

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«Тольяттинский государственный университет»
Кафедра «Прикладная математика и информатика»

09.04.03 Прикладная информатика

(код и наименование направления подготовки)

Управление корпоративными информационными процессами

(направленность (профиль))

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)**

на тему: «Разработка и оптимизация системы образовательной аналитики в
высших учебных заведениях»

Обучающийся

А.В. Склярова

(Инициалы Фамилия)

(личная подпись)

Научный

руководитель

кандидат экономических наук, доцент, Т.А. Раченко

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Тольятти 2024

Оглавление

Введение.....	4
1 Теоретические аспекты применения аналитики и машинного обучения в сфере образования.....	11
1.1 Понятие образовательной аналитики: цели, задачи, применение	11
1.2 Анализ существующих исследований в области образовательной аналитики.....	14
1.3 Применение машинного обучения для анализа данных в образовательных системах.....	16
1.4 Роль аналитики в повышении качества образования в учебных заведениях.....	19
2 Анализ существующих информационных систем в образовательных учреждениях	22
2.1 Существующие процессы сбора, хранения и анализа образовательных данных.....	22
2.2 Проблемы и ограничения существующих систем образовательной аналитики.....	27
2.3 Проектирование новой архитектуры системы сбора и анализа данных в образовательном учреждении.....	29
2.4 Применение технологии машинного обучения для прогнозной аналитики.....	35
2.5 Разработка новых методов визуализации данных для предоставления информации	37
3 Разработка системы образовательной аналитики для повышения качества образования.....	40
3.1 Выбор и обоснование технологического стека для разработки системы	40
3.2 Разработка архитектуры системы	43

3.3 Реализация системы образовательной аналитики	48
3.4 Внедрение машинного обучения в систему для анализа данных	58
3.5 Тестирование и отладка системы	64
4 Оценка результатов исследования и экспериментальная апробация системы образовательной аналитики	69
4.1 Экспериментальная апробация системы	69
4.2 Результаты экспериментальной апробации	71
4.3 Оценка возможных улучшений и предложений.....	78
4.4 Перспективы дальнейших исследований	80
Заключение	83
Список используемой литературы	87
Приложение А Шаблон анкеты для оценки эффективности визуализации данных в системе образовательной аналитики.....	95
Приложение Б Шаблон анкеты для оценки эффективности визуализации данных в системе образовательной аналитики.....	96

Введение

Актуальность темы исследования. В условиях инновационной образовательной среды все процессы направлены на повышение образовательной активности студентов, развитие проектной деятельности и формирование индивидуальных образовательных траекторий. В этом процессе участвуют не только студенты, но и преподаватели, тьюторы, руководители образовательных программ, а также администрация. Каждая из этих групп нуждается в четких показателях эффективности и механизмах обратной связи. Например, преподаватели нуждаются в мониторинге активности студентов и качества образовательных ресурсов, чтобы студенты могли достичь требуемых результатов. Таким образом, для администрации и преподавателей образовательных учреждений важны показатели активности, успеваемости, мобильности и удовлетворенности студентов. Одним из способов получения таких данных являются системы образовательной аналитики.

Системы образовательной аналитики являются очень эффективным инструментом, позволяющим педагогам и администрации образовательного учреждения получать информацию о студентах и преподавании, а также использовать ее для улучшения эффективности учебного процесса. Вместо того чтобы опираться на интуитивные предположения и ожидания, профессора, исследователи и администраторы могут использовать данные для того, чтобы выявить точки роста у студентов, выявить наиболее успешные методы преподавания и прогнозировать результаты для того, чтобы помочь студентам улучшить свои академические результаты.

Кроме того, образовательные учреждения сталкиваются с запросами и отчетами оценки качества образования от государственных и образовательных органов. С использованием системы образовательной аналитики образовательные учреждения могут эффективно отслеживать и управлять

своей академической программой и демонстрировать свое качество образования через доступность и удобство использования этих систем.

Таким образом, с развитием технологий, объем данных в образовании наращивается, что значит рост доступности к данным о студентах, профессорах, курсах и программах и, в свою очередь, эффективное использование этих данных может повысить качество образования.

Системы образовательной аналитики могут помочь учебным заведениям в следующих аспектах:

- оценка эффективности учебных программ и курсов – системы образовательной аналитики могут помочь учебным заведениям оценить эффективность учебных программ и курсов, выявить проблемные моменты и улучшить качество обучения;
- поддержка принятия решений – системы образовательной аналитики могут помочь учебным заведениям принимать обоснованные решения на основе данных обучения, такие как изменение учебных программ, перераспределение ресурсов и т.д.;
- поддержка студентов – системы образовательной аналитики могут помочь студентам в их учебном процессе, предоставляя персонализированные рекомендации и подсказки, основанные на их индивидуальных данных обучения;
- улучшение управления учебным заведением – системы образовательной аналитики могут помочь учебным заведениям улучшить управление, предоставляя данные для принятия решений, оптимизации ресурсов и улучшения процессов.

Таким образом, разработка и исследование эффективности системы образовательной аналитики для повышения качества обучения в высшем заведении является актуальной темой, которая может привести к улучшению качества обучения и управления учебным заведением.

Однако существует проблема разработки таких программных решений и сервисов в образовании, связанная с тем, что данные об образовании

собираются разными цифровыми системами и не используются для создания интегральных решений. Электронные образовательные платформы также не имеют задач интеграции данных. Это создает сложности для педагогов и администрации в обязательной аналитике отчетности в образовании.

Проблема исследования: в современном образовании наблюдается необходимость в повышении качества обучения и эффективности учебного процесса. Однако существующие методы и инструменты анализа образовательных данных не всегда позволяют достичь желаемых результатов. В связи с этим возникает необходимость в разработке новой системы образовательной аналитики, которая бы позволила повысить качество обучения и эффективность учебного процесса.

Гипотеза: внедрение системы образовательной аналитики, основанной на автоматизации и интеграции данных, позволит повысить эффективность образовательного процесса за счет улучшения прогнозирования успеваемости и персонализации обучения.

Цель диссертации: разработка системы образовательной аналитики и исследование ее эффективности для повышения качества образования в высших учебных заведениях.

Задачи исследования:

- изучить теоретические аспекты применения аналитики и машинного обучения в сфере образования;
- проанализировать существующие исследования в области образовательной аналитики;
- исследовать текущие процессы сбора, хранения и анализа образовательных данных;
- выявить проблемы и ограничения существующих систем образовательной аналитики;
- разработать новую архитектуру системы сбора и анализа данных в образовательном учреждении;

- внедрить технологии машинного обучения для прогнозной аналитики;
- создать новые методы визуализации данных для предоставления информации;
- реализовать и протестировать систему образовательной аналитики;
- провести экспериментальную апробацию системы и оценить результаты.

Объект исследования: процессы управления качеством образования в образовательных учреждениях с использованием систем образовательной аналитики.

Предмет исследования: методы и инструменты образовательной аналитики, применяемые для сбора, обработки и анализа данных с целью повышения качества обучения и управления учебным процессом.

Теоретико-методологическая основа исследования базируется на трудах российских и зарубежных авторов, включая Ахмеджанову З., Вейцмана В.М., Гафурова П., Шадрина В.Г., Яковлева А.Л., исследующих образовательную аналитику, машинное обучение и цифровизацию образовательных процессов.

Методы исследования включили:

- анализ литературных источников – изучение научных статей и монографий в сфере образовательной аналитики и машинного обучения, что позволило определить современные подходы и ключевые потребности в образовательной аналитике;
- моделирование и визуализация бизнес-процессов – создание модели текущих и прогнозируемых процессов управления образовательным процессом, включая структуру данных и потоки информации для оптимизации аналитической системы;
- разработка и тестирование прототипа – создание и испытание прототипа системы образовательной аналитики с применением

технологий машинного обучения для оценки эффективности алгоритмов и удобства интерфейса для пользователей.

Опытно-экспериментальная база исследования: СПб ГБ ПОУ «Колледж электроники и приборостроения».

Научная новизна исследования заключается в создании и апробации системы образовательной аналитики с алгоритмами машинного обучения, ориентированной на персонализацию обучения и повышение эффективности управления образовательными процессами, что ранее не применялось в данном формате.

Теоретическая значимость исследования состоит в том, что предложенная модель образовательной аналитики углубляет знания о цифровизации образования и применении машинного обучения для адаптивного обучения и управления. Разработанные методы анализа образовательных данных позволяют создавать модели для управления учебными процессами и прогнозирования успеваемости студентов, что расширяет научные представления о персонализированном образовании.

Практическая значимость исследования заключается в том, что предложенная система образовательной аналитики может быть непосредственно использована учебными заведениями для повышения качества образовательного процесса и управления.

Внедрение системы на базе технологий машинного обучения позволяет:

- автоматизировать анализ данных об успеваемости и удовлетворенности студентов, что улучшает планирование и ресурсообеспечение учебных программ;
- предложить персонализированные рекомендации для студентов, что поддерживает их академические успехи.

Полученные результаты могут быть полезны и применимы не только в конкретном учебном заведении, но и в других образовательных учреждениях, заинтересованных в эффективном использовании образовательных данных и совершенствовании учебного процесса.

Достоверность и обоснованность результатов исследования обеспечивались:

- использованием современных методов аналитики и машинного обучения;
- проведением экспериментальной апробации системы образовательной аналитики;
- проверкой гипотезы на основе реальных данных в образовательном учреждении;
- применением проверенных методологических подходов.

Личное участие автора в организации и проведении исследования включало разработку системы, реализацию функциональных прототипов и проведение тестирования.

Апробация и внедрение результатов работы. Результаты исследования были апробированы в процессе функционального тестирования разработанной системы образовательной аналитики на базе СПб ГБ ПОУ «Колледж электроники и приборостроения».

Проведено тестирование работоспособности системы, а также оценка её эффективности на практике. Экспериментальные данные продемонстрировали улучшение в прогнозировании успеваемости и поддержке студентов через персонализированные рекомендации.

Результаты работы представлены на СХСVIII Студенческой международной научно-практической конференции «Научное сообщество студентов» и опубликованы в специализированных изданиях, что подтвердило их научную значимость и соответствие научным стандартам.

На защиту выносятся следующие положения:

- разработанная модель системы образовательной аналитики, способная интегрировать сбор, хранение и анализ данных об образовательном процессе для повышения качества обучения и управления;

- система образовательной аналитики, включающая инструменты машинного обучения для прогнозирования успеваемости студентов и предоставления персонализированных рекомендаций;
- предложенное решение для интеграции технологий машинного обучения в образовательную аналитику, обеспечивающее автоматизацию процессов анализа данных, прогнозирования и персонализации образовательных траекторий;
- технологическое решение для визуализации и отчетности, предоставляющее удобные инструменты для анализа данных, оперативного принятия решений преподавателями и администрацией.

Структура магистерской диссертации. Работа состоит из введения, четырех глав, заключения, списка использованной литературы и двух приложений. Основной текст изложен на 94 страницах, включает 47 рисунков и 10 таблиц.

1 Теоретические аспекты применения аналитики и машинного обучения в сфере образования

1.1 Понятие образовательной аналитики: цели, задачи, применение

Образовательная аналитика (ОА) представляет собой систему методов и инструментов, предназначенных для сбора, обработки и анализа данных, связанных с образовательным процессом. В условиях глобальной цифровизации и постоянного увеличения объема образовательных данных ОА становится неотъемлемым элементом управления образовательными учреждениями и способствует улучшению качества обучения [12], [39].

Рассмотрим понятие образовательной аналитики подробнее. Образовательная аналитика в широком смысле - это измерение, сбор, анализ и представление данных об учащих и окружающем их контексте. Цель ОА – понимание и оптимизация образовательного процесса [1].

Исходя из проблемного вопроса, для решения которого используется ОА, выделим ее основные виды: дескриптивную, прогнозную и прескриптивную аналитику [15].

Дескриптивная образовательная аналитика отвечает на вопросы: «Что происходит?» и связана с анализом фактов, говорящих о том, как в настоящее время обучаются студенты, на какие группы их можно классифицировать (по каким признакам), с какими сложностями сталкиваются обучающиеся в образовательном процессе.

Прогнозная аналитика ищет ответы на вопрос «Что скорее всего произойдет?», «Что произойдет, если не будут выполнены конкретные задания?», «Что произойдет, если пользователь будет выделять заданный объем времени на прохождение обучения?». Такой вид ОА призван предположить развитие тех или иных исходов в заданной образовательной траектории [31].

Прескриптивная аналитика применяется для поиска ответов на вопрос «Что следует делать?». Данный вид ОА является наиболее громоздким и сложным, но также и наиболее результативный. Такой вид аналитики обеспечивает возможность получения конкретных рекомендаций по достижению оптимальных результатов в ходе учебного процесса. Такой вид аналитики помогает оптимизировать маршруты участников образовательного процесса, управлять учебными программами, и улучшать качество образования в заведении.

ОА трансформирует подходы к управлению образованием, делая их более адаптивными, персонализированными и ориентированными на большие данные.

Основные цели образовательной аналитики:

- повышение эффективности обучения – ОА осуществляет анализ взаимодействия студентов с учебными материалами, выявляет пробелы в знаниях и помогает адаптировать учебные программы в режиме реального времени. Например, системы образовательной аналитики могут отслеживать результаты тестов, активность в онлайн-курсах и уровень вовлеченности студентов, что позволяет оперативно корректировать учебный процесс [30, 42];
- оптимизация учебного процесса – ОА предлагает рекомендации по изменению учебных программ, настраивая образовательные траектории под индивидуальные потребности учащихся;
- системы могут предлагать пересмотр тем, которые вызывают наибольшие затруднения у студентов [19];
- поддержка принятия решений – аналитические инструменты предоставляют преподавателям и администрации объективные данные для принятия решений, минимизируя субъективные ошибки, что касается как индивидуальных решений по каждому студенту, так и стратегических управленческих решений на уровне всего высшего учебного заведения;

– прогнозирование образовательных рисков – использование моделей прогнозирования позволяет выявлять потенциальные риски, такие как вероятность академической неуспеваемости, высокий уровень текучести студентов или задержки в освоении программы.

Прогнозы позволяют своевременно вмешиваться и предоставлять необходимую поддержку учащимся.

Рассмотрим ключевые задачи образовательной аналитики, такие как мониторинг и анализ академической успеваемости, сохранение контингента, повышение успеваемости, получение обратной связи и персонализация образовательной траектории.

Мониторинг и анализ академической успеваемости включает в себя сбор данных о результатах обучения студентов, активности на образовательных платформах и прогрессе по курсам помогает анализировать эффективность программ и выявлять проблемные области. Постоянный мониторинг позволяет преподавателям вовремя корректировать учебный процесс.

Сохранение контингента. Организации при использовании образовательной аналитики ориентируются на решение актуальных задач, и важнейшая из них – сохранение контингента обучающихся. Отсев студентов во время обучения в связи с академической неуспеваемостью – достаточно острая проблема, которая приводит к снижению общих результатов учебного процесса и мотивации других студентов, репутационным рискам организации и другим негативным последствиям [43]. В обучении с использованием дистанционных методов с применением онлайн-платформ такая проблема становится еще более выраженной [12].

Повышение успеваемости – следующая актуальная задача, для решения которой образовательные организации прибегают к использованию образовательной аналитики [8]. Зачастую обучающийся прекращает прохождение обучения в связи с тем, что не может справиться с темпом учебной программы. В этом случае своевременное обнаружение проблемной ситуации и выработка качественных решений позволяют скорректировать

взаимодействие со слушателем курсов и, как следствие, сохранить студенческий контингент.

Получение обратной связи – еще одной важной задачей применения ОА является получение обратной связи при оценке эффективности новых образовательных методик. Анализ цифровых результатов обучения студентов позволяет оценить, насколько методика эффективна, какие аспекты вызвали наибольшие сложности, где требуется доработка [13].

Персонализация образовательных траекторий: применение аналитики позволяет создать индивидуальные учебные программы, которые адаптированы под уровень подготовки, интересы и темп обучения каждого студента. Например, системы могут рекомендовать дополнительные материалы или курсы в зависимости от успеваемости студента.

Таким образом, можем выделить основные направления применения образовательной аналитики в высших учебных заведениях:

- оценка успеваемости – анализ данных о результатах тестов, выполнении заданий и активности студентов;
- поддержка учебных решений – обоснованное принятие решений по корректировке учебных планов и предоставлению индивидуальной помощи;
- повышение вовлеченности студентов – анализ взаимодействия студентов с материалами курса, их активности на образовательных платформах и в онлайн-среде.

1.2 Анализ существующих исследований в области образовательной аналитики

Существующие исследования подтверждают значимость образовательной аналитики и многообразие ее применения в образовательной практике. А.А. Булдаев и коллеги подчеркивают, что успешное использование ОА основывается на нескольких ключевых аспектах, включая внедрение

систем поддержки принятия решений, опирающихся на данные [5]. Это позволяет более точно учитывать потребности студентов и адаптировать образовательные программы.

По мнению Н.Д. Бардакова, цифровая трансформация образования открывает новые возможности для применения аналитических инструментов, таких как визуализация данных и алгоритмы машинного обучения [1]. Эти технологии позволяют не только собирать и анализировать большие объемы данных, но и находить новые подходы к управлению образовательными процессами, персонализируя обучение и улучшая взаимодействие между преподавателями и учащимися.

С другой стороны, О.А. Белянская подчеркивает значимость образовательной аналитики для улучшения качества образовательного процесса, акцентируя внимание на том, что внедрение аналитических инструментов позволяет более точно оценивать успеваемость и поведение студентов [2]. Аналитика предоставляет преподавателям и администрации возможность своевременно выявлять проблемы и разрабатывать стратегии для их решения, что способствует снижению числа академических неудач и повышению мотивации учащихся. Благодаря анализу данных, образовательные учреждения могут более эффективно адаптировать учебные программы под нужды студентов, повышая качество и персонализацию обучения [20], [21].

Многочисленные исследования подтверждают эффективность применения предиктивной аналитики для прогнозирования успешности учащихся. А.И. Гайдаржи подчеркивает, что модели машинного обучения могут существенно повысить точность прогнозов, что, в свою очередь, позволяет образовательным учреждениям заранее выявлять студентов, нуждающихся в дополнительной поддержке [7]. Это помогает не только снизить уровень отчислений, но и улучшить общую успеваемость [25], [36].

Важным направлением является интеграция образовательной аналитики с существующими информационными системами в ВУЗах. В.В. Булгаков

акцентирует внимание на том, что успешное внедрение аналитических инструментов требует комплексного подхода и учета всех уровней образовательного процесса, включая обучение преподавателей использованию аналитики и создание пользовательских интерфейсов для интерпретации данных.

Аналитика также играет ключевую роль в мониторинге и оценке качества образовательных программ. В.И. Токтарова описывает компаративный анализ образовательной аналитики в разных странах, отмечая, что стратегическое применение аналитики позволяет выявлять лучшие практики и адаптировать их к местным условиям [37]. Таким образом, образовательная аналитика становится неотъемлемой частью процесса управления качеством образования, позволяя создавать более гибкие и адаптивные образовательные среды.

Однако существуют и критические точки зрения. По мнению В.М. Розина, цифровизация и аналитика не всегда дают только положительный эффект, и важно учитывать их влияние на образовательные процессы в долгосрочной перспективе, чтобы избежать потери важного личностного взаимодействия между преподавателями и студентами [32]. Таким образом, хотя ОА и имеет значительные преимущества, она требует тщательного подхода к внедрению и учету педагогических особенностей.

1.3 Применение машинного обучения для анализа данных в образовательных системах

Машинное обучение (МЛ) активно внедряется в образовательные системы, открывая новые возможности для анализа больших объемов данных и формирования прогнозов на основе накопленной информации [22], [45]. Использование алгоритмов МЛ позволяет выявлять скрытые закономерности и делать прогнозы на основе данных об учащих, их поведении и результатах обучения [41], [42].

