

[DOI 10.28925/2663-4023.2026.35.1402](https://doi.org/10.28925/2663-4023.2026.35.1402)

УДК 004.93'1:004.89

**Коротун Ольга Володимирівна**

кандидат педагогічних наук, доцент

доцент кафедри комп'ютерних наук

Державний університет «Житомирська політехніка», м. Житомир, Україна

ORCID: 0000-0003-2240-7891

korotun-o@zstu.edu.ua

**Марчук Галина Вікторівна**

старший викладач кафедри комп'ютерних наук

Державний університет «Житомирська політехніка», м. Житомир, Україна

ORCID: 0000-0003-2954-1057

pzs\_mgv@zstu.edu.ua

**Лисюк Аліса Олегівна**

здобувачка

Державний університет «Житомирська політехніка», м. Житомир, Україна

ORCID: 0009-0007-1990-5458

ist231\_lao@student.zstu.edu.ua

**МОДЕЛЮВАННЯ ТА ОЦІНЮВАННЯ РІВНЯ ПСИХОЕМОЦІЙНОГО НАВАНТАЖЕННЯ  
КОРИСТУВАЧІВ У ЦИФРОВОМУ СЕРЕДОВИЩІ**

**Анотація.** У статті розв'язано актуальне науково-прикладне завдання щодо підвищення об'єктивності та точності оцінювання психоемоційного стану користувачів у цифровому середовищі на основі застосування алгоритмів класифікації та кластеризації. У вступній частині обґрунтовано суперечність між зростанням рівня цифрового дистресу користувачів і обмеженістю традиційних діагностичних методів, що зумовлює необхідність розробки математичних моделей оцінювання на основі цифрових слідів. На етапі підготовки даних проведено очищення вибірки з 5005 спостережень із 25 параметрами цифрової поведінки, фізіології та ментального здоров'я, а також сформовано пайплайн препроцесингу з one-hot кодуванням категоріальних змінних та z-стандартизацією числових ознак. Запропоновано та математично формалізовано авторський інтегральний індекс емоційного навантаження (ІЕН), який балансує вагові коефіцієнти факторів ризику та компенсаторних ресурсів організму пропорційно їхнім кореляційним зв'язкам із рівнем стресу. За допомогою побудованої моделі множинної лінійної регресії ( $R^2 = 0,9389$ ) доведено, що основним тригером дистресу є тривалість використання соціальних мереж, а не загальний екранний час поза ними. На основі квантильного поділу рівня стресу на три класи (Low, Mid, High) проведено порівняльний аналіз моделей класифікації Naive Bayes, Decision Tree та SVM із RBF-ядром. За результатами 5-fold крос-валідації доведено, що модель SVM RBF має найвищу узагальнювальну здатність та стійкість до перенавчання (Accuracy = 0,9856). Застосування кластеризації методом k-means ( $k=2$ , Silhouette score = 0,42) у поєднанні з візуалізацією в просторі головних компонент дозволило ідентифікувати два поведінкові профілі: «Ресурсний» та «Цифрового ризику». Практична цінність результатів полягає у створенні підґрунтя для розробки програмних модулів автоматизованого моніторингу ментального здоров'я користувачів у реальному часі.

**Ключові слова:** психоемоційне навантаження; цифрове середовище; індекс емоційного навантаження; машинне навчання; класифікація; кластеризація; екранний час; цифрова поведінка

**ВСТУП**

**Постановка проблеми.** Стрімка цифровізація суспільства, масштабування дистанційних форм праці та навчання, а також інтенсивний розвиток онлайн-платформ сформували нове середовище функціонування сучасної людини. Поряд із беззаперечними перевагами (доступність інформації,



оперативність комунікації, автоматизація процесів), активна взаємодія з цифровими системами супроводжується значним інформаційним та психоемоційним тиском.

Явища екранної залежності, інформаційного перевантаження та постійної цифрової присутності зумовлюють зростання рівня дистресу, виснаження й зниження адаптаційних ресурсів користувачів. У контексті надійності функціонування інформаційно-комунікаційних систем цей чинник є критичним: втомлена людина або, яка знаходиться в стані стресу, частіше припускається помилок під час прийняття рішень, втрачає концентрацію уваги та стає вразливою до загроз соціальної інженерії.

Попри наявність численних концепцій дослідження стресу в медицині та психології, існує виражена науково-практична суперечність. Традиційні інструменти оцінювання стану людини спираються на суб'єктивні опитувальники або громіздкі фізіологічні вимірювання, що унеможлиблює їх оперативне застосування в реальному часі. Водночас постійне накопичення цифрового сліду створює підґрунтя для використання методів інтелектуального аналізу даних із метою безперервної моніторингової класифікації та оцінювання станів користувача.

Таким чином, актуальною науково-прикладною задачею є розробка формалізованих математичних моделей, інтегральних метричних показників і порівняльний аналіз алгоритмів інтелектуальної обробки даних, здатних прогнозувати рівні психоемоційного навантаження на основі поведінкових параметрів у цифровому просторі.

**Аналіз останніх досліджень і публікацій.** Проаналізуємо наукові праці вітчизняних та закордонних авторів, що присвячені вивченню психоемоційного стану, рівня стресу людей у цифровому середовищі. А також застосування алгоритмів машинного навчання для діагностики їх стресових станів.

Безпосередню оцінку факторів стресу та їхнього руйнівного впливу на організм молоді висвітлено у низці праць. В. В. Жекова [1] на основі соціологічного опитування доводить наявність прямого зв'язку між академічним навантаженням і виникненням психосоматичних розладів, тривожності та дефіциту сну. Це навантаження суттєво підсилюється воєнним контекстом в Україні, що детально розкрито у дослідженні С. Є. Білоусової [2]. Авторка показує, як хронічна тривога, систематичні повітряні тривоги та безпекові загрози створюють кумулятивний ефект, спричиняючи швидке емоційне виснаження та підвищуючи ризики дезадаптації.

