

ВИКОРИСТАННЯ ТЕХНОЛОГІЇ BIG DATA ДЛЯ РОЗВИТКУ ВИЩОЇ ОСВІТИ

Моторин Руслан Миколайович,

доктор економічних наук, професор,
Київський національний торговельно-економічний університет

Протягом тривалого періоду в системах освіти закладів вищої освіти (ЗВО) накопичується величезна кількість інформації про різні аспекти освітнього процесу: студентів, їх успішність і відвідуваність; викладачів, їх наукову, освітню, а також адміністративну діяльність; навчально-методичні матеріали (тексти лекцій, практичних занять, аудіо, відео записи та ін.) і т. д. Ці дані необхідно зберігати, обробляти і аналізувати для ефективного управління навчальним процесом. Для обробки великих архівів і великих потоків даних потрібні нові технології, серед яких особливе місце займають технології Big Data.

Термін Big Data (великі дані) стосується великих і складних наборів даних, які можуть бути структурованими або неструктурованими і займають дуже великий обсяг дискової пам'яті. Характеристики Big Data можуть бути описані за правилом «5V»:

- 1V (volume): обсяг фізичних даних значний. Великий обсяг даних означає інформацію про велику кількість студентів та навчальних закладів. Ці дані несуть інформацію, яка може бути використана для ефективного управління навчальним процесом.

- 2V (velocity): швидкість збирання даних і швидкість обробки результатів порівняно висока; наприклад, дані про оцінки за заняття вносяться не пізніше закінчення дня їх отримання; викладач після внесення даних майже відразу може ознайомитися з аналітикою успішності. Швидкість зміни великих даних дозволяє в інтерактивному режимі контролювати процес навчання і своєчасно реагувати на будь-які зміни навчального процесу. Використання інтерактивних тестів дозволяє викладачам виявити студентів, які дають неправильні відповіді на тестові запитання та в режимі реального часу дати їм необхідний матеріал для вивчення і кращого засвоєння навчального матеріалу.

- 3V (variety): варіативність алгоритмів обробки різних типів зібраних результатів; наприклад, результати виконання домашніх завдань студентами можуть бути представлені в розрізі статі, віку і т. д.

- 4V (veracity): висока вірогідність зібраних даних, що дозволяє формулювати репрезентативні результати; наприклад, після проведення національного дослідження якості освіти можна зробити висновок, що студенти 4 курсу мають значно вище оцінки, ніж студенти 1 курсу.

- 5V (value): цінність накопичуваних даних має полягати в можливості на їх основі формулювати корисні різноаспектні залежності системи освіти; наприклад, можна помітити, що при переході студентів з першого курсу на

другий кількість відмінників зменшується, при цьому спостерігається відповідна зміна структури оцінок, що може говорити про поетапне ускладнення навчального матеріалу.

Завдяки досягненням в області інформаційних технологій великі дані у сфері освіти тепер можна накопичувати й аналізувати. Для цього необхідно вивчати існуючі технології зберігання й обробки даних, що використовуються провідними країнами світу в освіті. Дослідження у вітчизняній і зарубіжній літературі з питання використання Big Data в системі освіти досить розрізнені.

Важливим аспектом дослідження Big Data є інфраструктура зібраних даних. Так, Ф. Алмейда Нето і А. Кастро, з огляду на онлайн платформи, на яких розміщуються освітні заходи, розробили модель, де дані, створені на основі взаємодії між користувачами і самою платформою, вибираються і зберігаються в локальних базах даних [1]. Потім локальні бази зводяться і групуються в глобальну базу. Окремими аспектами даного напрямку дослідження є питання, пов'язані з досягненням освітніх результатів. EDM (Educational Data Mining) описується як засіб підвищення ефективності електронного навчання. Так, М. Насирі, Б. Минаї, Ф. Вафа розробили модель для прогнозування академічної успішності студентів першого курсу [2].