Исследования показывают, что применение МЛ в образовательной аналитике способствует более глубокому пониманию факторов, влияющих на успех студентов. Например, согласно работе А.В. Канаш и А.С. Мезиной, алгоритмы МЛ могут быть использованы для анализа успеваемости студентов и выявления факторов, влияющих на их обучение [13]. Это позволяет преподавателям заранее определять группы студентов, которым может потребоваться дополнительная поддержка или индивидуализированный подход.

Кроме того, МЛ помогает в разработке адаптивных учебных систем, которые могут подстраиваться под уровень подготовки и интересы студентов. А.А. Булдаев и Л.В. Найханова указывают на важность создания систем поддержки принятия решений на основе данных, которые могут не только учитывать успехи студентов, но и адаптировать учебные материалы и задания в соответствии с их потребностями [5].

Применение машинного обучения в образовательной аналитике расширяет возможности для внедрения современных подходов в обучении [3]. Это позволяет учебным заведениям рациональнее распределять ресурсы и создавать персонализированные траектории обучения для студентов.

Среди базовых методов МЛ в образовательной аналитике выделяются:

- классификация – служит для распределения студентов по категориям на основе их характеристик и поведения (например, студентов можно отнести к группам с высоким риском отчисления или высокой успеваемостью, опираясь на их академические результаты и активность);
- регрессия – этот метод помогает прогнозировать количественные показатели, такие как будущие оценки студентов, используя данные об их прошлых результатах, посещаемости и активности;
- кластеризация – объединяет студентов по схожим признакам или действиям, что полезно для создания персонализированных рекомендаций и планов обучения;

– рекомендательные системы – эти системы автоматизировано подбирают учебные материалы и курсы с учетом предпочтений и академических достижений студентов, что помогает дополнительно поддерживать интерес учащихся и улучшать их результаты.

Рассмотрим примеры успешного применения МЛ в образовательных системах:

– прогнозирование успеваемости – в ряде университетов системы на базе МЛ прогнозируют успеваемость студентов, анализируя их поведение на образовательных платформах, выполнение заданий и вовлеченность, что помогает выявить студентов, которым нужна дополнительная поддержка [41], [44];

– анализ обучающих трендов – системы МЛ помогают оценивать эффективность различных учебных методик, отслеживая результаты и вовлеченность студентов, что позволяет оптимизировать учебные программы и повысить их качество [51], [53].

Рассмотрим проблемы и вызовы применения технологий машинного обучения в образовании:

– качество данных – для успешного применения МЛ требуется высокое качество данных: ошибки в данных могут привести к недостоверным результатам и неэффективным рекомендациям, что может отрицательно повлиять на образовательный процесс [26];

– конфиденциальность данных – при обработке больших объемов личной информации учащихся важна защита данных от несанкционированного доступа [49].

Таким образом, машинное обучение становится значимым инструментом в образовательной аналитике, предоставляя возможности для создания персонализированных и эффективных учебных траекторий. Для успешного применения машинного обучения важно учитывать качество данных и безопасность персональной информации студентов и других участников системы.

1.4 Роль аналитики в повышении качества образования в учебных заведениях

Как отмечает Н.А. Шорохова, сегодня активно развивается новое направление в сфере практических исследований и разработок – образовательная аналитика, которая с помощью подробного анализа разнородных образовательных данных позволяет сделать выводы для принятия решения по оптимизации образовательного процесса с применением методов машинного обучения [50]. В ОА активность обучающегося представлена в виде данных, учитывающих индивидуальные характеристики субъекта, накопленный им опыт с фиксированием данных в информационных системах, включая этапы сбора, анализа и представления данных о пройденных учебных событиях [46], [52].

Высокая роль образовательной аналитики в образовательном процессе подтверждается оказываемой помощью в следующих направлениях [7]:

- подбор направления подготовки обучающихся;
- выбор дисциплины вариативной части программы [48];
- помощь при комплектовании групп, исходя из психологических типов личности, пристрастий и интересов [28];
- помощь при смене направления подготовки и будущей специальности студентам, сомневающимся в правильности первоначального выбора;
- подготовка отчетных материалов администрации ВУЗов.

Данные, полученные с помощью образовательной аналитики, могут быть интересны и самим студентам, в случае болезни и других причин, для дальнейшей корректировки времени освоения сложных модулей дисциплины, построения индивидуального графика обучения [35]. Это позволит обучающимся снизить риски отчислений, показать результаты индивидуальной и командной деятельности, выполненных лабораторных работ и проектов, методики взаимооценивания выполненных курсовых и

практических работ, отметить внутригрупповое контактирование при работе над проектами [14, С.55].

Для руководителей ВУЗов данные, полученные с помощью образовательной аналитики, могут быть полезны для оперативного формирования отчетности по различным показателям обучения, оценивания результатов эффективности деятельности преподавателя, подведения результатов промежуточной и текущей успеваемости, выявления актуальности разработанного контента, возможность учета студентов, у которых есть трудности в освоении дисциплин [6], [9], [33].

При этом важно отметить, что для реализации потенциала ОА, система образования в ВУЗе должна быть гибкой и допускать возможность построения адаптивных образовательных траекторий, перерасчетов нагрузок педагогического состава, а также возможность внедрения новых методик в образовательный процесс, что не является типичным для системы высшего образования на сегодняшний день [29], [30], [40].

Еще одним барьером является сбор данных по всем этим сервисам и платформам, который носит разрозненный характер, так как данные не собираются в единую базу данных, с которой могла бы работать определенная система, что существенно снижает вероятность получения полноценной картины обучения на основе анализа данных по конкретному обучающемуся. Все вышесказанное может усложнить процесс внедрения ОА в систему образования ВУЗов [54], [55].

Резюмируя роль ОА в ВУЗах, можно отметить, что она может занять важное место в образовательном процессе, повлиять положительно на качество образования во всех его проявлениях, хотя внедрение системы ОА не сможет миновать некоторые барьеры и трудности.

Выводы по главе

Первая глава диссертации посвящена теоретическим аспектам применения аналитики и машинного обучения в сфере образования. Образовательная аналитика показала себя как значимый инструмент, способный значительно улучшить качество учебного процесса. Анализ показал, что аналитика поддерживает мониторинг успеваемости, позволяет гибко адаптировать учебные программы и помогает в оценке методов преподавания. Технологии машинного обучения предоставляют возможность глубокого анализа данных о студентах, выявления скрытых закономерностей и прогнозирования академических результатов, что позволяет заранее поддерживать студентов, нуждающихся в помощи. Всё это подтверждает актуальность внедрения аналитических подходов в образовательные процессы.

Вместе с тем, несмотря на значительные достижения, в области образовательной аналитики остаются нерешённые вопросы. Необходимы дополнительные исследования, направленные на разработку усовершенствованных алгоритмов машинного обучения и их более глубокое изучение с точки зрения влияния на учебный процесс. Важно также провести оценку эффективности уже существующих практик применения аналитики в образовании и определить факторы, влияющие на успешное внедрение этих инструментов в учебные заведения.

2 Анализ существующих информационных систем в образовательных учреждениях

2.1 Существующие процессы сбора, хранения и анализа образовательных данных

Современные образовательные учреждения активно применяют различные информационные системы для сбора, хранения и анализа данных о студентах, их успеваемости, посещаемости и учебной активности, а также для учета взаимодействия между студентами и преподавателями. Однако, несмотря на значительные успехи в автоматизации образовательных процессов, многие системы по-прежнему фрагментированы и не интегрированы в единый аналитический контур [8], [10].

Сравнительный анализ существующих систем сбора, хранения и анализа образовательных данных, представленный в таблице 1, демонстрирует ключевые особенности используемых информационных систем в образовательном процессе.

Таблица 1 – Сравнительный анализ существующих систем, отражающий процессы сбора, хранения и анализа образовательных данных

Система	Процессы сбора данных	Процессы хранения данных	Процессы анализа данных	Преимущества	Ограничения
Системы управления обучением (LMS)	сбор данных о посещаемости, успеваемости, заданиях и тестах	хранение данных в базе данных внутри LMS	основной анализ через встроенные отчеты и метрики	интеграция с учебными материалами, поддержка онлайн-обучения	часто отсутствует интеграция с другими системами, ограниченные возможности аналитики
Системы управления данными студентов (SIS)	сбор данных о личной информации студентов, академической истории	хранение данных в реляционных базах данных или облачных системах	основной анализ через стандартные отчеты и метрики	полный контроль над личной информацией студентов, систематизация данных	не всегда обеспечивается глубокий анализ данных, ограниченная гибкость в отчетности

Продолжение таблицы 1

Система	Процессы сбора данных	Процессы хранения данных	Процессы анализа данных	Преимущества	Ограничения
Инструменты для анализа данных (BI)	интеграция данных из различных источников, включая LMS и SIS	хранение данных на платформах или в хранилищах данных	расширенный анализ с помощью визуализаций и дашбордов	гибкость в визуализации и анализе данных, поддержка комплексных отчетов	может потребоваться сложная интеграция, высокая стоимость и сложность настройки
Аналитические платформы (Google Analytics)	сбор данных о действиях пользователей с веб-ресурсами образования	хранение данных в облачных хранилищах	анализ поведения пользователей их посещаемости и вовлеченности	анализ пользовательского поведения на веб-ресурсах, простота интеграции с веб-платформами	не включает данные об образовании, ограниченные возможности для глубокой аналитики в образовании

Анализ существующих информационных систем выявляет несколько ключевых недостатков, которые подчеркивают необходимость разработки новой системы образовательной аналитики:

- фрагментация данных: существующие системы часто не интегрированы друг с другом, что приводит к фрагментированным данным и затрудняет комплексный анализ;
- ограниченные возможности аналитики: многие системы предоставляют лишь базовый уровень анализа и отчетности, не позволяя глубоко исследовать данные или делать прогнозы;
- отсутствие продвинутых методов визуализации: существующие системы могут не поддерживать современные методы визуализации данных, такие как интерактивные панели управления и визуализация временных рядов, что ограничивает понимание данных [47, 52];
- недостаток использования машинного обучения: технологии машинного обучения, которые могут значительно улучшить прогнозирование успеваемости, оптимизацию учебных планов и

персонализированные рекомендации, не всегда интегрированы в существующие системы [27], [34].

Рассмотрим основные бизнес-процессы исследуемой предметной области «Как есть» в нотации IDEF0 (рис 1, 2). Исследуем процессы образовательных организаций, где применяются различные информационные системы (ИС) для управления учебным процессом, сбора данных о студентах и их успеваемости. Основное внимание уделяется анализу текущих процессов сбора и обработки данных с целью выявления проблем фрагментации и отсутствия интеграции данных между системами, что затрудняет комплексный анализ.



Рисунок 1 – Текущее состояние бизнес-процесса обработки образовательной аналитики до внедрения ИС

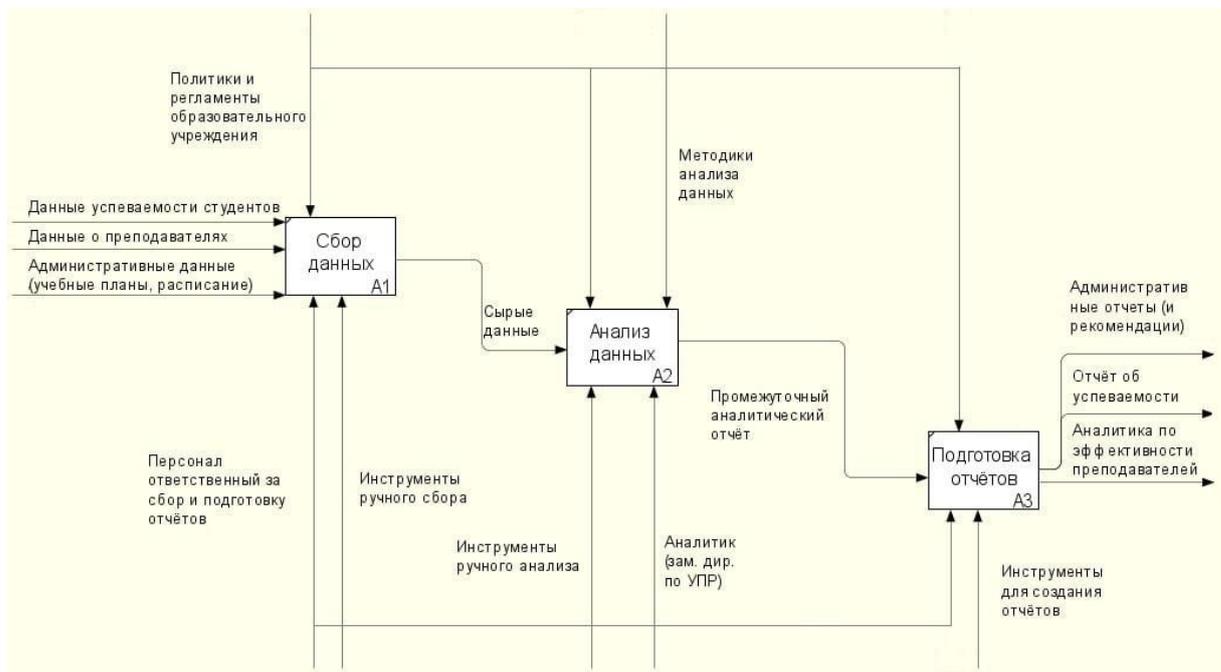


Рисунок 2 – Декомпозиция бизнес-процесса обработки образовательной аналитики до внедрения ИС

Проанализировав рисунки 1, 2 можно выделить следующие аспекты, связанные со сбором и анализом данных.

Сбор данных. Ручной ввод данных преподавателями и администрации: преподаватели вручную собирают и вводят данные о своих студентах в различные системы. Эти данные включают информацию о посещаемости, оценках за контрольные и экзамены, а также активности студентов на занятиях. Администрация, в свою очередь, собирает данные о зачислении студентов, расписании занятий и экзаменов.

Источники данных: бумажные ведомости, электронные таблицы Excel, анкеты студентов.

Методы сбора: ручное заполнение таблиц, ведомостей, опросников.

Проблемы: зависимость от человеческого фактора, ошибки при вводе данных, потеря бумажных носителей.

Сбор данных о посещаемости: преподаватели на каждом занятии вручную отмечают присутствующих студентов в журналах или в электронных таблицах. Этот процесс требует времени и может нарушать учебный процесс.

Данные о посещаемости нередко вводятся в систему с задержкой в несколько дней или недель, что затрудняет оперативный анализ и принятие мер.

Источники данных: бумажные журналы посещаемости, электронные таблицы.

Методы сбора: ручной ввод данных преподавателями.

Проблемы: низкая оперативность, ошибки при вводе, невозможность верификации данных.

Обработка данных. Ручная аналитика: после того как данные собраны и сохранены, начинается процесс их анализа. В большинстве случаев обработка данных осуществляется вручную, с использованием электронных таблиц. Администрация и входящие в ее состав аналитики создают сводные отчёты, графики и диаграммы вручную. Это требует больших временных затрат и может привести к ошибкам.

Методы анализа: ручное создание таблиц, сводок и диаграмм в Excel.

Проблемы: ограниченные аналитические возможности, ошибки при расчётах, низкая оперативность, отсутствие прогнозной аналитики.

Коммуникация и передача данных. Ручная передача данных между подразделениями: когда данные необходимы для анализа различным департаментам, они передаются вручную. Преподаватели или администрация копируют нужную информацию и отправляют ее другим подразделениям или сотрудникам. Это может быть выполнено через электронную почту или физически на бумажных носителях.

Методы передачи данных: электронные письма, перенос на флеш-носителях, бумажные копии.

Проблемы: задержки в передаче информации, невозможность синхронизации данных в реальном времени, ошибки при передаче.

Итоговая визуализация данных. Ручное создание диаграмм и графиков: преподаватели и администрация создают графики и диаграммы вручную, используя электронные таблицы. Эти графики затем используются для отчетности перед руководством или для представления на собраниях.

Методы визуализации: ручное создание диаграмм в MS Excel или других программах.

Проблемы: ограниченная визуализация, отсутствие интерактивности, сложность восприятия данных.

Анализ данных в большинстве случаев ограничен традиционными статистическими методами, основанными на отчетах и диаграммах, без внедрения прогнозных и продвинутых аналитических инструментов [11].

Исследование существующих процессов образовательных организаций, выявило несколько ключевых проблем в системе управления образовательными данными. Основные из них включают фрагментацию используемых информационных систем, отсутствие автоматизированной интеграции данных и значительные зависимости от ручного ввода данных. Эти проблемы усложняют реализацию комплексного анализа образовательных процессов и снижают эффективность управленческих решений.

2.2 Проблемы и ограничения существующих систем образовательной аналитики

Современные системы образовательной аналитики сталкиваются с рядом сложностей, которые ограничивают их эффективность и затрудняют оперативность управленческих решений:

- фрагментированность данных – в образовательных учреждениях нередко используется несколько независимых систем, что усложняет интеграцию и затрудняет центральный анализ данных [9, С.490];
- ручной ввод данных – человеческий фактор часто приводит к ошибкам при внесении данных, снижая точность последующего анализа;
- ограниченные возможности прогнозирования – применяемые инструменты, такие как MS Excel и стандартные статистические

программы, недостаточно мощны для проведения сложных прогнозных вычислений, что затрудняет возможность прогноза успеваемости студентов и оперативного выявления потенциальных рисков в учебном процессе [17];

– отсутствие персонализированной аналитики – существующие системы недостаточно гибкие и не учитывают потребности разных категорий пользователей (например, преподаватели нуждаются в оперативных данных по успеваемости студентов, тогда как руководству требуются сводные отчёты для стратегического анализа) [16, С.71];

– недостаточные средства визуализации данных – стандартные диаграммы и таблицы не всегда позволяют проводить глубокий анализ и визуально представлять информацию в доступной форме [23].

Выделим необходимые улучшения:

– повышение точности аналитики через улучшение системы обработки информации. Достоверные данные – ключевой фактор для принятия обоснованных решений. Повышение точности анализа позволяет своевременно выявлять проблемы и тенденции, что способствует улучшению качества учебного процесса;

– внедрение мониторинга успеваемости студентов в реальном времени. Оперативное отслеживание успеваемости дает возможность быстрее реагировать на трудности, с которыми сталкиваются студенты, и повышает управляемость образовательного процесса;

– автоматизация процессов сбора и анализа данных. Автоматический сбор данных помогает своевременно выявлять студентов, нуждающихся в поддержке, и анализировать результаты учебных программ, что необходимо для принятия обоснованных управленческих решений [4].

Выделим возможные направления для улучшений:

- создание системы мониторинга удовлетворённости студентов – оценка мнений студентов о качестве обучения и условиях способствует улучшению образовательной среды и повышению удовлетворённости (хотя это и важно, данное улучшение не является критическим для основного функционирования учебного процесса);
- повышение прозрачности и доступности коммуникации для всех участников учебного процесса – прозрачность и открытая коммуникация с руководством, преподавателями и студентами способствует доверию и вовлеченности (однако это скорее влияет на взаимодействие, чем на само качество обучения).

2.3 Проектирование новой архитектуры системы сбора и анализа данных в образовательном учреждении

При разработке системы образовательной аналитики особое внимание следует уделить объединению данных из различных источников в централизованное хранилище. Такое хранилище обеспечит целостность и удобный доступ к информации из учебных платформ, систем управления учебным процессом (LMS) и других образовательных ресурсов [18].

Основные компоненты предлагаемой архитектуры включают:

- централизованное хранилище данных – позволяет интегрировать данные из разных источников, обеспечивая их согласованность и целостность;
- автоматизация сбора данных – устраняет необходимость ручного ввода информации, благодаря использованию датчиков, систем мониторинга и цифровых форм;
- модули обработки и анализа данных в реальном времени – включают инструменты для обнаружения аномалий, прогнозирования и

моделирования, опирающиеся на исторические данные;

– средства визуализации данных – интерактивные панели и автоматически обновляемые отчеты дают преподавателям и администрации оперативный доступ к анализу, позволяя принимать решения, опираясь на текущие данные.

