

УДК 378.016:811.1+004.8
ББК Ш12/18-9+А66

ГРНТИ 14.35.07

Код ВАК 5.8.2

Ся Янь,

доктор философии, доцент, Цзилиньский педагогический университет; Китайская Народная Республика, провинция Цзилинь, г. Сыпин; e-mail: xiauyan539@163.com

Сунь Инвэй,

магистр гуманитарных наук по обучению русскому языку, инженер, Beijing Xingyou Engineering Project Management Co Ltd; Китайская Народная Республика, г. Пекин; e-mail: sunyingweixu@cnpc.com.cn

АНАЛИЗ ПРИМЕНЕНИЯ МОДЕЛИ ПРОГНОЗИРОВАНИЯ УСПЕВАЕМОСТИ СТУДЕНТОВ ПО ИНОСТРАННОМУ ЯЗЫКУ НА ОСНОВЕ АЛГОРИТМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

КЛЮЧЕВЫЕ СЛОВА: студенты; иностранные языки; методика преподавания иностранных языков; модели прогнозирования; успешность обучения; прогнозирование успеваемости; успеваемость студентов; учебная деятельность; искусственный интеллект; информационные технологии

АННОТАЦИЯ. В данной статье представлена модель прогнозирования успеваемости студентов по иностранным языкам, разработанная с использованием алгоритмов искусственного интеллекта. Модель применяет методы глубокого и машинного обучения для достижения высокой точности прогнозов. В статье рассматриваются различия в производительности модели среди различных групп студентов и предлагаются методы улучшения, такие как оптимизация предварительной обработки данных и корректировка архитектуры модели. Подчеркивается значимость применения модели для разработки персонализированных образовательных программ и мониторинга успеваемости, при этом особое внимание уделяется вопросам конфиденциальности и безопасности данных, а также точности модели. Предлагаются стратегии для решения этих проблем. Дополнительно обсуждаются будущие направления развития, включая технологическую интеграцию, персонализацию на основе данных и междисциплинарное взаимодействие, акцентируется внимание на важности этических норм и защиты конфиденциальности. В процессе внедрения технологий искусственного интеллекта в образование необходимо учитывать вопросы конфиденциальности и безопасности данных, точности и стабильности моделей, а также их популярности и принятия. Создание надежной системы защиты данных, оптимизация работы моделей, активное информационное продвижение технологий, разработка справедливых и разумных политик и норм позволят полностью реализовать потенциал и ценность технологий искусственного интеллекта в образовательной сфере, способствуя ее высококачественному развитию.

ДЛЯ ЦИТИРОВАНИЯ: Ся, Янь. Анализ применения модели прогнозирования успеваемости студентов по иностранному языку на основе алгоритма искусственного интеллекта / Ся Янь, Сунь Инвэй. – Текст : непосредственный // Педагогическое образование в России. – 2024. – № 6. – С. 221–231.

Xia Yan,

PhD, Associate Professor, Jilin Normal University, Jilin, People's Republic of China

Sun Yingwei,

Master in Russian Education, Engineer, Beijing Xingyou Engineering Project Management Co., LTD., Beijing, People's Republic of China

ANALYSIS OF APPLICATION OF A FORECASTING MODEL OF STUDENTS' PERFORMANCE IN A FOREIGN LANGUAGE BASED ON AN ARTIFICIAL INTELLIGENCE ALGORITHM

KEYWORDS: students; foreign languages; methods of teaching foreign languages; forecasting models; learning success; forecasting academic performance; student academic performance; educational activities; artificial intelligence; information technology

ABSTRACT. This paper mainly introduces the prediction model of students' foreign language achievement based on artificial intelligence algorithm. Through deep learning and machine learning technology, the accurate prediction of students' foreign language achievement is realized. The paper also analyzes the performance difference of the model in different student groups, and puts forward the improvement direction such as optimizing data preprocessing and adjusting the model structure. This paper emphasizes the application prospect of the model in personalized teaching scheme design, learning progress tracking and feedback, and pays attention to data privacy and security, model accuracy and other challenges, and puts forward corresponding countermeasures. In addition, the paper also looks forward to the future trend of technology integration, data-driven personalization and cross-disciplinary integration, and emphasizes the importance of ethics and privacy protection. When applying artificial intelligence technology, the education industry needs to comprehensively consider data privacy and security, model accuracy and stability, technology popularization and acceptance, and educational equity. By building a sound data protection mechanism, optimizing model performance, strengthening technical publicity and promotion, and formulating fair and reasonable policies and norms and other measures, we can give full play to the potential and

value of artificial intelligence technology in the field of education, and promote the education industry to achieve high-quality development.

FOR CITATION: Xia, Yan, Sun, Yingwei. (2024). Analysis of Application of a Forecasting Model of Students' Performance in a Foreign Language Based on an Artificial Intelligence Algorithm. In *Pedagogical Education in Russia*. No. 6, pp. 221–231.

Добыча данных об образовании – одна из важных областей поиска данных, в которой прогнозирование успеваемости находится в центре внимания исследователей, а данные об успеваемости являются важным отражением поведения учащихся в процессе обучения [10]. С быстрым развитием технологии искусственного интеллекта (далее – ИИ) в современной сфере образования ее применение в образовательных инновациях становится все более широким и глубоким, особенно в прогнозировании и управлении успеваемостью студентов по иностранным языкам, что демонстрирует значительный потенциал и ценность. Интеграция технологий ИИ не только обогащает учебные методы, но и способствует значительному улучшению эффективности преподавания и персонализации образовательного процесса. Создание модели прогнозирования успеваемости по иностранным языкам на основе алгоритмов ИИ позволяет использовать анализ больших данных и технологии машинного обучения для глубокого изучения данных о студентах, включая их предыдущие академические достижения, результаты тестов, учебные привычки и другую многомерную информацию. В этом процессе исключаются субъективность и ограничения традиционных методов прогнозирования, что позволяет достичь высокой точности прогнозирования благодаря алгоритмическим моделям и масштабной поддержке данных. Такие прогнозы не только помогают преподавателям заранее выявлять потенциальные проблемы в обучении, но и обеспечивают надежную количественную основу для принятия решений в образовательной сфере, гарантируя актуальность и эффективность стратегий обучения.

