

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«Тольяттинский государственный университет»

Кафедра Прикладная математика и информатика
(наименование)

09.04.03 Прикладная информатика
(код и наименование направления подготовки)

Управление корпоративными информационными процессами
(направленность (профиль))

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)

на тему «Исследование методов формирования индивидуальной траектории онлайн обучения с применением искусственного интеллекта»

Обучающийся

А.С. Тимербулатов

(Инициалы Фамилия)

(личная подпись)

Научный
руководитель

к.п.н., О.А. Крайнова

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

Тольятти 2024

Оглавление

Введение.....	3
Глава 1 Анализ состояния исследований в области формирования индивидуальной траектории онлайн обучения на основе современных образовательных и информационных технологий.....	7
Глава 2 Анализ методологий формирования индивидуальной траектории онлайн обучения на основе технологий искусственного интеллекта и машинного обучения.....	18
2.1 Технологии персонализированного онлайн обучения на основе машинного обучения.....	19
2.2 Методологии построения систем формирования индивидуальной траектории онлайн обучения на основе технологий искусственного интеллекта и машинного обучения.....	21
Глава 3 Разработка моделей и алгоритма системы формирования индивидуальной траектории онлайн обучения на основе технологий машинного обучения.....	29
3.1 Разработка логической модели системы формирования индивидуальной траектории онлайн обучения с применением технологий машинного обучения.....	29
3.2 Обзор и анализ алгоритмов машинного обучения для разработки системы формирования индивидуальной траектории онлайн обучения.....	36
3.3 Описание используемых алгоритмов машинного обучения.....	42
Глава 4 Апробация проектных решений и оценка их эффективности.....	48
4.1 Разработка элементов системы формирования индивидуальной траектории онлайн обучения.....	48
4.2 Оценка эффективности проектных решений.....	60
Заключение.....	65
Список используемой литературы и используемых источников.....	68

Введение

Одной из основных задач образования является удовлетворение индивидуальных запросов и потребностей каждого обучаемого. Система образования с годами становится гибкой, вариативной, и в этой сфере появляются новые тенденции. Они приходят на смену стандартным учебным планам. Личностный потенциал каждого ученика реализует индивидуальная образовательная траектория (ИОТ).

Основным принципом ИОТ является учет интересов, возможностей, особенностей и потребностей обучаемого. Процесс обучения направлен на приобретение новых компетенций в сферах, представляющих интерес для обучаемого, который становится субъектом своего образования [4]. Следует отметить, что в отечественных и зарубежных источниках в качестве ИОТ используются также такие понятия, как «индивидуальный образовательный маршрут», «индивидуальный план обучения», «персонализированное образование» и др. [18].

Процесс обучения уникален для каждого человека. Как показывает практика, важное значение для повышения успеваемости обучаемых и освоения ими новых компетенций имеют используемые методы обучения.

Существующая система онлайн обучения сосредоточена на создании инструментов для совместного обучения или облегчении проведения курсов для дискуссионного форума и прямого общения, например, в чате. Однако проблема повышения индивидуальных достижений каждого обучающегося остается актуальной.

Решение данной проблемы возможно за счет применения современных методов формирования индивидуальной траектории онлайн обучения.

Как показывают исследования, наилучшие результаты достигаются, если в процессе формирования ИОТ учитываются два основных ключевых фактора: персонализированное онлайн обучение и использование технологий искусственного интеллекта (ИИ). Данная тенденция обусловлена значительным интересом к современным образовательным парадигмам

благодаря быстрому прогрессу, достигнутому в области ИИ и образовательных онлайн технологий.

Таким образом, исследование методов формирования индивидуальной траектории онлайн обучения с применением ИИ актуально и представляет научно-практический интерес.

Объектом настоящего исследования является индивидуальная траектория онлайн обучения.

Предметом исследования являются методы формирования индивидуальной траектории онлайн обучения с применением ИИ.

Целью работы является исследование методов формирования индивидуальной траектории онлайн обучения с применением ИИ.

Для достижения цели необходимо решить следующие задачи:

- произвести анализ состояния исследований в области формирования индивидуальной траектории онлайн обучения на основе современных образовательных и информационных технологий;
- произвести анализ методологических подходов к формированию индивидуальной траектории онлайн обучения на основе технологий ИИ и машинного обучения;
- разработать модели и алгоритмы системы формирования индивидуальной траектории онлайн обучения на основе технологий ИИ и машинного обучения;
- провести апробацию и оценить эффективность предлагаемых проектных решений.

Гипотеза исследования: применение предлагаемых в диссертации моделей и алгоритмов системы формирования ИОТ на основе технологий ИИ и машинного обучения обеспечит повышение эффективности процесса онлайн обучения.

Методы исследования: в процессе исследования использованы следующие подходы и методы: методы формирования ИОТ, методы и технологии проектирования информационных систем, технологии ИИ и

машинного обучения.

Научная новизна исследования заключается в разработке новых моделей и алгоритмов формирования индивидуальной траектории онлайн обучения.

Практическая значимость исследования заключается в возможности практического применения предлагаемых моделей и алгоритмов формирования ИОТ для повышения эффективности онлайн обучения в образовательной организации.

Теоретической основой диссертационного исследования являются научные труды российских и зарубежных ученых, занимающихся проблемами формирования индивидуальной ИОТ с применением ИИ.

Основные этапы исследования: исследование проводилось с 2021 по 2023 год в несколько этапов.

На первом (констатирующем) этапе формулировалась тема исследования, выполнялся сбор информации по теме исследования из различных источников, проводилась формулировка гипотезы, определялись постановка цели, задач, предмета исследования, объекта исследования и выполнялось определение проблематики данного исследования.

Второй этап – теоретический. В ходе проведения данного этапа осуществлялся анализ методологических подходов к формированию индивидуальной траектории онлайн обучения на основе технологий ИИ и машинного обучения (МО), разработаны модели и алгоритмы системы формирования индивидуальной траектории онлайн обучения на основе технологий МО, опубликована статья по теме исследования в научном сборнике.

Третий этап – практический. В ходе проведения данного этапа проводилась апробация предлагаемых проектных решений, произведена оценка их эффективности, сформулированы выводы о полученных результатах по проведенному исследованию.

На защиту выносятся:

- модели и алгоритмы системы формирования индивидуальной траектории онлайн обучения на основе технологий машинного обучения;
- результаты апробации и оценки эффективности предлагаемых проектных решений.

По теме исследования опубликована 1 статья:

Тимербулатов А.С. Модель системы формирования индивидуальной траектории онлайн обучения // Вестник научных конференций. 2024. №2-1 (102). С. 85-86.

Диссертация состоит из введения, четырех глав, заключения, списка используемой литературы и используемых источников.

Во введении обоснована актуальность темы исследования, представлены объект, предмет, цели, задачи и положения, выносимые на защиту диссертации.

В первой главе дан анализ состояния исследований в области формирования индивидуальной траектории онлайн обучения на основе современных образовательных и информационных технологий.

Во второй главе дан анализ методологических подходов к формированию индивидуальной траектории онлайн обучения на основе технологий ИИ и МО.

Третья глава посвящена разработке моделей и алгоритмов системы формирования индивидуальной траектории онлайн обучения на основе технологий машинного обучения.

В четвертой главе выполнены апробация предлагаемых проектных решений и оценка их эффективности.

В заключении приводятся результаты исследования.

Работа изложена на 72 страницах и включает 31 рисунок, 11 таблиц и 41 источник.

Глава 1 Анализ состояния исследований в области формирования индивидуальной траектории онлайн обучения на основе современных образовательных и информационных технологий

Проблематике формирования ИОТ на основе современных образовательных и информационных технологий посвятили свои работы Е.А. Алешугина, Т.М. Шамсутдинова, К. К. Хоанг, Hengyao Tang, Milind Bhushan, Kulkatechol Kanokngamwitroj, Kelly Lake, Juan Carlos Muñoz-Carpio и другие.

В работе [15] рассмотрены теоретические и практические аспекты проектирования адаптивных учебных курсов, ориентированных на создание ИОТ, с учетом различных механизмов (ресурсов) адаптации. Отмечено, «что ИОТ является одним из эффективных средств реализации профессионально-образовательного потенциала личности и строится на основе определенных форм, методов, технологий и механизмов обучения. Современные технологии онлайн образования предоставляют широкие возможности по реализации концепции персонифицированного гибкого обучения, в частности, с использованием адаптивных учебных курсов (рисунок 1)» [15].

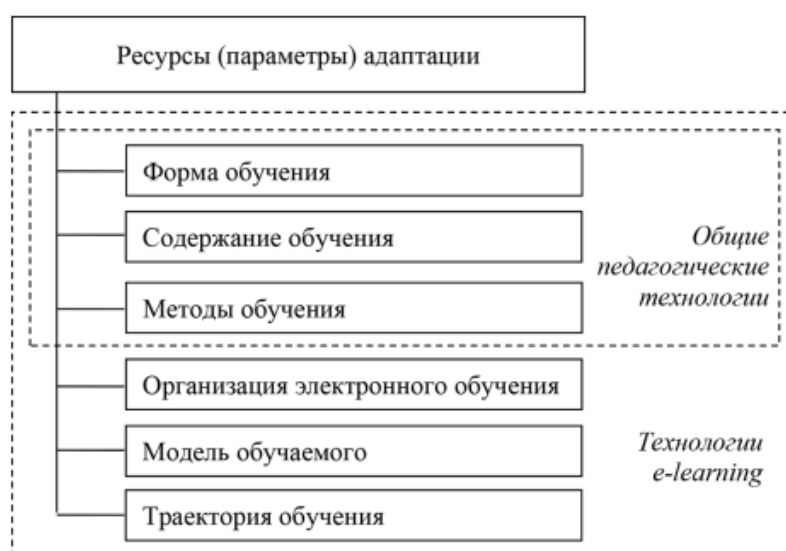


Рисунок 1 – Ресурсы адаптации в адаптивных системах обучения

В работе [1] показано, что реализация собственных образовательных целей студента по его способностям, возможностям и мотивации, возможна с использованием онлайн-курсов, доступ к которым осуществляется круглосуточно.

«Они могут использоваться как дополнительный инструмент в организации смешанного обучения, так и самостоятельно (при переводе какой-либо дисциплины в полностью электронный формат).

Для подготовки студента к формированию индивидуальной образовательной траектории необходимо прохождение нескольких этапов, после чего обучающемуся будет легче выбирать для себя цели и ставить задачи» [1].

В качестве примера приведена деятельность Нижегородского Государственного педагогического университета имени Козьмы Минина.

В зарубежных исследованиях особое внимание уделяется индивидуальному плану обучения и персонализированному образованию.

«Individual Learning Plan или индивидуальный план обучения (ИПО) — это инструмент планирования и мониторинга, ориентированный на обучающихся, который настраивает возможности обучения на протяжении всего периода обучения, расширяет их кругозор и способствует достижению целей» [23].

ИПО – это также инструмент профессионального развития, который помогает человеку определить сильные стороны и области для роста, а также установить достижимые цели.

ИПО уже много лет используется в качестве инструмента планирования карьеры в промышленности, государственной работе и школьном образовании.

В университетах США студентам и аспирантам было предложено использовать ИПО в качестве инструмента профессионального развития, который поможет им подготовиться к будущей академической и неакадемической карьере.

Процесс ИПО – это полезный способ целенаправленного разговора со своим наставником/консультантом о карьерных интересах, развитии навыков и планировании своего будущего на протяжении всего времени обучения в образовательной организации.

В ИПО документируются интересы, потребности, поддержка, выбор курсов (включая доступ к программам на уровне колледжа), места перехода и другой опыт обучения как в образовательной организации, так и за ее пределами.

Эта информация позволяет разработать продуманную программу обучения, ведущую к получению квалификации и необходимых для профессии навыков [23].

Персонализированное обучение – это опыт обучения, отвечающий уникальным потребностям и предпочтениям человека.

Это ориентированный на обучающихся подход, который предоставляет им более индивидуальный и гибкий путь обучения, а также возможности изучать темы, соответствующие их увлечениям.

По оценкам, к 2024 году мировой рынок персонализированного обучения может превысить 2 миллиарда долларов.

Поскольку организации по всему миру стремятся сделать обучение доступным для сотрудников со всего мира, персонализированное обучение все чаще используется для создания индивидуальных путей обучения, адаптированных к местоположению, ролям, карьерным целям и производительности персонала [27].

Отмечено, что персонализированное обучение с использованием ИИ делает обучение более эффективным, действенным и масштабируемым, позволяя L&D-руководителям оказывать адресную поддержку каждому обучающемуся, адаптироваться к уникальным стилям обучения и более внимательно отслеживать прогресс.

На рисунке 2 показаны способы использования ИИ для персонализации обучения [17].

How AI Is Used To Personalize Learning

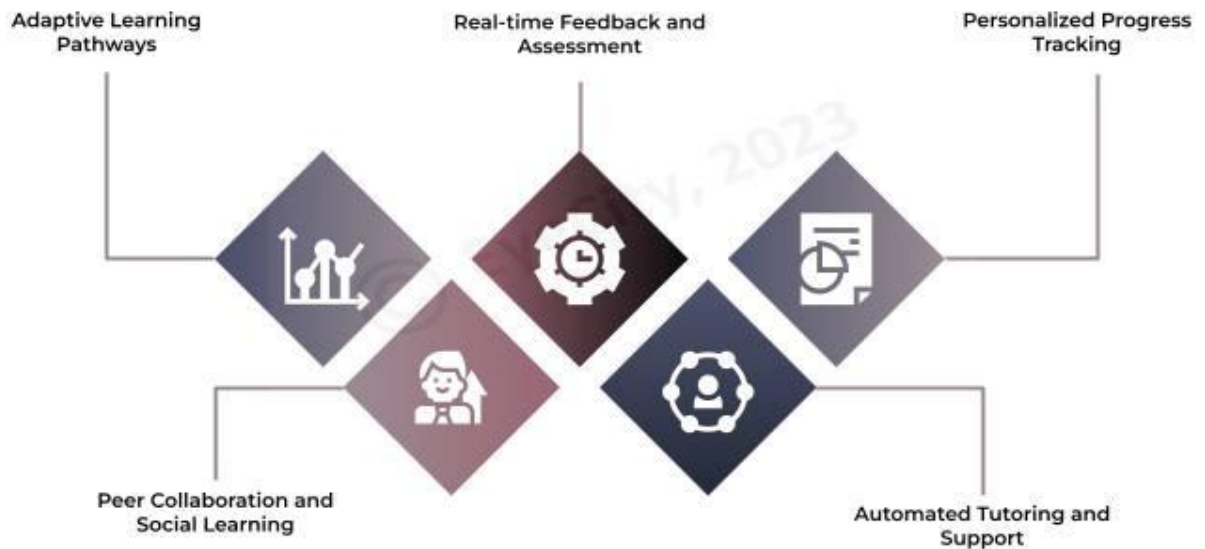


Рисунок 2 – Способы использования ИИ для персонализации обучения

Алгоритмы искусственного интеллекта могут анализировать огромные объемы данных, включая прошлые достижения, интересы и цели учащегося, чтобы генерировать персонализированные рекомендации по образовательным ресурсам, книгам, статьям, видео и другим учебным материалам. Это помогает обучающимся находить соответствующий контент, соответствующий их конкретным потребностям и предпочтениям [21].

В работе [13] представлены алгоритмы формирования ИОТ автоматизированному проектированию средствами интеллектуальной адаптивной системы, ядром которой является математическая модель, реализованная в виде нейронной сети для решения задач повышения эффективности обучения за счет автоматизированного управления процессом обучения (рисунок 3).

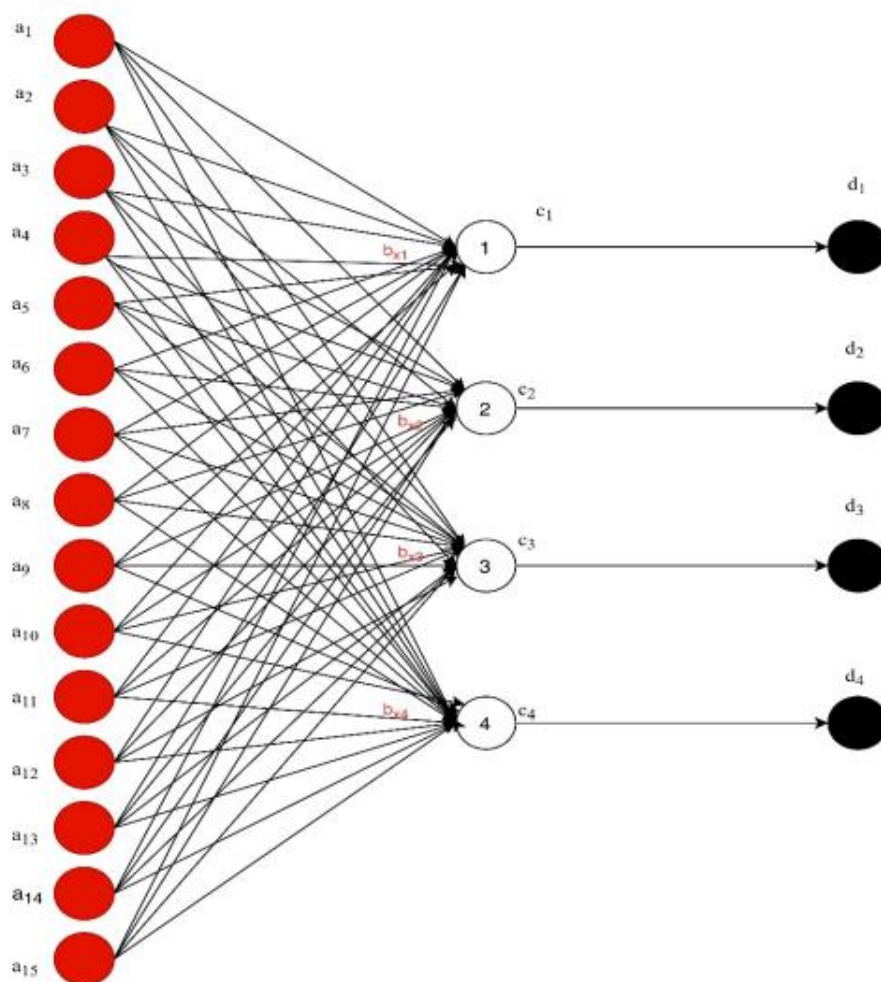


Рисунок 3 – Графическое отображение искусственной нейронной сети для 15 вопросов

В работе [16] предложена модель персонализированного обучения для поиска подходящих методов обучения на основе алгоритма глубокого обучения. «При построении и тестировании модели были учтены такие факторы персонализированного обучения, как адаптивное, индивидуализированное, дифференцированное и компетентностное обучение» [16].

