

РЕЦЕНЗІЯ

кандидата технічних наук, доцента кафедри інформаційних
та комп'ютерних технологій
Київського національного університету технологій та дизайну

Тетяни Іванівни ДЕМКІВСЬКОЇ

на дисертаційну роботу Пилипенка Владислава Ігоровича на тему
«Моделі та методи прогнозування успішності здобувачів освіти на основі
машинного навчання», представлену на здобуття наукового ступеня
доктора філософії (PhD) в галузі знань 12 – «Інформаційні технології», за
спеціальністю 122 – «Комп'ютерні науки»

Актуальність обраної теми дисертаційного дослідження

Сучасні заклади вищої освіти активно впроваджують системи управління навчанням (LMS), зокрема Moodle, які забезпечують організацію освітнього процесу, оцінювання, комунікацію та накопичення великих масивів освітніх даних. У таких умовах особливої ваги набувають задачі освітньої аналітики та прогнозування академічної успішності здобувачів освіти з метою раннього виявлення студентів групи ризику. Більшість існуючих підходів до прогнозування базуються переважно на традиційних освітніх показниках – оцінках, відвідуваності та результатах тестування. Водночас поведінкові дані про реальну взаємодію здобувачів із відеонавчальним контентом (тривалість перегляду, паузи, перемотування, повторні перегляди) використовуються недостатньо, що обмежує інформативність ознакового простору та точність моделей машинного навчання. Актуальність теми підсилюється розвитком

ансамблевих методів машинного навчання, поведінкової освітньої аналітики та потребою закладів освіти у практичних інструментах підтримки прийняття рішень безпосередньо в середовищі LMS. Обрана тема дисертаційного дослідження є своєчасною, науково обґрунтованою та має важливе прикладне значення для спеціальності 122 «Комп'ютерні науки».

Зв'язок дисертації з науковими програмами, планами, темами

Дисертаційна робота виконана відповідно до наукового напрямку кафедри комп'ютерної інженерії та електромеханіки Київського національного університету технологій та дизайну. Дослідження пов'язане з ініціативною темою «Дослідження Internet-технологій для побудови систем управління дистанційним навчанням» (державний реєстраційний номер 0122U200947, дата реєстрації 02.10.2022). Тема дисертації узгоджується із зазначеною науковою тематикою кафедри та спрямована на розвиток методів і засобів освітньої аналітики в електронних системах управління навчанням.

Ступінь обґрунтованості та достовірності наукових положень, висновків і рекомендацій

Основні наукові положення дисертації є достатньо обґрунтованими. Автор послідовно поєднав аналітичний огляд предметної області, порівняльне дослідження базових моделей машинного навчання, експериментальну оцінку впливу поведінкових відеоданих та розроблення дворівневої стекінгової моделі. Для розв'язання поставлених задач використано методи машинного навчання для класифікації та прогнозування, методи статистичного аналізу даних, підходи до аналізу великих освітніх даних, а також метрики якості класифікації (Accuracy, Precision, Recall/Sensitivity, Specificity, F1-score, ROC-AUC). Порівняльний аналіз моделей виконано на однаковому наборі даних із застосуванням єдиної стратегії формування навчальної та тестової вибірок, що підвищує об'єктивність висновків. Достовірність результатів підтверджується експериментальними

дослідженнями, програмною реалізацією плагіна VideoPlayer для LMS Moodle, засобами формування аналітичної звітності, апробацією на наукових конференціях (зокрема IEEE) та впровадженням у освітній процес закладів вищої освіти. Висновки дисертації логічно випливають із змісту розділів і є аргументованими.

Структура і зміст дисертаційної роботи. Відповідність дисертації встановленим вимогам