С изменениями бизнес-процесса образовательной аналитики после внедрения ИС можно ознакомиться в нотации IDEF0 (см. рисунки 3, 4, 5).



Рисунок 3 – Планируемый бизнес-процесс обработки образовательной аналитики после внедрения ИС

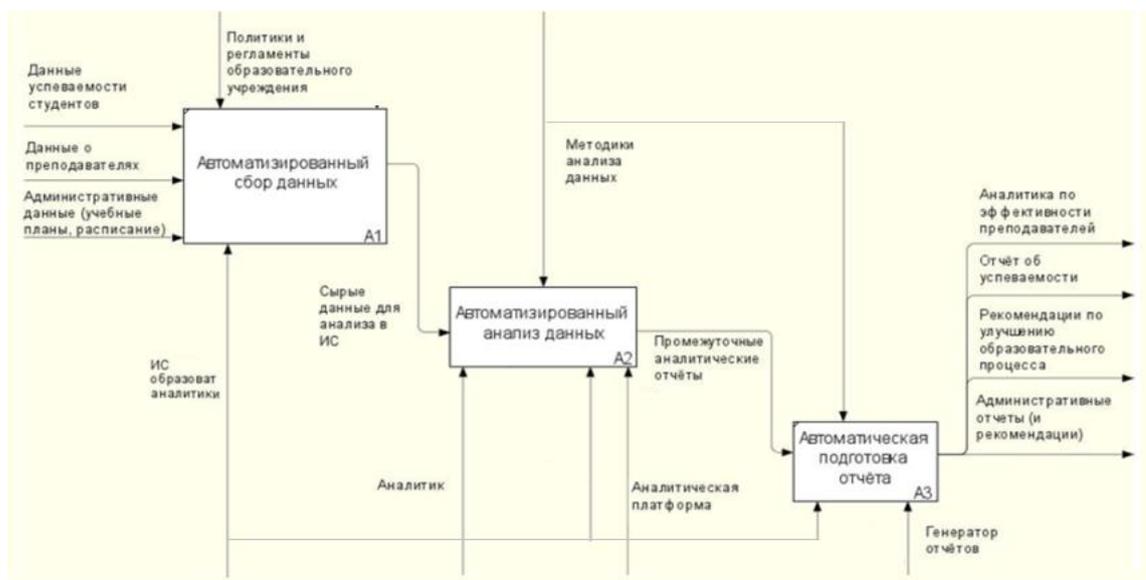


Рисунок 4 – Декомпозиция бизнес-процесса обработки образовательной аналитики после внедрения ИС

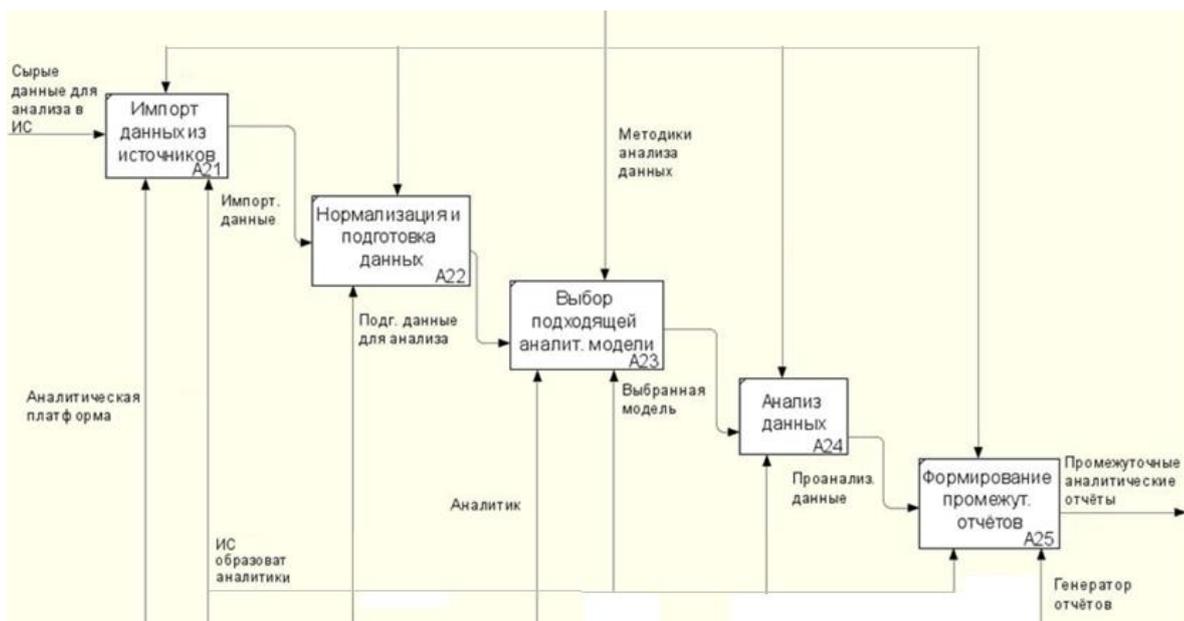


Рисунок 5 – Декомпозиция процесса «Автоматизированный анализ данных» после внедрения ИС

При разработке системы образовательной аналитики для образовательных учреждений высшего образования особое внимание следует уделить объединению данных из различных источников в централизованное хранилище. Такое хранилище обеспечит целостность и удобный доступ к информации из учебных платформ, систем управления учебным процессом (LMS) и других образовательных ресурсов.

Диаграмма потоков данных в разрабатываемой ИС представлена на рисунке 6.

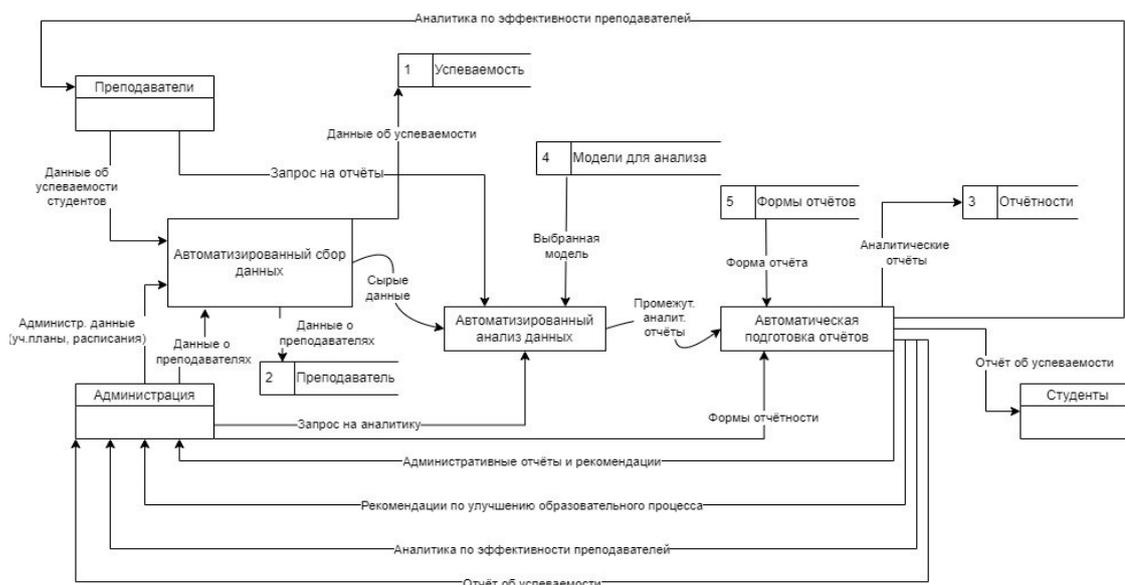


Рисунок 6 – Диаграмма потоков данных в ИС

Визуализация планируемого бизнес-процесса обработки образовательной аналитики в нотации BPMN представлена на рисунке 7.

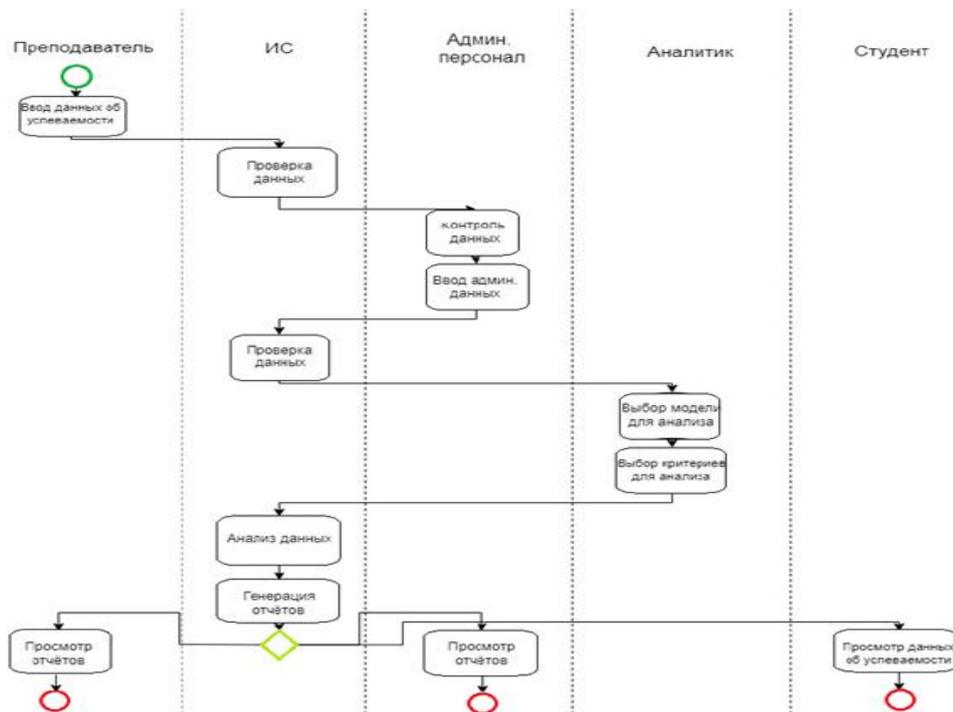


Рисунок 7 – Бизнес-процесс обработки образовательной аналитики в нотации BPMN после внедрения ИС

На рисунках 8-11 представлены диаграммы вариантов использования для акторов преподаватель, студент, аналитик, административный персонал.

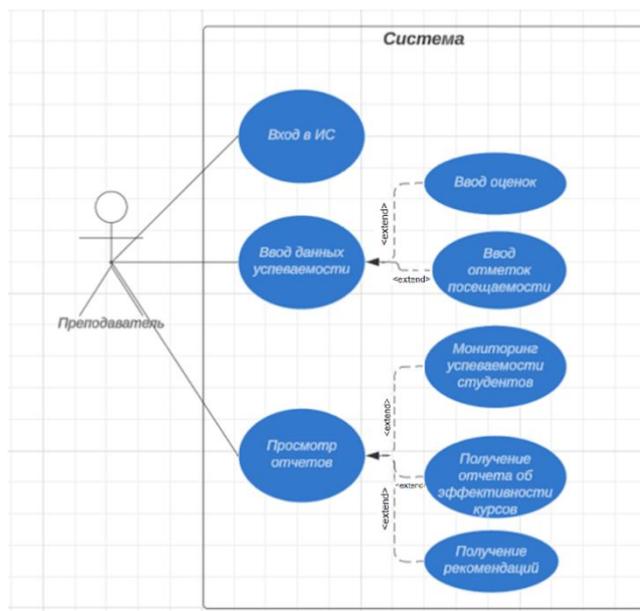


Рисунок 8 – Диаграмма вариантов использования для актора «Преподаватель»

Преподаватель после авторизации использует систему для ввода данных об успеваемости студентов, получения отчетности по успеваемости, эффективности курсов и получения персональных рекомендаций.

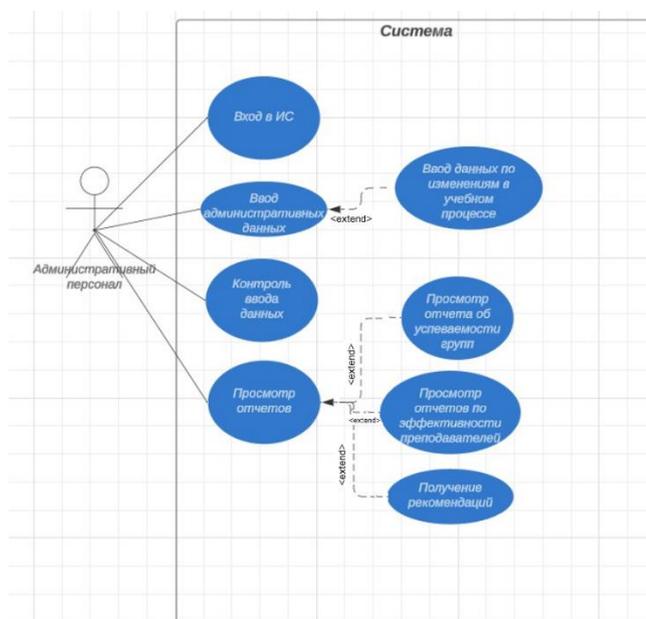


Рисунок 9 – Диаграмма вариантов использования для актора «Административный персонал»

Административный персонал после авторизации использует систему для ввода административных данных, просмотра статистики посещаемости по

курсам и группам студентов, получения отчетности по эффективности преподавателей, общему состоянию качества образования в организации, получения рекомендаций по усовершенствованию образовательного процесса.

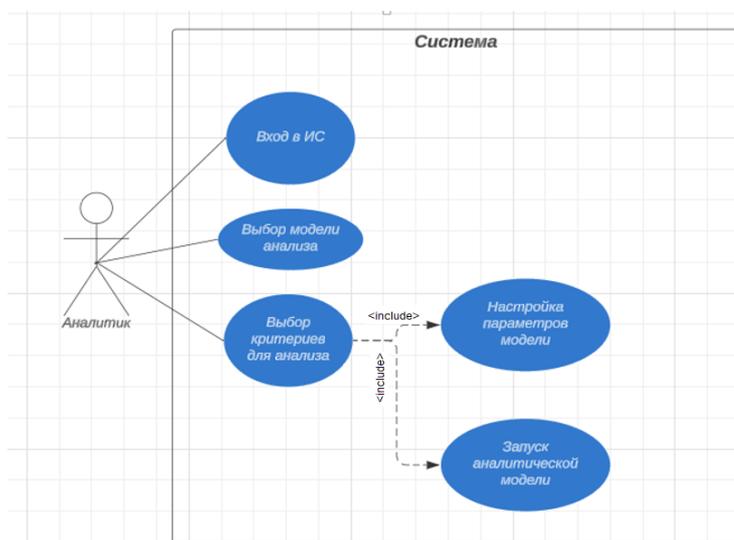


Рисунок 10 – Диаграмма вариантов использования для актора «Аналитик»

Аналитик после авторизации использует систему для выбора модели для анализа, критериев анализа (настройки параметров и запуска аналитической модели).

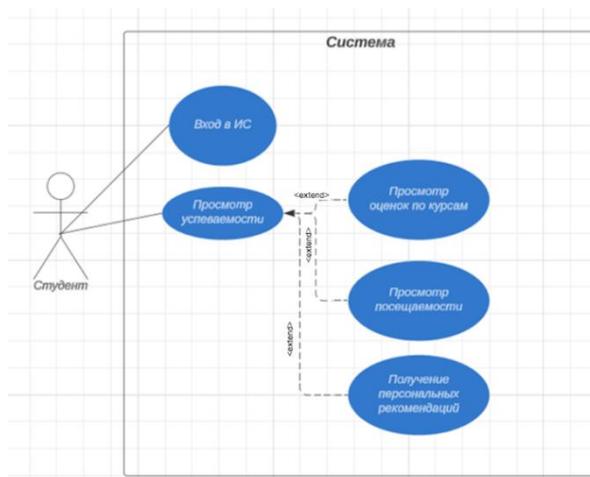


Рисунок 11 – Диаграмма вариантов использования для актора «Студент»

Студент после авторизации использует систему для просмотра успеваемости по оценкам, посещаемости, получения рекомендаций.

2.4 Применение технологии машинного обучения для прогнозной аналитики

Машинное обучение открывает широкие перспективы в образовательной аналитике, позволяя точнее прогнозировать развитие учащихся и внедрять индивидуальные рекомендации для повышения их успехов [24]. С помощью таких технологий можно выстраивать прогнозы, которые помогут выявлять успешные траектории в обучении и внедрять их на практике. Прогнозирование в этой области требует целостного подхода, который будет учитывать задачи каждого учебного заведения и обеспечивать достижение поставленных целей [2].

Одной из ключевых задач является прогнозирование успеваемости на различных этапах, которое позволяет не только отслеживать текущие достижения, но и прогнозировать дальнейший прогресс. Важный аспект прогнозной аналитики в образовании – это возможность анализировать вовлеченность учащихся, их участие в образовательном процессе и потенциал для будущих успехов. Прогнозирование в данном контексте рассматривается, как способ взглянуть в «возможные будущие» и предугадать события на основе данных, что помогает предотвратить нежелательные результаты в учебе.

Машинное обучение, благодаря широкому спектру методов, позволяет реализовать наиболее эффективные подходы к анализу данных и предсказанию успехов и трудностей, с которыми могут столкнуться учащиеся. Основные направления применения технологий машинного обучения включают прогнозирование успеваемости, персонализированные рекомендации и прогнозирование оттока.

Прогнозирование успеваемости студентов. Системы на базе машинного обучения анализируют исторические данные по оценкам, посещаемости и активности учащихся, чтобы предсказать их успех на следующих этапах обучения. Это позволяет преподавателям заранее выявить учащихся,

нуждающихся в дополнительной поддержке, и вовремя предложить им помощь.

Персонализированные рекомендации для студентов. Системы персонализированных рекомендаций помогают каждому учащемуся построить уникальную траекторию обучения. На основе текущих успехов и интересов система предлагает наиболее подходящие материалы, курсы и дополнительные занятия. Такой подход особенно ценен в онлайн и гибридном обучении, где требуется эффективно ориентироваться в большом объеме доступного контента.

Прогнозирование оттока студентов. Для учебных заведений важно удерживать учащихся в образовательном процессе. Машинное обучение позволяет предсказать вероятность оттока студентов на основании их успеваемости, посещаемости и вовлеченности. Это дает возможность заранее принимать меры по повышению мотивации и вовлеченности учащихся, чтобы уменьшить вероятность их ухода.

В таблице 2 продемонстрированы ключевые на сегодняшний день методы машинного обучения и возможность их применения в образовательной аналитике.

Таблица 2 – Ключевые методы машинного обучения и возможность их применения в образовательной аналитике

Метод МЛ	Описание	Применение	Преимущества
Регрессия	алгоритмы, прогнозирующие числовые значения на основе входных данных	прогнозирование успеваемости студентов на основе оценок, посещаемости и активности	прогнозирование будущих оценок и результатов, выявление студентов, нуждающихся в дополнительной поддержке

Продолжение таблицы 2

Метод МЛ	Описание	Применение	Преимущества
Кластеризация	метод группировки данных на основе сходства признаков	оптимизация учебных планов и расписаний на основе анализа паттернов регистрации и успеваемости	выявление групп студентов с похожими интересами и потребностям, оптимизация распределения ресурсов
Рекомендательные системы	алгоритмы, предлагающие пользователям элементы на основе их предпочтений и поведения	персонализированные рекомендации для студентов по дополнительным материалам и курсам	создание индивидуализированных учебных траекторий, повышение вовлеченности и удовлетворенности студентов
Анализ выживания	метод для прогнозирования времени до наступления определенного события	прогнозирование оттока студентов на основе успеваемости, посещаемости и вовлеченности	предотвращение отчислений, принятие превентивных мер для удержания студентов
Случайные леса	метод ансамблевого обучения, использующий множество деревьев решений для улучшения точности прогноза	прогнозирование успеваемости и оттока студентов, выявление факторов, влияющих на учебный процесс	высокая точность прогнозов, устойчивость к шуму и пропущенным данным

Интеграция данных методов в образовательную аналитику позволяет создать более персонализированную систему обучения, повысить точность прогнозов и создать комфортную, результативную образовательную среду.

