На правах рукописи

УДК 004.891.2 (004.89)

КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ им. И. РАЗЗАКОВА

Болотбек уулу Нурсултан

РАЗРАБОТКА НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ НА ОСНОВЕ АДАПТИВНЫХ ФУНКЦИЙ АКТИВАЦИИ И МНОГОВЕТВЕВЫХ УПРАВЛЯЕМЫХ НЕЙРОНОВ

710100 - Компьютерные и информационные технологии

ДИССЕРТАЦИЯ

на соискание ученой степени доктора (PhD)

Научные Руководители:

Доктор технических наук

Верзунов Сергей Николаевич

Dr. Professor

Марина Тропман-Фрик

Бишкек 2026

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

AI (Artificial Intelligence) – искусственный интеллект.

ANN (Artificial Neural Network) – искусственная нейронная сеть.

LLM (Large Language Model) – крупная языковая модель.

RL (Reinforcement Learning) – обучение с подкреплением.

MBGN (Multi-Branch Gated Neuron) – Нейрон с несколькими ветвями и управляющими (гейтовыми) механизмами.

SMOTE (Synthetic Minority Over-sampling Technique) – метод синтетического увеличения выборки для борьбы с дисбалансом классов.

RF (Random Forest) – метод случайного леса.

XGBoost (Extreme Gradient Boosting) – алгоритм градиентного бустинга.

LightGBM (Light Gradient Boosting Machine) – алгоритм градиентного бустинга, оптимизированный по скорости и памяти.

CatBoost – алгоритм градиентного бустинга, ориентированный на категориальные данные.

LOF (Local Outlier Factor) – алгоритм выявления локальных выбросов.

API (Application Programming Interface) – программный интерфейс приложения.

SAAF (Stochastic Adaptive Activation Function) – стохастическая адаптивная функция активации.

LMS (Learning Management System) – электронные образовательные системы

GPA (Grade Point Average) – усредненный балл оценок студента

Stacking – метод объединения моделей машинного обучения.

MAB (Multi-Armed Bandit) – вид задач метода обучения с подкреплением, заключающиеся в выборе эффективного варианта с максимальной наградой.

GP (Genetic Programming) – алгоритм машинного обучения, который использует методологию, вдохновленную «эволюцией».

MLP (Multi-Layer Perceptron) – (Многослойная полносвязная нейронная сеть)

AUC-ROC (Area Under the Receiver Operating Characteristic Curve) **-** метрика для оценки качества бинарных классификаторов

Baseline model – стартовая модель определяющая минимально приемлемый уровень качества следующих моделей.

ИНС – Искусственные нейросети

ВВЕДЕНИЕ

Актуальность работы. Актуальность, рассматриваемой проблемы, определяется потребностью в объективном и эффективном отборе абитуриентов в условиях постоянно растущего объема данных и высокой конкуренции среди поступающих. В традиционных методах к отбору используются фиксированные правила и простые статистические модели, которые не способны в полной мере отразить многофакторную природу образовательного процесса и выявить скрытые зависимости между результатами вступительных испытаний, социальными условиями и потенциальной академической успешностью студентов.

В этих условиях применение методов искусственного интеллекта (ИИ) и машинного обучения (МО) представляется наиболее перспективным направлением развития систем отбора абитуриентов. Использование интеллектуальных алгоритмов позволяет выявлять сложные нелинейные зависимости в данных, повышать точность прогнозирования и обеспечивать большую адаптивность к изменяющимся условиям образовательной среды. В сравнении с классическими моделями, интеллектуальные системы могут обучаться на основе собранных данных, что является большим преимуществом и делает их гибкими для разного рода входной информации.

Диссертационная работа выполнена в русле приоритетных направлений развития науки и образования Кыргызской Республики, связанных с цифровизацией образовательных процессов и внедрением интеллектуальных систем анализа больших данных.

Проведенные исследования являются продолжением проектов, направленных на применение нейронных сетей и интеллектуальных алгоритмов для задач классификации, прогнозирования и поддержки принятия решений в социальных и образовательных системах. В ходе диссертационной работы, полученные результаты, апробированы в рамках публикаций в международных конференциях и журналах, индексируемых в Scopus.

Объектом исследования является приемная кампания университета, которая представляет из себя совокупность информационных и управленческих процессов, обеспечивающая взаимодействие абитуриентов и университета. Приемная кампания выступает в качестве сложной организационной системы, характеризующаяся на множественных факторах принятия решений в условиях неопределенности и ограниченности времени, а также значительного информационного потока.

Предметом исследования является система поддержки принятия решений (СППР), рассматриваемая как инструмент, обеспечивающий повышение эффективности анализа информации, а также минимизацию случаев «неправильно выбранной профессии». Традиционные методы приема абитуриентов не позволяют учитывать накопленный опыт, тогда как интеллектуальные системы используют данные предыдущих лет для корректировки своих решений. Такой подход способствует более глубокому анализу информационных зависимостей и улучшению качества рекомендаций.

Цели и задачи исследования. Цель работы заключается в разработке и экспериментальном обосновании новых методов распределения абитуриентов по специальностям на основе нейросетевых технологий – адаптивных функций активаций и многоветвевых архитектур ИНС.

Для достижения цели решались следующие задачи:

1. Провести анализ существующих методов классификации в образовании и выявить их ограничения.
2. Провести анализ исходных данных и выявить взаимосвязь между характеристиками абитуриентов и специальностями
3. Разработать и исследовать адаптивную функцию активации (*Adaptive Activation Function*) для повышения точности и обобщающей способности нейросетей.
4. Разработать архитектуру *Multi-Branch Gated Neuron*, позволяющую учитывать сложные взаимосвязи между результатами ОРТ и выбранной специальностью.

5. Провести экспериментальные количественные исследования, разработанных архитектур, на реальных и синтетических данных, а также сравнить ее с классическими архитектурами.
6. Разработать ПО для практического применения предложенных архитектур в автоматизированных системах управления приемной кампанией.

Научная новизна работы:

1. Выявлены новые закономерности связывающие характеристики студентов и специальности и форму обучения на которые они поступили
2. Выявлены ключевые особенности применения адаптивных функций активации и многоветвевых нейронных архитектур в контексте анализа образовательных данных. Это открывает новые возможности для разработки более точных предиктивных моделей.
3. Разработана и эмпирически проверена функция активации, которая представляет собой адаптивную комбинацию нескольких базовых функций. В отличие от традиционных, неизменных активационных механизмов, предложенный подход способен самостоятельно подстраиваться под структуру обучающих данных. Это улучшает устойчивость модели к изменчивости выборок и способствует повышению ее точности. Экспериментальные исследования продемонстрировали рост обобщающих способностей нейросетей при использовании данной функции по сравнению с существующими решениями.
4. Предложена оригинальная архитектура *Multi-Branch Gated Neuron*, позволяющая гибко перераспределять информационные потоки между различными каналами вычислений. Такой подход обеспечивает возможность учитывать сложные нелинейные взаимосвязи между признаками.

Практическая значимость. Разработанные методы могут быть использованы в интеллектуальных системах поддержки принятия решений при приёмной кампании. Их внедрение позволит повысить прозрачность и объективность отбора, снизить влияние субъективных факторов и обеспечить более рациональное распределение абитуриентов по образовательным программам. Кроме того, предложенные архитектуры могут быть адаптированы для решения широкого круга задач в других областях, где требуется классификация данных с высокой вариативностью.

На основе полученных моделей реализована информационная система, которая на основе вводимых данных (баллов ОРТ) формирует рекомендации по поступлению на различные специальности университета. Система отображает процентное распределение вероятности зачисления на каждое направление, а также оценивает вероятность обучения на бюджетной основе. Это обеспечивает практическую ценность исследования, позволяя абитуриентам принимать более обоснованные решения, а образовательным учреждениям – повышать прозрачность и объективность приёмной кампании.

Предложенная архитектура нейронных сетей и полученные результаты внедрены в учебный процесс при изучении дисциплин, связанных с искусственным интеллектом и анализом данных в Кыргызском Государственном Техническом Университете им. И. Раззакова и Международного Университета в Центральной Азии.

Основные положения, выносимые на защиту.

1. Метод, обоснованной классификации абитуриентов, на основе адаптивной функции активации, позволяющий представлять подход, в котором для повышения точности распределения используется нейронная сеть с настраиваемой функцией активации, адаптирующейся под особенности входных данных.
2. Архитектура *Multi-Branch Gated Neuron* для выявления сложных взаимосвязей в образовательных данных. Эта архитектура представляет собой ИНС, где несколько параллельных ветвей с управляющими (*gating*) механизмами позволяют выделять и комбинировать различные типы признаков, что способствует выявлению сложных взаимосвязей в образовательных данных.

3. Экспериментальные результаты, подтверждающие эффективность предложенных подходов, в задачах прогнозирования и классификации. Применение разработанной адаптивной функции активации позволило повысить точность до 11.89%, что демонстрирует относительное улучшение результатов более, чем на 10% по сравнению с базовой моделью (*ReLU*). Применение архитектуры *Multi-Branch Gated Neuron* (*MBGN*) позволило повысить точность до 64.6%, что демонстрирует значительное улучшение относительно результатов, полученных классическими методами машинного обучения (62.9%).
4. Методика апробации и валидации разработанных моделей на реальных данных ОРТ, обеспечивающая их практическую применимость и достоверность получаемых результатов.

Личный вклад соискателя. Все ключевые результаты, включая разработку адаптивной функции активации, проектирование архитектуры Multi-Branch Gated Neuron, а также проведение экспериментов и анализ результатов, получены автором лично и под руководством научного руководителя.

В статье Bolotbek uulu, N., Verzunov, S.N., Tropmann-Frick, M., Saliev, A.B., Musina, I.R., Homenko, T.N. (2025). Intelligent Analysis of Applicant Classification Using Neural Networks Based on Adaptive Activation Function. In: Software Engineering: Emerging Trends and Practices in System Development. CSOC 2025. Lecture Notes in Networks and Systems, vol 1559. Springer, Cham. автору (**Болотбек уулу Н.**) принадлежат разработка концепции исследования, постановка задачи классификации абитуриентов, создание и реализация архитектуры нейронной сети с адаптивной функцией активации, проведение вычислительных экспериментов и анализ результатов. **Верзунову С. Н.** принадлежит научное руководство, формулировка целей исследования и корректировка методологии. **Тропманн-Фрик М.** – консультирование по вопросам архитектуры моделей и их интерпретации. **Салиеву А. Б.** – участие в обсуждении и анализе результатов. **Мусиной И. Р.** – постановка прикладной задачи классификации и предварительная подготовка образовательных данных. **Хоменко Т. Н.** – редактирование текста и оформление рукописи.

В статье Bolotbek uulu, N. et al. (2025). Machine Learning in Education: A Literature Review of Predictive Models and Recommender Systems. In: Software Engineering: Emerging Trends and Practices in System Development. CSOC 2025. Lecture Notes in Networks and Systems, vol 1563. Springer, Cham. автором (**Болотбек уулу Н.**) выполнены разработка протокола обзора (поисковые запросы, критерии включения/исключения), систематизация и агрегация источников, построение таксономии моделей (предиктивная аналитика и рекомендательные системы), написание основного текста и формулировка выводов. **Верзунову С. Н.** принадлежит научное руководство, корректировка целей и методологии исследования, критическое редактирование рукописи. **Мусиной И. Р.** – отбор и описание прикладных образовательных кейсов, валидация интерпретаций и проверка полноты выборки. **Мусабаеву Е. Б.** – библиометрический анализ корпуса публикаций, подготовка сводных таблиц и сравнительных матриц. **Ашымовой А. Ж.** – разметка и кодирование публикаций, data curation и контроль качества, согласование меж экспертных оценок. **Сабаевой К. К.** – статистическая обработка результатов обзора и визуализация (диаграммы, графы), оформление библиографии. **Турдалиеву Е. К.** – сбор и нормализация сведений по рекомендательным системам, подготовка приложений и дополнений (таблиц, перечней датасетов и метрик). **Ордобаеву А. Б.** – языковая вычитка и техническое редактирование, проверка соответствия требованиям серии LNNS (Springer), финальное оформление иллюстраций и таблиц.

Апробация результатов. Результаты диссертационной работы были представлены на международных конференциях и опубликованы в материалах конференции, индексируемых в базе данных Scopus. В частности, по теме диссертации опубликованы статьи:

1. Bolotbek uulu, N., Verzunov, S.N., Tropmann-Frick, M., Saliev, A.B., Musina, I.R., Homenko, T.N. (2025). Intelligent Analysis of Applicant Classification

Using Neural Networks Based on Adaptive Activation Function. In: Silhavy, R., Silhavy, P. (eds) Software Engineering: Emerging Trends and Practices in System Development. CSOC 2025. Lecture Notes in Networks and Systems, vol 1559. Springer, Cham. https://doi.org/10.1007/978-3-032-00712-4_29

2. Bolotbek uulu, N. *et al.* (2025). Machine Learning in Education: A Literature Review of Predictive Models and Recommender Systems. In: Silhavy, R., Silhavy, P. (eds) Software Engineering: Emerging Trends and Practices in System Development. CSOC 2025. Lecture Notes in Networks and Systems, vol 1563. Springer, Cham. https://doi.org/10.1007/978-3-032-00715-5_30

Основные научные результаты опубликованы в 2 печатных работах в SCOPUS. Дополнительно подготовлены материалы для участия в научных конференциях по проблемам применения искусственного интеллекта в образовании.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений. Общий объем диссертации составил 109 страниц компьютерного набора, содержит 31 рисунок и 2 таблицы.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

Современные научные исследования в области образовательных технологий и машинного обучения охватывают разнообразные направления от разработки новых активационных функций в нейросетях до прогностического анализа академической успеваемости студентов. Кроме того, активно исследуются рекомендательные механизмы, адаптированные к нуждам абитуриентов и образовательных учреждений, а также эффективные методы обработки данных, связанных с учебным процессом. Такие исследования способствуют не только повышению качества обучения, но и обеспечивают более точную персонализацию образовательных траекторий, что в конечном итоге положительно сказывается на результатах студентов.

Настоящий обзор литературы обобщает научные труды, посвященные анализу различных аспектов образовательных ИИ-систем, с целью выявления пересечений между подходами, их достоинств и ограничений, а также определения наиболее эффективных стратегий для решения задач академического планирования.

Основные акценты обзора сделаны на следующих направлениях:

- роль активационных функций в нейросетевых моделях и их влияние на точность прогнозов;
- алгоритмы оценки успеваемости на основе результатов вступительных испытаний и другой учебной информации;
- разработка и внедрение рекомендательных механизмов в образовательной среде;
- применение глубокого обучения и методов интеллектуального анализа для работы с академическими данными.

В рамках анализа рассматриваются и существующие подходы к классификации абитуриентов по направлениям подготовки. Отмечается, что одним из ключевых факторов повышения точности моделей является учет не только итоговых баллов, но и характера их распределения.

Таким образом, данный литературный обзор представляет собой попытку систематизировать современные разработки на стыке образования и ИИ, выявить актуальные тренды и предложить возможные направления дальнейших исследований.

1.1 Машинное обучение и активационные функции

Обучение нейросетей во многом определяется выбором функции активации, которая напрямую влияет на способность модели обучаться и обобщать. Несмотря на распространённость *ReLU*, известны её ограничения – например, «замирание» нейронов при нулевом градиенте. Современные исследования акцентируют внимание на гибридных и адаптивных функциях, которые способны динамически изменять свою форму в зависимости от входных данных. Это открывает путь к созданию более устойчивых архитектур и, как следствие, к повышению точности решений в задачах образовательной аналитики.

В публикации **Agarap, A. F. (2018). «Deep learning using rectified linear units (ReLU)»** рассматривается роль функции активации *ReLU* в обучении глубоких нейронных сетей. Автор подробно описывает, почему эта функция стала одной из наиболее распространённых: она проста в вычислении, помогает уменьшить эффект исчезающих градиентов и показывает высокую эффективность при работе с крупными архитектурами. Вместе с тем подчеркивается и ее ограничения – например, ситуация, при которой часть нейронов перестает реагировать на входные данные, так называемый эффект «мертвых нейронов».

В работе также проводится сравнительный анализ *ReLU* и других популярных функций активации, таких как сигмоидальная функция, tanh и *Softmax*. Исследование демонстрирует, что *ReLU* заметно влияет на динамику обучения моделей: по сравнению с сигмоидами обучение с градиентным спуском происходит быстрее, а слои сети формируют более разреженные представления. Однако автор обращает внимание на то, что риск полной деактивации некоторых нейронов делает необходимым использование дополнительных архитектурных решений.

Функция ReLU задаётся выражением $F(x) = max(0, x)$. Среди её ключевых свойств выделяются следующие особенности:

- отсутствие верхнего предела выходных значений, что снижает вероятность исчезновения градиента;
- простота вычислений, ускоряющая процесс обучения;
- способность порождать разреженные активации, уменьшая избыточную нагрузку на модель.

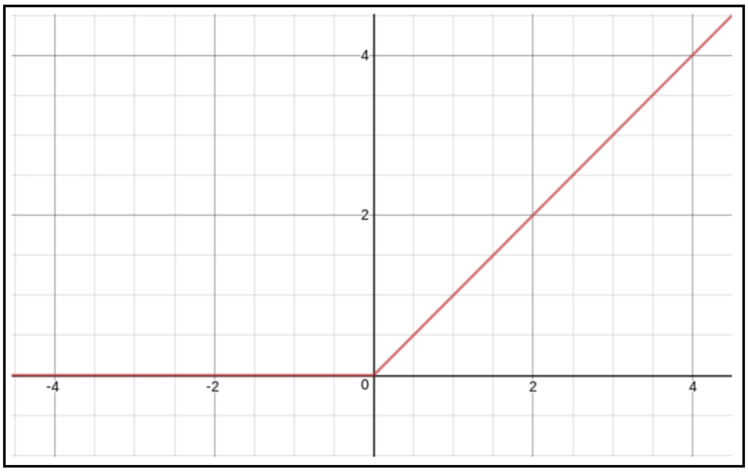

Рисунок 1.1 – Функция активации Rectified Linear Unit (ReLU).

График *ReLU* приведен на рисунке 1.1: при отрицательных входных значениях функция возвращает ноль, а при положительных – линейно возрастает.

Исследование рассматривает стандартные датасеты:

- *MNIST* – набор изображений рукописных цифр, традиционно используемый для тестирования моделей машинного обучения.
- *Fashion-MNIST* – усложненный аналог *MNIST*, содержащий изображения одежды, что делает задачу классификации более сложной.
- Wisconsin Diagnostic Breast Cancer (*WDBC*) – медицинский датасет, применяемый для диагностики злокачественных и доброкачественных опухолей.

