

Интеллектуальный анализ классификации абитуриентов с помощью нейронных сетей на основе адаптивной функции активации

Н. Болотбек уулу, С. Н. Верзунов, А. Б. Салиев, И. Р. Мусина

Кыргызский государственный технический университет им. И. Раззакова

Кафедра «Программное обеспечение компьютерных систем»

E-mail: n.bolotbekuulu@kstu.kg

Аннотация:

Выбор специальности при поступлении в вуз является важнейшим этапом в академической жизни абитуриента, так как этот выбор влияет на дальнейшее развитие карьеры и профессиональную реализацию. Однако в условиях ограниченной информации, с которой абитуриенты сталкиваются на этапе выбора, существует риск принятия необоснованных решений, что может привести к неудовлетворенности учебным процессом, снижению академической успеваемости и необходимости смены направления обучения. В данной работе рассматривается разработка системы интеллектуальных рекомендаций на основе нейросетевого подхода, где ключевым элементом является адаптивная функция активации. В отличие от стандартных функций активации, она динамически изменяет форму отклика нейронов, что повышает гибкость модели при обучении и улучшает точность предсказаний. В исследовании проводится сравнительный анализ базовой модели и модели с адаптивной активацией, демонстрирующий повышение точности классификации специальностей. Полученные результаты подтверждают эффективность предлагаемого подхода для поддержки абитуриентов в принятии решений.

Введение

Процесс выбора специальности в высшем учебном заведении является одной из ключевых задач для абитуриента, поскольку этот выбор оказывает непосредственное влияние на его академическое развитие, карьерные перспективы и профессиональную реализацию. Однако, несмотря на существующие формализованные методы выбора, многие студенты сталкиваются с рядом проблем, связанных с несоответствием их реальных способностей и выбранной специальности. Традиционные подходы, основанные на результатах Общереспубликанского тестирования (ОРТ), учитывают лишь поверхностные критерии, что нередко приводит к выбору, не соответствующему академическим возможностям студентов. В результате, в ходе обучения студенты часто сталкиваются с трудностями в освоении предметов, что ведет к низкой академической успеваемости, снижению мотивации и даже необходимости смены направления подготовки.

Таким образом, существует явная потребность в разработке более эффективных инструментов для поддержки абитуриентов в процессе выбора. В современных условиях машинное обучение представляет собой мощный инструмент, способный анализировать и интегрировать большой объем данных, чтобы предоставить более персонализированные и объективные рекомендации.

Текущие подходы и их ограничения

1.1 Экспертные системы – базируются на заранее заданных правилах и критериях, разработанных экспертами в области образования. Однако такие системы имеют ограниченную адаптивность, так как не учитывают динамические изменения в образовательной среде и результаты тестирования абитуриента.

1.2 Анализ исторических данных – предполагает использование статистических методов и корреляционного анализа на основе данных о прошлых наборах студентов. Такой подход позволяет выявить основные факторы, влияющие на успех в обучении, но не дает возможности персонализированного прогнозирования.

1.3 Совсем иной подход предполагает использование методов машинного обучения и искусственного интеллекта. Эти методы позволяют учитывать широкий спектр данных и выявлять сложные зависимости, что открывает новые возможности для создания более точных и персонализированных рекомендаций. Например, с помощью алгоритмов машинного обучения можно интегрировать данные о результатах вступительных экзаменов, исторических успехах студентов, а также других характеристиках, таких как средний балл, успехи в различных дисциплинах и даже дополнительные сведения о трудовых предпочтениях.

В данной работе рассматривается возможность применения методов машинного обучения для построения рекомендательной системы, которые могут помочь абитуриентам сделать более осознанный выбор специальности, исходя из их результатов тестирования. Используемый подход ориентирован на выявление подходящей специальности для абитуриентов и определения вероятности попадания на бюджетную форму обучения.

Методы

Описание данных

Разработка рекомендательной системы основывалась на комплексном подходе, включающем сбор, обработку и анализ данных, применение методов машинного обучения, а также исследование существующих подходов к образовательным рекомендациям. В ходе работы учитывались ключевые факторы, влияющие на успешность поступления абитуриентов и их распределение по направлениям подготовки.