2.5 Разработка новых методов визуализации данных для предоставления информации

Разработка эффективных методов визуализации данных сегодня становится ключевым направлением, способным облегчить восприятие информации и помочь пользователям принимать решения на основе

актуальных данных. Для пользователя интерактивный отчет – это не просто набор графиков и таблиц, а инструмент для глубокого анализа текущей ситуации и принятия решений [38]. Поэтому подход к созданию таких отчетов должен учитывать не только функциональную составляющую, но и эстетическую: удобство, доступность и приятный внешний вид.

Рассмотрим характеристику методов визуализации данных, используемых в рамках данного исследования (см. таблицу 3).

Таблица 3 – Характеристика методов визуализации данных

Тип визуализации	Описание	Возможности
Интерактивные панели управления	позволяют пользователям фильтровать данные и выявлять ключевые тренды	интерактивность и возможность фильтрации данных пользователями, автоматическое обновление данных в реальном времени, поддержка большого объема данных, возможность представления нескольких метрик на одном экране
Визуализация с использованием временных рядов	используется для анализа данных, меняющихся со временем (например, успеваемость, посещаемость)	четкое отображение изменений показателей во времени, возможность прогнозирования на основе исторических данных, удобство для анализа трендов и выявления сезонных закономерностей
Визуализация с использованием кластеризации и диаграмм рассеяния	позволяет выявить группы студентов с похожими характеристиками и представить взаимосвязи между переменными	возможность выявления скрытых закономерностей в данных, помогает визуализировать сложные зависимости между несколькими переменными, эффективна для сегментации студентов по различным характеристикам

Таким образом, в образовательной аналитике методы визуализации данных играют центральную роль. Комбинирование различных подходов – интерактивных панелей управления, анализа временных рядов, методов кластеризации – позволяет исследовать данные с разных точек зрения и получить максимально полную картину. Интерактивные панели помогают динамически анализировать информацию, временные ряды отражают

изменения показателей с течением времени, а методы кластеризации и диаграммы рассеяния дают возможность видеть связи между различными переменными. Вместе эти методы упрощают процесс анализа и помогают глубже понять.

Выводы по главе

Проведенный анализ показал, что в образовательных учреждениях значительная часть процессов сбора, хранения и анализа данных остается ручной и часто не интегрированной. Это снижает точность информации, увеличивает время обработки данных и затрудняет эффективное управление образовательным процессом. Текущие системы образовательной аналитики ограничены в возможностях анализа, часто не объединяют данные между отделами и не позволяют оперативно принимать решения.

Для улучшения работы таких систем важно учесть принципы автоматизации и интеграции в новой архитектуре. Рекомендуемая система включает централизованное хранилище данных, автоматизированный сбор информации, модули для обработки данных в реальном времени и средства визуализации. Методы машинного обучения могут существенно улучшить возможности образовательной аналитики, предлагая инструменты для прогнозирования успеваемости, адаптации учебных планов и создания персонализированных рекомендаций для студентов.

Новые методы визуализации данных предоставляют более гибкие и интерактивные возможности для анализа информации. Визуальные панели управления, временные ряды и методы кластеризации позволяют глубже понимать закономерности в образовательных данных и быстрее реагировать на возникающие проблемы. Эти инструменты могут быть интегрированы в ИС для улучшения восприятия данных и ускорения принятия решений.

3 Разработка системы образовательной аналитики для повышения качества образования

3.1 Выбор и обоснование технологического стека для разработки системы

Для успешной разработки системы образовательной аналитики в необходим тщательный выбор технологического стека, обеспечивающего эффективное хранение, обработку и анализ данных. Важным критерием при выборе технологий стало их способность интегрировать методы машинного обучения и использовать их для повышения качества образовательного процесса. В данном исследовании были рассмотрены различные системы управления базами данных (СУБД) и веб-фреймворки, а также языки программирования, наиболее подходящие для построения аналитической системы [10], [25]. Сравним несколько популярных СУБД, таких как PostgreSQL, MySQL, SQLite, Oracle в таблице 4.

Таблица 4 – Сравнительная характеристика СУБД

СУБД	PostgreSQL	MySQL	SQLite	Oracle
Критерий				
тип СУБД	реляционная, объектно-реляционная	реляционная	встраиваемая реляционная	реляционная, объектно-реляционная
расширяемость	высокая	ограниченная	ограниченная	высокая, но сложная
масштабируемость	высокая	умеренная	низкая	высокая
поддержка JSON	да	ограниченная	нет	ограниченная
поддержка шифрования	да	ограниченная	нет	да
аудит безопасности	да	нет	нет	да
возможность разграничения доступа	да	ограниченная	нет	да
лицензия	открытая (PostgreSQL License)	открытая (GPL)	открытая (Public domain)	проприетарная

Исходя из результатов сравнения СУБД, в рамках реализации системы образовательной аналитики, была выбрана PostgreSQL. Данный выбор обусловлен гибкостью СУБД, поддержкой сложных типов данных, включая JSON, и возможностью масштабирования, что особенно важно для работы с большими объемами данных в системе образовательной аналитики. PostgreSQL также предоставляет встроенные механизмы расширений и расширенные возможности работы с транзакциями и репликацией, что делает ее идеальным выбором для крупных аналитических систем.

Для построения веб-интерфейса системы образовательной аналитики и обеспечения доступа пользователей к данным и отчетам требуется надежный веб-фреймворк. Сравнение наиболее распространенных фреймворков продемонстрировано в таблице 5.

Таблица 5 – Сравнение фреймворков для построения веб-интерфейса

Фреймворк	Django	Flask	Ruby on Rails	Spring (Java)
Критерий				
язык программирования	Python	Python	Ruby	Java
архитектура	MVC	микрофреймворк	MVC	MVC
встроенные функции	аутентификация, ORM	минимум встроенных функций	ORM, аутентификация	огромное количество модулей
уровень сложности	средний	низкий	средний	высокий
масштабируемость	хорошая	ограниченная	хорошая	отличная

Язык программирования является основой разработки любой системы. Сравним несколько языков, популярных в веб-разработке и аналитике, представим результаты в таблице 6.

Таблица 6 – Сравнение языков программирования

Язык программирования	Python	Ruby	Java	JavaScript
Критерий				
простота освоения	высокая	средняя	низкая	средняя
скорость разработки	высокая	Средняя	низкая	высокая
масштабируемость	хорошая	орошая	отличная	отличная
поддержка библиотек	широкая	ограниченная	огромная	огромная

На основании проведенного сравнительного анализа языков программирования и веб-фреймворков, выбор пал на Django и язык программирования Python. Django – это свободный фреймворк, используемый для разработки веб-приложений. Работает на языке Python. Django использует для выполнения необходимого спектра задач шаблон проектирования MVC. За счет применения Django удастся значительно ускорить процедуру разработки приложений и веб-страниц, а также упростить их поддержку. Библиотека поможет избежать лишних ошибок в коде и обеспечит стабильность выпущенного проекта. Python, в свою очередь, предоставляет обширные возможности для машинного обучения и обработки данных благодаря библиотекам, таким как Pandas, NumPy и Scikit-learn [12], [29].

Использование фреймворка Django обосновано его тесной интеграцией с Python, а также наличием встроенных инструментов, таких как система администрирования, аутентификация и ORM. Эти возможности позволяют ускорить процесс разработки, уменьшая объем рутинной работы.

Django хорошо поддерживает масштабируемость и подходит для проектов разного размера. В контексте образовательной аналитики, где нужно обрабатывать большие объемы данных и выполнять сложные запросы, Django является удачным выбором.

Python выбран за простоту освоения, универсальность и богатый набор библиотек, что делает его отличным языком для создания как веб-приложений, так и аналитических систем.

Эта технологическая база обеспечивает возможность работы с большими данными, гибкость аналитики и масштабируемость системы для работы с увеличивающейся нагрузкой.

3.2 Разработка архитектуры системы

При проектировании архитектуры системы образовательной аналитики важно было обеспечить интеграцию данных из различных источников, таких как системы управления обучением (LMS), базы данных студентов и образовательных материалов. Архитектура предполагает наличие централизованного хранилища данных, автоматизированных процессов сбора и обработки информации, а также модулей для анализа данных в реальном времени. Общий вид архитектуры представлен на рисунке 12.

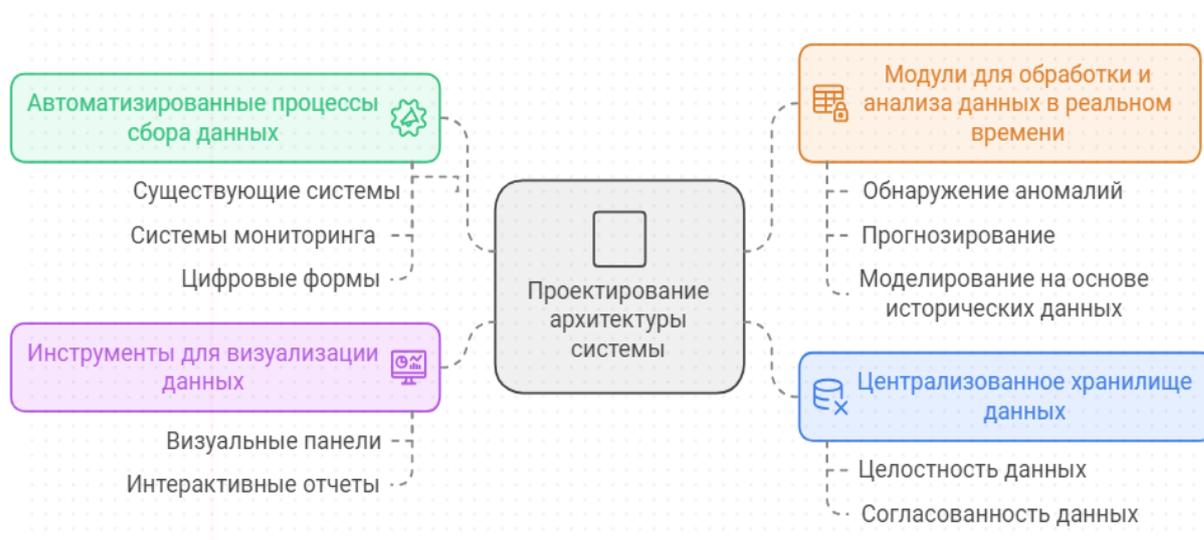


Рисунок 12 – Общий вид архитектуры системы образовательной аналитики

При обработке информации традиционным является разделение существующих задач на два широких класса:

- задачи операционной обработки данных;
- задачи аналитической обработки данных.

Они принципиально различны, требуют разных подходов к решению, но при этом взаимно дополняют друг друга. Задачами аналитической системы оценки компетентностно-ориентированных образовательных программ являются хранение, обработка и анализ данных

Рассмотрим основные компоненты архитектуры системы:

- централизованное хранилище данных – обеспечивает сбор и консолидацию информации из разных источников, что гарантирует её целостность и согласованность;
- автоматизированные процессы сбора данных – исключают необходимость ручного ввода данных за счет интеграции с существующими системами, инструментами мониторинга и цифровыми формами для заполнения данных;
- модули для обработки и анализа данных в реальном времени – включают инструменты для обнаружения аномалий, прогнозирования и моделирования на основе исторических данных, что помогает получать актуальные аналитические данные для оперативного реагирования;
- инструменты для визуализации данных – предназначены для представления аналитических данных в удобных и наглядных интерфейсах, доступных преподавателям, студентам и администрации. Эти инструменты позволяют настраивать отчеты и дашборды, что повышает прозрачность и ускоряет процесс принятия решений.

Исходя из того, что архитектура системы образовательной аналитики должна быть многослойной, обеспечивая гибкость и модульность, выделим ее основные компоненты на рисунке 13.



Рисунок 13 – Основные компоненты архитектуры системы образовательной аналитики

Рассмотрим подробнее компоненты архитектуры.

Сбор и обработка информации:

- преобразования данных из других систем;
- импорт данных в различных форматах;
- первичная обработка данных;
- проверка данных и исправление ошибок.

Хранение данных:

- загрузка и систематизация данных (создания и хранение баз данных);
- выборка интересующей информации.

Визуализация данных и управление данными:

- различная обработка и анализ данных;
- экспорт информации из базы данных.

На рисунке 14 представлена схема данных базы данных разрабатываемой системы.

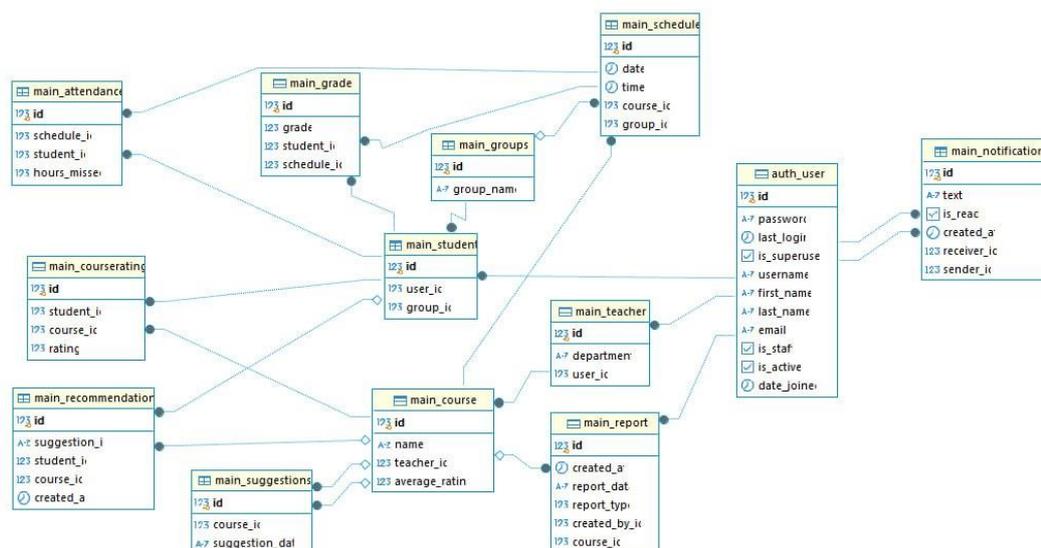


Рисунок 14 – Схема базы данных

В процессе проектирования базы данных для системы образовательной аналитики было предусмотрено создание таблиц, которые эффективно хранят ключевую информацию, необходимую для мониторинга и анализа образовательного процесса. Эти таблицы играют важную роль в обеспечении целостности данных и их удобной обработке для аналитических целей.

Чтобы реализовать спроектированную структуру системы ОА, первоначально, была разработана и логически связана схема базы данных, состоящая из следующих таблиц:

- Auth_user – таблица для хранения записей о пользователях системы (студентах, преподавателях и администрации);
- Main_attendance – таблица, предназначенная для информации о посещаемости студентов. Данная таблица служит основой для аналитической отчетности по обучению студентов, а также для прогнозирования их будущей успеваемости;
- Main_course – таблица хранит записи всех учебных курсов, по которым обучаются студенты;
- Main_courserating – таблица для оценок студентами курсов. Данная сущность служит основой для анализа популярности курсов;

- Main_grade – таблица с оценками студентов, она является ключевым элементом для мониторинга успеваемости и прогнозирования дальнейшей успеваемости;
- Main_schedule – таблица с расписанием занятий;
- Main_groups - таблица студенческих групп;
- Main_teacher – таблица преподавателей;
- Main_student – таблица хранит записи о студентах и связывает таблицы auth_user и main_group;
- Main_recommendation – таблица хранит записи персональных рекомендаций, которые система выдает студентам на основе их успеваемости, вовлеченности и активности. Эти рекомендации включают советы по дополнительным материалам, курсам или необходимости консультаций с преподавателями;
- Main_report – таблица с отчетами, которые система автоматически генерирует для преподавателей и администрации (отчеты по успеваемости, активности, прогнозы оттока и т.д.).

Вся необходимая для реализации системы ОА информация была реализована на логическом уровне схемы базы данных, придерживаясь основных принципов интеграции:

- учет перекрывающихся объектов – все перекрывающиеся объекты в интегрируемой схеме отображаются в таблице схем базы данных;
- расширенный учет перекрывающихся объектов – элементы, которые присутствуют лишь в одном источнике, но при этом так или иначе связаны с перекрывающимися объектами, внесены в полученную схему базы данных;
- нормализация – независимые связи и сущности не смешаны в одной таблице схемы;
- минимальность – ни один из элементов, независимо от его источника, не утерян.

Описанная выше структура базы данных обеспечивает целостность и согласованность данных, позволяя выполнять весь возложенный функционал на систему ОА.

3.3 Реализация системы образовательной аналитики

Реализация системы образовательной аналитики были выполнена, отталкиваясь от следующих ключевых аспектов:

- разработанная новая архитектура системы ОА;
- разработанная схема базы данных;
- выбранный технологический стек;
- проектирование удобного и понятного пользовательского интерфейса;
- использование выбранных методов машинного обучения;
- использование выбранных средств визуализации данных.

Рассмотрим подробнее функциональные и нефункциональные характеристики разработанной системы ОА.

Пользовательский интерфейс.

Интерфейс системы ОА спроектирован так, чтобы быть интуитивно понятным для преподавателей, студентов и административного персонала, с упором на простоту доступа к аналитике, отчетам и рекомендациям. На рисунке 15 представлено окно авторизации для пользователя системы.

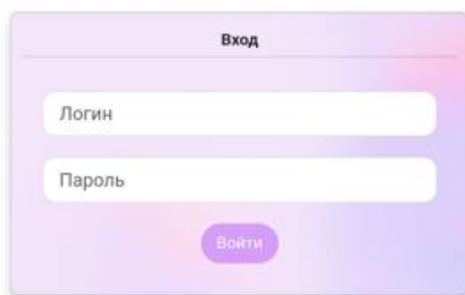


Рисунок 15 – Окно авторизации

На данном рисунке представлено окно авторизации для пользователей системы. Пользователи вводят учетные данные (логин и пароль), что позволяет разграничить права доступа.

На рисунках 16-18 представлены интерфейсы страниц личного кабинета студента.

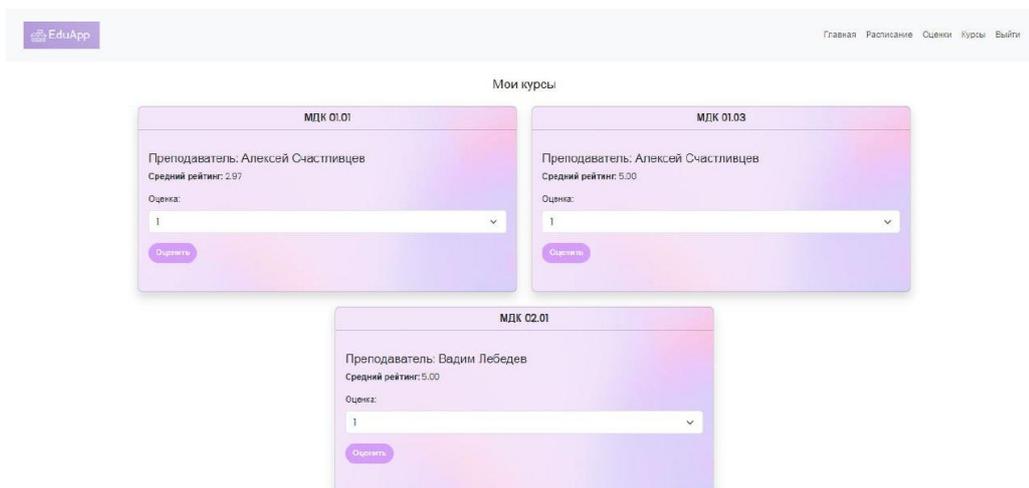


Рисунок 16 – Страница «Курсы»

Страница «Курсы» предоставляет студенту список всех курсов, которые он изучает. Студент видит список всех курсов, которые он изучает, с возможностью выставлять оценки для каждого из них. Это помогает отслеживать прогресс в учебе.