Для построения моделей использовались современные инструменты машинного обучения, такие как *Keras*, *TensorFlow*, *NumPy* и *scikit-learn*. Автор использовал алгоритм *Adam* для оптимизации нейросетей, что позволило добиться стабильной сходимости моделей.

В работе тестировались два типа архитектур:

- Полносвязная нейронная сеть (*FFNN*) – с использованием *ReLU* и *Softmax* на выходном слое.
- Сверточная нейронная сеть (*CNN*) – с несколькими сверточными и пулинговыми слоями, аналогичная архитектуре *VGG*.

Эксперименты показывают, что *FFNN* с *ReLU* демонстрирует схожую точность с моделями на *Softmax*, однако *CNN* с *ReLU* уступает *CNN* с *Softmax* из-за медленной сходимости. Стоит учитывать, что Softmax обеспечивает более плавную передачу градиентов на выходном слое, что делает его более подходящим для сверточных сетей.

Сравнение точности моделей на *MNIST* представлено на рис. 1.2 – 1.5, а на *Fashion-MNIST* – на рис. 1.6 – 1.9

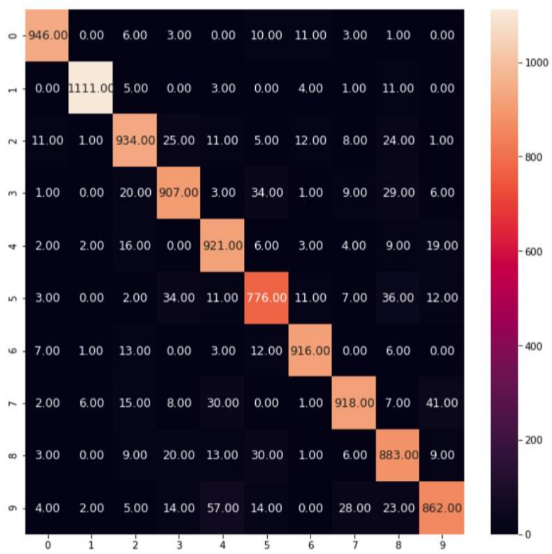

Рисунок 1.4 – Матрица ошибок для CNN-ReLU на классификации MNIST. Источник: Agarap A.F. (2018)

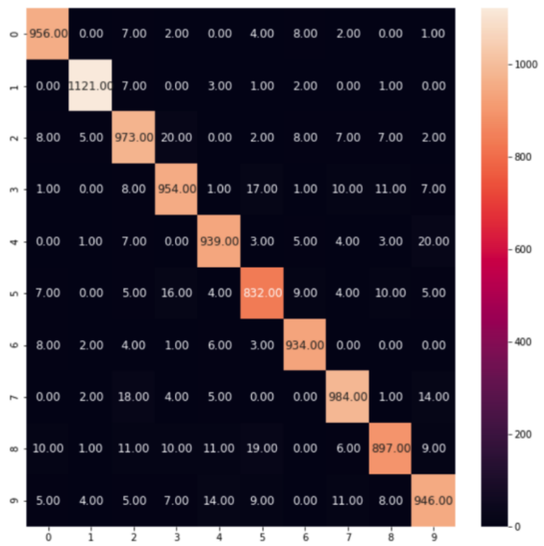

Рисунок 1.5 – Матрица ошибок для CNN-Softmax на классификации MNIST. Источник: Agarap A.F. (2018)

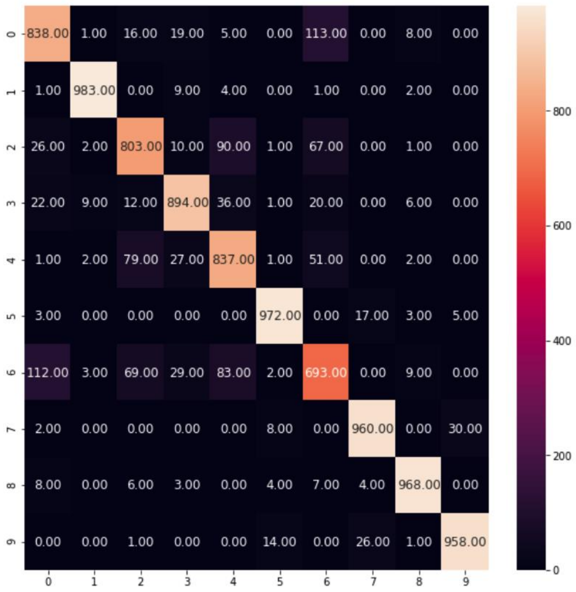

Рисунок 1.6 – Матрица ошибок для FFNN-ReLU на классификации Fashion-MNIST. Источник: Agarap A.F. (2018)

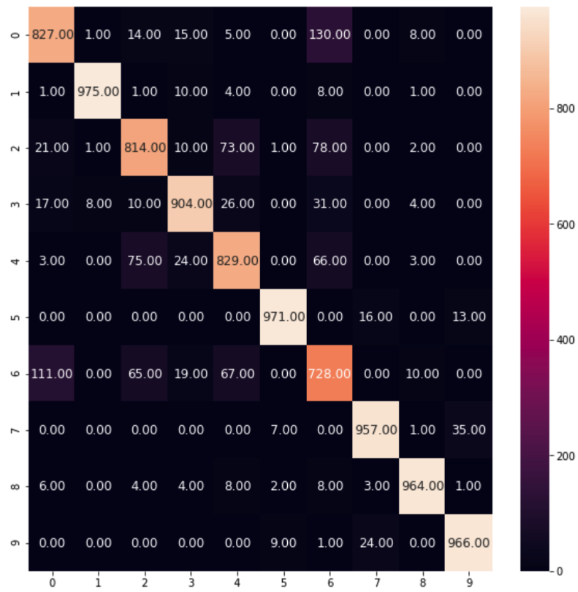

Рисунок 1.7 – Матрица ошибок для FFNN-Softmax на классификации Fashion-MNIST. Источник: Agarap A.F. (2018)

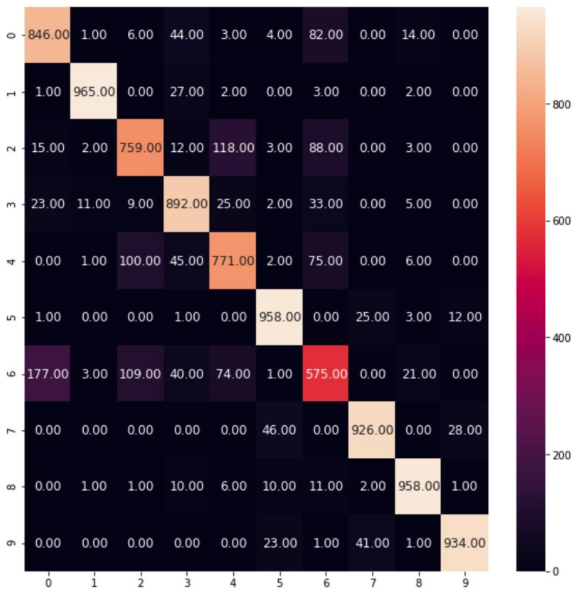

Рисунок 1.8 – Матрица ошибок для *CNN-ReLU* на классификации Fashion-MNIST. Источник: Agarap A.F. (2018)

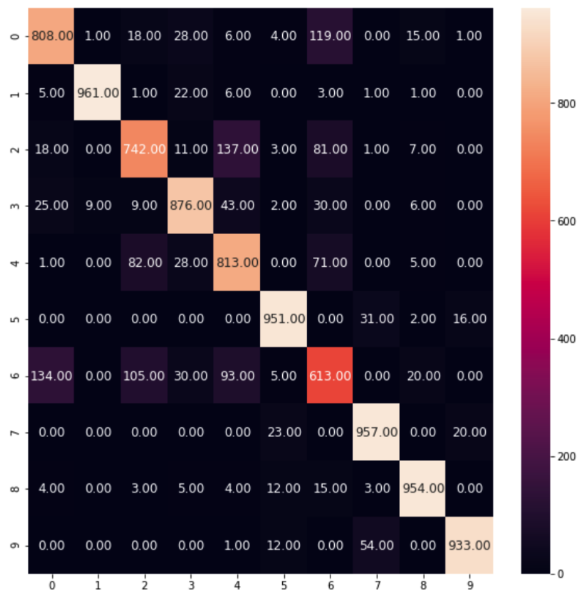

Рисунок 1.9 – Матрица ошибок для *CNN-Softmax* на классификации Fashion-MNIST. Источник: Agarap A.F. (2018)

Результаты тестирования на *MNIST*	Результаты тестирования на Fashion-MNIST
● FFNN-Softmax: 97,98%	● FFNN-Softmax: 89,35%
● FFNN-ReLU: 97,77%	● FFNN-ReLU: 89,06%
● CNN-Softmax: 95,36%	● CNN-Softmax: 86,08%
● CNN-ReLU: 91,74%	● CNN-ReLU: 85,84%

На основе анализа результатов автор приходит к выводу, что ReLU сохраняет высокую эффективность в полносвязных нейросетях, однако в сверточных архитектурах лучшие показатели демонстрирует функция активации Softmax. Связано это с тем, что данная функция более равномерно распределяет выходные активации, а это положительно влияет на поведение градиентного спуска и минимизирует риск «застывания» нейронов.

Выводы и практическая значимость

Проведенный анализ подтверждает важность выбора функции активации при обучении нейросетевых моделей. Особое внимание в статье уделяется функции *ReLU*, которая благодаря своей простоте и способности эффективно устранять проблему исчезающих градиентов, получила широкое распространение в глубоких архитектурах. Вместе с тем автор указывает на определённые ограничения *ReLU*, в частности – риск формирования неактивных (так называемых «мертвых») нейронов, что может негативно сказываться на качестве обучения и требует применения дополнительных механизмов.

Основные выводы исследования:

- Функция *ReLU* показывает хорошую эффективность в полносвязных нейросетях. Однако в сверточных архитектурах часто лучшую результативность демонстрирует Softmax, обеспечивающий более равномерное распределение выходов.
- При решении задач классификации изображений *ReLU* способствует созданию разреженных представлений, что способствует улучшению обобщающей способности модели.

- Улучшить работу *ReLU* можно за счёт использования её модификаций, таких как *Leaky ReLU* или *Parametric ReLU*, которые позволяют избежать полного «отключения» нейронов.
- Перспективным направлением исследований является изучение *ReLU* в контексте сложных архитектур – например, в трансформерах или рекуррентных сетях, – что поможет выявить условия ее оптимального применения.
- Введение регуляризующих методов, таких как *Dropout* и *Batch Normalization*, способствует компенсации слабых сторон *ReLU* и повышает общую устойчивость модели при обучении.

1.2 Преимущества адаптивных функций активации

Классические функции активации, такие как *ReLU, Sigmoid* и *Tanh* характеризуются неизменными параметрами, что ограничивает их возможности при моделировании сложных и нелинейных зависимостей. В отличие от них, адаптивные функции активации обладают настраиваемыми параметрами, которые изменяются в процессе обучения. Это делает их значительно более гибкими и позволяет нейросетям лучше адаптироваться к особенностям конкретных данных. Такая гибкость способствует улучшению способности модели к обобщению и снижает риск переобучения.

Apicella, A., Donnarumma, F., Isgrò, F., & Prevete, R. (2021). A survey on modern trainable activation functions.

Основные аспекты исследования

В статье рассматриваются различные виды адаптивных функций активации:

- *Swish* – предложенная Google функция, определяемая как $f(x) = x \cdot \sigma(\beta x)$, где $\sigma(x)$ – сигмовидная функция. Она показывает более высокую производительность по сравнению с *ReLU*, особенно в глубоких сетях.
- *PReLU (Parametric ReLU)* – модификация *ReLU*, в которой коэффициент наклона для отрицательных значений входа обучается, а не фиксируется.

- *APL (Adaptive Piecewise Linear function)* – использует несколько линейных сегментов для лучшей аппроксимации нелинейных зависимостей.
- *Maxout* – метод, в котором функция активации выбирает максимум среди нескольких линейных преобразований.

1.2.1 Влияние функций активации на обучение нейросетей

Авторы подчеркивают, что выбор функции активации оказывает заметное влияние на эффективность обучения нейросетей. В частности, он влияет на:

- Скорость обучения: адаптивные функции *Swish* и *PReLU*, позволяют модели быстрее обучаться, так как они обеспечивают лучшую передачу градиентов.
- Устойчивость градиентов: традиционные функции, такие как *Sigmoid* и *Tanh*, подвержены проблеме исчезающего градиента, эта особенность проявляется при обучении глубоких моделей, когда значения вектора частных производных становятся слишком малыми для корректного обновления весов. Адаптивные функции помогают решать эту проблему, улучшая передаваемый сигнал в процессе обратного распространения ошибки.
- Способность обобщать: адаптивные функции оказывают заметное влияние на способность нейросетей к обобщению за счет гибкого подстраивания под структуру данных.
- Экспериментальные результаты, представленные в статье, подтверждают это наблюдение: архитектуры, где используются адаптивные функции активации, стабильно превосходят модели с классическими активациями. Улучшения фиксируются как в задачах компьютерного зрения, так и в обработке естественного языка, что говорит о широком потенциале подобных подходов.

Теоретический анализ функций активации

Авторы обсуждают математические свойства различных функций активаций:

- Дифференцируемость: основной фактор для работы градиентных алгоритмов для нахождения оптимальной точки.
- Гладкость функции: *Swish*, в сравнение с *ReLU*, является гладкой функцией, что способствует более стабильному процессу оптимизации.
- Адаптивность: функции активации наподобие *PReLU* обладают настраиваемыми параметрами, которые обучаются совместно с остальными весами модели.

Результаты экспериментов

В во время проведения экспериментов авторы оценивают как различные функции активации влияют на:

- Качество итоговой классификации на наборах данных *CIFAR-10, ImageNet* и др.
- Производительность моделей – оценивается *accuracy* (доля верных ответов), скорость обучения и устойчивость к переобучению.
- Совместимость *ResNet, EfficientNet* и *Transformer* архитектурами.

Выводы и практическая значимость. Статья Apicella и соавторов (2021) показывает значимость корректного выбора активационной функции при обучении нейросетей. Адаптивные функции активации обладают рядом преимуществ: они способствуют повышению точности моделей, ускорению процесса сходимости, а также помогают снизить риск переобучения. Полученные результаты подтверждают перспективность данного направления для дальнейшего развития технологий глубокого обучения.

Lee, K., Yang, J., Lee, H., & Hwang, J. Y. (2022). Stochastic adaptive activation function

Статья Lee, K., Yang, J., Lee, H., & Hwang, J. Y. (2022). «Stochastic adaptive activation function» представляет собой исследование новой стохастической адаптивной функции активации, разработанной для повышения эффективности глубоких нейронных сетей. В отличие от традиционных функций активации, таких как *ReLU, Swish, PReLU*, предложенный метод динамически изменяет свои параметры в зависимости от входных данных, что делает модель более гибкой и способной адаптироваться к сложным распределениям данных.

Основные идеи и новизна исследования

1. **Стохастическая адаптивная активационная функция (SAAF)**

В представленной работе авторы вводят новую функцию активации, обладающую настраиваемыми параметрами и элементом стохастичности. Такой подход позволяет модели нейросети адаптировать характер активации в зависимости от входных данных. *SAAF*, в сравнении с традиционными активационными функциями, такие как *ReLU* или *Swish*, динамически изменяет свои параметры, что делает ее эффективной при работе с высокоразмерными данными и нерегулярными распределениями признаков.

Ключевые черты SAAF:

- Адаптация во время обучения: она изменяет параметры в процессе, что помогает учитывать разнообразие входных данных;
- Стохастичность: улучшает способность обобщать и помогает избежать застревания в локальных минимумах;
- Универсальность: может использоваться в задачах обработки изображений, анализа текста, прогнозирования на основе временных рядов и в моделях с высокоразмерными пространствами.

Преимущества SAAF по сравнению с классическими функциями

- Повышенная точность предсказаний: за счет адаптивных параметров достигаются лучшие результаты в задачах машинного обучения;
- Ускоренное обучение: стохастическая компонента способствует улучшенной передаче градиента, сокращая число эпох до сходимости;
- Снижение риска переобучения: случайные колебания во время обучения препятствуют подгонке под тренировочные данные;
- Лучшее распространение градиента: глубокие сети с использованием *SAAF* менее чувствительны к затуханию градиента.

2. **Влияние на процессы оптимизации**

Авторы статьи отмечают, что внедрение *SAAF* требует более тщательного подбора гиперпараметров. Присутствие стохастического компонента может вызвать нестабильность при градиентном спуске, особенно при неправильной конфигурации параметров.

Рекомендации по использованию:

- Контроль уровня случайности: адаптация параметров должна учитывать специфику решаемой задачи;
- Тонкая настройка регуляризации: важно сбалансировать влияние шумовых факторов и стабильности;
- Гибридный подход: возможно использование SAAF совместно с классическими функциями активации в различных слоях сети.

3. **Экспериментальные результаты**

Функция SAAF была протестирована на нескольких популярных наборах данных и в различных типах моделей:

- CIFAR-10 / CIFAR-100: демонстрирует устойчивый прирост точности в задачах распознавания изображений;
- ImageNet: в глубоких CNN архитектурах достигается высокая точность при сохранении сопоставимой вычислительной сложности;
- Задачи обработки естественного языка (NLP): улучшает качество работы с длинными текстами, повышает устойчивость к потере информации;
- Прогнозирование по временным рядам: показывает лучшую адаптацию к сложным и нелинейным зависимостям во входных данных.

Сравнение с другими методами

- На архитектуре *ResNet-50 SAAF* позволила уменьшить ошибки классификации на 3–5% по сравнению с *ReLU*.
- На *LSTM*-моделях для анализа последовательных данных стохастическая адаптация улучшила качество предсказаний на 6–8%.

При генерации изображений (*GANs*) *SAAF* позволила моделям достигнуть более высокой стабильности при обучении.

Выводы и практическое применение

Исследование Lee et al. (2022) показывает, что стохастическая адаптивная функция активации является перспективным направлением в развитии глубокого обучения. Она улучшает качество предсказаний, ускоряет процесс обучения и делает модели более устойчивыми к переобучению.

Образовательные технологии и анализ данных

Alyahyan, E., Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices

Статья Alyahyan, E., Düştegör, D. (2020). «Predicting academic success in higher education: literature review and best practices» представляет собой обзор существующих методов предсказания академической успеваемости студентов, а также анализ лучших практик в этой области. Авторы исследуют, какие факторы влияют на академический успех, какие алгоритмы машинного обучения наиболее эффективны для прогнозирования и какие данные позволяют улучшить точность предсказаний.

