– Вам знакомо понятие «интеллектуальный анализ данных»? Знаете, что такое «метакогнитивные навыки»? А «тяжелый режим» на Android? Нет? А ведь это главные ингредиенты для рецепта, который позволит предсказывать будущее образования. Кроме шуток, гадание на кофейной гуще тут ни при чем – данные и ничего кроме. Если хотите узнать больше, тогда Райан Бэйкер тот, кто вам нужен.

Он преподает в Высшей школе образования Пенсильванского университета и возглавляет Пенсильванский центр учебной аналитики, а, кроме того, основал Международное общество интеллектуального анализа образовательных данных. Исследует основы образовательного процесса.

Для нас большая честь, что он согласился рассказать обо всём этом подробнее. Слово Райану Бэйкеру.

– Спасибо. Спасибо, что пригласили. Сегодня я затрону тему возможностей интеллектуального анализа данных в образовании.

Итак, 2020 год пока что получается довольно необычным. Образование немного изменилось. До 2020 года уже наблюдался взрывной рост объема данных об обучающихся и обучении.

**Текст на слайде:**

**До 2020-го**

**Уже наблюдался взрывной рост объема данных об обучающихся и обучении**

**Так как для множества учащихся всех возрастов обучение перешло в онлайн, количество доступных данных существенно увеличивается**

И, так как для множества учащихся всех возрастов обучение перешло в онлайн, количество доступных данных существенно увеличивается.

**Текст на слайде:**

**Это создает возможности…**

**Использовать эти данные (не затрагивая персональные данные учащихся)**

**Повысить качество учебных материалов**

**Улучшить академический эффект и жизненные перспективы**

Всё это создает возможность использовать эти данные (не затрагивая персональные данные учащихся), чтобы повысить качество учебных материалов, улучшить академический эффект и жизненные перспективы.

**Текст на слайде:**

**Интерактивные учебные платформы**

Этот взрывной рост возможностей случился, отчасти, благодаря широкому использованию интерактивных учебных платформ – систем, которые позволяют учиться с большей интерактивностью, чем при обычной работе с учебником. Это может быть моделирование экспериментов, как на картинке в центре (в данном случае ученик воспроизводит опыт Галилея со сбрасыванием шаров с Пизанской башни). Или игровая платформа, вроде физической песочницы внизу слева. В этой игре ученик, чтобы решить головоломку, рисует объекты, которые, потом материализуются, и таким образом изучает действие разных видов сил. Или, как, например, на нижней правой картинке, система, которая встраивает математические задачки в художественный рассказ.

**Текст на слайде:**

**Зарегистрированная информация о студентах**

Такие системы создают обширнейшие объемы зарегистрированной информации, которые отражают все те действия, что ученик выполняет во время обучения.

**Текст на слайде:**

**Широкий масштаб данных**

**Данные, зарегистрированные в системах, которыми пользуются сотни тысяч или даже миллионы студентов**

**- Aleks, Matthias, Snappit, Edx, Coursera**

И масштаб этих данных постоянно растет, ведь в течение года системами вроде Aleks, Matthias, Snappit, Edx, Coursera пользуются сотни тысяч, а то и миллионы студентов.

**Текст на слайде:**

**Данные об образовании были:**

**Разрозненными**

**Их было трудно собирать**

**Несущественными**

Всё это меняет принцип существования данных об образовании. Раньше они были разрозненными, несущественными и их было трудно собирать.

**Текст на слайде:**

**Данные сегодня**

**Текст на слайде:**

**Даташоп PSLC (Кедингер и др., 2008, 2010)**

**>250 000 часов, проведенных учениками на образовательных платформах**

**>30 миллионов действий учащихся, ответов и примечаний**

А сегодня существуют площадки вроде даташопа Питтсбургского центра изучения образования, где хранится информация о более чем 250000 часов, проведенных учениками на образовательных платформах. Это более 30 миллионов действий учащихся, ответов образовательных программ и любых значимых примечаний, например, о соответствующем навыке, которые сделали эти программы.

**Текст на слайде:**

**Репозиторий MORF (Андрес и др., 2018; Гарднер и др., 2018)**

**Защищенное хранилище данных, где хранится информация МООК Пенсильванского, Мичиганского, Эдинбургского университетов**

**Позволяет исследовать весь объем данных без прямого доступа к ним, персональные данные студентов защищены**

**Дает возможность повторить (или не повторить) эксперимент, позволяющий предсказать, кто из студентов бросит МООК**

Другой пример – репозиторий MORF. Это защищенное хранилище данных, где хранится информация массивных открытых онлайн-курсов Пенсильванского и Мичиганского университетов, а также Эдинбургского университета в Шотландии. MORF позволяет исследовать весь объем этих данных без прямого доступа к ним. То есть, ученый может пользоваться этими данными, но не видит их, персональные данные студентов защищены. Это дает возможность, например, повторить (а иногда и не повторить) эксперимент, позволяющий предсказать, кто из студентов, проходящих МООК, бросит его, не закончив.

