УДК 004.9:378.147

ДОСЛІДЖЕННЯ ТА РОЗРОБКА МЕХАНІЗМІВ АНАЛІЗУ, ЗБОРУ ТА ОБРОБКИ ОСВІТНІХ ДАНИХ ПРИ ОНЛАЙН-НАВЧАННІ

Зовтун Т.І.[1]; Партас В.К.[2], к.т.н., доц.
[1,2]Національний університет кораблебудування імені адмірала Макарова
[1,2]Україна, Миколаїв
[1]tzovtun@gmail.com;[2]viktor.partas@nuos.edu.ua

***Анотація**. Виконано аналіз сучасних підходів до аналітики освітніх даних, розроблено концептуальну модель збору, обробки та візуалізації освітньої інформації. Проведено кластеризацію студентів за поведінковими характеристиками та створено прогностичну модель оцінювання успішності. Експериментальні результати показали, що використання розробленої системи дозволяє підвищити точність прогнозування успішності студентів, збільшити середній бал та зменшити частку незадовільних результатів.*

***Ключові слова:** онлайн-навчання; освітні дані; аналітика даних; машинне навчання; Moodle; Power BI; Learning Analytics; Data Mining; цифрова освіта.*

Вступна частина. У XXI столітті цифровізація освіти стала визначальним чинником модернізації освітніх процесів і трансформації підходів до навчання [1]. Масове впровадження систем дистанційного навчання (Learning Management Systems – LMS) дало змогу забезпечити безперервність освітнього процесу, створити нові формати взаємодії між викладачами та студентами й накопичити значні обсяги освітніх даних [2]. Утім, більшість таких даних залишається неструктурованою, що унеможливлює їхнє ефективне використання для прийняття педагогічних і управлінських рішень [3]. Це визначає потребу у створенні спеціалізованих механізмів збору, аналізу та інтерпретації освітніх даних, орієнтованих на принципи аналітики та наукового моделювання.

Метою роботи є підвищення ефективності онлайн-освіти шляхом розроблення та впровадження науково обґрунтованих механізмів збору, обробки та аналітичного використання освітніх даних у середовищах дистанційного навчання. Реалізація цієї мети потребує виконання наступних задач:

- Системний аналіз сучасних тенденцій цифровізації освіти;
- Класифікація основних типів освітніх даних, що генеруються у LMS;
- Розроблення методології інтеграції технологій Learning Analytics, Data Mining та Machine Learning у процес онлайн-навчання;
- Створення концептуальної моделі обробки даних та перевірити її ефективність на експериментальній вибірці.

Основна частина. У дослідженні використано системний та структурно-функціональний підходи, які дозволяють розглядати аналітику освітніх даних як комплекс взаємопов'язаних процесів: від збору та очищення інформації до її інтерпретації у вигляді аналітичних звітів. Для реалізації аналітичної моделі застосовано середовище Python (бібліотеки *Pandas*, *Scikit-learn*, *Matplotlib*), систему керування базами даних PostgreSQL та інструменти візуалізації Power BI [4]. Такий технологічний стек забезпечує повний цикл обробки освітніх даних – від первинного збору логів активності студентів у LMS до формування прогнозних моделей і інтерактивних інформаційних панелей.

Кожен запис у базі LMS подано у вигляді вектора подій:

$$x_i = \{t_i, a_i, r_i, s_i, o_i\},$$

де: t_i – час події;

a_i – тип дії;

r_i – результат;

s_i – студент;

o_i – об'єкт навчання.

Множина таких векторів формує освітній простір даних $D = \{x_1, x_2, \ldots, x_n\}$, у якому, за допомогою алгоритмів кластеризації (K-Means, DBSCAN), виявляються поведінкові патерни студентів, а методи класифікації (Random Forest, Gradient Boosting) [5] використовуються для прогнозування академічної успішності.

Розроблена модель дозволяє здійснювати аналітичну обробку даних за трьома рівнями:

1) описовим – аналіз поточної активності студентів;
2) діагностичним – виявлення причин академічних труднощів;
3) прогностичним – оцінювання майбутніх результатів навчання.

Отримані результати візуалізуються у Power BI, що забезпечує зрозуміле представлення аналітичних показників для викладачів та адміністраторів освітнього процесу.

Експериментальна перевірка моделі на основі реальних даних із LMS Moodle засвідчила, що використання запропонованих алгоритмів дозволяє підвищити точність прогнозування успішності студентів до 88–90 %, зменшити частку незадовільних результатів на 45 % і своєчасно виявляти студентів, які потребують додаткової підтримки. Отримані аналітичні залежності також виявили сильну кореляцію між рівнем активності студентів та їх підсумковими оцінками.

Розроблені механізми можуть бути інтегровані у системи LMS для реалізації підходу data-driven education [6], який ґрунтується на прийнятті педагогічних і управлінських рішень на основі об'єктивних даних. Це створює передумови для персоналізації освітніх траєкторій, підвищення мотивації до навчання та формування адаптивних освітніх екосистем.

Висновки. Запропонована модель аналітики освітніх даних є комплексним рішенням, що поєднує математичні методи, інформаційні технології та педагогічну аналітику. Її впровадження забезпечує якісно новий рівень управління освітнім процесом у закладах вищої освіти. Подальші дослідження спрямовані на використання методів штучного інтелекту для автоматичної адаптації контенту під індивідуальні особливості студентів та розроблення рекомендаційних систем у середовищі LMS.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Сидоренко, О.В., Гужва, В.М. (2024). Освітня аналітика та технології машинного навчання у системах дистанційного навчання. *Інформаційні технології та комп'ютерна інженерія,* 2(57), 45–54.

[2] Ferguson, R. (2023). Learning Analytics: Drivers, Developments and Challenges. *International Journal of Technology Enhanced Learning,* 2(15), 101–118. DOI: 10.1504/IJTEL.2023.128745.

[3] Siemens, G., Long, P. (2021) Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, 4 (56), 30–41.

[4] Power BI. Uncover powerful insights and turn them into impact. Retrieved October 20, 2025 from https://www.microsoft.com/en-us/power-platform/products/power-bi/

[5] Gradient Boosting vs Random Forest, (2024, April 09). Retrieved from https://www.geeksforgeeks.org/machine-learning/gradient-boosting-vs-random-forest/

[6] Data-Driven Education: Navigating the Compelling Debate, (2024, September 10). Retrieved from https://strobeleducation.com/blog/data-driven-education/

Zovtun T.I, Partas V.K.

Research and development of mechanisms for analysis, collection and processing of educational data in online learning.

Abstract. *An analysis of modern approaches to educational data analytics was performed, a conceptual model for collecting, processing and visualizing educational information was developed. Students were clustered according to behavioral characteristics and a predictive model for assessing success was created. Experimental results showed that the use of the developed system allows to increase the accuracy of predicting student success, increase the average score and reduce the share of unsatisfactory results*.

Keywords: *online learning; educational data; data analytics; machine learning; Moodle; Power BI; Learning Analytics; Data Mining; digital education.*