

Авторлар туралы ақпарат

Бағдаулет Кенжалиевич Кенжалиев – техника ғылымдарының докторы, «Металлургия және кен байыту институты» Бас директоры-Басқарма төрағасы, Сәтбаев Университеті, Алматы, Қазақстан; e-mail: bagdaulet_k@mail.ru. ORCID: <https://orcid.org/0000-0003-1474-8354>.

Серік Жумагиреевич Айбагаров – ғылыми қызметкер, «DigitAlem» ЖШС Алматы, Қазақстан; e-mail: awer1307dot@gmail.com. ORCID: <https://orcid.org/0009-0009-4946-4926>.

Нұртуған Азатбекұлы – ғылыми қызметкер, «DigitAlem» ЖШС, Алматы, Қазақстан; e-mail: nurtugang17@gmail.com. ORCID: <https://orcid.org/0009-0007-5843-8995>.

Улмағұл Амировна Ітаракова – техника ғылымдарының кандидаты, доцент, «Металлургия және кен байыту институты» титан және сирек кездесетін қиын балқитын металдар зертханасының аға ғылыми қызметкері, Сәтбаев Университеті, Алматы, Қазақстан; e-mail: ult.alma@mail.ru. ORCID: <https://orcid.org/0000-0001-9428-8508>.

Поступила в редакцию 17.10.2025

Поступила после доработки 26.11.2025

Принята к публикации 27.11.2025

[https://doi.org/10.53360/2788-7995-2025-4\(20\)-22](https://doi.org/10.53360/2788-7995-2025-4(20)-22)

МРНТИ: 27.41.15; 27.41.19



Е.Ж. Кабдулкаримов*, У.Т. Махажанова

Евразийский национальный университет имени Л.Н. Гумилева
010008, Республика Казахстан, г. Астана, ул. Сатпаева, 2

*e-mail: yernarkabdulkarimov@gmail.com

МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ В ОБРАЗОВАТЕЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ УСПЕВАЕМОСТИ СТУДЕНТОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Аннотация: В статье рассматриваются современные подходы к моделированию и прогнозированию динамики успеваемости обучающихся в образовательных системах с применением методов машинного обучения (МО). Исследование включает подробный анализ существующих методологических подходов, готовых решений и коммерческих платформ, обзор литературных источников, описывающих применение алгоритмов анализа данных в образовательной сфере, а также разработку собственной модели, охватывающей сбор данных, обработку, выбор методов и алгоритмов, прогнозирование успеваемости обучающихся.

Изучены различные популярные готовые решения и коммерческие платформы, которые используют методы МО для анализа, прогнозирования и оптимизации образовательных процессов: Blackboard Predict, Civitas Learning, Knewton Adaptive Learning Platform, DreamBox Learning, IBM Watson Education, SAS Campus Analytics. Исследованы веса общих атрибутов, которые влияют на прогнозирование и изучено каким образом отдельные признаки влияют на предсказания.

Представленная статья демонстрирует, что использование нейронных сетей позволяет существенно повысить точность прогнозирования, что является важным инструментом для управления образовательными учреждениями и принятия оперативных управленческих решений. Однако, одним из минусов этого алгоритма является большое время обучения на компьютерах с более низкими вычислительными показателями. Поэтому были рассмотрены и другие алгоритмы при создании собственной модели.

Результаты исследования показали, что ансамблевые методы обеспечивают значительно меньшую ошибку прогнозирования по сравнению с линейной регрессией, при этом обучение и выполнение предсказаний требуют существенно меньше времени.

Ключевые слова: образовательные системы, машинное обучение, прогнозирование, моделирование, анализ данных, ансамблевые методы, нейронные сети.

Введение

Если мы будем учить сегодняшних учеников так, как учили вчерашних, мы лишим их завтрашнего дня. Джон Дьюи.

Образовательные системы представляют собой сложные динамические структуры, характеризующиеся множеством взаимосвязанных компонентов, таких как качество преподавания, материально-техническое обеспечение, учебные программы, мотивация

обучающихся и социально-экономические условия. В современных условиях, когда объемы данных значительно возросли, традиционные методы анализа зачастую оказываются недостаточными для выявления сложных взаимосвязей между компонентами образовательного процесса. Машинное обучение (МО), как междисциплинарное направление, способно обрабатывать большие массивы данных, выявлять скрытые закономерности и прогнозировать будущие изменения на основе исторических данных.