Эти факторы будут служить источником информации для сети, а затем они будут обработаны, чтобы определить, какая модель обучения имеет наибольшие шансы на повышение успеваемости студентов (рисунок 4).

Считается, что модель приносит большую пользу как обучаемому, так и преподавателю, а также позволяет более удобно управлять процессом

обучения для обеих сторон.

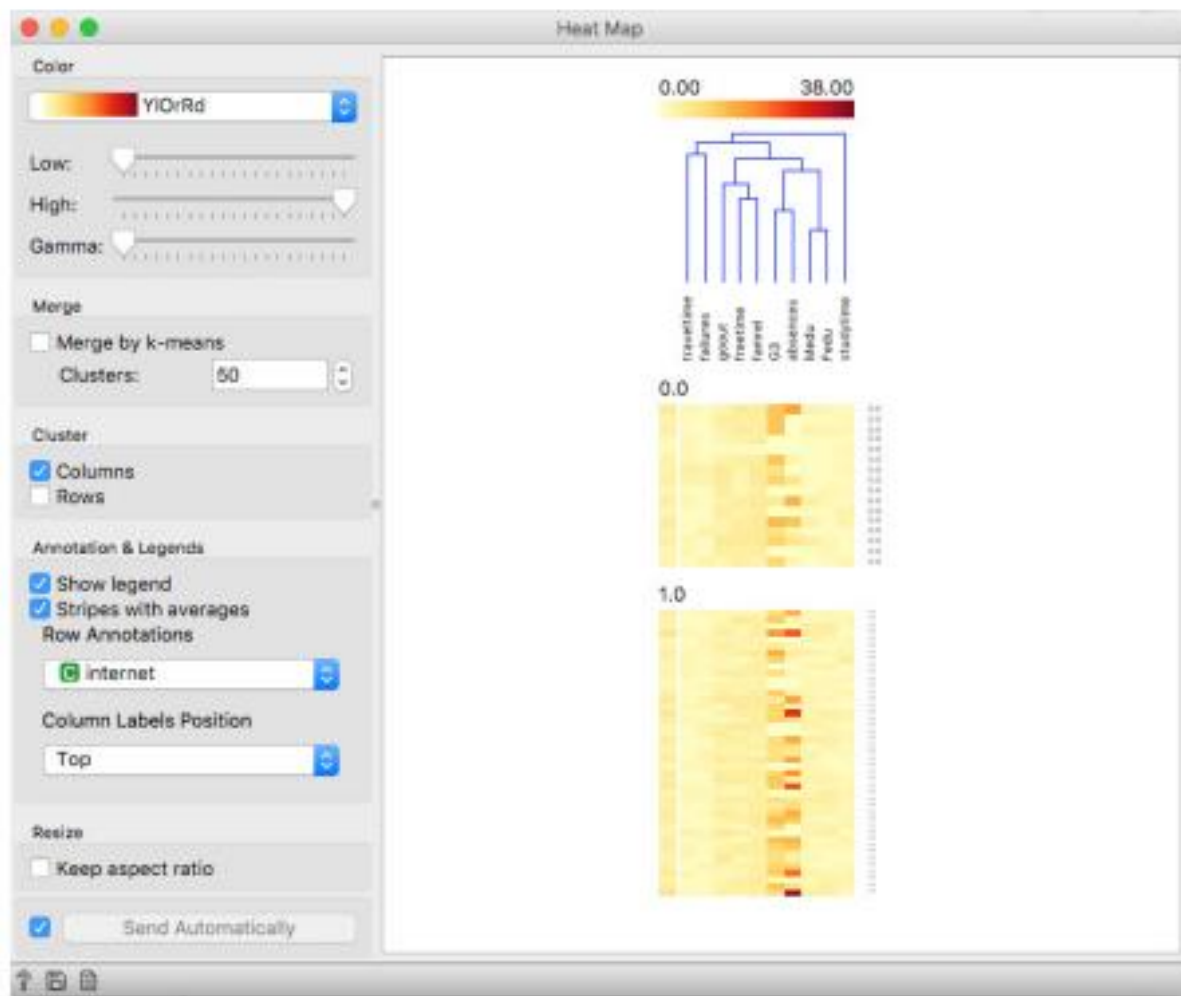


Рисунок 4 – Визуализация тепловой карты предлагаемой модели

Результат тестирования показал, что предложенная модель с подходом глубокого обучения превзошла другие классификаторы с вероятностью успеха 72%.

В этой работе [24] представлена модель прогнозирования успеваемости с использованием оптимизированного ансамблевого классификатора, который представляет собой тип модели машинного обучения (МО) для прогнозирования успеваемости учащихся с использованием несбалансированных наборов данных. Результаты сравниваются с

существующими современными ансамблевыми методами, включая бэггинг и бустинг, которые в настоящее время используются в литературе (рисунок 5).

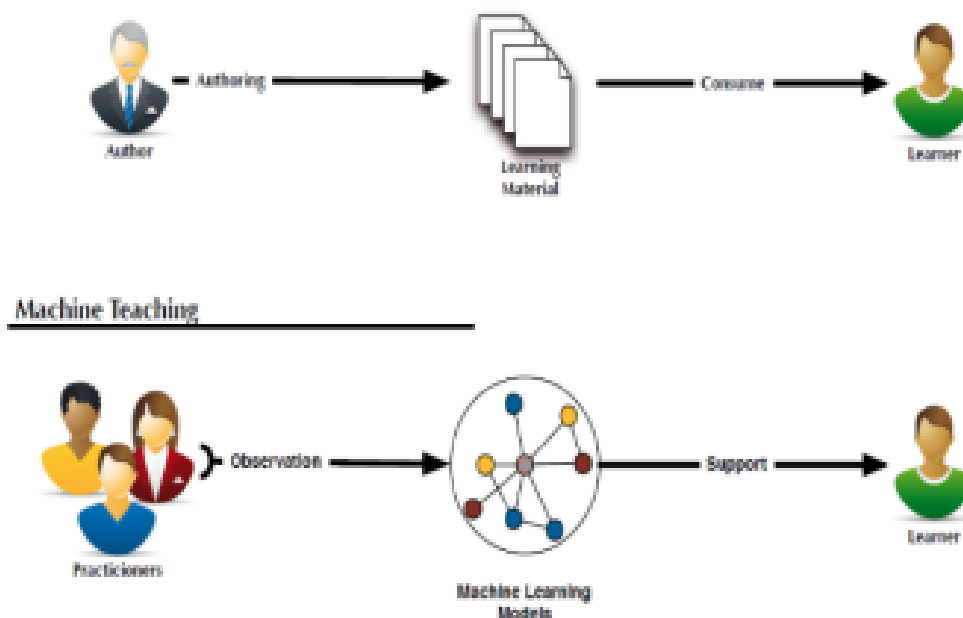


Рисунок 5 – Использование МО для персонализации обучения

Результаты предложенной модели с использованием набора данных студентов показывают точность более 80%.

В работе [26] используется модель машинного обучения и предлагается новый алгоритм «RSU ML-PL», который позволяет создать персонализированную систему самообучения.

Авторы исследования пришли к выводу, что обучаемые хорошо сдают выпускные экзамены.

На рисунке 6 показан структурная схема фреймворка, обеспечивающего реализацию предлагаемой модели.

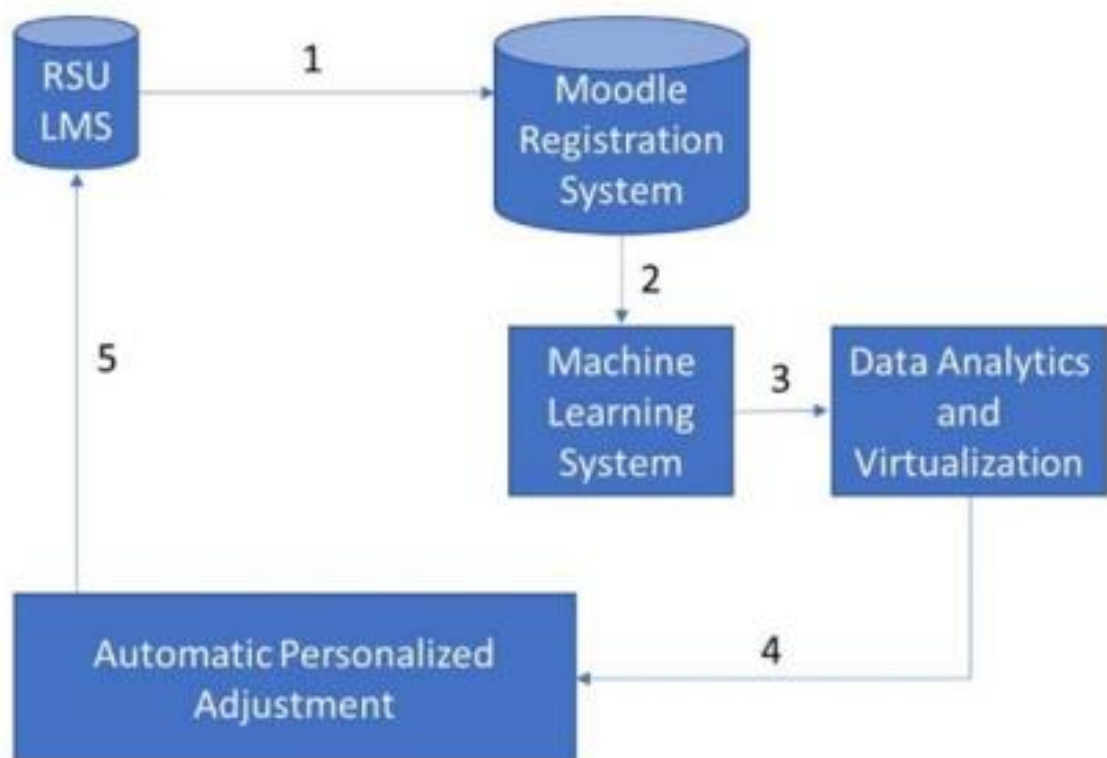


Рисунок 6 – Структурная схема фреймворка

По мнению авторов, «после программы обучения по группе риска предлагаемая модель прогнозирования по дереву решений превосходит другие. Модель станет ориентиром для следующего набора учащихся в следующем году» [26].

Алгоритм RSU-ML-PL приводит к созданию системы рекомендаций для персонализированных обучающих кластеров для будущей работы.

В работе [30] предлагается интеллектуальная веб-система обучения на основе искусственного интеллекта, которая по мнению авторов потенциально может произвести революцию в том, как студенты взаимодействуют с учебными материалами и получают к ним доступ.

Эта система может предоставлять персонализированные цели обучения, рекомендовать ресурсы и виды деятельности, а также оценивать прогресс с учетом потребностей каждого обучаемого, устраняя ограничения существующих систем обучения.

Архитектура предлагаемой системы показана на рисунке 7.

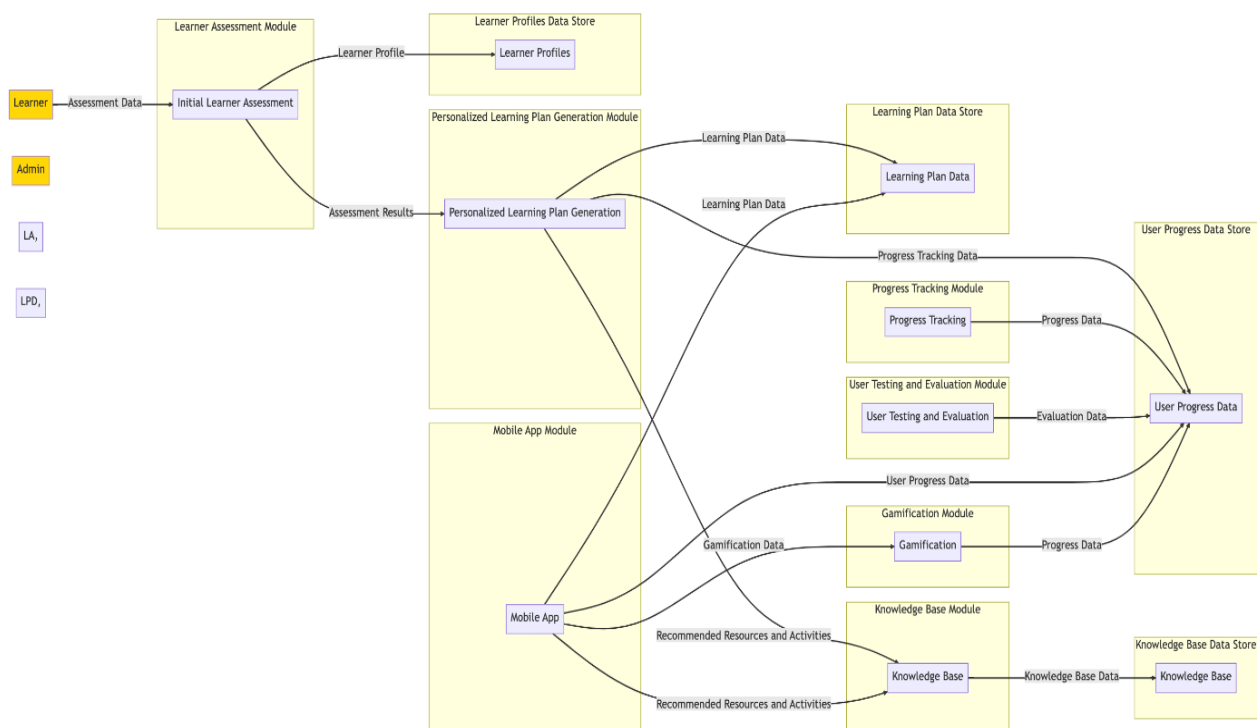


Рисунок 7 – Архитектура предлагаемой системы

Использование системы не только будет мотивировать студентов продолжать обучение, но и повысит эффективность данного процесса.

В работе [20] отмечено, что с точки зрения образовательных услуг, то, как реализовать оценку обучения, отвечающую потребностям персонализированного обучения, является важным вопросом, который необходимо изучить в области персонализированного обучения.

В этой работе большие данные, генерируемые студентами на платформе онлайн-обучения, используются в качестве цели исследования, и в соответствии с уровнем способности студентов к обучению создается глубокая нейронная сеть для кластеризации и группировки их в соответствии с методом когнитивного мышления.

Чтобы уменьшить избыточность данных и повысить эффективность обработки, для извлечения типичных признаков используется глубокая

нейронная сеть с пятью скрытыми слоями, чтобы получить более точные результаты оценки.

Наконец, модель нейронной сети используется для получения результатов кластеризации различных групп учебного поведения и кривых оценки баллов знаний студентов на разных уровнях по пяти курсам (рисунок 8).

