ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

УДК 004.51 DOI: https://doi.org/10.30837/2522-9818.2025.4.044

Є. Ігнатюк, А. Попов

ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ UX/UI-ДАНИХ МАСОВИХ ОПИТУВАНЬ КОРИСТУВАЧІВ

Предметом дослідження є впровадження методів машинного навчання для інтерпретації UX/UI-даних, зібраних способом масових опитувань користувачів освітніх онлайн-платформ. У роботі розглянуто гіпотезу про те, що узгоджене застосування різних моделей класифікації дає змогу виокремити поведінкові патерни, які мають прогностичне значення для оцінки використання окремих функцій продукту. **Метою** є порівняльний аналіз точності таких моделей на прикладі реального UX/UI-опитування. **Методика дослідження** передбачає етапи підготовки даних, кодування ознак, нормалізації, кластеризації та побудови моделей шести типів: дерева рішень, випадкового лісу, градієнтного бустингу, багатошарової нейромережі, логістичної регресії та методу *k*-найближчих сусідів. Основну увагу зосереджено на тому, як кожна з моделей поводиться в умовах UX/UI-даних, що мають обмежений обсяг, складну структуру й множину змішаних ознак. **Результати** моделювання демонструють, що навіть у межах невеликих вибірок моделі можуть виявляти значущі залежності між соціально-демографічними факторами, типами користувачів і функціональною активністю. **Висновки.** Застосування машинного навчання до UX/UI-даних є перспективним підходом до аналізу поведінки користувачів з можливістю подальшої інтеграції таких підходів у системи підтримки рішень. Запропонований підхід дає змогу виявляти стійкі поведінкові патерни користувачів на основі структурованих UX/UI-даних і має потенціал для подальшого розвитку методів прогнозування функціональної активності. Це відкриває перспективи інтеграції таких рішень у системи підтримки рішень в онлайн-освіті, охороні здоров'я, державному управлінні та інших сферах, де аналіз користувацьких даних відіграє критичну роль.

Ключові слова: UX/UI-аналітика; анкетні дані; машинне навчання; поведінкові сценарії; нейронні мережі; ансамблеві моделі; цифрові сервіси; досвід взаємодії (UX).

Вступ і постановка проблеми

У сучасному цифровому середовищі дизайн інтерфейсів відіграє ключову роль не лише в естетичному сприйнятті, а й у зручності використання цифрових сервісів [1] – від освітніх платформ і банківських застосунків до систем електронного врядування. Залежно від того, наскільки зручною, зрозумілою та послідовною є взаємодія користувача з інтерфейсом, безпосередньо залежить ефективність його застосування, рівень залученості аудиторії та показники утримання. Усе більше компаній – від стартапів до великих платформ – інтегрують UX/UI як окрему стратегічну функцію в процес розроблення [2].

Чимало дослідників і практиків вивчають оптимізацію та систематизацію покращення інтерфейсів та їх окремих аспектів, зокрема візуальних або структурних. Так, наприклад, у роботі *Heil* та *Gaedke* [3] запропоновано систему автоматичного аналізу візуальної складності вебінтерфейсів, що поєднує методи комп'ютерного зору, розпізнавання тексту й просторового структурування для обчислення метрик складності, які потім корелюються з оцінками користувачів. Автори дослідження *Computational Layout Perception using Gestalt Laws* [4] алгоритмічно змоделювали сприйняття інтерфейсу користувача (UI), упроваджуючи гештальт-принципи як евристику для автоматизованого групування елементів. Це дало змогу сформувати ієрархічну структуру, яка в 90% випадків узгоджувалась з очікуваною візуальною організацією інтерфейсу, що відкриває потенціал для застосування в завданнях адаптивного дизайну або автоматизованого UX-аудиту.

Використовуючи згорткові нейронні мережі, *Wang, Zhang, Du* та ін. у дослідженні [5] запропонували модель кількісного оцінювання якості кольорового оформлення інтерфейсів. Модель витягує та аналізує багатовимірні ознаки зображень інтерфейсів (насиченість, яскравість, відтінок) і зіставляє їх з оцінками користувачів, щоб уникнути суб'єктивності й забезпечити стабільну, формалізовану оцінку візуальної якості інтерфейсів.

Проте, щоб створювати не лише красиві, а й ефективні інтерфейси, дизайнери мають спиратися на дослідження користувацького досвіду, додатково беручи до уваги UI-складник. UX/UI-дослідження дають змогу не лише з'ясувати проблеми поточної взаємодії, а й виявити потреби, мотивації та бар'єри користувачів. У цьому сенсі вони стають аналітичною опорою для дизайнерських і продуктових рішень [6]:

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

допомагають сформулювати гіпотези, перевірити сценарії використання, виявити закономірності та обґрунтувати зміни. Дослідження бувають різними за форматом і обсягом [7] – від глибинних інтерв'ю та масових опитувань до п'ятисекундних тестів та A/B-тестування. І кожен із цих методів дає унікальний зріз реального користувацького досвіду. Однак розмаїття форматів і збільшення обсягів дослідницької інформації ускладнює її оброблення: результати можуть бути структурованими або ні, містити відкриті текстові відповіді, числові шкали, категоріальні змінні, поведінкові залежності, бути у форматі неструктурованого діалогу, містити відеозаписи користування інтерфейсом тощо. Показовим прикладом є актуальне дослідження, де *Zhang* та команда використали тематичне моделювання для автоматизованого аналізу тисяч текстових UX-відгуків на маркетплейсі, виявивши ключові сценарії поведінки користувачів і тренди взаємодії [8]. Ця та подібні роботи демонструють, що аналіз UX-даних активно розвивається, і дослідження в цій сфері з'являються протягом останніх 2–3 років. Водночас більшість з них зосереджені на неструктурованих або слабкоформалізованих джерелах, текстах, відгуках, відео, де застосування NLP або *vision*-підходів є найбільш очевидним. Формалізовані ж UX/UI-дані, зокрема результати опитувань, що мають структуру шкал, категорій і числових змінних, залишаються менш дослідженим полем. Утім, саме вони так само мають значний потенціал для порівнюваного, стандартизованого аналізу, що особливо цінно в умовах масштабування досліджень.

Особливо гостро це питання постає під час спроби працювати з повторюваними дослідженнями або вводити UX-аналітику як постійний процес, а не разову активність. Тоді на перший план виходять методи структурування, інтерпретування та порівняння інформації, незалежно від її масштабів. У реальних умовах усе це часто лягає на плечі дизайнера, який одночасно має будувати вигляд інтерфейсу та проєктувати досвід взаємодії [9], але не завжди володіє достатніми ресурсами для складного оброблення даних. Це створює запит на побудову систем, здатних частково або повністю брати на себе ці функції, знижуючи навантаження на команду та забезпечуючи стабільність аналізу з часом.

У науковій літературі, яку ширше розглянуто в наступному розділі, відзначено висхідну потребу в нових підходах до аналізу UX/UI-даних, зокрема в умовах повторюваних вимірювань, збільшенні проєктів і роботи з різними форматами даних (текстовими, візуальними чи поведінковими). Попри зростання кількості досліджень, що інтегрують UX-підхід з автоматизованим аналізом, більшість з них зосереджуються на нестандартизованих джерелах інформації. Натомість формалізовані UX/UI-дані, які дають змогу системно збирати стандартизовану інформацію про досвід користувачів, досі залишаються малодослідженим полем. Саме ця лакуна відкриває можливість для системного вивчення потенціалу формалізованих джерел інформації, наприклад UX/UI-опитування, для обчислювального аналізу, що може закласти підґрунтя для програмних стабільних підходів до UX-аналітики в цифрових продуктах. А загальна зацікавленість дизайн-спільноти у використанні нейромереж [10] підсилює ці дослідження, що формує засновок їх застосування для аналізу UX-даних зокрема.