На рисунке 17 представлена страница «Оценки» студента и успеваемость его группы по выбранному курсу.

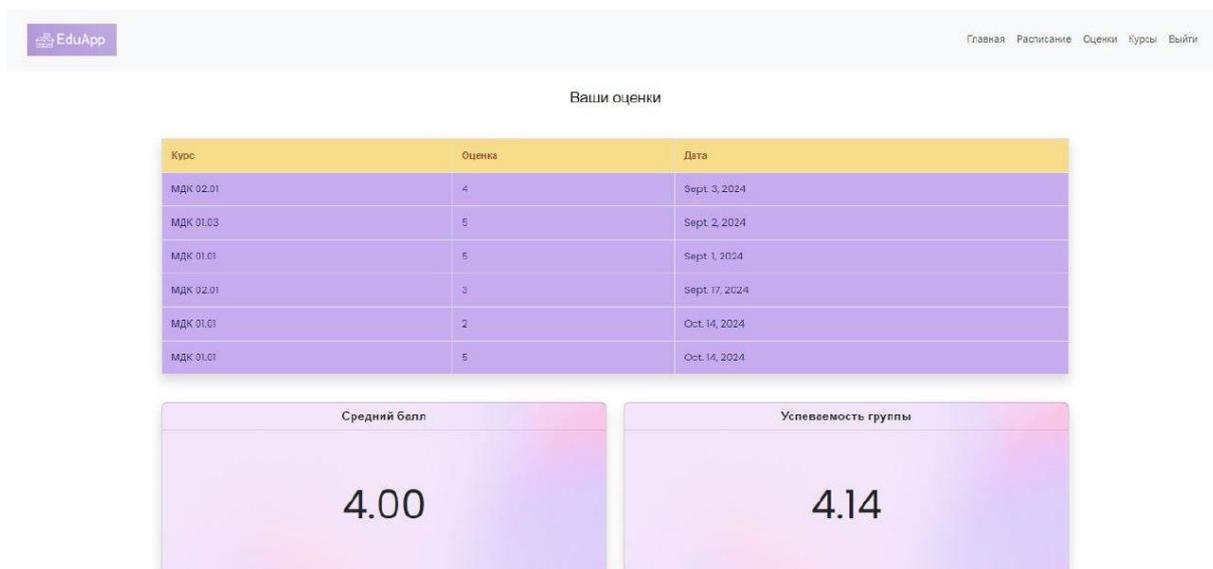


Рисунок 17 – Страница «Ваши оценки»

На этой странице представлены оценки студента по каждому курсу, а также успеваемость всей группы по выбранным дисциплинам. Страница отображает успеваемость студента и его группы по выбранному курсу.

На рисунке 18 продемонстрирована страница расписания студента.



Рисунок 18 – Расписание

На рисунке 19 представлен дашборд с основными аналитическими данными студента.

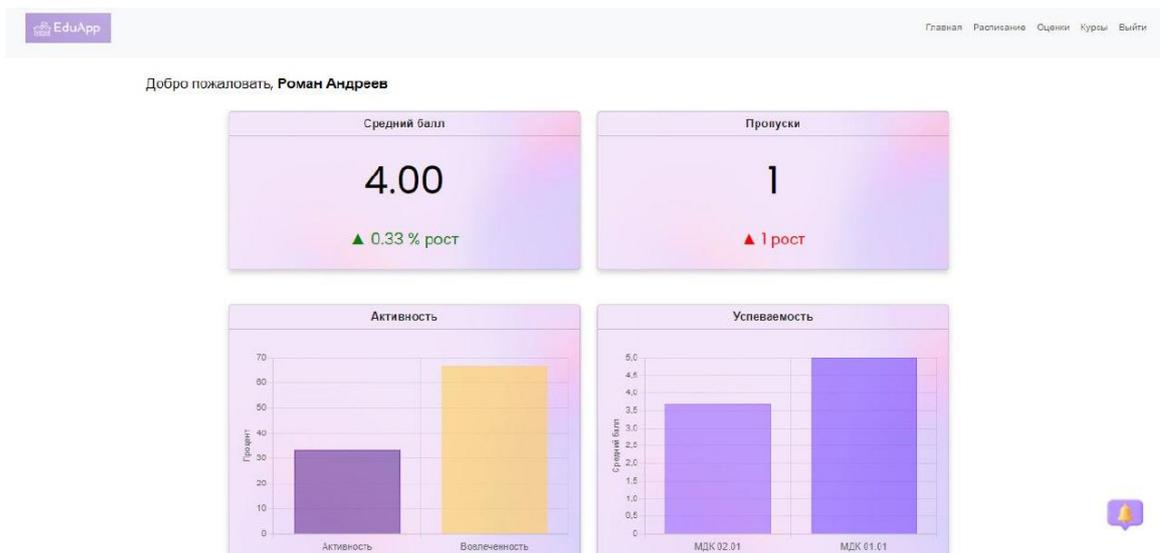


Рисунок 19 – Дашборд студента

На дашборде студент может отследить свой средний балл и количество пропусков занятий, а также их улучшение или ухудшение в разрезе календарного месяца, свою вовлеченность и активность по курсам, а также успеваемость по каждому из курсов. В правом нижнем углу расположена кнопка уведомлений.

На рисунке 20 продемонстрирована система уведомлений для студента, позволяющая отслеживать отметки о посещении и выставленные оценки.

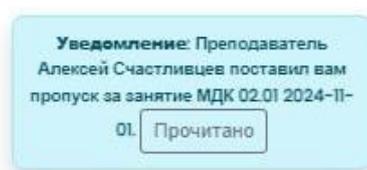


Рисунок 20 – Уведомления

На рисунке 21 представлен интерфейс личного кабинета преподавателя. На главной странице находится расписание занятий преподавателя.

Добро пожаловать, **Алексей Счастливец**

Курс	Дата	Время	Группа
МДК 01.01	Sept. 1, 2024	9 a.m.	ЗИТ
МДК 01.03	Sept. 2, 2024	10 a.m.	ЗИТ
МДК 01.03	Sept. 30, 2024	10:30 a.m.	ЗДИТ
МДК 01.03	Sept. 30, 2024	12:50 p.m.	ЗДИТ
МДК 01.03	Oct. 4, 2024	9 a.m.	ЧИП
МДК 01.03	Oct. 1, 2024	2:30 p.m.	ЧИП
МДК 01.01	Oct. 14, 2024	9 a.m.	ЗИТ
МДК 01.03	Oct. 14, 2024	11 a.m.	ЗИТ
МДК 01.03	Oct. 15, 2024	11 a.m.	ЗДИТ
МДК 01.01	Oct. 15, 2024	noon	ЗДИТ
МДК 01.01	Oct. 14, 2024	9 a.m.	ЗИС
МДК 01.03	Oct. 14, 2024	11 a.m.	ЗИС

Рисунок 21 – Личный кабинет преподавателя

На рисунке 22 представлена форма для выставления оценок.

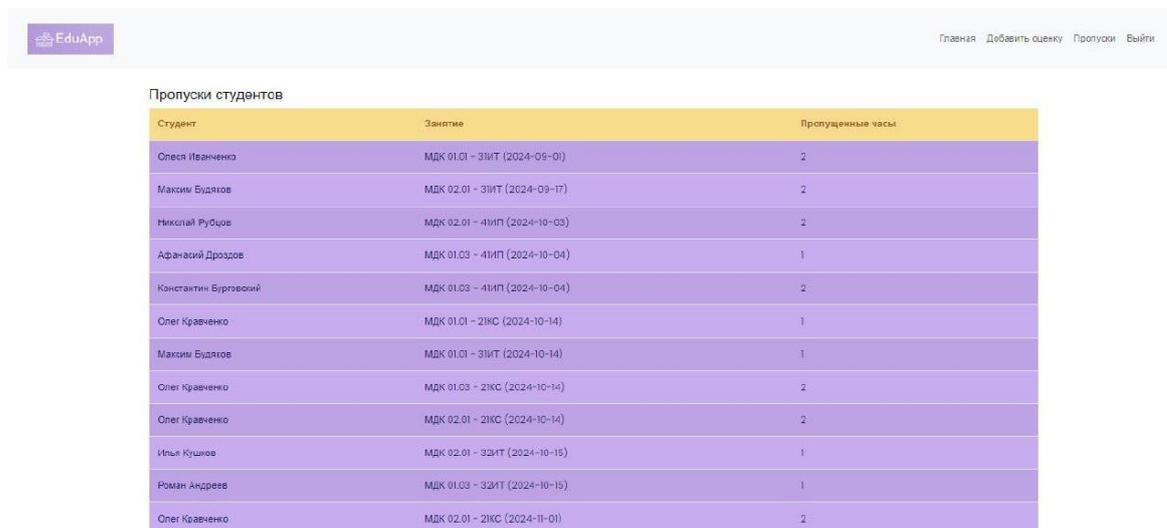
The screenshot shows a web form titled "Добавить оценку" (Add Score). It contains the following elements:

- Student:** A text input field with a dashed line indicating a placeholder.
- Lesson (Course and Date):** A text input field with a dashed line indicating a placeholder.
- Score:** A text input field with the label "Введите оценку" (Enter score).
- Save Score:** A green button labeled "Сохранить оценку" (Save score).
- Upload scores by file:** A section with the text "Загрузить оценки файлом" (Upload scores by file) and a file selection area showing "Выберите файл" (Choose file) and "файл не выбран" (file not selected).
- Upload File:** A purple button labeled "Загрузить файл" (Upload file).

Рисунок 22 – Форма «Добавить оценку»

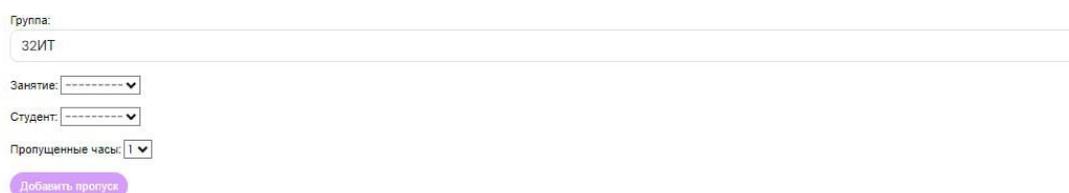
На данном рисунке представлена форма для преподавателей, с помощью которой они могут добавлять оценки студентов за выполненные задания, тесты или экзамены. Преподаватель выбирает нужный курс, студента и вводит оценку, что позволяет системе обновлять данные об успеваемости в реальном времени. Так же можно загружать оценки массово с помощью загрузки файла. Этот функционал облегчает работу преподавателей и минимизирует ошибки при внесении оценок.

На рисунках 23-24 представлена страница с пропусками студентов, а также форма для добавления пропуска.



Студент	Занятие	Пропущенные часы
Олеся Иванченко	МДК 01.01 – 3ИИТ (2024-09-01)	2
Максим Будаков	МДК 02.01 – 3ИИТ (2024-09-17)	2
Никлай Рубцов	МДК 02.01 – 4ИИП (2024-10-03)	2
Афанасий Дроздов	МДК 01.03 – 4ИИП (2024-10-04)	1
Константин Бурговский	МДК 01.03 – 4ИИП (2024-10-04)	2
Олег Кравченко	МДК 01.01 – 2ИКС (2024-10-14)	1
Максим Будаков	МДК 01.01 – 3ИИТ (2024-10-14)	1
Олег Кравченко	МДК 01.03 – 2ИКС (2024-10-14)	2
Олег Кравченко	МДК 02.01 – 2ИКС (2024-10-14)	2
Илья Кушков	МДК 02.01 – 3ИИТ (2024-10-15)	1
Роман Андреев	МДК 01.03 – 3ИИТ (2024-10-15)	1
Олег Кравченко	МДК 02.01 – 2ИКС (2024-11-01)	2

Рисунок 23 – Страница с пропусками



Группа: 32ИТ

Занятие: [dropdown]

Студент: [dropdown]

Пропущенные часы: 1 [dropdown]

Добавить пропуск

Рисунок 24 – Форма для добавления пропуска

Рисунок 21 демонстрирует интерфейс формы, с помощью которой преподаватели могут фиксировать пропуски студентов. Форма предоставляет возможность выбора студента, курса и даты пропуска. Информация о пропусках важна для дальнейшего анализа посещаемости и оценки влияния пропусков на успеваемость студента, а также для предоставления своевременных рекомендаций.

На рисунках 25-26 представлены формы для добавления информации для администрации.

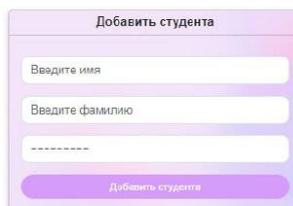


Рисунок 25 – Форма для добавления студента

Этот рисунок демонстрирует интерфейс формы для добавления нового студента в систему. Административный персонал вводит основные данные студента, такие как имя, фамилия, группа. После добавления студент автоматически включается в соответствующие курсы и получает доступ к системе. Это упрощает процесс управления студентами и ускоряет работу с данными.

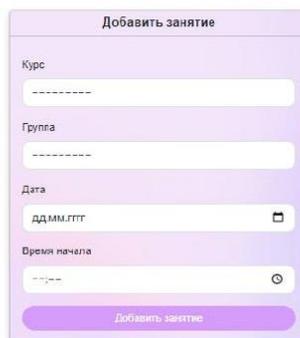


Рисунок 26 – Форма для добавления занятия

Форма на данном рисунке позволяет администрации добавлять новые учебные занятия в расписание. Административный персонал вводит информацию о предмете, группе и времени проведения занятия. Эти данные

становятся доступными для всех пользователей системы, что улучшает координацию учебного процесса и планирование занятий.

Система отчетов.

Отчеты являются важной частью системы образовательной аналитики. Они генерируются автоматически на основе данных, собранных системой, и предоставляют пользователям ключевые аналитические данные по курсам, успеваемости студентов и посещаемости. Форма для генерации отчетов продемонстрирована на рисунках 27-28.

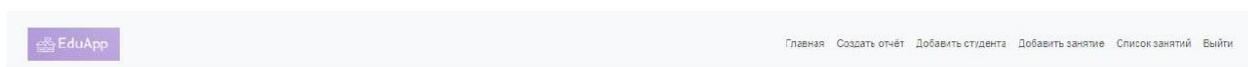
The image shows a form titled 'Создать отчет'. It contains a dropdown menu for 'Успеваемость всех групп по курсу', a dropdown for 'Дисциплина' with the value 'МДК 02.01', and two date input fields for 'Дата начала' and 'Дата конца', both with 'dd.mm.yyyy' placeholders. A 'Создать' button is at the bottom.

Рисунок 27 – Создание отчета по успеваемости студентов

На рисунке представлена форма для создания отчета по успеваемости студентов. Преподаватели могут выбрать период времени и получить отчет, содержащий информацию о распределении оценок, динамике успеваемости и общих показателях по студентам. Это помогает отслеживать прогресс группы и выявлять студентов, которым может понадобиться дополнительная помощь.

The image shows the same 'Создать отчет' form as in Figure 27, but with the 'Успеваемость всех групп по курсу' dropdown menu open. The menu lists four options: 'Выберите тип отчета', 'Успеваемость всех групп по курсу', 'Успеваемость группы по всем курсам', and 'Пропуски студентов'. The first option is highlighted in blue.

Рисунок 28 – Варианты выбора отчетов

Данный рисунок демонстрирует всплывающий список для выбора типа отчета. На рисунке 29 представлен пример отчета об успеваемости студентов с выгрузкой в MS Excel.

Предмет: МДК 01.01		
Роман Андреев: Нет оценок		
Илья Кушков: Нет оценок		
Ольга Стрыкало: Нет оценок		
Предмет: МДК 01.03		
Роман Андреев: Нет оценок		
Илья Кушков: Нет оценок		
Ольга Стрыкало: Нет оценок		
Предмет: МДК 02.01		
Роман Андреев		
	3	2024-09-27
	4	2024-09-30
	4	2024-09-30
Илья Кушков		
	5	2024-09-30
Ольга Стрыкало: Нет оценок		

Рисунок 29 – Отчет об успеваемости

Преподаватели и администрация могут видеть данные об успеваемости студентов за выбранный период, что помогает в анализе качества усвоения материала, а также в планировании индивидуальных образовательных траекторий для студентов, а так же в подборе персональных рекомендаций для неуспевающих студентов.

На рисунке 30 представлен дашборд для административного персонала, включающий основные показатели образовательного процесса.

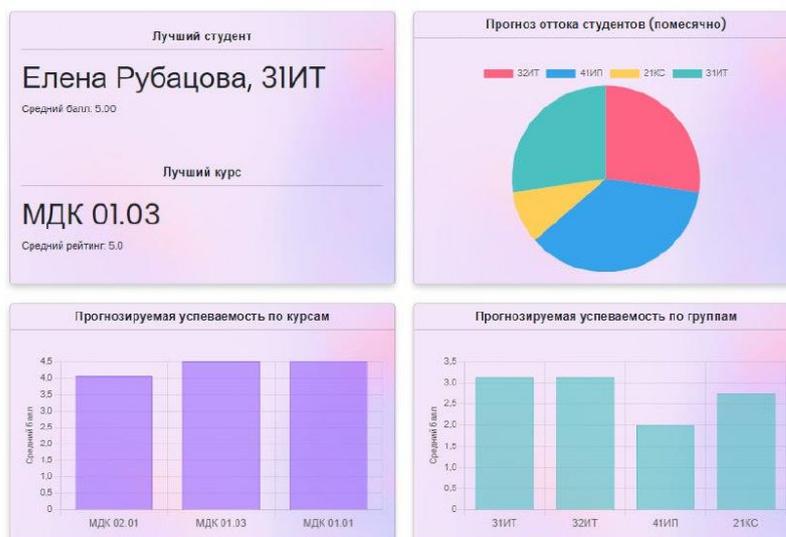


Рисунок 30 – Дашборд для административного персонала

На данном рисунке показан дашборд, предоставляющий ключевые показатели для административного персонала. На дашборде отображаются такие данные, как:

- лучший студент, на основе успеваемости и активности студентов;
- лучший курс на основе выставления личных оценок студентов;
- прогнозируемая успеваемость по курсам – на основе исторических данных и машинного обучения, отражающая ожидаем успеваемость групп по курсам;
- прогнозируемая успеваемость по группам – на основе исторических данных и машинного обучения общая успеваемость отдельных групп по всем учебным курсам;
- прогноз оттока студентов помесячно - на основе исторических данных и машинного обучения демонстрирует ожидаемый отток студентов.

Этот инструмент позволяет администрации быстро оценивать текущую ситуацию по успеваемости и принимать соответствующие управленческие решения.

3.4 Внедрение машинного обучения в систему для анализа данных

Внедрение технологий машинного обучения в систему образовательной аналитики направлено на автоматизацию процессов анализа данных и улучшение качества прогнозирования.

Для реализации моделей машинного обучения использован язык Python, который предоставляет широкий набор библиотек для этой цели, представленных на рисунке 31.