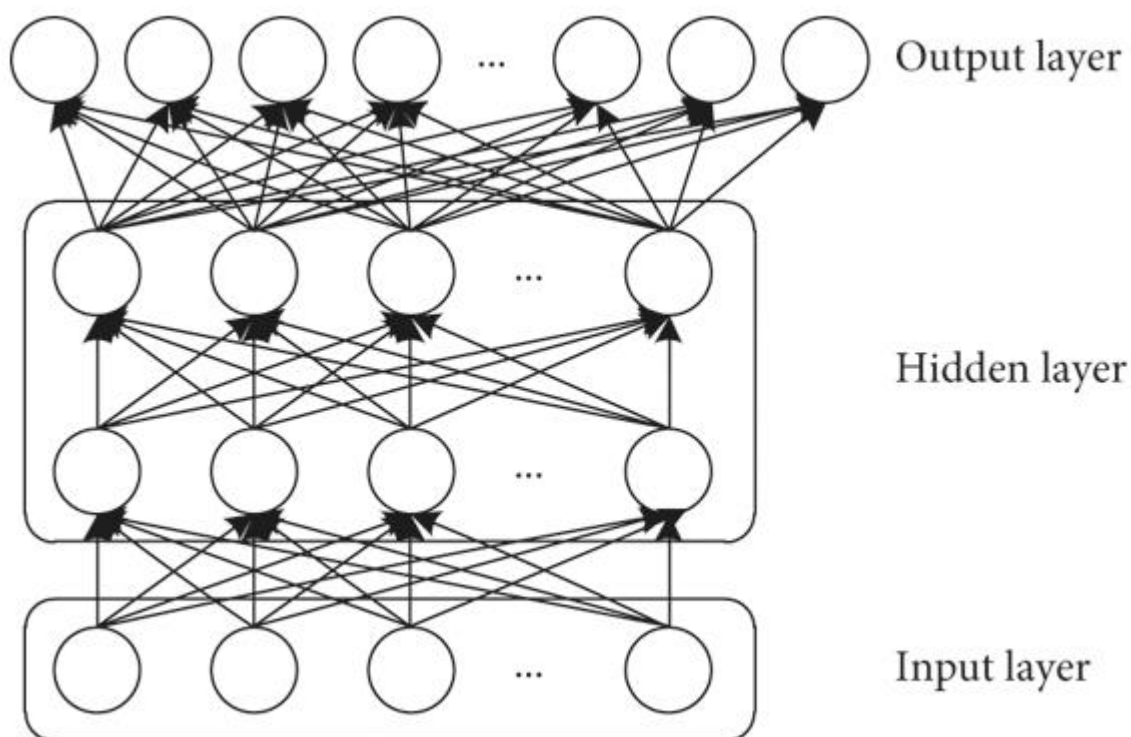


Рисунок 8 – Структура нейросети

Судя по результатам эксперимента, предлагаемый метод персонализированной оценки может эффективно анализировать различия в обучении между студентами с разными уровнями способностей и в основном соответствует стандартам оценки искусственных экспертов.

Выводы по главе 1

В результате проделанной работы были сделаны следующие выводы:

- проблематика формирования ИОТ на основе современных образовательных и информационных технологий широко рассмотрена в работах отечественных и зарубежных ученых;
- в отечественных и зарубежных источниках используются также такие понятия, как «индивидуальный образовательный маршрут», «индивидуальный план обучения», «персонализированное образование» и др.;
- как показывают исследования, наилучшие результаты достигаются, если в процессе формирования ИОТ учитываются два основных ключевых фактора: персонализированное онлайн обучение и использование технологий ИИ;
- персонализированное обучение с использованием ИИ делает обучение более эффективным, действенным и масштабируемым, позволяя L&D-руководителям оказывать адресную поддержку каждому обучающемуся, адаптироваться к уникальным стилям обучения и более внимательно отслеживать прогресс;

Таким образом, анализ литературы и источников подтвердил интерес ученых и специалистов к проблеме формирования индивидуальной траектории онлайн обучения на основе ИИ и МО, что подтверждает актуальность темы исследования.

Глава 2 Анализ методологий формирования индивидуальной траектории онлайн обучения на основе технологий искусственного интеллекта и машинного обучения

В наш цифровой век интеграция искусственного интеллекта (ИИ) в образовательный ландшафт вызвала смену парадигмы в том, как мы воспринимаем преподавание и обучение.

Внедрение алгоритмов ИИ не только изменило образовательный опыт, но и проложило путь к персонализированному обучению и моделям прогнозирования, адаптированным к конкретным потребностям учащихся.

Платформы на базе искусственного интеллекта могут создавать индивидуальные пути обучения, исходя из их целей, интересов и предшествующих знаний. Адаптируя учебную программу и темп обучения к потребностям каждого учащегося, ИИ гарантирует, что они получат соответствующие задачи и поддержку, максимизируя результаты обучения.

Важно отметить что, хотя ИИ предлагает огромный потенциал для персонализации образования, учителя-люди по-прежнему играют решающую роль в создании благоприятной и вдохновляющей среды обучения. Сочетание искусственного интеллекта и человеческого опыта может привести к более эффективному и персонализированному образованию для каждого учащегося.

Машинное обучение - это подразделение искусственного интеллекта, которое использует алгоритмы для обучения на основе данных. В секторе образования оно помогает преподавателям точно определять академические проблемы учащихся и решать их, импровизировать при выборе учебной программы и адаптировать индивидуальные методики преподавания.

«МО сыграло жизненно важную роль в преобразовании традиционных методов обучения, позволяя преподавателям анализировать обширные наборы данных и выявлять закономерности, которые помогают ускорить процесс обучения» [29].

Используя прогностическую аналитику, алгоритмы МО могут оценивать успеваемость и поведение учащихся, одновременно облегчая создание персонализированных траекторий обучения и адаптивных структур учебных программ.

2.1 Технологии персонализированного онлайн обучения на основе машинного обучения

Анализ литературы и источников по теме исследования позволил выделить следующие технологии персонализированного онлайн обучения, созданные на основе ИИ и МО:

- адаптивное обучение. Платформы адаптивного обучения, основанные на алгоритмах МО, сыграли жизненно важную роль в переосмыслении образовательной среды. Эти платформы способны оценивать успеваемость обучающихся и соответствующим образом адаптировать учебную программу. Обучающиеся, испытывающие трудности, получают дополнительную поддержку, в то время как отличникам предлагаются более продвинутые учебные материалы. Такой персонализированный подход не только способствует более глубокому пониманию предмета, но и повышает общую вовлеченность и результаты обучения;
- создание персонализированного контента. В эпоху цифровых технологий МО внедрило сложные системы рекомендаций, которые играют ключевую роль в настройке учебных материалов в соответствии с уникальным стилем обучения и интересами каждого учащегося. Подобно тому, как Netflix предлагает фильмы или Amazon рекомендует продукты, эти системы предлагают обучающимся специальные учебные ресурсы, такие как обучающие видеоролики, литературу или упражнения, которые соответствуют их предпочтениям и способствуют более глубокому и

увлекательному обучению;

- предоставление индивидуальной обратной связи в режиме онлайн. Одним из наиболее значительных вкладов машинного обучения в образование является возможность предоставлять обучающимся персонализированную обратную связь в режиме реального времени. Модели машинного обучения тщательно анализируют работу обучающихся, выявляя индивидуальные сильные и слабые стороны. Предлагая индивидуальную обратную связь, обучающиеся получают всестороннее представление о своей успеваемости, способствуют постоянному совершенствованию и способствуют созданию проактивной среды обучения;
- гибкий подход к обучению. Все мы знаем, что у каждого обучающегося свой уникальный темп обучения, поэтому алгоритмы МО используют очень гибкий и адаптивный подход к обучению. Благодаря этому обучающиеся могут продвигаться по учебной программе с собственной скоростью, обеспечивая при этом полное усвоение сложных концепций, прежде чем двигаться дальше. Этот персонализированный подход способствует более глубокому пониманию предмета, что способствует более всестороннему и целостному обучению;
- обеспечение лучшей вовлеченности. Предоставление индивидуальных учебных материалов и ресурсов, соответствующих интересам и способностям обучающихся, эффективно способствует повышению их вовлеченности. Такой подход гарантирует, что обучающиеся постоянно мотивированы и заинтригованы содержанием обучения, делая образовательный процесс более приятным и вознаграждаемым. В результате обучающиеся активно участвуют в учебном процессе, что способствует формированию культуры энтузиазма и любознательности;
- совершенствование учебной программы с помощью данных. МО

дает преподавателям возможность анализировать обширные наборы данных об успеваемости обучающихся для выявления ценных идей и тенденций. Это может послужить основой для разработки учебных программ, методик преподавания и персонализированных планов обучения. Используя информацию, основанную на данных, преподаватели могут адаптировать учебные стратегии в соответствии с конкретными потребностями и учебными предпочтениями обучающегося. «Такой подход к образованию, ориентированный на данные, не только повышает успеваемость обучающихся, но и предоставляет преподавателям инструменты для постоянного совершенствования и оптимизации учебной среды» [32].

Анализ показал, что наилучшего результата можно достичь при использовании комплексного подхода при реализации концепции персонализированного обучения на основе ИИ и МО.

2.2 Методологии построения систем формирования индивидуальной траектории онлайн обучения на основе технологий искусственного интеллекта и машинного обучения

С помощью инструментов на базе искусственного интеллекта учителя могут лучше понимать сильные и слабые стороны своих учеников и разрабатывать персонализированные планы обучения, которые помогут каждому ученику добиться успеха.

Таким образом, раннее выявление пробелов в обучении имеет решающее значение для обеспечения учащимся прочной основы для дальнейшего развития. ИИ может стать отличным инструментом для выявления этих пробелов и предоставления адресной поддержки.

Платформы на базе ИИ могут собирать и анализировать данные учащихся о взаимодействии с учебными материалами, времени выполнения

упражнений, результатах тестов и общей успеваемости, чтобы понять отношения и потребности каждого учащегося. Опираясь на эти данные, инструменты искусственного интеллекта могут разрабатывать персонализированные маршруты обучения и адаптировать их в режиме реального времени к прогрессу учащегося. Например, если система на базе искусственного интеллекта обнаруживает, что у учащегося возникают трудности с выполнением определенного задания (включая тесты и упражнения), она может рекомендовать ему дополнительный материал для изучения. В то же время система также может рекомендовать тесты для самопроверки на основе изученного обучающимся материала.

Как показал анализ, для формирования ИОТ с помощью ИИ и МО используются большие образовательные данные (Big Data in Education).

Большие данные трансформируют образовательный ландшафт, предоставляя ценную информацию о поведении учащихся, моделях обучения и академической успеваемости. С помощью аналитики больших данных теперь возможно разрабатывать персонализированные учебные планы, которые удовлетворяют индивидуальные потребности учащихся, улучшают взаимодействие учителя и ученика и повышают общую вовлеченность учащихся (рисунок 9) [31], [32].

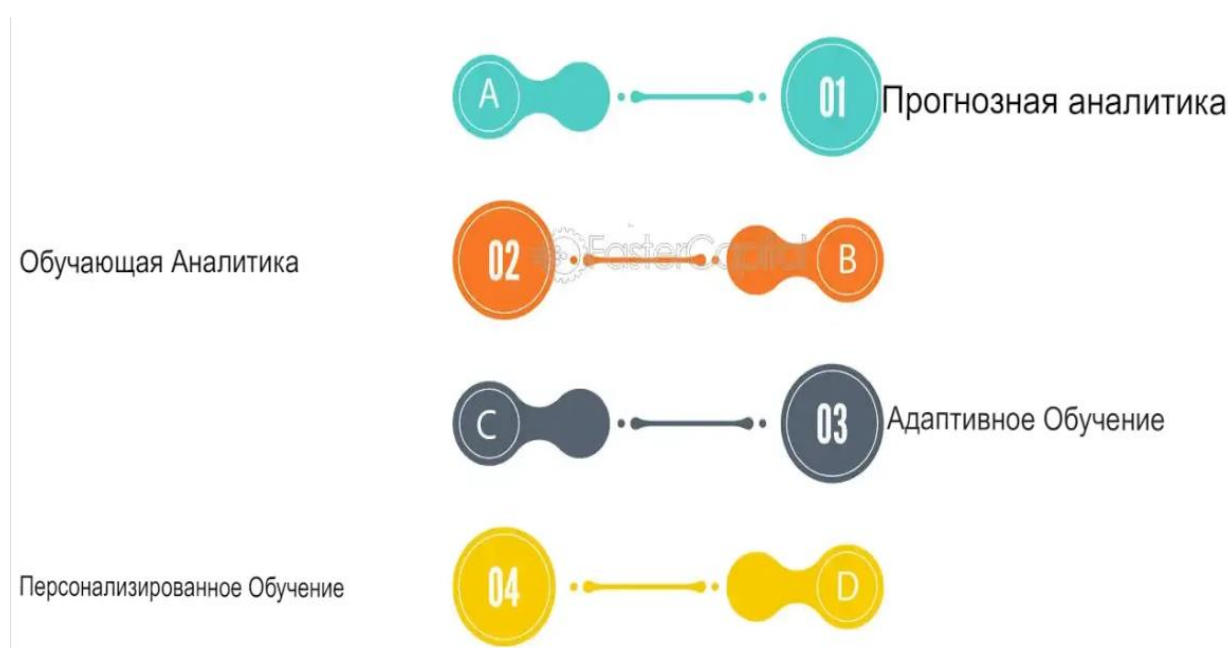


Рисунок 9 – Применение Big Data в образовании

В работе [22] описан алгоритм формирования маршрута персонализированного обучения с помощью механизма ИИ, блок-схема которого показана на рисунке 10.

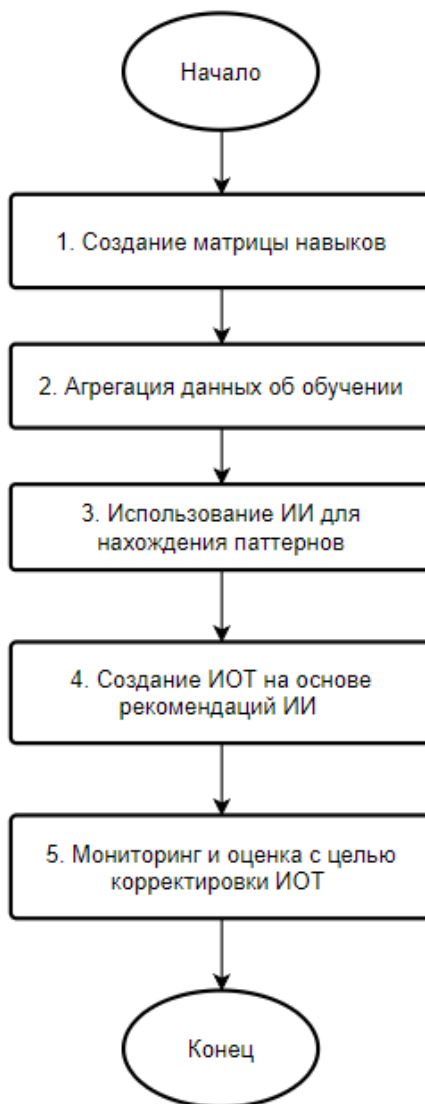


Рисунок 10 – Блок-схема алгоритма формирования маршрута персонализированного обучения с помощью механизма ИИ

Алгоритм состоит из следующих шагов:

Шаг 1. Создание матрицы навыков. Чтобы определить знания учащегося и затем построить профиль обучения, начните с создания матрицы навыков. Для этого навыки должны быть описаны человеком один раз. Это можно сделать, проанализировав резюме сотрудников с помощью текстовой

аналитики. Или сотрудники могут заполнить простую форму, указав навыки в строках и уровни знаний в столбцах. А затем эти данные будут проанализированы машиной. В результате у каждого учащегося будет индивидуальный профиль, отражающий его текущий уровень знаний и навыков.

Шаг 2. Агрегация данных об обучении. Следует учесть, что знание только навыков не даст полной картины. ИИ фиксирует, собирает и анализирует информацию об обучении сотрудников из различных источников внутри компании. «Это можно легко сделать с помощью xAPI. xAPI (сокращенно от Experience API, ранее Tin Can API) - это открытая спецификация, которая описывает формат передачи статистики между курсом / сайтом / приложением / любым другим провайдером учебной активности и базой данных LRS (Learning Record Store, хранилище учебных активностей)» [2]. С помощью xAPI вы будете собирать информацию о предыдущем и текущем опыте обучения сотрудников в LMS, LXP, TMS, HRIS и т. д. с высокой степенью детализации.

Шаг 3. Использование ИИ для нахождения паттернов. Объединив все данные, ИИ определяет тенденции и выстраивает закономерности для каждого учащегося. Эти данные дают администраторам или команде L&D ценную информацию об успеваемости учащихся и образовательных предпочтениях. На следующем этапе эта информация будет использоваться для предложения наиболее релевантного контента.

Шаг 4. Создание ИОТ на основе рекомендаций ИИ. Определяя тенденции обучения, система может предоставлять учащимся учебные материалы в тех форматах и с тем уровнем сложности, которые лучше всего соответствуют их предпочтениям и навыкам в обучении. Например, ИИ сообщает администратору, что определенный сотрудник использует короткие видеоролики в качестве основного контента и склонен смотреть темы «Техническое SEO». На основе этих данных система рекомендаций начинает предлагать учащемуся больше видео вместо текстового формата по

предпочтительной теме. Благодаря получению актуального контента в предпочтительном формате больше учащихся преуспевают в учебном курсе.

Шаг 5. Мониторинг и оценка с целью корректировки ИОТ. Учащиеся начинают обучение с использованием LMS или LXP. Успеваемость каждого учащегося регулярно измеряется и оценивается. Механизм ИИ учитывает последние результаты ученика и корректирует курс. И этот цикл постоянно повторяется. Любой бизнес, который внедряет ИИ в учебные курсы, получает стратегическое преимущество благодаря возможности обнаружить недостающие навыки у сотрудников и организовать эффективное повышение квалификации и переквалификацию для своевременного решения этой проблемы.

МО может быть использовано для формирования индивидуальных планов обучения, учитывая различные факторы, такие как уровень знаний ученика, его стиль обучения, предпочтения и прогресс.