Важливе значення мають публікації, присвячені аналізу поведінкових факторів та способу життя як ключових детермінантів стресостійкості. Дослідження Д. Батрони [3] виявляє високу поширеність гіподинамії, незбалансованого харчування та системного недосипання, що знижує фізичні й психічні резерви організму. Шляхи подолання цієї проблеми пропонує О. С. Бородіна [4], обґрунтовуючи стратегії систематичної рухової активності як потужного природного буфера, який знижує рівень кортизолу та стабілізує емоційний стан.

У працях інших дослідників розглядається специфіка впливу тривалого використання гаджетів і соціальних мереж на ментальний стан. У [5] зазначено, що помірне (до 2 годин на день) використання пристроїв не викликає депресії чи тривожності, що доводить недоцільність оцінювання лише загальної тривалості перебування в мережі. Водночас масштабне дослідження ABCD Study [6] уточнює, що тривалий екранний час статистично пов'язаний із депресивними симптомами, розладами поведінки та гіперактивністю (особливо під час взаємодії з відеочатами, повідомленнями та іграми). Це підкреслює критичну важливість аналізу саме якісних патернів і типів цифрової активності.

Низка авторів аналізує психологічні причини залежності користувачів від смартфонів та шляхи її подолання. Дослідження Дж. Д. Елхая [7] доводить, що залежність від смартфонів виникає через порушення емоційної регуляції, коли низька толерантність до дистресу та низький рівень усвідомленості змушують людину використовувати гаджет для «втечі» від тривоги. Робота Р. А. Бонда [8] у межах моделі I-PACE уточнює, що ключовим медіатором між соціальною тривожністю та проблемним використанням смартфона є експресивне пригнічення емоцій. Щодо подолання цих наслідків, у [9] визначено, що «цифровий детокс» (пауза тривалістю від одного дня до кількох тижнів) суттєво зменшує симптоми депресії, тривожності та психологічного дистресу. Втім, для досягнення довготривалого ефекту його слід поєднувати із системним формуванням здорових цифрових звичок. У роботі І. Ломачинської, І. Волинець та С. Хрипко [10] аналізуються соціопсихологічні аспекти «цифрового здоров'я» та виклики дистанційного навчання, де «цифрова втома» й інформаційне перевантаження стають окремими джерелами стресу. З іншого боку, Р. К. Морока [11] акцентує на позитивному потенціалі інновацій: грамотне застосування інструментів штучного інтелекту в навчанні допомагає оптимізувати тайм-менеджмент, зменшує рутинне навантаження та знижує рівень тривожності, пов'язаної з дефіцитом часу.

Окремі дослідження присвячені використанню цифрових технологій та алгоритмів машинного навчання для діагностики й моніторингу стресового стану користувачів. Дослідження М. О. Верчинської та О. О. Войцеховської [12] спрямоване на застосування розвідувального аналізу даних для створення інформаційних технологій моніторингу стресу. Публікації щодо класифікації стресу за допомогою



алгоритму опорних векторів і аналізу електрокардіограм [13], а також праці з класифікації психоемоційних станів за допомогою моделей машинного навчання [14] доводять, що алгоритми (SVM, Random Forest, Decision Trees, KNN) за фізіологічними показниками та цифровими слідами досягають 85-95% точності класифікації. Дослідження [15] підтвердило, що під час аналізу 28 анкетних параметрів життєдіяльності саме класифікатор SVM демонструє високу точність (до 95%) для прогнозування та профілактики стресових станів.

Узагальнюючи розглянуті праці, можна зробити висновок, що сучасне психоемоційне навантаження користувачів у цифровому середовищі має комплексну природу. Ефективне моделювання та оцінювання цього стану вимагає комплексної системи заходів: від дотримання норм здорового способу життя й цифрової гігієни до впровадження точних моделей машинного навчання для автоматизованого та раннього виявлення рівня стресу.

**Мета та завдання дослідження.** Метою статті є підвищення об'єктивності та точності оцінювання психоемоційного стану користувачів у цифровому середовищі шляхом розробки інтегрального індексу емоційного навантаження та порівняльного аналізу алгоритмів машинного навчання для його прогнозування.

Для досягнення поставленої мети сформульовано такі завдання: проаналізувати та підготувати датасет поведінкових та психофізіологічних показників користувачів із застосуванням методів очищення, стандартизації та кодування категоріальних змінних; запропонувати та математично формалізувати інтегральний індекс емоційного навантаження (ІЕН), що враховує баланс між чинниками дистресу та компенсаторними ресурсами організму; побудувати та порівняти моделі класифікації рівнів навантаження на основі алгоритмів Naive Bayes, Decision Tree та SVM із застосуванням k-fold крос-валідації та аналізом матриць помилок; провести кластеризацію даних методом k-means для виділення типових поведінкових профілів користувачів та оцінити їх стійкість.

## МЕТОДИКА ДОСЛІДЖЕННЯ

Методика дослідження базується на інтелектуальному аналізі даних вибірки з 5005 спостережень із 25 показниками цифрової поведінки, фізіології та психоемоційного стану. На етапі попередньої обробки проведено очищення даних, one-hot кодування категоріальних змінних та z-стандартизацію числових ознак. Для комплексної оцінки сформовано авторський ІЕН, що балансує фактори ризику (соцмережі, тривожність) та компенсаторні ресурси (сон, ментальне здоров'я, фізичну активність). Силу впливу факторів оцінено множиною лінійною регресією, для класифікації користувачів за трьома рівнями стресу застосовано алгоритми Naive Bayes, Decision Tree та SVM RBF із перевіркою стійкості через 5-fold крос-валідацію, а приховані поведінкові профілі ідентифіковано методом кластеризації k-means із візуалізацією у просторі головних компонент.

## РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для проведення аналізу чинників формування дистресу в цифровому середовищі використано відкритий публічний датасет Tech Use & Stress Wellness з платформи Kaggle. Початковий масив даних містив 5013 спостережень та 25 змінних, які охоплюють поведінкові, демографічні, фізіологічні та психологічні показники.

Усі атрибути набору даних за своїм аналітичним призначенням розподілено на чотири ключові групи.

1. Показники цифрової поведінки – загальний щоденний екранний час (daily\_screen\_time\_hours), тривалість активності у соціальних мережах (social\_media\_hours), час роботи за пристроями (телефон, ноутбук, планшет, TV), а також час, витрачений на ігри, розваги та робочі/навчальні завдання.

2. Показники психоемоційного стану – підсумковий рівень стресу (stress\_level, шкала 1-10), оцінка настрою (mood\_rating), показники ментального здоров'я (mental\_health\_score), а також тижневі показники тривожності (weekly\_anxiety\_score) та депресивних симптомів (weekly\_depression\_score).

3. Фізіологічні та поведінкові чинники – тривалість (sleep\_duration\_hours) та суб'єктивна якість сну (sleep\_quality), тижневий обсяг фізичної активності (physical\_activity\_hours\_per\_week), щоденне споживання кофеїну, практика майндфулнес, використання wellness-застосунків та дотримання здорового харчування;

4. Соціально-демографічні характеристики – вік, стать, тип місцевості проживання.



На етапі первинного аналізу проведено перевірку цілісності та структури даних. Початковий набір даних містив 5013 спостережень. У результаті перевірки виявлено 18 рядків, що відповідали 10 унікальним записам, які дублювалися. За допомогою процедури вилучення повторів було збережено по одному оригінальному запису з кожної групи та видалено 8 дубльованих копій. Також проведено аналіз вибірки на наявність пропущених значень (NaN/Null), за результатами якого неструктуровані або відсутні дані виявлені не були. Таким чином, фінальний обсяг очищеної вибірки склав 5005 спостережень без втрати її аналітичної репрезентативності.

Категоріальні змінні (gender, location\_type) було приведено до стандартного рядкового формату для подальшої коректної кодової обробки.

Обчислення мір центральної тенденції та розсіювання дозволило встановити базові характеристики вибірки:

Середнє значення рівня стресу становить  $5,72 \pm 2,92$  (медіана – 6,0), що вказує на переважання у вибірці респондентів із помірним та підвищеним рівнем психоемоційного напруження.

Вимір цифрового навантаження показав, що середній щоденний екранний час становить  $5,04 \pm 1,84$  години, а середній час у соціальних мережах –  $3,28 \pm 1,20$  години.

Компенсаторні ресурси мають наступні значення: середня тривалість сну становить  $7,37 \pm 0,54$  години з малим інтерквартильним розмахом (0,7), що свідчить про високу щільність показників у межах норми (7-8 годин). Середня якість сну становить 4,01 (за шкалою 1-5).

Аналіз показників форми розподілу (асиметрії та ексцесу) засвідчив, що більшість числових змінних мають коефіцієнти асиметрії у межах  $\pm 0,7$ , що свідчить про наближеність їхніх розподілів до нормального. Виражену правосторонню асиметрію виявлено для змінних фізичної активності (physical\_activity\_hours\_per\_week, асиметрія 0,712) та майндфулнес-практик (0,598), що відображає нерівномірність розподілу здорових поведінкових стратегій серед респондентів.

Для підготовки даних до залучення в алгоритмах машинного навчання застосовано уніфікований пайплайн препроцесингу.

Для категоріальних змінних застосовано One-Hot кодування із бінарним перетворенням унікальних значень.

Усі числові ознаки було стандартизовано за допомогою z-оцінок. Стандартизація звела всі числові показники до єдиного масштабу із середнім значенням  $\mu=0$  та дисперсією  $\sigma^2=1$ , що усунуло проблему домінування великомасштабних змінних (наприклад, споживання кави в мг чи віку) над дрібномасштабними показниками в процесі навчання регресійних та класифікаційних моделей.

З метою агрегації різномірних поведінкових і психоемоційних параметрів у єдину нормалізовану метрику розроблено індекс емоційного навантаження. В основу індексу закладено ідею розрахунку балансу між ваговими коефіцієнтами факторів ризику та захисних ресурсів користувача. До складу індексу відібрано 5 ключових змінних:

1. S – час у соціальних мережах (social\_media\_hours, фактор ризику);
2. A – рівень тривожності (weekly\_anxiety\_score, фактор ризику);
3. Q – якість сну (sleep\_quality, захисний ресурс);
4. M – оцінка ментального здоров'я (mental\_health\_score, захисний ресурс);
5. P – фізична активність (physical\_activity\_hours\_per\_week, захисний ресурс).