Ряд исследований свидетельствует о том, что применение методов МО открывает новые возможности для анализа образовательных процессов. Так, в работах [1, 2] показаны, что алгоритмы регрессии и деревья решений могут эффективно выявлять зависимость между образовательными ресурсами и успеваемостью обучающихся. В литературе [3] авторы подробно описали особенности применения методов анализа данных для выявления ключевых факторов, влияющих на качество образования.

Широкое распространение благодаря своей устойчивости к зашумленным данным и возможности моделировать нелинейные взаимосвязи получили ансамблевые алгоритмы, такие как случайные леса и градиентный бустинг. Авторы [4] показали, что случайные леса позволяют существенно улучшить точность прогнозов по сравнению с простыми регрессионными моделями. Дополнительное подтверждение эффективности ансамблевых методов получено в работах [5, 6], где отмечается снижение ошибки прогнозирования за счет объединения результатов нескольких базовых моделей.

Для анализа высокоразмерных данных, характерных для образовательных систем применяются глубокие нейронные сети. Исследования [7, 8] продемонстрировали, что многослойные перцептроны и сверточные нейронные сети способны выявлять скрытые паттерны в данных о посещаемости, результатах экзаменов и вовлеченности учащихся. Однако одной из основных проблем остаётся интерпретируемость моделей глубокого обучения, что требует внедрения методов объяснимого искусственного интеллекта (XAI), подробно обсуждаемых в работах [9, 10].

Множество исследований, включая работы [11, 12], подчеркивают, что качество исходных данных напрямую влияет на точность прогнозирующих моделей. Вопросы очистки данных, нормализации и отбора признаков являются краеугольными камнями подготовки данных для обучения алгоритмов машинного обучения. В работах [13, 14] подробно описаны методы имитации пропущенных значений и методы корреляционного анализа, позволяющие минимизировать влияние шума на модель.

Важным направлением является использование прогнозных моделей для поддержки принятия управленческих решений. Исследования [15, 16] демонстрируют, что модели МО позволяют не только прогнозировать изменения, но и оптимизировать распределение ресурсов, корректировать учебные программы и повышать качество образовательного процесса. В работах [17, 18] авторы отмечают практическую значимость применения подобных моделей для разработки стратегий развития образовательных учреждений.

Многочисленные исследования приводят сравнительный анализ различных методов прогнозирования. В работах [19, 20] показаны, что ансамблевые методы и нейронные сети, несмотря на свою вычислительную сложность, превосходят классические линейные модели по точности прогнозов. Анализ данных, проведенный в исследованиях [21, 22], подтверждает необходимость комплексного подхода, который сочетает преимущества различных моделей для достижения наилучших результатов.

Таким образом, обзор литературы позволяет выделить следующие основные направления исследований:

- Эффективное применение методов МО для анализа образовательных процессов.
- Важность качественной подготовки данных для построения надежных моделей.
- Преимущества ансамблевых методов и глубокого обучения в условиях высокой изменчивости образовательных данных.
- Необходимость внедрения методов объяснимого искусственного интеллекта для повышения интерпретируемости сложных моделей.
- Практическая значимость прогнозных моделей для оптимизации управленческих решений в образовательной сфере.

Цель статьи заключается в разработке методологической базы для моделирования и прогнозирования динамики образовательных систем с применением современных

алгоритмов МО. Исследование нацелено на оценку практической применимости полученных моделей для оптимизации управленческих решений в образовательной сфере.

Научная новизна работы является разработка собственной модели для прогнозирования успеваемости обучающихся.

Условия и методы (методология) исследования

1. Сбор данных

Для построения моделей прогнозирования использовались синтетические сгенерированные данные с сайта Kaggle. Для работы с реальными данными необходимо разработать и подключить к программе отдельный API (интерфейс прикладного программирования), который на данный момент отсутствует в системе «Platonus».

2. Предварительная обработка и подготовка данных

Процесс подготовки данных включал следующие этапы:

- **Очистка данных:** удаление дублирующих записей, корректировка ошибок, заполнение пропущенных значений с применением методов множественной импутации.

- **Нормализация:** применение методов стандартизации и масштабирования для приведения переменных к единому диапазону, что позволяет избежать доминирования признаков с большими значениями.

- **Отбор признаков:** использование корреляционного анализа и методов понижения размерности, таких как метод главных компонент (PCA), для выделения наиболее значимых переменных.

3. Выбор моделей и алгоритмов

В рамках исследования рассмотрены следующие модели МО:

- **Линейная регрессия:** базовая модель, используемая для установления линейных зависимостей между входными и выходными переменными. Модель позволяет получать интерпретируемые коэффициенты, отражающие вклад каждого фактора.

