

УДК 004.8

Овчарук О.М.

Хмельницький національний університет

МОДЕЛЬ АНАЛІЗУ ПСИХІЧНОГО СТАНУ ГРОМАДЯН ІЗ ПОСТТРАВМАТИЧНИМ СТРЕСОВИМ РОЗЛАДОМ ЗА ПОВІДОМЛЕННЯМИ

У роботі розглянуто модель аналізу психічного стану громадян із посттравматичним стресовим розладом за текстами цифрових повідомлень. Запропонований підхід поєднує контекстно-орієнтовану трансформерну модель типу DeBERTa, бінарну детекцію ПТСР, мультилейблову класифікацію психічних розладів і формалізацію ризику соціально небезпечної поведінки на основі розширеної множини психолінгвістичних та поведінкових ознак.

The paper presents a model for analysing the mental state of citizens with post-traumatic stress disorder based on their digital text messages. The proposed approach integrates a context-aware transformer model of the DeBERTa family, binary PTSD detection, multi-label classification of mental disorders, and a formalised risk index of socially dangerous behaviour built on an extended set of psycholinguistic and behavioural features.

Посттравматичний стресовий розлад та інші психічні порушення, що виникають унаслідок травматичних подій, є серйозною проблемою охорони психічного здоров'я та соціальної безпеки [1]. Своєчасне виявлення ознак ПТСР дозволяє проводити ефективні психологічні втручання, знижувати ризики соціально небезпечної поведінки та сприяти реабілітації постраждалих [2]. Традиційні методи оцінювання психічного стану, що базуються на опитуваннях і клінічних інтерв'ю, є трудомісткими, суб'єктивними та обмеженими за масштабом [3, 4]. У цьому контексті обробка природної мови відкриває нові можливості для автоматизованого аналізу текстів цифрових комунікацій, таких як соціальні мережі, месенджери та форуми, де люди часто виявляють симптоми психічного дискомфорту [5].

Сучасні трансформерні моделі, зокрема архітектури типу DeBERTa, забезпечують ефективне представлення контексту та врахування складних семантичних і синтаксичних залежностей у тексті [6]. Це дозволяє виявляти навіть латентні прояви ПТСР та інших психічних розладів, що можуть проявлятися у формулюваннях, стилі написання або частоті певних мовних конструкцій [7]. Можливості NLP у цьому напрямку включають бінарну класифікацію на

наявність/відсутність ПТСР, мультилейблову ідентифікацію суміжних психічних станів, а також формалізацію ризику соціально небезпечної поведінки на основі комплексних психолінгвістичних та поведінкових ознак, таких як використання негативної лексики, ознаки тривожності, ізоляції чи агресивності [8].

Використання NLP дозволяє масштабувати процес оцінювання, аналізувати великі об'єми повідомлень у режимі реального часу та створювати інструменти підтримки рішень для психологів, соціальних служб і органів безпеки [9]. Перспективним напрямком є інтеграція моделі з іншими джерелами даних, наприклад мультимодальними сигналами поведінки, що підвищує точність прогнозів і робить систему більш адаптивною до різноманітних комунікаційних форматів. Додатково, можливості інтерпретації результатів трансформерних моделей відкривають шлях до прозорих і пояснюваних рішень, що підвищує довіру фахівців до автоматизованих систем оцінки психічного стану. Таким чином, NLP є ключовим інструментом для розробки сучасних інтелектуальних систем моніторингу психічного здоров'я та соціальної безпеки, здатних виявляти ПТСР і суміжні психічні стани на ранніх етапах прояву [10].

Модель аналізу психічного стану громадян із посттравматичним стресовим розладом за повідомленнями ґрунтується на інтеграції трансформерних нейромережових моделей, мульти-лейблової класифікації психічних розладів та розширеної множини психолінгвістичних і поведінкових ознак користувачького контенту. У дисертаційному дослідженні цей підхід реалізовано як послідовність трьох взаємопов'язаних методів: аналізу проявів ПТСР, виявлення інших психічних розладів та оцінювання ризику соціально небезпечної поведінки на основі розширеної множини ознак.

Текстові повідомлення подаються на вхід контекстно-орієнтованої трансформерної моделі типу DeBERTa, яка перетворює кожне повідомлення x на векторну ознакову репрезентацію $h(x)$. На відміну від класичних моделей, DeBERTa розділяє представлення значення слова і позиції в реченні, що посилює здатність моделі вловлювати тонкі контекстні й конотативні відтінки, характерні для текстів користувачів із ПТСР. Кінцевий вектор $h(x)$ можна трактувати як компактне числове відображення психолінгвістичних характеристик повідомлення, у якому акумульовано інформацію про лексику, синтаксис, стиль, емоційне забарвлення та приховані патерни мовлення.

