

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ BIG DATA В ДИСТАНЦИОННОМ ОБУЧЕНИИ

Хоронеко М.П., Медунецкий М.А., Прахоцкий М.Г.

Белорусский государственный университет информатики и радиоэлектроники, г Минск, Беларусь, maxim.horoneko@gmail.com, mmed.log@gmail.com, prahotsky96@gmail.com

Abstract. Educational system continuously accumulates vast amounts of data which significantly increase need to develop and change approaches used to store and distribute knowledge to the students. Big Data technologies can be powerful tool for identifying common education system patterns and use it to implement those changes quickly and effectively.

Благодаря развитию информационных технологий, образовательная система давно вышла за рамки классической модели передачи знаний от учителя к ученику. Стремительно набирает популярность электронное образование, растет потребность в технологиях, которые могли бы точно определить необходимые меры по изменению подходов к обучению для более эффективного и персонализированного доступа к информации. В этой среде существует большое количество структурированных и неструктурированных типов данных, каждый из которых представлен в больших объемах, таким образом использование традиционных методов обработки информации становится неэффективным. Одним из возможных решений этой проблемы может стать использование технологий Big Data.

В рамках дистанционного обучения существует возможность накапливать, хранить и обрабатывать большое количество информации о студентах, используя ее для построения эффективного учебного процесса. Так как дистанционное обучение предполагает активное взаимодействие учащихся и преподавателей при помощи информационных технологий, внедрение механизмов для сбора такой информации не является проблемой.

Скорость работы современных алгоритмов Big Data позволяет в реальном времени контролировать процесс обучения. Например, использование онлайн тестов позволяет установить проблемные области в знаниях студентов и скорректировать как общий, так и персональный учебный план.

Еще одной проблемой, которую позволяют решить Big Data алгоритмы, является определение новых взаимосвязей в больших данных или Data Mining. Использование таких технологий в области дистанционного обучения позволит определить предметы и темы, вызывающие у студентов наибольшие затруднения, с какими задачами они справляются лучше всего, какую форму проведения занятий они предпочитают, и как можно оптимально построить учебную программу, чтобы студент приобрел именно те навыки, которые будут полезны в его будущей профессии. Можно выделить следующие основные задачи такого подхода.

Классификация – отнесение объектов к одному из заранее заданных классов [1]. К примерам использования таких задач в дистанционном обучении можно отнести:

– классификацию электронных текстовых ресурсов (по характеру информации, форме изложения, назначению);

– классификацию тестовых заданий (по сложности, тематике, временным затратам).

Регрессия – моделирование влияния, которое оказывают друг на друга параметры при изменении, определение зависимости выходных значений от входных параметров. Решение задач регрессии может помочь спрогнозировать результаты тестирования, возможности трудоустройства выпускника после окончания обучения, уровень заработной платы. Это в свою очередь позволит установить влияние таких факторов как уровень кадрового и технического обеспечения учреждения образования, востребованность специалистов на рынке труда, заработная плата преподавателей на учебный процесс.

Кластеризация – разделение множества объектов на кластеры, основываясь на схожести ряда признаков и свойств [2] (Рис. 1). Решение задачи кластеризации может позволить обнаруживать наличие студентов со схожими поведенческими и интеллектуальными особенностями. Руководствуясь этими данными, можно объединять учащихся в группы, наиболее эффективно разрабатывать учебные планы для них.

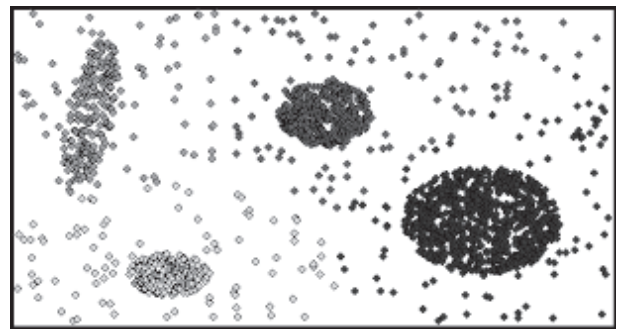


Рисунок 1 – Визуализация кластеров объектов

Для обработки и персонализации больших объемов данных в среде электронного образования, в том числе дистанционного обучения, предлагается использовать технологии MapReduce, Hadoop, NoSQL и другие. Эти технологии могут быть применены и к электронному образованию, включая вопросы принятия решений обработки данных, финансовое планирование и мониторинг успеваемости учащихся.

Литература

1. Дюк В.А., Самойленко А.П. Data Mining: учебный курс – 2001.
2. Профессиональный информационный ресурс, посвященный машинному обучению. [Электронный ресурс]. – Режим доступа: <https://clck.ru/FmWDj>.