Процесс сбора и подготовки данных включал несколько этапов. На первом этапе осуществлялся сбор исходных данных, полученных из официальных источников - Информационный портал AVN KSTU, содержащих сведения об абитуриентах, их академических достижениях и условиях поступления. В выборку были включены показатели в виде баллов за общереспубликанское тестирование (OPT) и предметные тесты, а также информация о форме финансирования обучения (бюджет или контракт).

Далее проводился этап обезличивания данных, в ходе которого персональная информация абитуриентов удалялась или

шифровалась в целях обеспечения конфиденциальности. В процессе стандартизации данные приводились к единому формату, устранялись пропущенные значения, а также исправлялись возможные ошибки ввода. Процесс описан на рис. 1.



Рисунок 1 - Схема получения и подготовки данных

В данном исследовании использовались реальные данные о поступивших абитуриентах за 2024 год. Датасет состоит из двух частей:

- *contract.csv* – данные абитуриентов, поступающих на контрактной основе;
- *budget.csv* – данные абитуриентов, поступающих на бюджетной основе.

Каждая запись содержит следующие атрибуты:

- *OPT* – балл за общереспубликанский тест;
- *Балл по предметам*: биология, химия, физика, английский, математика;
- *Category* – категория абитуриента (город, село, горная местность и др.);
- *Department* – выбранное направление подготовки;
- *Budget* – индикатор поступления (1 – бюджет, 0 – контракт).

Данные предварительно объединялись, при этом для каждого набора добавлялась соответствующая метка поступления. Гистограмму значений дополнительных параметров можно посмотреть на рис. 2.

Распределение значений дополнительных параметров

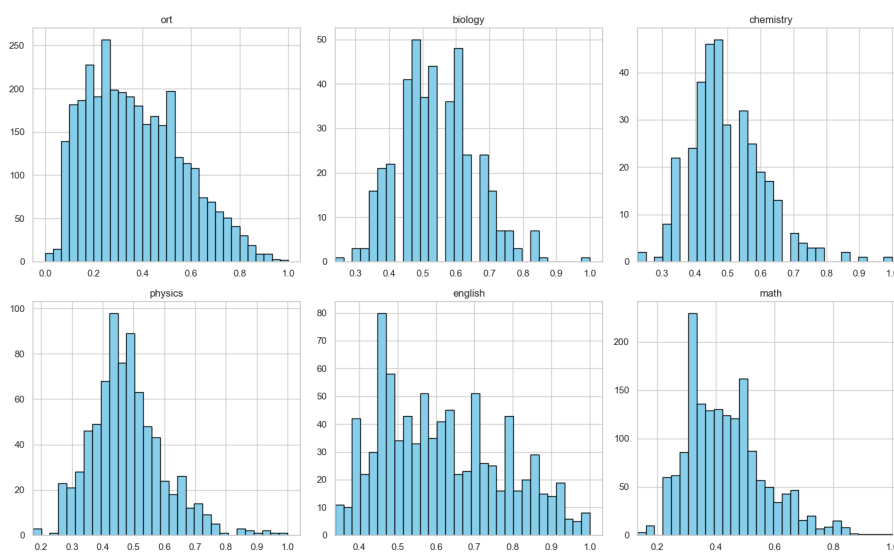


Рисунок 2 - Гистограмма распределений основного и дополнительных предметов по OPT

Построение моделей

Для решения задачи рекомендации специальностей на основе интеллектуального анализа данных были разработаны две модели с использованием библиотеки *TensorFlow/Keras*:

Базовая Модель

Модель представлена в виде сети, состоящей из двух полносвязных слоев (128 и 64 нейрона, соответственно) с функцией активации *ReLU*. Для повышения устойчивости к переобучению применялись слои *Batch Normalization* и *Dropout*. Выход модели разделяется на два блока:

• **Department:** классификация по выбранному направлению подготовки с использованием *softmax*-активации;

• **Budget:** бинарная классификация поступления (бюджет/контракт) с использованием сигмоидальной функции активации.

В качестве функции потерь использовались *sparse_categorical_crossentropy* для классификации специальностей и *binary_crossentropy* для определения типа поступления.

Модель с адаптивной активацией

Для оценки возможности улучшения результатов была построена модель, использующая разработанный нами слой *AdaptiveActivation*. Данный слой комбинирует функции активации *sigmoid* и *tanh*, где коэффициенты (*alpha*) обучаются в процессе

тренировки. Архитектура модели аналогична базовой, но вместо стандартной функции активации применяются адаптивные функции активации, что позволяет модели лучше подстраиваться под особенности данных. В данной модели адаптивная функция активации задается формулой:

$$f(x) = \alpha * \sigma(x) + (1 - \alpha) * \tanh(x),$$

где $\sigma(x) = \frac{1}{1+e^{-x}}$ – сигмоидальная функция активации; $\tanh(x)$ – гиперболический тангенс; α – настраиваемый параметр, значение которого оптимизируется во время обучения.

