

# НОВАЯ НОРМА ОБРАЗОВАНИЯ



Обработка и анализ больших массивов данных стали по-настоящему революционным прорывом, серьезно повлиявшим на многие отрасли экономики, в том числе и на образование. Появление многочисленных обучающих информационных систем и технологий позволило этой сфере получать огромный объем информации, касающейся процессов обучения, поведения и успеваемости, и, соответственно, интерпретировать и анализировать ее. Возникло множество видов учебной аналитики: предиктивная, поведенческая, визуальная, диагностическая, описательная, операционная, с помощью которых можно не только автоматизировать многие рутинные процедуры, но и определять существующие проблемы, а также своевременно их упреждать.

## Парадигма Data Driven

Аналитика данных считается одним из ключевых факторов успеха деятельности, причем наметился переход от традиционной описательной ее формы к новому подходу Data Driven Decision Making – аналитике для принятия решений, или информационно обоснованному решению. Стоит отметить, что и до внедрения цифровых технологий компании всегда накапливали различные сведения

в виде финансовой, административной, бухгалтерской и другой отчетности, с тем чтобы на их основе совершенствовать свою работу. В сфере образования это были анкетирование и опросы, которые чаще всего становились вспомогательным инструментом для принятия решений, основанных на субъективных суждениях преподавателей, методистов, заказчиков обучения и т.д. Как правило, эта информация отличалась фрагментарностью и ограниченно-

стью, поэтому об эффективном реформировании образовательной системы на ее базе говорить не приходилось. Концепция Data Driven исправила этот пробел, позволив непрерывно агрегировать данные о результатах обучения, поведении слушателей, их взаимодействии с контентом и на этой основе строить и проверять гипотезы, оптимизировать процессы, а впоследствии автоматизировать принятие решений. Существует определенная иерархия уровней анализа данных, на которых решаются различные задачи (рисунки).

На уровне оптимизации происходит трансформационный переход от предиктивной аналитики к предписывающей – тому ее виду, который позволяет компании совершать действия, полностью основанные на данных.

В целом субъекты хозяйствования, взявшие на вооружение парадигму Data Driven, руководствуются несколькими базовыми принципами, как то: тестирование гипотез, эксперименты; режим непрерывного улучшения, итерационной оптимизации; предсказательное моделирование; взвешенный выбор между альтернативными вариантами на основе значимых в конкретной ситуации переменных [2]. По мнению директора по аналитике американской компании Warby Parker Карла Андерсона, в организации во главу угла должны быть поставлены не столько определенные инструменты, ресурсы и инфраструктура, сколько формирование особой культуры, направленной на убеждение сотрудников в необходимости информированных решений на основе данных. К тому же следует интегрировать

в процессы деятельности фирмы цифровые персонализированные системы, использовать данные при расчете тактических и долгосрочных задач, разработать стратегию управления и удобную систему для взаимодействия с ними. Как отмечают многие авторитетные эксперты, истинные изменения начинаются с людей и процессов, а не с технологий. Поэтому концептуальной основой Data Driven выступает коммуникация между всеми участниками, заинтересованными в поиске закономерностей и принятии верного решения, где каждый осознает свою ценность и вклад в общий успех [1].

С точки зрения лидеров бизнеса, использование этого подхода в образовании позволяет направить силу аналитики для генерации и проверки прорывных идей и предпринимать действия, соответствующие стратегическим целям организации. Все игроки на этом поле получают значительные преимущества: ответствен-

ные за обучение – платформу для расстановки приоритетов, принятия решения и эффективного распределения ресурсов; методисты – базу для оптимизации программы; обучающиеся – наиболее эффективный образовательный опыт.

## Разновидности и выбор метрик в образовании

Объективным индикатором успешности подхода Data Driven служат метрики – качественные или количественные показатели, отражающие характеристики изучаемого параметра. По их изменениям проверяется состоятельность гипотез. Поскольку современные платформы позволяют собирать большое количество данных о действиях пользователей, арсенал учебной аналитики поистине безграничен. Обычно изменения в обучающей программе проводятся для повышения уровня образования (помогает ли материал получить новые навыки, сменить специальность и т.д., соот-

ветствует ли он уровню подготовки обучающегося); для оптимизации учебного кейса (как достигаются учебные цели, какие разделы следует доработать); для повышения эффективности организации процесса (окупаются ли ресурсы, вложенные в программу).

Для того чтобы выбрать необходимые источники данных и показатели, изменения которых будут оцениваться, нужно установить цели оптимизации обучающей программы. Их достижение соотносится с определенными метриками, которые подразделяются на несколько категорий.