У попередній науковій публікації [11] запропоновано концептуальну архітектуру комплексного рішення для оброблення результатів UX/UI-досліджень. У ній окреслено модулі роботи з різними форматами інформації (структурованими, візуальними й текстовими) й описано методичний підхід до поєднання аналітики з дизайнерським процесом. Зокрема порушено питання про необхідність моделей, здатних працювати на повторюваних наборах даних різного походження, що збираються протягом життєвого циклу сервісу.

Однак, попри наявні теоретичні основи досліджень, що демонструють роботу таких моделей на реальних UX/UI-даних, наприклад, з опитувальників, майже немає. Саме цей розрив між концептуальною можливістю та практичною реалізацією і визначає наукову новизну цієї роботи. Відповідно, аналітика анкетних UX/UI-даних із застосуванням моделей машинного навчання – потенційна можливість стандартизації UX-аналітики як повторюваного процесу.

Ця стаття присвячена одному з напрямів – формалізованій інформації з масових опитувань. В її основі лежить спроба опрацювати реальні анкетні UX/UI-дані користувачів цифрової освітньої платформи, трансформувати їх у формат, придатний для обчислювального аналізу, й оцінити ефективність різних моделей машинного навчання в завданні

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

прогнозування поведінкових сценаріїв. Отже, ця робота не лише розглядає конкретний приклад використання анкетної інформації, але й продовжує реалізацію загального дослідницького підходу до роботи з UX/UI-даними в межах запропонованої структури.

Аналіз останніх досліджень і публікацій

Зі зростанням складності цифрових інтерфейсів і варіативності користувацьких сценаріїв UX/UI-дослідження набули статусу системної практики, необхідної для створення інтерфейсів, що дійсно відповідають потребам користувачів. Сучасні дослідження у сфері UX-фахової аналітики охоплюють широкий спектр підходів – від методів збору інформації до їх інтерпретації та подальшого застосування в проєктних рішеннях. У межах цієї роботи увагу привертають наукові спроби поєднати результати UX-досліджень з алгоритмами машинного навчання, зокрема для побудови адаптивних або персоналізованих систем взаємодії.

У згаданому раніше дослідженні *Zhang, Chen* і *Huang* демонструють використання моделей для передбачення потреб нових користувачів у дизайнерських середовищах. Автори доводять, що навіть обмежені за обсягом UX-дані можуть бути корисними для побудови практичних предикативних сценаріїв. Це свідчить про потенціал простих опитувальних даних у завданнях індивідуалізації досвіду. *Zender, Humm* і *Holzheuser* [12] вивчали інтеграцію UX-принципів у дизайн взаємодії з AI-системами. Результати дослідження вказують на те, що адаптивні інтерфейси, побудовані на основі користувацьких шаблонів поведінки, підвищують довіру та задоволеність від системи. Зібрана інформація є не лише аналітичним ресурсом, а й основою для проєктування сценаріїв взаємодії. У роботі аналізується застосування пояснюваного штучного інтелекту (XAI) для UX-оцінювання. Запропонований підхід не лише класифікує проблеми користувацького досвіду, а й пояснює ці рішення, що сприяє зростанню довіри з боку кінцевих користувачів.

Duan, Chen та *Li* [13] запропонували застосовувати методи комп'ютерного зору для автоматичного виявлення елементів інтерфейсу у відеозаписах. Це стандартизує UX-аналіз на ранніх етапах без потреби в повній залученості людини як експерта. Показовим є приклад *Čejka, Smutný, Vondrák* та колег [14], які використовували *YOLOv7* для побудови динамічних теплових карт взаємодії. Модель відстежує події в реальному часі, аналізуючи зміну елементів інтерфейсу на відео. Ці підходи демонструють прогрес у напрямі мультимодального UX-аналізу.

Водночас, як свідчить публікація *Batch, Ji, Fan*, дослідники визнають важливість переходу від суто відео- або аудіоаналізу до комбінованих методів, здатних працювати з простішими типами даних. У роботі [15] запропоновано систему *uxSense*, що поєднує технології комп'ютерного зору, оброблення природної мови та базове класифікування на основі UX-метрик, хоча застосування нейронних мереж виявилось обмеженим. А в огляді [16] у журналі *Human-Centric Intelligent Systems* узагальнено сучасні напрями застосування нейромереж для поліпшення користувацького досвіду. Поряд із мультимодальними підходами, у роботі наголошено на недостатності досліджень, спрямованих на структури на кшталт анкет чи табличних даних, які залишаються найбільш доступними в умовах обмежених ресурсів.

Dou, Zheng і *Sun* у праці [17] запропонували модель *Webthetics* – глибоку нейронну мережу, розроблену для обчислювального оцінювання естетичності вебсторінок. Ця система навчається на реальних оцінках користувачів і застосовує трансферне навчання зі стилістичної класифікації зображень для підвищення точності. Замість ручних ознак, таких як колір чи симетрія, система самостійно виокремлює багатовимірні репрезентації візуального оформлення, що відповідають оцінкам користувачів. Це уможливлює об'єктивне порівняння варіантів UI-дизайну на основі глибоко інтегрованих параметрів. Хоча модель працює з візуальною інформацією, ключовим є те, що саме формалізовані оцінки користувачів були застосовані як навчальні дані.

Використанню простих, але формально структурованих UX-даних у поєднанні з машинним навчанням присвячено дослідження *Zhang et al.* [8], у якому запропоновано інструмент для автоматизованої класифікації користувацьких потреб на основі моделі *Kano*. Автори об'єднують анкетні UX-дані, глибинні інтерв'ю та зібрані онлайн-відгуки, щоб сформувати повноцінний навчальний набір для моделей глибокого навчання. Найуспішнішою виявилась модель RCNN (*Recurrent Convolutional Neural*

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Network), яка дала змогу класифікувати текстові фрагменти за типами потреб. Важливо, що результатом стала не лише класифікація, а й інтерактивний інструмент з графічним інтерфейсом, орієнтований на початківців у UX. Тобто аналітика складних моделей стала доступною навіть у навчальному контексті.

Проте питання автоматизації UX-дизайну не зводиться лише до передбачення реакцій користувача. У праці [18] порушено глибшу проблему: UX-дизайн є процесом із непередбачуваним результатом, який містить не лише логіку, а й інтуїцію, досвід, гнучке реагування на контекст. Автор пропонує розглядати оптимізацію як підтримку на етапах зворотного зв'язку та генерації варіантів для створення нових рішень або аналізу даних. Однак ці моделі мають уміти адаптуватись до неоднозначних сценаріїв і коригуватися людиною в разі нових або нестандартних запитів. Навіть за умови успішного дизайну, на думку автора, користувацький досвід може розвиватися несподівано.

Власна попередня публікація присвячена архітектурі модульної системи для підтримки продуктового та дизайнерського прийняття рішень на основі UX/UI-досліджень. Запропонована структура передбачає інтеграцію моделей, що працюють з різними форматами інформації: текстовий (наприклад, відкриті відповіді), числовий (наприклад, анкети), візуальний (відеозаписи інтерв'ю чи з екранів) та поведінковий (переходи, теплові карти, перегляди). Для кожного з цих типів передбачається використання спеціалізованих моделей машинного навчання, які можуть бути інтегровані в єдину систему через центральний шар логіки. У роботі наголошено на важливості формування бази знань, яка не лише проводить аналітику системи, а й формулює рекомендації щодо покращення інтерфейсних рішень.

