

## 90. РАЗВИТИЕ ОБРАЗОВАТЕЛЬНЫХ СИСТЕМ НА ОСНОВЕ ТЕХНОЛОГИИ BIG DATA

*Бусько Н.П., студент гр. 272301, Полоско Е.И., аспирант,  
Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Ефремов А.А. – канд. экон. наук, доцент каф. ЭИ*

Рассматриваются особенности внедрения технологии Big Data в образовательные системы на основе найденных примеров. Уделяется внимание способам применения технологий, таких как сбор и классификация данных из большого количества источников, их анализ, а также прогнозирование осуществления процесса на будущее.

Образовательные системы, как отдельный комплекс социальных институтов, направлены на подготовку подрастающего поколения к самостоятельной жизни в современном социуме. При помощи интернета учреждения образования, организаторы онлайн курсов, ученики, преподаватели собирают и обмениваются информацией с чрезвычайно большой скоростью, а значит актуальность и целесообразность использования технологий хранения, обработки и реализации этой информации в данной сфере обоснована.

Хорошая организация и обработка всех этих данных позволяет не только повысить качество получения информации, но и сам процесс обучения. Использование технологий Big Data позволяет постоянно совершенствовать существующие подходы в образовании людей, а также создавать новые, более эффективные, предоставлять комфортные условия для самого процесса обучения, а также делать прогнозы на будущее.

Основных типов данных в сфере образования существует всего 5:

- личная информация;
- информация о взаимодействии студентов с электронными системами обучения (электронными учебниками, онлайн-ресурсами);
- данные об эффективности учебных материалов;

- административные (общесистемные) данные;
- прогнозные данные [1].

Характеристики Big Data во многих источниках описываются по правилу 5V:

1. volume: значительный объем физических данных;
2. velocity: регулярное обновление данных, сравнительно высокая скорость их сбора и скорость обработки результатов;
3. variety: вариативность алгоритмов обработки различных типов собранных результатов – текста, видео, изображения и т. д.
4. veracity: высокая достоверность собранных данных;
5. value: ценность накапливаемых данных, которая заключается в возможности на их основе формулировать полезные разноаспектные зависимости [2].

Каждое образовательное учреждение принимает и обучает большое количество учеников и преподавателей, и каждый год их количество пополняется новыми людьми с новыми данными, а значит и обработка этих данных требует постоянного совершенствования и автоматизации. В Беларуси идея использования больших данных была реализована единой образовательной платформой schools.by, которая предлагает бесплатный сервис электронных дневников и журналов. На данный момент платформа объединяет более 1,4 млн. пользователей почти из 4 тыс. учреждений образования страны, занимает 1-е место по посещаемости среди образовательных порталов и входит в ТОП-10 интернет-ресурсов Беларуси по версиям рейтингов Alexa Rank и Similarweb. Мобильное приложение Schools.by представлено на всех мобильных платформах, входит в ТОП-10 приложений в категории «Образование» в Беларуси и имеет высокие оценки от пользователей (4,4 на январь 2023-го года). Основная категория посетителей по возрасту — школьники до 18 лет, а также люди активного возраста от 25 до 34 лет [3].

Технология Big Data активно применяется в проведении централизованного экзамена (ранее тестирования), обязательного для поступления в ВУЗы Беларуси экзамена. Подсчет результата выполнения теста осуществляется в два этапа. Сперва определяется первичный балл, набранный участником тестирования. После определения первичных баллов каждому его возможному значению в автоматическом режиме с учетом статистических данных ставится в соответствие тестовый балл на нормализованной шкале, который и является итоговым результатом тестирования [4]. Все баллы автоматически заносятся в одну большую базу данных, в которой хранятся и в соответствии с которой выдаются сертификаты о прохождении теста. С данными сертификатами абитуриент может идти подавать документы в желаемый ВУЗ. Таким образом, обработка результатов тестирования и занесение данных в базу позволяет эффективно и справедливо оценить уровень знаний участников, а впоследствии осуществить прием на обучение в образовательное учреждение. Кроме того, данные о баллах поступивших в конкретный ВУЗ могут использоваться составления статистики поступления в различные учебные заведения на различные специальности, анализировать, где самые высокие и самые низкие баллы, в какие ВУЗы и на какие специальности подают документы в большем количестве, куда подают больше лиц мужского пола, а куда женского и т.д.