Основные аспекты исследования

1. Факторы, которые влияют на академическую успеваемость

Авторы статьи отмечают ключевые факторы, оказывающие влияние на успешность студентов в вузах:

- Исторические академические сведения – общий средний балл обучения в школе, результаты вступительных тестов, а также предыдущий образовательный опыт учащегося.
- Социально-экономическое положение – семейный доход, уровень образованности родителей, наличие дополнительной поддержки со стороны ближайшего образовательного окружения.
- Когнитивные и поведенческие особенности – регулярность посещения занятий, степень вовлеченности в образовательный процесс и устойчивость к стрессу или другим психологическим нагрузкам.
- Демографические характеристики – возрастные особенности, пол, этническая и культурная принадлежность, а также их связь с мотивацией к обучению.

- Цифровая активность обучающегося – поведение в онлайн-среде: участие в образовательных платформах (LMS), активность в онлайн-курсах, на форумах и других цифровых ресурсах, отражающих взаимодействие с образовательной средой.

Объединение вышеперечисленных признаков значительно повышает точность предсказательных моделей, особенно при применении алгоритмов машинного обучения. В работе рассматриваются следующие методы:

- **Линейная регрессия.** Простой и интерпретируемый алгоритм, позволяющий выявить линейные зависимости между признаками и академической успешностью. Однако ограничен в возможностях при наличии сложных, нелинейных структур.
- **Деревья решений.** Обеспечивают визуализацию иерархии факторов, хорошо справляются с категориальными данными. Однако подвержены переобучению, особенно при наличии шумов в данных.
- **Случайный лес (Random Forest).** Ансамблевый метод, объединяющий множество деревьев решений. Повышает точность модели и устойчив к переобучению. Дополнительно позволяет ранжировать значимость признаков.
- **Градиентный бустинг (XGBoost, LightGBM).** Более сложный и производительный вариант деревьев решений. Отличается высокой точностью, но требует тщательной настройки параметров и значительных вычислительных ресурсов.
- **Нейросетевые подходы.** Самые мощные инструменты для анализа сложных и многомерных данных. Хорошо работают с большими выборками и способны выявлять глубокие зависимости. Минус — низкая интерпретируемость, что может затруднять использование в контексте образовательной политики.

Авторы проводят серию экспериментов, демонстрируя, что модели машинного обучения заметно превосходят традиционные статистические методы в прогнозировании академической успешности:

- Точность линейной регрессии: 65–75%
- Точность деревьев решений: 70–80%
- Случайный лес: 75–85%
- Градиентный бустинг: 80–90%
- Нейросети: 85–95%

Наиболее высокую точность демонстрируют гибридные предсказательные модели, в которых совмещаются различные подходы и одновременно учитываются такие факторы, как предшествующие академические результаты, социально-экономические характеристики и цифровая активность студентов. Применение подобных моделей позволяет университетам оперативно выявлять студентов, находящихся в группе риска, и разрабатывать индивидуальные меры поддержки для повышения их успеваемости.

В исследовании, проведённом Alyahyan и Düştegör (2020), показано, что наилучшие результаты достигаются при использовании мощных моделей машинного обучения в сочетании с мультимодальными наборами данных. Авторы делают акцент на том, что наибольшую точность обеспечивают ансамблевые методы и нейросетевые архитектуры, использующие совокупность входных данных. Хотя средний балл (GPA) и итоги экзаменов остаются основными предикторами, значительное влияние также оказывает и социальная среда учащегося. Дополнительный учет активности в системах электронного обучения и онлайн-платформах существенно повышает точность прогностических алгоритмов.

На основе результатов анализа авторы рекомендуют внедрение систем предиктивной аналитики в образовательный процесс. Такие системы позволяют заранее идентифицировать учащихся, испытывающих сложности, и оказывать им своевременную поддержку. Это способствует снижению уровня отсева, улучшению образовательных результатов и более рациональному использованию академических ресурсов.

Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V. (2023). Predicting Learning Success: Research Problems and Challenges

Статья Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V. (2023). «Predicting Learning Success: Research Problems and Challenges» посвящена изучению сложностей, возникающих при прогнозировании академической успеваемости студентов. Авторы анализируют основные проблемы, с которыми сталкиваются исследователи в этой области, а также предлагают методы их решения с использованием современных подходов обработки данных и машинного обучения.

Неполнота данных. Одной из наиболее значимых трудностей в сфере образовательной аналитики является наличие пропусков в данных. В Университетских информационных системах нередко отсутствуют критически важные сведения о студентах, такие как их предыдущий учебный опыт, социально-финансовое положение, уровень учебной мотивации и степень вовлеченности в образовательные активности. Подобная неполнота данных возникает по разным причинам: технические ограничения систем хранения, организационные ошибки при вводе информации, а также нежелание студентов раскрывать персональные данные.

Авторы отмечают, что классические модели прогнозирования плохо справляются с подобными ситуациями, поскольку отсутствие ключевой информации может серьезно исказить результаты анализа и снизить точность предсказаний. В связи с этим предлагается несколько подходов для компенсации неполных данных:

- методы восстановления пропусков, включающие статистическое вменение недостающих значений или предиктивное заполнение пропусков с использованием алгоритмов машинного обучения;
- использование устойчивых к пропускам алгоритмов, например деревьев решений и ансамблевых методов, которые способны справляться с отсутствующими значениями без необходимости их предварительного заполнения;

- удаление студентов с неполными данными (но только в случаях, когда их доля невелика, чтобы избежать значительных потерь информации).

В этой связи важную роль играет цифровой образовательный профиль студента, содержащий информацию о его академической истории и активности в цифровой образовательной среде:

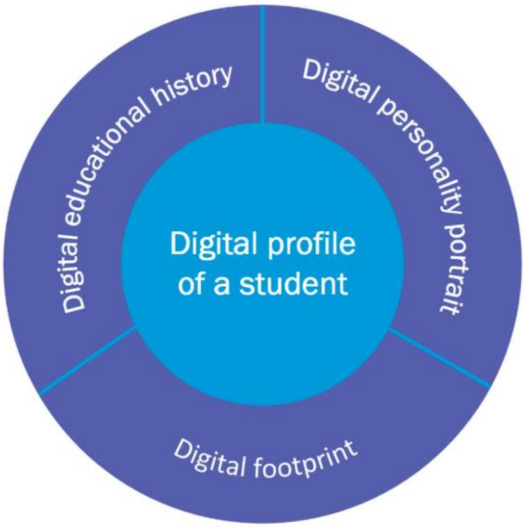

Рисунок 1.10 – Цифровой образовательный профиль студента. Источник: Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V. (2023).

2. Субъективность оценок

Еще одна важная проблема связана с субъективностью оценивания. В работе подчеркивается, что традиционные академические оценки не всегда объективно отражают знания студентов, так как они могут зависеть от методологии преподавания, личных предпочтений преподавателей и специфики образовательных учреждений. Например, один и тот же уровень знаний может быть оценен по-разному в разных университетах или даже у разных преподавателей в рамках одной программы.

Для улучшения объективности анализа успеваемости студента исследователи предлагают следующий подход:

- Опора на стандартизированные показатели. Помимо классических оценок, рекомендуется учитывать результаты тестов и экзаменов, что позволяет получить более сопоставимые данные.

- Учет учебной активности. Предлагается анализировать поведенческие индикаторы, такие как количество посещенных занятий, участие в учебных проектах и т. д., активность в цифровых образовательных системах.
- Оценка динамики оценок. Вместо одноразовых значений важно отслеживать изменения в академических результатах обучающегося на протяжении времени. Данный подход помогает выявить устойчивый прогресс или, наоборот, снижение мотивации.
- Нормализация оценок. Для уменьшения влияния различий между учебными заведениями и индивидуальными подходами преподавателей нужно применять методы нормализации, выравнивающие шкалы оценивания.

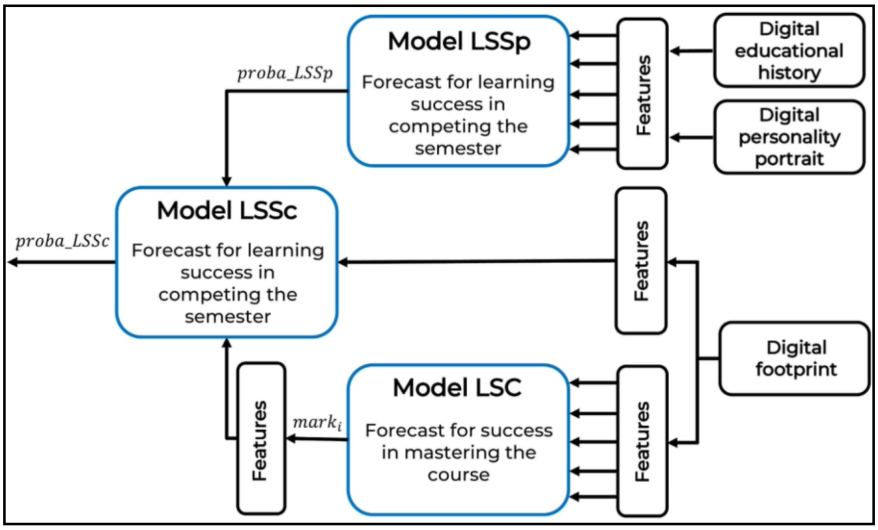

Рисунок 1.11 – Схема гибридного прогноза.
Источник: Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V. (2023).

3. Трудности учета разнообразных факторов

Прогнозирование академической успеваемости – это задача высокой сложности, так как требует взаимодействия разнородных параметров, включая как количественные, так и качественные показатели. Наиболее значимыми из них являются:

- Когнитивные способности обучающегося, включающие IQ, скорость обработки информации, логического мышления и т. д.
- Психофакторы, которые включает: уровень мотивации, стрессоустойчивость и способность к самостоятельному обучению.
- Социально-экономический статус (финансовая поддержка, образование родителей, доступ к учебным ресурсам).
- Демографические характеристики (пол, возраст, семейное положение).
- Цифровые следы (использование образовательных онлайн-платформ, активность в электронных курсах, частота взаимодействия с преподавателями через цифровые системы).

Однако многие из этих факторов трудно измерить количественно, что создает проблемы при построении математических моделей. Для устранения данной проблемы авторы предлагают использовать многомерные модели, которые способны одновременно учитывать числовые и категориальные признаки, а также разрабатывать методы, позволяющие системам адаптироваться к разнородной информации.

С целью повышения точности предсказательных моделей, Kustitskaya и соавторы рассматривают различные методы предварительной обработки данных, уделяя особое внимание проблеме дисбаланса классов. В рамках образовательной аналитики часто наблюдается ситуация, при которой число студентов, демонстрирующих успешную академическую успеваемость, значительно превосходит количество обучающихся, сталкивающихся с трудностями.

Методы устранения дисбаланса классов

Для решения этой проблемы исследователи предлагают следующие меры:

- Oversampling: увеличение представительности недостаточно представленного класса путём дублирования существующих примеров или генерации новых, аналогичных записей;
- Undersampling: сокращение численности доминирующего класса для достижения баланса между категориями;

- Генеративные методы: использование техник, таких как SMOTE (Synthetic Minority Over-sampling Technique), для синтетического создания новых наблюдений, приближенных к малочисленному классу, что позволяет улучшить качество обучения модели;
- Взвешенные функции потерь: применение штрафов к ошибкам предсказания с учетом соотношения классов, что снижает перекос в сторону более часто встречающегося класса.

Дополнительно, авторы поднимают вопрос обработки выбросов, которые представляют собой нетипичные или аномальные значения в данных. Например, единичные случаи исключительно низких или необычно высоких оценок. Наличие подобных аномалий в образовательных датасетах способно значительно исказить выводы модели, поэтому их корректная идентификация и устранение (или адаптация моделей к ним) играет важную роль в обеспечении достоверности аналитики.

Для выявления и корректировки подобных значения исследователи предлагают применять:

- статистические методы, такие как метод трех сигм или интерквантильных размах, которые помогают выделить типичные точки;
- использовать алгоритмы детекции выбросов, включая *Isolation Forest* и *Local Outlier Factor (LOF)*;
- проверять подозрительные записи вручную, чтобы исключить ошибки ввода данных.

Применения перечисленных подходов доказало свою эффективность: точность моделей заметно возрастает, что особенно хорошо видно по динамике прогнозирования в течении семестра.

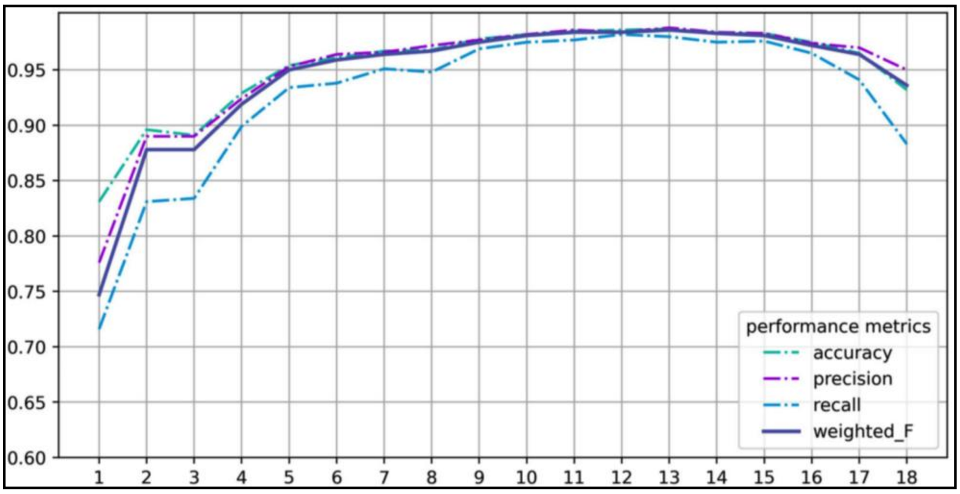

Рисунок 1.12 – Динамика качества прогнозирования модели в течение семестра. Источник: Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V. (2023).

Наконец, в статье подробно рассматривается использование ансамблевых моделей машинного обучения. Авторы показывают, что методы как случайный лес, градиентный бустинг (*XGBoost, LightGBM*) стекинг-подходы – позволяют добиться повышенной точности предсказаний благодаря объединению возможностей различных алгоритмов. Они выделяют три наиболее эффективные стратегии:

- Случайный лес – алгоритм, который формирует множество деревьев решений, а затем агрегирует их результаты. Такой подход хорошо справляется с переобучением и уверенно работает с разнородными наборами данных
- Градиентный бустинг – более сложная и мощная техника, где деревья обучаются последовательно с целью минимизации остаточной ошибки предыдущих моделей.
- Стекинг – метод комбинирования нескольких различных моделей (например, линейной регрессии, случайного леса и нейросети),

предсказания которых подаются на вход метамодели для получения финального результата.

Работы Kustitskay и её коллег подчёркивают важность интегрированного подхода к задаче прогнозирования академических достижений студентов. Исследователи приходят к выводу, что высокая точность предсказаний возможна только при комплексном учете множества факторов, применении современных методов подготовки данных и использовании продвинутых алгоритмов машинного обучения. Они также отмечают, что дальнейшие научные изыскания должны быть направлены на разработку ещё более гибких и адаптивных моделей, способных эффективно обрабатывать разнородную и сложно структурированную информацию в образовательной сфере.

Рекомендательные системы в образовании.

Belotsk, E.A., Suetin. A.V. (2024). Building a recommender system for selecting higher education institutions for applicants.

Статья Belostky и Sueting (2024) «Building a recommender system for selecting higher education institutions for applicants» – посвящена разработке и исследованию рекомендательной системы, предназначенной для помощи абитуриентам в выборе подходящих вузов.

В публикации подробно анализируется использование гибридных методов фильтрации, которые объединяют преимущества контентной фильтрации и коллаборативной фильтрации. Такой синтез позволяет значительно повысить персонализации рекомендаций, учитывая одновременно индивидуальные предпочтения пользователей и коллективные паттерны поведения других абитуриентов. Авторы детально описывают принципы функционирования этих методов, а также демонстрируют, как их совместное применение способствует построению более точных и релевантных рекомендаций.

Контентная фильтрация – это подход, основанный на анализе характеристик, параметров и атрибутов объектов рекомендаций (в данном случае – высших учебных заведений). Такой подход позволяет рекомендовать абитуриенту учебные заведения, которые схожи с теми, что он выбирал ранее, или которые соответствуют его текущим предпочтениям. В контексте выбора учебного заведения это может включать такие характеристики, как:

- Специализация вуза (например, технические, гуманитарные, медицинские направления и т. д.),
- Местоположение учебного заведения (город, регион, страна),
- Условия обучения (предоставление стипендий, стоимость обучения, возможность обучаться на бюджетной основе),
- Условия проживания (общежитие, аренда жилья вблизи учебного заведения),
- Рейтинг учебного заведения (например, международные или национальные рейтинги вузов).

Используя данные об этих характеристиках, система может предсказать, какие учебные заведения будут наиболее подходящими для абитуриента, основываясь на его предпочтениях и истории выбора.

Коллаборативная фильтрация, в свою очередь, основывается на анализе поведения других пользователей (абитуриентов), которые делали похожий выбор. Подход предполагает, что если группа студентов с похожими характеристиками, интересами или паттернами поведения выбирала определенные образовательные программы, то эти программы могут быть релевантны и для текущего абитуриента.

Коллаборативный подход включает несколько вариантов реализации:

- Пользовательская фильтрация – система выявляет предпочтения других абитуриентов, схожих с текущим, и рекомендует ему те учебные заведения, которые были выбраны похожими поведением.
- Предметно-ориентированная фильтрация – алгоритм анализирует, какие учебные заведения выбирают пользователи, отдавшие предпочтение программам, имеющим похожие характеристики (например, с похожей специализацией и уровнем образования).

Благодаря этому подходу система предлагает образовательные программы, которые могут не выделяться по содержательным признакам, но часто выбираются определенными группами студентов со схожими интересами.

Гибридный подход

Авторы отмечают, что наиболее результативным решение при создании рекомендательных систем является гибридный метод, объединяющий идеи конечной и коллаборативной фильтрации. В отличии от классических моделей, основанных только на одном типе анализа, комбинированные системы способны одновременно учитывать особенности объектов и поведенческие данные пользователей, что в итоге повышает точность выдаваемых рекомендаций.

Как функционирует гибридная система:

- **Совмещение контентного и коллаборативного анализа.** На первом этапе алгоритм может применить контентную фильтрацию, чтобы отсечь заведения, не соответствующий базовым критерии, а затем подключает коллаборативный компонент, который уточняет выбор исходя из поведения абитуриентов с аналогичными предпочтениями.
- **Переходные механизмы.** Для объединения различных типов данных в гибридных системах используются специальные переходные функции. Они задают вес каждому из факторов. Например, контентная фильтрация может выполнять грубую селекцию, после чего коллаборативная часть дополняет список вариантами, популярными среди близких по профилю пользователей.
- **Модели смешивания.** В данном случае оба подхода работают одновременно, а итоговый список рекомендаций формируется путем объединения их результатов. Такой метод снижает влияние слабых сторон каждого из алгоритмов и усиливает их достоинства.

