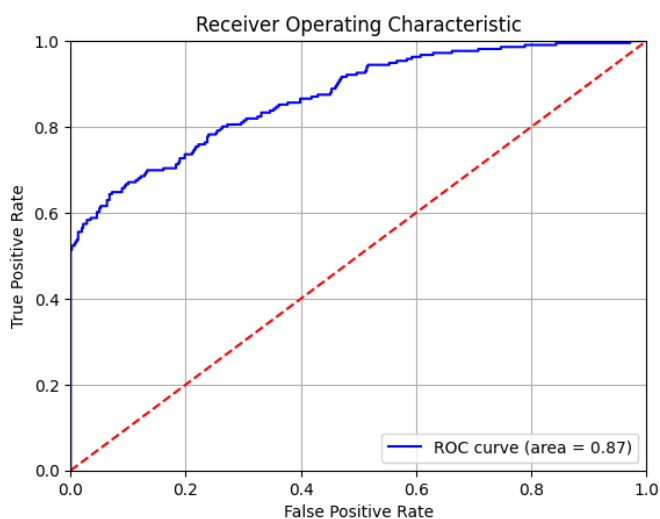


Рис. 2. ROC-криві нейронної мережі, відсоток інгібуючих нейронів 60

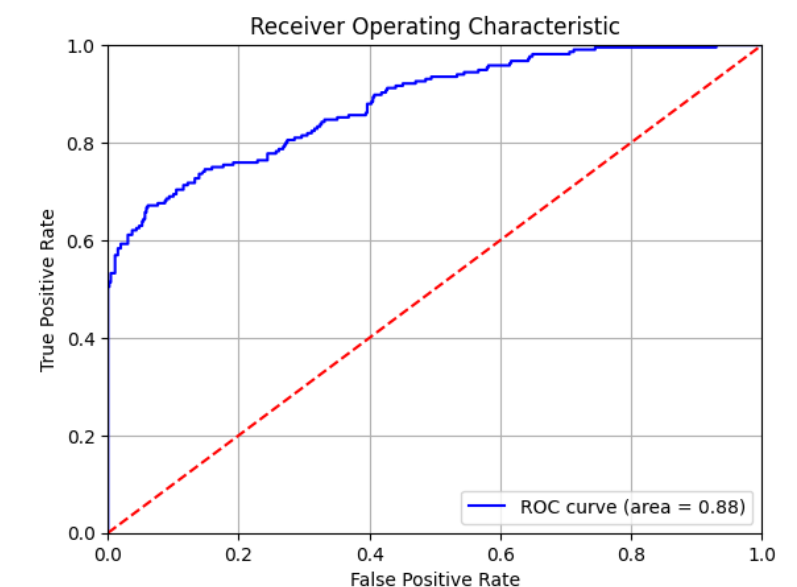


Рис. 3. ROC-криві нейронної мережі, відсоток інгібуючих нейронів 75

Після навчання моделі на запропонованому наборі даних були отримані такі результати: точність (Accuracy) – 0.85, точність передбачень (Precision) – 0.87, і площа під кривою (AUC) – 0.88. Навчання здійснювалося протягом 100 епох із застосуванням механізму ранніх зупинок для уникнення перенавчання та збереження найкращої версії моделі. Розмір навчальної партії (batch size) становив 64, а словник охоплював 2000 токенів.

## ВИСНОВКИ

Розроблена архітектура нейронної мережі показала значне покращення результатів порівняно з існуючими аналогами. Це особливо важливо для задач виявлення ПТСР у текстовому контенті. Встановлено підвищення точності і зменшення ризику сплутування з іншими психічними розладами. Однак метод діагностування проявів посттравматичного стресового розладу за нейромережовим аналізом текстового контенту має певні обмеження: він застосовується лише до текстів англійською мовою. Це обмеження можна подолати шляхом розширення навчального набору даних текстами іншими мовами.

Подальші дослідження будуть зосереджені на розширенні набору навчальних даних і тестуванні альтернативних нейромережових архітектур для підвищення ефективності моделі та точності виявлення симптомів ПТСР.

## Список посилань.

11. Krak I., Zalutska O., Molchanova M., Mazurets O., Bahrii R., Sobko O., Barmak O. *Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network*. *CEUR Workshop Proceedings*, 2024, vol. 3688, pp. 16-28. <https://doi.org/10.31110/COLINS/2024-3/002>
12. Park, A. H., Patel, H., Mirabelli, J., Eder, S. J., Steyrl, D., Lueger-Schuster, B., Scharnowski, F., O'Connor, C., Martin, P., Lanius, R. A., McKinnon, M. C., & Nicholson, A. A. (2023). *Machine learning models predict PTSD severity and functional impairment: A personalized medicine approach for uncovering complex associations among heterogeneous symptom profiles*. *Psychological Trauma: Theory, Research, Practice, and Policy*. Advance online publication. <https://doi.org/10.1037/tra0001602>
13. Zalutska O., Molchanova M., Sobko O., Mazurets O., Pasichnyk O., Barmak O., Krak I. *Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network*. *CEUR Workshop Proceedings*, 2023, vol. 3387, pp. 344–356.
14. 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. <https://doi.org/10.20944/preprints202409.1313.v1>

15. 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*. *Sci Rep* 14, 12468 (2024). <https://doi.org/10.1038/s41598-024-61557-7>
16. 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»*. 16-18.09.2024. Львів – 2024. с. 192-198.
17. Набір даних «AyaPTSD». URL: <https://www.kaggle.com/datasets/abdelrahmanahmed3/aya-ptsd>
18. Набір даних «HumanStressPrediction». URL: <https://www.kaggle.com/datasets/kreeshrajani/human-stress-prediction/data>