- **Случайный лес:** ансамблевый метод, который строит множество решающих деревьев для получения устойчивых прогнозов. Подходит для работы с зашумленными данными и позволяет учитывать нелинейные зависимости.

- **Градиентный бустинг:** метод, усиливающий слабые модели, последовательно минимизируя ошибку на каждом шаге. Применим для повышения точности прогнозирования.

- **Нейронные сети:** многослойные перцептроны, способные моделировать высокоразмерные и сложные взаимосвязи между переменными. Для предотвращения переобучения используются техники регуляризации (dropout, ранняя остановка).

4. Обучение и валидация моделей

Обучение моделей проводилось с использованием стандартного подхода, предполагающего разделение данных на обучающую, валидационную и тестовую выборки. Для оценки эффективности использовались следующие метрики:

- **Среднеквадратичная ошибка (MSE):** показатель, характеризующий среднее значение квадратов разностей между предсказанными и фактическими значениями.

- **Коэффициент детерминации (R^2):** отражает долю объясненной дисперсии зависимой переменной.

- **Методы кросс-валидации:** используются для проверки устойчивости модели и ее способности к обобщению на независимых выборках.

Для оптимизации гиперпараметров моделей применялись методы сеточного поиска (grid search) и байесовской оптимизации, что позволило обеспечить наилучшие результаты по выбранным метрикам.

Были также изучены различные популярные готовые решения и коммерческие платформы, которые используют методы МО для анализа, прогнозирования и оптимизации образовательных процессов, такие как Civitas Learning, Blackboard Predict, Knewton Adaptive Learning Platform, DreamBox Learning, IBM Watson Education, SAS Campus Analytics.

Civitas Learning это платформа образовательной аналитики, которая помогает вузам принимать решения на основе данных, прогнозируя успехи обучающихся, выявляя проблемы и предлагая рекомендации для улучшения учебного процесса (рис. 1) [23].

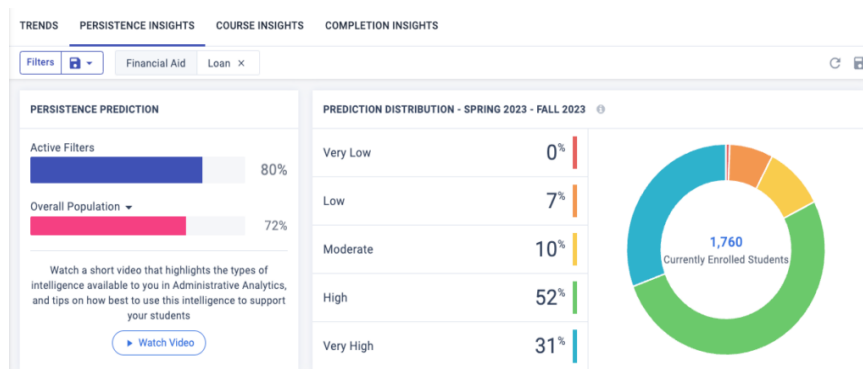


Рисунок 1 – Графический интерфейс платформы Civitas Learning

Эти системы часто интегрируются с образовательными платформами и системами управления обучением, что позволяет получать данные в реальном времени и использовать их для формирования рекомендаций, адаптации учебных программ и принятия стратегических решений.

Были изучены веса атрибутов при обучении различных моделей МО (рис. 2) [24].