Результаты экспериментов

В статье приведены детальные результаты тестирования гибридных моделей на реальных данных абитуриентов. Информация включала их получения, интересы и сведения об учебных заведениях. Анализ показал, что комбинированные алгоритмы дают заметно более высокие показатели точности по сравнению с системами, основанными только на одном методе.

Дополнительно установлено, что комбинированные рекомендательные модели показывают высокую эффективность при решении проблемы «холодного старта», когда информация о новых пользователях практически отсутствует. Эти модели показывают устойчивость в условиях ограниченности данных о конкретных образовательных учреждениях.

Исследователи Белоцкий и Суетин акцентируют внимание на важности полученных результатов. По их мнению, гибридные рекомендации являются мощным инструментом в профориентации абитуриентов, потому что могут одновременно учитывать личные характеристики кандидата и поведение других пользователей с аналогичными интересами. Сочетание контентного и коллаборативного подходов приводит к улучшению точности рекомендаций, компенсируя ограничения каждого метода в отдельности. Помимо этого, подобные алгоритмы могут быть успешно адаптированы для смежных задач, включая выбор образовательных курсов, программ и специализаций.

Авторы подчеркивают, что важным направлением дальнейших исследований является развитие гибридных моделей – создание новых схем фильтрации и повышение качества рекомендаций в условиях работы с разнородными и неполными данными.

Mityanina, A., Gazha, K. (2024). Recommender system «Abiturient» for university admission committee

В работе Mityanina и Gazha (2024) рассматривается создание и внедрение рекомендательной системы «Abiturient», предназначенной для помощи абитуриентам при выборе университетов. Разработчики используют современные методы машинного обучения позволяющие формировать персональные рекомендации, сопоставляя характеристики кандидата с требованиями образовательных учреждений. Такая система облегчает работу приемной комиссии и помогает преуспевающим найти учебные заведения, наиболее подходящие к их профилю.

Система «Abiturient» базируется на нескольких ключевых алгоритмах анализа данные абитуриента. Среди них:

1. Алгоритмы классификации, которые могут:
   - Оценивать профиль каждого студента на основе собранных факторов.
   - Анализировать такие параметры, как баллы на вступительных экзаменах, академические достижения, участие во внешкольных и общественных активностях.
   - Учитывать предпочтения абитуриента в выборе специализации и места учебы.
2. Методы машинного обучения, которые позволяют:
   - Формировать прогноз наиболее подходящего университета для каждого кандидата.
   - Сопоставлять профиль абитуриента с требованиями учебных заведений.
   - Определять вероятность успешного поступления.
3. Методы ранжирования, которые обеспечивают:
   - Оптимизацию процесса выбора вузов.
   - Сортировку учебных заведений по степени соответствия интересам и данным абитуриента.
   - Анализ исторических данных о поступлении и успеваемости студентов с похожими характеристиками.
4. Анализ социально-экономического контекста, включающий:
   - Финансовое положение студента.
   - Рекомендации вузов, предоставляющих стипендии или доступные условия обучения.

Использование таких методов машинного обучения в разрабатываемой авторами системе помогает:

- Повысить точность рекомендаций, обеспечивая более персонализированный подход к каждому абитуриенту.
- Учитывать динамичные изменения в требованиях университетов (изменения стоимости обучения, новых академических программ и вступительных экзаменов).
- Сделать систему гибкой и адаптируемой к изменениям в образовательной среде.

Авторы статьи подчеркивают, что разрабатываемая система предоставляет значительные преимущества:

- **Для студентов:**
    - Гарантирует рекомендации, соответствующие их академическим и личным интересам.
    - Повышает шансы на успешное поступление.
- **Для университетов:**
    - Позволяет привлекать наиболее подходящих кандидатов.
    - Улучшает качество набора студентов.
    - Снижает вероятность академической неуспеваемости студентов.

Внедрение таких технологий в работу приемных комиссий способствует:

- Ускорению процесса поступления.
- Повышению прозрачности отбора студентов.
- Улучшению планирования набора в университетах.
- Эффективному взаимодействию с кандидатами, включая информирование о возможностях (стипендии, образовательные программы).

Использование методов машинного обучения в системе «Abiturient» позволяет заметно повысить точность выдаваемых рекомендаций, обеспечивая более индивидуализированный подход к каждому абитуриенту. Кроме того, применяют алгоритмы способны учитывать изменения, происходящие в требованиях вузов – например, корректировку стоимости обучения, появление новых образовательных программ или обновление правил вступительных испытаний. Благодаря этому платформа остается гибкой и способной быстро подстраиваться под трансформации в образовательной среде.

Рисунок 1.13 – Предложенная авторами архитектура системы.

Авторы отмечают, что система «Abiturient» дает ощутимые преимущества как будущим студентам, так и самим вузам. Абитуриенты получают рекомендации, которые лучше соответствуют их личным целям, уровня подготовки и интересам, что повышает вероятность успешного поступления. Вузы же применять такую систему для отбора наиболее мотивированных и подходящих кандидатов, тем самым повышая качество контингента и увеличивая шансы на успешную интеграцию студентов в учебный процесс.

Подводя итоги, исследование Mityanina и Gazha подчеркивает значимость использования алгоритмов машинного обучения при формировании персонализированных рекомендаций в высшем образовании. «Abiturient» не только упрощает и совершенствует процесс выбора университета для абитуриентов, но и помогает учебным заведениями точнее прогнозизрвать успешность потенциальных студентов и выбирать наиболее подходящих кандидатов. Авторы подчеркивают, что дальнейшее развитие рекомендательных систем способно вывести такие инструменты на новый уровень точности и эффективности, что в будущем может существенно улучшить качество образовательных процессов.

5. Алгоритмы классификации и обработка данных

Even-Dar, E., Mannor, S., Mansour, Y. (2002). Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems

Статья Even-Dar, E., Mannor, S., Mansour, Y. (2002). «Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems» исследует задачи, связанные с многоруким бандитом (Multi-Armed Bandit, MAB), а также стратегии, направленные на оптимизацию выбора действий в условиях неполной и ограниченной информации. Это одна из ключевых проблем в теории принятия решений и области машинного обучения, которая имеет широкий спектр применений, включая рекомендательные системы и, в частности, образовательные системы.

Основные положения статьи:

1. **Задача многорукого бандита:**
    - Агент должен выбирать одно из нескольких возможных действий (или «рук»), не имея полной информации о потенциальных выигрышах.
    - Цель – максимизировать суммарный выигрыш, обучаясь на накопленном опыте.
    - Применимость в рекомендательных системах: необходимость предлагать пользователям релевантные рекомендации при неполных данных.
2. **Методы устранения неэффективных действий:**
    - Исключение вариантов, демонстрирующих низкую эффективность.
    - Оптимизация процесса обучения путем сосредоточения на наиболее перспективных альтернативах.
    - В образовательных рекомендательных системах это может означать отказ от предложений вузов и курсов, не вызывающих интереса у пользователей.
3. **Условия остановки в процессе обучения:**

- Определение момента, когда агент накопил достаточную информацию для принятия оптимального решения.
- В реальных приложениях, таких как рекомендательные системы, важно избегать избыточного тестирования и излишних вычислительных затрат.
- В образовательных технологиях алгоритмы остановки помогают своевременно сформировать персонализированные рекомендации, избегая продолжительных циклов экспериментов.

4. **Применение в образовательных рекомендательных системах:**
   - Анализ данных абитуриентов: вступительные баллы, академические предпочтения, социально-экономический контекст.
   - Исключение наименее релевантных вариантов и фокусировка на более перспективных предложениях.
   - Определение оптимального момента завершения тестирования рекомендаций и формирования окончательного списка вузов.

В задаче многорукого бандита агент должен выбрать одно из нескольких возможных действий (или «рук»), каждое из которых имеет свою степень неопределенности. Проблема заключается в том, что агент не знает заранее, какое действие приведет к наибольшей отдаче, и должен на основе опыта выбирать те действия, которые будут максимизировать общий выигрыш. Данная ситуация аналогична задаче в рекомендательных системах, где система должна прогнозировать пользователю наиболее подходящие рекомендации, основываясь на его предпочтениях, но без полной информации о том, какие рекомендации окажутся наилучшими.

Основная цель работы Even-Dar и соавторов заключается в разработке подходов, позволяющих эффективно отсеивать не результативные действия и формировать корректные условия остановки при поиске оптимальной стратегии. Проблема исключения неэффективных действий заключается в том, чтобы вовремя исключить те варианты, которые показывают плохие результаты, и сосредоточиться на тех, которые приносят наибольшую прибыль. Если рассматривать это в контексте рекомендательных систем, то речь идет о своевременном отказе от рекомендаций, которые не вызывают интереса у пользователей, и переход к более релевантным или интересным вариантам.

Вторым ключевым аспектом работы является формирование критериев остановки. В классической задаче многорукого бандита агент продолжает проверять различные варианты до тех пор, пока не получит достаточно данных для выбора наилучшей стратегии. Однако в реальных приложениях, таких как рекомендательные системы, часто существуют ограничения по времени и объему информации, из-за чего необходимо в определенный момент прекратить экспериментирование и принять решение. Алгоритмы остановки позволяют решить эту проблему, определяя момент, когда можно завершить тестирование доступных альтернатив и сделать вывод о том, какие из них показывают наилучшую эффективность.

Применение подобных методик в рекомендательных системах, ориентированных на образование, может оказаться особенно ценным. Например, при подборе для абитуриентов возможных учебных заведений или программ система может использовать алгоритмы отсечения неэффективных варианте, исключая те вузы или курсы, которые не вызывали интереса у предыдущих пользователей. Это позволяет сосредоточиться на предложениях, которые с большей вероятностью привлекут внимание студента. Не менее важно задействовать и механизмы остановки: система должна уметь определять момент, когда собранных данных достаточно для формирования точны и индивидуальных рекомендаций, избегая при этом бессмысленного повторного тестирования одних и тех же альтернатив.

Методы, предложенные Eve-Dar и коллегами, способны существенно повысить результативность рекомендательных систем, особенно в условиях ограниченных или неполных данных, а также при наличии пользователей с разнородными и изменяющимся предпочтениями. Благодаря сочетанию доходов по исключению слабых решений и алгоритмов остановки, система может гибко адаптироваться к обновлениям и оперативно предлагать наиболее релевантные варианты.

В итоге внедрение таких подходов позволяет создавать рекомендательные системы, которые не только эффективно функционируют при дефиците данных, но и способны быстро реагировать на новые условия, улучшая качество и точность рекомендаций. В сфере образовательных технологий это может способствовать более индивидуализированным стратегиями поступления, повышая удовлетворенность абитуриентов и общую эффективность образовательных процессов.

Jabeen, H., Baig, A.R. (2010). Review of Classification Using Genetic Programming.

Статья Jabeen и Biag (2010) представляет развернутый обзор применения генетического программирования (*GP*) в задачах классификации; Генетическое программирование рассматривает как мощный инструмент эволюционной оптимизации, основанный на механизмах естественного отбора, генетических мутаций и эволюционного развития. В отличие от многих классических методов машинного обучения, которые стремятся лишь подобрать оптимальные параметры модели, GP оптимизирует одновременно и структуру модели, и ее параметры. По сути, метод генерирует новые программы или классификаторы, способные самостоятельно подстраивать под особенности данных ри решаемые задачи. Данный метод позволяет создавать более сложные и адаптивные решения, чем при использовании классических алгоритмов.

Преимущества использования генетического программирования в классификации:

- Высокая гибкость – GP может подстраиваться под различные структуры данных.
- Формирование нелинейных моделей – метод способен моделировать неочевидные взаимосвязи между признаками
- Масштабируемость – при увеличении объема данных метод показывает лучшие результаты по сравнению с другими методами.

- Универсальность метода – способность обрабатывать как числовые, так и категориальных данных.

В то же время метод имеет несколько значительных ограничений:

- Высокая вычислительная стоимость – требует много ресурсов из-за большого количества итераций.
- Медленный процесс обучения – большое количество циклов эволюции увеличивает время обучения.
- Ограниченная применимость в ресурсозависимых задачах – сложно использовать в реальных приложениях с жесткими временными ограничениями.

Процесс эволюции в GP требует от алгоритма выполнения большого количества циклов, в каждом из которых оцениваются различные программы, и в итоге выбираются лучшие из них. Эта итерационная природа процесса означает, что обучение может быть весьма медленным, что ограничивает использование генетического программирования в реальных приложениях с ограниченными ресурсами. Для решения данной проблемы ученые разрабатывают методы ускорения эволюционного процесса, такие как параллельные вычисления и улучшенные алгоритмы отбора.

Но, несмотря на существующие вычислительные ограничения, генетическое программирование остается весьма перспективным подходом для решения задач классификации, особенно когда требуется анализировать сложные, многомерные и разнородные данные. Данный метод может применяться в разных направлениях: например, в медицинской диагностике, где важна высокая точность при прогнозировании заболеваний; в финансовой сфере – для выявления динамики и предсказания рыночных колебаний; в обработке изображений – для распознавания объектов: а также в других случаях, когда классические алгоритмы машинного обучения оказываются недостаточно эффективными.

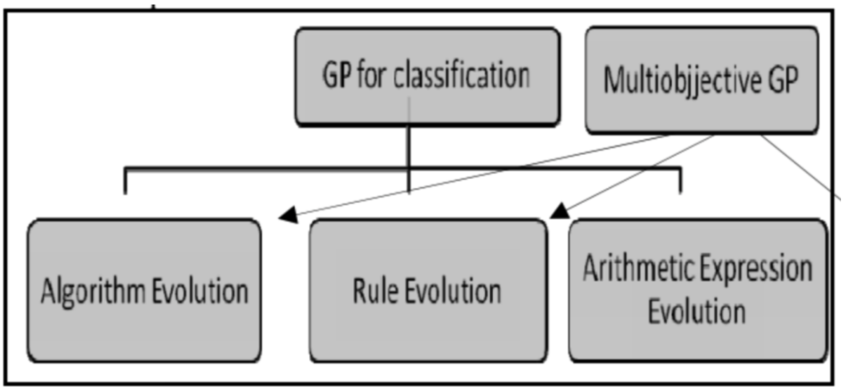

Рисунок 1.14: Классификация алгоритмов GP.
Источник: Jabeen, H., Baig, A.R. (2010).

С учетом стремительного роста вычислительных ресурсов GP может стать более доступным и широко применимым в практических приложениях, обеспечивая высокую точность в различных прикладных областях, включая образование, медицину, финансы и промышленную аналитику.

Таким образом, статья Jabeen и Baig показывает, что хотя генетическое программирование и требует значительных вычислительных ресурсов, его преимущества в создании высокоэффективных и гибких классификаторов делают его важным методом в области машинного обучения и классификации данных.

Автоматизация образовательных процессов

Molchanov, A.N., Burmistrov, A.V., Grishunov, S.S., Chukhraev, I.V. (2016). Some issues of data collection and processing in the information system of a higher educational institution

Статья посвящена анализу проблем сбора, обработки и интеграции данных в информационных системах высших учебных заведений. Данная статья дает всестороннее представление о ключевых аспектах функционирования описанных систем, затрагивая вопросы стандартизации, унификации и автоматизации процессов управления образовательной информацией. Авторы обращают внимание на то, что результативность работы информационных платформ образовательных учреждений во много зависит от корректной организации процессов по сбору, хранению и анализу данных. При этом важно учитывать, что работа с образовательной информацией предъявляет высокие требования к точности, надежности и способности системы адаптироваться к изменяющимся запросам учебных организаций и регуляторов.

В статье описаны трудности, возникающие при сборе данных в сфере образования. Одной из ключевых проблем является разрозненность источников информации: данные поступают из различных подразделений – кафедр, деканатов, библиотек, приемных комиссий, а также платформ дистанционного обучения. При отсутствии единых стандартов хранения и обмена данными возникают такие проблемы как дублирование, ненормализация данных и т.д., что увеличивает вероятность ошибок и снижает уровень доверия к информации.

Для эффективного управления образовательными данными необходимо учитывать ряд факторов, включая:

- разработку единого стандартизированного формата хранения информации для всех подразделений;
- внедрение автоматизированных механизмов проверки корректности и валидации данных;
- обеспечение целостности, актуальности и согласованности информации;

- контроль версионности, позволяющий избежать использования устаревших данных;
- защиту персональных данных обучающихся и сотрудников вуза.

В рамках исследования был проведён анализ популярных информационных систем, применяемых в образовательных учреждениях, таких как «1С: Университет» и федеральная информационная система ГИА и приема. Несмотря на широкое распространение и функциональное разнообразие этих решений, они обладают рядом существенных недостатков. К их числу относятся высокая стоимость внедрения и сопровождения, сложность интеграции с существующей инфраструктурой вузов, а также ограниченные возможности адаптации к изменениям нормативно-правовой базы и образовательных стандартов.

Данные факторы указывают на актуальность создания более гибких и целенаправленных программных решений, разработанных с учетом специфики отдельных учебных заведений и их внутренних процессов.

В работе также рассмотрены современные подходы к повышению эффективности функционирования образовательных информационных систем. Отмечено, что с развитием технологий анализа данных и интеллектуальных алгоритмов открываются новые перспективы в области управления образовательной деятельностью. Среди приоритетных направлений оптимизации авторы выделяют:

- использование облачных технологий для хранения и обработки данных;
- применение блокчейн-технологий для повышения безопасности и прозрачности учета данных;
- внедрение предиктивной аналитики для прогнозирования академической успеваемости студентов;
- автоматизация обработки заявок, экзаменационных результатов и аттестационных данных;
- интеграция локальных информационных систем вузов с федеральными и международными образовательными платформами.

Другой ключевой аспект – создание единой цифровой образовательной среды, объединяющей различные источники данных в рамках централизованной платформы. Авторы статьи предлагают концепцию универсальной платформы, обеспечивающей удобный доступ к образовательной информации, автоматизированный документооборот и эффективное взаимодействие между структурными подразделениями вуза.

Важными модулями такой платформы являются:

- централизованные базы данных, содержащие структурированную и достоверную информацию;
- интерфейсы API, обеспечивающие взаимодействие между различными модулями системы;
- средства аналитической обработки и визуализации образовательных данных;

Разработка и внедрение специализированной подсистемы

Практическая часть исследования была посвящена проектированию и внедрению специализированного модуля для сбора и обработки данных абитуриентов. Отличительной чертой разработанного решения является модульный подход, позволяющий адаптировать подсистему под потребности конкретного вуза. Архитектура системы включает в себя базу данных и клиентское приложение с понятным пользовательским интерфейсом, что значительно упрощает работу сотрудников приемных комиссий.