Іншими питаннями в дослідженнях учених стають аспекти Big Data, пов'язані із взаємодією суб'єктів навчання. Г. Мобашер, А. Шавіш, О. Ібрахім описують структуру великої бази даних в освіті, яка серед іншого містить демографічні дані учнів, психологічні характеристики студентів [3]. У роботах В. Там зі співавторами описаний підхід до організації спільного навчання, що дозволяє виявляти освітні закономірності, засновані на різноманітному наборі освітніх онлайн-ресурсів [4]. С. Двіведі та В. Рошні на основі аналітичного підходу описують технологію відбору найбільш прийнятних курсів за вибором [5].

Великі дані в освіті дозволяють викладачам отримати різноманітну інформацію про рівень підготовки студентів, засвоєння навчальної інформації, виконаних контрольних завдань і лабораторних робіт. Іншою важливою проблемою освіти є питання виявлення нових, часом прихованих, взаємозв'язків у великих даних, нових знань. Для цього можна використовувати методи data mining з метою поліпшення організації освітнього процесу та підвищення ефективності його управління.

Ще одним важливим напрямком досліджень є питання, пов'язані з внутрішньою взаємодією. Прогнозування академічної успішності – одна з ключових тем досліджень у сфері Big Data в освіті. Оцінювання успішності є складним завданням, оскільки вона залежить від різних факторів. Взаємозв'язок між факторами успішності для прогнозування продуктивності навчання характеризується складними нелінійними зв'язками, тому напрями збирання даних повинні бути всеосяжними. Управління великими даними дає можливість обробки інформації для аналізу ключових показників навчальної ефективності.

Важливо також відзначити користь використання великих даних для адміністративного персоналу ЗВО. Успішність, відвідуваність, стипендії та інша персональна інформація про студентів підлягає постійному збиранні, обробці, аналізу. Робота з таким обсягом даних вимагає значних витрат праці. Автоматизація рутинної роботи призведе до заощадження фінансових і трудових ресурсів.

У сфері освіти виділяються п'ять основних типів даних:

- 1) персональні дані;
- 2) дані про взаємодію студентів з електронними системами навчання (електронними підручниками, онлайн курсами);
- 3) дані про ефективність навчальних матеріалів;
- 4) адміністративні дані;
- 5) прогнозні дані.

Для структуризації процесів управління Big Data в освіті може бути виділено п'ять взаємопов'язаних груп процесів:

- 1) цілепокладання: визначення мети і завдань дослідження;
- 2) планування: підбір джерел інформації, процедур отримання даних, алгоритмів обробки інформації;
- 3) зведення даних: організація зведення даних в єдину базу;
- 4) аналіз показників: аналіз отриманих даних, визначення способів представлення результатів;
- 5) коригування: розробка практичних заходів регулювання;

Сайт OnlineUniversities виділив десять напрямів, в яких буде змінюватися вища освіта під впливом великих даних.

1. Зміна методики роботи в групах: наприклад, на одному з курсів в Гарварді студентів з різними відповідями об'єднують в пари, щоб вони змогли прийти до єдиного рішення, відстоюючи свою позицію.

2. Досвід навчання стане більш особистісним: технології дозволяють індивідуально підбирати не тільки курси, а й також домашні завдання.

3. Студенти отримуватимуть більше рекомендацій: вже зараз програми вміють передбачати, наскільки успішно буде пройдений курс, ще до того, як він почався.

4. Дані відіграють важливу роль у виборі вузу: передбачається, що абітурієнтам навіть не доведеться подавати заяви, тому що роботи самі підберуть для них кращі місця.

5. Чи зміниться маркетинг: навчальні заклади зможуть заздалегідь дізнаватися про перспективних кандидатів.

6. Більше студентів будуть добиратися до кінця навчання: вже зараз технології допомагають виявити учнів, які перебувають в групі ризику, і допомогти їм.

7. Оптимізується управління ЗВО: заклади різного типу зможуть отримувати більш точні рекомендації.

8. Викладачі зможуть краще допомагати відстаючим студентам: програми дозволять дізнатися, в яких саме сферах є проблеми.

9. Буде простіше вибирати кар'єру: цифрові портфоліо розкажуть всю вашу історію за вас.

