УДК 371.261

В. А. Мамонова,
факультет психолого-педагогического и специального образования,
Саратовский государственный национальный исследовательский университет им. Н. Г. Чернышевского
Научный руководитель: канд. пед. наук, доц. А. В. Букушева

Анализ успеваемости студентов методами машинного обучения

Аннотация. Рассматривается актуальность применения методов машинного обучения для анализа успеваемости обучающихся. Оценка результатов проведенного опроса среди студентов показала, что дерево решений и K-ближайших соседей имеют более высокую точность. Делается вывод о том, что применение алгоритмов машинного обучения позволяет исследователям и преподавателям распознавать ключевые факторы, которые существенно влияют на успеваемость обучающихся, и находить более эффективные способы повышения качества обучения.

Ключевые слова: машинное обучение, анализ образовательных данных, прогнозирование успеваемости.

Применение методов машинного обучения (МО) в сфере образования привлекло большое внимание исследователей [1; 2; 3]. Алгоритмы МО стали эффективным инструментом, используемым для выявления скрытых закономерностей в образовательных данных, прогнозирования академических достижений и улучшения среды обучения. В образовательном секторе анализ данных с помощью машинного обучения может помочь преподавателям принимать обоснованные решения и улучшать академические стратегии.

Аналитика обучения охватывает различные аспекты сбора информации о студентах, лучшего понимания среды обучения путем ее изучения и анализа [4]. Другим измерением аналитики обучения является прогнозирование успеваемости учащихся.

Цель данного исследования — прогнозирование успеваемости (среднего балла за сессию) студентов методами машинного обучения.

Нами был проведен опрос среди студентов 1–4-х курсов Саратовского государственного университета имени Н. Г. Чернышевского.

Использовались следующие входные данные: средний балл за последнюю и предыдущую сессию; промежуточные результаты за предыдущие семестры (выполнение лабораторных работ, тестов, домашних заданий, контрольных работ, защиту рефератов); возраст; пол; место проживания; результаты вступительных испытаний; удовлетворенность своей подготовкой к сессии [1; 5]. Обработка данных опроса проводилась на Python.

© Мамонова В. А., 2025

В опросе приняли участие 163 студента. Распределение по полу и возрасту представлено на рисунке 1.

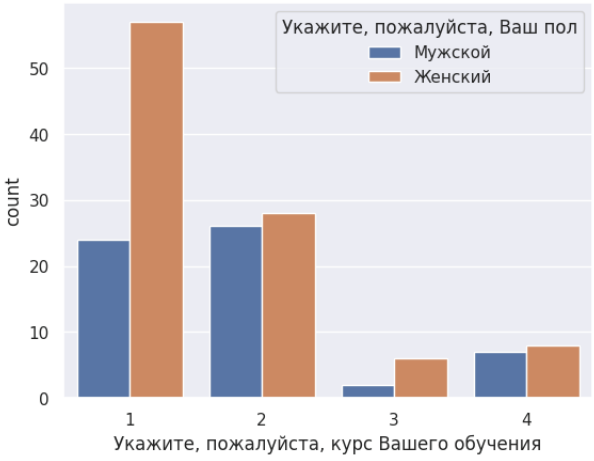


Рис. 1. Распределение студентов по курсу и полу

По трем вопросам нами построены облака слов:

– *Как Вы обычно готовитесь к экзаменам?* (рис. 2)

– *Какие факторы, по вашему мнению, наиболее сильно влияют на вашу успеваемость на экзаменах*? (укажите 2-3 самых значимых) (рис. 3а)

– *Какие эмоции вы обычно испытываете перед экзаменами?* (рис. 3б)

Целевым признаком являлась переменная «средняя оценка за сессию». Использовались

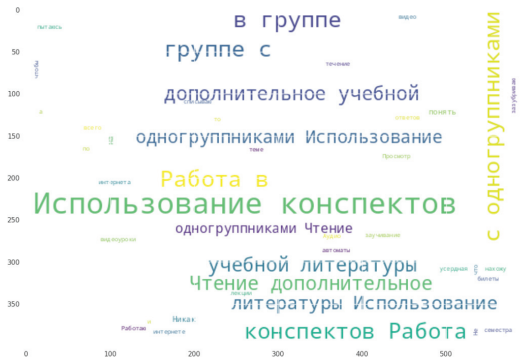


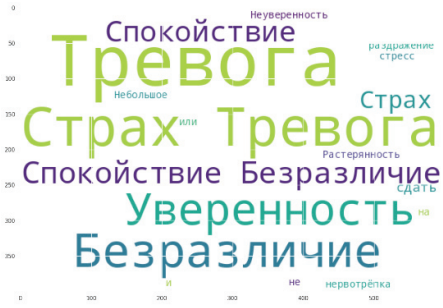
Рис. 2. Визуализация ассоциативных связей респондентов для категории «подготовка к экзаменам»

следующие методы машинного обучения: регрессия, дерево решений, K-ближайших соседей (KNN), случайный лес.

Дерево решений и KNN, используемые в работе, достигли точности 85,12 % и 89,14 % соответственно. В контексте данного исследования KNN применяется для прогнозирования категорий успеваемости обучающихся («отлично», «хорошо», «удовлетворительно») на основе их характеристик.

Академическая успеваемость студентов является важным показателем образовательного прогресса. Применение алгоритмов машинного обучения позволяет исследователям и преподавателям распознавать ключевые факторы, которые сильно влияют на успеваемость обучающихся, и находить более эффективные способы повышения качества обучения.

а)



б)

Рис. 3. Визуализация ассоциативных связей респондентов для категории «факторы, влияющие на экзамен» и «эмоции перед экзаменом»

Сравнительный анализ алгоритмов машинного обучения подтверждает их эффективность в прогнозировании результатов обучающихся [2]. Эти результаты согласуются с аналогичными исследованиями, подтверждая универсальность и точность этих классификаторов.

1. Прогнозирование результатов обучения студентов с использованием инструментов машинного обучения / Ю. Ю. Якунин, В. Н. Шестаков, Д. И. Ликсонова, А. А. Даничев // Информатика и образование. — 2023. — № 38(4). — С. 28–43.

2. Esmael A. Student performance prediction using machine learning algorithms // Applied Computational Intelligence and Soft Computing. — 2024. — Vol. 2024. — DOI: 10.1155/2024/4067721

3. Jin X. Predicting academic success: machine learning analysis of student, parental, and school efforts // Asia Pacific Educ. Rev. — 2023. — DOI: 10.1007/s12564-023-09915-4

4. Siemens G., Long P. Penetrating the fog: analytics in learning and education // EDUCAUSE Review. — 2011. — Vol. 46, no. 5. — P. 31–40.

5. Students Performance in Exams // Kaggle : [сайт]. — URL: https://www.kaggle.com/datasets/spscientist/students-performance-in-exams/data (дата обращения: 05.05.2025).