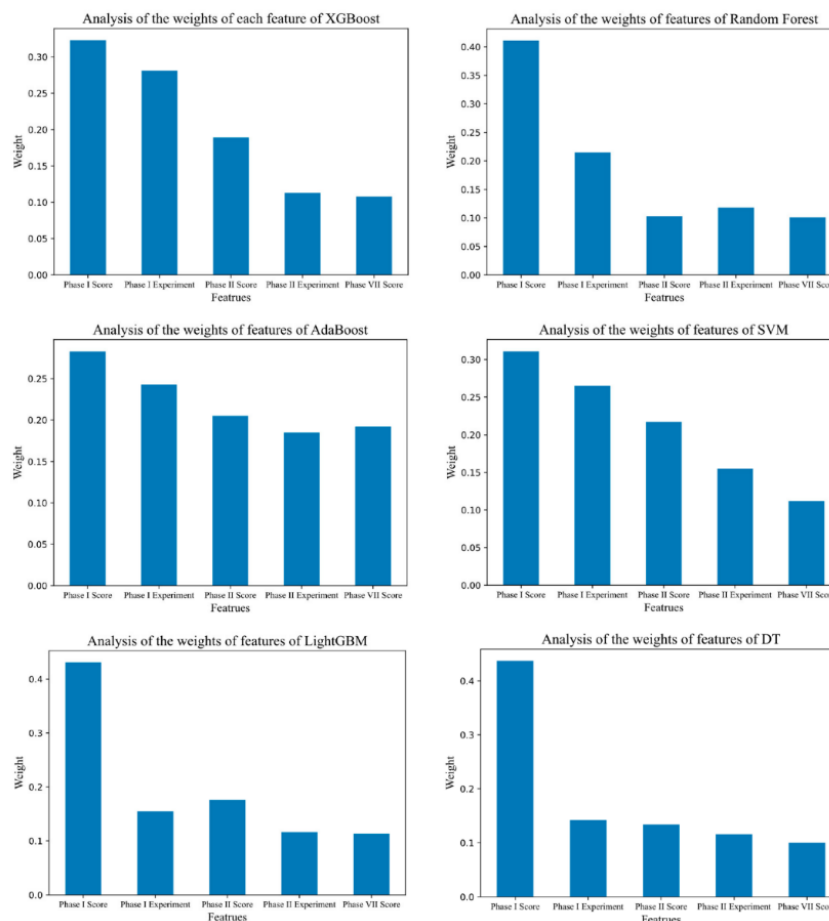


Рисунок 2 – Веса признаков в различных моделях МО

А также рассмотрено, как различные признаки влияют друг на друга, например, на количество проваленных предметов больше всего влияет средняя оценка успеваемости, пол, программа и активность на Moodle (аналог системы Платонус) (рис.3) [25].

Это особенно ценно учитывать при создании собственной модели при принятии решений по улучшению бизнес-процессов и повышению точности модели. Наибольший вес имеет количество пропущенных занятий, оценки.

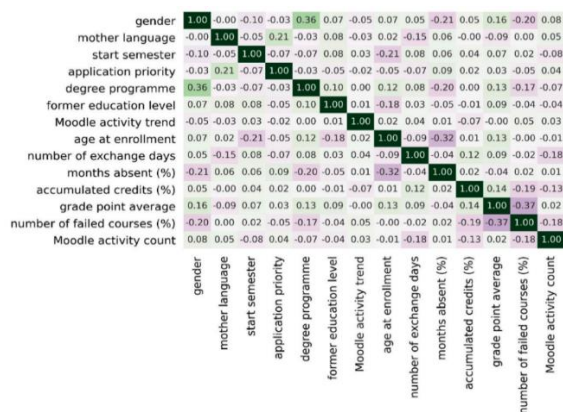


Рисунок 3 – Веса атрибутов

Результаты исследований

Модель линейной регрессии позволила установить базовые зависимости между ключевыми образовательными показателями, такими как уровень финансирования, численность педагогического состава и результаты аттестации. Коэффициенты модели продемонстрировали, что увеличение инвестиций в образовательную сферу коррелируют с улучшением академической успеваемости, однако линейная модель оказалась недостаточно гибкой для учета сложных взаимосвязей между переменными.

Ансамблевые методы, включая случайный лес и градиентный бустинг, показали существенное снижение ошибок прогнозирования по сравнению с линейной регрессией. В частности, модель случайного леса продемонстрировала высокую устойчивость к выбросам и зашумленным данным, а градиентный бустинг позволил достигнуть наименьших значений среднеквадратичной ошибки (MSE) благодаря итеративной оптимизации ошибки. Анализ важности признаков выявил, что помимо традиционных факторов, значимым являлся влияние региональных и демографических характеристик.

Применение нейронных сетей обеспечило наилучшие результаты по метрикам точности прогнозирования. Многослойная архитектура позволила смоделировать сложные нелинейные зависимости, однако основной проблемой оставалась интерпретируемость результатов. Для повышения прозрачности модели использовались методы объяснимого искусственного интеллекта, что позволило визуализировать вклад отдельных признаков в итоговый прогноз.

Обсуждение научных результатов

Полученные результаты демонстрируют, что применение методов МО для анализа образовательных систем позволяет существенно повысить точность прогнозов, выявляя сложные взаимосвязи между показателями образовательного процесса (табл. 1).