На основі вектора $h(x)$ будується бінарний класифікатор ПТСР. Для кожного повідомлення x модель повертає оцінку $s_PTSD(x)$ у діапазоні від 0 до 1, яку інтерпретують як ймовірнісну міру прояву ПТСР у даному фрагменті тексту. Якщо $s_PTSD(x)$ перевищує заздалегідь обраний поріг τ_PTSD , то повідомлення відносять до групи з ознаками ПТСР; якщо нижче, до групи без виявлених ознак.

Навчання такої моделі здійснюється на спеціально сформованій вибірці, де кожне повідомлення має експертну позначку «ПТСП / не ПТСП», а цільова функція мінімізує середню логарифмічну втрату між $s_PTSD(x)$ та еталонною міткою. У дисертації показано, що включення квантового шару з інгібувальними нейронами та тонка настройка DeBERTa дають змогу досягти точності близько 0,85, точності за позитивним класом (Precision) до 0,87 та площі під ROC-кривою AUC близько 0,88, що перевищує базові архітектури, для яких точність утримується на рівні близько 0,80.

Друга компонента моделі пов'язана з мультилейбловою класифікацією психічних розладів. Замість того, щоб примусово обирати один домінуючий розлад, вводиться множина діагностичних категорій $C = \{c_1, c_2, \dots, c_K\}$, до якої входять ПТСП, депресивні, тривожні, суїцидальні та інші розлади. Для кожного повідомлення x формується вектор міток $y(x) = (y_1(x), \dots, y_K(x))$, де $y_k(x)$ дорівнює 1, якщо текст містить ознаки k -го розладу, і 0 в іншому випадку. Мультилейблова модель повертає вектор оцінок $s(x) = (s_1(x), \dots, s_K(x))$, де $s_k(x) \in [0;1]$ є мірою впевненості в наявності відповідного розладу. Навчання здійснюється за узагальненою логістичною функцією втрат, що підсумовує внесок кожного класу, з можливістю вагування рідкісних розладів для компенсації дисбалансу даних. Така постановка природно відображає коморбідність: один і той самий текст може одночасно демонструвати, наприклад, ПТСП та депресивні прояви, що принципово недоступно в класичних мультикласових моделях.

Третя компонента пов'язана з формалізацією ризику соціально небезпечної поведінки на основі розширеної множини психолінгвістичних та поведінкових ознак. Для кожного повідомлення x , окрім вектора $h(x)$ та оцінок $s_PTSD(x)$, $s_1(x)$, \dots , $s_K(x)$, обчислюється вектор «традиційних» ознак $\varphi(x) = (\varphi_1(x), \dots, \varphi_p(x))$. До нього входять інтегральна тональність тексту, частка негативних і позитивних речень, амплітуда коливань полярності між сусідніми реченнями, кількість емодзі, питома вага агресивної та ворожої лексики, оцінки базових емоцій (страх, гнів, смуток, відчай) за емоційними лексиконами, а також ознаки, пов'язані з граматичними спотвореннями, різкими обривами фраз, фіксацією нав'язливих тем. Такі ознаки побудовано так, щоб вони підсилювали нейромережеві представлення, надаючи більш прозорі, інтерпретовані виміри поведінки користувача в тексті.

Інтегральний ризик соціально небезпечної поведінки для одного повідомлення задається деякою функцією $r_risk(x) = g(s_PTSD(x), s_1(x), \dots, s_K(x), \varphi(x))$, де $g(\cdot)$ реалізується у вигляді логістичної регресії або компактної нейромережі з невеликою кількістю параметрів. На виході $r_risk(x)$ інтерпретується як числова оцінка у діапазоні від 0 до 1 (або у відсотках), що показує, наскільки поєднання ознак ПТСП, коморбідних розладів та емоційно-поведінкових патернів у даному

тексті наближається до типових профілів підвищеного ризику. Важливо, що така функція дозволяє явно моделювати внесок кожної компоненти: збільшення $s_PTSD(x)$ при стабільно низьких значеннях агресивної лексики може не давати високого $r_risk(x)$, тоді як поєднання помірного ПТСР, суїцидального ризику та різко негативних емоційних показників $\varphi(x)$ спричинює значне зростання інтегрального індексу.

Для переходу від аналізу одиничних повідомлень до оцінювання психічного стану конкретної особи використовується агрегування показників за часовим вікном T . Нехай для користувача u розглядається множина повідомлень за проміжок $[t_0; t_1]$. Тоді можна визначити середню інтенсивність проявів ПТСР як $R_PTSD(u, T) =$ середнє значення $s_PTSD(x)$ по всіх повідомленнях за обраний період. Аналогічно, для кожного розладу c_k обчислюється $R_k(u, T)$ як середнє по $s_k(x)$, а інтегральний індекс ризику $R_risk(u, T)$ як середнє по $r_risk(x)$.