Прогнозирование успеваемости позволяет отсеять учащихся с трудностями в обучении на ранней стадии. Своевременная помощь таким учащимся очень важна для ежедневного управления преподаванием и хороших результатов учащихся [9]. На основании результатов прогностической модели ИИ преподаватели могут точнее оценивать состояние и потребности каждого студента в изучении иностранного языка, а затем разрабатывать индивидуальные учебные планы и стратегии наставничества. Например, для студентов с низкими результатами можно разработать целевые курсы интенсивной подготовки для укрепления их слабых сторон, а для студентов с высокими

результатами предложить расширенные учебные ресурсы для дальнейшего развития. Такой подход не только повышает интерес и мотивацию студентов к обучению, но и способствует сбалансированному повышению общего уровня владения иностранным языком. Учащиеся, заранее зная об ожидаемых результатах своих предметов, могут повысить свое внимание к соответствующим предметам и предотвратить академический кризис [2]. Применение технологий ИИ в прогнозировании успеваемости по иностранным языкам также стимулирует школы и образовательные учреждения к оптимизации распределения образовательных ресурсов. Комплексный анализ и прогнозирование результатов учащихся позволяют учебным заведениям точно определять текущее распределение ресурсов и потенциальный спрос, а затем рационально планировать учебные программы и корректировать распределение преподавательского состава. Например, в ответ на увеличение спроса на конкретные языки, выявленное в прогнозе, учебные заведения могут предложить дополнительные курсы или привлечь иностранных преподавателей. Для групп студентов с высокими результатами можно организовать углубленные курсы или программы международного обмена для удовлетворения их образовательных потребностей на более высоком уровне. Такое распределение ресурсов на основе данных значительно повышает эффективность и результативность использования образовательных ресурсов.

Модель прогнозирования ИИ также предоставляет ценные данные для лиц, принимающих решения в сфере образования. Углубленный анализ оценок студентов и всестороннее рассмотрение результатов прогнозирования позволяют более научно обоснованно формулировать образовательную политику и планы развития. Например, при разработке политики в области образования на иностранных языках можно оценить эффект текущих образовательных стратегий на основе результатов прогнозирования и скорректировать направление и цели политики. При планировании развития образования можно предсказать будущие тенденции и изменения спроса, а также заранее подготовить образовательные ресурсы и преподавательский состав. Такой подход к принятию решений на основе данных не только повышает актуальность и эффективность образовательной политики,

но и закладывает прочный фундамент для модернизации и качественного развития образования. Применение технологии искусственного интеллекта в прогнозировании и управлении успеваемостью студентов по иностранным языкам не только повышает точность прогнозов, способствует персонализации обучения и оптимизирует распределение образовательных ресурсов, но и улучшает процесс принятия решений в сфере образования¹. С развитием технологий и расширением их применения ИИ будет играть все более важную и значимую роль в образовательной сфере².

Обоснование и техническая поддержка

1. Теоретические основы.

Добыча данных – это междисциплинарная область исследований, объединяющая результаты исследований новейших технологий, таких как технологии баз данных, искусственный интеллект, машинное обучение, статистика, инженерия знаний, объектно ориентированные методы, информационный поиск, высокопроизводительные вычисления и визуализация данных [8]. С развитием технологий больших данных в образовании Educational Data Mining (далее – EDM) предоставляет уникальные подходы и методы для построения моделей прогнозирования успеваемости студентов по иностранным языкам. EDM систематически извлекает ценную информацию из различных источников образовательных данных, таких как задания студентов, результаты тестов и записи взаимодействий в классе. Это помогает выявлять потенциальные закономерности и ключевые факторы, влияющие на успеваемость. Процесс требует не только эффективного сбора и организации данных, но и глубокого анализа для выделения характеристик, способствующих прогнозированию. Благодаря EDM модель прогнозирования строится на надежном фундаменте данных, обеспечивая точность и достоверность прогнозов.

Машинное обучение как основная технология в области ИИ обеспечивает мощные инструменты для построения моделей прогнозирования успеваемости. С использованием алгоритмов машинного обучения, таких как деревья решений, случайные леса и нейронные сети, можно анализировать

сложные взаимосвязи между успеваемостью студентов и различными переменными. Построение дерева принятия решений можно описать как рекурсивный процесс, суть которого заключается в определении критерия ветвления, т. е. в выборе оптимального атрибута разбиения из большого числа атрибутивных переменных [4]. Случайный лес использует только две случайные характеристики: выбор выборки и выбор атрибута, что значительно повышает успешность эксперимента [5]. Искусственная нейронная сеть – это система обработки информации, состоящая из абстракции, упрощения и моделирования структуры и функции нейронной сети человеческого мозга [13]. Эти алгоритмы выявляют скрытые закономерности в данных и сохраняют высокую вычислительную мощность и способность к обобщению даже при работе с крупными и многомерными данными. Машинное обучение позволяет обеспечить точные прогнозы успеваемости и предоставить преподавателям значимую поддержку в принятии решений.

Теория когнитивной психологии играет важную роль в построении моделей прогнозирования успеваемости. Исследования в этой области помогают глубже понять процесс изучения иностранных языков и выявить психологические факторы, влияющие на обучение, такие как мотивация, стратегии обучения и память. Эти факторы оказывают решающее влияние на точность прогнозирования. Интеграция когнитивной психологии в модели прогнозирования позволяет учитывать особенности и индивидуальные различия студентов, делая модели более реалистичными и практичными. Это также дает преподавателям психологические перспективы для лучшего понимания потребностей студентов и разработки целенаправленных стратегий обучения.

В целом интеграция EDM, машинного обучения и когнитивной психологии обеспечивает комплексный подход к построению моделей прогнозирования успеваемости по иностранным языкам. Использование соответствующих моделей для прогнозирования успеваемости учащихся не только стимулирует их интерес к учебе, но и позволяет снизить нагрузку на педагогов и выработать более целенаправленные стратегии обучения [1]. Эти дисциплины дополняют друг друга, способствуя интеллектуальному развитию образовательной сферы. В будущем, с продолжением технологического прогресса и углублением теоретических исследований, результаты в этой области будут способствовать повышению качества образования и оптимизации учебного процесса.