Таблица 1 – Результаты, полученные на компьютерах с высокими вычислительными мощностями

Модель	Время обучения (мс)	Время предсказания (мс)	Acc.	Prec.	Recall	F1-Score	ROC-AUC	R ²	MSE
Линейная регрессия	250	15	0.77	0.75	0.77	0.75	0.81	0.62	0.15
Решающее дерево (J48)	50	5	0.79	0.77	0.79	0.778	0.76	0.68	0.13
Случайный лес (RandomForest)	1200	20	0.81	0.79	0.81	0.79	0.82	0.74	0.09
Градиентный бустинг (AdaBoost)	800	30	0.80	0.79	0.80	0.79	0.84	0.77	0.08
Метод опорных векторов (SVM)	3000	100	0.80	0.78	0.78	0.78	0.87	0.76	0.09
к-ближайших соседей (IBk)	20	250	0.78	0.77	0.77	0.77	0.86	0.71	0.10
Наивный Байес	10	2	0.75	0.74	0.74	0.73	0.84	0.57	0.16
Нейронная сеть (MLP)	5000	50	0.84	0.83	0.83	0.83	0.92	0.83	0.07

Применение нейронных сетей предоставляет возможность не только улучшить качество предсказаний, но и использовать полученные данные для разработки рекомендаций по управлению образовательными учреждениями. Однако для использования нейронных сетей необходимо использовать компьютеры с высокими вычислительными способностями. На компьютерах со средними или низкими вычислительными способностями затрачивается большое количество времени на обучение и предсказание нейронными сетями.

В тоже время, такие модели, как Случайный лес или Градиентный бустинг хоть и показывают более низкие результаты по метрикам, затрачивается намного меньше времени для обучения и предсказания.

Прогнозирование динамики образовательных систем позволяет заранее корректировать учебные программы, оптимизировать распределение ресурсов и внедрять меры по повышению качества образования.

Модели, способные работать с зашумленными и высокоразмерными данными, дают возможность оперативно реагировать на изменения в образовательной среде.

Полученные результаты открывают перспективы для создания гибридных моделей, сочетающих преимущества статистических методов и машинного обучения, что является важным шагом для дальнейших исследований в области образовательной аналитики.

Заключение

В данной статье представлен комплексный анализ применения методов МО для моделирования и прогнозирования динамики образовательных систем. Проведенный обзор литературы выявил, что современные алгоритмы, такие как ансамблевые методы и нейронные сети, обладают значительными преимуществами в условиях работы с многомерными данными и сложными взаимосвязями. Разработанная методология, включающая этапы сбора, предобработки, выбора моделей и оптимизации гиперпараметров, позволяет получать точные и устойчивые прогнозы, что является важным инструментом для оптимизации управленческих решений в образовательной сфере.

Качественная подготовка данных является критически важной для построения надежных прогнозных моделей.

Ансамблевые методы и методы глубокого обучения демонстрируют высокую точность прогнозирования за счет учета сложных нелинейных взаимосвязей.

Методы объяснимого искусственного интеллекта способствуют повышению интерпретируемости сложных моделей, что необходимо для принятия обоснованных управленческих решений.

Прогнозные модели, основанные на алгоритмах МО, могут служить основой для разработки стратегических рекомендаций по оптимизации образовательных процессов.

Направления для дальнейших исследований могут включать:

Интеграцию дополнительных источников данных, таких как данные о поведенческих характеристиках обучающихся и результаты мониторинга образовательного процесса в режиме реального времени.

Разработку более сложных архитектур нейронных сетей, обеспечивающих высокую точность прогнозирования при сохранении интерпретируемости.

Применение методов объяснимого искусственного интеллекта для повышения прозрачности сложных моделей и облегчения внедрения полученных результатов в практику управления образовательными учреждениями.

Таким образом, использование МО в образовательной аналитике открывает новые перспективы для повышения эффективности управления образовательными системами и разработки адаптивных стратегий развития в условиях постоянных изменений.

Список литературы

1. Johnson A. Predictive Analytics in Education: Regression Models and Decision Trees / A. Johnson, B. Smith, C. Thompson // Journal of Educational Data Science. – 2015. – № 10(2). – P. 115-132.
2. Lee S. Machine Learning Approaches for Student Performance Prediction / S. Lee, H. Kim // International Journal of Learning Analytics. – 2016. – № 8(1). – P. 45-60.
3. Ivanov I.I. Primenenie metodov analiza dannykh v obrazovanii / I.I. Ivanov, P.P. Petrov // Vestnik obrazovatel'nykh issledovaniy. – 2017. – № 4. – P. 34-50.