Для усунення різниці у масштабах вимірювання всі компоненти попередньо стандартизовано у вигляді z-оцінок.

$$S_z = \frac{S - \mu_S}{\sigma_S}, \quad A_z = \frac{A - \mu_A}{\sigma_A}, \quad Q_z = \frac{Q - \mu_Q}{\sigma_Q}, \quad M_z = \frac{M - \mu_M}{\sigma_M}, \quad P_z = \frac{P - \mu_P}{\sigma_P} \quad (1)$$

Вагові коефіцієнти  $w_j$  для кожного фактора розраховано пропорційно абсолютним значенням їхніх коефіцієнтів кореляції Пірсона  $|\rho_j|$  із рівнем стресу за умови нормування їх суми до одиниці:

$$\bar{w}_j = |\rho_j|, \quad w_j = \frac{\bar{w}_j}{\sum_{k=1}^5 \bar{w}_k} \quad (2)$$

Формула «сирого» індексу ( $IEH_{raw}$ ) має вигляд:

$$IEH_{raw} = w_S * S_z + w_A * A_z - w_Q * Q_z - w_M * M_z - w_P * P_z \quad (3)$$

Для зручності практичної інтерпретації та приведення показника до шкали [0; 100] застосовано Мін-Макс нормування:

$$IEH = 100 * \frac{IEH_{raw} - \min(IEH_{raw})}{\max(IEH_{raw}) - \min(IEH_{raw})} \quad (4)$$

Аналіз розрахованого індексу показав сильний позитивний лінійний зв'язок із реальним рівнем стресу ( $r \approx 0.89$ ). Значення  $IEH > 70$  ідентифікують критичний стан цифрового виснаження та високого



ризик дистресу, тоді як значення  $IEH < 40$  характерні для респондентів із високим рівнем компенсаторних ресурсів та ефективним цифровим саморегулюванням.

Для оцінювання сумарного та «чистого» внеску факторів у формування дистресу побудовано модель множинної лінійної регресії МНК:

$$\hat{y} = \beta_0 + \beta_1 * S + \beta_2 * D_{scr} + \beta_3 * Q + \beta_4 * M \quad (5)$$

де  $\hat{y}$  – прогнозований рівень стресу (stress\_level),  $S$  – час у соцмережах,  $D_{scr}$  – загальний щоденний екранний час,  $Q$  – якість сну,  $M$  – ментальне здоров'я.

Отримані параметри моделі:

- Вільний член  $\beta_0 = 3,9172$ ;
- $\beta_1 = + 1,5850$  ( $p < 0,001$ );
- $\beta_2 = + 0,1172$  ( $p < 0,001$ );
- $\beta_3 = - 0,0970$  ( $p < 0,05$ );
- $\beta_4 = - 0,0675$  ( $p < 0,001$ ).

Модель продемонструвала високу пояснювальну здатність: коефіцієнт детермінації  $R^2 = 0,9389$  (скоригований  $R^2_{adj} = 0,9388$ ).

Аналіз часткових регресійних діаграм підтвердив, що час у соціальних мережах зберігає виражений негативний вплив на психоемоційний стан навіть після контролю рівня ментального здоров'я та якості сну. Водночас загальний екранний час поза соціальними мережами має значно слабший прямий вплив, що вказує на вирішальну роль саме контенту та характеру цифрової взаємодії, а не просто тривалості роботи з пристроями.

Для побудови моделей розпізнавання станів психоемоційного навантаження здійснено категоризацію неперервної змінної stress\_level ( $y \in [1; 10]$ ) за квантильним поділом на три збалансовані класи:

- Low ( $y \leq q_{0,33}$ ) – низький рівень стресу (стан адаптації);
- Mid ( $q_{0,33} < y \leq q_{0,66}$ ) – помірний рівень стресу (компенсоване навантаження);
- High ( $y > q_{0,66}$ ) – високий рівень стресу (стан дистресу, група ризику).

Розподіл класифікованих спостережень виявився збалансованим (Low: ~38%, Mid: ~29%, High: ~33%), що усуває загрозу упередженості моделей та дозволяє застосовувати класичні метрики якості класифікації без штучного зважування вибірки.

Для експериментальної перевірки вибірку розділено у співвідношенні 70/30 (стратифіковане розбиття  $X_{train}$ ,  $X_{test}$  із фіксацією random\_state = 23). Навчання здійснювалося за уніфікованим препроцесинговим пайплайном: стандартизація числових векторів та one-hot кодування категоріальних ознак.

Досліджено три алгоритми машинного навчання з різною математичною природою поділу простору ознак.

1. Наївний Байєсівський класифікатор – імовірнісна модель, яка виходить із припущення незалежності ознак.
2. Дерево рішень – непараметричний метод, що будує ієрархічні правила поділу на основі зменшення ентропії (параметри: max\_depth = 8, min\_samples\_leaf = 20).
3. Метод опорних векторів з RBF-ядром – алгоритм максимізації розділювальної смуги в багатовимірному просторі для виявлення нелінійних залежностей.

Отримані метрики ефективності моделей на тестовій вибірці ( $n = 1502$ ) наведено в Таблиці 1.

Таблиця 1

Порівняльний аналіз точності моделей класифікації рівнів стресу

| Модель класифікації | Accuracy | Precision (macro) | Recall (macro) | F1-score (macro) |
|---------------------|----------|-------------------|----------------|------------------|
| Naive Bayes         | 0,9687   | 0,9667            | 0,9700         | 0,9683           |
| Decision Tree       | 1,0000   | 1,0000            | 1,0000         | 1,0000           |
| SVM (RBF)           | 0,9893   | 0,9883            | 0,9897         | 0,9890           |

Аналіз матриць помилок дав змогу деталізувати характер роботи алгоритмів.

Naive Bayes показав загальну точність 96,87%. Помилки класифікації виникали виключно між суміжними класами (наприклад, 22 випадки змішування Low та Mid), що зумовлено кореляційною залежністю між окремими ознаками, яка порушує базове припущення алгоритму про їхню незалежність.