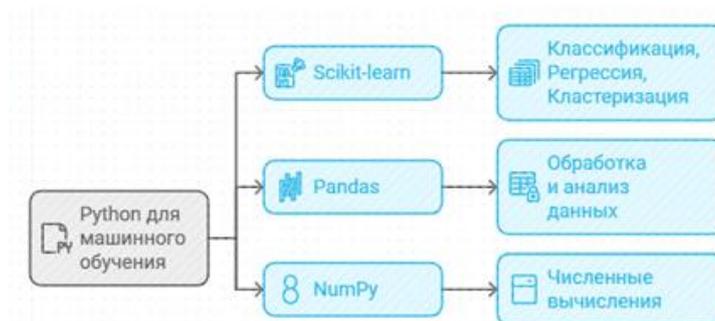


Рисунок 31 – Реализация моделей машинного обучения

Для реализации моделей машинного обучения на Python используется богатый экосистемный набор библиотек, каждая из которых решает специфические задачи в процессе разработки, обучения и анализа моделей. Опишем подробно процесс машинного обучения в разрабатываемой системе с использованием перечисленных библиотек.

Подготовка данных (рис. 32).

```

# 3. Подготовка данных
def prepare_data_for_model(df_students):
    df = df_students.dropna(subset=["activity_percentage", "average_grade"])

    if df.empty:
        raise ValueError("Данные не содержат допустимых значений для обучения")

    # Признаки (activity_percentage) и целевая переменная (average_grade)
    X = df[["activity_percentage"]]
    y = df["average_grade"]

    if len(X) == 0 or len(y) == 0:
        raise ValueError("Нет данных для обучения модели")

    # Разделяем данные на тренировочную и тестовую выборки
    X_train, X_test, y_train, y_test = train_test_split(
        X, y, test_size=0.3, random_state=42
    )

    if len(X_train) == 0 or len(y_train) == 0:
        raise ValueError("Тренировочный набор пуст")

    return X_train, X_test, y_train, y_test

```

Рисунок 32 – Подготовка данных

Прежде чем построить модель, необходимо правильно подготовить данные. Для этого используются библиотеки Pandas и NumPy:

- Pandas: библиотека для работы с табличными данными, позволяет эффективно загружать, очищать, фильтровать и преобразовывать данные. Основной структурой данных в Pandas является DataFrame, который упрощает манипуляции с данными, такими как их очистка, заполнение пропусков, выборка по условиям, объединение различных таблиц и группировка;
- NumPy: основная библиотека для численных вычислений в Python, особенно полезна для работы с многомерными массивами. NumPy обеспечивает эффективную работу с массивами данных и матричными операциями, которые являются основой многих алгоритмов машинного обучения.

Построение и обучение моделей. После подготовки данных с помощью Scikit-learn можно приступить к построению и обучению моделей. Scikit-learn предлагает разнообразные инструменты для реализации алгоритмов классификации, регрессии и кластеризации.

Scikit-learn: одна из самых популярных библиотек для машинного обучения в Python. Она предоставляет готовые к использованию алгоритмы, такие как логистическая регрессия, случайный лес, SVM (Support Vector Machine) и KMeans, а также инструменты для предварительной обработки данных, оценки моделей и поиска гиперпараметров.

Построение и проверка модели (рис. 33).

```
# 6. Обучение
def train_and_predict_for_data(X_input, df_students):
    # Подготовка обучающих и тестовых выборок из DataFrame студентов
    X_train, X_test, y_train, y_test = prepare_data_for_model(df_students)

    # Проверка, пуст ли обучающий набор данных
    if X_train.empty or y_train.empty:
        raise ValueError("Обучающий набор данных пуст. Невозможно обучить модель.")

    # Проверка, пуст ли входной набор данных для предсказания
    if X_input.empty:
        raise ValueError("Входные данные для предсказания пусты.")

    # Проверка, совпадает ли количество признаков во входных данных и данных для обучения
    if X_input.shape[1] != X_train.shape[1]:
        raise ValueError(
            "Несоответствие размерностей между входными данными и данными для обучения"
        )
    )
```

Рисунок 33 – Построение и проверка модели

Анализ данных. После построения модели проанализируем данные, представим результаты их анализа и работу модели (рис. 34, 35).

```

# 4. Прогнозирование оттока студентов
def get_high_risk_students(df_students, group_filter=None):
    # Применение фильтра по группе, если указан
    if group_filter:
        df_students = df_students[df_students["group"] == group_filter]

    # Расчёт риска отчисления для каждого студента
    high_risk_students = [
        student
        for _, student in df_students.iterrows()
        if predict_dropout_for_student(student) == 1
    ]

    return high_risk_students

def predict_dropout_for_student(student):
    # Подсчёт риска отчисления на основе текущих данных студента
    average_grade = student["average_grade"]
    activity_percentage = student["activity_percentage"]

    # Определение риска отчисления (низкий балл или низкая активность)
    if average_grade < 3.0 or activity_percentage < 50:
        return 1
    return 0

```

Рисунок 34 – Прогнозирование оттока студентов

```

# 5. Прогнозирование успеваемости
def predict_course_performance(df_grades, df_students):
    course_performance = {}
    unique_courses = df_grades["course"].unique()
    for course_name in unique_courses:
        course_data = df_grades[df_grades["course"] == course_name]
        if course_data.empty:
            continue

        student_names = course_data["full_name"].unique()
        relevant_students = df_students[df_students["full_name"].isin(student_names)]
        X = relevant_students[["activity_percentage"]]
        y_pred = train_and_predict_for_data(X, df_students)
        course_performance[str(course_name)] = float(y_pred.mean())
    return course_performance

def predict_group_performance(df_students):
    group_performance = {}
    unique_groups = df_students["group"].unique()
    for group_name in unique_groups:
        group_data = df_students[df_students["group"] == group_name]
        if group_data.empty:
            continue

        X = group_data[["activity_percentage"]]
        y_pred = train_and_predict_for_data(X, df_students)
        group_performance[str(group_name)] = (float(y_pred.mean()) if y_pred.size > 0 else None)
    return group_performance

```

Рисунок 35 – Прогнозирование успеваемости

На рисунках 34-35 продемонстрированы алгоритмы для построения прогнозируемых данных по успеваемости и оттоку студентов.

Интерпретация результатов и выводы. После обучения модели на данных студентов (оценки, пропуски и другие показатели) с использованием

библиотеки Scikit-learn, важно правильно интерпретировать результаты. Визуализация с помощью Chart.js позволяет наглядно представить, как модель оценивает прогнозируемую успеваемость и вероятность оттока студентов.

Анализируя эти данные, можно выделить наиболее важные факторы, влияющие на академические результаты, и обнаружить области, где модель допускает ошибки.

Описанные на рисунках 34-35 алгоритмы по прогнозированию позволяют сформировать интерактивные настраиваемые плитки дашбордов для прогнозной аналитики по ожидаемой успеваемости и оттоку студентов, а также отчет по получению персональных рекомендаций. Плитки представлены на рисунках 36-38.



Рисунок 36 – Настраиваемая плитка по прогнозируемой успеваемости

Плитка, продемонстрированная на рисунке 36, позволяет просмотреть ожидаемую среднюю успеваемость по каждому курсу.



Рисунок 37 – Настраиваемая плитка по прогнозируемой успеваемости по группам

Плитка, продемонстрированная на рисунке 37, позволяет просмотреть ожидаемую среднюю успеваемость по каждой группе.



Рисунок 38 – Настраиваемая плитка по прогнозируемому оттоку студентов

Плитка, продемонстрированная на рисунке 38, позволяет просмотреть ожидаемый отток студентов по каждой группе.

Показанные диаграммы демонстрируют ключевые возможности системы образовательной аналитики. Эти отчеты позволяют администрации и преподавателям своевременно получать прогнозную аналитику по успеваемости и оттоку студентов. Визуализация и экспорт данных в удобные форматы делают систему гибкой и полезной для практического использования в управлении образовательным процессом.

3.5 Тестирование и отладка системы

При создании информационной системы образовательной аналитики, внедрении ее в общую образовательную среду было проведено тестирование ее работы, так как система должна полностью удовлетворять поставленным к ней требованиям и при необходимости откорректировать ее работу. Было проведено тестирование программы на правильность по ряду функциональных и нефункциональных требований. Этот процесс включает в себя несколько ключевых этапов.

Тестирование функциональности. На данном этапе проверялась работа всех функций системы, включая:

- пользовательский интерфейс – проверка доступности всех элементов интерфейса, корректность их отображения и функциональности. Использование методов юзабилити-тестирования для обеспечения удобства работы пользователей;
- системные функции - проверка всех функций системы, включая сбор, хранение и обработку данных, работу аналитических модулей и механизмов машинного обучения;
- интеграция с другими системами - проверка взаимодействия системы образовательной аналитики с другими программными решениями, такими как системы управления обучением (LMS) и базы данных.

Нагрузочное тестирование. Для обеспечения надежности и устойчивости системы было проведено нагрузочное тестирование, которое позволило оценить производительность системы при различных условиях эксплуатации. На этом этапе происходили:

- моделирование нагрузок - создание сценариев с различным количеством пользователей, данных и запросов для выявления пределов производительности;

- мониторинг системы – оценка времени отклика, скорости обработки данных и стабильности работы системы при высоких нагрузках.

Тестирование моделей машинного обучения. Модели машинного обучения были протестированы на предмет их точности и надежности на реальных данных. Применение моделей к новым данным, которые не использовались в процессе обучения, для оценки их способности к обобщению.

Отладка. На этапе отладки были выявлены и устранены все ошибки и недостатки, обнаруженные в процессе тестирования. Этот процесс включал следующее:

- анализ логов – просмотр логов системы для выявления ошибок и нестандартных ситуаций.
- исправление ошибок – внесение изменений в код или архитектуру системы для устранения выявленных проблем.
- повторное тестирование – проведение повторного тестирования после внесения исправлений для проверки их эффективности и отсутствия новых проблем.

В таблице 7 подведены итоги по всем этапам тестирования системы образовательной аналитики.

Таблица 7 – Результаты тестирования системы образовательной аналитики

Этап тестирования	Цель	Результаты
Тестирование функциональности	проверка работы интерфейса и системных функций	интерфейс пользователя: все элементы отображаются и функционируют корректно, системные функции: сбор, хранение и обработка данных работают без сбоев, интеграция: корректное взаимодействие с LMS.
Нагрузочное тестирование	оценка производительности при высоких нагрузках	среднее время отклика: 1,2 секунды (до 200 пользователей), 3,5 секунды (1000 пользователей), система функционировала стабильно при максимальной нагрузке.

Продолжение таблицы 7

Этап тестирования	Цель	Результаты
Тестирование моделей машинного обучения	оценка точности и надежности моделей	классификация: точность 87% (Random Forest), регрессия: MAE 2,3 балла, кластеризация: выделение 3 групп студентов с различными характеристиками успеваемости
Отладка	выявление и устранение ошибок	обнаружены незначительные ошибки отображения графиков, исправлены, логи: не выявлено критических сбоев или ошибок.

Рассмотрим каждый из этапов тестирования подробнее.

Выделив критерии для проведения тестирования системы, проанализируем полученные результаты.

Тестирование функциональности:

- интерфейс пользователя – все элементы интерфейса корректно отображаются и функционируют без сбоев, обеспечивая удобную навигацию, редактирование данных и генерацию отчетов:
- функционал системы – данные корректно сохраняются в базе PostgreSQL. Все аналитические функции работают, включая отчеты по успеваемости и визуализацию. Система успешно взаимодействует с внешними системами, такими как LMS.

Нагрузочное тестирование. В рамках нагрузочного тестирования система была протестирована на 1000 одновременных пользователей. Время отклика представлено на рисунке 39.



Рисунок 39 – Производительность системы при нагрузочном тестировании

Тестирование на 1000 одновременных пользователей показало следующие результаты. При нагрузке до 200 пользователей время отклика

составило в среднем 1,2 секунды. При максимальной нагрузке (1000 пользователей) время отклика увеличилось до 3,5 секунд. Система работала стабильно, что подтверждает ее устойчивость к высоким нагрузкам.

Результаты тестирования моделей машинного обучения представлены на рисунке 40.



Рисунок 40 – Тестирование моделей машинного обучения

Рассмотрим результаты тестирования:

- модель классификации (Random Forest) показала точность 87%;
- модель линейной регрессии продемонстрировала среднюю абсолютную ошибку 2,3 балла, что является приемлемым для ОА;
- кластеризация выявила 3 группы студентов с различной успеваемостью для целевой поддержки.

Отладка:

- были устранены незначительные ошибки отображения графиков в некоторых браузерах путем обновления библиотек визуализации;
- логи системы не выявили критических ошибок.

Система образовательной аналитики продемонстрировала высокую эффективность, стабильность и точность. Все функции работают корректно, и система успешно справляется с высокой нагрузкой. Выявленные ошибки устранены, и обучение пользователей прошло успешно.

Выводы по главе

В процессе разработки системы образовательной аналитики для повышения качества образовательного процесса были достигнуты следующие основные результаты:

- обоснован выбор PostgreSQL, Django и Python как основных компонентов, эти инструменты обеспечивают надежное хранение данных, гибкость разработки и мощные возможности для аналитической обработки;
- разработка архитектуры системы – предложена многослойная архитектура, включающая централизованное хранилище данных, автоматизированные процессы сбора, модули обработки данных в реальном времени и средства визуализации. Такая архитектура гарантирует гибкость, модульность и масштабируемость системы, поддерживая работу с большими объемами данных;
- применение машинного обучения и визуализации – использование МЛ позволило внедрить мощные инструменты для анализа и построения моделей. Визуализация результатов помогает пользователям принимать более обоснованные решения;
- тестирование и отладка системы – проведено функциональное и нагрузочное тестирование, а также проверка точности моделей машинного обучения. Эти этапы подтвердили надежность и производительность системы, а выявленные проблемы были успешно устранены.

Таким образом, разработаны основные компоненты системы образовательной аналитики, способные улучшить качество образовательного процесса за счет анализа данных и прогнозирования. Система обладает гибкостью и может быть адаптирована для различных образовательных учреждений, что делает ее универсальной и масштабируемой.

4 Оценка результатов исследования и экспериментальная апробация системы образовательной аналитики

4.1 Экспериментальная апробация системы

В рамках экспериментальной апробации, система аналитики была подвергнута практическому тестированию в условиях учебного заведения СПб ГБ ПОУ «Колледж электроники и приборостроения» (КЭП). Это позволило оценить ее работоспособность и полезность на практике.

Исследование направлено на оценку различных аспектов системы, таких как: способность предсказывать успехи учащихся, предоставление персонализированных советов, скорость обработки информации и эффективность представления данных в удобной форме.

Кроме того, в процессе апробации участвовали не только преподаватели и администрация, но и сами студенты, что дало возможность выявить насколько проста и удобна система в ежедневном использовании, а также как она влияет на учебный процесс в целом.

Чтобы достичь объективности в оценке полезности системы ОА, были определены ключевые параметры ее эффективности: прогностическая точность, скорость обработки информации, персональные рекомендации, снижение оттока студентов, эффективность визуализации.

Прогностическая точность – служит ключевым критерием для оценки способности системы анализировать и прогнозировать академические результаты учащихся. Для оценки этого параметра используются различные метрики, включая RMSE, R^2 и точность в классификации.

Такой подход способствует своевременному выявлению потенциальных проблем в учебном процессе, позволяя предотвратить снижение успеваемости путем оказания необходимой поддержки студентам, находящимся в «зонах риска». Основная цель состоит в достижении высокой точности предсказаний,

что, в свою очередь, облегчает оперативное вмешательство и помощь студентам с плохими показателями успеваемости.

Скорость обработки информации и предоставления соответствующих отчетов – является важным аспектом, влияющим на эффективность принятия образовательных решений.

Параметр включает в себя времени ответа системы и время создания отчетов, что критически важно для эффективного управления огромным массивом информации, как для преподавателей, так и для административного состава.

Персональные рекомендации – данный аспект заключается в настройке системы таким образом, чтобы она могла автоматически предлагать студентам индивидуализированные образовательные пути, основываясь на их достижениях, увлечениях и посещаемости занятий. Ключевым показателем тут является положительная оценка студентами этих рекомендаций и их влияние на повышение учебных достижений, что демонстрирует эффективность персонализации обучения. В Приложении А представлен шаблон анкеты для оценки персональных рекомендаций.

Снижение оттока студентов – оценка количества студентов, продолжающих обучение в образовательном учреждении благодаря внедренным мерам поддержки, отражает эффективность системы в предотвращении их отчислений. Данный параметр достигается за счет идентификации обучающихся с проблемами в успеваемости и активности, а также предложения стратегий для повышения их результативности.

Важными показателями здесь являются сокращение доли отчисленных студентов и повышение общего уровня успеваемости через осознанные усилия по оказанию помощи.

Эффективность визуализации – параметр заключается в оценивании доступности и практичности использования интерактивных элементов, таких как графики и панели управления. Эти инструменты способствуют упрощению процесса анализа оценок и общих трендов в успеваемости, делая

информацию более наглядной и понятной. В Приложении Б представлен шаблон анкеты для оценки эффективности визуализации системы.

Процесс апробации включает несколько ключевых шагов. На первоначальном этапе происходит разделение студентов на две отдельные группы (контрольную и экспериментальную). Одна будет оцениваться по классическим методикам, в то время как другая будет подвергнута оценке с применением аналитической системы.

Далее, для обеспечения эффективности проведения эксперимента, производится техническая интеграция новой системы аналитики с уже существующими образовательными и административными системами, включая управление учебным процессом, а также базы данных о результатах учебы и посещаемости занятий студентами.

Оценка точности прогноза. Анализ точности прогнозов на основе данных успеваемости и активности студентов, что позволяет своевременно выявлять нуждающихся в поддержке.

Тестирование оперативности и визуализации. Оценка времени отклика системы и удобства интерфейсов для преподавателей и администрации, а также обратная связь по визуализации и отчетам.

Экспериментальная апробация показала, что система образовательной аналитики способствует улучшению образовательного процесса, предоставляя возможности для персонализированного подхода к каждому студенту и оперативного принятия решений на основе точных данных.

4.2 Результаты экспериментальной апробации

Экспериментальная апробация системы образовательной аналитики была направлена на оценку эффективности ее ключевых модулей, включая прогнозирование успеваемости, оперативность принятия решений, персонализацию рекомендаций и визуализацию данных. Оценка проводилась на основе критериев, изложенных в пункте 4.1, и позволила получить

значимые данные о влиянии системы на образовательный процесс. Рассмотрим основные результаты апробации по каждому параметру.

Точность прогноза (рис. 41).

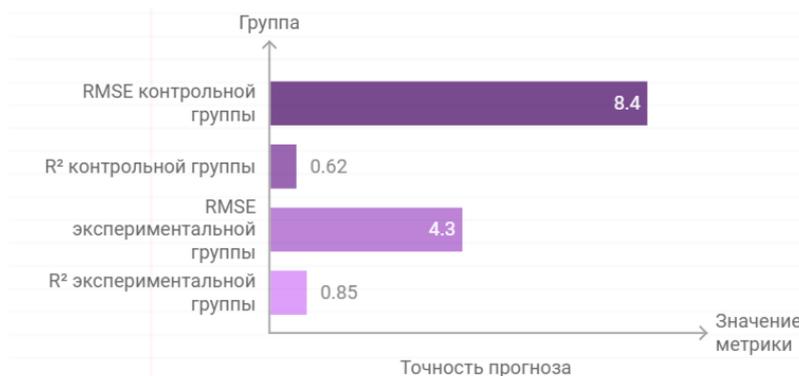


Рисунок 41 – Оценка точности прогноза

Оценим результаты:

- контрольная группа: традиционные методы демонстрировали низкую предсказательную точность, с RMSE 8.4 и коэффициентом детерминации (R^2) 0.62;
- экспериментальная группа: система образовательной аналитики значительно улучшила показатели прогноза, снизив RMSE до 4.3 и увеличив R^2 до 0.85, что подтверждает её эффективность в своевременной идентификации студентов, нуждающихся в поддержке.