¹ Высококачественные данные способствуют прогрессу в системах искусственного интеллекта // Новостях ZAKER. 17 июля 2024 г. URL: <http://www.myzaker.com/article/669741d78e9f091c693db088> (дата обращения: 20.12.2024).

² Искусственный интеллект может изменить парадигму образования: стандартизированное преподавание становится персонализированным обучением // Китайская новостная сеть. 29 декабря 2023 г. URL: <https://www.chinanews.com/m/sh/2023/12-29/10137650.shtml> (дата обращения: 20.12.2024).

2. Техническая поддержка.

Прогнозирование успеваемости студентов на основе их поведения в режиме онлайн может помочь педагогу в динамическом понимании процесса обучения и разработке дифференцированных стратегий обучения. Однако в рамках гибридных курсов сложно прогнозировать успеваемость учащихся только на основе данных, полученных в режиме онлайн, что все еще находится в стадии изучения [6]. При прогнозировании успеваемости студентов по иностранному языку ключевыми этапами являются сбор и обработка данных, применение алгоритмов машинного обучения, выбор и извлечение признаков, оценка и оптимизация моделей, а также визуализация результатов.

Первая задача при прогнозировании успеваемости студентов по иностранному языку заключается в создании полной и точной базы данных. Для этого необходимо собрать данные по множеству параметров, включая, в частности, базовую информацию о студентах (возраст, пол, уровень образования), данные об учебном поведении (продолжительность обучения, частота общения в Интернете, предпочтение учебных ресурсов), результаты предыдущих тестов и оценки преподавателей. Методы сбора данных применяются для автоматического извлечения информации с различных платформ онлайн-обучения, систем управления школами и других источников. Очистка данных направлена на выявление и исправление ошибок, пропусков и выбросов, что обеспечивает их точность и согласованность. Далее данные стандартизируются с использованием методов преобразования данных для упрощения последующего анализа и обработки. Наконец, информация сохраняется в эффективной и масштабируемой базе данных, что гарантирует надежную поддержку для построения прогностических моделей. Особое внимание уделяется защите конфиденциальности данных, что предполагает соблюдение всех соответствующих законов и норм.

Алгоритмы машинного обучения являются основным инструментом для построения модели прогнозирования успеваемости студентов по иностранному языку. Ключевым моментом в улучшении возможностей машинного обучения является предоставление машинам возможности наследовать и развивать знания, полученные в прошлом, т. е. способность к инкрементальному обучению [7]. Используется комбинация алгоритмов линейной регрессии, деревьев решений, случайного леса и нейронных сетей для достижения высокой точности прогнозирования. Алгоритм линейной регрессии

позволяет выявить линейные зависимости между переменными и подходит для предварительного анализа взаимосвязей между факторами обучения и успеваемостью. Алгоритмы деревьев решений и случайного леса демонстрируют свои преимущества при работе со сложными нелинейными взаимосвязями благодаря интуитивной интерпретации и способности к обобщению. Нейронные сети, особенно методы глубокого обучения, обладают мощными возможностями для извлечения признаков и распознавания образов, что позволяет автоматически выявлять глубокие и скрытые особенности, влияющие на успеваемость, и, таким образом, повышать точность прогнозирования.

При построении модели прогнозирования выбор и извлечение признаков имеют ключевое значение. Чрезмерное количество признаков не только увеличивает вычислительную сложность, но и может внести шум, влияя на эффективность прогнозирования. В связи с этим применяются передовые методы, такие как анализ главных компонент (PCA) и факторный анализ (FA), для уменьшения размерности данных и выделения признаков, оказывающих значительное влияние на результаты прогнозирования. Этот процесс не только снижает размерность данных и улучшает эффективность обучения модели, но и повышает ее интерпретируемость, делая результаты прогнозирования более релевантными образовательной реальностью.

Для обеспечения точности и надежности модели прогнозирования мы используем такие строгие методы оценки, как перекрестная валидация и метод «оставления без внимания», чтобы проверить модель с разных точек зрения и на разных уровнях. Перекрестная валидация позволяет эффективно избежать переобучения путем разделения набора данных на обучающий и проверочный и многократного повторения процесса обучения и проверки для оценки эффективности модели на различных подмножествах данных. Метод «оставления без внимания» служит в крайних случаях для повышения строгости оценки. На основе результатов оценки проводятся постоянная корректировка параметров модели и оптимизация структуры алгоритма для достижения наилучшей эффективности прогнозирования. Одновременно обеспечивается высокая способность модели к обобщению для поддержания стабильной точности прогнозирования даже на неизвестных данных.

Технологии визуализации играют важную роль в интуитивном представлении и глубоком анализе результатов прогнозирования. Применяются современные инструменты визуализации для представления

сложных моделей и данных в виде графиков, тепловых карт, линий тренда и других визуальных форматов. Это позволяет преподавателям и студентам интуитивно понять результаты прогнозирования и лежащую в их основе логику. Визуализация помогает выявить потенциальные проблемы и закономерности в обучении, а также предоставляет научную основу для корректировки стратегий преподавания и обучения. С помощью визуального анализа можно более точно определить слабые стороны учащегося, разработать индивидуальные планы обучения и повысить эффективность и качество образовательного процесса.

Построение модели для прогнозирования успеваемости студентов по иностранному языку

1. Источники данных и предварительная обработка.

Для всесторонней оценки успеваемости студентов мы используем разнообразные источники данных. Основным источником является университетская система учета успеваемости, где фиксируются оценки студентов по различным видам учебной деятельности в течение учебного года. Это включает обычные оценки, оценки за выпускные экзамены, передачи экзаменов и другие формы оценки. Эти данные непосредственно отражают результаты обучения и академические способности студентов. Кроме того, мы собрали значительное количество данных о поведении студентов на персональных учебных платформах, таких как системы онлайн-обучения и системы управления обучением. Эти данные включают продолжительность учебных занятий, частоту доступа к образовательным ресурсам, степень участия в интерактивных дискуссиях и другие аспекты, которые предоставляют важную информацию для анализа учебных привычек, отношения и предпочтений студентов. Для получения более глубокой обратной связи мы также разработали и внедрили анкету, охватывающую различные аспекты, такие как методы обучения, самооценка и учебное давление. Эта анкета позволила собирать субъективные ощущения и потребности студентов. Обобщив данные из этих источников, мы создали многомерный и многоуровневый набор данных об успеваемости студентов.