Дисертаційна робота має чітку й логічну структуру та складається зі вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. Повний обсяг становить 209 сторінок, з яких 163 сторінки основного тексту. Робота містить 48 рисунків, 27 таблиць, список літератури з 116 найменувань та 3 додатки. У вступі обґрунтовано актуальність теми, сформульовано мету й задачі, визначено об'єкт і предмет, методи дослідження, наукову новизну, практичне значення, особистий внесок, апробацію та публікації. У першому розділі виконано аналіз сучасних підходів до прогнозування академічної успішності в системах дистанційного навчання, особливостей накопичення даних у LMS Moodle, методів формування ознакового простору та застосування машинного навчання. Обґрунтовано потребу в інтеграції поведінкових відеоданих. У другому розділі проведено порівняльний аналіз моделей Logistic Regression, Naive Bayes, Support Vector Machines, Random Forest і Neural Network на традиційних освітніх ознаках. Показано переваги випадкового лісу та нейронних мереж і водночас обмеження щодо розпізнавання студентів групи ризику. У третьому розділі досліджено вплив поведінкових характеристик взаємодії з відеоматеріалами, сформовано розширений ознаковий простір і запропоновано дворівневу стекінгову модель із метамоделлю на основі градієнтного бустингу. Досягнуто показників Accuracy 90,2 %, Sensitivity 97,5 %, збалансованої точності 85 % та ROC-AUC 92,6 %. У четвертому розділі представлено програмну

реалізацію системи збору та аналізу поведінкової освітньої аналітики, інтеграцію плагіна відеоплеєра в Moodle і засоби візуалізації результатів прогнозування.

За структурою, змістом, обсягом і оформленням дисертація відповідає вимогам до робіт на здобуття ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» та вимогам Дисертація загальним обсягом 209 сторінок, з яких 163 сторінки основного тексту. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел (116 найменувань) та 3 додатків. Робота містить 48 рисунків, 27 таблиць та 3 додатки, у які винесені фрагменти програмного коду та два акти впровадження. Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

Повнота відображення наукових положень дисертаційної роботи в опублікованих працях

Основні результати дисертації достатньо повно відображено в наукових публікаціях автора. За матеріалами дослідження опубліковано 23 наукові праці, зокрема: 10 статей у наукових фахових періодичних виданнях України; 1 статтю у науковому періодичному виданні іншої держави, індексованому в Scopus (Q3); 12 публікацій у матеріалах міжнародних і всеукраїнських конференцій. Опубліковані праці висвітлюють ключові положення дисертації: порівняльний аналіз моделей прогнозування, розширення ознакового простору поведінковими відеоданими, розроблення плагіна для Moodle, побудову стекінгової моделі та результати експериментальних досліджень. Особистий внесок здобувача у працях зі співавторством визначено коректно. Кількість публікацій є достатньою для представлення дисертації до захисту. Наукові публікації відповідають вимогам п. 8, 9 Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової

установи про присудження ступеня доктора філософії (Постанова Кабінету Міністрів України від 12 січня 2022 р. № 44).

Наукова новизна одержаних результатів

Дисертаційна робота містить ряд сформульованих та обґрунтованих наукових положень, що мають теоретичну та практичну значущість. Отримані результати є вагомим внеском у розвиток методів освітньої аналітики та машинного навчання для прогнозування академічної успішності і характеризуються належним рівнем наукової новизни.

Найбільш вагомі наукові результати:

уперше:

- розроблено дворівневу стекінгову модель прогнозування академічної успішності здобувачів освіти, яка відрізняється використанням на першому рівні ансамблю алгоритмів логістичної регресії, випадкового лісу та нейронної мережі, а на другому рівні – алгоритму градієнтного бустингу, що забезпечує підвищення достовірності прогнозування;

удосконалено:

- метод розширення простору ознак для задач прогнозування академічної успішності за рахунок включення характеристик взаємодії здобувачів освіти з навчальними відеоматеріалами в LMS;
- методи збору та підготовки даних для задач прогнозування успішності, що передбачають автоматизоване отримання і накопичення даних про взаємодію користувачів з відеоматеріалами та їх інтеграцію у процес побудови моделей машинного навчання;

набули подальшого розвитку:

- методи аналізу освітніх даних щодо оцінювання впливу поведінкових характеристик роботи користувачів електронного освітнього середовища на результати прогнозування академічної успішності;
- підходи до побудови інтелектуальних сервісів освітньої аналітики в системах управління навчанням на основі ансамблевих моделей машинного навчання.

Наукова новизна дисертаційної роботи є переконливою, підтверджена результатами експериментальних досліджень і достатньо повно висвітлена у публікаціях автора.

Наукове та практичне значення результатів дослідження

Наукове значення роботи полягає у розвитку методів і моделей прогнозування академічної успішності на основі інтеграції освітніх і поведінкових відеоданих, а також у побудові ефективної ансамблевої стекінгової архітектури для задач класифікації в освітній аналітиці.