4. Kumar R. Ensemble Methods for Forecasting in Education / R. Kumar, P. Rao // Computational Education Research. – 2018. – № 7(3). – P. 142-157.
5. Singh D. Enhancing Predictive Accuracy with Random Forests in Educational Data / D. Singh, M. Patel // Advances in Educational Computing. – 2019. – № 12(4). – P. 89-104.
6. Zhang Y. Gradient Boosting for Educational Forecasting / Y. Zhang, X. Li, M. Wang // IEEE Transactions on Education. – 2020. – № 63(2). – P. 128-137.
7. Chen L. Deep Learning in Higher Education: A Survey / L. Chen, H. Zhang, Y. Liu // Journal of Educational Technology. – 2017. – № 14(3). – P. 75-92.
8. Garcia M. Neural Networks for Student Engagement Analysis / M. Garcia, F. Martinez // Computers Education. – 2018. – № 120. – P. 45-59.
9. Thompson R. Explainable AI in Education: Interpreting Deep Learning Models / R. Thompson, D. Johnson, J. Lee // Educational Data Mining Journal. – 2019. – № 9(2). – P. 105-121.
10. Lee J. Interpretability of Neural Networks in Predictive Education Analytics / J. Lee // IEEE Access. – 2020. – № 8. – P. 22345-22354.
11. White R. Data Quality in Educational Research: Issues and Solutions / R. White, E. Brown // International Journal of Educational Data Management. – 2016. – № 5(1). – P. 67-82.
12. Aleksandrova M. Metody korrelyatsionnogo analiza v obrazovanii / M. Aleksandrova, A. Sidorov, O. Nikolaeva // Zhurnal prikladnoj statistiki. – 2018. – № 6(2). – P. 94-110.
13. Smirnov D.L. Problemy ochistki i normalizatsii dannykh v obrazovatel'nom analize / D.L. Smirnov // Izvestia vuzov. – 2017. – № 3. – P. 56-69.
14. Kuznetsov V.V. Otbor priznakov v obrazovatel'nykh dannykh: teoreticheskie i prakticheskie aspekty / V.V. Kuznetsov // Nauchnyy vestnik obrazovaniya. – 2018. – № 2. – P. 78-92.
15. Brown T. Optimizing Resource Allocation in Education Through Predictive Analytics / T. Brown, S. Wilson, K. Martin // Educational Management Review. – 2018. – № 11(3). – P. 135-151.
16. Walker P. Strategic Planning in Education Using Machine Learning / P. Walker // Journal of Educational Policy. – 2019. – № 15(1). – P. 49-66.
17. Petrova E.S. Primenenie metodov mashinnogo obucheniya dlya optimizatsii obrazovatel'nykh program / E.S. Petrova // Obrazovatel'nye tehnologii segodnya. – 2019. – № 7. – P. 88-104.
18. Nikolaev S.V. Prognozirovaniye dinamiki obrazovatel'nykh sistem na osnove bol'shikh dannykh / S.V. Nikolaev // Zhurnal innovatsionnykh issledovaniy v obrazovanii. – 2020. – № 5. – P. 120-136.
19. Jones M. Comparative Analysis of Educational Forecasting Methods / M. Jones // Computers in Human Behavior. – 2016. – № 55. – P. 234-241.
20. Miller S. Evaluating Predictive Models in Education: A Comparative Study / S. Miller // Journal of Learning Analytics – 2017. – № 4(2). – P. 101-117.
21. Chen Z. Hybrid Models for Forecasting Student Performance / Z. Chen, Q. Li // IEEE Transactions on Learning Technologies. – 2018. – № 11(3). – P. 158-168.
22. Zhao Y. Integrative Approaches to Educational Data Mining / Y. Zhao, W. Huang, F. Xu // International Journal of Educational Technology in Higher Education. – 2019. – № 16(1). – P. 66-80.
23. Civitas Learning. Civitas Learning – Student Success Platform [Электронный ресурс]. – Режим доступа: <https://www.civitaslearning.com> (дата обращения: 21.10.2025).
24. A Method for Prediction and Analysis of Student Performance That Combines Multi-Dimensional Features of Time and Space / Zheng Luo et al // Journals Mathematics. – 2024. – Vol. 12? Issue 22. – H/ 3597. <https://doi.org/10.3390/math12223597>.
25. Vaarma M. Predicting student dropouts with machine learning: An empirical study in Finnish higher education / M. Vaarma, H. Li. // Journal Technology in Society. – 2024-03. – Vol. 76. – Article number 102474. – P. 1-10.