Вот некоторые методы, которые могут быть применены [14]:

- кластеризация. Этот метод может быть использован для группировки обучающихся на основе их профилей обучения. Затем для каждой группы можно разработать специфические планы обучения;
- рекомендательные системы. Эти системы могут предложить индивидуальные ресурсы для обучения на основе предыдущих взаимодействий ученика и схожести с другими учениками.
- обучение с подкреплением. Этот метод может быть использован для адаптации плана обучения в реальном времени, учитывая обратную связь от ученика.
- прогнозирование успеваемости. Модели машинного обучения могут предсказать, как ученик будет справляться с определенными задачами или темами, что позволяет адаптировать план обучения для максимизации успеха.

В работе [28] предлагается методология, основанная на построении многоалгоритмической совместной персонализированной модели

рекомендаций по траектории обучения, которая предоставляет рекомендации по обучению для учащихся платформ онлайн-обучения.

Модель учащегося строится с четырех точек зрения: когнитивный уровень, способность к обучению, стиль обучения и интенсивность обучения.

Алгоритм правил ассоциации используется для генерации последовательности точек знаний и планирования последовательности обучения учащихся.

Алгоритм роевого интеллекта используется для обеспечения сопоставления каждой точки знаний с персонализированными учебными ресурсами с более высокой степенью адаптивности, чтобы учащиеся могли учиться, используя более целенаправленный подход.

На рисунке 11 показана архитектура фреймворка формирования персонализированного маршрута обучения.

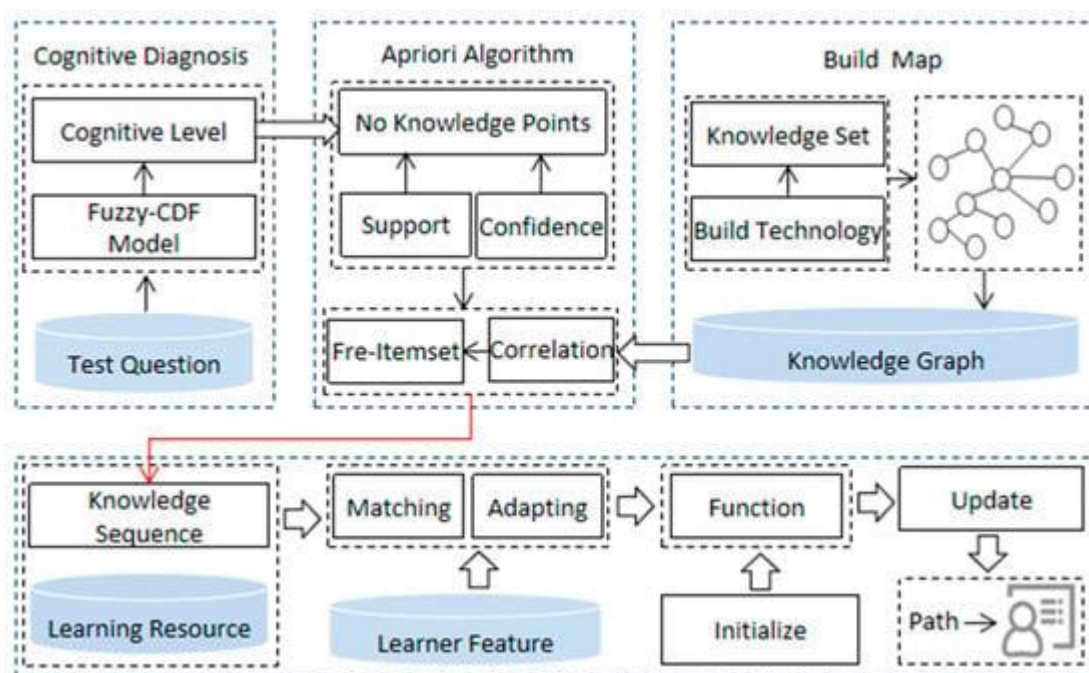


Рисунок 11 – Архитектура фреймворка формирования персонализированного маршрута обучения

Результаты экспериментов показывают, что результаты исследования

данной модели могут в определенной степени рекомендовать реальные маршруты обучения целевым пользователям, эффективно повышать точность рекомендуемых ресурсов и, таким образом, улучшать качество обучения и эффект обучения пользователей.

В работе [34] описано применение образовательной технологии (EdTech) для создания персонализированного маршрута обучения с помощью ИИ и МО.

Компании, занимающиеся образовательными технологиями, используют большие образовательные данные для создания своих собственных моделей МО, чтобы воспользоваться преимуществами множества готовых сервисов искусственного интеллекта, доступных в Amazon Web Services (AWS) облаке.

Поскольку искусственный интеллект и машинное обучение становятся все более повсеместными и простыми в использовании, компании и стартапы EdTech по всему миру раскрывают преимущества этих технологий, особенно когда речь идет о персонализации учебного контента для конкретных потребностей учащегося или преподавателя.

В качестве примера приведена программа Lalia – интерактивная образовательная онлайн-платформа, базирующаяся в Сингапуре, которая помогает изучать английский язык лицам, не являющимся носителями английского языка, предоставляя адаптивные материалы, адаптированные к уровню владения языком каждого отдельного учащегося. Laila персонализирует свои учебные мероприятия для каждого отдельного учащегося с помощью ИИ, чтобы помочь учащимся расширить свой словарный запас и грамматику и улучшить их общение в темпе, соответствующем их потребностям в обучении.

Таким образом, ИИ и МО довольно широко применяются при построении систем формирования ИОТ. Вместе с тем более перспективными представляется подход, основанный на применении МО.

Для описания данного подхода используем диаграмму деятельности

UML, показанную на рисунке 12.

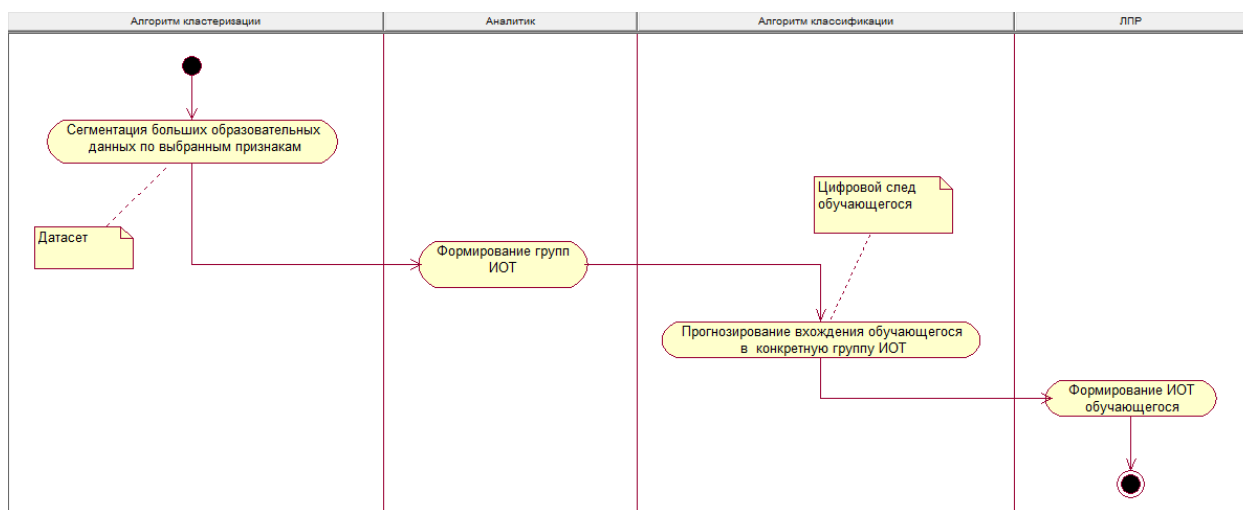


Рисунок 12 – Диаграмма деятельности системы формирования ИОТ на основе технологий МО

Данный подход выбран для построения системы формирования ИОТ обучающегося.

Выводы по главе 2

В результате проделанной работы были сделаны следующие выводы:

- наилучшего результата можно достичь при использовании комплексного подхода при реализации концепции персонализированного обучения на основе ИИ и МО;
- для формирования ИОТ с помощью ИИ и МО используются большие образовательные данные.

Основой объектной модели СФИОО является класс «Модель МО», который на логическом уровне представляет модель МО, используемую для создания прогноза вхождения обучающегося в конкретную поведенческую группу и последующего формирования его индивидуальной траектории онлайн обучения.

Глава 3 Разработка моделей и алгоритма системы формирования индивидуальной траектории онлайн обучения на основе технологий машинного обучения

3.1 Разработка логической модели системы формирования индивидуальной траектории онлайн обучения с применением технологий машинного обучения

Для разработки логической модели системы формирования индивидуальной траектории онлайн обучения (СФИТОО) с применением технологий машинного обучения используем объектно-ориентированный подход (ООП) к анализу и проектированию информационных систем [12].

Согласно ООП для создания логической модели системы достаточно разработать ее диаграммы вариантов использования и классов, которые отражают функциональный и структурный аспекты системы, соответственно.

Диаграммы разработаны с помощью CASE-средства Rational Rose [8].

Диаграммы вариантов использования обладают рядом преимуществ при разработке программного обеспечения и проектировании систем:

- сбор функциональных требований. Диаграммы вариантов использования помогают фиксировать функциональные требования системы. Визуально представляя взаимодействие между участниками (пользователями или внешними системами) и системой, они дают четкое представление о том, как система должна вести себя;
- прослеживаемость. Варианты использования легко отслеживаемы. Каждый вариант использования соответствует определенной функциональности или взаимодействию с пользователем. Такая прослеживаемость гарантирует соблюдение требований во время разработки и тестирования;
- основа для оценки и планирования. Варианты использования служат

основой для оценки, планирования и проверки затрат. Они помогают менеджерам проектов распределять ресурсы и планировать циклы разработки;

- эволюционное развитие. Варианты использования могут меняться на протяжении всего процесса разработки. Они начинаются как метод фиксации требований, затем служат руководством для программистов во время разработки, служат тестовыми примерами и, наконец, становятся частью пользовательской документации;
- альтернативные пути. Варианты использования позволяют фиксировать альтернативные пути или исключительное поведение. Эти пути повышают надежность системы за счет учета различных сценариев;
- удобное общение. Варианты использования легко понятны бизнес-пользователям. Они служат мостом между разработчиками программного обеспечения и конечными пользователями, способствуя эффективному общению;
- профессиональное моделирование. Вы можете создавать профессиональные модели вариантов использования с помощью инструментов моделирования UML.

Подводя итог, можно сказать, что диаграммы вариантов использования предоставляют мощный способ визуализации функциональности системы, передачи требований и обеспечения успешной разработки программного обеспечения. Они являются ценным инструментом в наборе инструментов для разработки программного обеспечения.

В процессе разработки диаграммы вариантов использования были выделены следующие акторы: Отв. сотрудник учебной методического управления (УМУ) образовательной организации, Обучающийся, Аналитик, Экспертная комиссия УМУ.

Описание вариантов использования СФИТОО представлено в таблицах 1-6.

Таблица 1 – Описание прецедента: Создать модель МО

«Элемент диаграммы	Описание
Прецедент	Создать модель МО
ID	1
Краткое описание	Создание модели МО
Главный актер	Аналитик
Второстепенный актер	Нет
Предусловие	Нет
Основной поток	Аналитик создает модель МО
Постусловие	Нет
Альтернативные потоки	Нет» [8]

Таблица 2 – Описание прецедента: Подготовить датасет

«Элемент диаграммы	Описание
Прецедент	Подготовить датасет
ID	2
Краткое описание	Подготовка датасета обучающей выборки
Главный актер	Аналитик
Второстепенный актер	Нет
Предусловие	Нет
Основной поток	Аналитик готовит датасет
Постусловие:	Нет
Альтернативные потоки	Нет» [8]

Таблица 3 – Описание прецедента: Сформировать поведенческие группы обучающихся

«Элемент диаграммы	Описание
Прецедент	Сформировать поведенческие группы обучающихся
ID	3
Краткое описание	Формирование поведенческих групп обучающихся
Главный актер	Экспертная комиссия УМУ
Второстепенный актер	Нет
Предусловие:	Нет
Основной поток	Экспертная комиссия УМУ формирует поведенческие группы обучающихся
Постусловие	Нет
Альтернативные потоки	Нет» [8]

Таблица 4 – Описание прецедента: Сформировать цифровой след обучающегося

«Элемент диаграммы	Описание
Прецедент	Сформировать цифровой след обучающегося
ID	4
Краткое описание	Формирование цифрового следа обучающегося
Главный актер	Аналитик
Второстепенный актер	Нет
Предусловие	Нет
Основной поток	Аналитик формирует цифровой след обучающегося
Постусловие	Нет
Альтернативные потоки	Нет» [8]

Таблица 5 – Описание прецедента: Спрогнозировать вхождение обучающегося в поведенческую группу

«Элемент диаграммы	Описание
Прецедент	Спрогнозировать вхождение обучающегося в поведенческую группу
ID	5
Краткое описание	Создать прогноз вхождения обучающегося в конкретную поведенческую группу
Главный актер	Аналитик
Второстепенный актер	Нет
Предусловие	Нет
Основной поток	Аналитик создает прогноз вхождения обучающегося в конкретную поведенческую группу
Постусловие:	Нет
Альтернативные потоки	Нет» [8]

Таблица 6 – Описание прецедента: Сформировать ИОТ

«Элемент диаграммы	Описание
Прецедент	Сформировать ИОТ
ID	6
Краткое описание	Формирование ИОТ обучающегося
Главный актер	Отв. сотрудник УМУ
Второстепенный актер	Нет
Предусловие	Прогноз вхождения обучающегося в группу ИОТ
Основной поток	Отв. сотрудник УМУ формирует ИОТ обучающегося
Постусловие:	Нет
Альтернативные потоки	Нет» [8]

Диаграмма вариантов СФИТОО показана на рисунке 13.

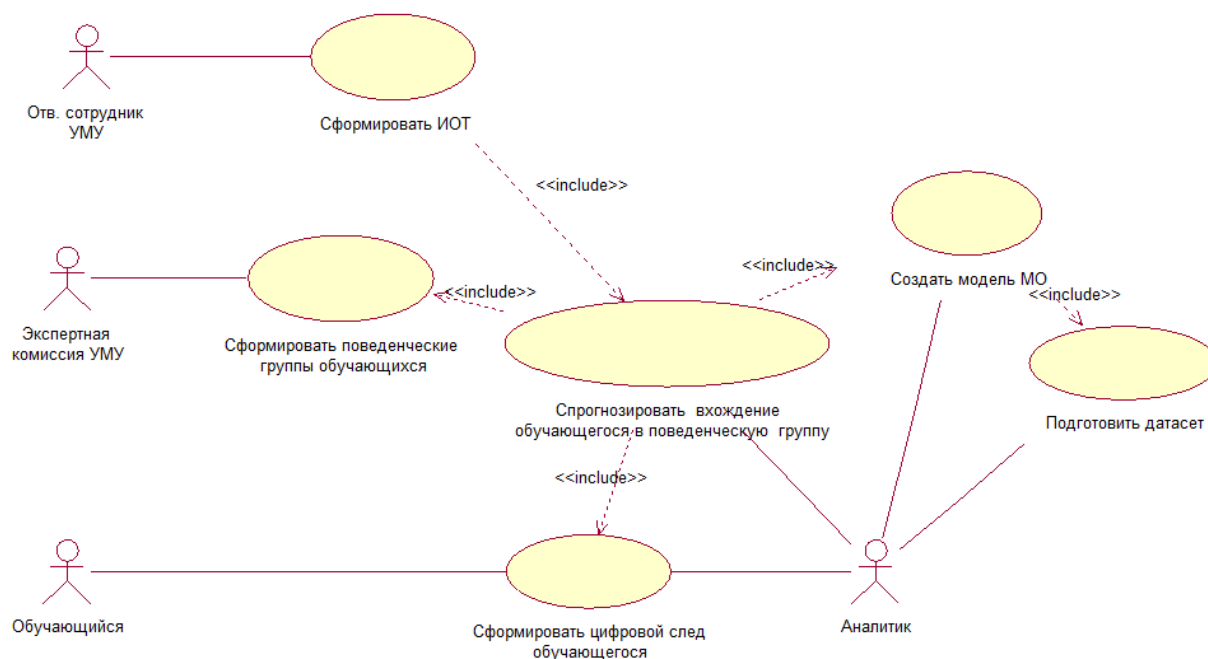


Рисунок 13 – Диаграмма вариантов использования СФИОО

Диаграмма классов – ценный инструмент при разработке программного обеспечения и проектировании систем.

Рассмотрим их преимущества:

- иллюстрация модели данных. Диаграммы классов позволяют иллюстрировать модели данных для информационных систем, независимо от их сложности. Они обеспечивают визуальное представление того, как классы, атрибуты и отношения сочетаются друг с другом;
- обзор системы. Создавая диаграммы классов, вы можете лучше понять общий обзор структуры приложения. На этих диаграммах показаны основные компоненты и их взаимодействие;
- коммуникация и распространение. Диаграммы классов помогают визуальным образом выразить конкретные потребности системы. Вы можете распространить эту информацию по всей организации, гарантируя,

что все понимают структуру системы.

- подробное руководство по программированию. При создании диаграммы классов позволяют сгенерировать конкретный код, необходимый для реализации. Они направляют программистов во время разработки и следят за тем, чтобы система соответствовала намеченной структуре;
- описания, независимые от реализации. Диаграммы классов предоставляют независимое от реализации описание типов, используемых в системе. Эти типы впоследствии можно будет передавать между различными компонентами системы.