Очистка данных является критически важным этапом предварительной обработки, от которого напрямую зависят точность и эффективность последующих анализов. Очистка данных означает удаление или идентификацию изолированных точечных данных на предмет аномальных или отсутствующих данных [14]. Мы применяем стратегию многоступенчатой очистки дан-

ных, в рамках которой сначала выявляются и удаляются дубликаты с помощью уникальных идентификаторов для обеспечения чистоты набора данных. Для обработки недостающих значений мы используем различные подходы, включая замену средними значениями, множественную импутацию, интерполяцию или удаление записей, в зависимости от характеристик данных и модели пропусков. Эти методы направлены на снижение погрешностей, вызванных отсутствующими значениями, при сохранении целостности данных. Выявление и коррекция выбросов также являются важными аспектами очистки данных. Мы применяем статистические методы (например, Z-оценку, интерквартильный размах и другие) в сочетании с бизнес-контекстом для выявления и обработки выбросов, что позволяет обеспечить соответствие данных реальному положению дел и избежать недостоверных результатов анализа.

Из-за разнообразия источников данных интеграция данных представляет собой сложную задачу. Мы разработали детализированный план интеграции данных, который включает унификацию форматов и стандартов кодирования для обеспечения совместимости и сопоставимости данных из различных источников. Затем данные из академической системы, учебных платформ и анкет интегрированы с использованием технологий хранения данных, таких как хранилище данных или озеро данных, для создания единого хранилища данных об успеваемости студентов. В процессе интеграции мы уделяем особое внимание проверке целостности и непротиворечивости данных, а также обеспечению качества интегрированного набора данных, устанавливая правила сопоставления ключевых полей и логику проверки данных.

Преобразование данных является ключевым этапом для повышения эффективности анализа. Чтобы исключить влияние различных масштабов и величин на обучение моделей, мы проводим стандартизацию и нормализацию данных. Нормализация данных, например Z-оценка, приводит к тому, что обработанные данные имеют среднее значение 0 и стандартное отклонение 1, что подходит для алгоритмов, требующих расчета расстояния или анализа ковариации. Нормализация в диапазоне $[0,1]$ масштабирует данные для ускорения сходимости в специфических алгоритмах, таких как нейронные сети. В соответствии с требованиями анализа мы также выполняем отбор признаков и снижение размерности данных для уменьшения избыточности и повышения эффективности обучения модели. Эти преобразования данных создают

прочную основу для последующего анализа академической успеваемости.

2. Извлечение и выбор признаков.

В процессе построения модели прогнозирования успеваемости студентов по иностранному языку ключевыми аспектами являются выбор и оценка признаков, так как они напрямую влияют на точность и достоверность прогноза модели. Извлечение признаков является основой построения модели, направленной на выделение высоко репрезентативных показателей способностей студентов из разнообразных учебных данных. Эти признаки охватывают не только базовый словарный запас, скорость чтения и восприятие речи на слух, но и косвенные факторы, такие как учебные привычки студентов, стратегии обучения и даже эмоциональное состояние. Благодаря комплексному использованию обработки естественного языка, анализа данных и других передовых технологий мы можем точно фиксировать успеваемость студентов в различных сценариях обучения и формировать многомерный и детализированный набор характеристик.

Отбор признаков, с другой стороны, представляет собой процесс выбора и оптимизации извлеченных признаков с целью удаления избыточной информации и сохранения тех, которые вносят наибольший вклад в результаты прогнозирования. Этот процесс может быть реализован параллельно с использованием различных методов, например статистических методов, таких как тест хи-квадрат или взаимная информация, для быстрого выявления признаков, которые сильно коррелируют с успеваемостью по иностранному языку. В то же время с помощью моделей машинного обучения, таких как деревья решений и случайные леса, набор признаков уточняется путем оценки их важности. Совместное использование этих методов не только улучшает научный подход к выбору признаков, но и обеспечивает эффективность и результативность последующего построения модели.

Оценка признаков является необходимым шагом для проверки обоснованности и значимости выбранных признаков. С помощью перекрестной валидации и экспериментов по удалению признаков мы можем интуитивно оценить вклад каждого признака в прогностическую способность модели, чтобы скорректировать и оптимизировать набор признаков. В сочетании с практическими сценариями обучения разумная интерпретация и применение признаков также являются ключом к повышению интерпретируемости и прикладной ценности модели. В процессе оценки признаков мы также обращаем внимание на

взаимодействие и потенциальные корреляции между признаками, чтобы обнаружить новые паттерны и тенденции в обучении и обеспечить более точное руководство и поддержку в изучении иностранных языков.

Благодаря научно обоснованному процессу извлечения, отбора и оценки признаков нам удалось построить эффективную и точную модель прогнозирования успеваемости студентов по иностранному языку, которая придала новый импульс развитию образования в этой области. Этот исследовательский опыт предоставляет ценные идеи и рекомендации для дальнейших исследований.

3. Разработка и внедрение модели прогнозирования.

С быстрым развитием технологий проникновение искусственного интеллекта, особенно больших данных и технологий машинного обучения, в сферу образования становится все более глубоким, открывая новые пути для повышения качества преподавания и оптимального распределения образовательных ресурсов. Прогностические модели как один из основных инструментов обеспечивают мощную информационную поддержку для персонализированного преподавания, учебных мероприятий и раннего предупреждения академических проблем путем точного анализа многомерной информации, такой как поведение студентов в процессе обучения и данные об успеваемости.