Метрики вовлеченности и активности слушателей отражают параметры взаимодействия обучающихся с образовательным контентом и представляют собой описательный уровень анализа: статистику и логи сайта/платформы (IP, местоположение); количество зачисленных слушателей, выполненных заданий, длительность просмотров видеороликов и сессий. С помощью этой метрики можно получить «цифровой эскиз» образовательного продукта, составить о нем первичное представление и ответить на вопросы: какая аудитория собралась для обучения, в каком контексте оно происходит и т.д.

Метрики эффективности обучения служат для оценки качества учебного материала, квалификации преподавателей и уровня подготовки слушателей. В свою очередь эти метрики подразделяются на следующие подкасты:

- индекс лояльности слушателей (NPS), который рассчитывается как разница между долями сторонников



Рисунок. Уровни аналитики данных и интеллектуального анализа

и критиков курса в результате опроса обучающихся;

- индекс удовлетворенности слушателей (CSI), дающий представление о наиболее сильных и слабых элементах программы;
- уровень прохождения (COR), демонстрирующий процент завершивших образовательный процесс от числа зачисленных;
- показатель оттока (CR), отражающий количество слушателей, отказавшихся от обучения в течение определенного периода времени.

С помощью упомянутых метрик дается оценка отдельных компонентов программы, ее наиболее трудных элементов и заинтересованности слушателей в окончании курса.

Метрики сопровождения процесса обучения очень важны для оптимизации программы и измеряют процент трудоустроенных после завершения обучения, скорость и качество обратной связи по проектам, оценку активности в учебных чатах. На их базе оценивается эффективность взаимодействия между слушателями и командой образовательной программы, возможность новых форматов коммуникации.

Метрики эффективности образовательного продукта предназначены для многосторонней его оценки и включают в себя многие упомянутые индикаторы, призванные ответить на весьма важные вопросы: какова реакция слушателей на образовательный продукт; получили ли они планируемые знания и навыки; применяют ли они их в жизни и работе; насколько полезно обучение для компании и др.

## Сбор и интерпретация данных

Стратегия Data Driven способна совместить цели обучающей программы с эффективностью бизнеса и запросами слушателей, оценить состояние образовательной системы, запланировать желаемые результаты и необходимые изменения. Для генерации и проверки гипотез следует подобрать соответствующие метрики и определиться с источниками информации. Для ее оценки широко используются два метода: учебная аналитика и интеллектуальный анализ образовательных данных (EDM). Первый заключается в измерении и сборе сведений об обучающихся и образовательной среде для понимания особенностей процесса и максимальной его оптимизации; второй представляет собой спектр методов исследования данных и поиска в них скрытых закономерностей, используемых с целью принятия решений в данной сфере. Оба этих метода способствуют выявлению взаимосвязей в массиве информации и скрытых знаний, которые при грамотном подходе помогают усовершенствовать обучение. Основное различие между образовательной аналитикой и интеллектуальным анализом образовательных данных состоит в том, что в первом случае полезные находки генерирует специалист, а во втором многие инновационные решения являются порождением автоматизированных алгоритмов. Можно выделить несколько категорий данных, которые обычно используются для учебной аналитики (табл. 1).

### Административные данные

Информация о преподавателе, методисте, наличии и опциях поддержки, опыте автора, тематике программы или курса и т.д.

### Предпочитаемые обучающие медиа или жанры

Ретроспективные показатели предпочтительных для обучающегося медиа или жанров в тех случаях, когда выбор был возможен, например длительность видео, дослушиваемость подкаста, дочитываемость лонгрида и т.д.

### Взаимодействие с образовательными ресурсами

Показатели взаимодействия во время обучения, в том числе манера навигации, ответы на упражнения и тесты, количество попыток, типы совершаемых ошибок, временные характеристики, связанные с деятельностью студента во время обучающих мероприятий.

### Прошлая деятельность

Ретроспективные показатели прошлой деятельности обучающегося, раскрывающие усвоение идей, навыков или компетенций на текущий момент.

### Временная история

Показатели ближайшего контекста, представляющие временную историю действий обучающегося, данные о которой доступны в конкретный день.

### Социальные показатели

Показатель взаимодействия учащегося с другими обучающимися и преподавателем в процессе обучения или с записанной речью (со всеми ее разнообразными свойствами, например семантическим содержанием, просодией и т.д.).

### Демографическая информация

Показатели периферийного контекста: регион, возраст, пол, уровень подготовки и т.д.

### Социальные связи

Показатели ближайшего окружения: количество связей, их сила и активность.

### Тип мышления

Данные из анкеты или самоотчета о том, как учащийся устанавливает связь между своими стратегическими усилиями во время обучения и развитием компетенций, а также о том, как происходит индивидуальный процесс обучения.

### Эмоциональное состояние

Психофизиологические показатели, относящиеся к обучению, например эмоциональное состояние, качество сна, показатели питания.