Диаграмма классов СФИТОО показана на рисунке 14.

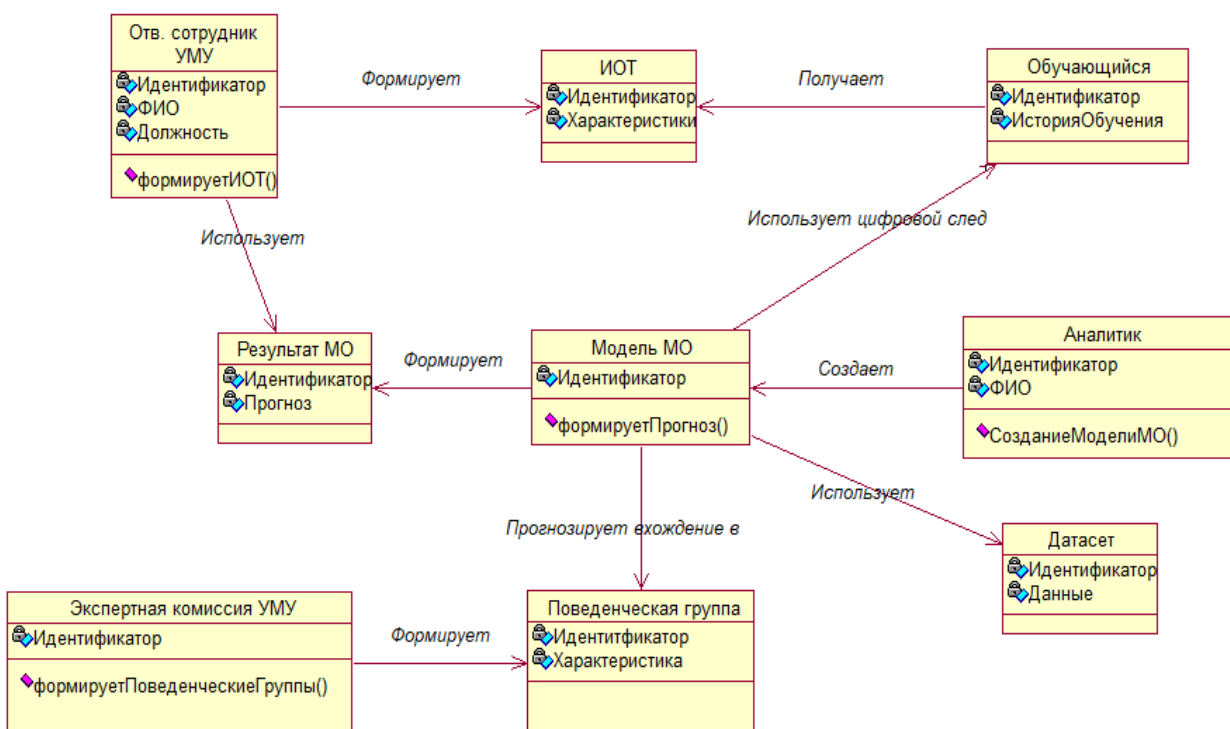


Рисунок 14 – Диаграмма классов СФИОО

Спецификация классов диаграммы показана в таблице 7.

Таблица 7 – Спецификация классов СФИОО

Класс	Описание
«Отв. сотрудник УМУ	Класс объектов, представляющих на логическом уровне сотрудников, формирующих ИОТ
Обучающийся	Класс объектов, представляющих на логическом уровне обучающихся
ИОТ	Класс объектов, представляющих на логическом уровне ОИТ
Модель МО	Класс объектов, представляющих на логическом уровне модели МО
Экспертная комиссия УМУ	Класс объектов, представляющих на логическом уровне экспертные комиссии УМУ
Поведенческая группа	Класс объектов, представляющих на логическом уровне поведенческие группы
Датасет	Класс объектов, представляющих на логическом уровне наборы данных для обучения модели МО
Аналитик	Класс объектов, представляющих на логическом уровне сотрудников-аналитиков
Результат МО	Класс объектов, представляющих на логическом уровне результаты выполнения алгоритмов МО» [8]

Основой объектной модели СФИОО является класс «Модель МО», который на логическом уровне представляет модель МО, используемую для создания прогноза вхождения обучающегося в конкретную поведенческую группу и последующего формирования его индивидуальной траектории онлайн обучения.

Таким образом, диаграммы классов служат мощным средством визуализации структуры системы, передачи требований и облегчения успешной разработки программного обеспечения.

Они являются важными строительными блоками в наборе инструментов UML.

3.2 Обзор и анализ алгоритмов машинного обучения для разработки системы формирования индивидуальной траектории онлайн обучения

Для разработки системы формирования индивидуальной траектории онлайн обучения с помощью технологий машинного обучения необходимо решить следующие задачи:

- сформировать поведенческие группы обучающихся с ИОТ;
- спрогнозировать вхождение конкретного обучающегося в определенную поведенческую группу по его цифровому следу.

Рассмотрим и проанализируем свойства популярных алгоритмов МО на предмет возможности решения указанных задач.

Для формирования поведенческих групп обучающихся с ИОТ необходимо произвести сегментацию данных.

«Сегментация данных – это процесс разделения большого и разнородного набора данных на более мелкие и более однородные подмножества на основе некоторых критериев или особенностей» [6].

Сегментация данных может быть полезна для многих целей, таких как улучшение обслуживания клиентов, таргетинг маркетинговых кампаний, оптимизация бизнес-процессов, повышение качества данных и т.д.

Алгоритмы кластеризации являются одним из основных методов сегментации данных, поскольку они могут группировать точки данных на основе их сходства или несходства, не требуя каких-либо предварительных знаний или меток.

Алгоритмы кластеризации - это методы машинного обучения без учителя, которые позволяют группировать данные по их сходству или различию. Цель кластеризации - найти скрытую структуру данных и выделить естественные группы или кластеры.

Кластеризация может быть полезна для многих задач, таких как сегментация рынка, анализ социальных сетей, сжатие изображений,

обнаружение аномалий и т.д. [9].

Существует множество различных алгоритмов кластеризации, каждый из которых имеет свои преимущества и недостатки.

Одним из наиболее популярных алгоритмов является алгоритм k-means, который разделяет данные на k кластеров, учитывая их близость друг к другу.

Другим распространенным алгоритмом является агломеративная кластеризация, которая основана на иерархической кластеризации, и формирует древовидную структуру, представляющую иерархию кластеров.

Есть и другие типы алгоритмов кластеризации, такие как спектральная кластеризация, DBSCAN, OPTICS, BIRCH и т.д.

Для оценки качества кластеризации можно использовать различные метрики, такие как индекс Дэвиса-Болдуина, коэффициент силуэта, индекс Данна, индекс Колдера-Харабаса и т.д.

Эти метрики учитывают внутрикластерное и межкластерное расстояние, а также количество кластеров. Однако нет единого правила для выбора оптимального алгоритма или метрики кластеризации, поскольку это зависит от цели и характера данных [40].

Следует учесть, что разные алгоритмы кластеризации могут иметь разные предположения, цели и результаты в зависимости от характеристик данных и целей сегментации.

Поэтому важно сравнивать и оценивать производительность различных алгоритмов кластеризации для сегментации данных.

Как показал анализ, на сравнение алгоритмов кластеризации для сегментации данных могут повлиять следующие факторы:

- тип данных: числовые, категориальные, смешанные, многомерные и т. д.;
- форма и размер кластеров: шаровидные, вытянутые, перекрывающиеся и т.д.;
- количество кластеров: заданное, неизвестное, фиксированное,

- переменное и т.д.;
- метод кластеризации: иерархический, секционный, по плотности, спектральный и т.д.;
- критерий кластеризации: расстояние, плотность, связность, модульность и т.д.;
- проверка кластеризации: внутренняя, внешняя, относительная, на стабильность и т. д.;
- оценка кластеризации: точность, чистота, энтропия, силуэт и т. д.

Для сравнения методов и выбора алгоритма кластеризации используем таблицу 8 [37].

Таблица 8 – Сравнение характеристик алгоритмов кластеризации для сегментации данных

Метод	Преимущества	Недостатки	Алгоритмы
Кластеризация на основе центроидов	Простой, быстрый и понятный. Может включать нечеткость.	Не существует естественного способа обнаружить выбросы. Ограничено данными, где центры (центроиды) легко различимы. Не может соответствовать произвольному (невыпуклому) формы кластера	K-means Нечеткий C-means
Кластеризация на основе плотности	Обнаружение выбросов как побочный эффект кластеризации. Количество кластеров определяется алгоритмом. Произвольные формы кластеров.	Определенное количество кластеров может не соответствовать ожиданиям (например, определено слишком много кластеров).	DBSCAN OPTICS

Продолжение таблицы 8

Метод	Преимущества	Недостатки	Алгоритмы
Кластеризация на основе подключения	Обнаруживает иерархическую структуру	Обнаружение выбросов довольно проблематично	Метод Уорда

По результатам сравнения выбираем для сегментации данных алгоритм K-means как наиболее простой для реализации и понятный.

Для прогнозирования вхождение конкретного обучающегося в определенную поведенческую группу по его цифровому следу используем алгоритмы классификации.

Алгоритмы классификации — это контролируемые методы машинного обучения, которые присваивают метку класса входному экземпляру на основе некоторых функций.

Существует множество типов алгоритмов классификации, каждый из которых имеет свои сильные и слабые стороны. Некоторые из наиболее распространенных из них [36]:

- «ближайшие соседи» (KNN): этот алгоритм присваивает метку класса входному экземпляру на основе меток классов его ближайших соседей в пространстве объектов. Он прост, интуитивно понятен и может обрабатывать нелинейные границы, но может быть медленным, чувствительным к шуму и требовать много памяти;
- линейный SVM: этот алгоритм находит линейную гиперплоскость, разделяющую классы с максимальным запасом. Он быстрый, надежный и может обрабатывать многомерные данные, но может обрабатывать только линейно разделимые данные и может быть чувствителен к выбросам;
- RBF SVM: этот алгоритм похож на линейный SVM, но он использует ядро радиальной базисной функции для отображения данных в многомерное пространство, где их можно линейно разделить. Он может обрабатывать нелинейные границы и сложные данные, но

- может быть медленным, склонным к переобучению и требовать настройки гиперпараметров;
- Гауссов процесс: этот алгоритм представляет собой вероятностную модель, которая присваивает метку класса входному экземпляру на основе апостериорного распределения класса с учетом функций. Он может обрабатывать неопределенность, шум и нелинейные границы, но может быть очень медленным, дорогостоящим в вычислительном отношении и требовать большого количества данных;
 - дерево решений: этот алгоритм создает древовидную структуру, которая разбивает данные на основе некоторых критериев на каждом узле. Его легко интерпретировать, он может обрабатывать категориальные признаки и фиксировать нелинейные отношения, но он может быть нестабильным, склонным к переобучению и чувствительным к выбору критериев разделения;
 - случайный лес: этот алгоритм представляет собой ансамблевый метод, который объединяет несколько деревьев решений и усредняет их прогнозы. Он может уменьшить дисперсию и переобучение одного дерева решений, а также обрабатывать пропущенные значения, выбросы и несбалансированные данные, но он может быть медленным, трудным для интерпретации и требовать настройки гиперпараметров;
 - нейронная сеть: этот алгоритм представляет собой сеть взаимосвязанных узлов, которые учатся сопоставлять объекты с метками классов посредством нескольких уровней нелинейных преобразований. Он может обрабатывать сложные, многомерные и нелинейные данные и может учиться на любых типах данных, но он может быть очень медленным, его трудно обучать и требовать большого количества данных и настройки гиперпараметров;
 - AdaBoost: этот алгоритм представляет собой еще один ансамблевый метод, который итеративно обучает слабые классификаторы и

комбинирует их с весами в зависимости от их производительности. Он может повысить точность и надежность одного классификатора, а также обрабатывать шум и несбалансированные данные, но может быть чувствителен к выбросам и переобучению и требовать настройки гиперпараметров;

- наивный Байес: этот алгоритм представляет собой вероятностную модель, которая присваивает метку класса входному экземпляру на основе априорной вероятности класса и условной вероятности свойств данного класса. Он быстрый, простой и может обрабатывать многомерные данные, но предполагает независимость между функциями, что может не соответствовать действительности, и может плохо работать с несбалансированными данными;
- QDA: этот алгоритм похож на простой байесовский алгоритм, но он ослабляет предположение о независимости и позволяет функциям иметь разные распределения для каждого класса. Он может фиксировать ковариацию между функциями и обрабатывать нелинейные границы, но может быть чувствителен к выбросам и переобучению и требовать большого количества данных и параметров.

Лучший алгоритм классификации для конкретной задачи зависит от многих факторов, таких как размер, качество и характер данных, сложность и линейность границ, доступные вычислительные ресурсы и время, а также желаемая точность и интерпретируемость модели.

Не существует универсального решения, и часто рекомендуется опробовать несколько алгоритмов и сравнить их производительность, используя соответствующие показатели и статистические тесты.

Для сравнения популярных методов и выбора алгоритма классификации используем таблицу 9 [37].

Таблица 9 – Сравнение характеристик алгоритмов классификации и прогнозирования

Алгоритм	Преимущества	Недостатки
KNN	Простой и гибкий алгоритм, который может обрабатывать как числовые, так и категориальные данные. Он легко адаптируется к новым данным и менее чувствителен к выбросам.	Для больших наборов данных может быть медленным и дорогостоящим с точки зрения вычислений, поскольку требует расчета расстояния между каждой парой точек. Это также зависит от выбора количества соседей (K) и метрики расстояния.
Дерево решений	Графический и интуитивно понятный алгоритм, который может выполнять задачи как классификации, так и регрессии. Он может обрабатывать нелинейные отношения и взаимодействия функций, его легко интерпретировать и объяснять.	В зависимости от глубины и сложности дерева оно может быть подвержено переобучению или недостаточному оснащению. Оно также страдает от нестабильности, поскольку небольшие изменения в данных могут привести к большим изменениям в структуре дерева.
SVM	Мощный и надежный алгоритм, который может решать линейные и нелинейные задачи, а также многомерные и разреженные данные. Он может достичь высокой точности и производительности обобщения, используя трюк с ядром и методы регуляризации.	Его может быть сложно понять и настроить, поскольку он включает в себя выбор подходящей функции ядра и гиперпараметров. Также может быть медленным и ресурсоемким для больших наборов данных, поскольку требует решения задачи квадратичной оптимизации.

По результатам сравнения выбираем для решения задачи прогнозирования алгоритм KNN как наиболее простой и гибкий [25].

3.3 Описание используемых алгоритмов машинного обучения

3.3.1 Алгоритм K-means

Блок-схема алгоритма K-means показана на рисунке 15.

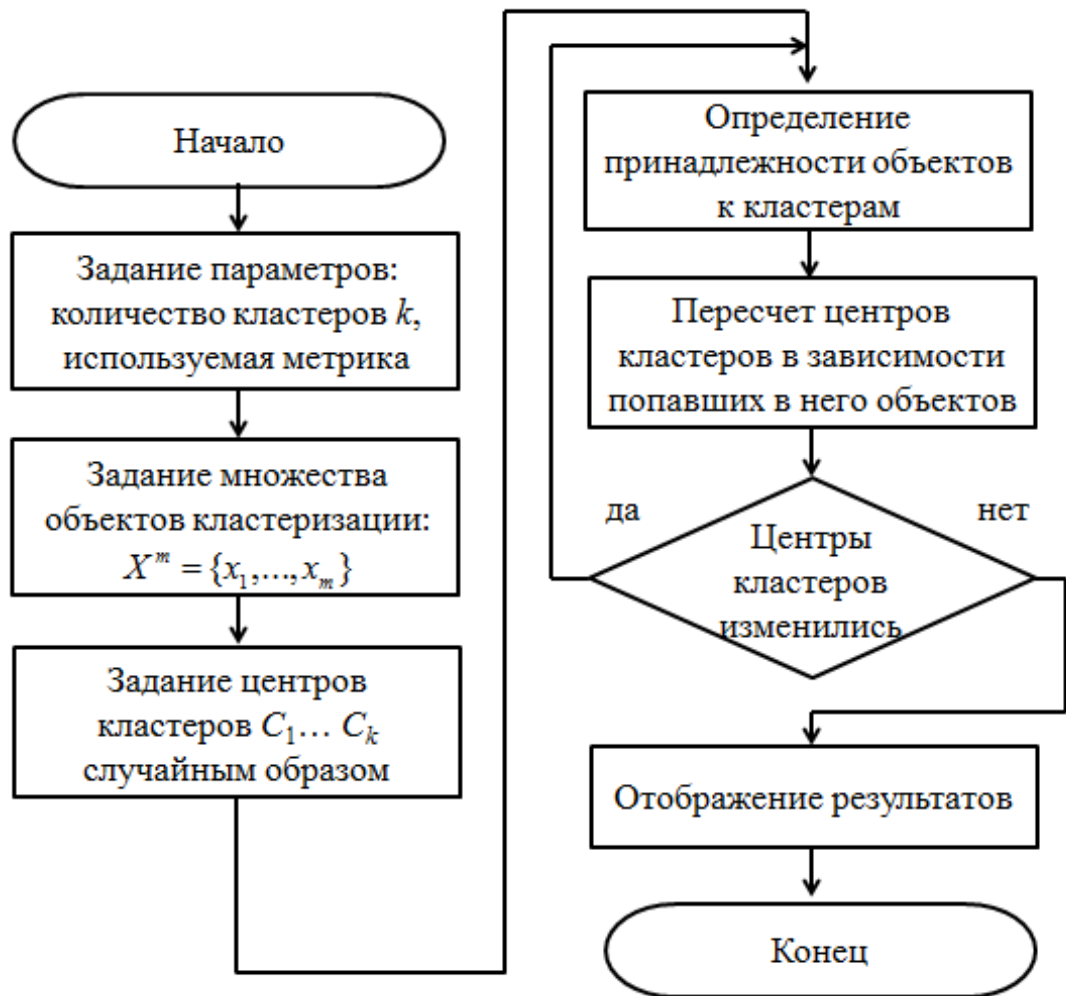


Рисунок 15 – Блок-схема алгоритма K-means

«Алгоритм K-means состоит из следующих шагов:

Шаг 1. Случайным образом выбирается k объектов обучающей выборки, которые будут служить начальными центрами кластеров.