Попри висхідну кількість досліджень на перетині UX-аналітики та машинного навчання, переважає увага до аналізу мультимедійних або складноструктурованих даних [19]. Натомість можливості нейромережевих моделей щодо оброблення табличних результатів масових UX/UI-опитувань залишаються недостатньо вивченими. У цій роботі висувається гіпотеза про те, що навіть в умовах обмежених, але формалізованих UX-даних можливо досягти ефективної класифікації користувачів із застосуванням сучасних моделей машинного навчання. Наукова новизна полягає в апробації цього підходу на реальних UX/UI-даних опитування, що відтворюють взаємозв'язок між користувачами, поведінковими шаблонами й важливими для продукту метриками. Такий підхід дає змогу розширити інструментарій UX-досліджень у ресурсно обмежених проєктах, де недоступні складні або коштовні методи збирання та оброблення інформації. Він також може бути екстрапольований на інший освітній або сервісний продукт суміжних галузей і застосовуватись за схожим принципом, який описано в наступних розділах.

Мета й завдання дослідження

У межах загальної концепції побудови інструментарію для аналізу UX/UI-досліджень у цій статті необхідно вивчити можливості використання моделей машинного навчання для роботи з формалізованими UX/UI-даними, отриманими внаслідок масових опитувань користувачів. Особливий інтерес становить дослідження потенціалу цих моделей у контексті обмежених вибірок, властивих для цифрових сервісів із вузькопрофільною аудиторією, а також для стартапів і сервісів з короткостроковими функціоналами, де час і обсяг інформації обмежені. На відміну від мультимодальних чи поведінкових джерел, анкети становлять один з найбільш доступних форматів UX-досліджень, а тому потребують адаптації сучасних інструментів аналізу. Ключовою метою роботи є оцінювання ефективності застосування моделей на основі дерев рішень і перцептронів до структурованих UX/UI-даних масових опитувань для їх класифікації та визначення придатності для долучення до модульної системи UX-аналітики.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

1) підготувати UX/UI-дані для подальшого аналізу, зокрема з перейменуванням полів, кодуванням ознак, заповненням пропущених значень, бінаризацією відповідей і нормалізацією числових шкал;

2) здійснити попередній дослідницький аналіз даних, зокрема визначити основні взаємозв'язки за допомогою кореляційного аналізу, оцінити залежності між змінними, а також виконати кластеризацію за поведінковими характеристиками;

3) визначити ключову цільову змінну, що може бути використана як прогнозний індикатор у моделі, та обґрунтувати її вибір;

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

4) побудувати моделі машинного навчання, як базові (*Decision Tree*), так і ансамблеві (*Random Forest,* XGBoost), а також модель на основі багатошарового перцептрона (MLP) для подальшого порівняння їх результативності;

5) оцінити ефективність цих моделей з використанням метрик точності, а також зробити порівняльний аналіз продуктивності на вибірці даних.

Поставлені завдання визначають логіку наступного розділу, в якому поетапно подано реалізацію кожного з етапів – від формування вибірки до побудови й тестування моделей.

Методика проведення UX/UI-опитування

У сфері дослідження користувацького досвіду застосовується широкий спектр підходів – від якісних глибинних інтерв'ю до поведінкового відстежування та аналізу метрик взаємодії. Ці дослідження розрізняють [20, 21] за різними критеріями: за метою (оцінювання ставлення, поведінки чи очікувань), типом даних (якісні, кількісні), джерелом (пряме опитування, непряме спостереження, автоматизований запис взаємодії), методом збору тощо. У цьому дослідженні обрано кількісний підхід, а саме масове анкетування, спрямоване на вивчення ставлення користувачів до обраної платформи та їхнього досвіду взаємодії. Попри потенційні обмеження в глибині відповідей, цей підхід є цінним інструментом для пошуку загальних закономірностей у поведінці користувачів і прийнятті рішень.

Опитування було реалізовано в межах цифрової платформи CASES [22] – екосистеми для навчання, кар'єрного розвитку та пошуку роботи у сфері креативних індустрій. CASES має власну навчальну платформу, живі та записані курси, додаткові навчальні формати, базу студентських профілів, можливості спілкування та пошуку роботи. Платформа орієнтована на українську аудиторію, зокрема молодих фахівців і фахівців середнього рівня, які навчаються або підвищують кваліфікацію у сфері дизайну, медіа, маркетингу та суміжних напрямів. UX/UI-дослідження в такому контексті є важливим інструментом підтримки гіпотез щодо поведінки, зацікавлень і мотивацій користувачів, зокрема для оптимізації підписок, структури курсів, інтерфейсних рішень і контентного подання матеріалу.

Анкета масового опитування була сформована через сервіс форм *Google* та поширена за допомогою соціальних мереж платформи у *Facebook, Telegram, Instagram* та електронної пошти. В опитуванні взяли участь понад 170 респондентів. Було запропоновано відповісти на серію запитань щодо їхнього стилю навчання, частоти використання платформи, професійного контексту, рівня доступу до техніки, самооцінки досвіду та інтересу до платних форматів. Відповіді були попередньо підготовлені, після цього експортовані у форматі CSV і завантажені в середовище для оброблення, що підтримує мову програмування *Python*, і саме з цього етапу було розпочато аналіз, оброблення та побудову моделей.

Попереднє оброблення інформації

Отримані результати були трансформовані в структуру табличного типу для подальшого опрацювання. У процесі попереднього оброблення неопрацьований набір даних поступово був трансформований у формат, придатний для машинного аналізу. У вихідній таблиці було 22 ознаки. Серед них можна було виокремити кілька ключових груп, перелік яких поданий у табл. 1. Інформація містила окремі пропущені значення (до 8–10% у певних стовпцях), а також комбіновані формати відповідей, що вимагали структурного поділу.

Усі етапи оброблення було реалізовано на *Python*. Розглянемо їх.

1. Розбиття складних змінних на окремі логічні ознаки. Наприклад, ознаку *active-subscription*, що відповідала одразу на два запитання (Чи була колись у користувача підписка? Чи є вона наразі?), було трансформовано у дві. Це ознаки *ever_had_subscription* та *currently_has_subscription*, які відтворюють відповіді на ці запитання окремо. Аналогічно ознаку *auditor-mode* було поділено на *knows_auditor_mode* та *used_auditor_mode*, а *employment-status* – на *is_employed* і *works_in_creative*.

2. Кодування впорядкованих категорій (*Ordinal Encoding*). Для ознак на кшталт *income-range, age, usage-time, skill-level* було зафіксовано логічну послідовність значень і закодовано їх як числові ранги. Наприклад, для ознаки *income-range* – послідовність від "не маю доходу" до "понад 40 тис. грн".

3. Кодування невпорядковних категорій за підходом *One-hot Encoding*. Ознаки *preferred-learning, preferred-online-courses, marital-status* тощо були

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

перетворені на множину бінарних змінних. Це усуває ризик того, що модель помилково сприйматиме номінальні значення як кількісні.

Наприклад, ознака *preferred-learning*, що містила значення на кшталт "самостійне навчання за власним графіком" або "навчання в групі з дедлайнами під наглядом куратора", була перетворена на кілька окремих бінарних стовпців, кожен з яких позначає наявність певного варіанта у відповіді.

Таблиця 1. *Групи ознак в табличній структурі даних*

№	Групи	Назва	Опис
1	Кількісні	*platform-usability*	Оцінка зручності користування
		search-difficulty	Оцінка складності пошуку
		feedback-quality	Оцінка якості зворотного зв'язку
2	Категоріальні впорядковані	*skill-level*	Рівень навичок у відповідній галузі
		age	Вік користувача
		income-range	Оцінка доходу
		usage-time	Час проведення на платформі
		earning-frequency	Як часто навчається на платформі?
3	Категоріальні невпорядковані	*preferred-online-courses*	Типи вподобаних онлайн-курсів
		archetype-purpose	Мотивація і мета користування платформою (кар'єра, хобі, перекваліфікація)
		gender	Стать користувача
		marital-status	Сімейний стан
		preferred-learning	Вподобаний формат навчання
		employment-status	Статус зайнятості
4	Складні комбіновані	*active-subscription*	Чи була колись активною підписка на освітній контент і чи активна підписка наразі?
		auditor-mode	Чи знає користувач про функцію "вільний слухач" та чи послуговується нею?
5	Мультивибіркові поля	*used-features*	Функції платформи, які активно використовуються
		three-main-functions	Основні функції платформи, які застосовує користувач
6	Бінарні	*platform-issues*	Наявність проблем користування
		children	Наявність дітей
		job-search	Чи перебуває в пошуку роботи?
		freelance-projects	Чи має фриланс-проєкти?