**Текст на слайде:**

**Международное общество интеллектуального анализа образовательных данных**

**«Замер, сбор, анализ и распространение данных об учащихся и их окружении в целях понимания и оптимизации учебного процесса и платформ, на которых он ведется»**

Распространение такого рода данных привело к созданию двух научных обществ: Международного общества интеллектуального анализа образовательных данных и Общества исследования учебной аналитики, работа которых сосредоточена на замере, сборе, анализе и распространении таких данных.

**Текст на слайде:**

**Цели**

**Общая цель – исследовать «большие данные» об обучающихся и обучении**

**Это позволит**

**- Совершать новые научные открытия и продвигать науку об образовании**

**- Лучше оценивать учащихся по нескольким направлениям**

**- Социальному, когнитивному, эмоциональному, метакогнитивному**

**- Индивидуальному, групповому, институциональному**

**- Улучшать поддержку учащихся в режиме реального времени лучше, делая обучение по-настоящему адаптивным**

Два этих общества объединены одной целью – исследовать, в кавычках, большие данные об обучающихся и обучении, которые теперь доступны, чтобы совершать новые научные открытия и двигать вперед науку об образовании, лучше оценивать учащихся по нескольким направлениям, включая социальное, когнитивное, эмоциональное, метакогнитивное (я буквально через минуту объясню, что это такое), а также на разных уровнях, исследовать не только отдельных учащихся, но и группы, и даже институты, в которых они учатся. Это делает поддержку учащихся в режиме реального времени лучше, обучение становится по-настоящему адаптивным.

**Текст на слайде:**

**Некоторые ключевые методы (Бэйкер и Сименс, 2014; на осн. Бэйкер и Ясеф, 2009)**

Возможно, стоит упомянуть несколько ключевых методов.

**Текст на слайде:**

**Прогнозирование**

**Разработка модели, которая может вывести один аспект данных (прогнозируемую переменную) из комбинации других аспектов данных (заданной переменной)**

**Кому из студентов скучно?**

**Кто завалит курс?**

**Кто бросит учебу в средней школе?**

**Узнаем что-то важное, чтобы исправить ситуацию**

Один из них – прогнозное моделирование, его еще называют предиктивная аналитика. Мы разрабатываем модель, которая может вывести один аспект данных, мы называем его прогнозируемой переменной, из заданных переменных, то есть комбинаций других аспектов данных. Так, например, система может спрогнозировать или понять, кому из студентов скучно, или кто, скорее всего, завалит курс, или кто может бросить учебу в средней школе. Цель – узнать что-то важное, чтобы исправить ситуацию.

**Текст на слайде:**

**Структурный анализ**

**Находим в полученных данных структуру и паттерны, которые складываются «естественным образом»**

**Нет конкретных целевых показателей или заданной переменной**

**Какие задачи требуют одних и тем же навыков для выполнения?**

**Есть ли группы студентов, которые подойдут к программе по-своему?**

**Какие именно студенты наладят больше социальных связей с помощью дискуссионных форумов?**

Еще один крупный категорный метод – структурный анализ. Здесь мы уже знаем, что хотим предсказать. Мы пытаемся найти в полученных данных структуру и паттерны, которые складываются, так сказать, естественным образом. У нас нет конкретных целевых показателей, но делаем мы примерно вот что. Допустим, у нас есть большая база математических задач, собранная тысячами учителей. Какие задачи требуют одних и тем же навыков для выполнения? Или, к примеру, у нас большая группа студентов, которые учатся по одной программе. Есть ли в этой группе те, кто подойдут к программе по-своему? И, быть может, достигнут необычного результата? Или, например, мы изучаем большие дискуссионные форумы, вроде тех, что используются в массовых онлайн-курсах. Какие именно студенты наладят больше социальных связей?

**Текст на слайде:**

**Интеллектуальный анализ связей**

**Обнаруживает связи между переменными в массиве данных с большим количеством переменных**

**Существуют ли более эффективные траектории освоения учебных программ (наборов курсов, образовательных объектов и т. д.)?**

**Какая компоновка образовательной системы привлечет больше студентов?**

Третья важнейшая категория – интеллектуальный анализ связей. Здесь исследователи или специалисты пытаются обнаружить связи между переменными в массиве данных, который содержит множество переменных. К примеру, если у вас есть набор курсов или образовательных объектов, то существуют ли более эффективные траектории их освоения? Или, если у нас есть много образовательного контента, мы можем с уверенностью сказать, какая именно компоновка сделает его более или менее привлекательным для студентов.

**Текст на слайде:**

**Множество способов применения**

**Прогноз успеха/неудачи**

**Автоматическое отслеживание успеваемости, вовлеченности, эмоций, стратегий – для лучшей адаптивности**

**Улучшение отчетов для преподавателей и других заинтересованных лиц**

**Фундаментальные исследования образования**

Эта технология может применяться множеством способов. В том числе, для прогнозирования того, кто, скорее всего, бросит учебу или кто, вероятно, добьется успеха. Или для отслеживания успеваемости студентов в режиме реального времени, или их вовлеченности, эмоций, которые они испытывают, или стратегий, которые применяют. Это повышает адаптивность. Еще одно практическое применение это создание улучшенных отчетов для преподавателей и других заинтересованных лиц об успехах студентов. И, наконец, для фундаментальных исследований образования. Мы пытаемся выяснить фундаментальные основы того, как люди учатся.