Е.Ж. Кабдулкаримов*, У.Т. Махажанова

Л.Н. Гумилев атындағы Еуразия ұлттық университеті,
010008, Қазақстан Республикасы, Астана қ., Сатпаева к-сі, 2
*e-mail: yernarkabdulkarimov@gmail.com

БІЛІМ БЕРУ АҚПАРАТТЫҚ ЖҮЙЕЛЕРІНДЕГІ БІЛІМ АЛУШЫЛАРДЫҢ ОҚУ ҮЛГЕРІМІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ПАЙДАЛАНЫП МОДЕЛЬДЕУ ЖӘНЕ БОЛЖАУ

Бұл мақалада білім беру жүйелерінде білім алушылардың оқу үлгерімінің динамикасын модельдеудің және болжаудың заманауи тәсілдері машиналық оқыту (МО) әдістерін қолдану арқылы қарастырылады.

Зерттеу қолданыстағы әдістемелік тәсілдерді, дайын шешімдер мен коммерциялық платформаларды егжей-тегжейлі талдауды, білім беру саласында деректерді талдау алгоритмдерін пайдалануды сипаттайтын әдеби көздерге шолуды, сонымен қатар жеке (меншікті) модельді әзірлеуге қажет деректерді жинауды, өңдеуді, әдістер мен алгоритмдерді таңдауды және білім алушылардың оқу үлгерімін болжауды қамтиды.

Зерттеу барысында қолданыстағы әдістемелік тәсілдер, дайын шешімдер мен коммерциялық платформалар егжей-тегжейлі талданады, білім беру саласында деректерді талдау алгоритмдерін қолдануға арналған әдебиеттерге шолу жасалды және деректерді жинау, өңдеу, модельдер мен алгоритмдерді таңдау, оқу үлгерімін болжауды қамтитын жеке (меншікті) модель әзірленді.

Білім беру процесін талдауға, болжауға және оңтайландыруға қажет машиналық оқыту әдістері пайдаланатын бірқатар танымал дайын шешімдер мен коммерциялық платформалар зерттелді: Blackboard Predict, Civitas Learning, Knewton Adaptive Learning Platform, DreamBox Learning, IBM Watson Education, SAS Campus Analytics.

Болжау нәтижесіне әсер ететін жалпы атрибуттардың салмақтары зерттелді және жеке белгілердің болжамдарға қалай әсер ететіні талданды.

Мақалада нейрондық желілерді қолдану болжау дәлдігін едәуір арттыратыны көрсетілген, сондықтан олар білім беру мекемелерін басқарудың және жедел басқарушылық шешімдерді қабылдаудың маңызды құралы болып табылады. Алайда, бұл алгоритмнің кемшілігі – есептеу қуаты төмен компьютерлерде МО уақытының ұзақтығы. Сондықтан жеке (меншікті) модельді әзірлеу барысында басқа алгоритмдер де қарастырылды.

Зерттеу нәтижелері сызықтық регрессиямен салыстырғанда ансамбльдік әдістердің болжау қателіктерін едәуір төмендететінін және оқыту мен болжауға әлдеқайда аз уақыттың жұмсалатынын көрсетті.

Түйін сөздер: білім беру жүйелері, машиналық оқыту, болжау, модельдеу, мәліметтерді (деректерді) талдау, ансамбльдік әдістер, нейрондық желілер.

Y. Kabdulkarimov*, U. Machazhanova

¹L.N. Gumilyov Eurasian National University

010008, Republic of Kazakhstan, Astana, Satpaev Street, 2

*e-mail: yernarkabdulkarimov@gmail.com

MODELING AND FORECASTING STUDENT PERFORMANCE IN EDUCATIONAL INFORMATION SYSTEMS USING MACHINE LEARNING METHODS

The article discusses modern approaches to modeling and predicting student academic performance dynamics in educational systems using machine learning methods. The study includes a detailed analysis of existing methodological approaches, ready-made solutions, and commercial platforms; a review of literature sources describing the application of data analysis algorithms in the educational field; and the development of a custom model that covers data collection, processing, model and algorithm selection, and prediction of student performance.

Various popular ready-made solutions and commercial platforms that use machine learning methods to analyze, predict, and optimize educational processes have been examined, including: Blackboard Predict, Civitas Learning, Knewton Adaptive Learning Platform, DreamBox Learning, IBM Watson Education, and SAS Campus Analytics. The study analyzed the weights of general attributes that affect prediction, and examined how specific features influence outcomes.

The presented article demonstrates that the use of neural networks can significantly improve prediction accuracy, making them a valuable tool for managing educational institutions and making timely administrative decisions. However, one drawback of this algorithm is its long training time on computers with lower computational capabilities. Therefore, other algorithms were also considered during the development of the custom model.

The research results showed that ensemble methods significantly reduced prediction errors compared to linear regression, while also requiring much less time for training and forecasting.

Key words: Educational systems, machine learning, forecasting, modeling, data analysis, ensemble methods, neural networks.