Шаг 2. Для каждого объектов обучающей выборки определяется ближайший к ней центр кластера. Для этого вычисляется расстояние между объектами и центрами кластеров. Считается, что объект принадлежит тому кластеру, к которому он ближе. В качестве формулы для оценки близости объектов в многомерном пространстве признаков используется одна из известных метрик.

Шаг 3. Как только состав кластеров на данной итерации известен,

производится расчёт новых центров кластеров. Это делается путем расчета средних значений для каждого числового признака по всем объектам рассматриваемого кластера. Например, в двухмерном пространстве координаты центр кластера на основе вошедших в него t объектов рассчитывается следующим образом (1):

$$(P_{2ц}, P_{1ц}) = \left(\frac{\sum_1^t P_1(t)}{t}, \frac{\sum_1^t P_2(t)}{t} \right) \quad (1)$$

Шаг 4. Шаги 2 и 3 повторяются до тех пор, пока не выполнятся один из двух критериев остановки:

- границы кластеров и расположения центров кластеров не перестанет изменяться от итерации к итерации, т.е. на каждой итерации в каждом кластере будет оставаться один и тот же набор записей. На практике алгоритм k-means обычно находит набор стабильных кластеров за несколько десятков итераций;
- достигнут критерий сходимости. Чаще всего используется критерий суммы квадратов ошибок между центром кластера и всеми вошедшими в него объектами (2):

$$E = \sum_{i=1}^k \sum_{p \in C_i} (p - m_i)^2 \quad (2)$$

где $p \in C_i$ - произвольная точка данных, принадлежащая кластеру C_i , m_i – центр данного кластера. Иными словами, алгоритм остановится тогда, когда ошибка E достигнет достаточно малого значения» [6].

3.3.2 Алгоритм KNN

Блок-схема алгоритма KNN показана на рисунке 16 [7].

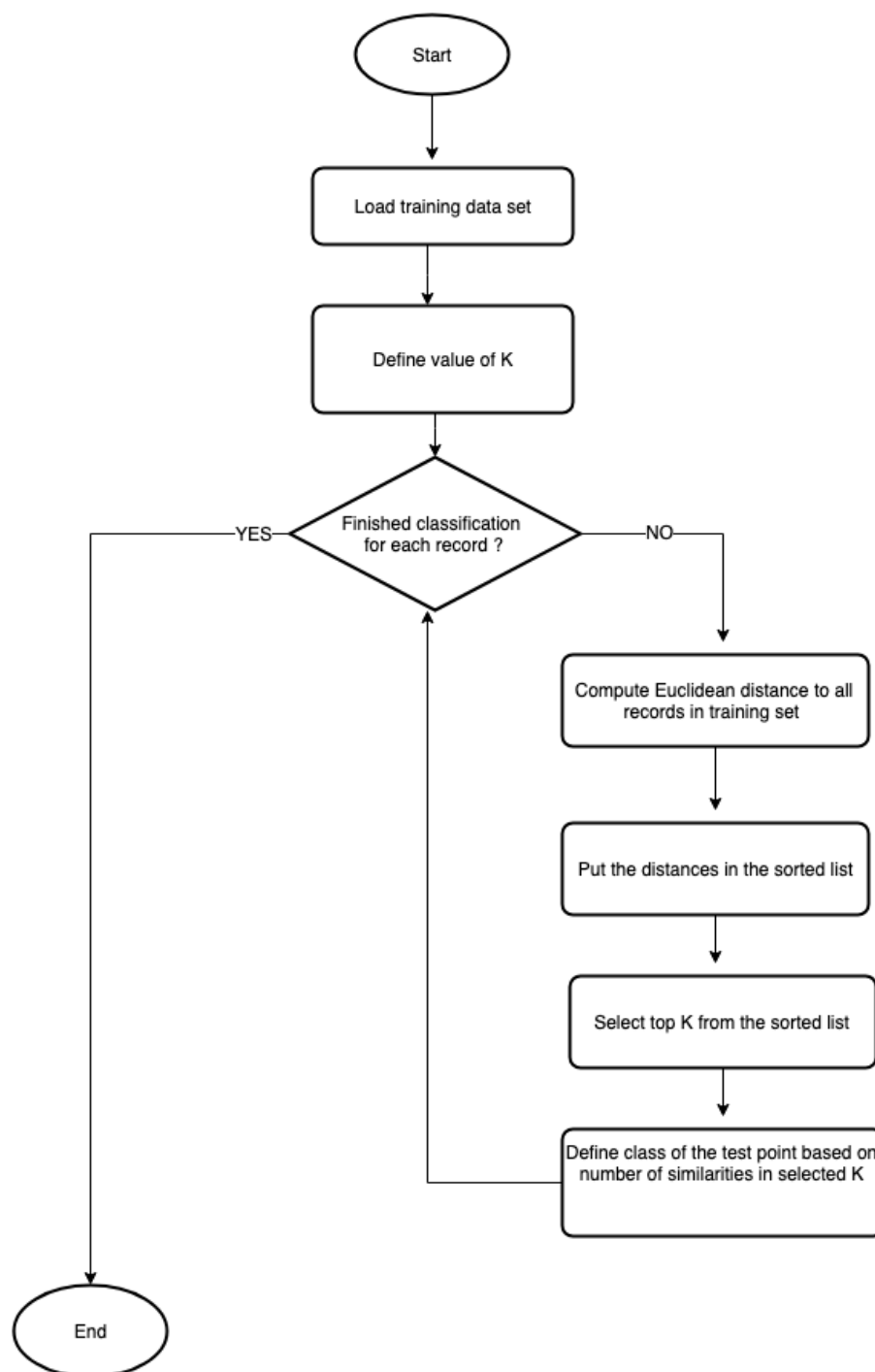


Рисунок 16 – Блок-схема алгоритма KNN

Алгоритм KNN, или k-ближайшие соседи, представляет собой метод контролируемого машинного обучения, который можно использовать для

решения задач классификации и регрессии.

Он работает путем поиска k точек данных, наиболее похожих на новую точку данных, на основе метрики расстояния, такой как евклидово расстояние. Затем он присваивает класс или значение новой точке данных на основе большинства голосов или среднего значения k соседей.

KNN — простой и гибкий алгоритм, который не делает никаких предположений о распределении данных, но может быть неэффективным и чувствительным к выбору k и метрики расстояния.

Алгоритм KNN состоит из следующих шагов:

Шаг 1: Выберите значение k — количество ближайших соседей, которые следует учитывать.

«Шаг 2. Рассчитайте расстояние между новой точкой данных и всеми существующими точками данных, используя метрику расстояния, например, евклидово расстояние.

Чтобы определить, какие точки данных наиболее близки к данной точке запроса, необходимо вычислить расстояние между точкой запроса и другими точками данных. Эти метрики расстояния помогают сформировать границы принятия решений, которые разделяют точки запроса на разные регионы.

Шаг 3: Отсортируйте расстояния в порядке возрастания и выберите k точек данных с наименьшими расстояниями.

Шаг 4: Подсчитайте количество точек данных в каждом классе среди k соседей» [5].

Шаг 5: Присвойте новую точку данных классу с самой высокой частотой.

Результатом является определение класса конкретного объекта.

Выводы к главе 3

В результате проделанной работы были сделаны следующие выводы:

– согласно ООП для создания логической модели системы

достаточно разработать ее диаграммы вариантов использования и классов, которые отражают функциональный и структурный аспекты системы, соответственно;

- алгоритмы кластеризации являются одним из основных методов сегментации данных, поскольку они могут группировать точки данных на основе их сходства или несходства, не требуя каких-либо предварительных знаний или меток;
- по результатам сравнения для сегментации данных выбран алгоритм K-means как наиболее простой для реализации и понятный;
- для прогнозирования вхождение конкретного обучающегося в определенную поведенческую группу по его цифровому следу используются алгоритмы классификации.

По результатам сравнения для решения задачи прогнозирования выбран алгоритм KNN.

KNN — простой и гибкий алгоритм, который не делает никаких предположений о распределении данных, но может быть неэффективным и чувствительным к выбору k и метрики расстояния.

Глава 4 Апробация проектных решений и оценка их эффективности

4.1 Разработка элементов системы формирования индивидуальной траектории онлайн обучения

Рассмотрены и проанализированы аналоги системы формирования индивидуальной траектории онлайн обучения (СФТИОО).

Система формирования персонализированного плана обучения по специальности «Управление Big data» через Интернет вещей и алгоритм кластеризации предназначена для построения систем знаний различных специальностей, обеспечения персонализированных подходов к обучению студентов, повышения эффективности преподавания преподавателей и учебных интересов студентов, а также для помощи студентам точно диагностировать свои проблемы в обучении [41].

Сначала анализируются существующие проблемы системы обучения и знаний студентов по дисциплине.

Затем строится модель персонализированного обучения на основе метода распределенных вычислений Интернета вещей и алгоритма кластеризации глубокого обучения.

Эта модель реализует эффективную классификацию учащихся с помощью алгоритма глубокого обучения в соответствии с данными обучения учащихся и разрабатывает разнообразное индивидуальное содержание обучения в соответствии с производительностью Интернета вещей.

Диаграмма формирования персонализированного плана обучения на основе ИИ в представленной системе показана на рисунке 17.

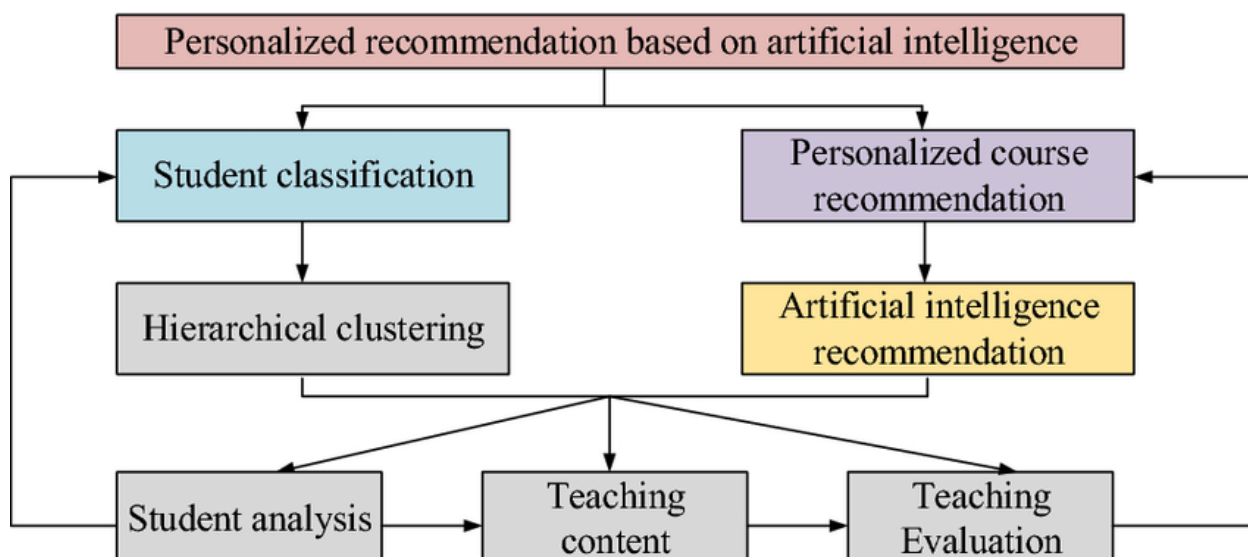


Рисунок 17 – Диаграмма формирования персонализированного плана обучения на основе ИИ

Онлайн ресурс StudyPlannr – это ИИ-компаньон на базе чат-бота, который генерирует персонализированные планы обучения для оптимизации повторения знаний для учащихся (рисунок 18).

Create Plan

1) Please enter your subjects for revision sequentially. Enter subjects with earliest exam dates first.
 2) Enter topics comma separated.
 3) There is a limit of 4 subjects and 10 topics per subject.

Plan Title
 Peter's plan ← Add plan title

Start Date
 dd/mm/yyyy

Subject

Subject Exam Date
 dd/mm/yyyy

Subject Topics
 Enter your topics here. Separate each topic with a comma.

Рисунок 18 – Окно для создания персонализированного плана обучения StudyPlannr

«Учитывая предметы, темы и даты оценивания, StudyPlannr создает индивидуальные планы обучения, которые помогают учащимся саморегулироваться и повышать эффективность повторения» [41].

Основные функции StudyPlannr:

- создание персонализированных планов обучения;
- учет предметов, тем и дат оценивания;
- саморегулирование и повышение эффективности пересмотра;
- загрузка в формате PDF для удобства использования;
- поддержка студентов в постановке целей и развитии хороших привычек в учебе.

Благодаря возможности скачивать планы в формате PDF или распечатывать их для удобства использования, StudyPlannr является ценным ресурсом для родителей и преподавателей, которые хотят помочь учащимся в постановке целей и формировании хороших учебных привычек.

Анализ показал, что рассмотренные решения ориентированы на особенности обеспечения индивидуального обучения в зарубежных странах.

Это снижает их возможности для использования в российских образовательных организациях и подтверждает актуальность разработки СФТИОО.

В таблице 10 представлен бизнес-план по разработке СФТИОО.

Таблица 10 – Бизнес-план по разработке СФТИОО

Раздел	Описание
Введение	Цель – составление плана по разработке СФТИОО.
Описание программного проекта	Назначение проекта: разработка СФТИОО. Обоснование: необходимость автоматизации процесса формирования ИОТ обучающихся. Основная бизнес-потребность - повышение эффективности процесса обучения в образовательной организации.

Продолжение таблицы 10

Раздел	Описание
Требования к ПО	<p>Функциональные требования: автоматизация процесса формирования ИОТ обучающихся.</p> <p>Требования к интерфейсу: дружелюбный интуитивный интерфейс.</p> <p>Требования к надежности: допустимая частота/периодичность сбоев: 1 раз в 300 часов, режим работы: рабочий день образовательной организации.</p> <p>Требования к производительности: допустимое количество одновременно работающих пользователей: 5</p> <p>Требования к поддержке: средствами ИТ-службы образовательной организации.</p> <p>Проектные требования: разработка ПО на языке Python.</p>
План качества	<p>Для обеспечения высокого качества ПО перед запуском СФТИОО в эксплуатацию сначала нужно создать прототип и протестировать его.</p>
План тестирования	<p>Тестирование состоит из следующих этапов: Функциональное тестирование; Бета-тестирование.</p>
План документации	<p>Техническая документация должна быть разработана в составе, указанном в разделе 5 «Состав и содержание работ по созданию СФТИОО» ТЗ, и должна удовлетворять требованиям комплекса стандартов и руководящих документов. Документация оформляется в соответствии с требованиями ГОСТ 2.105-95. Вся документация должна быть подготовлена и передана как в печатном, так и в электронном виде (в форматах DOC, DOCX или PDF) Заказчику.</p>
Результаты	<p>Результатом успешного выполнения проекта является приемка СФТИОО в опытную эксплуатацию.</p> <p>Контроль и приемка СФТИОО проводятся в процессе опытной эксплуатации в соответствии с принятой программой и методикой испытаний.</p>

Продолжение таблицы 10

Раздел	Описание
Предположения проекта	Основные ресурсы будут поставлены согласно графику; Участники проекта будут выполнять требования и соблюдать сроки выполнения проекта; Заказчик понимает необходимость начала и завершения проекта; Проект имеет организационную поддержку со стороны руководства Заказчика; В ОО имеется возможность выделить персонал для обеспечения работ по проекту; Заказчик и Исполнитель понимают необходимость обеспечения высокой организационной дисциплины по проекту.
План бета-тестирования	Бета-тестирование проводится сотрудниками УМУ образовательной организации. Производится проверка работоспособности СФТИОО, выявление возможных ошибок и их документирование созданием скриншотов. Отзывы тестировщиков передаются программисту для доработки. По результатам бета-тестирования составляется протокол по установленной форме.
Проблемы, возникающие при планировании разработки программного обеспечения	Нарушение сроков выполнения работ по разработке ПО.

Для разработки элементов СФТИОО используем язык Python, среду Jupyter Notebook и библиотеку машинного обучения Scikit-learn [33], [35].

Среда Jupyter Notebook обладает рядом преимуществ, которые делают ее популярным выбором среди специалистов по обработке данных и исследователей.

