

УДК 004.8

Мазурець О.В., Віт Р.В.

Хмельницький національний університет

МОДЕЛЬ ВИЯВЛЕННЯ ЦИФРОВОЇ ВТОМИ У ТЕКСТОВОМУ КОНТЕНТІ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

У роботі розглянуто проблему цифрової втоми, що виникає внаслідок тривалого перебування користувачів у цифрових середовищах, інформаційного перевантаження та емоційного виснаження. Метою є розроблення нейромережевої моделі виявлення цифрової втоми у текстовому контенті користувачів, здатної автоматично оцінювати рівень втоми для окремих повідомлень і формувати інтегральні показники для користувача за певний період. Запропоновано конвеєр, у якому контекстно-орієнтована трансформерна модель типу BERT формує векторні подання текстів, на основі яких навчається бінарний класифікатор «втома / відсутність втоми». Додатково використано тематичне моделювання та виділення цільових об'єктів предметної області, що дає змогу деталізувати цифрову втому за темами й ключовими поняттями та будувати інтерпретований профіль користувача.

The paper addresses the problem of digital fatigue that emerges from prolonged exposure to online environments, information overload and emotional exhaustion. The aim is to develop a neural network-based model for detecting digital fatigue in user text content, capable of automatically assessing fatigue levels for individual messages and deriving integral indicators for a user over a given time period. The proposed pipeline employs a context-aware transformer model of the BERT family to obtain vector representations of texts, on top of which a binary “fatigue / no fatigue” classifier is trained. Additionally, topic modelling and extraction of domain-specific target objects are used to break down digital fatigue by themes and key concepts, forming an interpretable user profile.

Інтенсивне зростання часу, який користувачі проводять у цифрових середовищах, поєднання дистанційної роботи, онлайн-навчання, новинних стрічок і постійної присутності в соціальних мережах призводять до накопичення явищ цифрової втоми [1, 2]. Вона проявляється у перевантаженні інформацією, зниженні концентрації, емоційному виснаженні та формальному, «механічному» стилі взаємодії з цифровим контентом [3]. Традиційні підходи до діагностики ґрунтуються на анкетуванні та самооцінці, потребують активної участі користувача й не масштабуються до великих потоків текстових даних [4]. В умовах зростання інформаційного тиску та потреби в ранньому виявленні негативних тенденцій стає актуальним перехід до автоматизованого аналізу текстового контенту із застосуванням сучасних моделей штучного інтелекту [5].

Мета роботи полягає в розробленні нейромережевої моделі виявлення цифрової втоми у текстовому контенті користувачів, яка забезпечує автоматичну оцінку рівня цифрової втоми для окремих текстів і інтегральних показників для

користувача за визначений період, а також дозволяє деталізувати структуру цієї втоми за тематичними сегментами та ключовими об'єктами предметної області.

Запропонована модель будується як конвеєр, що поєднує глибокий нейромережеві подання тексту [6], тематичне моделювання [7] й аналіз цільових об'єктів [8, 9]. Вхідними даними є тексти повідомлень користувача x (пости, коментарі, чат-репліки), виходом – оцінка рівня цифрової втоми $d(x)$ у діапазоні від 0 до 1 та її розклад за основними темами. Обчислювальне ядро становить контекстно-орієнтована трансформерна модель типу BERT [10], яка перетворює кожен текст x на векторну репрезентацію $h(x)$, чутливу до семантики, емоційного забарвлення та когнітивної складності мовлення. Вектор $h(x)$ інтерпретується як компактний числовий опис стану користувача, відображаючи лексичні, синтаксичні й стилістичні особливості тексту.

На основі вектора $h(x)$ будується бінарний детектор цифрової втоми. Для кожного тексту x модель повертає ймовірнісну оцінку $d(x) = P(\langle \text{цифрова втома} \mid x \rangle)$ у межах $[0;1]$. Якщо $d(x) \geq \tau$, де τ – обраний поріг чутливості, текст інтерпретується як такий, що містить ознаки цифрової втоми; якщо $d(x) < \tau$, ознаки втоми вважаються відсутніми або недостатньо вираженими. Навчання моделі здійснюється на розміченому корпусі текстів з експертними мітками «втома / немає втоми» із використанням логістичної (крос-ентропійної) функції втрат, що мінімізує розбіжність між $d(x)$ та еталонною міткою.

Для семантичної деталізації стану цифрової втоми текст x додатково подається до модуля тематичного моделювання, який розкладає його на суміш тем $T = \{t_1, \dots, t_m\}$ з вагами $w_j(x)$ у діапазоні $[0;1]$, де сума всіх $w_j(x)$ дорівнює 1. Паралельно формується множина цільових об'єктів $O(x)$ – характерних слів, стійких словосполучень та іменованих сутностей, виділених за поєднанням TF-IDF, виявлення іменованих сутностей та статистичного аналізу. Таким чином, кожен текст описується трійкою $(h(x), T, O(x))$, що дає змогу пов'язати факт цифрової втоми не лише з загальним стилем, а й з конкретними темами та поняттями, у межах яких виникає перевантаження.