Эта функция представляет собой линейную комбинацию сигмоидальной функции и гиперболического тангенса. Параметр определяет вклад каждой из функций:

Если приближается 1, функция активации больше напоминает $\sigma(x)$;

Если приближается к 0, преобладает $\tanh(x)$.

В противном случае функция активации представляет собой комбинацию двух этих функций

Таким образом, адаптивная функция активации позволяет модели динамически выбирать оптимальную форму нелинейности для каждого признака, что может способствовать более эффективному обучению и лучшей обобщающей способности сети.

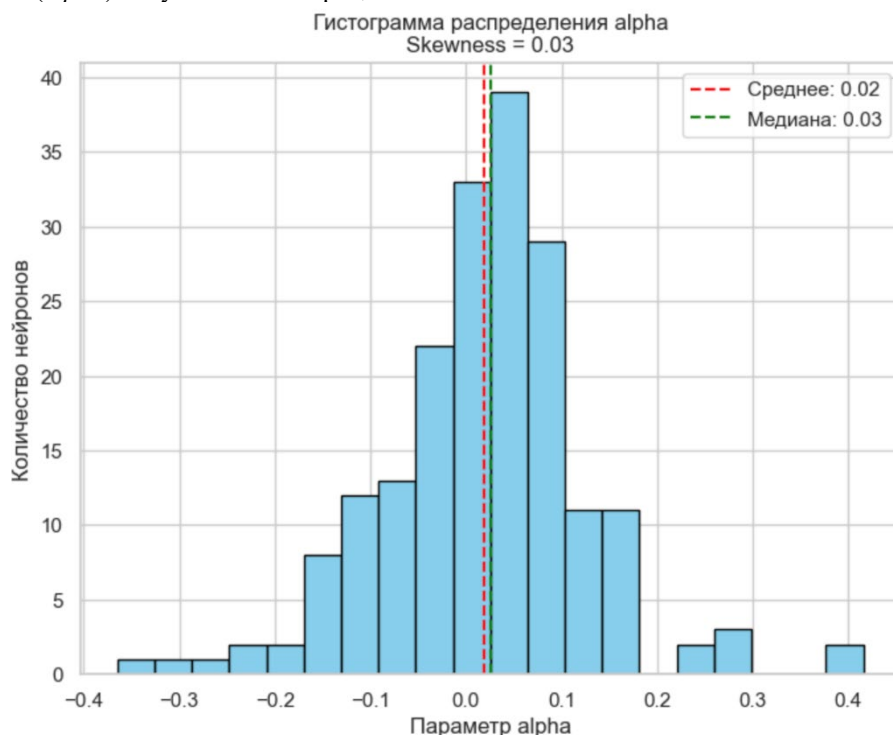


Рисунок 3 - Гистограмма распределения параметра Alpha

Обучение и оценка моделей

Обе модели обучались в течение 50 эпох с использованием оптимизатора Adam и размера пакета, равного 32. Для каждой модели проводилось обучение с одновременной оптимизацией двух выходов – классификации специальностей и предсказания типа поступления (контракт или бюджет). В качестве метрики для оценки использовалась точность (accuracy).

После обучения модели были протестированы на валидационной выборке, после чего проведена оценка по следующим параметрам:

- Общая функция потерь (*Loss*);
- Точность предсказания специальности (*Department Accuracy*);
- Точность определения типа поступления (*Budget Accuracy*).

Для дополнительного анализа были вычислены предсказания для специальностей и рассчитана итоговая точность классификации (рис 4).

Результаты

Графики зависимости точности и потерь от количества эпох (рис. 4) показывает, что модель относительно хорошо справляется с

предсказанием специальностей и бюджетных мест. Видно, что точность предсказания специальности (синий и оранжевый) остается довольно низкой (~10-15%), что говорит о сложности задачи. Однако, по сравнению со случайным определением специальности (3.9%), она оказывается достаточно высокой

• Точность предсказания бюджетных мест (зелёный и красный) значительно выше (~60-65%), что указывает на лучшее качество предсказаний для этой задачи.