Decision Tree демонструє формальну точність 100% на контрольному зрізі. Проте в умовах реальних аналітичних систем абсолютна точність є чітким маркером перенавчання, оскільки дерево надто детально адаптувалося до локальних особливостей вибірки й втрачає робастність при надходженні зашумлених даних.

SVM RBF продемонстрував точність 98,93% ( $F_1 = 0,9890$ ). Усі помилки (поодинокі 11 випадків на вибірці з 1502 об'єктів) зафіксовано лише на межі суміжних рівнів `stress_level`, що свідчить про високу гладкість розділювальної гіперповерхні.

Для остаточної валідації та перевірки узагальнювальної здатності проведено 5-fold стратифіковану крос-валідацію та серію стрес-тестів зі зміною розміру тестової вибірки (`test_size` ∈ [0,05; 0,95]). Результати крос-валідації: Naive Bayes показав середнє значення `Accuracy` = 0,9646 (дисперсія за фолдами [0,9540; 0,9740]); Decision Tree показав середнє значення `Accuracy` = 1,0000 (стійкий відгук на перенавчання); SVM RBF показав середнє значення `Accuracy` = 0,9856 (діапазон [0,9810; 0,9880]).

Аналіз динаміки помилок при збільшенні частки тестової вибірки підтвердив, що модель SVM з RBF-ядром зберігає точність вище 95% навіть при `test_size` = 0,80, що обґрунтовує її вибір як найбільш стійкої, стабільної та оптимальної для практичної імплементації в системах підтримки прийняття рішень і моніторингу ментального здоров'я користувачів.

Для виявлення прихованих поведінкових паттернів та сегментації користувачів без попередньої розмітки цільової змінної застосовано кластеризацію за методом `k-means`. Для усунення впливу різниці в масштабах ознак процедуру виконано на попередньо `z`-стандартизованому просторі даних.

Для визначення оптимальної кількості кластерів `k` проведено розрахунок внутрішньокластерної суми квадратів відстаней та середнього коефіцієнта силуету для `k` ∈ [2; 10]: при `k` = 2 зафіксовано максимальне значення силуетного коефіцієнта на рівні 0,42, що свідчить про чітку міжкластерну відокремлюваність та компактність сформованих груп; збільшення кількості кластерів (`k` ≥ 3) призводило до суттєвого падіння значення метрики силуету (<0,28) та розмивання меж між сегментами. Таким чином, для подальшого аналізу обрано двокластерну структуру (`k` = 2).

Аналіз центроїдів виділених кластерів дозволив ідентифікувати два кардинально протилежні поведінкові профілі користувачів (рис.1).

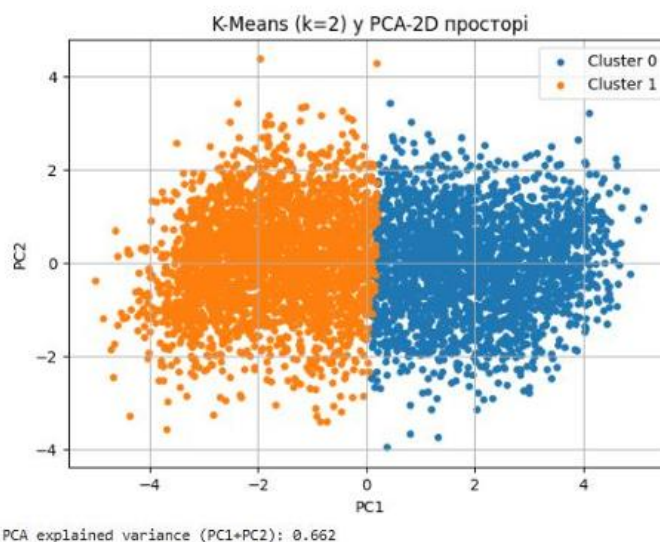


Рис. 1. Кластери `k-means` (`k` = 2) у просторі перших двох головних компонент

Кластер 0 «Ресурсний профіль (Низький ризик)» (обсяг: ~58% вибірки) характеризується помірним рівнем екранного часу (`daily_screen_time_hours` ≈ 4,1 год) та відносно низькою тривалістю використання соціальних мереж (`social_media_hours` ≈ 2,1 год); демонструє високі показники захисних ресурсів – висока якість сну (`sleep_quality` ≥ 4,2), тривала тижнева фізична активність (`physical_activity_hours_per_week` ≈ 5,4 год) та висока оцінка ментального здоров'я (`mental_health_score` ≈ 7,8); середнє значення рівня стресу в цій групі становить  $3,2 \pm 1,1$ , а розрахований ІЕН < 40.

Кластер 1 «Профіль цифрового ризику (Група дистресу)» (обсяг: ~42% вибірки) характеризується критичним рівнем занурення в цифровий простір, тобто щоденний екранний час перевищує 6,8 годин, із яких понад 4,6 години припадає на соціальні мережі; супроводжується дефіцитом компенсаторних



ресурсів – низька якість сну ( $<3,1$ ), суттєве скорочення фізичної активності ( $<1,8$  год/тиждень) та підвищені показники тривожності й депресивних симптомів; середній рівень стресу становить  $8,1 \pm 1,2$ , а значення ІЕН  $> 70$ , що відповідає стану вираженого цифрового виснаження.

Для наочної графічної інтерпретації та підтвердження відокремлюваності кластерів у багатовимірному просторі ознак (25 параметрів) застосовано метод аналізу головних компонент. Зниження розмірності дозволило виділити дві ортогональні головні компоненти: перша головна компонента ( $PC_1$ ) акумулює  $\approx 41,2\%$  загальної дисперсії даних і найбільше корелює з показниками цифрового навантаження та тривожності; друга головна компонента ( $PC_2$ ) пояснює  $\approx 18,5\%$  дисперсії і відображає фізіологічні та ресурсні фактори (сон, фізична активність).