При построении моделей прогнозирования в сфере образования первой задачей является выбор подходящей архитектуры модели, исходя из характеристик образовательных данных и целей прогнозирования. Линейные регрессионные модели подходят для анализа линейной связи между факторами и успеваемостью, в то время как деревья решений и машины опорных векторов лучше справляются с многомерными нелинейными задачами. В условиях сложных образовательных сценариев методы интегрированного обучения, такие как случайный лес (Random Forest) и градиентный бустинг (Gradient Boosting), демонстрируют отличные результаты, значительно повышая точность и стабильность прогноза за счет объединения нескольких слабых моделей. Также необходимо учитывать интерпретируемость модели, чтобы преподаватели и администраторы образовательных учреждений могли интуитивно понимать логику, лежащую в основе прогнозов модели, и принимать научные и рациональные решения.

Обучение модели – ключевое звено, обеспечивающее качество прогнозирования. На этом этапе набор данных должен быть строго разделен на обучающий, про-

верочный и тестовый наборы, чтобы модель могла усвоить универсальные закономерности и эффективно избежать переобучения. Параметры модели настраиваются в ходе нескольких итераций, а для постоянной оптимизации используются перекрестная валидация и другие техники. Кроме того, важную роль играет разработка признаков, включающая такие этапы, как предварительная обработка данных, выбор признаков и их преобразование, направленные на улучшение качества данных и повышение способности модели улавливать ключевую информацию. Реализация этого усовершенствованного процесса обучения закладывает прочный фундамент для надежности модели в практических приложениях.

Для всесторонней оценки эффективности модели необходимо использовать различные метрики, такие как средняя квадратичная ошибка (MSE), среднеквадратичная ошибка (RMSE) и средняя абсолютная ошибка (MAE), чтобы количественно определить отклонение между предсказанными и фактическими значениями. Для задач классификации такие показатели, как матрица путаницы и ROC-кривая, могут эффективно отражать эффективность модели. На основе результатов оценки можно скорректировать параметры модели, оптимизировать стратегию выбора признаков или ввести новые признаки для повышения точности прогнозирования и обобщающей способности модели. Регулярная переоценка эффективности модели для поддержания ее оптимального состояния в меняющихся образовательных условиях также является важным аспектом.

В образовательной практике сценарии применения прогностического моделирования обширны и далеко идущи. В персонализированном обучении модели могут рекомендовать индивидуальные учебные маршруты и ресурсы, основанные на уровне способностей и стилях обучения студентов, чтобы повысить эффективность обучения. В области вмешательства в процесс обучения модели могут своевременно выявлять трудности студентов, инициируя целевое консультирование и меры вмешательства. Системы раннего оповещения об успеваемости используют модели для прогнозирования динамики успеваемости студентов, предоставляя администраторам информацию для принятия превентивных мер. Использование методов интеллектуального анализа данных в управлении успеваемостью учащихся для извлечения полезных знаний и правил для постоянного совершенствования управления обучением – это тенденция в развитии учебного менеджмента в учебных заведениях [12]. Более то-

го, модели прогнозирования помогают оптимизировать распределение образовательных ресурсов, содействовать равенству в образовании и обеспечивать научную основу для разработки политики путем точного анализа образовательных потребностей различных регионов, школ и отдельных учащихся.

Прогнозирование последующих оценок через успеваемость учащихся на ранних этапах обучения способствует повышению качества обучения и воспитания [3]. Глубокое применение прогностических моделей в сфере образования не только повышает качество преподавания и эффективность обучения, но и способствует развитию образования в направлении большей интеллектуальности и персонализации. С постоянным прогрессом технологий и углублением их применения можно с уверенностью утверждать, что в будущем прогностические модели будут играть еще более важную роль, способствуя созданию высококачественной образовательной системы.

Анализ перспектив и проблем применения

1. Идеи по расширению сценариев применения.

Использование технологий ИИ в обучении иностранным языкам имеет огромный потенциал. На основе алгоритмов ИИ модели прогнозирования успеваемости учащихся предоставляют беспрецедентную персонализированную и точную поддержку обучения. Эти модели могут значительно повысить эффективность преподавания, оптимизировать распределение ресурсов и облегчить сотрудничество между школой и домом.

Прогнозирование результатов тестирования учащихся и даже основных экзаменов – это в основном суждение педагогов, основанное на опыте [15]. Реализация персонализированных решений для обучения: модели ИИ способны точно определить стиль обучения, уровень способностей и предпочтения каждого студента через глубокий анализ данных об их обучении. На основе этих данных модель может разрабатывать индивидуальные рекомендации и планы репетиторства, полностью соответствующие потребностям учащихся. Такой подход значительно повышает мотивацию и эффективность обучения, помогая студентам преодолеть трудности и улучшить результаты. Например, для студентов со слабыми навыками понимания прочитанного модель может рекомендовать более целенаправленные материалы и упражнения для чтения, а также давать советы по развитию этих навыков. Для студентов с недостаточными устными навыками модель может сосредоточиться на фонологической подго-

товке и практике устного диалога¹.

Модель ИИ может отслеживать прогресс учащихся в режиме реального времени, выявляя слабые и узкие места в их обучении. Это позволяет учителям быстро корректировать стратегии преподавания и учебные планы, чтобы поддерживать актуальность и эффективность обучения. Постоянный мониторинг учебных траекторий учащихся позволяет своевременно вмешиваться в процесс обучения, предотвращая накопление проблем и обеспечивая постоянный прогресс студентов. Такой подход, основанный на данных, также способствует научной оценке преподавания, помогая учителям более полно понимать эффективность своих методов и оптимизировать учебные программы.

В условиях ограниченных образовательных ресурсов модель ИИ может прогнозировать, какие учащиеся могут нуждаться в дополнительной поддержке в будущем, основываясь на прогнозах их успеваемости. Это позволяет школам и учебным заведениям точнее распределять ресурсы, такие как учителя и учебные помещения, обеспечивая их максимальное использование. Такое распределение не только повышает справедливость образования, но и способствует общему улучшению его качества.

Модель прогнозирования успеваемости на основе ИИ строит новый мост для сотрудничества между школой и домом. Родители могут видеть успеваемость, прогресс и потенциальные проблемы своих детей через результаты прогнозов модели, что позволяет им более точно и эффективно общаться с учителями. Такая коммуникация на основе данных укрепляет доверие и взаимопонимание между школой и домом, позволяя обеим сторонам совместно разрабатывать целевые учебные планы и способствуя всестороннему развитию ребенка. В то же время родители могут лучше понимать процесс обучения и активно участвовать в образовании своих детей, создавая хорошую ситуацию совместного обучения «дом-школа».