Практичне значення підтверджується:

- розробленням програмного модуля VideoPlayer для LMS Moodle;
- реалізацією засобів підготовки даних і формування розширеного набору ознак;
- створенням програмного забезпечення для візуалізації результатів прогнозування та формування аналітичних звітів у форматах PDF і XLS;
- підвищенням точності прогнозування порівняно з базовими моделями (зокрема на 14,3 % порівняно з логістичною регресією, на 4,56 % порівняно з нейронною мережею та на 2,29 % порівняно з випадковим лісом);
- впровадженням результатів у освітній процес Київського національного університету технологій та дизайну і Хмельницького національного університету (підтверджено актами впровадження).

Отримані результати можуть бути використані при створенні та модернізації інформаційних систем освітньої аналітики і систем підтримки прийняття рішень у закладах освіти.

Відсутність порушення академічної доброчесності

У дисертації не виявлено ознак академічного плагіату, самоплагіату, фабрикації, фальсифікації чи інших порушень, що могло б поставити під сумнів самостійний характер виконаної роботи. У дисертації є відповідні посилання на інформаційні джерела у тих випадках, де використано результати досліджень інших авторів. Проведений аналіз дисертаційної роботи та наукових публікацій свідчить, що всі положення дослідження викладено із дотриманням принципів академічної доброчесності. Дисертація Пилипенка Владислава Ігоровича повністю відповідає вимогам Наказу МОН України № 40 від 12.01.2017 «Про затвердження вимог до оформлення дисертації» (зі змінами від 31.05.2019 р.). За результатами аналізу представлених матеріалів не виявлено фактів порушень академічної доброчесності.

Зауваження та дискусійні положення щодо змісту дисертації

Поряд з охарактеризованими вище науковими здобутками та значимими прикладними розробками дисертація не позбавлена певних дискусійних моментів:

1. У роботі детально описано розширення ознакового простору поведінковими відеоданими, проте доцільно було б чіткіше формалізувати процедуру відбору найбільш інформативних ознак і оцінити їх відносну важливість у підсумковій моделі.
2. У порівняльному аналізі моделей варто детальніше описати процедуру налаштування гіперпараметрів і стратегію крос-валідації для підвищення відтворюваності експериментів.

3. Практична реалізація орієнтована переважно на LMS Moodle. Перспективним напрямом подальших досліджень є узагальнення архітектури збору поведінкових даних для інших систем управління навчанням.

4. Враховуючи високу чутливість моделі (97,5 %), доцільно додатково проаналізувати баланс хибнопозитивних і хибнонегативних результатів у контексті прийняття рішень щодо студентів групи ризику, а також оцінити стабільність моделі на даних різних дисциплін і семестрів.

Зазначені зауваження мають рекомендаційний характер і не знижують загальної позитивної оцінки дисертаційної роботи.

Загальні висновки

Дисертаційна робота Пилипенка Владислава Ігоровича на тему: «Моделі та методи прогнозування успішності здобувачів освіти на основі машинного навчання» є авторським, багатоаспектним, виконаним на актуальну тему і на високому теоретико-практичному рівні завершеним науковим дослідженням. У роботі вирішено важливе науково-прикладне завдання щодо підвищення точності прогнозування академічної успішності студентів на основі методів машинного навчання та даних взаємодії з відеонавчальним контентом у системах управління навчанням, зокрема шляхом розроблення дворівневої стекінгової моделі та програмних засобів збору поведінкової освітньої аналітики в LMS Moodle. Наукові результати дисертації є достатньо обґрунтованими, достовірними та характеризуються належним рівнем наукової новизни. Практична цінність роботи підтверджується програмною реалізацією, експериментальною апробацією, впровадженням результатів дослідження у освітній процес закладів вищої освіти та можливістю їх широкого використання у системах освітньої аналітики і підтримки прийняття рішень. Дисертаційна робота відповідає спеціальності 122 «Комп'ютерні науки». Та відповідає вимогам п. 6–9 «Порядку присудження ступеня доктора філософії та скасування

рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44, а її автор Пилипенко Владислав Ігорович заслуговує на присудження ступеня доктора філософії в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Рецензент:

Кандидат технічних наук, доцент

кафедри інформаційних та

комп'ютерних технологій

Київського національного університету

технологій та дизайну

Тетяна ДЕМКІВСЬКА