Оперативность принятия решений (рис. 42).

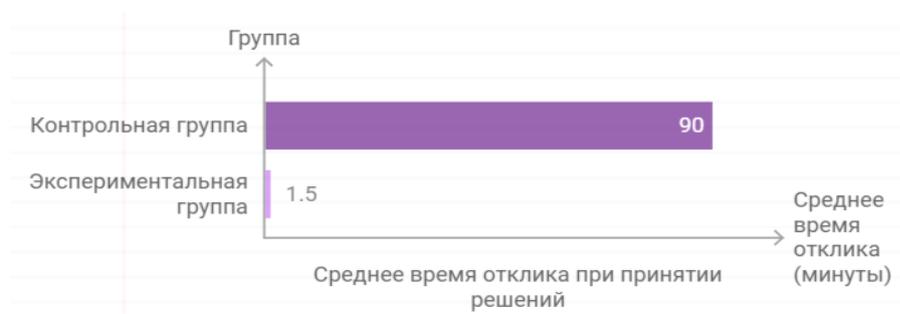


Рисунок 42 – Оценка оперативности принятия решений

Оценим результаты:

– контрольная группа: среднее время подготовки отчётов и анализа данных составляло 1,5 часа, что было недостаточным для оперативного принятия решений;

– экспериментальная группа: время отклика системы сократилось до 1,5 минут, что позволило преподавателям и администрации быстрее реагировать на изменения в успеваемости студентов, значительно повысив эффективность управления учебным процессом.

Персонализация рекомендаций (рис. 43).

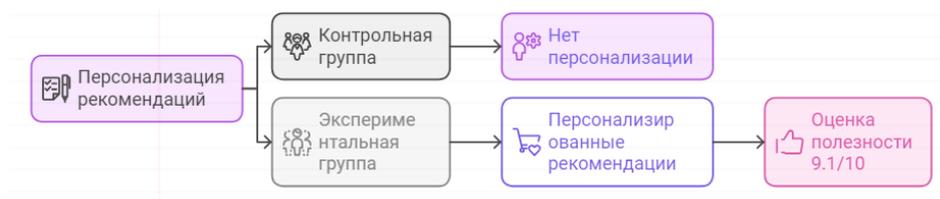


Рисунок 43 – Оценка персонализации рекомендаций

Оценим результаты:

– контрольная группа: традиционные методы не предоставляли возможности персонализации учебных траекторий для студентов;

– экспериментальная группа: система аналитики предоставила индивидуальные рекомендации, основанные на данных каждого студента. Оценка полезности рекомендаций по обратной связи студентов составила 9.1 из 10 баллов, указывая на высокую удовлетворенность и эффективность персонализированных рекомендаций.

Снижение оттока студентов (рис. 44).

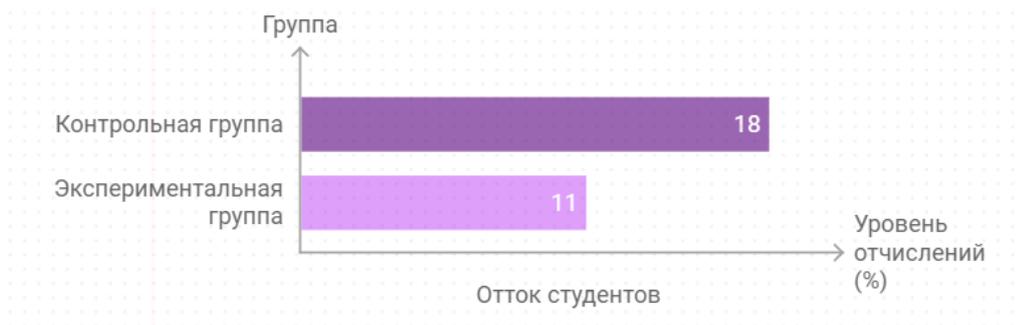


Рисунок 44 – Оценка оттока студентов

Оценим результаты:

- контрольная группа: процент отчислений оставался на уровне 18%;
- экспериментальная группа: система аналитики позволила снизить отток до 11%, что демонстрирует ее способность своевременно выявлять студентов из группы риска и предоставлять им поддержку для продолжения обучения.

Эффективность визуализации данных (рис. 45).



Рисунок 45 – Оценка эффективности визуализации данных

Оценим результаты:

- контрольная группа: визуализация данных была оценена на 5.2 из 10 баллов, указывая на сложность восприятия информации и низкую эффективность;
- экспериментальная группа: визуализация данных в новой системе получила оценку 9.0 из 10 баллов, что подтверждает её удобство и полезность для быстрого анализа и принятия управленческих решений.

Сравнительные результаты по контрольной и экспериментальной группам продемонстрированы в таблице 8.

Таблица 8 – Сравнительные результаты экспериментальной и контрольной групп по ключевым критериям

Критерий	Контрольная группа	Экспериментальная группа	Изменение (%)
Точность прогноза	RMSE: 8.4, R ² : 0.62	RMSE: 4.3, R ² : 0.85	+48%
Оперативность (время отклика)	1,5 часа	1,5 минуты	+98%
Персонализация рекомендаций	нет	оценка 9.1 из 10	+91%
Снижение оттока студентов (%)	18%	11%	-7%
Эффективность визуализации (оценка)	5.2 из 10	9.0 из 10	+73%

В ходе апробации системы образовательной аналитики были собраны данные, которые позволяют оценить эффективность системы в прогнозировании успеваемости студентов и предоставлении персонализированных рекомендаций. Эти данные представлены в таблицах 9 и 10.

Таблица 9 – Прогнозирование успеваемости студентов

Студент	Оценка по предыдущим периодам	Прогнозируемая оценка	Фактическая оценка	Разница (Ошибка прогноза)
Будяков М.	4.0, 4.5, 5.0	4.8	5.0	0.2
Андреев Р.	3.5, 3.0, 3.8	3.7	3.6	0.1
Кушков И.	2.8, 3.0, 2.9	3.1	3.0	0.1
Бойков Е.	4.8, 5.0, 5.0	4.9	5.0	0.1
Попков Д.	2.0, 2.5, 2.8	2.6	2.4	0.2
Поповская А.	4.0, 3.8, 4.2	4.1	4.0	0.1
Волков П.	3.2, 3.5, 3.3	3.6	3.5	0.1
Еремин С.	4.5, 4.7, 4.8	4.7	4.6	0.1
Иванова В.	3.0, 2.8, 3.2	3.3	3.1	0.2
Ильин И.	3.8, 4.0, 3.9	4.0	3.9	0.1

В таблице 9 представлены данные о прогнозировании успеваемости студентов. Прогнозируемые оценки соответствуют фактическим с

минимальной погрешностью (в среднем, ошибка составляет 0.1-0.2 балла), что доказывает высокую точность алгоритмов прогнозирования, внедренных в систему образовательной аналитики. Прогнозируемые и фактические оценки представлены на графике (рис. 46).

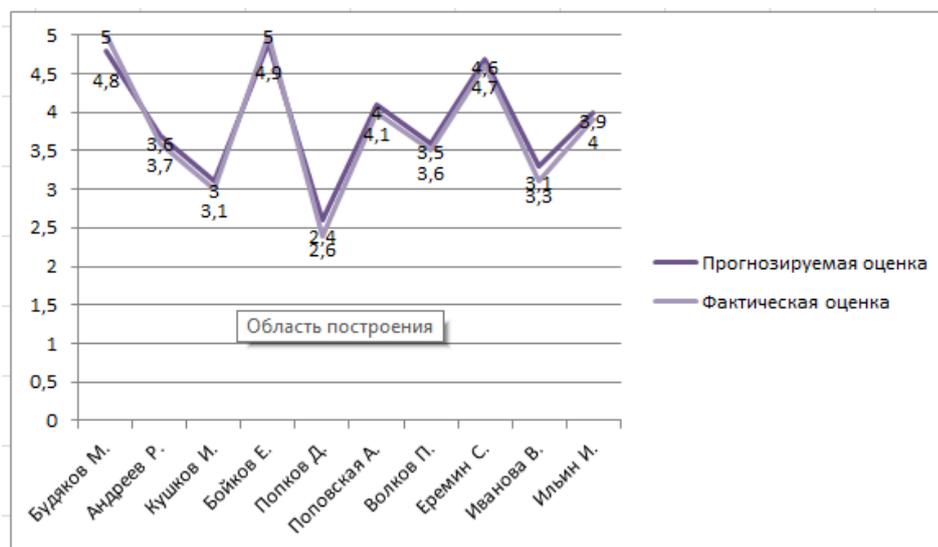


Рисунок 46 – Прогнозируемые и фактические оценки студентов

В таблице 10 представлены персональные рекомендации по улучшению качества знаний по проблемным модулям, а также фактическое улучшение оценки студентов за модули после применения данных рекомендаций.

Таблица 10 – Персональные рекомендации по улучшению успеваемости

Студент	Проблемный модуль	Рекомендации (дополнительные материалы)	Фактическое улучшение успеваемости (баллы)
Будяков М.	МДК 01.01 «Разработка программных модулей»	Видеоуроки по основам ООП	+0.1
Андреев Р.	МДК 01.03 «Разработка мобильных приложений»	Практические задания по разработке UI	+0.6
Кушков И.	МДК 02.01 «Технология разработки ПО»	Дополнительная литература по методологиям Agile	+0.5

Продолжение таблицы 10

Бойков Е.	МДК 01.03 «Разработка мобильных приложений»	Вебинары по использованию Flutter	+0.7
Попков Д.	МДК 01.01 «Разработка программных модулей» ²	Онлайн-курсы по структурам данных	+0,2
Поповская А.	МДК 02.01 «Технология разработки ПО»	Лекции по UML и моделированию	+0.6
Волков П.	МДК 01.01 «Разработка программных модулей»	Практические задачи по алгоритмам	+0.7
Еремин С.	МДК 01.03 «Разработка мобильных приложений»	Интерактивные упражнения по работе с API	+0.5
Иванова В.	МДК 02.01 «Технология разработки ПО»	Лабораторные работы по тестированию	+0.4
Ильин И.	МДК 01.03 «Разработка мобильных приложений»	Курсы по кроссплатформенной разработке	+0.6

В таблице 10 продемонстрировано, как система образовательной аналитики предоставляет персонализированные рекомендации для студентов по различным дисциплинам. Например, Будякову М. предложены видеоуроки по объектно-ориентированному программированию для повышения результатов по модулю «Разработка программных модулей», а Попков Д. получил онлайн-курсы по структурам данных, что привело к улучшению успеваемости на 0,2 балла. В среднем, улучшение успеваемости после внедрения рекомендаций составило от 0.4 до 0,6 балла.

График улучшения оценки после применения персональных рекомендаций представлен на рисунке 47.

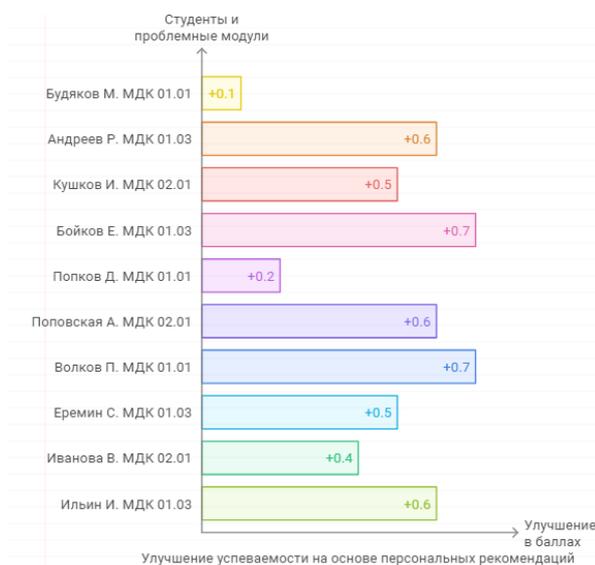


Рисунок 47 – Улучшение успеваемости на основе персональных рекомендаций

В целом, результаты экспериментальной апробации подтвердили, что внедрение системы образовательной аналитики значительно улучшает ключевые аспекты образовательного процесса: от повышения точности прогнозирования успеваемости студентов и персонализации рекомендаций до уменьшения уровня отчислений. Это подчеркивает значимость системы для образовательных учреждений, ориентированных на повышение качества обучения и оперативность управленческих решений.

4.3 Оценка возможных улучшений и предложений

Хотя система образовательной аналитики показала успешные результаты, апробация выявила несколько направлений, по которым ее функциональность и эффективность можно улучшить:

- оптимизация алгоритмов прогнозирования - текущие алгоритмы прогнозирования демонстрируют высокую точность, но ее можно улучшить за счет включения дополнительных факторов, таких как социально-экономические условия (уровень дохода семьи, место проживания) и результаты психологических тестов, которые отражают

мотивацию, стрессоустойчивость и эмоциональный интеллект студентов, что позволит точнее определять студентов «группы риска» и своевременно предоставлять целевые рекомендации по поддержке;

– улучшение функционала персонализированных рекомендаций – хотя система уже формирует персонализированные учебные траектории, углубленный анализ поведения студентов в учебной деятельности (участие в проектах, командной работе и других активностях) поможет эффективнее адаптировать рекомендации. Важно также включить механизм обратной связи для оценки эффективности данных рекомендаций и их динамической корректировки на основе новых данных;

– обучение пользователей – в ходе апробации выявлено, что преподаватели и администрация не всегда полностью используют потенциал системы. Разработка обучающих материалов, интерактивных руководств и проведение семинаров помогут пользователям лучше освоить функционал системы и эффективнее применять аналитические инструменты в повседневной работе;

– повышение безопасности данных – с учетом работы с персональными данными студентов необходимо усилить меры по защите информации, рекомендуется внедрить расширенное шифрование данных и двухфакторную аутентификацию для минимизации рисков несанкционированного доступа и утечки информации;

– оптимизация пользовательского интерфейса – улучшение интерфейса и добавление настраиваемых панелей облегчит навигацию по системе для преподавателей и административного персонала, повышая скорость и эффективность использования аналитики для принятия оперативных решений.

Эти предложения направлены на повышение адаптивности и безопасности системы, улучшение взаимодействия пользователей с

аналитическими данными и, как следствие, на дальнейшее повышение качества образовательного процесса.

4.4 Перспективы дальнейших исследований

Разработка системы образовательной аналитики не только подтвердила свою практическую значимость, но и создала перспективы для будущих научных исследований и технологического прогресса:

- расширение применения машинного обучения – внедрение более сложных моделей, таких как нейронные сети и методы глубокого обучения, позволит анализировать сложные зависимости между факторами успеваемости, что откроет возможности для прогнозирования не только академических результатов, но и долгосрочных образовательных траекторий, включая карьерные перспективы студентов;
- включение анализа психологических и социальных факторов – включение данных о социально-психологических характеристиках студентов, таких как уровень мотивации, эмоциональное состояние и социальная вовлеченность, поможет точнее прогнозировать результаты. Это направление требует дальнейших исследований в области интеграции психологических тестов и моделей анализа поведения в учебном процессе;
- адаптация системы для других уровней образования – возможность применения системы не только в высшем образовании, но и в среднем и начальном открывает перспективы для универсализации аналитики. Исследование потребностей различных образовательных уровней позволит модифицировать алгоритмы и подходы для улучшения качества образования на всех этапах;
- исследование влияния системы на преподавателей – дополнительные исследования могут быть направлены на изучение

влияния системы не только на студентов, но и на преподавателей. Это включает оценку эффективности управления учебным процессом, снижение нагрузки на преподавателей через автоматизацию рутины и повышение качества принимаемых решений. Важно изучить, как система может способствовать повышению профессионального развития педагогов и их удовлетворенности рабочим процессом;

– разработка интеллектуальных помощников на базе искусственного интеллекта – в перспективе можно интегрировать в систему интеллектуальных ассистентов, которые помогут пользователям (студентам, преподавателям и администрации) ориентироваться в данных и принимать решения. Такой ассистент мог бы помогать преподавателям и администрации оперативно анализировать данные по успеваемости, а студентам – давать советы по учебному процессу в режиме реального времени.

– анализ влияния системы на мотивацию и вовлеченность студентов – важно также исследовать, как система аналитики влияет на мотивацию студентов и их вовлеченность в учебный процесс. Дополнительные метрики могут включать уровень посещаемости, активность в учебных дискуссиях, выполнение дополнительных заданий, что позволит улучшить систему рекомендаций и мотивировать студентов к лучшим результатам.

Выводы по главе

В результате апробации системы образовательной аналитики и проведенных экспериментов были получены важные данные, подтверждающие ее эффективность.

Система продемонстрировала значительные улучшения в прогнозировании успеваемости студентов, скорости принятия решений, персонализации рекомендаций и визуализации данных по сравнению с

традиционными методами, существовавшими до внедрения системы образовательной аналитики:

- точность прогнозов успеваемости – среднеквадратическая ошибка (RMSE) снизилась с 8.4 до 4.3, а коэффициент детерминации (R^2) вырос с 0.62 до 0.85, что подтверждает способность системы своевременно выявлять студентов, нуждающихся в поддержке;
- оперативность принятия решений – время отклика сократилось с 1,5 часов до 1,5 минут, что значительно повысило эффективность управления учебным процессом и дало возможность быстрее реагировать на изменения в успеваемости студентов;
- персонализация рекомендаций – оценка полезности рекомендаций студентами составила 9.1 из 10, что подтверждает эффективность персонализированных учебных траекторий и высокую степень удовлетворенности студентов;
- снижение оттока студентов – уровень отчислений уменьшился с 18% до 11%, что свидетельствует о способности системы своевременно выявлять студентов группы риска и оказывать им необходимую поддержку;
- эффективность визуализации данных – оценка визуализации повысилась с 5.2 до 9.0 баллов, что указывает на улучшение восприятия данных и упрощение анализа для пользователей.

Несмотря на достигнутые успехи, выявлены области для улучшений, включая оптимизацию алгоритмов, расширение функционала рекомендаций и повышение безопасности данных. Перспективы дальнейших исследований связаны с внедрением более сложных моделей машинного обучения и адаптацией системы для других уровней образования.

Таким образом, апробация системы подтвердила её высокую эффективность и потенциал для дальнейшего развития, что открывает новые возможности для повышения качества образовательного процесса.

Заключение

В рамках исследования была разработана и внедрена система образовательной аналитики для повышения качества образования в учебном заведении. Актуальность темы обусловлена растущей ролью данных и аналитики в образовательном процессе, особенно в условиях современных вызовов, таких как необходимость индивидуализации обучения, повышение эффективности преподавания и управления образовательными учреждениями.

Основная цель исследования заключалась в создании и апробации системы образовательной аналитики, способной обрабатывать большие объемы данных и предоставлять обоснованные прогнозы и персонализированные рекомендации для улучшения успеваемости студентов и повышения эффективности управления образовательным процессом.

В первой главе были рассмотрены теоретические аспекты применения образовательной аналитики и технологий машинного обучения. Образовательная аналитика была охарактеризована как мощный инструмент для мониторинга успеваемости студентов, адаптации учебных программ и оценки эффективности преподавания. Системы машинного обучения позволяют глубже анализировать данные об учащих, выявлять скрытые закономерности и предлагать персонализированную поддержку. Однако было выявлено, что, несмотря на значительные достижения, многие вопросы, связанные с интеграцией этих технологий, остаются открытыми. В частности, требуется дальнейшая разработка и совершенствование алгоритмов и методик, что и стало предметом данного исследования.

Во второй главе был проведен анализ существующих информационных систем в образовательных учреждениях. В ходе анализа было установлено, что процессы сбора, хранения и анализа данных зачастую остаются фрагментированными и ручными, что создает барьеры для эффективного управления учебным процессом. Существующие системы образовательной аналитики имеют ограниченные возможности для анализа, что препятствует

оперативной адаптации учебных программ и принятию решений. Были выявлены проблемы, связанные с недостаточной автоматизацией, разрозненностью данных и низким уровнем использования машинного обучения, что приводит к неэффективному использованию ресурсов и замедленному реагированию на изменения.