З математичної точки зору, повний конвеєр можна описати як відображення від множини текстів до множини показників стану. Для кожного користувача u його потік повідомлень $M(u)$ спершу перетворюється нейромережею на послідовність векторів $h(x)$, потім через модулі класифікації ПТСР і мультилейблової діагностики на оцінки $s_PTSD(x)$ та $s_k(x)$, далі через функцію $g(\cdot)$ на значення $r_risk(x)$, після чого усі ці величини агрегуються у вектор $\Psi(u, T)$. Таким чином реалізується двоетапне відображення $M(u) \rightarrow \{s_PTSD(x), s(x), r_risk(x)\} \rightarrow \Psi(u, T)$, де перший етап відповідає глибинному семантико-психолінгвістичному аналізу тексту, а другий – побудові інтерпретованих індикаторів стану.

Практична цінність запропонованої моделі полягає в тому, що вона поєднує сильні сторони сучасних трансформерних архітектур, здатних працювати з неструктурованими текстами реальних соціальних мереж, із прозорими числовими індексами, які можна використовувати в автоматизованих системах моніторингу психічного здоров'я, телемедицині сервісах та превентивних механізмах у сфері громадської безпеки. Застосування мультилейблової схеми замість жорсткого мультикласового поділу дозволяє коректно відобразити коморбідність та перехресні симптоми, характерні для ПТСР, тоді як облік розширеної множини психолінгвістичних ознак зменшує залежність моделі від зміщень у даних та дозволяє частково пояснювати, які саме патерни поведінки спричинили підвищення індексу ризику. Це створює передумови для інтеграції моделі в етичні та юридично обґрунтовані сценарії підтримки фахівців із психічного здоров'я без заміщення їхніх професійних рішень, але з істотним посиленням можливостей раннього виявлення кризових станів у цифрових середовищах.

Перелік посилань

1. Inamdar, S., Chapekar, R., Gite, S. et al. Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing. *Hum-Cent Intell Syst* 3, 80–91 (2023). <https://doi.org/10.1007/s44230-023-00020-8>
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. *Sci Rep* 14, 12468 (2024). <https://doi.org/10.1038/s41598-024-61557-7>
3. Papini S, Norman SB, Campbell-Sills L, et al. Development and Validation of a Machine Learning Prediction Model of Posttraumatic Stress Disorder After Military Deployment. *JAMA Netw Open*. 2023;6(6):e2321273. doi:10.1001/jamanetworkopen.2023.21273
4. L. F. Ramos-Lima et al., 'Identifying posttraumatic stress disorder staging from clinical and sociodemographic features: a proof-of-concept study using a machine learning approach', *Psychiatry Research*, vol. 311, p. 114489, 2022. <https://doi.org/10.1016/j.psychres.2022.114489>.
5. Овчарук О.М., Мазурець О.В. Нейромережевий метод діагностування психологічних розладів за аналізом повідомлень на основі роздільного підходу до класифікації. Міжнародний науково-технічний журнал «Вимірвальна та обчислювальна техніка в технологічних процесах». № 1, 2025. с. 210-216.
6. Мазурець О.В., Овчарук О.М. Дослідження ефективності методу аналізу прояву посттравматичного стресового розладу в користувацькому контенті. Інформаційні технології і автоматизація. Матеріали XVII міжнародної науково-практичної конференції. 31 жовтня – 1 листопада 2024 р. Одеса, ОНТУ. 2024. С.653-655.
7. Овчарук О.М., Мазурець О.В. Підхід до виявлення ознак психічних розладів людини за аналізом користувацьких дописів ансамблем нейромереж-трансформерів. Збірник наукових праць за матеріалами XVI Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2024». 15-16 листопада 2024. Хмельницький, 2024. с. 389-393.
8. Mazurets O., Ovcharuk O. Efficiency Research of Method for Detecting Mental Disorders by Analysis of User Content. *Information Technology and Implementation (Satellite). Proceedings 11th International Conference*. November 21, 2024. Kyiv, Ukraine. 2024. Pp. 46-47.
9. Овчарук О.М., Мазурець О.В. Метод нейромережевого виявлення ознак психічних розладів людини за аналізом її дописів. Тези доповідей XII Міжнародної науково-практичної конференції «Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій». 10-12 грудня 2024. Запоріжжя, Національний університет «Запорізька політехніка». 2024. с. 380-383.
10. Мазурець О.В., Овчарук О.М. Діагностування проявів посттравматичного стресового розладу за нейромережевим аналізом текстового контенту. Збірник наукових праць XXIII Міжнародної наукової конференції «Нейромережні технології та їх застосування НМТІЗ-2024». 11-12 грудня 2024. Краматорськ-Тернопіль, ДДМА. 2024. с. 85-89.