10. Аналіз даних стане ключовим елементом життя ЗВО: використовуючи аналіз даних на всіх рівнях, адміністрація зможе ефективніше приймати рішення.

Наведемо приклад процесу управління Big Data у вищій освіті на прикладі Київського національного торговельно-економічного університету (КНТЕУ).

1. Мета – експертно-аналітична оцінка кадрового потенціалу КНТЕУ.

Завдання:

- виявити вікові характеристики викладачів КНТЕУ;
- визначити, чи існують об'єктивні вакансії або має місце перенасиченість КНТЕУ викладачами;
- дати оцінку кваліфікації діючих викладачів.

2. Підбір джерел інформації, процедур отримання даних, алгоритмів обробки інформації.

Підбір джерел інформації:

- Додаток 1. Положення КНТЕУ «Рейтингова оцінка діяльності науково-педагогічних працівників університету»;
- річна форма статистичного спостереження 2-3нк «Звіт закладу вищої освіти на початок 2018–19 навчального року»;
- анкетна форма оцінки викладачів студентами.

Процедура отримання даних:

- відбір необхідної інформації, що міститься в Додатку 1. Положення КНТЕУ «Рейтингова оцінка діяльності науково-педагогічних працівників університету»;
- відбір необхідної інформації, що міститься у формі статистичного спостереження 2-3нк;
- проведення анкетування студентів за формою, розробленою КНТЕУ.

Алгоритм обробки даних: для аналізу існуючих кадрових проблем в розрізі навчального предмета анкета оцінки викладачів студентами містила детальну інформацію про викладачів. За результатами дослідження вікової структури побудуємо розподіл викладачів за віком порівняно з нормальним розподілом. Нормальний розподіл сприяє стабільному функціонуванню системи вищої освіти при дотриманні балансу між можливістю розвитку і збереженням традиції.

3. Збирання та структуризація даних.

Етап передбачав збирання і структуризацію даних, які надаються у КНТЕУ, а також експорт даних, зібраних за формою 2-3нк.

4. Аналіз отриманих даних, визначення способів подання результатів, фіксація закономірностей.

Порівняння розподілу викладачів за віком з нормальним розподілом у розрізі навчального предмета. Побудова аномальних графіків розподілу.

5. Розробка практичних заходів з регулювання процесів.

Після прийняття підсумкового звіту про негативні тенденції повідомляють управлінським органам КНТЕУ, відбувається затвердження та реалізація плану заходів щодо кадрової політики на 2018–2021 роки.

Методики з використанням великих даних дозволяють сформуванню взаємозв'язок між факторами успішності й оцінити потенціал і прогрес студента протягом усієї його навчальної історії. Подібний підхід може полегшити формування індивідуального освітнього маршруту з урахуванням особливостей кожного студента.

Список використаних джерел

1. De Almeida Neto F. A., Castro A. A reference architecture for educational data mining // Proceedings Frontiers in Education Conference, FIE. October, 2017. P. 1–8. doi: 10.1109/FIE.2017.8190728.

2. Nasiri M., Minaei B., Vafaei F. Predicting GPA and academic dismissal in LMS using educational data mining: A case mining // Proceedings of the 6th National and 3rd International Conference of E-Learning and E-Teaching, ICeLeT. 2012. P. 53–58. doi: 10.1109/ICELET.2012.6333365.

3. Mobasher G., Shawish A., Ibrahim O. Educational data mining rule based recommender systems // Proceedings of the 9th International Conference on Computer Supported Education. 2017. № 1. P. 292–299.

4. Enhancing educational data mining techniques on online educational resources with a semi-supervised learning approach / V. Tam et al. // Proceedings of 2015 IEEE International Conference on Teaching, Assessment and Learning for Engineering, TALE 2015. 2016. P. 203–206. doi: 10.1109/TALE.2015.7386044.

5. Dwivedi S., Roshni V. S. K. Recommender system for big data in education // Proceedings of the 5th National Conference on E-Learning & E-Learning Technologies (ELELTECH). 2017. doi: 10.1109/ELELTECH.2017.8074993.