Преимущества Jupyter Notebook:

- комплексная среда: Jupyter Notebook предоставляет интерактивную веб-среду, которая сочетает в себе код, форматированный текст, изображения, видео, анимацию, математические уравнения, графики, карты, интерактивные рисунки, виджеты и графические элементы пользователя. интерфейсы в единый документ. Эта

- универсальность позволяет легко интегрировать код и документацию в одном интерфейсе;
- удобство совместного использования и совместной работы. Блокноты сохраняются в виде структурированных текстовых файлов (в формате JSON), что упрощает совместное использование. Вы можете сотрудничать с коллегами, предоставляя общий доступ к своим записным книжкам, позволяя им воспроизводить вашу работу или дополнять ее;
 - воспроизводимость. Блокноты Jupyter облегчают воспроизводимые эксперименты. Вы можете вести подробный учет своей работы, что облегчает повторное посещение и воспроизведение анализа позже. Простота использования гарантирует, что вам не придется беспокоиться о воспроизводимости; просто сосредоточьтесь на интерактивном исследовании.
 - поддержка выполнения кода в реальном времени: Jupyter позволяет выполнять ячейки кода в реальном времени, что делает его удобным для итеративной разработки и отладки. Вы можете запускать отдельные блоки кода и немедленно наблюдать за результатами;
 - документация в формате RTF. Вы можете включать в свой код форматированный текст, изображения и математические уравнения. Это упрощает создание комплексной документации в одном блокноте;
 - независимость от языка: изначально разработанный для Python (как IPython), Jupyter теперь поддерживает несколько языков программирования, включая Python, R, Julia и другие. Такая гибкость позволяет беспрепятственно работать с разными языками;
 - интерактивная визуализация данных. Блокноты Jupyter хорошо интегрируются с такими библиотеками, как Matplotlib, Seaborn и Plotly, обеспечивая интерактивную визуализацию данных

непосредственно в вашем блокноте;

- сегментированное выполнение кода. Вы можете запускать сегменты кода по отдельности, что полезно для понимания и отладки сложных рабочих процессов. Эта функция помогает поддерживать порядок и управляемость вашего кода;
- расширяемость. Экосистема Jupyter включает в себя различные интерфейсы (например, JupyterLab и nteract), библиотеки интерактивной визуализации и инструменты разработки, совместимые с ноутбуками. Это универсальная платформа, которую можно настроить в соответствии с индивидуальными потребностями;
- сообщество и финансирование. Jupyter приобрел огромную популярность среди исследователей, инженеров, специалистов по обработке данных и преподавателей. Он проводит собственную конференцию (JupyterCon) и получает финансирование от таких организаций, как Фонд Альфреда П. Слоана и Фонд Гордона и Бетти Мур.

Подводя итог, можно сказать, что Jupyter Notebook предоставляет мощную и гибкую среду для интерактивных вычислений, анализа данных и совместной работы. На рисунке 19 представлен программный код подготовки датасета для кластерного анализа.

```
Ввод [8]: 1 #Подготовка датасета для кластерного анализа
2 from sklearn.cluster import KMeans
3 import pandas as pd
4 import pathlib
5 import matplotlib.pyplot as plt
6 import seaborn as sns
7 p = pathlib.Path('leandataset.csv')
8 from sklearn.cluster import KMeans
9 indata = pd.read_csv(p)
10 indata.describe()
```

Рисунок 19 – Программный код подготовки датасета для кластерного анализа

Обучающей выборкой для кластеризации являются данные прохождения входных тестов обучающимися, желающими получить ИОТ.

Обучающая выборка представляет собой файл типа CSV, полученного из тестового XLSX-файла.

На рисунке 20 показаны статистические данные обучающего датафрейма.

Out[1]:

	x1	x2	x3	x4	x5
count	50.00000	50.00000	50.00000	50.00000	50.00000
mean	2.80000	3.60000	3.60000	4.20000	3.60000
std	1.17803	0.808122	1.030158	0.989743	1.030158
min	1.00000	3.00000	2.00000	3.00000	2.00000
25%	2.00000	3.00000	3.00000	3.00000	3.00000
50%	3.00000	3.00000	4.00000	5.00000	4.00000
75%	4.00000	4.00000	4.00000	5.00000	4.00000
max	4.00000	5.00000	5.00000	5.00000	5.00000

Рисунок 20 – Статистические данные обучающего датафрейма для кластеризации

Здесь x1-x5 – результаты тестирования обучающихся по выбранным дисциплинам.

На рисунке 21 представлен программный код для создания тепловой карты.

```
Ввод [2]: 1 #Тепловая карта
          2 plt.figure(figsize=(10,5))
          3 c= indata.corr()
          4 sns.heatmap(c, cmap="YlGnBu", annot=True)
          5 plt.title('Тепловая карта признаков набора данных')
          6 c
```

Рисунок 21 – Программный код для создания тепловой карты

Тепловая карта – это графическое представление данных, в котором для представления различных значений используется система цветового кодирования. Тепловые карты используются в различных формах аналитики, но чаще всего используются для отображения поведения пользователей на определенных веб-страницах или в шаблонах веб-страниц.

Что еще более важно, тепловые карты помогают классифицировать разделы, которые работают не на должном уровне и нуждаются в оптимизации [39].

Тепловая карта признаков обучающего набора данных показана на рисунке 22.



Рисунок 22 – Тепловая карта признаков обучающего набора данных

На рисунке 23 представлен программный код для выполнения кластерного анализа с помощью алгоритма K-means.


```

Ввод [6]: 1 #Кластерный анализ, алгоритм Kmeans
          2 nc=5
          3 kmeans = KMeans(nc)
          4 kmeans.fit(indata.values)
          5 print(kmeans.cluster_centers_)
          6

```

Рисунок 23 – Программный код для выполнения кластерного анализа с помощью алгоритма K-means

В результате кластеризации выборки по алгоритму K-means получены следующие кластеры (рисунок 24):

```

[[4. 5. 3. 3. 4.]
 [4. 4. 5. 5. 3.]
 [1. 3. 4. 5. 2.]
 [2. 3. 2. 3. 4.]
 [3. 3. 4. 5. 5.]]

```

Рисунок 24 – Результаты кластеризации обучающего набора данных

На рисунке 25 представлен программный код создания трехмерного графика представления кластеров x_1 , x_2 , x_3 , показанного на рисунке 26.

```

Ввод [53]: 1 import matplotlib.pyplot as plt
          2 fig = plt.figure(figsize=(13,13))
          3 font = {'fontname':'Book Antiqua'}
          4 ax = fig.add_subplot(projection='3d')
          5 ax.set_ylabel('x1', font,fontSize=10)
          6 ax.set_xlabel('x2', font,fontSize=10)
          7 ax.set_zlabel('x3',font, fontSize=10)
          8 ax.scatter(indata['x1'], indata['x2'],indata['x3'], c=kmeans.labels_,
          9           cmap='inferno',s=(100*kmeans.labels_+100),alpha=1)
          10
Out[53]: <mpl_toolkits.mplot3d.art3d.Path3DCollection at 0x212c91834d0>

```

Рисунок 25 – Программный код создания трехмерного графика результата кластеризации по признакам x_1 , x_2 , x_3

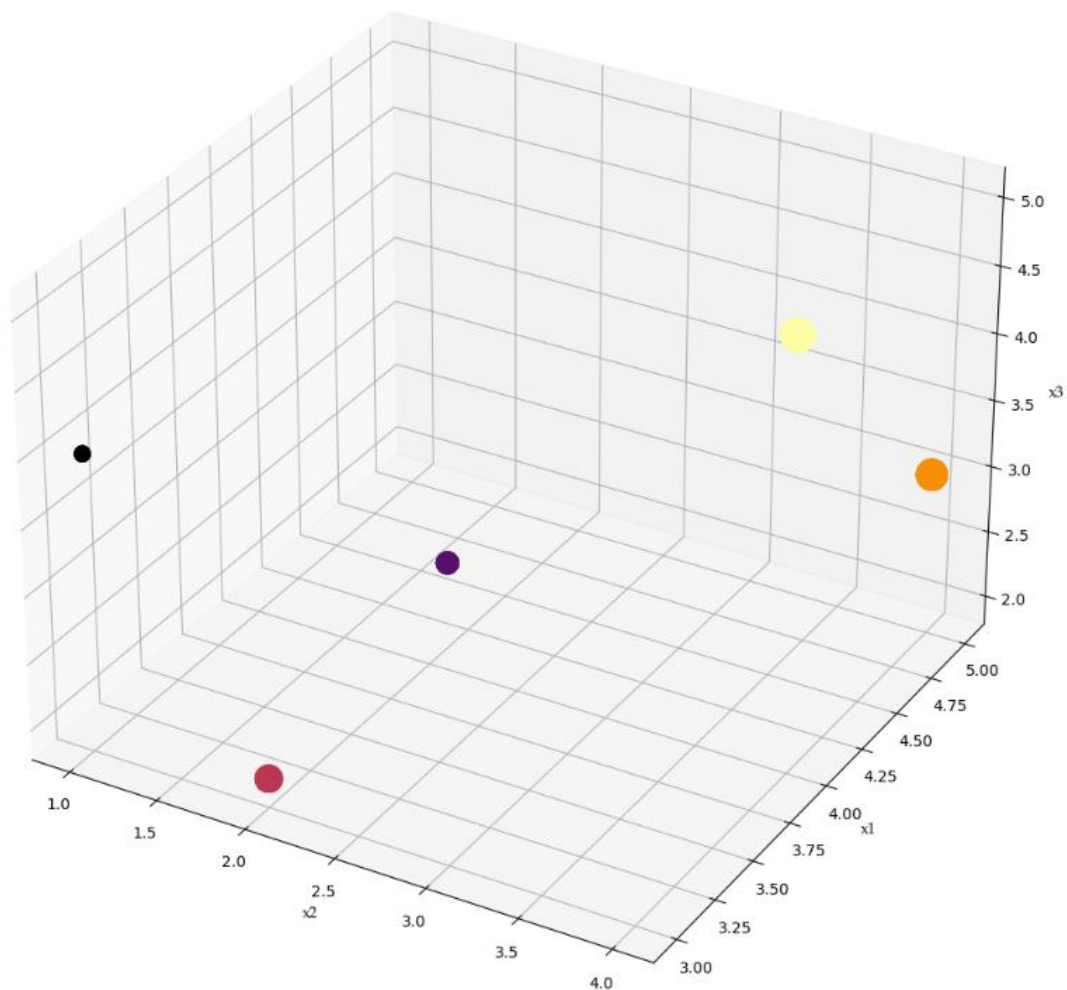


Рисунок 26 – График результата кластеризации по признакам x_1 , x_2 , x_3

Далее экспертная комиссия УМУ формирует поведенческие группы обучающихся.

На рисунке 27 представлен программный код подготовки обучающей выборки для классификатора.

```

Ввод [7]: 1 #Анализ датасета
          2 import pandas as pd
          3 import pathlib
          4 p2 = pathlib.Path('leandataset2.csv')
          5 dcf=pd.read_csv(p2)
          6 dcf.describe()

```

Рисунок 27 – Программный код подготовки обучающей выборки для классификатора

Результат формирования является обучающей выборкой для классификатора (рисунок 28):

Out[4]:

	k1	k2	k3	k4	k5	IOT
count	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000
mean	2.800000	3.600000	3.600000	4.200000	3.600000	3.000000
std	1.30384	0.894427	1.140175	1.095445	1.140175	1.581139
min	1.000000	3.000000	2.000000	3.000000	2.000000	1.000000
25%	2.000000	3.000000	3.000000	3.000000	3.000000	2.000000
50%	3.000000	3.000000	4.000000	5.000000	4.000000	3.000000
75%	4.000000	4.000000	4.000000	5.000000	4.000000	4.000000
max	4.000000	5.000000	5.000000	5.000000	5.000000	5.000000

Рисунок 28 – Статистические данные датафрейма для классификатора

Для прогнозирования вхождения обучающегося в конкретную группу использован алгоритм KNN.

На рисунке 29 представлен программный код классификатора на основе алгоритма KNN.

```
Ввод [5]: 1 # Классификтор
2 import pandas as pd
3 train_df = pd.read_csv(p2)
4 test_df = pd.read_csv('E:\\vkr\\ContrSet.csv')
5 print(test_df)
6 from sklearn.neighbors import KNeighborsClassifier
7 predictors = ['k1','k2','k3','k4','k5']
8 outcome = 'IOT'
9 new_record = test_df.loc[0:4, predictors]
10 print('Обучающая выборка:')
11 #print(new_record)
12 x = train_df.loc[0:, predictors]
13 y = train_df.loc[0:, outcome]
14 kNN = KNeighborsClassifier(n_neighbors=1)
15 kNN.fit(x, y)
16 print(x)
17 gr=kNN.predict(new_record)
18 print('')
19 print('Обучающая выборка',kNN.predict_proba(new_record))
20 print('Рекомендуемая IOT:',gr)
21
```

Рисунок 29 – Программный код классификатора на основе алгоритма KNN

Входными данными для прогнозирования является цифровой след обучающегося, который формирует аналитик, используя результаты тестирования компетенций конкретного обучающегося.

На рисунке 30 показан пример результата прогнозирования.

	k1	k2	k3	k4	k5
0	3	4	4	5	4
Обучающая выборка:					
	k1	k2	k3	k4	k5
0	4	4	5	5	3
1	4	5	3	3	4
2	1	3	4	5	2
3	2	3	2	3	4
4	3	3	4	5	5

Обучающая выборка [[0. 0. 0. 0. 1.]]
Рекомендуемая ИОТ: [5]

Рисунок 30– Результат прогнозирования по алгоритму KNN

Таким образом, в предлагаемом примере обучающемуся на основе результатов входного тестирования рекомендована ИОТ группы 5.

4.2 Оценка эффективности проектных решений

«Для оценки экономической эффективности проектных решений используем методику сравнения затрат на разработку СФТИОО внешним программистом по договору аутсорсинга (базовый вариант) и программистом образовательной организации (проектный вариант), соответственно» [10].

«В калькуляцию себестоимости заказной разработки СФТИОО включаются следующие статьи затрат:

- зарплата исполнителя проекта по трудовому договору (ЗБ₁);
- социальные страховые взносы (ЗБ₂);
- прочие прямые расходы (ЗБ₃);

– накладные расходы ($ЗБ_4$).

В заказной доработке задействован внешний программист» [10].

Средняя стоимость часа работы программиста по договору составляет 1500 руб [11].

Ориентировочное время разработки составляет 100 час.

Итого затраты базового варианта $C_{баз}$ составят (3):

$$C_{баз} = ЗБ_1 + ЗБ_2 + ЗБ_3 + ЗБ_4 = 1500*100 + 0,271*1500*100 + 0 + 0 = 190650 \text{ руб (3)}$$

В самостоятельной разработке СФТИОО задействованы программист и аналитик образовательной организации» [10].

«В калькуляцию себестоимости собственной разработки СФТИОО включаются следующие статьи затрат:

- зарплата исполнителей проекта с учетом затраченного времени 100 час ($ЗП_1$);
- социальные страховые взносы ($ЗП_2$);
- прочие прямые расходы ($ЗП_3$);
- накладные расходы ($ЗП_4$)» [10].

Итого затраты проектного варианта $C_{пр}$ составят (4):

$$C_{пр} = ЗП_1 + ЗП_2 + ЗП_3 + ЗП_4 = (50000+40000) \text{ руб} + 0,3*(50000+40000) + 0 + 0 = 117000 \text{ руб (4)}$$

На основании полученных результатов выполнена оценка экономической эффективности проектных решений.

Сформирована таблица показателей экономической эффективности проектных решений (таблица 11).

Построен график показателей экономической эффективности проектных решений (рисунок 31).