Для активного вовлечения студентов в процесс обучения и улучшения качества образования ВУЗы также прибегают к аналитике больших данных и методам машинного обучения. Университет Пердью в США разработал систему предиктивной аналитики, называемая Course Signal. Система помогает прогнозировать академические и поведенческие проблемы и оповещать как преподавателей, так и учеников о необходимости принятия мер. Course Signal комбинирует прогностическое моделирование с интеллектуальным анализом данных на Blackboard (Blackboard Learn – это приложение для интерактивного преподавания, обучения, создания сообществ и обмена знаниями).

Система использует различные источники данных, такие как информация о студентах и системы управления курсами. Уже на вторую неделю семестра такой инструмент по анализу данных способен понять академическую подготовку студента, вовлеченность и прилагаемые усилия в рамках курса (сессии, викторины, обсуждения, время необходимое для выполнения задания), а также учебную успеваемость в определенный момент времени. Алгоритм прогнозирует для каждого студента профиль риска на основе простой для понимания системы: зеленый (высокая вероятность успеха в конкретном курсе), желтый (есть потенциальные проблемы) и риск (неуспеха). Тот факт, что этот прогноз предоставляется уже на второй неделе семестра, дает студентам широкие возможности для улучшения своих результатов. Signal Course предоставляют учителям обратную связь, когда они запускают программное обеспечение, позволяя учителям мгновенно реагировать на студентов, когда могут возникнуть проблемы. Учителя могут запускать программу так часто, как они хотят, а обновление прогноза происходит только при ее запуске. Система моментально предоставляет учащимся различные ресурсы, которые могут помочь им совершенствоваться в курсе. Профиль риска может быть скорректирован для каждого курса в зависимости от того, что учителя сочтут нужным [5].

Однако, большие данные в образовании - это не только анализ того, как люди взаимодействуют с программным обеспечением, чтобы улучшить способы обучения. Исследовательские материалы,

миллионы медиа-материалов и текстовые файлы по определенным темам также ежедневно перемещаются в облачное хранилище. Человечество пополняет большие данные информацией, относящейся к конкретной предметной области, с помощью первоклассных программных решений. В качестве примера можно взять астрономию. Программные решения, такие как решение для дистанционного наблюдения за космосом для астрономов США, позволяют исследовать космос, сохраняя уникальные данные о космических исследованиях. Они предназначены для обмена наблюдениями и обучения у профессионалов. Геймификация - это еще одна дополнительная деталь, которая делает решение для дистанционного наблюдения за космосом гораздо более полезным и интерактивным.

В будущем астрономы и астрофизики могли бы анализировать собранные и сохраненные данные, чтобы понять природу небесных явлений и совершать открытия в этой области. В то время как преподаватели могут оценить, как люди относятся к геймификации во время освоения космоса и что можно улучшить, чтобы сделать изучение астрономии особенным и увлекательным для людей по всему миру.

Система образования регулярно накапливает огромные объемы данных, поэтому вопрос систематической обработки этих данных является одним из самых актуальных на сегодняшний день. Большие данные в образовании позволяют переосмыслить подходы, устранить давние пробелы и адаптировать учебный опыт для повышения эффективности всей системы. Анализ данных о качестве образования помогает участникам процесса выбрать профессию, наиболее соответствующую их интересам и личностным качествам, а анализируя данные о том, какие профессии пользуются наибольшим спросом на рынке труда, можно сделать выводы о том, какие специальности наиболее перспективны для молодых людей. Актуальной также остается задача описания технологии работы с объемными массивами данных, направленная на выявление сформировавшихся закономерностей в обучении и дальнейшее моделирование развития системы.

**Список использованных источников:**

1. *Development of educational systems based on Big Data technology* [Электронный ресурс]. – Режим доступа: <https://e-koncept.ru/en/2018/181039.htm>
2. *Технологии Big Data: ключевые характеристики, особенности и преимущества* [Электронный ресурс]. – Режим доступа: <https://aicongference.com.ua/ru/news/tehnologii-big-data-klyuchevie-harakteristiki-osobennosti-i-preimushchestva-97883>
3. *Реклама* [Электронный ресурс]. – Режим доступа: <https://schools.by/comm>.
4. *Результаты тестирования* [Электронный ресурс]. – Режим доступа: <https://rikc.by/pretrialtesting/553-rezultaty-rt.html>
5. *Purdue University Achieves Remarkable Results With Big Data* [Электронный ресурс]. – Режим доступа: <https://datafloq.com/read/purdue-university-achieves-remarkable-results-data/>