4. Кодування мультивибіркових полів (*MultiLabel Binarization*). Для полів *used-features* і *three-main-functions* використано *MultiLabelBinarizer*.

5. Перетворення бінарних змінних. Ознаки на зразок *platform-issues, job-search, freelance-projects* були уніфіковані до числового формату 0 чи 1.

6. Заповнення пропущених значень. Для логічних змінних, де *NaN* означало "ні / не знаю / не користуюсь", значення було замінено на 0. Кількісні змінні заповнювалися середнім по стовпцю. Це виконано для збереження всіх спостережень без вилучення рядків.

Після всіх трансформацій таблиця розширилась з 22 до 47 стовпців, що свідчить про значне зростання інформаційної щільності. Усі змінні набули цифрової форми – цілочисельної (*int*8, *int*64), логічної (*bool*) або з рухомою комою (*float*64) – і були повністю готові до моделювання.

Дослідницький аналіз (EDA)

На етапі дослідницького аналізу (*exploratory data analysis*, EDA) ставилась мета – попередньо виявити залежності між ознаками, оцінити якість і структуру зібраних UX/UI-даних, а також перевірити припущення щодо обраної цільової змінної. Саме тут аналітик може знайти неочевидні зв'язки між поведінковими характеристиками

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

користувачів і результатами їхньої взаємодії з платформою, а також сформулювати попередні гіпотези для подальшого моделювання. На цьому етапі виконано роботу зі складними структурами респондентських відповідей, які важко оцінити вручну. Деякі зв'язки є слабкими або прихованими, інші залежать від багатьох змінних одночасно.

Перший крок – аналіз зв'язків між числовими змінними. Було перевірено, чи наявні залежності між такими параметрами, як, наприклад, суб'єктивна оцінка платформи, тривалість використання та кількість функцій, з якими користувач мав досвід. У процесі кореляційного аналізу виявлено лише окремі помірні залежності: оцінка якості платформи була слабко пов'язаною з частотою навчання, а тривалість використання – із впевненістю у власному рівні. Загальний рівень кореляцій між числовими змінними був низьким, що є типовим для UX-даних, де прямі лінійні зв'язки, як правило, не виражені. З огляду на обмежену інформативність такого аналізу подальший етап було зосереджено на категоріальних змінних. Для цього застосовано коефіцієнт Крамера (*Cramér's V*), який дає змогу оцінити силу асоціації між парами номінативних змінних. Сила зв'язку подана числами за шкалою від 0 (відсутній зв'язок) до 1 (сильний зв'язок). Цей підхід вважається доцільним у разі, коли аналізу підлягають поведінкові установки, мотивації або переваги, властиві для UX-опитувань. Значущі зв'язки між категоріальними ознаками було візуалізовано за допомогою матриці значень *Cramér's V*, що зображена на рис. 1. Цей і наступні візуальні матеріали статті подані у вигляді знімків з екрана безпосередньо з робочого середовища дослідження, що дає змогу зберегти точність відображення елементів та уникнути спотворень.

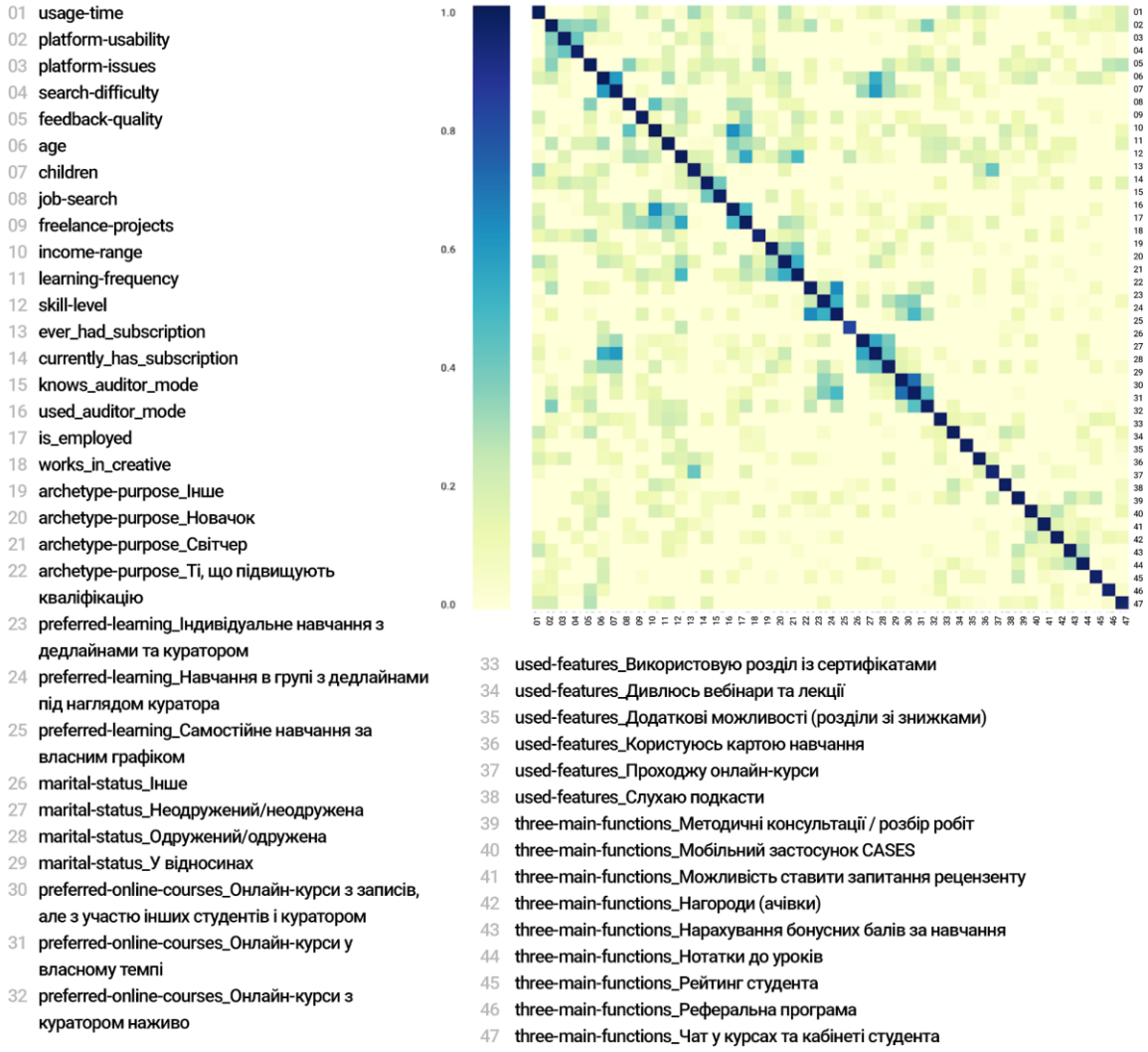

Рис. 1. Матриця Крамера для категоріальних ознак

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

На основі матриці виявлено кілька суттєвих залежностей, зокрема між частотою навчання й типом обраних курсів, архетипом і роллю користувача на платформі, а також між форматом зайнятості та участю у фриланс-проєктах. Отже, навіть в умовах обмеженої вибірки вдалося виявити значущі поведінкові шаблони, які не визначалися в межах числового аналізу й потребували категоріального підходу.