• Разница между результатами на обучающей и валидационной выборке минимальна, что свидетельствует об отсутствии переобучения.

Точность предсказания специальности остается на уровне (~10-15%), что может указывать на необходимость улучшения модели за счет сбора дополнительных данных. Однако, нет явных признаков переобучения.

При рассмотрении графика ошибки (рис. 5) можно увидеть, что происходит ее снижение, но при этом график ошибки не снижается до нуля, что говорит о том, что имеющихся данных оказывается недостаточно.

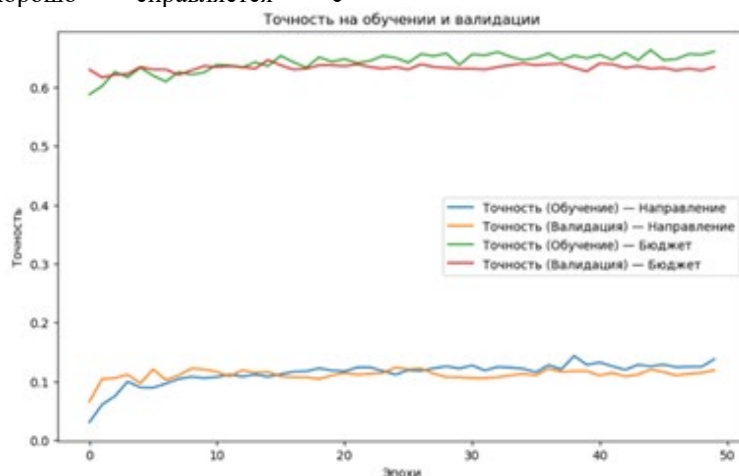


Рисунок 4 - График точности на обучении и валидации

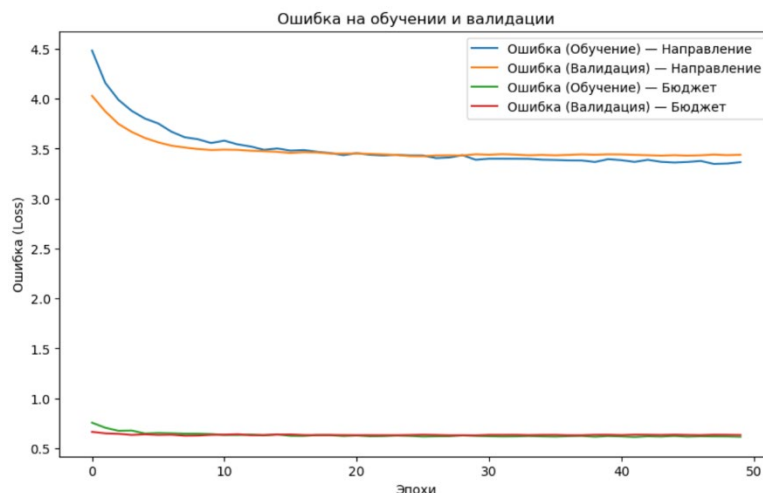


Рисунок 5. График ошибки на обучении и валидации

Результаты эксперимента показали, что использование адаптивной функции активации (*AdaptiveActivation*) привело к улучшению точности предсказания направления подготовки (*Department Prediction Accuracy*) по сравнению с базовой моделью, использующей ReLU.

Ключевые показатели точности:

- Базовая модель (ReLU): 10,70 %
- Модель с *AdaptiveActivation*: 11,89 %.

Абсолютное улучшение составило 1.19 процентных пункта, а относительное — около 11% по сравнению с базовой моделью: Этот результат демонстрирует, что адаптивная функция активации может повышать качество классификации, динамически подстраивая отклик слоя к различным паттернам в данных. Однако, несмотря на улучшение, абсолютное значение точности остается низким, что указывает на сложность задачи классификации и необходимость дальнейшего совершенствования модели за счёт сбора дополнительных данных.

Адаптивная функция активации играет важную роль в улучшении точности модели за счет динамического комбинирования двух различных функций — сигмоиды и гиперболического тангенса. Это позволяет каждому нейрону самостоятельно определять, какая функция активации лучше подходит для конкретного набора входных данных. Однако, для иных задач, можно комбинировать подобным образом и другие функции активации.