В качестве основы хранения информации была выбрана СУБД Firebird, зарекомендовавшая себя как надежная и производительная платформа. Разработка включала следующие этапы:

- проектирование базы данных для хранения персональных и заявочных данных абитуриентов;
- реализация модуля обработки заявлений и отслеживания этапов их прохождения;
- внедрение системы учёта представленных документов;
- интеграция блока анализа результатов вступительных испытаний;

- автоматизация формирования отчётности;
- обеспечение совместимости с государственными и федеральными системами мониторинга абитуриентов.

Эффект от внедрения

Внедрение разработанной подсистемы позволило:

- повысить эффективность работы приёмной кампании;
- снизить нагрузку на сотрудников, задействованных в обработке данных;
- уменьшить количество ошибок за счёт автоматизации операций;
- минимизировать влияние человеческого фактора.

Выводы и значимость работы

В статье дается развернутый анализ актуальных задач и эффективных решений в области цифровизации управления образовательными данными. Подчеркивается, что автоматизация и стандартизация процессов информационного взаимодействия являются ключевыми факторами повышения качества образовательной среды вуза.

В числе основных результатов исследования можно выделить:

- рост точности и оперативности обработки образовательной информации;
- упрощение процессов администрирования и документооборота;
- снижение затрат на эксплуатацию ИТ-инфраструктуры;
- обеспечение безопасности и целостности данных;
- расширение возможностей интеграции с государственными и международными информационными системами.

Разработанная и протестированная подсистема успешно показала свою применимость в условиях реального вуза. Работа Molchanov и др. (2016) подчёркивает необходимость системного подхода к обработке образовательной информации и демонстрирует перспективные направления развития цифровых образовательных решений.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

2.1 Введение

Выбор специальности в вузе является одним из наиболее значимых этапов для абитуриента, поскольку именно он определяет дальнейшую траекторию академического роста, будущие карьерные возможности и профессиональное развитие. Несмотря на наличие формальных процедур и критерии отбора, многие пустующие сталкиваются с проблемами, связанными с несовпадением между их реальными навыками и выбранным направлением подготовки. Традиционные механизмы, опирающиеся преимущественно на результаты Общереспубликанского тестирования (ОРТ), оценивают лишь ограниченный набор показателей, что зачастую приводит к выбору, не отражающему истинный потенциал поступающего. В итоге во время обучения нередко возникают сложности с освоением учебного материала, снижением успеваемости, потерей мотивации и даже необходимостью сменить специализацию.

Поэтому возникает объективная потребность в создании более продвинутых инструментов, способных эффективно поддерживать абитуриентов при принятии решения. В этих условиях технологии машинного обучения выступают как мощный механизм анализа и интеграции больших массивов данных, что позволяет формировать персонализированные и более обоснованные рекомендации.

Текущие подходы и их ограничения:

2.1.1 ***Экспертные системы***.

Подобные системы опираются на набор правил и рекомендаций, заранее сформированных специалистами в сфере образования. Однако их главный недостаток заключается в низкой гибкости, такие решения плохо подстраиваются под изменения в образовательной среде и не способны в полной мере учитывать индивидуальные результаты тестирования абитуриентов.

2.1.2 ***Анализ исторических данных.***

Этот подход использует статистические методы и анализ корреляций, основанный на данных о предыдущих потоках студентов. Он позволяет определить ключевые факторы, влияющие на академический успех, но при этом остается недостаточно персонализированными, потому что выводы строятся на общих тенденциях, а не на специфике каждого конкретного абитуриента.

2.1.3 Подход на основе методов машинного обучения.

В отличие от традиционных методик, алгоритмы машинного обучения могут обрабатывать гораздо более широкий набор признаков и выявлять сложные нелинейные зависимости. Это создает условия для построения более точных и индивидуализированных рекомендаций. Модели ИИ способны объединять данные о результатах ОРТ, академической успеваемости прошлых студентов, среднем балле, достижениях по отдельных дисциплинам, а при необходимости – и дополнительную информацию, отражающую профессиональные склонности абитуриента.

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.

2.2 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 (ОРТ) 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. 2.1

Рисунок 2.1 – 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:

- *ОРТ* – 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 2.2.

2 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.

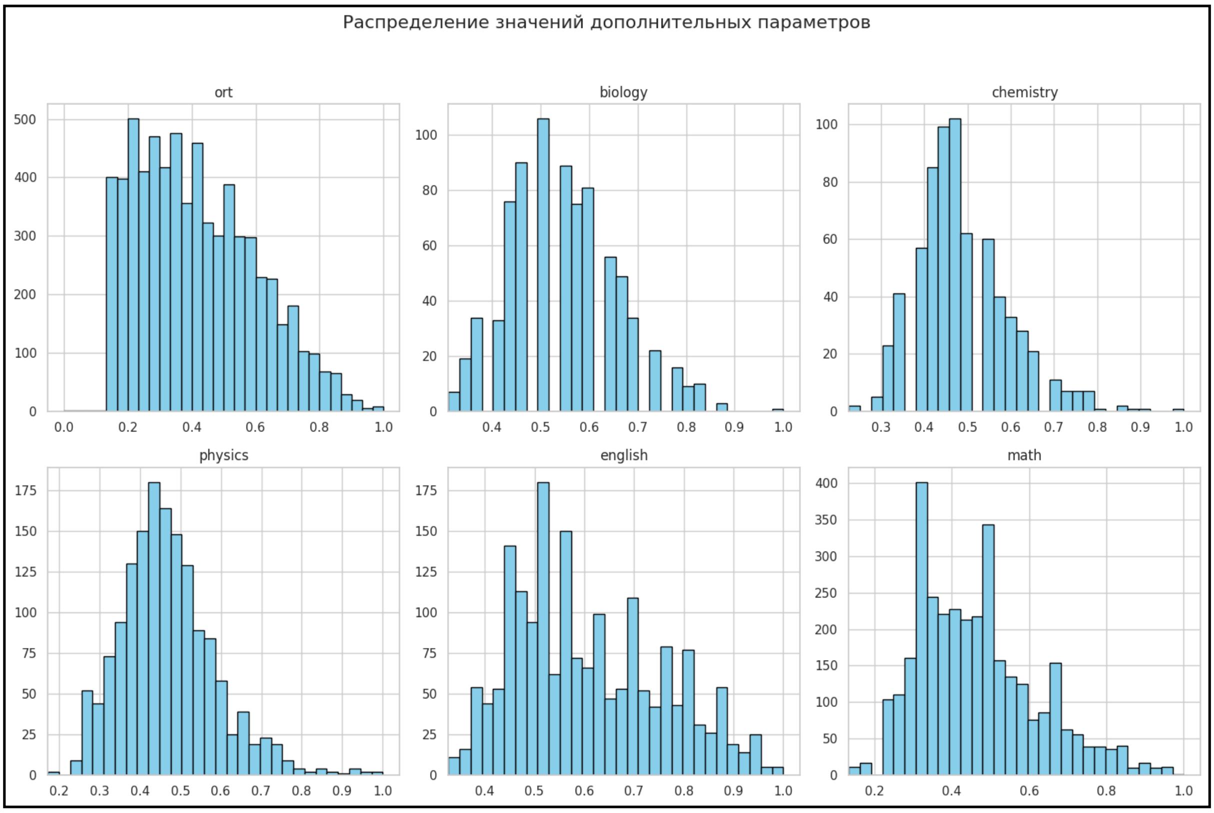

Рисунок 2.2 – Гистограмма распределений основного и дополнительных предметов по ОРТ

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:

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

- $\sigma(x) - \frac{1}{1+e^{-x}}$ – 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 $tahn(x)$.
- 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.

По мере того, как нейронные сети становятся все более сложными для решения сложных задач в области обработки естественного языка (NLP), компьютерного зрения и других областях ИИ, появляются новые архитектурные элементы для повышения их эффективности и гибкости. Одним из таких элементов современных нейронных архитектур является нейрон с многоразветвлённой структурой и динамическим гейтингом – *Multi-Branch Gated Neuron (MBGN).* Эта архитектура сочетает возможности параллельной обработки данных с адаптивным управлением информационными потоками, что

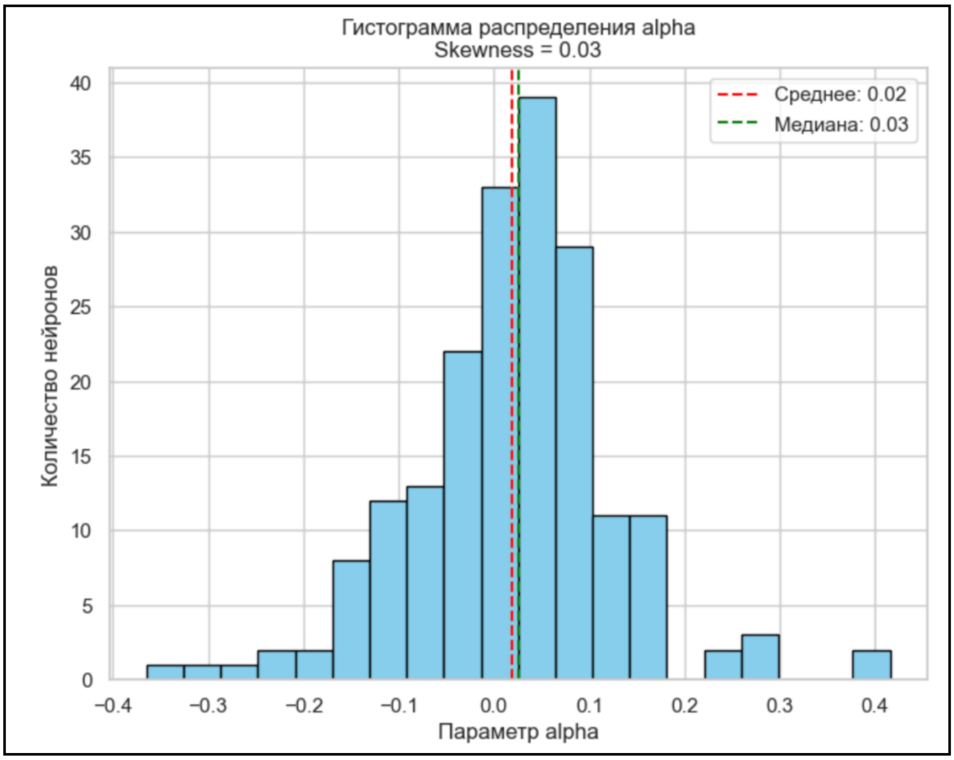

Рисунок 2.3 – Гистограмма распределения параметра Alpha

позволяет более эффективно моделировать сложные зависимости и структуры в данных.

Основные принципы

Нейрон с много разветвленной системой управления объединяет две ключевые идеи:

1. **Многоразветвлённая структура обработки**: Каждый нейрон включает несколько параллельных ветвей, каждая из которых независимо обрабатывает входной сигнал. Эти ветви могут быть реализованы с использованием различных методов обработки:
   - специализированные фильтры;
   - альтернативные трансформации входных признаков;
   - механизмы внимания (attention) для акцентирования на значимых аспектах данных.
2. **Механизм гейтинга (стробирование)**: Каждая ветвь оснащена отдельным обучаемым гейтом (затвором), который регулирует её вклад в итоговый

отклик нейрона. Эти гейты, как правило, реализуются с использованием сигмоидальной функции активации, возвращающей значение в диапазоне от 0 до 1, что интерпретируется как:

  - 0 – ветвь исключена из обработки (заблокирована),
  - 1 – ветвь полностью активна.

3. **Объединение результатов**: после обработки и взвешивания выходов всех ветвей происходит их интеграция в единый отклик нейрона. В зависимости от архитектурных решений могут использоваться различные методы агрегации:
  - суммирование (для объединения признаков одной размерности),
  - конкатенация (для увеличения размерности представления),
  - или другие агрегирующие операции (например, усреднение или макс-пулинг).

Архитектура

Предположим, что у нас есть входной вектор $x \in R^d$. . Мы хотим создать один **нейрон *MBGN***, состоящий из K параллельных ветвей. Каждая ветвь k имеет свой собственный весовой вектор $w_k \in R^d$ и член смещения b_k.

4.1 Линейная компонента (ветви)

Для каждой ветви k мы вычисляем линейную комбинацию:

$$z_k = w_k x + b_k, k = 1, \dots, K.$$

Это дает нам скалярные значения K : $z_1, z_2 \dots . z_k$

4.2 Активация в каждой ветви

Мы применяем функцию активации (например, ReLU или tanh) к каждой ветви z_k

$$a_k = \sigma(z_k).$$

Таким образом, каждая ветвь производит активированный выходa_k . (Можно использовать разные функции активации для разных ветвей, что добавляет еще один уровень гибкости и специализации).

4.3 Гейтинг (управление вкладом ветвей)

Чтобы позволить модели динамически определять, какой вклад должна вносить каждая ветвь для данного входа, мы вводим обучаемые коэффициенты гейтинга α_k. Они зависят от входа x и нормируются для обеспечения контролируемого вклада.

Простой подход заключается в том, чтобы ввести отдельный **слой стробирования**, который производит ненормированные логиты $\tilde{\alpha}_k$ из x . Затем мы нормализуем их с помощью softmax:

$$\alpha_k = \frac{exp(\tilde{\alpha}_k)}{\sum_{j=1}^{K} exp(\tilde{\alpha}_j)}$$

Значения $\widetilde{a}_k$ вычисляются аналогично z_k (например, с отдельными весами и смещениями), или даже с помощью небольшой нейронной сети для более сложного поведения, похожего на внимание.

В более простых вариантах □$_k$ могут быть **обучаемыми константами** (не зависящими от x) или простыми линейными функциями x.

4.4 Агрегация ветвей

Наконец, выход нейрона *MBGN* — это:

$$y = \sum_{k=1}^{K} \alpha_k * a_k.$$

Это **взвешенная комбинация** активаций всех параллельных ветвей.

В итоге вместо одной операции $y = \sigma(w^\top x + b)$ мы получаем комбинацию нескольких «траекторий обработки» для x , что приводит к созданию гораздо более мощного и гибкого блока.

Как показано на рисунке 2.4: входной сигнал x обрабатывается параллельно несколькими ветвями, каждая из которых состоит из линейного слоя, за которым следует отдельная функция активации. Выходы $a_1, a_2, \ldots a_n$ объединяются с помощью выученных коэффициентов активации (через softmax) для формирования взвешенной суммы, дающей конечный выход y.

4.5 Архитектурная новизна

Предложенный нейрон с несколькими ветвями вводит новый архитектурный элемент в нейронные сети, обрабатывая входные данные через

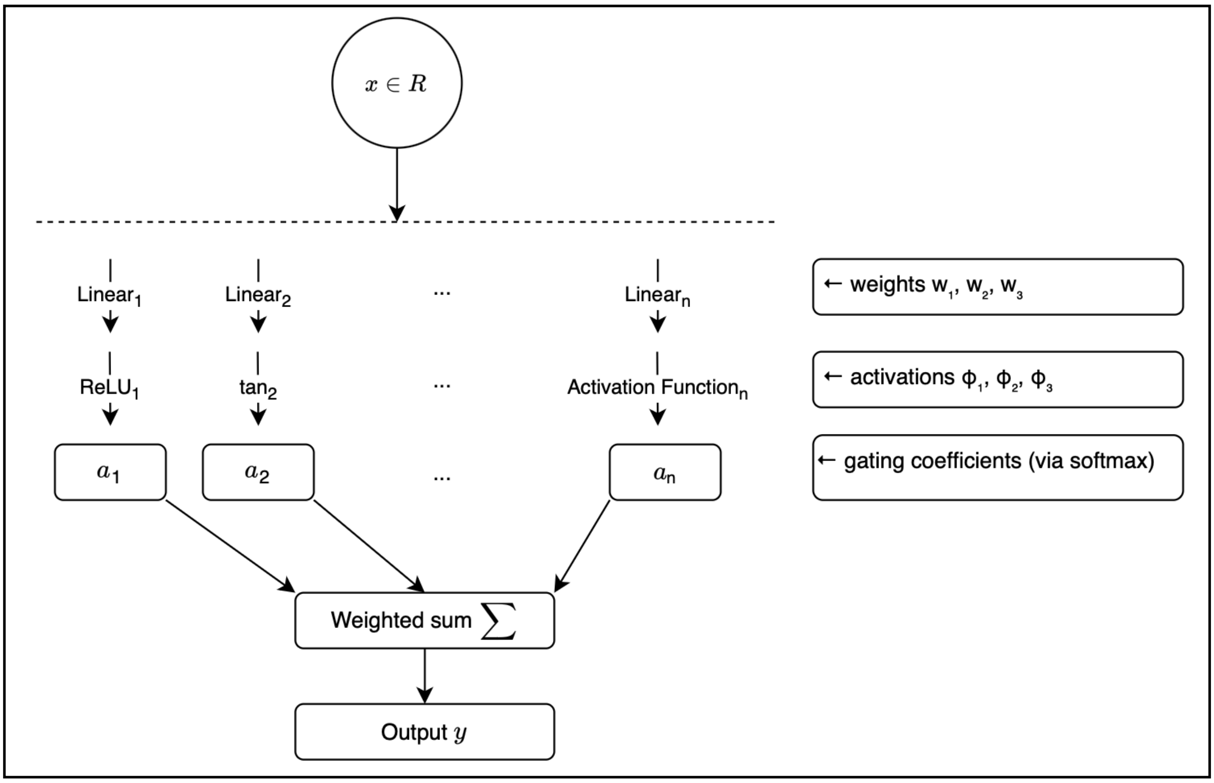

Рисунок 2.4 – Архитектура предлагаемого нейрона с несколькими ветвями (*Multi-Branch Gated Neuron, MBGN*).

несколько параллельных ветвей, каждая из которых имеет свое линейное преобразование и функцию активации (например, *ReLU*, *tahn*, *sigmoid* и т. д.). Эти ветви генерируют различные представления признаков, которые затем динамически объединяются с помощью набора доступных коэффициентов активации, получаемых с помощью функции softmax.

В отличие от традиционных нейронов, которые используют единственный путь активации (см. рис. 2.5). *MBGN* позволяет модели адаптивно выбирать и взвешивать несколько преобразований одного и того же входного сигнала в зависимости от контекста. Это позволяет сети использовать сильные стороны различных функций активации, улучшая репрезентативные возможности и гибкость. Общий выход вычисляется, как взвешенная сумма всех выходов ветвей, при этом веса оптимизируются в процессе обучения.

Такой подход повышает способность модели улавливать сложные закономерности и улучшает обобщение, особенно в сценариях, где полезны различные преобразования признаков.