У фокусі аналізу була цільова змінна *currently_has_subscription*, яка відтворює, чи має користувач активну передплату на платформі. Це бінарна ознака, яка розглянута як така, що позначає загальну залученість користувача. Змістовно вона добре репрезентує основну мету – зрозуміти, що відрізняє платних користувачів від безплатних. На рис. 2 зображено частотний розподіл цієї змінної. Очікувано, що більшість користувачів не мають активної підписки. Такий дисбаланс вплине на метрики моделювання, тому його враховано надалі у виборі стратегії тренування.

Щоб зрозуміти потенційні сегменти користувачів без попередньої прив'язки до цільової змінної, здійснено кластеризацію методом *KMeans*. Завдання кластеризації – виділити архетипи на основі реальної поведінки й контекстів використання, не нав'язуючи готових категорій. Щоб визначити оптимальну кількість кластерів, застосовано два підходи: *Elbow*-метод та оцінку якості за *Silhouette Score*. Результати обох визначень подано на рис. 3.

Обидва графіки демонструють, що оптимальна кількість кластерів – два. Це означає, що в популяції користувачів справді є чіткий бінарний розподіл: умовно активні / пасивні або залучені / ознайомчі. Після побудови кластеризації та зниження розмірності (через PCA) отримано візуалізацію на рис. 4, де видно окремі хмари точок.

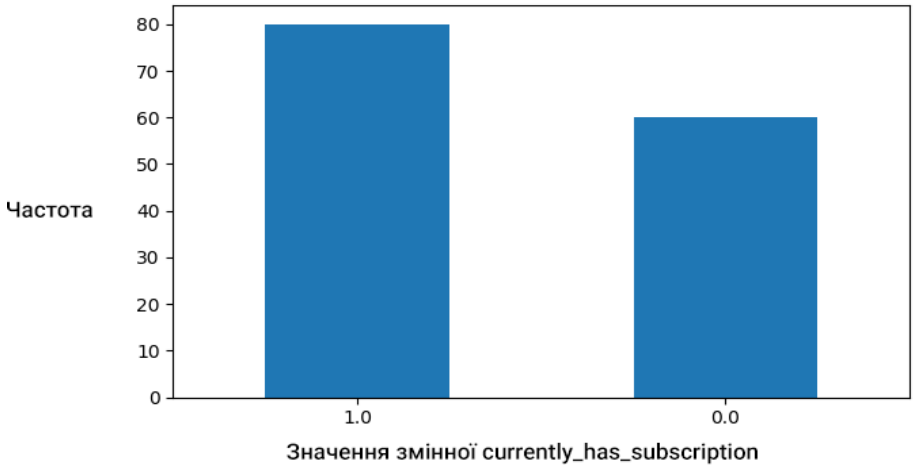

Рис. 2. Частотний розподіл активної підписки

Унаслідок кластеризації користувачів на основі числових ознак виявлено ключові розбіжності між двома групами. Найсильніше на формування кластерів вплинули змінні, пов'язані з користувацькою активністю, типом взаємодії з платформою та залученістю. Зокрема одна з найважливіших ознак досвіду наявності передплати: у першому кластері значно більше користувачів, які вже мали або мають активну підписку. Це вказує на більшу готовність платити за освітній сервіс. Подібним чином позначилась і змінна *currently_has_subscription*, яка ще більше закріплює кластер "залучених" користувачів. Також сильний вплив мало те, чи працює респондент у креативній індустрії. Ця змінна часто корелює з вищим інтересом до додаткових функцій платформи та самоосвіти. Водночас показовими стали ознаки використання конкретних функцій платформи, таких як режим вільного слухача *used_auditor_mode* та застосування сертифікатів, через що можна припустити, що активні користувачі глибше досліджують можливості та прагнуть підтвердження свого навчання. Висока частка користувачів із досвідом фриланс-проєктів також була притаманна кластеру з більшою залученістю. Нарешті, навіть суб'єктивна оцінка зручності платформи виявилася вищою в тій групі, де загальна активність і оплата були більшими, а отже, помітний причинно-наслідковий зв'язок між зручністю та конверсією.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