Проекція об'єктів на площину ( $PC_1, PC_2$ ) продемонструвала виражене просторове розмежування двох кластерів уздовж осі  $PC_1$ . Це підтверджує, що саме баланс між цифровим навантаженням та компенсаторними ресурсами є головним латентним фактором, який розділяє користувачів за рівнем психоемоційного ризику.

### ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У статті вирішено актуальне науково-прикладне завдання щодо підвищення об'єктивності та точності оцінювання психоемоційного стану користувачів у цифровому середовищі на основі застосування алгоритмів класифікації й кластеризації.

На етапі підготовки та якісного аналізу даних проведено первинну обробку й очищення набору даних. Після видалення 8 дубльованих копій зафіксовано репрезентативну вибірку обсягом 5005 спостережень із 25 параметрами. За допомогою one-hot кодування категоріальних змінних та z-стандартизації числових ознак сформовано уніфікований препроцесинговий пайплайн, що усунув проблему різномасштабності чинників.

У межах формалізації авторського індексу розроблено та математично обґрунтовано ІЕН, який інтегрує вагові коефіцієнти факторів ризику (час у соцмережах, рівень тривожності) та компенсаторних ресурсів (якість сну, ментальне здоров'я, фізична активність) пропорційно їхнім кореляційним зв'язкам із рівнем стресу. Нормалізація індексу до шкали  $[0; 100]$  дозволила встановити чіткі межні критерії: значення ІЕН  $> 70$  ідентифікують критичний стан цифрового виснаження, тоді як ІЕН  $< 40$  відповідає високому рівню компенсаторної стійкості.

Регресійне оцінювання чинників за допомогою побудованої моделі множинної лінійної регресії ( $R^2 = 0,9389$ ) та аналізу часткових регресійних діаграм довело, що саме тривалість використання соціальних мереж ( $\beta_1 = +1,5850$ ,  $p < 0,001$ ), а не загальний екранний час поза ними ( $\beta_2 = +0,1172$ ), є основним деструктивним тригером зростання дистресу.

Під час порівняльного аналізу моделей класифікації здійснено квантильний поділ рівня стресу на три збалансовані класи (Low, Mid, High). Зіставлення алгоритмів Naïve Bayes (Accuracy = 96,87%), Decision Tree (100% – маркер перенавчання) та SVM з RBF-ядром (98,93%) за умов 5-fold крос-валідації довело, що модель SVM RBF є найбільш стійкою та робастною (Accuracy<sub>CV</sub> = 0,9856), зберігаючи точність понад 95% навіть при суттєвому зменшенні навчальної вибірки.

Результати сегментації поведінкових профілів методом кластеризації k-means ( $k=2$ , Silhouette score = 0,42) у поєднанні з візуалізацією через метод головних компонент дозволили виділити дві ключові групи: «Ресурсний профіль» (58% вибірки) та «Профіль цифрового ризику» (42% вибірки). Це підтвердило, що превалювання факторів цифрового навантаження над компенсаторними ресурсами є головним латентним розділювальним чинником у багатовимірному просторі ознак.

Отримані результати створюють методичне підґрунтя для розробки програмних модулів автоматизованого моніторингу психоемоційного стану користувачів. Подальші науково-прикладні розвідки будуть спрямовані на створення інтелектуального веб-застосунку/мобільного модуля на основі розробленої алгоритмічної моделі SVM RBF та індексу ІЕН, що дозволить надавати користувачам персоналізовані рекомендації з цифрової гігієни та оптимізації режиму праці й відпочинку в реальному часі.