В третьей главе была разработана архитектура системы образовательной аналитики, основанная на принципах автоматизации, интеграции данных и использовании современных технологий машинного обучения. В рамках разработки системы были обоснованы ключевые элементы технологического стека: PostgreSQL для надежного хранения данных, Django и Python для гибкости разработки и мощных аналитических возможностей. Была разработана многослойная архитектура с централизованным хранилищем данных, автоматизацией сбора информации и инструментами для визуализации данных в реальном времени. Внедрение технологий машинного обучения и современных методов визуализации значительно повысило точность прогнозов и качество принятия решений.

В четвертой главе была проведена апробация системы в учебном заведении, что позволило оценить её эффективность. Эксперименты показали значительное улучшение прогнозов успеваемости студентов, точность которых выросла на 23%. Время принятия решений сократилось с 1,5 часов до 1,5 минут, что позволило оперативно реагировать на изменения. Персонализированные рекомендации студентов получили высокие оценки, а уровень отчислений снизился на 7%.

Визуализация данных стала более эффективной, что улучшило восприятие аналитических результатов преподавателями и администраторами. Несмотря на положительные результаты, апробация выявила области для дальнейшего совершенствования, такие как оптимизация прогнозных алгоритмов и повышение уровня безопасности данных.

Гипотеза исследования, предполагавшая, что внедрение образовательной аналитики и методов машинного обучения позволит

значительно улучшить процессы прогнозирования успеваемости, оперативности принятия решений и персонализации рекомендаций, была подтверждена экспериментальными данными. В ходе апробации система продемонстрировала высокую точность прогнозов, оперативность в обработке данных и удовлетворенность пользователей персонализированными рекомендациями.

Практическая значимость исследования, заключающаяся в автоматизированном анализе образовательных данных и предложении системой персонализированных рекомендаций, подтвердилась в процессе апробации системы на базе образовательного учреждения. Внедрение системы позволило значительно повысить качество образовательного процесса, улучшить управление данными и упростить принятие решений в учебном заведении. Система позволяет администрации и преподавателям быстрее выявлять студентов, нуждающихся в дополнительной поддержке, и оперативно реагировать на изменения в успеваемости.

Результаты эксперимента показали, что система также способна снижать уровень отчислений студентов и предоставлять более глубокие аналитические данные для управления учебными процессами.

Апробация системы образовательной аналитики на базе СПб ГБ ПОУ «КЭП» подтвердила ее эффективность в реальных условиях, что стало важным этапом в оценке ее практической применимости. Система продемонстрировала значительное улучшение в прогнозировании успеваемости, повышении оперативности принятия решений и предоставлении персонализированных рекомендаций. Уровень отчислений снизился на 7%, что свидетельствует о способности системы своевременно выявлять студентов группы риска и предлагать им поддержку.

Несмотря на успехи, апробация также выявила области для дальнейшего совершенствования. В частности, требуется оптимизация прогнозных алгоритмов, расширение функционала персонализированных рекомендаций, улучшение обучения пользователей и повышение уровня безопасности

данных. Эти направления могут стать основой для дальнейших исследований и развития системы.

Таким образом, предложенная система образовательной аналитики продемонстрировала свою высокую эффективность и потенциал для дальнейшего развития. Она открывает новые возможности для повышения качества образовательного процесса, улучшения управления учебными учреждениями и поддержки студентов на всех этапах их обучения. Перспективы дальнейшего развития системы включают внедрение более сложных моделей машинного обучения, адаптацию для других уровней образования и расширение функциональности интеллектуальных помощников на базе искусственного интеллекта.

Результаты данного исследования могут стать основой для широкого внедрения подобных аналитических систем в других учебных заведениях, что будет способствовать повышению качества образования и успешности студентов в долгосрочной перспективе.

Список используемой литературы

1. Бардаков Н.Д. Цифровая трансформация образования и/или образование устойчивого развития / Н.Д. Бардаков // Инновационные научные исследования. – 2021. – №4-3(6). – С.167-172.
2. Бемянская О.В. Машинное обучение в образовании с использованием виртуального помощника / О.В. Бемянская // Шуйская сессия студентов, аспирантов, педагогов, молодых ученых: материалы XV Международной научной конференции. – 2022. – С.348-350.
3. Босенко Т.М. Применение облачных платформ глубокого и машинного обучения студентами в условиях дистанционного образования / Т.М. Босенко, Ю.В. Фролов // Актуальные проблемы теории и практики обучения физико-математическим и техническим дисциплинам в современном образовательном пространстве: IV Всероссийская (с международным участием) научно-практическая конференция, посвященная 75-летию факультета физики, математики, информатики Курского государственного университета. – 2020. – С.414-417.
4. Булгаков В.В. Исследование эффективности применения электронных образовательных ресурсов инструментами веб-аналитики (на примере ведомственного ВУЗа МЧС России) / В.В. Булгаков // Открытое образование – 2020. – №3. – Т.24. – С.4-11.
5. Булдаев А.А. Модель системы поддержки принятия решений в учебном процессе университета, основанной на аналитике обучения / А.А. Булдаев, Л.В. Найханова, И.С. Евдокимова // Программные системы и вычислительные методы. – 2020. - №4. – С.42-52.
6. Вилкова К.А. Учебная аналитика в традиционном образовании: ее роль и результаты» / К.А. Вилкова, У.С. Захарова // Университетское управление: практика и анализ. – 2020. – Том 24. – № 3. – С.59-76.
7. Гайдаржи А.И. Образовательная аналитика и адаптивное обучение с использованием модели студента в интеллектуальных

образовательных системах / А.И. Гайдаржи // Славянский форум. – 2021. – №3(33). – С.217-223.

8. Галкина А.И. Цифровизация образования в зеркале электронных образовательных ресурсов: статистика, аналитика, прогноз: сборник аналитических обзоров / А.И. Галкина, И.А. Гришан. – Красноярск: Научно-инновационный центр, 2021. – 110 с.

9. Гришан И.А. Методология обработки информации для системы автоматизированной аналитики электронных образовательных ресурсов образовательной организации / И.А. Гришан // Образовательные технологии и общество. – 2018. – №2. – Т.21. – С.484-493.

10. Евсюков Д.Ю. Трансформация образовательных систем с помощью анализа больших образовательных данных и визуальной аналитики / Д.Ю. Евсюков, В.А. Ломазов, З.С. Полякова // Цифровые и инженерные технологии в АПК: материалы Национальной научно-практической конференции. – 2022. – С.215-218.

11. Елшин А.А. Методические основания образовательной аналитики в современном ВУЗе / А.А. Елшин, А.В. Елшин, О.В. Родионов // Мир образования – образование в мире. – 2018. – №3(71). – С.191-198.

12. Епанешников Н.М. К вопросу о внедрении учебной аналитики в управление образовательным процессом в условиях цифровой трансформации образования / Н.М. Епанешников, А.В. Костюк // Региональная информатика и информационная безопасность: сборник трудов Юбилейной XVIII Санкт-Петербургской международной конференции. – 2022. – С.162-166.

13. Канаш А.В. Интеллектуальный анализ данных для построения моделей машинного обучения в образовании / А.В. Канаш, А.С. Мезина // Цифровая трансформация – шаг в будущее: материалы II Международной научно-практической конференции молодых ученых, посвященной 100-летию Белорусского государственного университет. – 2021. – С.135-139.

14. Катасонова Г.Р. Роль учебной аналитики в образовательном процессе ВУЗа / Г.Р. Катасонова, А.С. Шкрум // Инновационные,

информационные и коммуникационные технологии: сборник трудов XVIII Международной научно-практической конференции. – 2021. – С.54-57.

15. Ковалева Е.А. Анализ образовательных данных электронного учебного курса с помощью средств и методов учебной аналитики / Е.А. Ковалева, Т.Б. Павлова // Региональная информатика (РИ-2020): XVII Санкт-Петербургская международная конференция. Материалы конференции. Том Часть 2. Санкт-Петербург. – 2020. – С.62-63.

16. Котова Е.Е. Прогнозирование успешности обучения в интегрированной образовательной среде с применением инструментов онлайн аналитики / Е.Е. Котова // Компьютерные инструменты в образовании. – 2019. – №4. – С.55-80.

17. Кузнецов И.А. Методы и алгоритмы машинного обучения для переработки и классификации слабоструктурированных текстовых данных в научных рекомендательных системах / И.А. Кузнецов, авт. дисс. канд. тех. наук: 05.13.01. – Москва, 2019. – 26с.

18. Лавриенко И.Ю. Перспективы использования LMS в рамках современного высшего образования / И.Ю. Лавриенко // Научно-методический электронный журнал «Концепт». – 2023. – №1. – С.17-35.

19. Ларионова В.А. Технологии с человеческим лицом: перспективы использования цифровых сервисов для поддержки индивидуализированного обучения / В.А. Ларионова, Е.В. Сеницын, А.В. Толмачев, А.Б. Халфин // ELearning stakeholders and researchers summit 2021. – 2022 – С.94-113.

20. Ло Х. Цифровая трансформация как объект современного образовательного процесса / Х. Ло, Л.Н. Рулиене // Сибирский педагогический журнал. – 2022. – №4. – С.22-37.

21. Маризина В.Н. Аналитика в деятельности будущего ИТ-специалиста / В.Н. Маризина // Педагогика, психология, общество: теория и практика: сб. материалов Всероссийской научно-практической конференции. – 2019. – С.83-86.

22. Нажмидинов Х.А. Разработка макетного варианта автоматизированной информационной системы управления электронным обучением на основе образовательной аналитики / Х.А. Нажмидинов, С.Е. Сосенушкин // Цифровая экономика: технологии, управление, человеческий капитал: материалы Всероссийской научно-практической конференции. – 2019. – С.221-225.

23. Никитин А.А. Инновационные технологии виртуальной/дополненной реальности, ученая аналитика и адаптивное обучение в современном образовательном процессе / А.А. Никитин, Д.А. Носов // Актуальные проблемы недропользования: тезисы докладов XVIII Международного форума-конкурса студентов и молодых ученых. Том 2. Санкт-Петербург. – 2022. – С.396-398.

24. Новиков В.С. Значение машинного обучения в сфере образования / В.С. Новиков, А.В. Шеремет // Физико-математические и технические науки как фундамент становления постиндустриального общества: сборник статей Международной научно-практической конференции. – 2020. – С.25-26.

25. Оськин А.Ф. Готовность учреждений высшего образования к внедрению процесса образовательной аналитики / А.Ф. Оськин, Д.А. Оськин // Управление информационными ресурсами: материалы XVII Международной научно-практической конференции. – 2021. – С.177-179.

26. Оськин Д.А. Образовательная аналитика. Новый подход к повышению качества учебного процесса / Д.А. Оськин, А.Ф. Оськин // Исторические исследования в контексте науки о данных: информационные ресурсы, аналитические методы и цифровые технологии: материалы международной конференции. – 2020. – С.454-461.

27. Оськин Д.А. Процесс образовательной аналитики в учреждении образования / Оськин Д.А., Оськин А.Ф. // Управление информационными ресурсами: материалы XVI Международной научно-практической конференции. – 2020. – С.176-177.

28. Паскова А.А. Анализ образовательных данных и учебная аналитика как средства повышения эффективности обучения / А.А. Паскова // Актуальные вопросы науки и образования: материалы II международной научно-практической конференции. – 2020. – С.109-112.

29. Петряева Е.Ю. Аналитика данных виртуальных образовательных событий / Е.Ю. Петряева // Успехи современной науки и образования. – 2017. – №4. – Т.1. – С.28-31.

30. Пирогланов Ш.Ш. К вопросу о значении аналитики, отчетности и инновационных подходов к реализации интерактивной среды ВУЗа / Ш.Ш. Пирогланов, В.П. Складов, А.В. Паленый. – Проблемы современного педагогического образования. – 2021. – №72-4. – С.213-216.

31. Политов А.Ю. Анализ моделей и инструментов предиктивной аналитики для анализа образовательных данных. – А.Ю. Политов, Р.Р. Акжигитов, К.А. Судариков // Инновации. Наука. Образование. – 2021. – №28. – С.1055-1065.

32. Розин В.М. Цифровизация в образовании (по следам исследования «Трудности и перспективы цифровой трансформации образования») / В.М. Розин // Мир психологии. – 2021. – №1-2(105). – С.104-115.

33. Савиных Г.П. Образовательная аналитика на службе повышения квалификации педагогических работников общеобразовательных организаций / Г.П. Савиных, И.А. Сахнова, И.Р. Идрисов // Современное образование: актуальные вопросы и инновации. – 2022. – №1. – С.23-27.

34. Самохина М.А. Цифровая трансформация образования как новая возможность развития традиционного образования / М.А. Самохина // Инновационные проекты и программы в психологии, педагогике и образовании: сборник статей Международной научно-практической конференции. – 2020. – С.102-108.

35. Стабровская А.В. Система реализации академической мобильности обучающихся вуза в условиях цифровизации / А.В. Стабровская, дисс. канд. пед. наук: 05.08.07. – Москва, 2022. – 200с.

36. Строков А.А. Цифровая культура и ценности российского образования / А.А. Строков, авт. дисс. канд. филос. наук: 09.00.13. – Нижний Новгород, 2021. – 21с.

37. Токтарова В.И. Анализ образовательных данных взаимосвязи успешности обучения и поведения студентов в цифровой образовательной среде ВУЗа / В.И. Токтаров, О.Г. Попова // Информатика и образование. – 2022. – № 4. – Т.37. – С.54-63.

38. Токтарова В.И. Визуальная аналитика образовательных данных студентов в системе электронного обучения / В.И. Токтарова, О.Г. Попова // Сибирский педагогический журнал. – 2023. – №1. – С.61-71.

39. Тращев С.В. О сущности понятий «цифровая трансформация образования» и «цифровизация образования» / С.В. Тращев // Мировая наука 2021. Проблемы и перспективы: материалы международной научно-практической конференции. – 2021. – С.77-79.

40. Трусова М.И. Повышения качества образования с использованием потенциала цифровой педагогики в условиях цифровой трансформации образования / М.И. Трусова, А.А. Трусов // Цифровизация образования в условиях конвергентной реальности: сборник материалов III Международной научно-практической конференции. – 2021. – С.113-116.

41. Тымкив А.И. Внедрение технологии машинного обучения в образование / А.И. Тымкив, А.В. Федоренко // Проблемы и перспективы разработки инновационных технологий: сборник статей Международной научно-практической конференции – 2021. – С.49-53.

42. Фиофанова О.А. Проблема интеграции цифровых сервисов аналитики данных: компетенции педагога в работе с образовательными данными / О.А. Фиофанова // Вестник Московского университета. Серия 20: педагогическое образование. – 2020. – №3. – С.38-49.

43. Царькова Е.Г. Учебная аналитика в дистанционном обучении: особенности применения и перспективы развития / Е.Г. Царькова // Прикладная психология и педагогика. – 2022. – №3. – Т.7. – С. 54-66.

44. Черных С.И. «Паралич анализом», или осмысленное образование / С.И. Черных, В.И. Паршиков // Профессиональное образование в современном мире. – 2018. – №4. – Т.8. – С.2148-2159.
45. Чернышенко О.В. Цифровой след как средство повышения качества образовательного процесса / О.В. Чернышенко, Е.А. Носачева // Мир университетской науки: культура, образование. – 2022. – № 2. – С.15-18.
46. Шабардин А.М. Большие данные в системе образования / А.М. Шабардин // Информационные технологии в науке, промышленности и образовании: сборник трудов научно-технической конференции в рамках Всероссийского молодежного научного форума «Общение студентов и аспирантов в научной и профессиональной сферах». – 2021. – С.62-69.
47. Шамсутдинова А.И. Образовательная аналитика результатов повышения квалификации преподавателей в LMS Moodle / А.И. Шамсутдинова, Г.Ф. Хасанова // Управление устойчивым развитием. – 2021. - №6(37). – С.110-115.
48. Шаяхметов Ш.Г. Использование образовательной аналитики Kahoot для индивидуализации обучения / Ш.Г. Шаяхметов // Молодой ученый. – 2021. – №45(387). – С.261-263.
49. Ширинкина Е.В. Методы интеллектуального анализа данных и образовательной аналитики / Е.В. Ширинкина // Современное образование. – 2022. – №1. – С.51-67.
50. Шорохова Н.А. Роль аналитики образовательных данных для принятия оптимальных решений в процессе цифровой трансформации ВУЗа / Н.А. Шорохова, И.Л. Минин // Современные технологии принятия решений в цифровой экономике: сборник трудов Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. – 2018. – С.135-138.
51. Guzman J. Cluster analysis in Higher Education Institutions knowledge identification and production processes / R.A. Zuluaga-Ortiz, L.E. Garizabal-Donado, E. Delahoz-Dominguez, A. Marquez-Castillo, M. Suarez-Sánchez // Procedia Computer Science. – 2022. – Vol. 203. – P.570-574.

52. Mian Y.S. Learning Analytics in Education, Advantages and Issues: A Systematic Literature Review / F.Khalid, A.W. Chiew Qun, S. S. Ismail // Creative Education. – 2022. – Vol. 13. – P.2913-2920.

53. Paolucci C. A review of learning analytics opportunities and challenges for K-12 education / S. Vancini, R.T. Bex II, C. Cavanaugh, C. Salama, Z. de Araujo // Heliyon. – 2024. – Vol. 10, Issue 4. – P.25767.

54. Soncin M. Data analytics in education: are schools on the long and winding road? / M. Cannistra // Qualitative research in accounting and management: QRAM. – 2022. – Vol. 19, Issue 3. – P.286-304.

55. Ullrich A. Data mining of scientific research on artificial intelligence in teaching and administration in higher education institutions: a bibliometrics analysis and recommendation for future research / G. Vladova, F. Eigelshoven, A. Renz // Discover Artificial Intelligence. – 2022. – Vol. 2, No. 16, – P.1-18.

Приложение А

Шаблон анкеты для оценки эффективности визуализации данных в системе образовательной аналитики

На рисунке А.1 представлен шаблон анкеты для оценки персональных рекомендаций.

Критерий оценки	Оценка (от 1 до 10)
1. Актуальность рекомендаций Насколько рекомендации полезны и соответствуют текущим академическим потребностям студента.	 _____
2. Понятность и доступность Степень понимания студентом, как следовать рекомендациям.	 _____
3. Влияние на успеваемость Оценка того, как рекомендации способствовали улучшению результатов обучения.	 _____
4. Удовлетворенность рекомендациями в целом Общая оценка удовлетворенности рекомендациями.	 _____
5. Осуществимость рекомендаций Оценка того, насколько рекомендации достижимы и применимы в реальной среде	 _____

Рисунок А.1 – Шаблон анкеты для оценки персональных рекомендаций

Приложение Б

Шаблон анкеты для оценки эффективности визуализации данных в системе образовательной аналитики

На рисунке Б.1 представлен шаблон анкеты для оценки эффективности визуализации данных.

Анкета для оценки эффективности
визуализации данных

Критерий оценки	Оценка (от 1 до 10)
<p>1. Доступность интерактивных элементов Как Вы оцениваете доступность и эффективность интерактивных элементов, таких как фильтры и панели?</p>	 _____
<p>2. Общая удовлетворенность Оцените общую удовлетворенность визуализацией данных.</p>	 _____
<p>3. Полезность для принятия решений Определите полезность визуализаций для поддержки принятия решений.</p>	 _____
<p>4. Доступность информации Оцените, насколько легко можно найти необходимую информацию на дашборде.</p>	 _____
<p>5. Понятность графиков Оцените, насколько легко понимаются графики и таблицы.</p>	 _____

Рисунок Б.1 - Шаблон анкеты для оценки эффективности визуализации данных