Модель прогнозирования успеваемости учащихся по иностранному языку, основанная на алгоритме искусственного интеллекта, привнесла новую жизнь в обучение иностранным языкам благодаря своим персонализированным, точным и эффективным характеристикам. Она не только повышает эффективность обучения и оптимизирует распределение ресурсов, но и углубляет сотрудничество между домом и школой и за-

кладывает прочный фундамент для выращивания высококачественных талантов в области иностранных языков, обладающих международной конкурентоспособностью².

2. Проблемы и стратегии их решения.

С быстрым развитием технологий ИИ возникают определенные проблемы, которые необходимо решать для успешного внедрения ИИ в образование. Рассмотрим основные из них и стратегии их решения в четырех ключевых аспектах: конфиденциальность и безопасность данных, точность и стабильность моделей, популярность и принятие технологии, а также справедливость образования.

Внедрение искусственного интеллекта в образовательную сферу создает новые возможности для персонализированного обучения, однако сопряжено с рядом проблем, требующих тщательного рассмотрения и решения. Прежде всего, обеспечение конфиденциальности и безопасности данных студентов является критически важной задачей. Сбор, обработка и анализ данных о студентах являются основой для повышения качества образования, но этот процесс связан с риском утечки и неправомерного использования данных. Для защиты данных образовательные учреждения должны разработать и внедрить всесторонний механизм защиты. Важными мерами являются усиленное шифрование данных, ограничение доступа к ним, а также обеспечение безопасности данных в процессе их передачи. Необходимо также разработать прозрачную политику использования данных, которая четко информирует студентов и родителей о способах сбора, использования и защиты данных. Это повысит уровень доверия к образовательным учреждениям и технологии ИИ. Регулярные аудиты и оценка систем защиты данных помогут своевременно обнаруживать и устранять потенциальные уязвимости в системе безопасности.

Кроме того, точность и стабильность ИИ-моделей имеют критическое значение для их успешного применения в образовании. Эффективность моделей, таких как те, которые предоставляют индивидуальные рекомендации по обучению и прогнозируют академическую успеваемость, зависит от их способности точно предсказывать и адаптироваться к изменениям в образовательном процессе. Для повышения точности моделей образовательные учреждения должны постоянно оптимизировать алгоритмы, вводя большее количество характеристик данных студентов и улучшая пред-

¹ Как искусственный интеллект может расширить возможности дельты реки Янцзы, на которую приходится одна треть промышленных масштабов страны? // Байцзяхао. 09 июля 2024 г. URL: <https://baijiahao.baidu.com/s?id=1804086401471275088&wfr=spider&for=pc> (дата обращения: 20.12.2024).

² Рекорды образования каждый день! // Китайская сеть образовательных новостей. 21 июля 2024 г. URL: http://paper.jyb.cn/zgjyb/html/2024-07/21/content_642205.htm (дата обращения: 20.12.2024).

сказательную способность моделей. Использование передовых технологий машинного обучения, таких как глубокое обучение, может значительно повысить сложность и адаптивность моделей. Создание системы непрерывного обучения и оптимизации модели, а также регулярная корректировка параметров на основе обратной связи и данных в реальном времени являются ключевыми для обеспечения долгосрочной эффективности.

Хотя технология искусственного интеллекта демонстрирует огромный потенциал в сфере образования, ее популярность и признание все еще связаны с определенными ограничениями. Чтобы ускорить применение технологии ИИ в сфере образования, необходимо усилить рекламу и продвижение технологии, а также повысить осведомленность и доверие учителей, родителей и учеников к технологии ИИ. Организуя обучающие семинары и разрабатывая учебные кейсы, мы можем помочь учителям освоить использование учебных инструментов ИИ и улучшить их преподавательские способности. В то же время мы создаем удобную технологическую платформу и дизайн интерфейса, чтобы снизить порог использования технологии, чтобы больше людей могли легко воспользоваться изменениями в образовании, вызванными технологией ИИ.

Применение алгоритмов ИИ в образовании не должны усугублять существующие образовательные неравенства, а должны служить мощным инструментом для обеспечения равенства в обучении. Для этого необходимо разработать справедливую и разумную образовательную политику и технические нормы применения, чтобы обеспечить справедливость и прозрачность разработки алгоритмов. В то же время следует уделять внимание образовательным потребностям учащихся из неблагополучных групп и использовать технологии ИИ для оказания адресной поддержки в обучении и сокращения разрыва в образовании. Усиление регулирования и проведение регулярных оценок образовательных алгоритмов помогут выявлять и устранять предвзятости и дискриминацию, что является необходимым шагом для обеспечения справедливости в образовательной сфере.

Перспективы прогнозирования будущего алгоритмов искусственного интеллекта. Успеваемость учащихся по программе – один из важнейших показателей для оценки качества обучения в учебном заведении [11]. В условиях стремительного развития технологий искусственного интеллекта модели прогнозирования успеваемости студентов по иностранным языкам переживают важный этап эволюции.

Основные тенденции этого процесса включают интеграцию различных технологий, управление данными, междисциплинарное взаимодействие, а также повышение уровня этики и защиты конфиденциальности.

1. Интеграция технологий и инновации.

Будущее моделей прогнозирования успеваемости студентов будет характеризоваться интеграцией передовых технологий, что позволит создать более интеллектуальные и адаптивные системы. Одной из ключевых областей является развитие технологий обработки естественного языка, что обеспечит моделям возможность более точного анализа языкового материала и выявления тонких нюансов и потенциальных трудностей в изучении языка. Параллельно интеграция технологий компьютерного зрения позволит моделям анализировать невербальные сигналы, такие как выражения лиц и движения студентов в видеозаписях, для оценки их состояния и эмоционального фона в процессе обучения. Ожидается, что использование методов обучения с подкреплением обеспечит динамическую корректировку стратегий обучения и моделей прогнозирования на основе обратной связи от студентов, что повысит точность и адаптивность таких систем. Эта тенденция интеграции нескольких технологий значительно повысит точность и адаптивность моделей прогнозирования.