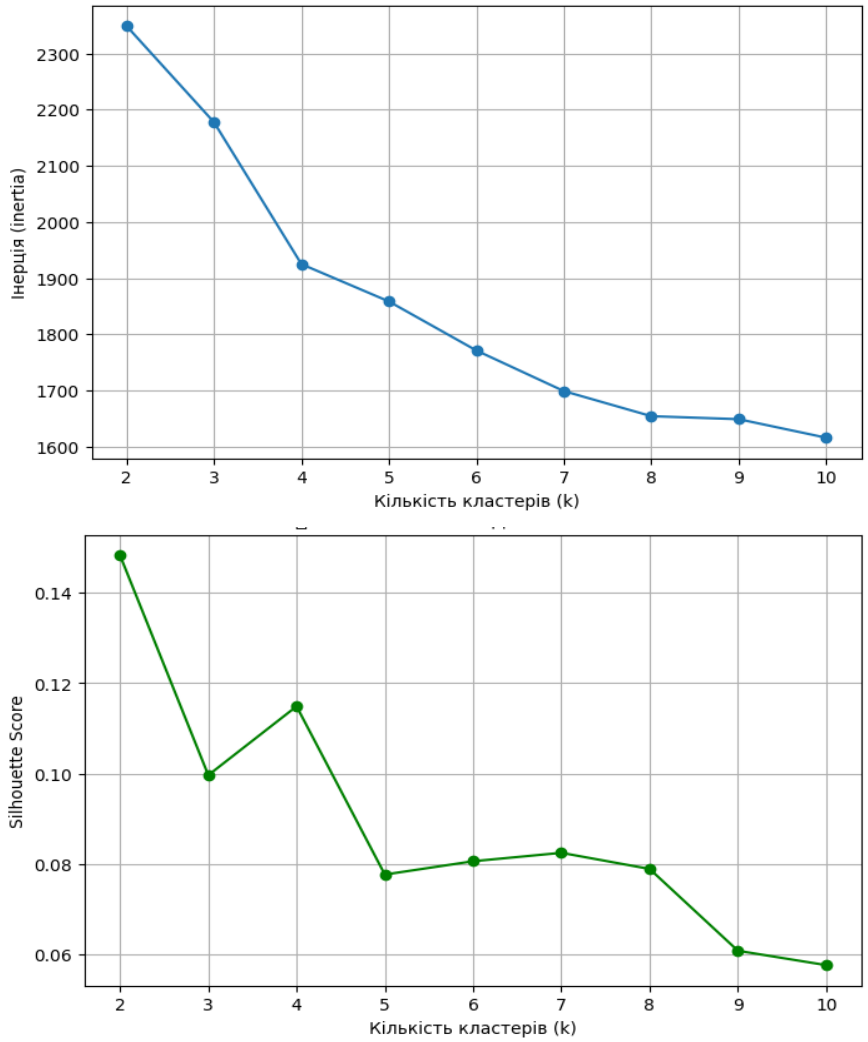

Рис. 3. *Elbow*-метод і *Silhouette Score* для вибору кількості кластерів

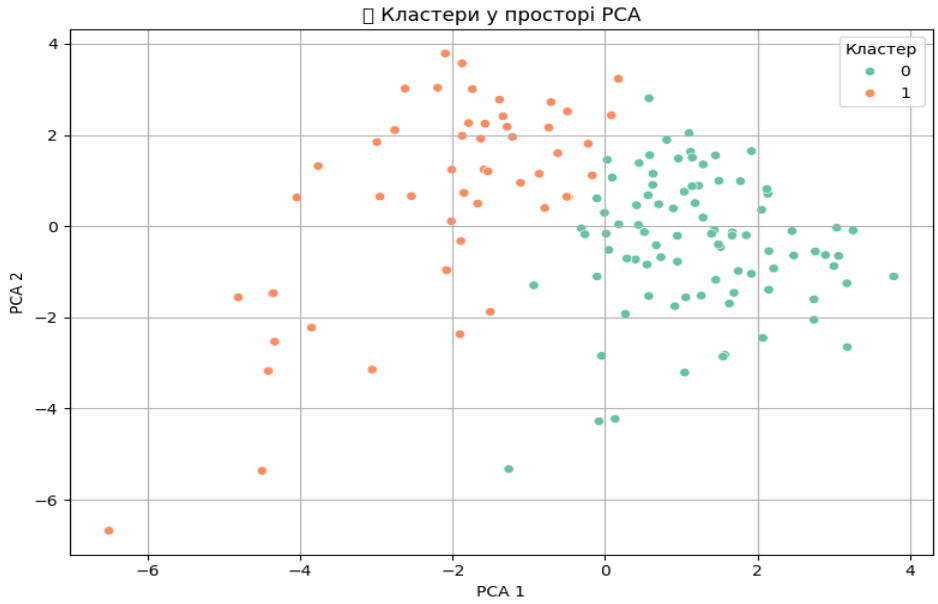

Рис. 4. Кластери користувачів у просторі PCA

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Така сегментація не лише виявляє латентні шаблони поведінки, а й задає напрями для подальшої персоналізації комунікацій і ціннісних пропозицій. Дослідницький аналіз показав, що навіть на невеликій вибірці можна виявити сильні залежності та розбіжності між користувачами. Комбінація кореляційного аналізу, методу Крамера та кластеризації потрібна для виявлення залежності між ознаками, які не завжди очевидні, та підтвердження релевантності вибору цільової змінної. Після цього стала можливою сегментація для персоналізованої аналітики.

Цей підхід може бути масштабований на будь-який інший освітній або сервісний продукт, який збирає змішані UX-дані – опитування, поведінкові записи, підписки. У наступному розділі розглянуто реалізацію моделей для автоматичного прогнозування залученості користувача.

Обґрунтування цільової змінної

У подальшому аналізі як цільову ознаку використано змінну *currently_has_subscription*, що позначає, чи має користувач активну підписку. Ця ознака є бінарною, однозначно інтерпретованою та має практичне значення: саме вона безпосередньо репрезентує один із ключових поведінкових результатів взаємодії. З погляду прикладного застосування, це критично важливий індикатор для прогнозування платоспроможності, готовності до купівлі або продовження передплати. Крім того, *currently_has_subscription* має збалансовану вибірку: у наборі достатньо прикладів як позитивного, так і негативного класу, через що можна уникнути типових проблем у навчанні моделей. Під час EDA також було підтверджено, що ця змінна пов’язана з низкою інших характеристик (зокрема оцінкою зручності, частотою навчання та використанням функцій), що додатково доводить її аналітичну цінність як результативного показника в нашому моделюванні.

Побудова моделей

У межах цього етапу дослідження побудовано й оцінено кілька моделей, мета яких – передбачення наявності активної підписки користувача (*currently_has_subscription*) на основі структурованих UX/UI-даних. Моделювання дає змогу виявити чинники, що впливають на рішення користувача оформити або продовжити передплату, а також описати потенційні способи їх використання на практиці, наприклад для побудови механізмів персоналізованих рекомендацій або підтримки в інтерфейсі. Моделі добирались з огляду на їх здатність працювати з табличними, гетерогенними UX/UI-даними. Для навчання й тестування дані були розподілені в пропорції 80/20. Точність моделей оцінювалась за чотирма метриками: *accuracy, precision, recall* і *F1-score*, – що комплексно аналізує поведінку класифікаторів в умовах помірного дисбалансу класів. Для їх порівняння сформовано консолідовану таблицю результатів (табл. 2), що містить основні показники алгоритмів.

Перша з моделей є дерева рішень. У цьому дослідженні використано дерево завглибшки до 6 з балансуванням класів для врахування нерівномірного розподілу ознаки-підписки. За результатами оцінювання дерево досягло точності 0.61, демонструючи помірну збалансованість між *recall* і *precision*, але слабшу загальну точність, якщо порівнювати з ансамблевими методами. Зі структури дерева було помітно, що найважливішими є критерії, які розподіляють сталі ознаки на кшталт *used-features*__Користуюсь картою навчання, *age, feedback-quality* тощо. На основі цього можна не лише прогнозувати, а й інтерпретувати логіку рішення для кожного випадку.

Інша ансамблева модель – "випадковий ліс" (*Random Forest*) – поєднує значну кількість дерев рішень і зменшує ризик перенавчання, притаманний окремим деревам. Завдяки використанню випадкового підбору ознак та агрегації результатів ця модель забезпечує високу стійкість до шуму й варіативності даних. У поточному дослідженні застосовано 100 дерев завглибшки до 8. За результатами тестування модель (за шкалою від 0 до 1) продемонструвала точність 0.88 та високі значення *F1-score* для обох класів. Вона також аналізує важливість ознак (рис. 5). Найбільш значущими стали *ever_had_subscription, used-features*__Користуюсь картою навчання, *income-range, usage-time, learning-frequency* – основні поведінкові та соціальні предиктори продовження підписки. Зокрема підписку частіше продовжують ті, хто вже мав її раніше, активно користується інструментами навчання (особливо мапами навчання), має вищий рівень доходу, знайомий більше з платформою та навчається регулярно.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

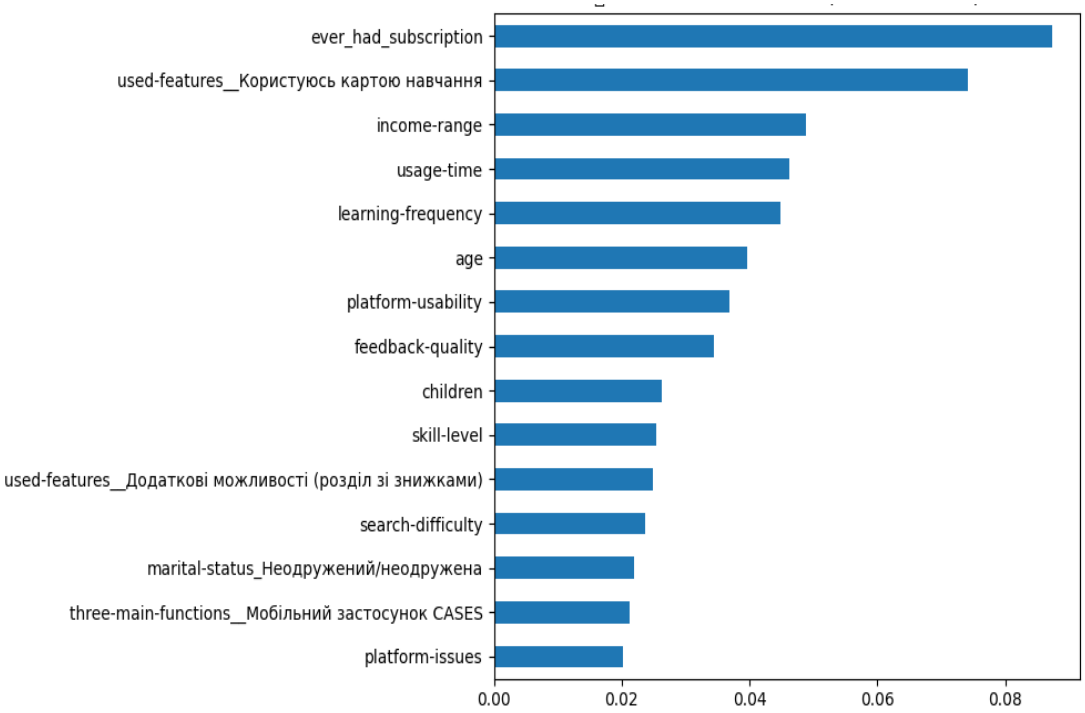

Рис. 5. Важливість ознак для *Random Forest*

Таблиця 2. *Порівняльна таблиця метрик моделей*

Модель	**Метрики**			
	Accuracy	**Precision (avg)**	**Recall (avg)**	**F1-score (avg)**
Decision Tree	0.61	0.61	0.61	0.6
Random Forest	0.88	0.86	0.85	0.86
XGBoost	0.87	0.85	0.3	0.84
MLP	0.89	0.89	0.88	0.87
kNN	0.61	0.62	0.6	0.59

На основі цього профілю можна сформувати цільові сегменти для маркетингової чи продуктової стратегії. Також ці ознаки допомагають ідентифікувати групи ризику, а саме користувачів, які з меншою ймовірністю продовжать підписку. Для них доцільно впроваджувати персоналізовані механіки утримання: тригери, нагадування, освітні рекомендації чи бонуси.