Таблица 1. Категории данных

Метод	Цель	Применение
Интеллектуальный анализ причинно-следственных связей	Выявить причинно-следственные связи	Понять, какие черты поведения слушателей приводят к обучению, неуспеваемости, отказу от обучения и т.д.
Кластеризация	Выявить группы с похожими свойствами	Объединить учебный материал или слушателей в группе, основываясь на найденных паттернах
Поиск и выявление взаимосвязей при помощи моделей машинного обучения	Использовать существующую модель в качестве компонента для текущего анализа	Выявить взаимосвязи между характеристиками слушателей или контекстуальными переменными и их поведением. Внедрить психометрические фреймворки в модель машинного обучения
Дистилляция данных (очистка данных для оценки человеком)	Представить данные в виде, пригодном для визуализации и взаимодействия с ними	Предоставить методистам данные, по которым удобно отслеживать и анализировать активность слушателей и их продвижение по программе
Отслеживание знаний	Оценить овладение знаниями при помощи логов ответов слушателей на задания и когнитивной модели, в которой построены взаимосвязи между заданиями и навыками, которые требуются для их решения	Отследить динамику прогресса
Неотрицательное матричное разложение	На основе оценок слушателей за тестовые задания создать матрицу положительных чисел, которая затем декомпозируется на две матрицы с заданиями и навыками для упрощения анализа	Оценить навыки слушателей
Выявление выбросов	Выявить нетипичных слушателей	Определить, какие слушатели не успевают за темпом программы или, наоборот, значительно опережают ее
Методы прогнозирования	Сделать умозаключение о целевой переменной исходя из комбинаций других переменных при помощи классификации, регрессии, оценки плотности и других методов	Спрогнозировать прогресс слушателя и определить его модель поведения
Интеллектуальный анализ процессов	Определить процесс на основе логов событий	Построить умозаключения о поведении студентов
Рекомендации	Предсказать, насколько предпочтительным для слушателя будет являться тот или иной компонент	Дать рекомендации студентам на основе их предыдущей зарегистрированной активности в программе
Интеллектуальный анализ взаимосвязей	Изучить взаимосвязи между переменными и выявить закономерности при помощи методов поиска ассоциативных правил, корреляционного анализа, интеллектуального анализа причинно-следственных связей	Выявить взаимосвязи между поведенческими паттернами слушателей и трудностями, которые они испытывают при обучении
Статистические методы	Описать происходящее и сделать выводы при помощи математической статистики	Проанализировать, проинтерпретировать и сделать выводы из данных учебной аналитики
Анализ социальных сетей	Проанализировать социальные взаимосвязи между элементами сети	Изучить структуру и взаимосвязи в групповой деятельности, а также взаимодействие слушателей в учебном чате
Интеллектуальный анализ текста	Извлечение данных из текстовой информации	Проанализировать содержимое чатов, домашних заданий и другого текстового цифрового следа слушателей
Визуализация	Отобразить данные в графическом виде	Представить данные в визуальном понятном виде

Таблица 2. Методы анализа данных в образовании

Следует отметить, что приведенный в таблице перечень не исчерпывающий, поскольку список источников данных адаптируется под цели и потребности конкретной организации. В современном образовании задействовано много технологий и платформ. Это система управления обучением (LMS), в частности Moodle, Canvas, iSpring, open EdX и т.д.; платформа опыта обучения, или агрегаторы образовательных проектов (LXP) – Degreed, EdCast; система управления корпоративным обучением (TMS), например ТренингСпейс; сервисы видеоконференций – Zoom, Skype, MS Teams; платформы проведения вебинаров; сервисы для проведения опросов; курсы, созданные через конструкторы, VR- и AR-платформы и др.

Для преобразования данных, полученных из них, в полезную информацию и принятия на ее основе наиболее оптимальных решений относительно образовательного процесса существуют определенные методы (табл. 2).

Таким образом, новый Data Driven подход дает возможность не только создать портрет учащегося, но и разработать точные модели поведения, выявить его слабые и сильные стороны, установить закономерности, характеризующие процесс обучения, и получить инсайты, способные значительно улучшить образовательную среду. ■

#### СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Аналитика обучения: о чем молчат данные // [https://sberuniversity.ru/upload/iblock/1f9/Edu\\_Tech\\_46\\_web\\_demo.pdf](https://sberuniversity.ru/upload/iblock/1f9/Edu_Tech_46_web_demo.pdf).
2. Creating a Data-Driven Organization // <https://oracle-patches.com/libr/book/is/creating-a-data-driven-organization>.
3. Educational data mining and learning analytics: An updated survey // <https://bookdown.org/chen/la-manual/files/Romero%20and%20Ventura%20-%202020.pdf>.

Ирина ЕМЕЛЬЯНОВИЧ