2. Управление данными и персонализация.

Важнейшим аспектом будущего развития моделей прогнозирования являются управление данными и персонализация. Качество и масштаб данных оказывают прямое влияние на точность прогнозов. В будущем модели будут уделять особое внимание сбору и анализу высококачественных многомерных данных, включая поведенческие характеристики студентов (например, время, затраченное на учебу, частота занятий), интересы и предпочтения (собранные через краудсорсинговые платформы, такие как Open Assistant и Chatbot Arena), а также когнитивные способности (например, понимание языка, память и др.). Такой углубленный анализ позволит моделям создавать индивидуальные профили студентов, обеспечивая более точные прогнозы успеваемости и рекомендации по обучению. Персонализированный подход на основе данных значительно повысит эффективность обучения и удовлетворение студентов от учебного процесса.

3. Междисциплинарное пересечение и интеграция.

Исследование модели прогнозирования успеваемости по иностранному языку уже не изолировано от области информатики, а

тесно пересекается и интегрируется с психологией, педагогикой и другими дисциплинами. Психологические теории могут помочь модели понять влияние мотивации, эмоционального состояния и других психологических факторов студентов на эффективность обучения; педагогические теории предоставляют научные стратегии обучения и методы оценки для модели, способствуя органичному сочетанию обучения и прогнозирования¹. Эта междисциплинарная исследовательская парадигма не только обогащает прогностическое измерение модели, но и способствует интеллектуальным изменениям в сфере образования и достижению более комплексной и научной оценки и руководства в сфере образования.

4. Этика и защита частной жизни.

С увеличением применения ИИ-моделей возрастает внимание к вопросам

этики и защиты конфиденциальности. В будущем важным аспектом станет повышение прозрачности и интерпретируемости моделей для обеспечения справедливости и точности прогнозов, а также для минимизации предвзятости алгоритмов. Усиленные исследования и применение технологий безопасности данных, таких как шифрование и анонимизация, станут необходимыми мерами для защиты личной информации студентов². Создание эффективного механизма этического контроля и строгого соблюдение этических стандартов в процессе разработки и применения моделей ИИ будут способствовать их здоровому и ответственному развитию. Реализация этих мер обеспечит прочную этическую и правовую основу для устойчивого применения моделей прогнозирования успеваемости студентов по иностранным языкам.

¹ Краткий обзор хода работы аспиранта Принстона Тяню Гао над проектом «Тонкая настройка инструкции: данные, алгоритмы и оценка». 22 января 2024 г. URL: http://www.360doc.com/content/24/0122/15/47115229_1111906779.shtml (дата обращения: 20.12.2024).

² Укрепление этического управления искусственным интеллектом требует как «жесткого», так и «мягкого» подхода. 24 июля 2024 г. URL: <https://baijiahao.baidu.com/s?id=1805242468079098214&wfr=spider&for=pc> (дата обращения: 20.12.2024).

ЛИТЕРАТУРА

1. Ван, Ди. Применение взвешенного случая скрытого наивного байесовского алгоритма в прогнозировании успеваемости учащихся / Ван Ди, Юй Лянцзюнь. – Текст : непосредственный // Учебный журнал Хубэйского второго педагогического университета. – 2023. – № 8. – С. 101–108.
2. Ван, Чжэндуо. Исследование прогнозирования успеваемости учащихся в учебных заведениях на основе LightGBM и алгоритма Stacking Fusion / Ван Чжэндуо, Ма Шилай. – Текст : непосредственный // Сборник научных трудов 27-й ежегодной конференции по новым сетевым технологиям и приложениям / Китайская ассоциация пользователей компьютеров – отделение сетевых приложений. – 2023. – № 11. – С. 90–94.
3. Ву, Синхуэй. Исследование по оценке и прогнозированию успеваемости учащихся с помощью алгоритма случайного леса / Ву Синхуэй, Чжоу Юйпин, Син Хайхуа. – Текст : непосредственный // Компьютерные знания и технологии. – 2020. – № 2. – С. 254–258.
4. Ву, Тун. Применение алгоритма дерева принятия решений в предиктивном анализе успеваемости учащихся / Ву Тун, Ван Сюйкун. – Текст : непосредственный // Микрокомпьютерная информация. – 2010. – № 1. – С. 209–211.
5. Линь, Ифэн. Применение алгоритма дерева принятия решений для прогнозирования успеваемости учащихся : выпускная магистерская работа / Линь Ифэн. – Нинбонский университет, 2018. – Текст : непосредственный.
6. Луо, Янъян. Исследование модели прогнозирования успеваемости студентов для гибридных учебных программ, основанной на алгоритме инкрементального обучения / Луо Янъян, Хань Сибинь. – Текст : непосредственный // Исследование в области электрохимического образования. – 2021. – № 7. – С. 83–90.
7. Се, Синьюй. Прогнозирование рейтинга учащихся на основе улучшенного алгоритма TrAdaboost / Се Синьюй, Чжан Инлу. – Текст : непосредственный // Вычислительная техника и модернизация. – 2016. – № 2. – С. 122–126.
8. Сюань, Гуоцин. Прогнозирование успеваемости учащихся на основе нейросетевого алгоритма кросс-покрытия : выпускная магистерская работа / Сюань Гуоцин. – Аньхойский университет, 2011. – Текст : непосредственный.
9. Фан, Тао. Исследование модели прогнозирования успеваемости студентов на основе алгоритма интегрированного обучения / Фан Тао. – Текст : непосредственный // Гуйлиньский университет электронной науки и техники. – 2023. – № 2.
10. Цянь, Тао. Реализация прогнозирования успеваемости учащихся на основе алгоритма случайного леса / Цянь Тао. – Текст : непосредственный // Новости науки и техники. – 2024. – № 8. – С. 37–39.
11. Чжуан, Цяохуэй. Исследование и применение алгоритма IRFC для прогнозирования оценок учащихся по предмету / Чжуан Цяохуэй. – Текст : непосредственный // Информационная технология и информатизация. – 2020. – № 6. – С. 192–194.
12. Чэн, Сюэ. Применение улучшенного алгоритма дерева принятия решений при анализе оценок старшеклассников / Чэн Сюэ. – Текст : непосредственный // Информация и вычислительная техника. – 2013. – № 10. – С. 139–140.
13. Юань, Цзэнжэнь. Нейронные сети и их применение / Юань Цзэнжэнь. – Пекин : Издательство Университета Цинхуа, 1999. – Текст : непосредственный.