Застосовано ще один ансамблевий підхід XGBoost (*Extreme Gradient Boosting*). Через послідовність навчання дерев це підвищує точність у задачах з високою складністю структури ознак. У моделі використано 100 дерев завглибшки до 5 та *learning rate* 0.1. За результатами класифікації точність моделі становила 0.82, зі значеннями *F1-score* до 0.87, особливо добре працюючи з позитивним класом (наявність підписки). Також було сформовано графік важливості ознак, де лідирують *age, children, is_employed, learning-frequency*, що підтверджує перехресну значущість з *Random Forest*.

Крім того, створено базовий тип штучної нейронної мережі MLP (*Multilayer Perceptron*). Ця модель була обрана, бо здатна виявляти складні нелінійні залежності, що часто не доступні для простіших алгоритмів, зокрема дерев рішень. У нашому випадку модель містить два приховані шари по 128 і 64 нейрони відповідно, використовує функцію активації *ReLU* і регуляризацію через параметр *alpha* зі значенням 0.0005 для запобігання перенавчанню. Модель навчалась на тих самих розбитих даних, що й інші моделі. Результати показали найвищу точність серед усіх протестованих підходів 0.89 з досить симетричними значеннями *precision* та *recall*. Цей результат підтверджує ефективність MLP у задачах, де є багатофакторні зв’язки та помірна кількість ознак. Високі показники точності моделі свідчать про її здатність брати до уваги складні зв’язки, які не розпізнаються лінійними або ієрархічними моделями.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Ще один алгоритм – пам'яткова модель *k*-ближніх сусідів (kNN). Модель не вимагає етапу навчання, але сильно залежить від масштабу ознак. Саме тому перед класифікацією застосовано *StandardScaler*. Використано п'ять сусідів, але результати були найнижчими серед усіх моделей: точність 0.61, що вказує на слабшу здатність kNN до генералізації в умовах високої кількості ознак і невеликого обсягу вибірки. Така модель може бути корисною в інших контекстах, але в нашій задачі вона поступилася іншим підходам.

Для додаткової перевірки якості протестовано кілька моделей на випадковому прикладі з тестової вибірки. Моделі MLP, *Random Forest* та XGBoost класифікували один і той самий випадок. У цьому прикладі лише модель XGBoost давала неправильні відповіді, тоді як решта мали позитивний результат. Варто зазначити, що такий підхід не є заміною загальній оцінці, але перевіряє, наскільки узгоджені або суперечливі висновки моделей на індивідуальних рівнях.

Отримані результати підтверджують науковий засновок про те, що навіть прості, формалізовані UX/UI-дані з масових опитувань можуть слугувати надійним джерелом для побудови точних і стабільних моделей класифікації користувацької поведінки. У контексті сучасних досліджень, що здебільшого зосереджені на мультимедійних або складних структурованих даних, це становить помітну наукову новизну. Зокрема на тлі зростання зацікавленості в застосуванні нейромереж до відео, текстів і мультимодальних сигналів тема автоматизованого аналізу опитувальних таблиць залишається малоопрацьованою. Отже, це дослідження пропонує вагоме доповнення до наявної парадигми, доводячи, що високоякісні моделі можуть бути побудовані й на базі простих, дешево отримуваних UX-даних без необхідності в складних сенсорних системах або великих обсягах мультимедійної інформації. Це відкриває шлях до створення адаптивних, масштабованих і пояснюваних систем підтримки дизайну, здатних працювати в реальних прикладних умовах.

Висновки

Проведене дослідження підтверджує гіпотезу про те, що навіть у межах невеликої, частково заповненої, але структурованої UX/UI-вибірки можна досягти достатньої точності в задачах передбачення поведінкових сценаріїв користувачів. Аналіз на основі ознаки *currently_has_subscription* дав змогу виявити змінні, що корелюють із користувацькою залученістю та готовністю сплачувати підписку за контент. Цей підхід продемонстрував свою валідність і в технічному, і в практичному вимірах.

Моделі на основі ансамблевих алгоритмів (*Random Forest,* XGBoost) забезпечили найвищі метрики точності та збалансованості, водночас даючи змогу інтерпретувати важливість окремих ознак. До основних предикторів у цьому наборі даних належать: попередній досвід підписки, активне застосування навчальних функцій, рівень доходу, тривалість користування та частота навчання. Ці ознаки можуть лягти в основу сценаріїв для сегментації аудиторії та побудови систем утримання. Зокрема виявлені шаблони допомагають ідентифікувати як активних користувачів, так і групи ризику, до яких можуть бути застосовані персоналізовані стратегії втримання. Нейронна мережа (MLP) також показала високі результати: попри вищу чутливість до оброблення даних, вона підтвердила свою придатність до задач із комплексними ознаками. Водночас моделі на кшталт дерева рішень або kNN можуть застосовуватись у разі, коли пріоритетом є простота або інтерпретованість.

Наукова новизна полягає в демонстрації застосовності глибинних і ансамблевих моделей до табличних UX/UI-даних, отриманих із масових опитувань, а також у створенні відтворюваного підходу до роботи з UX/UI-опитуваннями, який можна масштабувати на інші домени. Зокрема методика може бути використана в освітніх платформах для виявлення ознак відтоку студентів, прогнозування інтересу до нових курсів або автоматичного формування персоналізованих рекомендацій на основі змін у поведінці. У сфері онлайн-медіа цей підхід дає змогу аналізувати й передбачати падіння залученості читачів до певних рубрик або форматів, а також сегментувати аудиторію за ймовірністю підписки. У сервісах підписного типу – допомагає вчасно ідентифікувати користувачів із ризиком відмови та запустити адаптивні механіки утримання. Отже, модель дає змогу не лише аналізувати наявні шаблони, але й проактивно впливати на розвиток користувацьких сценаріїв і поведінкову динаміку в автоматизованому режимі.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Описаний у статті підхід є лише частиною більш широкої дослідницької роботи, спрямованої на розроблення системного підходу до оброблення, аналізу та інтерпретації UX/UI-даних. На наступних етапах заплановано розширити класифікаційний підхід на мультимодальні джерела інформації (відкриті відповіді, неструктуровані природні дані, поведінкові патерни, RNN), побудувати рекомендаційні механізми та формалізувати систему як окремий модуль для UX-аналітики. Цей напрям буде розгорнуто в наступних статтях й дослідницьких напрацювань серії.