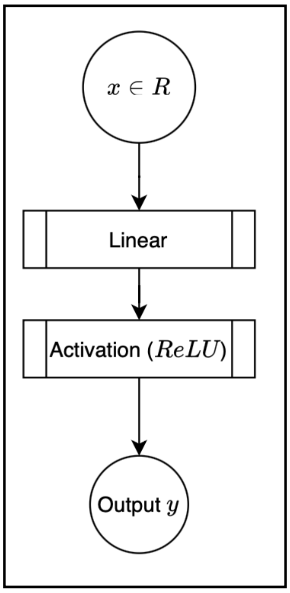

Рисунок 2.5 – Одиночный путь активации.

На рисунке 2.5 изображен одиночный путь активации: входной сигнал x обрабатывается с помощью единственного линейного преобразования, за которым следует фиксированная функция активации. Выходной сигнал генерируется без разветвлений или стробирования.

4.6. *MBGN* для решения задач классификации

Разные ветви для разных паттернов

Каждая ветвь может специализироваться на распознавании различных паттернов или признаков. Например, одна ветвь может сильно реагировать на определенные характеристики входного сигнала, в то время как другая может быть более чувствительна к другим паттернам. Когда механизм стробирования зависит от x , сеть учится «активировать» только наиболее релевантные ветви для данного входа и подавлять менее релевантные.

Это позволяет модели адаптироваться к различным типам входных данных, эффективно распределяя вычисления между ветвями, которые лучше всего подходят для данного случая.

Гейтинг как форма внимания

Коэффициенты стробирования $\square_k$ можно интерпретировать как распределение внимания по ветвям. Для одного входа может доминировать α_1, в то время как для другого входа приоритет может иметь $\square_2$ или $\square_3$. Такая гибкость помогает сети динамически адаптироваться.

Таким образом, механизм стробирования действует как механизм выбора контекста: он позволяет модели сосредоточиться на наиболее информативных преобразованиях для каждого отдельного входа.

Регуляризация и экспрессивность

Наличие множества параллельных ветвей увеличивает выразительность каждого блока (даже одного нейрона), что позволяет снизить необходимость в чрезмерной глубине сети. В то же время нормализованные по softmax гейты гарантируют, что не все ветви будут «полностью активны» одновременно, что вносит эффект регуляризации: некоторые ветви могут эффективно отключаться (когда $\square_k \approx 0$) для определенных входов.

Это делает *MBGN* не только более гибкой, но и более устойчивой к перестройке, поскольку модель учится использовать только самые полезные ветви и игнорировать нерелевантные. В результате достигается хороший баланс между точностью и способностью к обобщению.

Потенциальные варианты и расширения

- **Различные функции активации**: Одна ветвь может использовать *ReLU*, другая - *tanh*, третья - *LeakyReLU* и т. д. Модель может «выбирать» наиболее подходящий путь обработки в зависимости от входных данных.
- **Небольшая нейронная сеть стробирования**: Вместо линейного стробирования небольшая *MLP* может вычислять $\tilde{\alpha}_k$ для более сложного контекстно-зависимого управления.
- **Пропускные соединения**: Добавление пропускных соединений от входа к выходу нейрона *MBGN* может еще больше улучшить динамику обучения.

7. Конвейерная модель для прогнозирования вероятности поступления на бюджет

Для прогнозирования формы обучения (бюджет/контракт) используется многоступенчатый конвейер, включающий два последовательно связанных предсказательных модуля: нейронную сеть с адаптивной функцией активации и архитектуру *MBGN*. Общий процесс можно представить в виде следующей схемы:

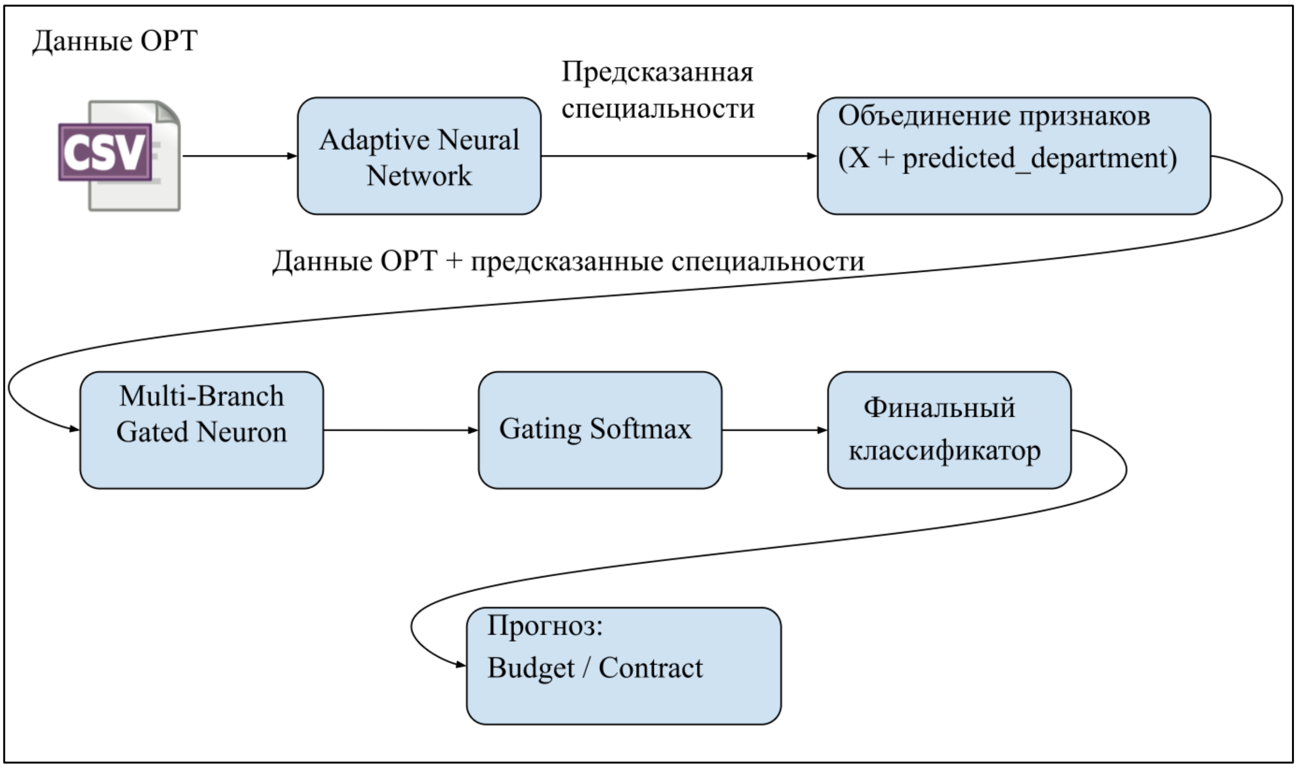

Рисунок 2.6 – Схема конвейерной модели.

Входные данные

На вход конвейера подаются числовые признаки абитуриента:

- результат основного теста ОРТ (математика, физика, химия, биология, английский язык)
- результаты профильных предметов
- категориальный признак региона

Все признаки предварительно масштабируются с помощью StandardScaler

Модуль 1. Нейронная сеть с адаптивной функцией активации

Первый модуль архитектуры решает задачу многоклассовой классификации специальности.

Основные элементы:

- полносвязные слои
- разработанная автором адаптивная функция активации, представляющая собой выпуклую комбинацию *sigmoid* и *tahn* с обучаемым коэффициентом *a*
- выходной слой *softmax*

Результат модуля:

- вероятностное распределение по всем специальностям
- выбранный класс специальности (*argmax*)

Предсказанное значение специальности добавляется как дополнительный информативный признак.

Формирование расширенного признакового пространства. К исходным масштабированным числовым признакам добавляется:

- предсказанная специальность (целочисленный класс)

Таким образом, конвейер увеличивает число признаков и позволяет второй стадии использовать знания первой модели. Полученная матрица объекты-признаки передается в *MBGN*.

Модуль 2. Модель MBGN

На втором этапе применяется архитектура *MBGN*, предназначенная для извлечения нелинейных зависимостей между признаками. *MBGN* состоит из 3-х параллельных ветвей:

1. *Branch 1* – нелинейная трансформация исходных числовых признаков
2. *Branch 2* – альтернативная нелинейная трансформация тех же признаков
3. *Branch 3* – комбинированная функция, обрабатывающая все признаки вместе

Каждая ветвь формирует свое скрытое представление о данных.

Gating механизм

Все три скрытых представления подаются в *gating*-слой, генерирующий веса $(\square_1, \square_2, \square_3)$ через *softmax*. Эти веса определяют вклад каждой ветви в финальное решение: $\square = \square_1\square_1 + \square_2\square_2 + \square_3\square_3$.

Такой подход адаптивно выбирает оптимальную комбинацию ветвей для каждого объекта.

Финальный выход

Полученное агрегированное представление подается в последний линейный слой с сигмоидой, который выдает прогноз вероятности обучения на бюджетной основе:

$p(budget) = \sigma(Wy + b)$. После порога *0.5* получается бинарная классификация (*budget/contract*)

ГЛАВА 3. РЕЗУЛЬТАТЫ СОБСТВЕННЫХ ИССЛЕДОВАНИЙ

3. 1 Набор данных и предварительный анализ

Набор данных состоит из 7315 записей, объединяющих абитуриентов поступивших, как на бюджетную, так и на контрактную основы обучения. Данные охватывают весь спектр специальностей и категорий тестирования, включая результаты основного теста ОРТ, дополнительные предметы и информацию о выбранных направлениях подготовки.

В данном разделе представлены ключевые закономерности, выявленные в ходе анализа данных, позволяющие установить связи между результатами тестирования абитуриентов, выбранными специальностями и формой обучения.

Установленные закономерности в данных

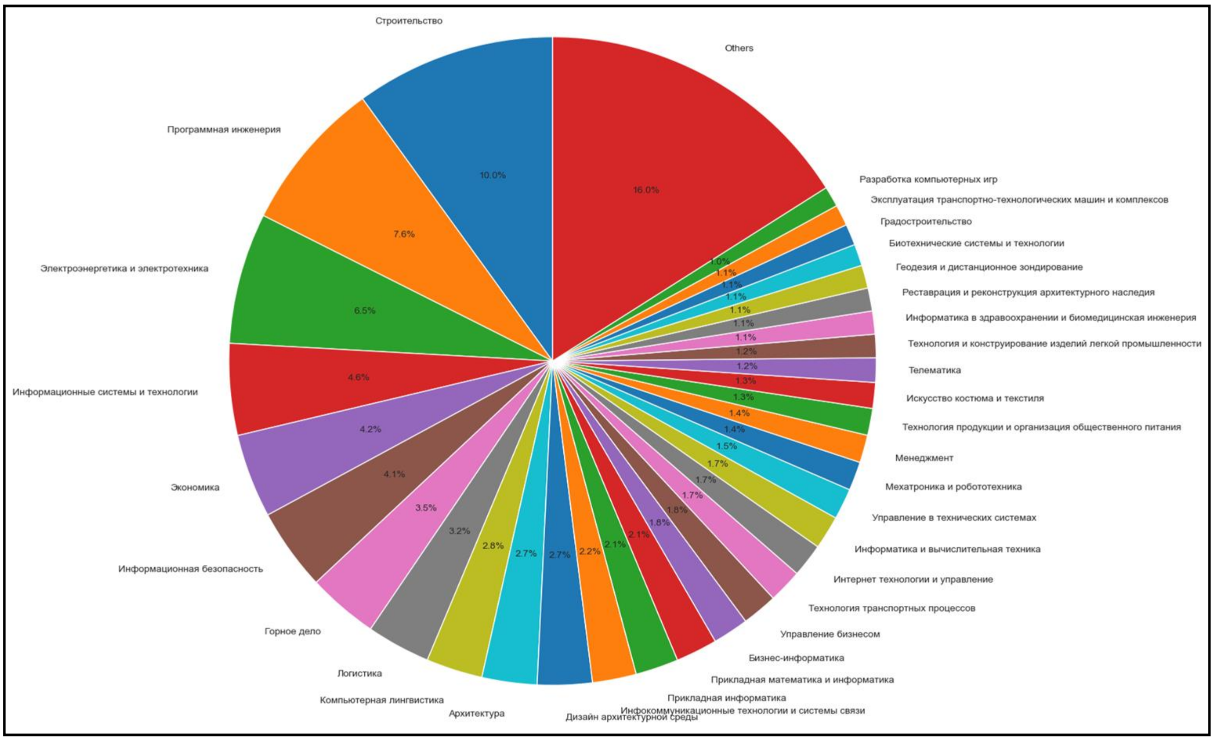

Рисунок 3.1 – Распределение специальностей за 2024 и 2025 годы

На рисунке 3.1 представлено распределение специальностей по доле поступивших абитуриентов за 2024 и 2025 год. Наибольшую долю занимает категория **Others** (16%), представляющая сумму всех направлений, представляющих менее массовые направления. Среди отдельных специальностей наибольшими являются **Строительство** (10%), **Программная инженерия** (7.6%), Информационные системы и технологии (4.6%), **Экономика**

(4.2%) и **Информационная безопасность** (4.1%). Заметную, но меньшую долю занимают специальности такие, как **Горное дело**, **Логистика**, **Компьютерная лингвистика**, **Архитектура** и другие специальности, каждая из которых набирает от 1 до 3,5 % от общего числа абитуриентов.

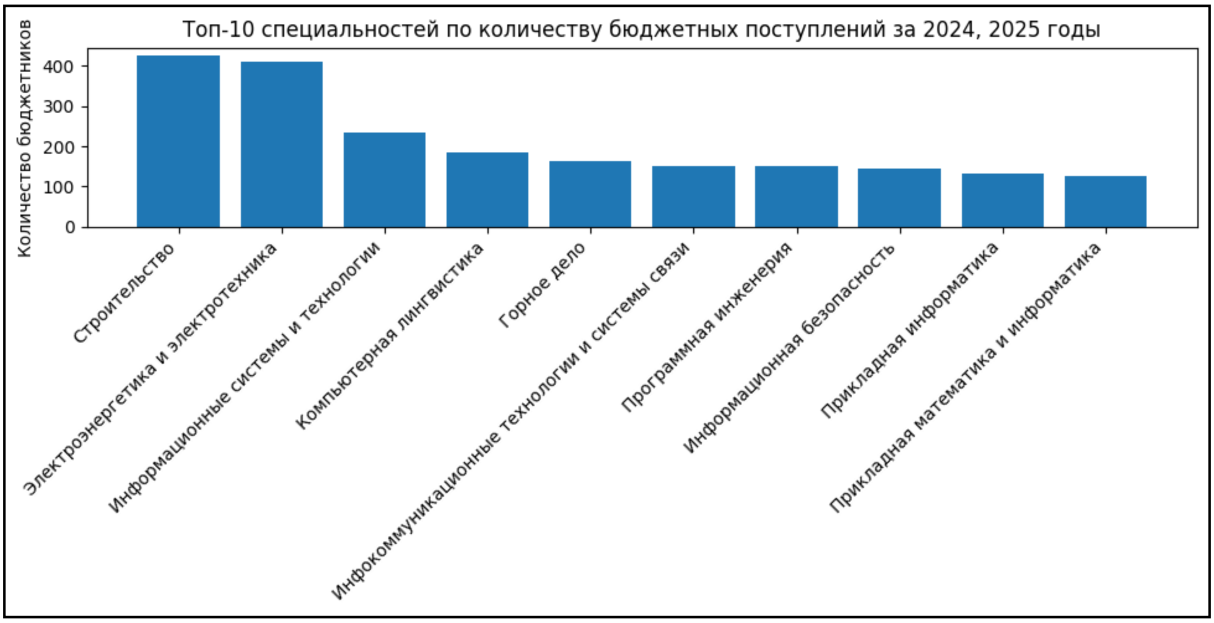

Рисунок 3.2 – Распределение специальностей по кол-ву бюджетных мест за 2024 и 2025 годы.

На рисунке 3.2 представлено распределение абитуриентов, поступивших на бюджет, по 10 наиболее популярным специальностям. Наибольшее количество бюджетников сосредоточено в направлениях: **Строительство** (425 человек) и **Электроэнергетика и электротехника** (410 человек). Значительное число студентов также поступило на специальности, связанные с информационными технологиями: **Информационные системы и технологии** (234), **Компьютерная лингвистика** (185), **Инфокоммуникационные технологии и системы связи** (152) и **Программная инженерия** (149). Остальные специальности из топа - 10 набирают от 123 до 163 студентов.

На рисунке 3.3 представлено распределение результатов Национального теста ОРТ среди студентов, поступивших на бюджетную и контрактную основы обучения. Гистограмма показывает, что большинство бюджетников (голубой цвет) имеют более высокие баллы ОРТ по сравнению с теми, кто поступил на контракт. Пик распределения бюджетных абитуриентов сосредоточен примерно в диапазоне 140–180 баллов, в то время как контрактники имеют более широкий разброс и смещенный в левую сторону центр нормального распределения.

Это подтверждает тенденцию, что доступ к бюджетным местам требует более высокого уровня подготовки абитуриентов и отражает конкуренцию за ограниченные бюджетные места.

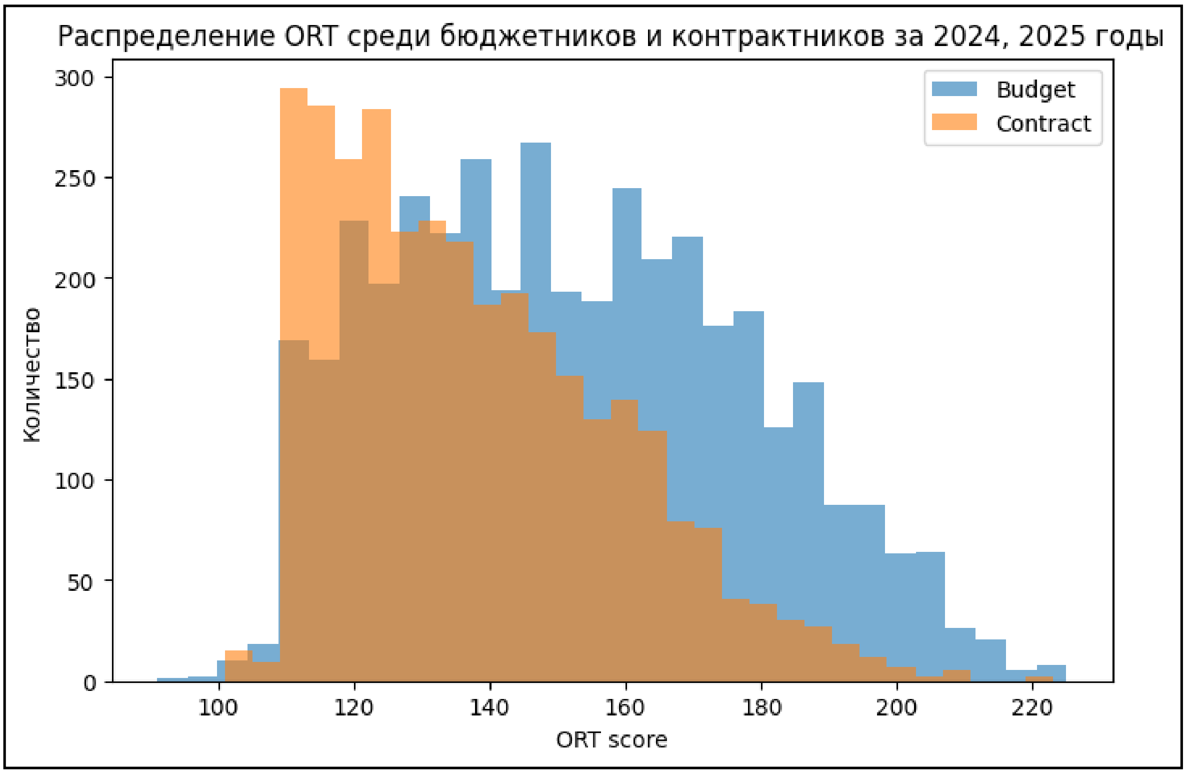

Рисунок 3.3 – Распределение баллов ОРТ специальностей бюджетников и контрактников за 2024 и 2025 годы.