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Zhekova, V. V. (2024). *Otsinka rivnia stresu u studentiv pid chas navchannia ta ioho vplyv na stan zdorov'ia: Rezultaty opytuvannia* [Assessment of student stress levels during studying and its impact on health status: Survey results]. Odesa National Medical University Repository. Світ наукових досліджень. Випуск 35: матеріали Міжнародної мультидисциплінарної наукової інтернет-конференції (м. Тернопіль, Україна, м. Ополь, Польща, 20-21 листопада 2024 р.) / за ред. : О. Патряк та ін. ГО "Наукова спільнота", WSZIA w Opolu. Тернопіль: ФО- П Ішпак В.Б. 2024. 230-232.
2. Bilousova, S. Y. (2024). *Riven stresu u studentiv pid chas navchannia ta ioho vplyv na zdorov'ia v umovakh viiny* [Stress level in students during study and its impact on health under conditions of war]. Odesa National Medical University Repository. The 2nd International scientific and practical conference "Science in the modern world: innovations and challenges" (October 24-26, 2024) Perfect Publishing, Toronto, Canada. 2024. 71-75
3. Batrona, D., Latyshev, M., Petrova, N., Savka, M., & Tsimbolynets, N. (2025). Dotrymanna zdorovoho sposobu zhyttia zdobuvachamy vyshchoi osvity: Kharchuvannia, son, rukhova aktivnist [Adherence to a healthy lifestyle by higher education applicants: Nutrition, sleep, physical activity]. *Osvita. Innovatyka. Praktyka*, 13(7), 20–25.
4. Borodina, O. S., Svatiev, A. V., Laichuk, A. M., Husarevych, O. V., & Burkovskiy, V. V. (2025). Rukhova aktivnist i psykholohichne zdorov'ia: Pedahohichni stratehii pidtrymky balansu u zdobuvachiv vyshchoi osvity [Physical activity and psychological health: Pedagogical strategies for maintaining balance in higher education students]. *Olimpiiskiy ta Paralimpiiskiy Sport*, (3), 7–11.
5. Karim F, Oyewande AA, Abdalla LF, Chaudhry Ehsanullah R, Khan S. Social Media Use and Its Connection to Mental Health: A Systematic Review. *Cureus*. 2020 Jun 15;12(6):e8627. doi: 10.7759/cureus.8627.
6. Smith, A. B., & Johnson, C. D. (2022). Screen time and mental health symptoms among young adults: A longitudinal study. *Developmental Psychology*, 58(9), 1740-1752.
7. Elhai, J. D., Tiamiyu, M. F., & Weeks, J. W. (2018). Distress tolerance and mindfulness mediate relations between depression and anxiety sensitivity with problematic smartphone use. *Computers in Human Behavior*, 84, 182-188.
8. Bond RA, Elhai JD, Montag C. Exploring the Mediating Role of Emotion Regulation and Distress Intolerance in the Relationship Between Social Anxiety and Problematic Smartphone Use: A Cross-Sectional Study. *Health Sci Rep*. 2026 Apr 16;9(4):e72356. doi: 10.1002/hsr.2.72356.
9. Ramadhan, R. N., Rampengan, D. D., Yumnانشa, D. A., Setiono, S. B., Tjandra, K. C., Ariyanto, M. V., ... Empitu, M. (2024). Impacts of digital social media detox for mental health: A systematic review and meta-analysis. *Narra J*, 4(2), e786. <https://doi.org/10.52225/narra.v4i2.786>.
10. Lomachinska, I., Volynets, I., & Khrypko, S. (2025). Socio-psychological aspects of students' digital health in the context of digital age challenges. *Osvitnolohichni Diskurs*, 50(3), 94-106.
11. Moroka, R. K., Kropivnyi, M. V., Kainara, V. M., & Trehub, P. O. (2024). *Vykorystannia AI v zhytti studentiv ta pokrashchennia pratsi za dopomohoiu AI v sferakh okhorony zdorov'ia* [The use of AI in students' lives and work improvement using AI in healthcare]. Громадське здоров'я в Україні: проблеми та способи їх вирішення «Томілінські читання»: матеріали VII наук.-практ. конф. з міжнар. участю (Харків, 30 жовтня 2024 р.) / Ред. кол.: В. В. М'ясоєдов, В. А. Огнєв, К. М. Сокол та ін. Харків, ХНМУ, 2024. <https://doi.org/10.5281/zenodo.14194418>.
12. Verchynska, M. O., & Voitsekhovska, O. O. (2025, June 15-16). Rozviduvalnyi analiz danykh dlia informatsiinoi tekhnolohii analizu ta prohnozuvannia stresu u studentiv zakladiv vyshchoi osvity metodamy mashynnoho navchannia [Exploratory data analysis for information technology of stress analysis and forecasting in higher education students using machine learning methods]. In *Molod v nautsi: Doslidzhennia, problemy, perspektyvy (MN-2025)* [Conference proceedings] (pp. 45-48). Vinnitsa.
13. Karim F, Oyewande A A, Abdalla L F, et al. (June 15, 2020) Social Media Use and Its Connection to Mental Health: A Systematic Review. *Cureus* 12(6): e8627. doi:10.7759/cureus.8627.
14. Verma, P., & Kumar, S. (2023). Stress assessment using machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, 14(5), 320-328.
15. Patel, R., Sharma, A., & Gupta, S. (2024). *Machine learning algorithms for detecting mental stress in college students* (arXiv Preprint No. arXiv:2412.07415v1). arXiv. <https://arxiv.org/html/2412.07415v1>

**Olha Korotun**

PhD in Pedagogical Sciences, Associate Professor,  
Associate Professor of the Computer Science Department  
Zhytomyr Polytechnic State University, Zhytomyr, Ukraine  
ORCID: 0000-0003-2240-7891  
korotun-o@ztu.edu.ua

**Galyna Marchuk**

Senior Lecturer of the Computer Science Department  
Zhytomyr Polytechnic State University, Zhytomyr, Ukraine  
ORCID: 0000-0003-2954-1057  
pzs\_mgv@ztu.edu.ua

**Alisa Lysiuk**

Student  
Zhytomyr Polytechnic State University, Zhytomyr, Ukraine  
ORCID: 0009-0007-1990-5458  
ist231\_lao@student.ztu.edu.ua

**MODELING AND EVALUATING THE LEVEL OF PSYCHO-EMOTIONAL LOAD OF USERS IN THE DIGITAL ENVIRONMENT**

**Abstract.** The article addresses an urgent scientific and practical problem of increasing the objectivity and accuracy of assessing the psycho-emotional state of users in the digital environment by applying classification and clustering algorithms. The introduction substantiates the contradiction between the growing level of user digital distress and the limitations of traditional diagnostic methods, which determines the need to develop mathematical models based on digital traces. At the data preparation stage, a dataset of 5,000 observations with 25 parameters covering digital behavior, physiology, and mental health was cleaned, and a preprocessing pipeline was formed using One-Hot encoding of categorical variables and z-score standardization of numerical features. The authors proposed and mathematically formalized an original integrated Emotional Load Index (IEH), which balances the weight coefficients of risk factors and compensatory body resources proportionally to their correlation with stress levels. Using a multiple linear regression model ( $R^2 = 0.9389$ ), it was proven that the primary trigger of distress is the duration of social media use rather than total screen time outside social platforms. Based on quantile division of stress levels into three classes (Low, Mid, High), a comparative analysis of Naive Bayes, Decision Tree, and SVM classification models with an RBF kernel was conducted. The results of 5-fold cross-validation demonstrated that the SVM RBF model possesses the highest generalization ability and robustness against overfitting (Accuracy = 0.9856). Applying k-means clustering ( $k=2$ , Silhouette score = 0.42) combined with visualization in the Principal Component Analysis (PCA) space allowed identifying two behavioral profiles: "Resourceful" and "Digital Risk". The practical value lies in creating a foundation for developing software modules for real-time automated monitoring of user mental health.