References

1. Farooqui T., Rana T., Jafari F. Impact of Human-Centered Design Process (HCDP) on Software Development Process. *Proceedings of the 2019 2nd International Conference on Communication, Computing and Digital Systems*. 2019. DOI: https://doi.org/10.1109/C-CODE.2019.8680978
2. Berni A., Borgianni Y., Basso D., Carbon C.-C. Fundamentals and issues of user experience in the process of designing consumer products. *Design Science*. 2023. DOI: https://doi.org/10.1017/dsj.2023.8.
3. Heil S., Gaedke M. Auto-Extraction and Integration of Metrics for Web User Interfaces. *Journal of Web Engineering*. 2018. DOI: https://doi.org/10.13052/jwe1540-9589.17676
4. Koch J., Oulasvirta A. Computational Layout Perception using Gestalt Laws. *CHI Extended Abstracts*. 2016. DOI: https://doi.org/10.1145/2851581.2892537
5. Wang S., Zhang R., Du J., Hao R., Hu J. A Deep Learning Approach to Interface Color Quality Assessment in HCI. 2025. DOI: https://doi.org/10.48550/arXiv.2502.09914
6. Stige Å., Zamani E.D., Mikalef P., Zhu Y. Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. *Information Technology & People*. 2024; 37(6):2324–2352. DOI: https://doi.org/10.1108/ITP-07-2022-0519
7. Rohrer C. When to Use Which User-Experience Research Methods. Nielsen Norman Group. URL: https://www.nngroup.com/articles/which-ux-research-methods (дата звернення: 27.04.2025).
8. Zhang Z., Chen H., Huang R., Zhu L., Ma S., Leifer L., Liu W. Automated Classification of User Needs for Beginner User Experience Designers: A Kano Model and Text Analysis Approach Using Deep Learning. *Information*. 2024. DOI: https://doi.org/10.3390/ai5010018
9. Perrig S.A.C., Aeschbach L.F., Scharowski N., von Felten N., Opwis K., Brühlmann F. Measurement practices in user experience (UX) research: a systematic quantitative literature review. *Frontiers in Computer Science*. 2024. DOI: https://doi.org/10.3389/fcomp.2024.1368860
10. Ігнатюк Є. О., Попов А. В. Методико-інструментальні засоби автоматизованого аналізу даних результатів UX-досліджень з використанням систем штучного інтелекту. *Системи управління, навігації та зв'язку*. 2024. DOI: https://doi.org/10.26906/SUNZ.2025.1.96-106
11. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y., Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*. Vol. 8, No. 4. P. 49–59. DOI: https://doi.org/10.20998/2522-9052.2024.4.07
12. Zender A., Humm B.G., Holzheuser A. Successfully Improving the User Experience of an Artificial Intelligence System. *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*. 2024. DOI: https://doi.org/10.15439/2024F2707
13. Duan P., Chen C.-Y., Li G., Hartmann B., Li Y. UICrit: Enhancing Automated Design Evaluation with a UI Critique Dataset. *Proceedings of the 2024 CHI Conference*. ACM. 2024. DOI: https://doi.org/10.1145/3654777.3676381
14. Čejka M., Masner J., Jarolímek J., Šimek P. UX and Machine Learning – Preprocessing of Audiovisual Data Using Computer Vision to Recognize UI Elements. *Agris on-line Papers in Economics and Informatics*. 2023; 15(3):35–44. DOI: https://doi.org/10.7160/aol.2023.150304
15. Batch A., Ji Y., Fan M., Zhao J., Elmqvist N. uxSense: Supporting User Experience Analysis with Visualization and Computer Vision. *IEEE*. 2023. DOI: https://doi.org/10.1109/TVCG.2023.3241581
16. Fan M., Yang X., Yu T.T., Liao Q.V., Zhao J. Human-AI Collaboration for UX Evaluation: Effects of Explanations and Synchronization. 2021. DOI: https://doi.org/10.48550/arXiv.2112.12387
17. Dou Q., Zheng X.S., Sun T., Heng P.-A. Webthetics: Quantifying webpage aesthetics with deep learning. *International Journal of Human-Computer Studies*. 2018. DOI: https://doi.org/10.1016/j.ijhcs.2018.11.006
18. Choudhury N. Can Artificial Intelligence be Used to Improve Productivity by Automating Elements of the User Experience Design Processes? 2023. DOI: https://doi.org/10.20944/preprints202202.0057.v2
19. Bezkorovainyi, V., Kolesnyk, L., Gopejenko, V., Kosenko, V. (2024), "The method of ranking effective project solutions in conditions of incomplete certainty", *Advanced Information Systems*, Vol. 8, No. 2, P. 27–38. DOI: https://doi.org/10.20998/2522-9052.2024.2.04

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

20. Nielsen J. Usability 101: Introduction to Usability. Nielsen Norman Group. 2012. URL: https://www.nngroup.com/articles/usability-101-introduction-to-usability (дата звернення: 27.04.2025).
21. Malyeyeva, O., Lytvynenko, D., Kosenko, V., Artiukh, R. Models of harmonization of interests and conflict resolution of project stakeholders. *Ceur Workshop Proceedings*, 2020. 2565. P. 24–35. URL: https://www.scopus.com/pages/publications/85082134474?origin=resultslist (дата звернення: 27.04.2025)
22. CASES. Соцмережа та EdTech для креативних індустрій. URL: https://cases.media (дата звернення: 11.05.2024).

Received (Надійшла) 01.02.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Ігнатюк Євгеній Олексійович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: y.o.ihnatiuk@khai.edu; ORCID ID: https://orcid.org/0009-0002-0568-829X

Попов Андрій В'ячеславович – кандидат технічних наук, доцент кафедри комп'ютерних наук та інформаційних технологій, Національний аерокосмічний університет "Харківський авіаційний інститут", Харків, Україна; e-mail: a.popov@khai.edu; ORCID ID: https://orcid.org/0000-0001-8984-731X; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=58558782100

Ihnatiuk Yevhenii – National Aerospace University "Kharkiv Aviation Institute", PhD Student, Department of Computer Science and Information Technologies, Kharkiv, Ukraine.

Popov Andrii – PhD (Engineering Sciences), Associate Professor, Department of Computer Science and Information Technologies, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

APPLICATION OF MACHINE LEARNING METHODS FOR ANALYSIS OF UX/UI DATA FROM MASS USER SURVEYS

The subject of this article is the application of machine learning methods to the interpretation of UX/UI data collected through mass user surveys on digital platforms. The paper explores the hypothesis that coordinated use of various c lassification models allows for the identification of behavioral patterns that hold predictive value for assessing users' interactions with product features. **The goal** is to perform a comparative analysis of classification accuracy using real-world UX/UI survey data. **The methodology** includes data preprocessing, feature encoding, normalization, clustering, and training of six model types: decision trees, random forest, gradient boosting, multilayer perceptron (MLP), logistic regression, and k-nearest neighbors. Particular attention is paid to how these models perform on small-scale, mixed-type UX/UI datasets. The modeling results demonstrate that even with limited data, it is possible to uncover significant relationships between socio-demographic variables, user types, and feature usage. **These findings suggest** that machine learning can be a promising approach for analyzing user behavior, with the potential for further integration into decision support systems. This approach can be adapted to various domains where structured user data is available, including online education, healthcare, public administration, urban services, and internal organizational platforms.

Keywords: UX/UI analytics; survey data; machine learning; behavioral scenarios; neural networks; ensemble models; digital services; user experience (UX).

Бібліографічні описи / Bibliographic descriptions

Ігнатюк Є. О., Попов А. В. Застосування методів машинного навчання для аналізу UX/UI-даних масових опитувань користувачів. *Сучасний стан наукових досліджень та технологій в промисловості.* 2025. № 4 (34). С. 44–57. DOI: https://doi.org/10.30837/2522-9818.2025.4.044

Ihnatiuk, Y., Popov, A. (2025), "Application of machine learning methods for analysis of UX/UI data from mass user surveys", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 44–57. DOI: https://doi.org/10.30837/2522-9818.2025.4.044