14. Юэ, Цзинпэн. Исследование и применение алгоритма прогнозирования успеваемости учащихся на основе алгоритма дерева принятия решений C5.0 : выпускная магистерская работа / Юэ Цзинпэн. – Цзилиньский университет, 2020. – Текст : непосредственный.
15. Mei, N. Energy efficiency prediction of screw chillers on BP neural network optimized by improved genetic algorithm / N. Mei, F. Qian, L. Yan et al. – Text : immediate // International Computers, Signals and Systems Conference. – Dalian, 2018. – P. 527–531.

REFERENCES

1. Wang, Di, Yu, Liangjun. (2023). Primenenie vzveshennogo sluchaya skrytogo naivnogo baiesovskogo algoritma v prognozirovani uspevaemosti uchashchikhsya [Application of Case-weighted Implicit Naive Bayes Algorithm in Student Achievement Prediction]. In *Uchebnyi zhurnal Khubeiskogo vtorogo pedagogicheskogo universiteta*. No. 8, pp. 101–108.
2. Wang, Zhentuo, Ma, Shilai. (2023). Issledovanie prognozirovaniya uspevaemosti uchashchikhsya v uchebnykh zavedeniyakh na osnove LightGBM i algoritma Stacking Fusion [Research on Student Achievement Prediction Based on the Fusion Algorithm LightGBM and Stacking]. In *Sbornik nauchnykh trudov 27-i ezhegodnoi konferentsii po novym setevym tekhnologiyam i prilozheniyam*. No. 11, pp. 90–94.
3. Wu, Xinghui, Zhou, Yuping, Xing, Haihua. (2020). Issledovanie po otsenke i prognozirovaniyu uspevaemosti uchashchikhsya s pomoshch'yu algoritma sluchainogo lesa [Research on Evaluation and Prediction of Student Achievement using Random Forest Algorithm]. In *Komp'yuternye znaniya i tekhnologii*. No. 2, pp. 254–258.
4. Wu, Tong, Wang, Xiukun. (2010). Primenenie algoritma dereva prinyatiya reshenii v prediktivnom analize uspevaemosti uchashchikhsya [Application of Decision tree Algorithm in Student Achievement Prediction Analysis]. In *Mikrokompyuternaya informatsiya*. No. 1, pp. 209–211.
5. Lin, Yifeng. (2018). Primenenie algoritma dereva prinyatiya reshenii dlya prognozirovaniya uspevaemosti uchashchikhsya [Application of Decision Tree Algorithm in Student Achievement Prediction]. *Ninbonskii universitet*.
6. Luo, Yangyang, Han, Xibin. (2021). Issledovanie modeli prognozirovaniya uspevaemosti studentov dlya gibridnykh uchebnykh programm, osnovannoi na algoritme inkremental'nogo obucheniya [Research on the Prediction Model of Mixed Curriculum Students' Achievement Based on Incremental Learning Algorithm]. In *Issledovanie v oblasti elektrokhimicheskogo obrazovaniya*. No. 7, pp. 83–90.
7. Xie, Xingyu, Zhang, Yinglu. (2016). Prognozirovanie reitinga uchashchikhsya na osnove uluchshennogo algoritma TrAdaboost [Prediction of Student Performance Ranking Based on Improved TrAdaboost Algorithm]. In *Vychislitel'naya tekhnika i modernizatsiya*. No. 2, pp. 122–126.
8. Xuan Guoqing. (2011). Prognozirovanie uspevaemosti uchashchikhsya na osnove neirosetevogo algoritma kross-pokrytiya [Student Achievement Prediction based on Neural Network Cross Coverage Algorithm]. *An'khoiskii universitet*.
9. Fang, Tao. (2023). Issledovanie modeli prognozirovaniya uspevaemosti studentov na osnove algoritma integrirovannogo obucheniya [Research on Student Achievement Prediction Model based on Ensemble Learning Algorithm]. In *Guilin'skii universitet elek-tronnoi nauki i tekhniki*. No. 2.
10. Qian, Tao. (2024). Realizatsiya prognozirovaniya uspevaemosti uchashchikhsya na osnove algoritma sluchainogo lesa [Implementation of Student Achievement Prediction based on Random Forest Algorithm]. In *Novosti nauki i tekhniki*. No. 8, pp. 37–39.
11. Zhuang, Qiaohui. (2020). Issledovanie i primenenie algoritma IRFC dlya prognozirovaniya otsenok uchashchikhsya po predmetu [Research and Application of IRFC Algorithm Based on Students' Course Achievement Prediction]. In *Informatsionnaya tekhnologiya i informatizatsiya*. No. 6, pp. 192–194.
12. Cheng, Xue. (2013). Primenenie uluchshennogo algoritma dereva prinyatiya reshenii pri analize otsenok starsheklassnikov [The Application of Improved Decision Tree Algorithm in the Analysis of Higher Vocational Grades]. In *Informatsiya i vychislitel'naya tekhnika*. No. 10, pp. 139–140.
13. Yuan, Zengren. (1999). *Neironnye seti i ikh primenenie* [Neural Networks and Their Applications]. Beijing, Izdatel'stvo Universiteta Tsinkhua.
14. Yue. Jingpeng. (2020). Issledovanie i primenenie algoritma prognozirovaniya uspevaemosti uchashchikhsya na osnove algoritma dereva prinyatiya reshenii C5.0 [Research and Application of Student Achievement Prediction Algorithm Based on C5.0 Decision Tree Algorithm]. *Tszilin'skii universitet*.
15. Mei, N., Qian, F., Yan, L. et al. (2018). Energy Efficiency Prediction of Screw Chillers on BP Neural Network Optimized by Improved Genetic Algorithm. In *International Computers, Signals and Systems Conference*. Dalian, pp. 527–531.