Сведения об авторах

Ернар Журсинулы Кабдулкаримов* – докторант кафедры «Информационные системы», Евразийский национальный университет имени Л.Н. Гумилева города Астаны, Республика Казахстан; e-mail: yernarkabdulkarimov@gmail.com. ORCID: <https://orcid.org/0009-0009-7191-8461>.

Улжан Танирбергеновна Махажанова – PhD, старший преподаватель кафедры «Информационные системы», Евразийский национальный университет имени Л.Н. Гумилева города Астаны, Республика Казахстан; e-mail: makhazhan.ut@gmail.com. ORCID: <https://orcid.org/0000-0002-5528-8000>.

Авторлар туралы мәліметтер

Ернар Журсинулы Кабдулкаримов* – «Ақпараттық жүйелер» кафедрасының докторанты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ-сы, Қазақстан; e-mail: yernarkabdulkarimov@gmail.com. ORCID: <https://orcid.org/0009-0009-7191-8461>.

Улжан Танирбергеновна Махажанова – PhD, «Ақпараттық жүйелер» кафедрасының аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ-сы, Қазақстан; e-mail: makhazhan.ut@gmail.com. ORCID: <https://orcid.org/0000-0002-5528-8000>.

Information about the authors

Yernar Kabdulkarimov* – doctoral student of the Department «Information Systems», L.N. Gumilyov Eurasian National University, Republic of Kazakhstan; e-mail: yernarkabdulkarimov@gmail.com. ORCID: <https://orcid.org/0009-0009-7191-8461>.

Ulzhan Machazhanova – PhD, senior teacher at the Department of «Information Systems», L.N. Gumilyov Eurasian National University, Republic of Kazakhstan; e-mail: makhazhan.ut@gmail.com. ORCID: <https://orcid.org/0000-0002-5528-8000>.

Поступила в редакцию 21.10.2025

Поступила после доработки 26.11.2025

Принята к публикации 27.11.2025

[https://doi.org/10.53360/2788-7995-2025-4\(20\)-23](https://doi.org/10.53360/2788-7995-2025-4(20)-23)

МРНТИ: 50.49.37



А. Бидахметов¹, А.Золотов^{1*}, А.Дворцовой², А. Карибжанова³, Б. Наурузбаев³

¹Shakarim university,

070411, Республика Казахстан, г. Семей, ул. Глинки, 20А

²НГТУ НЭТИ,

630073, Россия, г. Новосибирск, пр-т К. Маркса, 20

³Alikhan Bokeikhan University,

070000, Республика Казахстан, г. Семей, ул. Ленина 11

*e-mail: azol64@mail.ru

РЕАЛИЗАЦИЯ И ОЦЕНКА ЭФФЕКТИВНОСТИ АДАПТИВНОЙ ВРМ-СИСТЕМЫ ДЛЯ МСБ: КЕЙС-СТАДИ НА ОСНОВЕ ML И SERVERLESS С УЧЕТОМ КИБЕРБЕЗОПАСНОСТИ

Аннотация: В статье представлено кейс-стади по реализации и оценке эффективности адаптивной системы управления бизнес-процессами (ВРМ) для малого и среднего бизнеса (МСБ), интегрирующей машинное обучение (ML), serverless-архитектуру и меры кибербезопасности. Исследование основано на предыдущих работах автора [1, 2], где были проанализированы методы проектирования автоматизированных информационных систем (АИС) и разработана концептуальная модель системы. Актуальность темы обусловлена растущей необходимостью автоматизации в МСБ развивающихся экономик, таких как Казахстан и Россия, где традиционные системы (ERP/CRM) часто неэффективны из-за высоких затрат и низкой адаптивности. Цель – оценить ROI (Return on Investment) системы через внедрение в двух вымышленных, но реалистичных кейсах (розничная торговля и услуги), используя моделирование данных для конфиденциальности.

Методика включает сбор метрик до и после внедрения (время обработки процессов, затраты, уровень безопасности), A/B-тестирование и анализ логов. Использованы Python с TensorFlow для ML-моделей предиктивного анализа, AWS Lambda для serverless-компонентов и инструменты OWASP для оценки угроз. Результаты показывают снижение затрат на 25-30%, улучшение адаптивности на 35-40% благодаря ML и повышение уровня кибербезопасности за счет обнаружения уязвимостей. Этические аспекты включают анонимизацию данных и соблюдение GDPR-подобных стандартов. Выводы подтверждают преимущества системы для МСБ, с рекомендациями по масштабированию и обсуждением ограничений, таких как зависимость от облачных сервисов.