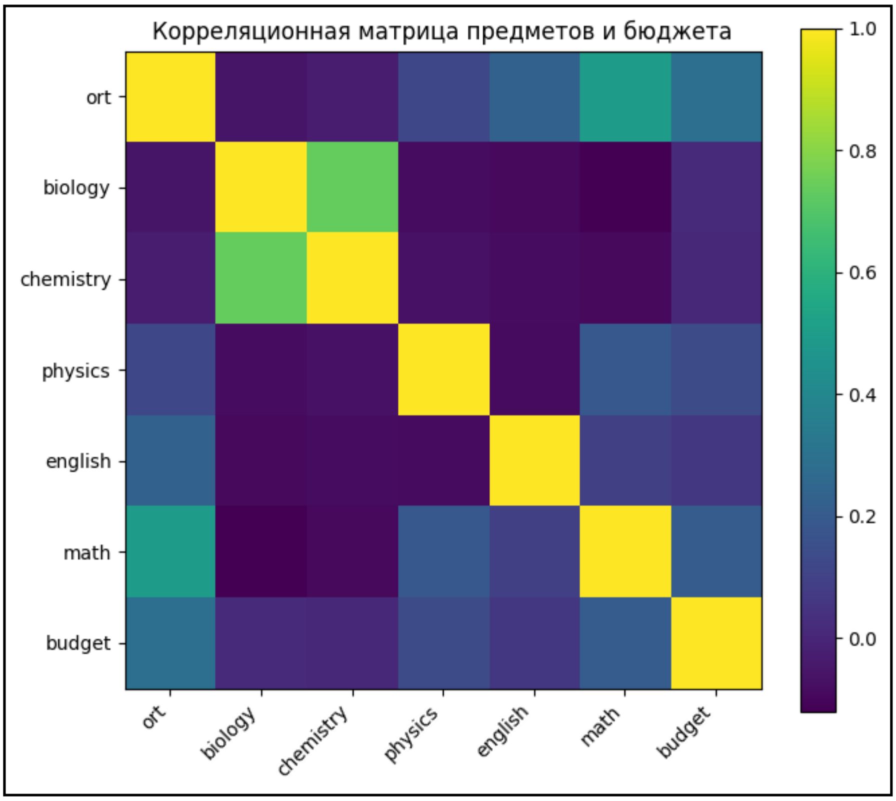

Рисунок 3.4 – Корреляционная матрица предметов и бюджета за 2024 и 2025 годы

Рисунок 3.4 иллюстрирует корреляционную матрицу между баллами по основному тесту, дополнительным предметам (биология, химия, физика, английский, математика) и формой обучения (бюджет/контракт). Анализ матрицы показывает, что наибольшую положительную корреляцию с поступлением на бюджет имеют общие результаты ОРТ ($r \approx 0.289$) и математика ($r \approx 0.205$). Другие предметы демонстрируют слабую корреляцию с формой обучения, что указывает на меньшую значимость этих дисциплин при распределении бюджетных мест. Данный результат подтверждает, что основным критерием при отборе на бюджет являются результаты баллов по общему тесту ОРТ и дополнительном тесте по математике.

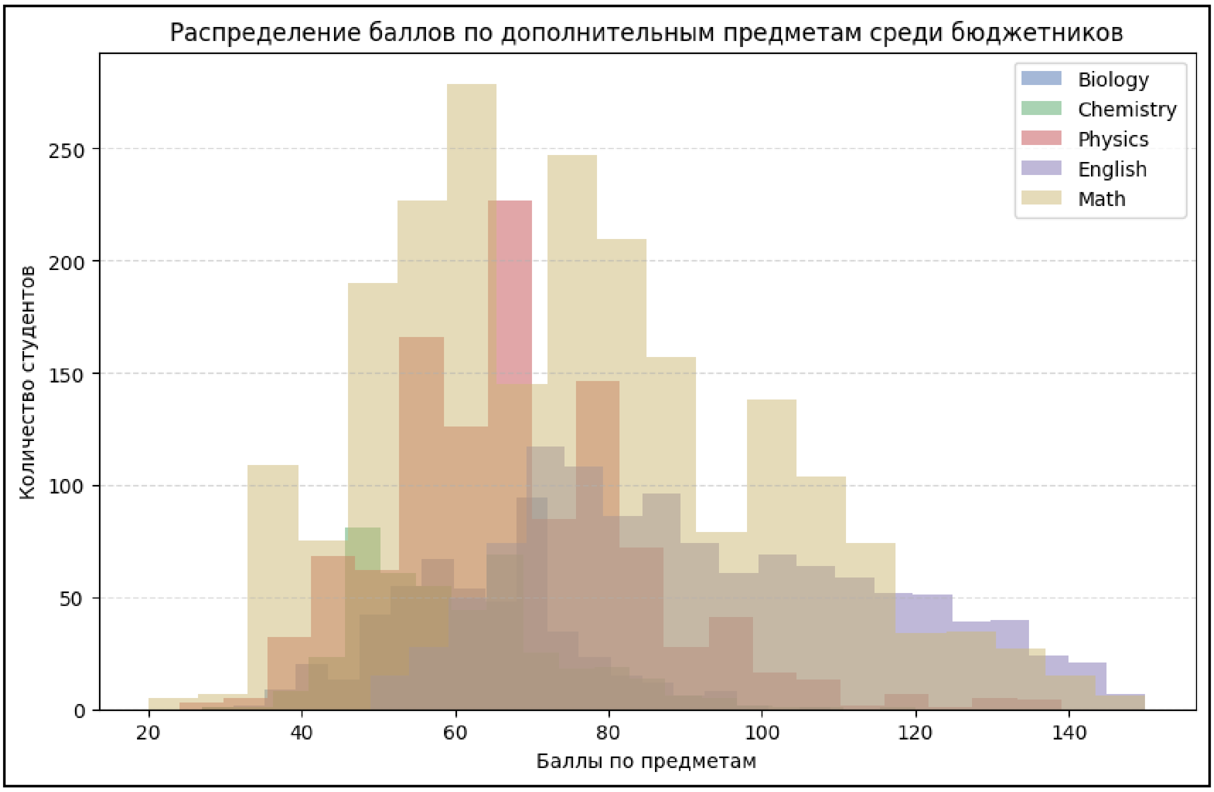

Рисунок 3.5 – Распределение баллов по дополнительным предметам у бюджетников за 2024 и 2025 годы

На рисунке 3.5 показано распределение результатов по пяти дополнительным предметам среди студентов, поступивших на бюджет. По данной гистограмме можно увидеть, что дополнительный предмет математика характеризуется широким разбросом с заметной концентрацией высоких баллов.

- Биология: диапазон баллов от 57 до 90, среднее 69.8 медиана 66.6
- Химия: диапазон 50–83, среднее 62.1, медиана 61
- Физика: диапазон 53–102, среднее 74.4, медиана 72
- Английский: диапазон 63–138, среднее 92.3, медиана 85
- Математика: диапазон 39–113, среднее 82.8, медиана 87

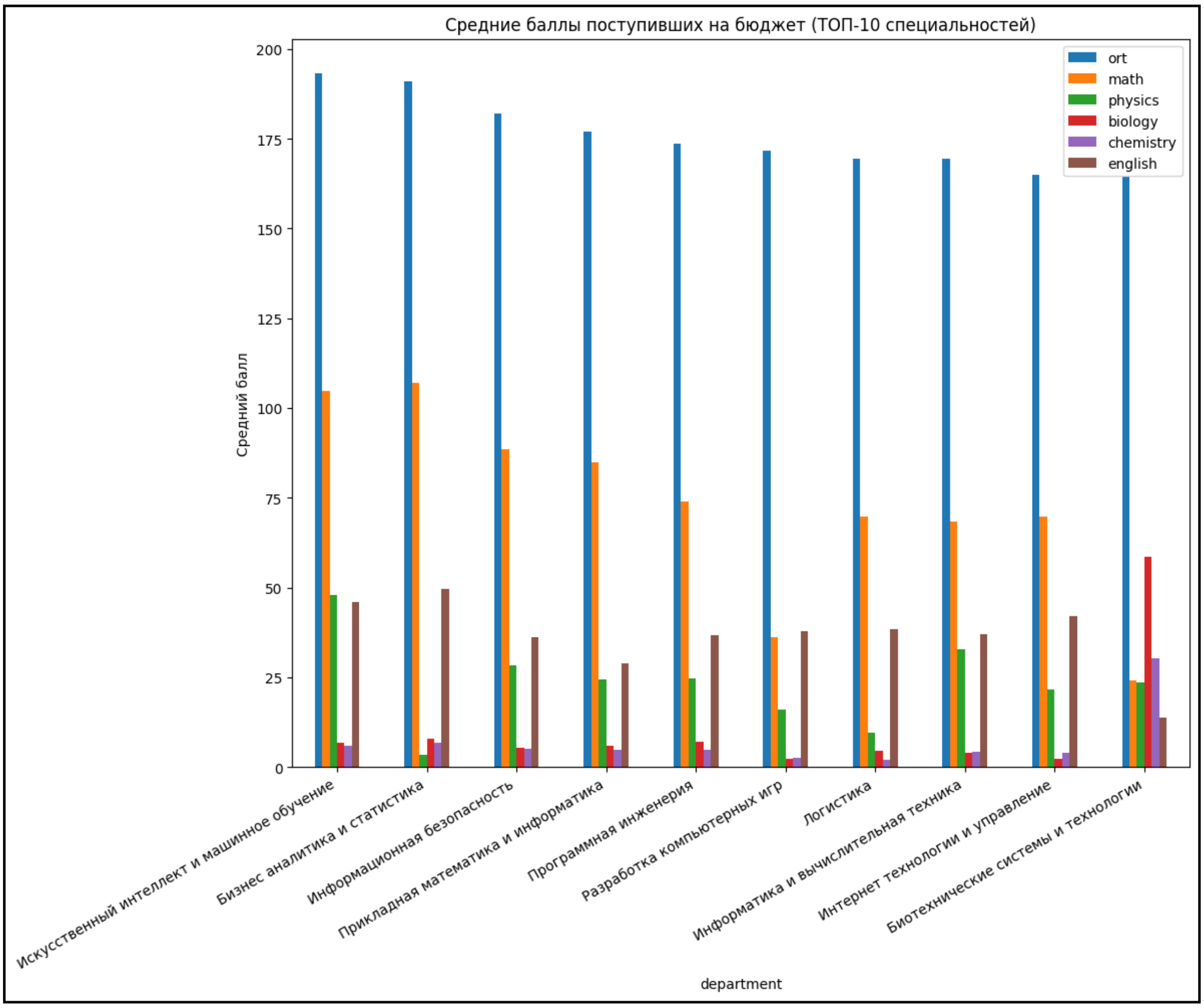

Рисунок 3.6 – Средние баллы у бюджетников по специальностям за 2024 и 2025 годы.

Анализ средних итоговых баллов среди абитуриентов, поступивших на бюджетные места на рисунке 3.6 показал существенное различие в зависимости от выбранной специальности. Наивысшие значения среднего суммарного балла демонстрируют направления, связанные с анализом данных и с искусственным интеллектом. Так, специальность Искусственный интеллект и машинное обучение характеризуется самым высоким средним баллом - 193.2 что указывает на высокую конкуренцию значительную привлекательность данного направления среди сильных абитуриентов. Схожие показатели наблюдаются и по специальности Бизнес-аналитика и статистика (190.9), где также требуются развитые математические компетенции.

Следующую группу формируют направления в области компьютерной безопасности и прикладной математики. Специальность Информационная безопасность показывает средний балл 182.1, а Прикладная математика и информатика - 177, что указывает на устойчивый интерес абитуриентов со сравнительно высокими результатами ОРТ.

Средние показатели наблюдаются в инженерно-технических направлениях Программная инженерия 173.7, Разработка компьютерных игр (171.6), Логистика (169.4) и Информатика и вычислительная техника (169.4), где конкуренция также достаточно заметна, однако уступает направлениям в области анализа данных.

Нижнюю часть распределения занимают Интернет-технологии и управление (164.9), и Биотехнические системы и технологии (164.3) Относительно более низкие значения могут объясняться меньшим конкурсом либо менее выраженным спросом среди абитуриентов с высокими результатами тестирования.

В целом наблюдается отчетливая тенденция: максимальные средние баллы у бюджетников приходятся на направления, связанные с анализом данных, машинным обучением и статистикой, что отражает современный тренд роста интереса к IT-специальностям с уклоном в математическое моделирование и работу с большими данными.

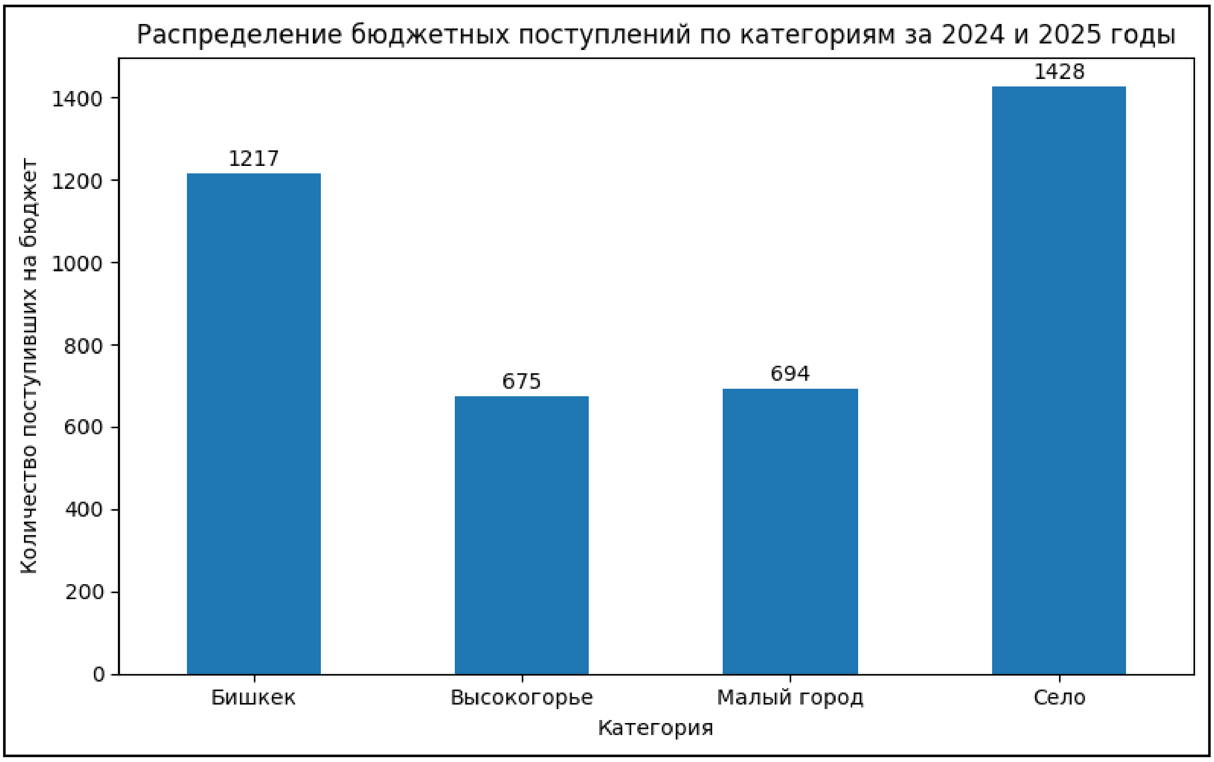

Рисунок 3.7 – Распределение бюджетных поступлений по категориям за 2024 и 2025 годы.

Распределение бюджетных поступлений по категориям показывает, что наибольшее количество абитуриентов, зачисленных на бюджет, приходится на категорию Село (1428 человек), далее следует Бишкек (1217). Значительно меньше поступивших из малых городов (694) и высокогорных районов (675). Это указывает на преобладающий вклад сельских выпускников в общий поток бюджетных абитуриентов.

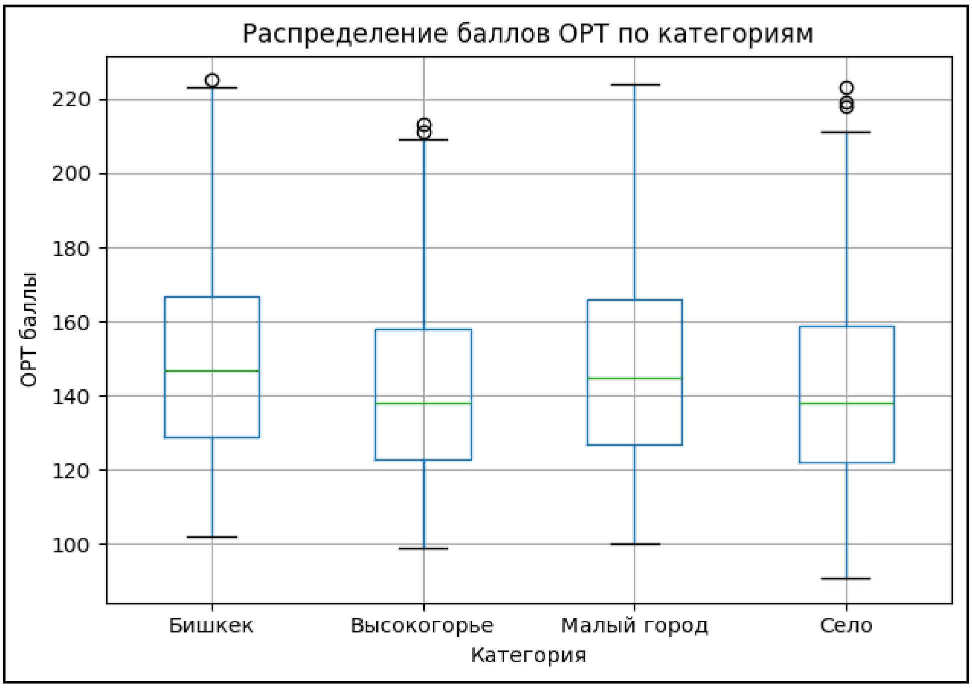

Рисунок 3.8 – Распределение основного балла ОРТ по категориям

График в виде *boxplot* на рисунке 3.8 демонстрирует различия в распределении баллов ОРТ среди абитуриентов из различных категорий населенных пунктов: Бишкек, Высокогорье, Малый город и Село. Категория Бишкек обладает наивысшим средним баллом (149.7) и медианой (147), а также наиболее широким верхним квартилем (Q3=167). Максимальные значения достигают 226, что указывает на присутствие значительной доли абитуриентов с высокими результатами. Схожие показатели наблюдаются и у категории Малый город, где средний балл составляет 158, медиана - 145, а верхний квартиль достигает 166; это свидетельствует о достаточно хорошем уровне абитуриентов проживают в небольших городах, сопоставимом с уровнем столицы.