Таблица 11 – Показатели эффективности проекта разработки СФТИОО

«Затраты»		Абсолютное изменение затрат	Коэффициент относительного снижения затрат	Индекс снижения затрат
Базовый вариант	Проектный вариант			
$C_{\text{баз}}$ (руб.)	$C_{\text{пр}}$ (руб.)	$\Delta C = C_{\text{баз}} - C_{\text{пр}}$ (руб.)	$K_C = \Delta C / C_{\text{баз}} \times 100\%$	$Y_C = C_{\text{баз}} / C_{\text{пр}}$
190650	117000	73650	39	1,6» [10]

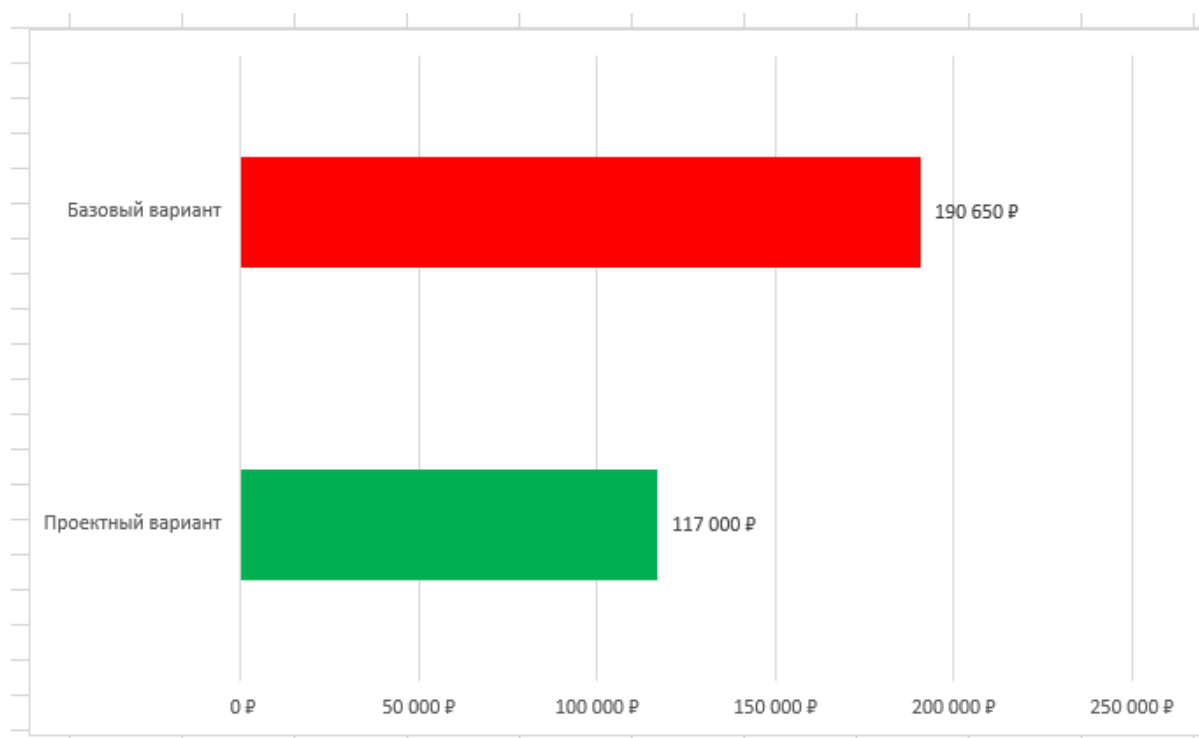


Рисунок 31 – Гистограмма сравнения затрат на разработку СФТИОО

Таким образом, затраты при проектном варианте разработки СФТИОО сократились в 1,6 раза.

«Срок окупаемости затрат на внедрение проектного решения ($T_{\text{ок}}$) определяется по формуле (5):

$$T_{\text{ок}} = K_{\text{П}} / \Delta C \text{ (мес.)}, \quad (5)$$

где K_{Π} – затраты на реализацию проектных решений (проектирование и внедрение СФТИОО).

Следовательно, срок окупаемости СФТИОО равен (6):

$$T_{\text{ок}} = 117000/73650 \approx 1,6 \text{ мес.} \quad (6)$$

Представленные расчеты подтвердили существенное снижение затрат на проектирование и эффективность проектного решения» [10].

«Для оценки эффективности управления СФИОО используем формулу расчета показателя функциональной эффективности управления (7):

$$K_{\text{эу}} = \frac{\sum_{i=1}^n P_{yi}}{n} \quad (7)$$

где n - количество функций управления, реализуемых СФИОО;

P_{yi} - вероятность выработки СФИОО эффективного управляющего воздействия при реализации i -й функции управления» [3].

«Для формирования ИОТ в СФИОО используются следующие функции:

- спрогнозировать вхождения обучающихся в определенную группу ИОТ и формирование рекомендаций для принятия решения;
- принятие решения ЛПР.

Как показывает практика, на выполнение функции «Принятие решения ЛПР» может негативно повлиять человеческий фактор.

Пусть вероятность выработки эффективного управляющего воздействия для данной функции равна 0.5.

В этом случае значение показателя функциональной эффективности управления СФИОО будет равно (8):

$$K_{\text{эу}} = 1.5/2 = 0,75 \quad (8)$$

Таким образом, коэффициент эффективности управления предлагаемой СФТИОО $K_{эу} > 0,5$, что свидетельствует о высокой функциональной эффективности процесса формирования ИОТ в образовательной организации» [3].

Выводы по главе 4

Результаты проделанной работы позволили сделать следующие выводы:

- для разработки элементов СФТИОО использованы язык Python, сред Jupyter Notebook и библиотеку машинного обучения Scikit-learn;
- обучающей выборкой для кластеризации являются данные прохождения входных тестов обучающимися, желающими получить ИОТ. Обучающая выборка представляет собой файл типа CSV, полученного из тестового XLSX-файла;
- тепловые карты помогают классифицировать разделы, которые работают не на должном уровне и нуждаются в оптимизации;
- входными данными для прогнозирования вхождения обучающихся в группу ИОТ является цифровой след обучающегося, который формирует аналитик, используя результаты тестирования компетенций конкретного обучающегося.

Расчеты подтвердили экономическую эффективность предлагаемых проектных решений и высокую функциональную эффективность управления СФТИОО.

Заключение

Процесс обучения уникален для каждого человека. Как показывает практика, важное значение для повышения успеваемости обучаемых и освоения ими новых компетенций имеют используемые методы обучения.

Существующая система онлайн обучения сосредоточена на создании инструментов для совместного обучения или облегчении проведения курсов для дискуссионного форума и прямого общения, например, в чате. Однако проблема повышения индивидуальных достижений каждого обучающегося остается актуальной.

Решение данной проблемы возможно за счет применения современных методов формирования индивидуальной траектории онлайн обучения.

Магистерская диссертация посвящена актуальной проблеме исследования методов формирования индивидуальной траектории онлайн обучения с применением ИИ и МО.

В процессе выполнения магистерской диссертации были решены следующие задачи:

- произведен анализ состояния исследований в области формирования индивидуальной траектории онлайн обучения на основе современных образовательных и информационных технологий. Как показал анализ, проблематика формирования ИОТ на основе современных образовательных и информационных технологий широко рассмотрена в работах отечественных и зарубежных ученых. В отечественных и зарубежных источниках используются также такие понятия, как «индивидуальный образовательный маршрут», «индивидуальный план обучения», «персонализированное образование» и др. Как показывают исследования, наилучшие результаты достигаются, если в процессе формирования ИОТ учитываются два основных ключевых фактора: персонализированное онлайн обучение и использование технологий

ИИ. Персонализированное обучение с использованием ИИ делает обучение более эффективным, действенным и масштабируемым, позволяя L&D-руководителям оказывать адресную поддержку каждому обучающемуся, адаптироваться к уникальным стилям обучения и более внимательно отслеживать прогресс. Анализ литературы и источников подтвердил интерес ученых и специалистов к проблеме формирования индивидуальной траектории онлайн обучения на основе ИИ и МО, что подтверждает актуальность темы исследования;

- произведен анализ методологий формирования индивидуальной траектории онлайн обучения на основе технологий искусственного интеллекта и машинного обучения. Как показал анализ, наилучшего результата можно достичь при использовании комплексного подхода при реализации концепции персонализированного обучения на основе ИИ и МО. Для формирования ИОТ с помощью ИИ и МО используются большие образовательные данные. Основой объектной модели СФИОО является класс «Модель МО», который на логическом уровне представляет модель МО, используемую для создания прогноза вхождения обучающегося в конкретную поведенческую группу и последующего формирования его индивидуальной траектории онлайн обучения;
- разработаны модели и алгоритмы системы формирования индивидуальной траектории онлайн обучения на основе технологий машинного обучения. Согласно ООП для создания логической модели системы достаточно разработать ее диаграммы вариантов использования и классов, которые отражают функциональный и структурный аспекты системы, соответственно. Алгоритмы кластеризации являются одним из основных методов сегментации данных, поскольку они могут группировать точки данных на основе их сходства или несходства, не требуя каких-либо предварительных

знаний или меток. По результатам сравнения для сегментации данных выбран алгоритм K-means как наиболее простой для реализации и понятный. Для прогнозирования вхождение конкретного обучающегося в определенную поведенческую группу по его цифровому следу используются алгоритмы классификации. По результатам сравнения для решения задачи прогнозирования выбран алгоритм KNN, как наиболее простой и гибкий;

- выполнены апробация и оценка эффективности проектных решений. для разработки элементов СФТИОО использованы язык Python, сред Jupyter Notebook и библиотеку машинного обучения Scikit-learn. Обучающей выборкой для кластеризации являются данные прохождения входных тестов обучающимися, желающими получить ИОТ. Входными данными для прогнозирования вхождения студента в группу ИОТ является цифровой след обучающегося, который формирует аналитик, используя результаты тестирования компетенций конкретного обучающегося. Расчеты подтвердили экономическую эффективность предлагаемых проектных решений и высокую функциональную эффективность управления СФТИОО.

Гипотеза исследования подтверждена.

Список используемой литературы и используемых источников

1. Алешугина Е. А., Ваганова О. И., Кутепов М. М. Проектирование индивидуальных траекторий обучения с использованием онлайн-курсов // БГЖ. 2019. №1 (26). С. 185-187.
2. Буйлов О. Стандарт электронного обучения xAPI (Tin Can API) [Электронный ресурс]. URL: <https://levellab.ru/blog/xapi-1> (дата обращения: 12.02.2024).
3. Вдовин В.М., Суркова Л.Е., Шурупов А.А. Предметно-ориентированные экономические информационные системы. М.: Дашков и К, 2016. 388 с.
4. Индивидуальная образовательная траектория. 5 шагов до создания [Электронный ресурс]. URL: <https://www.doma.uchi.ru/blog/individualnaya-obrazovatel'naya-traektoriya-5-shagov-do-sozdaniya/#chto-takoe-individualnaya-obrazovatel'naya-traektoriya-iot> (дата обращения: 12.02.2024).
5. Классификатор kNN [Электронный ресурс]. URL: <https://habr.com/ru/articles/149693/> (дата обращения: 17.02.2024).
6. Котов К., Красильников Н. Кластеризация данных [Электронный ресурс]. URL: <https://logic.pdmi.ras.ru/~yura/internet/02ia-seminar-note.pdf> (дата обращения: 17.02.2024).
7. Ланский А. Обзор методов классификации в машинном обучении с помощью Scikit-Learn [Электронный ресурс]. URL: <https://tproger.ru/translations/scikit-learn-in-python> (дата обращения 20.02.2024).
8. Леоненков А. В. Объектно-ориентированный анализ и проектирование с использованием UML и IBM Rational Rose : учебное пособие. М. : ИНТУИТ, Ай Пи Ар Медиа, 2020. 317 с. [Электронный ресурс]. URL: <https://www.iprbookshop.ru/97554.html> (дата обращения: 12.02.2024).
9. Обзор алгоритмов кластеризации данных [Электронный ресурс]. URL: <https://habr.com/ru/articles/101338/> (дата обращения 20.02.2024).
10. Поршкевич Н.Ю., Огнева А.Ю., Чинчукова Е.П. Оценка

экономической эффективности применения информационных систем // Экономика и социум. 2016. №7 (26). URL: <https://cyberleninka.ru/article/n/otsenka-ekonomicheskoy-effektivnosti-primeneniya-informatsionnyh-sistem> (дата обращения: 09.03.2024).

11. Сколько стоят услуги программистов? Цены студий и фрилансеров
Источник: <https://www.kadrof.ru/articles/46641> (дата обращения: 10.03.2024).

12. Сунгатуллина А. Т. Системный анализ и проектирование информационных систем на основе объектно-ориентированного подхода : учебно-методическое пособие по дисциплине «Методы и средства проектирования информационных систем». М. : Российский университет транспорта (МИИТ), 2020. 118 с. [Электронный ресурс]. URL: <https://www.iprbookshop.ru/115990.html> (дата обращения: 12.02.2024).

13. Хоанг К. К., Авксентьева Е. Ю., Федосов Ю. В. Формирование индивидуальной траектории обучения автоматизированному проектированию средствами интеллектуальной адаптивной системы // Эргодизайн. 2021. №1 (11). С. 41-48.

14. Что такое машинное обучение и как оно работает [Электронный ресурс]. URL: <https://trends.rbc.ru/trends/industry/60c85c599a7947f5776ad409> (дата обращения: 12.02.2024).

15. Шамсутдинова Т.М. Формирование индивидуальной образовательной траектории в адаптивных системах управления обучением // Открытое образование. 2021. № 25(6). С.36-44.

16. Andi Besse Firdausiah Mansur, Norazah Yusof, Ahmad Hoirul Basori “Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic”, Procedia Computer Science, Volume 163, 2019, Pages 125-133.

17. Anuvrat Dutta “Using AI For Personalized Learning” [Электронный ресурс]. URL: <https://www.evelynlearning.com/using-ai-for-personalized-learning/> (дата обращения: 12.02.2024).

18. Create an Individual Development Plan [Электронный ресурс]. URL:

<https://ctl.wustl.edu/writing-a-teaching-development-plan-tdp/> (дата обращения: 12.02.2024).

19. Educational Data Mining [Электронный ресурс]. URL: <https://educationaldatamining.org/> (дата обращения: 10.03.2024).

20. Hengyao Tang, Guosong Jiang, and Qingdong Wang “Personalized Learning Behavior Evaluation Method Based on Deep Neural Network” [Электронный ресурс]. URL: <https://www.hindawi.com/journals/sp/2022/9993271/> (дата обращения: 12.02.2024).

21. How Artificial Intelligence Can Be Used for Personalized Learning [Электронный ресурс]. URL: <https://www.eyecity.africa/post/how-artificial-intelligence-can-be-used-for-personalized-learning> (дата обращения: 12.02.2024).

22. How Does AI Change Our Approach To Personalized Learning? [Электронный ресурс]. URL: <https://wdhb.com/blog/how-does-ai-change-our-approach-to-personalized-learning/> (дата обращения: 12.02.2024).

23. Individual Learning Plan (ILP) Framework Purpose [Электронный ресурс]. URL: https://ride.ri.gov/sites/g/files/xkgbur806/files/Portals/0/Uploads/Documents/Students-and-Families-Great-Schools/Educational-Programming/Counseling/ILP_Framework_Revised_2017.pdf (дата обращения: 12.02.2024).

24. Juan Carlos Muñoz-Carpio, Zohaib Jan and Angelo Saavedra “Machine learning for learning personalization to enhance student academic performance”, CEUR Workshop Proceedings, 2020.

25. K-Nearest Neighbor(KNN) Algorithm - GeeksforGeeks. [Электронный ресурс]. URL: <https://www.geeksforgeeks.org/k-nearest-neighbours/> (дата обращения 20.02.2024).

26. Kulkatechol Kanokngamwitroj, Chetneti Srisa-An “Personalized Learning Management System using a Machine Learning Technique”, TEM Journal. Volume 11, Issue 4, pages 1626 -1633.

27. Lake K. How To Personalize Learning Using AI [Электронный ресурс]. URL: <https://elearningindustry.com/how-to-personalize-learning-using-ai> (дата обращения: 12.02.2024).

28. Ma Y., Wang L., Zhang J., Liu F., Jiang Q. A Personalized Learning Path Recommendation Method Incorporating Multi-Algorithm. Appl. Sci.2023, 13, 5946. <https://doi.org/10.3390/app13105946>.

29. Machine Learning : Impact on Personalized Learning And Adaptive Assessments [Электронный ресурс]. URL: <https://www.ksolves.com/blog/machine-learning/machine-learning-impact-on-personalized-learning-and-adaptive-assessments> (дата обращения: 12.02.2024).

30. Milind Bhushan, Rajat Shingate, Tanvi Shah, Naman Vyas, Shraddha Shirsath “Intelligent tutoring system: personalised learning plans with AI”, 2023, DOI: 10.13140/RG.2.2.36573.59369.

31. Personalized Learning Plans [Электронный ресурс]. URL: <https://fastercapital.com/keyword/personalized-learning-plans.html> (дата обращения: 12.02.2024).

32. Personalized Learning Using AI [Электронный ресурс]. URL: <https://belitsoft.com/custom-elearning-development/personalized-e-learning-action>

33. Project Jupyter [Электронный ресурс]. URL: <https://jupyter.org/> (дата обращения: 15.02.2024).

34. Rachel Van Dinter “How EdTechs use artificial intelligence and machine learning to create personalized learning experiences” [Электронный ресурс]. URL: <https://aws.amazon.com/ru/blogs/publicsector/how-edtechs-use-artificial-intelligence-machine-learning-create-personalized-learning-experiences/> (дата обращения: 12.02.2024).

35. Scikit-learn [Электронный ресурс]. URL: <https://blog.skillfactory.ru/glossary/scikit-learn/> (дата обращения: 15.02.2024).

36. Shetha V., Tripathia U., Sharma A. A Comparative Analysis of Machine Learning Algorithms for Classification Purpose, 4th International Conference on

Innovative Data Communication Technology and Application, *Procedia Computer Science* 215 (2022), 422–43.

37. Skovajsa Š. Review of clustering methods used in data-driven housing market segmentation, *Real estate management and valuation*, 2023. Vol. 31(3). P.67-74.

38. StudyPlannr [Электронный ресурс]. URL: (дата обращения: 15.02.2024).

39. The complete guide to heatmaps [Электронный ресурс]. URL: <https://www.hotjar.com/heatmaps/> (дата обращения: 15.02.2024).

40. Vidden C., Vriens M., Chen S. Comparing clustering methods for market segmentation: A simulation study, *Journal of Marketing Analytics*, 2016. Vol 2. P. 225-238.

41. Wang M., Lv Z. Construction of personalized learning and knowledge system of chemistry specialty via the internet of things and clustering algorithm. *J Supercomput* 78, 10997–11014 (2022). <https://doi.org/10.1007/s11227-022-04315-8> (дата обращения: 15.02.2024).