Преимущества Adaptive Activation

Активационные функции с адаптивными параметрами обладают рядом преимуществ по сравнению с традиционными фиксированными функциями, такими как ReLU, сигмоида и tanh. Основное преимущество — **гибкость**: нейроны могут адаптироваться к различным типам входных данных, что повышает универсальность и точность модели.

Снижение ограничений традиционных функций:

- **ReLU** может приводить к возникновению "мертвых нейронов", когда все значения меньше нуля обнуляются и нейрон перестает обучаться.

- **Сигмоида** и **tanh** страдают от проблемы затухания градиентов, особенно в глубоких сетях.

- **Adaptive Activation** объединяет сильные стороны этих функций и снижает их недостатки, обеспечивая стабильное распространение градиентов.

Автоматическая настройка

Модель самостоятельно подбирает оптимальное соотношение между различными формами активации (например, между сигмоидой и тангенсом), избавляя от необходимости вручную выбирать функции активации для каждого слоя.

Анализируя результаты, можно предположить, что *AdaptiveActivation* помогла модели лучше различать классы за счет более гибкой трансформации входных признаков. Это особенно полезно в сложных задачах, где фиксированные функции активации могут быть недостаточно выразительными.

Несмотря на то, что увеличение точности относительно небольшое, это подтверждает гипотезу о том, что адаптивные функции активации могут играть важную роль в улучшении обучения нейросетей.

Выводы

Проведенное исследование подтвердило, что использование адаптивной функции активации в нейронной сети позволяет улучшить точность классификации специальностей абитуриентов по сравнению с базовой моделью. Анализ результатов показал, что *AdaptiveActivation* способствует более гибкому преобразованию входных признаков, что позволяет модели эффективнее разделять классы.

Несмотря на относительно небольшое увеличение точности, данное улучшение подтверждает гипотезу о значимости адаптивных функций активации в машинном обучении. Это открывает перспективы дальнейших исследований, направленных на оптимизацию подобных функций, их комбинирование с другими методами обучения и применение в более сложных рекомендательных системах.

Кроме того, адаптация предложенного подхода для обработки больших объемов данных с использованием современных архитектур нейронных сетей может существенно повысить его практическую применимость. В дальнейшем исследования могут быть направлены на оптимизацию и адаптацию данного метода для работы с большими данными, а также на его применение в задачах прогнозирования и классификации.

Литература

1. Agarap, A. F. (2018). *Deep learning using Rectified Linear Units (ReLU)* [arXiv preprint arXiv:1803.08375]. <https://arxiv.org/abs/1803.08375>
2. Alyahyan, E., & Düşteğör, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education*, 17(3). <https://doi.org/10.1186/s41239-020-0177-7>
3. Apicella, A., Donnarumma, F., Isgro, F., & Prevete, R. (2021). A survey on modern trainable activation functions. *Neural Networks*, 138, 14–32. <https://doi.org/10.1016/j.neunet.2021.01.026>