**Keywords:** psycho-emotional load; digital environment; emotional load index; machine learning; support vector machine; classification; clustering; screen time; digital behavior.

**REFERENCES (TRANSLATED AND TRANSLITERATED)**

1. Batrona, D., Latyshev, M., Petrova, N., Savka, M., & Tsimbolynets, N. (2025). Dotrymannia zdorovoho sposobu zhyttia zdobuvachamy vyshchoi osvity: Kharchuvannia, son, rukhova aktyvnist [Adherence to a healthy lifestyle by higher education applicants: Nutrition, sleep, physical activity]. *Osvita. Innovatyka. Praktyka*, 13(7), 20–25.
2. Bilousova, S. Y. (2024). Riven stresu u studentiv pid chas navchannia ta ioho vplyv na zdorov'ia v umovakh viiny [Stress level in students during study and its impact on health under conditions of war]. In *The 2nd International scientific and practical conference "Science in the modern world: innovations and challenges"* (pp. 71–75). Perfect Publishing.



3. Bond, R. A., Elhai, J. D., & Montag, C. (2026). Exploring the mediating role of emotion regulation and distress intolerance in the relationship between social anxiety and problematic smartphone use: A cross-sectional study. *Health Science Reports*, 9(4), Article e72356. <https://doi.org/10.1002/hsr.2.72356>
4. Borodina, O. S., Svatiev, A. V., Laichuk, A. M., Husarevych, O. V., & Burkovskiy, V. V. (2025). Rukhova aktivnist i psykhologichne zdorov'ia: Pedagogichni stratehii pidtrymky balansu u zdobuvachiv vyshchoi osvity [Physical activity and psychological health: Pedagogical strategies for maintaining balance in higher education students]. *Olimpiyskiy ta Paralimpiyskiy Sport*, (3), 7–11.
5. Elhai, J. D., Tiamiyu, M. F., & Weeks, J. W. (2018). Distress tolerance and mindfulness mediate relations between depression and anxiety sensitivity with problematic smartphone use. *Computers in Human Behavior*, 84, 182–188. <https://doi.org/10.1016/j.chb.2018.02.029>
6. Karim, F., Oyewande, A. A., Abdalla, L. F., Chaudhry Ehsanullah, R., & Khan, S. (2020). Social media use and its connection to mental health: A systematic review. *Cureus*, 12(6), Article e8627. <https://doi.org/10.7759/cureus.8627>
7. Lomachinska, I., Volynets, I., & Khrypko, S. (2025). Socio-psychological aspects of students' digital health in the context of digital age challenges. *Osvitnolohichniy Diskurs*, 50(3), 94–106.
8. Moroka, R. K., Kropivnyi, M. V., Kainara, V. M., & Trehub, P. O. (2024). Vykorystannia AI v zhytti studentiv ta pokrashchennia pratsi za dopomohoiu AI v sferakh okhorony zdorov'ia [The use of AI in students' lives and work improvement using AI in healthcare]. In V. V. Miasoiedov et al. (Eds.), *Public Health in Ukraine: Problems and Ways to Solve Them "Tomilin Readings": Proceedings of the VII Scientific and Practical Conference* (pp. 112–115). Kharkiv National Medical University. <https://doi.org/10.5281/zenodo.14194418>
9. Patel, R., Sharma, A., & Gupta, S. (2024). *Machine learning algorithms for detecting mental stress in college students* (arXiv Preprint No. arXiv:2412.07415v1). arXiv. <https://arxiv.org/abs/2412.07415>
10. Ramadhan, R. N., Rampengan, D. D., Yumnانشa, D. A., Setiono, S. B., Tjandra, K. C., Ariyanto, M. V., & Empitu, M. (2024). Impacts of digital social media detox for mental health: A systematic review and meta-analysis. *Narra J*, 4(2), Article e786. <https://doi.org/10.52225/narra.v4i2.786>
11. Smith, A. B., & Johnson, C. D. (2022). Screen time and mental health symptoms among young adults: A longitudinal study. *Developmental Psychology*, 58(9), 1740–1752. <https://doi.org/10.1037/dev0001382>
12. Verchynska, M. O., & Voitsekhovska, O. O. (2025, June 15–16). Rozviduvalnyi analiz danykh dlia informatsiinoi tekhnolohii analizu ta prohnozuvannia stresu u studentiv zakladiv vyshchoi osvity metodamy mashynnoho navchannia [Exploratory data analysis for information technology of stress analysis and forecasting in higher education students using machine learning methods]. In *Youth in Science: Research, Problems, Prospects (MN-2025): Conference Proceedings* (pp. 45–48).
13. Verma, P., & Kumar, S. (2023). Stress assessment using machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, 14(5), 320–328. <https://doi.org/10.14569/IJACSA.2023.0140536>
14. Zhekova, V. V. (2024). Otsinka rivnia stresu u studentiv pid chas navchannia ta ioho vplyv na stan zdorov'ia: Rezultaty opytuvannia [Assessment of student stress levels during studying and its impact on health status: Survey results]. In O. Patryak et al. (Eds.), *World of Scientific Research: Materials of the 35th International Multidisciplinary Scientific Internet Conference* (pp. 230–232). NGO "Scientific Community", WSZIA w Opolu.

Отримано редакцією журналу / Received: 18.06.26

Прорецензовано / Revised: 28.07.26

Схвалено до друку / Accepted: 26.10.26



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.