Індекс цифрової втоми деталізується за темами у вигляді набору $d_j(x) = d(x) w_j(x)$, $j = 1 \dots m$, де $d_j(x)$ показує, яку частку загальної втоми пов'язано з кожною темою t_j . Це дозволяє побудувати тематичний профіль цифрової втоми для одного тексту й візуалізувати «зони напруження» у вигляді карти, на якій для окремих речень або фрагментів відображаються локальні значення $d_j(x)$ та пов'язані з ними цільові об'єкти з $O(x)$. У такий спосіб модель забезпечує не лише числову оцінку $d(x)$, а й інтерпретацію того, «де» і «внаслідок чого» формується стан перевантаження.

Для оцінювання стану конкретного користувача u у динаміці розглядається множина його повідомлень $M(u, T)$ за період T . Інтегральний рівень цифрової втоми визначається як $D(u, T) = \text{середнє значення } d(x) \text{ для всіх } x \in M(u, T)$. Аналогічно, для кожної теми t_j обчислюється тематичний індекс $D_j(u, T) = \text{середнє значення } d_j(x) \text{ за період } T$. Сукупність показників $\Psi(u, T) = \{D(u, T), D_1(u, T), \dots, D_m(u, T)\}$ описує як загальну інтенсивність цифрової втоми, так і її розподіл за темами, формуючи придатний до інтерпретації профіль користувача.

З математичної точки зору, повний конвеєр можна описати як відображення від множини текстів до множини показників стану: потік повідомлень $M(u, T)$ спершу перетворюється нейромережею на послідовність векторів $h(x)$, далі через модуль класифікації – на значення $d(x)$, потім через тематичний модуль і аналіз цільових об'єктів – на $d_j(x)$ і набори $O(x)$, після чого всі ці величини агрегуються у вектор $\Psi(u, T)$. Така модель поєднає гнучкість сучасних трансформерних архітектур із прозорими цифровими індикаторами, придатними для інтеграції в системи моніторингу ментального стану, адаптивні інтерфейси та сервіси підтримки користувачів, не замінюючи професійної оцінки фахівців, але підсилюючи можливості раннього виявлення цифрової втоми.

Перелік посилань

1. Watkins J. Alleviating digital fatigue through embodied artistic practice and green space. *International Journal of Performance Arts and Digital Media*. 2024. P. 1–16.
1. Fatigue-Related Changes of Daily Function: Most Promising Measures for the Digital Age / W. Maetzler et al. *Digital Biomarkers*. 2024. P. 30–39.
2. Byrne M. Special Topic on Reducing Technology Related Stress and Burnout: Digital Compassion Fatigue as an Emerging Phenomenon for Registered Nurses Experiencing Technostress. *Applied Clinical Informatics*. 2025. URL: <https://doi.org/10.1055/a-2564-8809>
3. Expressions of pandemic fatigue on digital platforms: a thematic analysis of sentiment and narratives for infodemic insights / B. K. White et al. *BMC Public Health*. 2024. Vol. 24, no. 1. URL: <https://doi.org/10.1186/s12889-024-17718-4>
4. Віт Р.В., Мазурець О.В. Підхід до тематичної класифікації текстової інформації засобами обробки природної мови. *Науковий журнал «Наукові праці Донецького національного технічного університету», серія «Проблеми моделювання та автоматизації проектування»*. 2025. №1 (21). С. 94-99.
5. Мазурець О.В., Віт Р.В. Дослідження ефективності методу виявлення цільових об'єктів предметної області. *Інформаційні технології і автоматизація. Матеріали XVII міжнародної науково-практичної конференції*. 31 жовтня – 1 листопада 2024 р. Одеса, ОНТУ. 2024. С.650-653.
6. Віт Р.В., Мазурець О.В. Тематична класифікація текстової інформації засобами обробки природної мови. *Збірник наукових праць XXIII Міжнародної наукової конференції «Нейромережні технології та їх застосування НМТІЗ-2024»*. 11-12 грудня 2024. Краматорськ-Тернопіль, ДДМА. 2024. с. 63-66.
7. Мазурець О.В., Віт Р.В. Метод виявлення цільових об'єктів предметної області у текстовому контенті. *Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький*, 2024. №6, Т.1 (343). С. 152-157.
8. Віт Р.В., Мазурець О.В. Метод формування множин цільових об'єктів предметних областей у цифрових текстах засобами машинного навчання. *Науковий журнал «Наука і техніка сьогодні»*. Київ, 2024. №13 (41). С. 926-937.
9. Віт Р.В., Мазурець О.В. Метод виявлення психологічного цифрового перевантаження за аналізом текстових даних нейромережевими моделями глибокого навчання. *Науковий журнал «Вісник Херсонського національного технічного університету»*. 2025. №2 (93). Т. 2. С. 107-114.