4. Drachsler, H., Hummel, H. G. K., & Koper, R. (2008). Personal recommender systems for learners in lifelong learning networks: The requirements, techniques and model. *International Journal of Learning Technology*, 3(4), 404–423. <https://doi.org/10.1504/IJLT.2008.019376>
5. Even-Dar, E., Mannor, S., & Mansour, Y. (2006). Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 7, 1079–1105. (Original work published as arXiv preprint: [cs/0206020](https://arxiv.org/abs/cs/0206020))
6. Fong, S., & Biuk-Aghai, R. P. (2009). An automated university admission recommender system for secondary school students. In *Proceedings of the 6th International Conference on Information Technology and Applications (ICITA)* (pp. 37–42). Hanoi, Vietnam. Retrieved from <https://www.researchgate.net/publication/221154329>
<https://www.researchgate.net/publication/221154329>
[An Automated University Admission Recommender System for Secondary School Students](https://www.researchgate.net/publication/221154329)
7. Jabeen, H., & Baig, A. R. (2010). Review of classification using genetic programming. *International Journal of Engineering Science and Technology*, 2(2), 94–103. Retrieved from <https://www.researchgate.net/publication/275022403>
[Review of classification using genetic programming](https://www.researchgate.net/publication/275022403)
8. Lee, K., Yang, J., Lee, H., & Hwang, J. Y. (2022). *Stochastic adaptive activation function* [arXiv preprint arXiv:2210.11672]. <https://arxiv.org/abs/2210.11672>
9. Lin, X., Zhong, G., Chen, K., Li, Q., & Huang, K. (2021). Attention-Augmented Machine Memory. *Cognitive Computation*, 13(3), 751–760. <https://doi.org/10.1007/s12559-021-09854-5>
10. Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (1st ed., pp. 73–105). Springer. https://doi.org/10.1007/978-0-387-85820-3_3
11. Ma, C., Wu, J., Si, C., & Tan, K. C. (2024). *Scaling supervised local learning with augmented auxiliary networks* [arXiv preprint arXiv:2402.17318]. <https://arxiv.org/abs/2402.17318>
12. Molina, M., & Blasco, G. (2003). A multi-agent system for emergency decision support. In *Lecture Notes in Computer Science* (Vol. 2669, pp. 43–51). https://doi.org/10.1007/978-3-540-45080-1_6
13. Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
14. Sakho, A., Malherbe, E., & Scornet, E. (2024). *Do we need rebalancing strategies? A theoretical and empirical study around SMOTE and its variants* [arXiv preprint arXiv:2402.03819]. <https://arxiv.org/abs/2402.03819>
15. Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1), 1–47. <https://doi.org/10.1145/505282.505283>
16. Miftahul, J. M., Sumi, K., & Mohammad, S. A. (2020). A content-based recommender system for choosing universities. *Turkish Journal of Electrical Engineering & Computer Sciences*, 28(4), 2128–2142. <https://doi.org/10.3906/elk-1911-37>
17. Misra, D. (2020). *Mish: A self regularized non-monotonic activation function* [arXiv preprint arXiv:1908.08681]. <https://arxiv.org/abs/1908.08681>

Болотбек уулу Н., Верзунов С. Н., Салиев А. Б., Мусина И. Р. Интеллектуальный анализ классификации абитуриентов с помощью нейронных сетей на основе адаптивной функции активации. Выбор специальности при поступлении в вуз является важнейшим этапом в академической жизни абитуриента, так как этот выбор влияет на дальнейшее развитие карьеры и профессиональную реализацию. Однако в условиях ограниченной информации, с которой абитуриенты сталкиваются на этапе выбора, существует риск принятия необоснованных решений, что может привести к неудовлетворенности учебным процессом, снижению академической успеваемости и необходимости смены направления обучения. В данной работе рассматривается разработка системы интеллектуальных рекомендаций на основе нейросетевого подхода, где ключевым элементом является адаптивная функция активации. В отличие от стандартных функций активации, она динамически изменяет форму отклика нейронов, что повышает гибкость модели при обучении и улучшает точность предсказаний. В исследовании проводится сравнительный анализ базовой модели и модели с адаптивной активацией, демонстрирующий повышение точности классификации специальностей. Полученные результаты подтверждают эффективность предлагаемого подхода для поддержки абитуриентов в принятии решений.

Ключевые слова: машинное обучение, глубокое обучение, техника увеличения объема данных, системы рекомендации, поступление в вуз, адаптивная активация

Bolotbek uulu N., Verzunov S. N., Saliev A. B., Musina I. R. Intelligent analysis of applicant classification using neural networks based on an adaptive activation function. *The process of selecting a specialization during university admission represents a critical decision in an applicant's academic trajectory, directly influencing future career prospects and professional development. However, due to the constraints of limited information available at the time of selection, applicants are susceptible to suboptimal decision-making, which may result in academic dissatisfaction, diminished performance and potential changes in their field of study. In recent years, machine learning methodologies have been extensively investigated for their potential in developing intelligent recommendation systems capable of integrating a broad spectrum of factors to enhance decision-making personalization. This study presents the development of an advanced recommendation system leveraging a neural network-based framework, wherein the core innovation lies in the implementation of an adaptive activation function. Unlike conventional activation functions, this approach dynamically adjusts the response characteristics of neurons, thereby augmenting model flexibility during training and improving predictive accuracy. A comparative evaluation between a baseline model and the proposed adaptive activation model is conducted, demonstrating a statistically significant improvement in classification accuracy for specialization recommendation. The empirical results substantiate the efficacy of the proposed approach in optimizing applicant decision support within the university admission process.*

Keywords: *machine learning, deep learning, data augmentation, recommendation systems, university admission, adaptive activation*

*Статья поступила в редакцию 18.05.2025
Рекомендована к публикации профессором Павлышом В. Н.*