В то же время категории Высокогорье и Село демонстрируют более низкие значения баллов по ОРТ. В обоих случаях медиана составляет 138, а средние - 142.2 и 142 соответственно. Нижние квартильные значения (Q1=123) указывают на более широкое присутствие низких баллов в данных категориях по сравнения с городскими. Однако максимальное значения в этих группах также достаточно высоки (до 213–223), что говорит о наличии отдельных высоко балльных абитуриентов. В целом *boxplot* отражает отчетливую тенденцию, уровень ОРТ у абитуриентов из городских категорий (Бишкек и малые города) выше, чем у абитуриентов из сельских высокогорных регионов. Это проявляется как в средних значениях, так и в медианах и межквартальных интервалах. Данные могут указывать на различия в доступе к качественному образованию и подготовке среди разных территориальных групп.

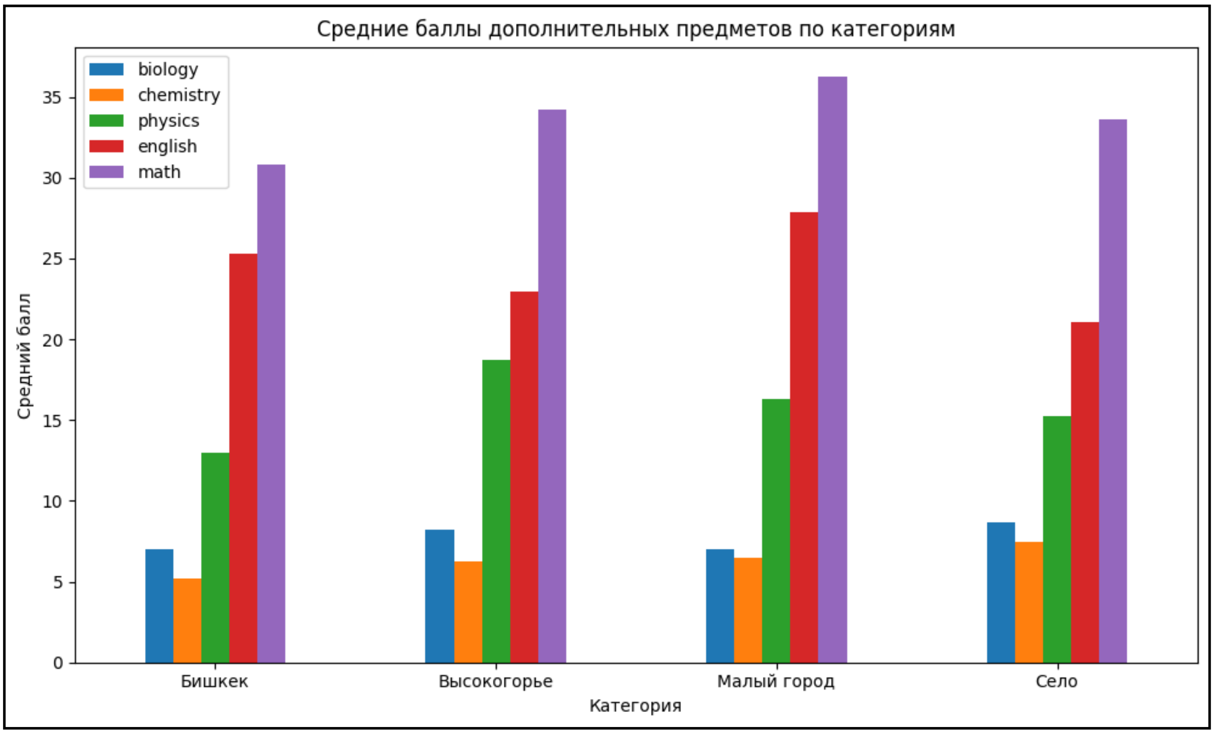

Рисунок 3.9 – Средние баллы дополнительных предметов по категориям за 2024 и 2025 годы.

Как можно видеть на рисунке 3.9: наиболее высокие показатели по математике получают абитуриенты из малых городов (36.3) и высокогорья (34.2), тогда как в категориях Бишкек и Село значения средних баллов по дополнительным предметам ОРТ составляют 30.8 и 33.6 соответственно. По английскому языку лидируют малые города (27.9) и Бишкек (25.3). В то же время по биологии и химии более высокие средние баллы характерны для села (8.6 и 7.5) и высокогорья (8.2 и 6.23). По физике максимальный средний балл наблюдается в высокогорных районах (18.7), затем следуют малые города (16.3) и село (15.22). Такие различия указывают на неоднородность подготовки абитуриентов в зависимости от категории населенного пункта.

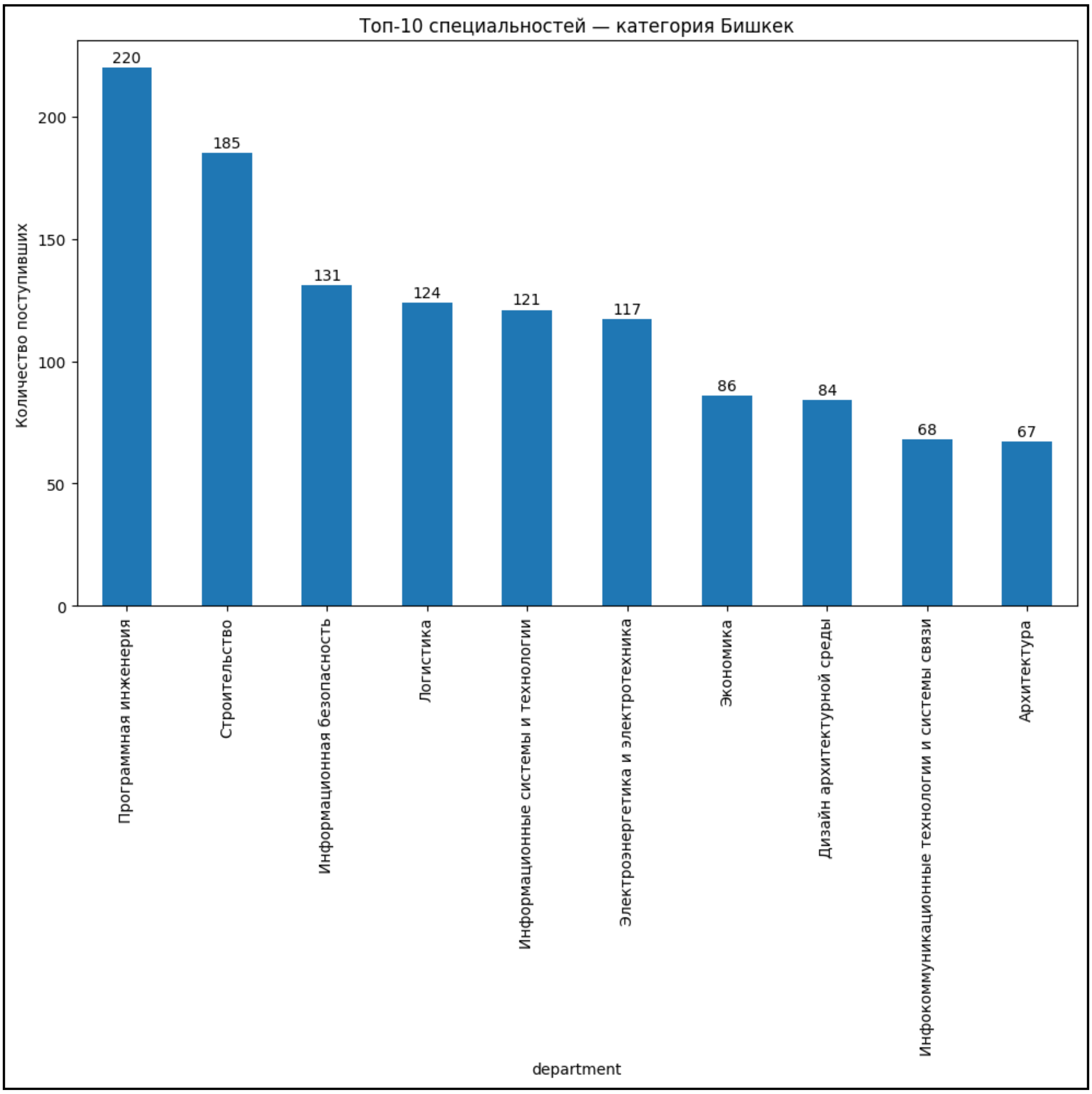

Рисунок 3.10 – Топ специальностей в категории: Бишкек за 2024 и 2025 годы

Анализ распределения специальностей среди абитуриентов категории Бишкек на рисунке 3.10 показал, что наибольшей популярностью пользуются направления, связанные с информационными технологиями и инженерно-техническими профилями. Лидером является специальность Программная инженерия (220) поступивших, что отражает растущий спрос на IT-кадры.

На втором месте – Строительство (185 человек), что подтверждает устойчивый интерес к инженерно-строительной отрасли; Высокие позиции также занимают направления Информационная безопасность (131) и Логистика

(124), демонстрирующая расширение образовательного спроса в областях цифровой безопасности и управления цепями поставок.

В целом структура поступлений для категории Бишкек показывает ярко выраженный интерес к ИТ, инженерным и техническим специальностям, что соответствует современной структуре спроса на рынке труда.

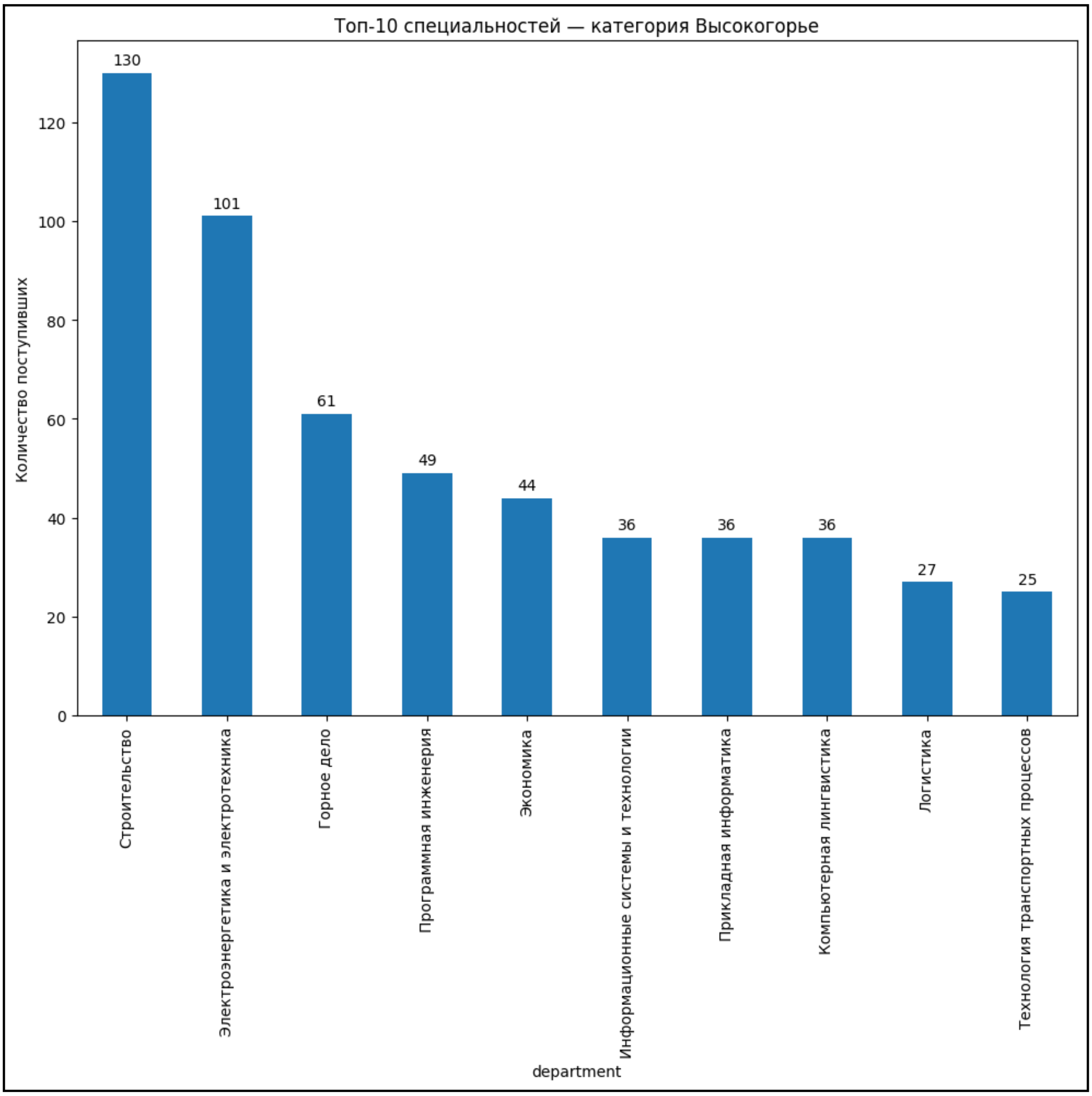

Рисунок 3.11 – Топ специальностей в категории: Высокогорье за 2024 и 2025 годы

Распределение специальностей среди абитуриентов категории Высокогорье на рисунке 3.11 показывает выраженную ориентацию на инженерно-технические и инфраструктурные направления. Лидирующую позицию занимает специальность Строительство (13) поступивших, что может отражать потребность регионов в развитии строительно-дорожной инфраструктуры; на втором месте находится Электроэнергетика и электротехника (101 человек), также являющаяся ключевой для высокогорных территорий. Существенно выделяется интерес к Горному делу (61), что соответствует экономической специфике данных регионов;

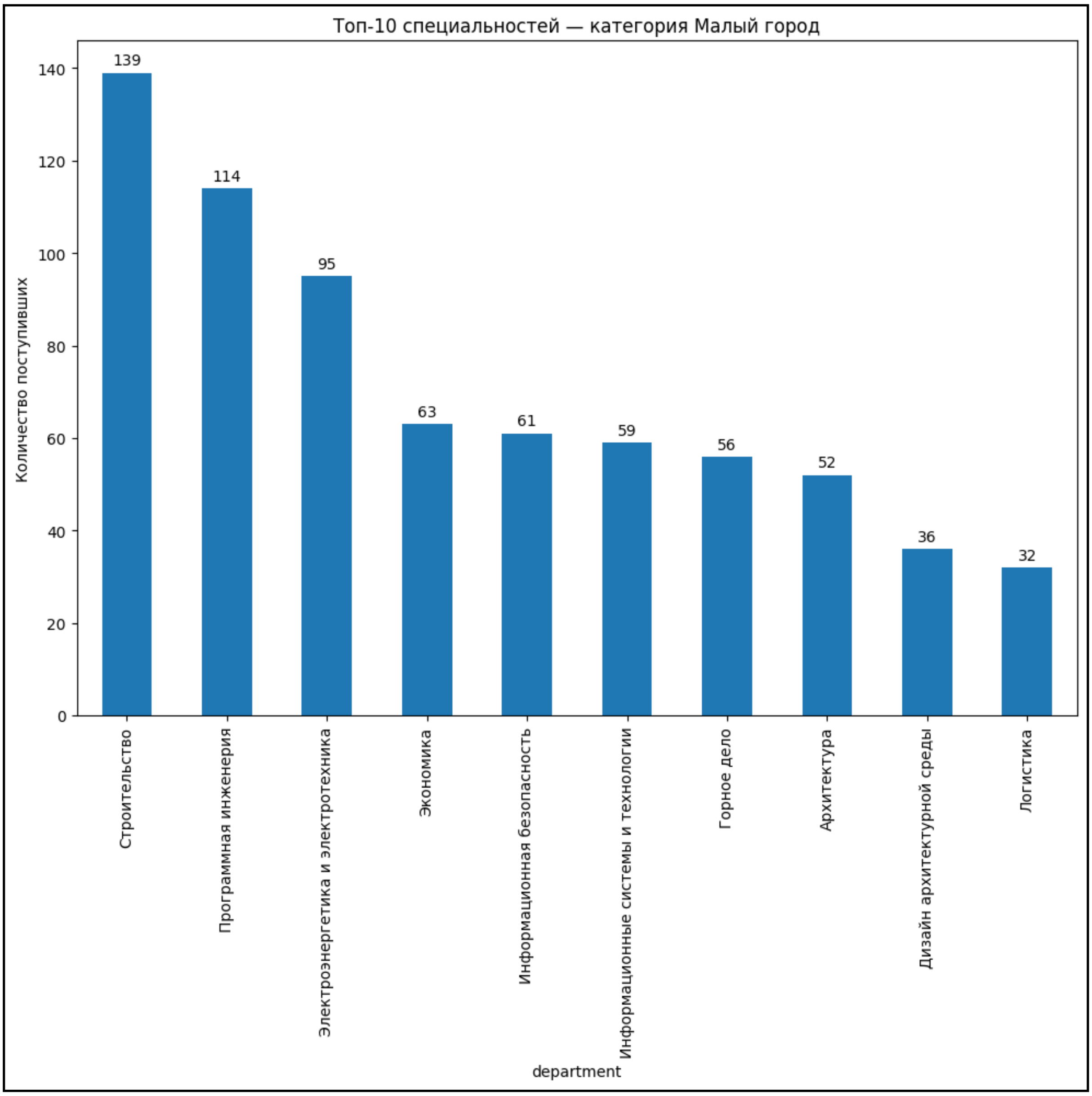

Рисунок 3.12 – Топ специальностей в категории: Малый город за 2024 и 2025 годы.

Лидирующие позиции занимают инженерно-технические направления: строительство – 139 человек, программная инженерия – 114, электроэнергетика и электротехника – 95. Далее следуют специальности экономического и IT-профиля: экономика – 63, информационная безопасность – 61, информационные

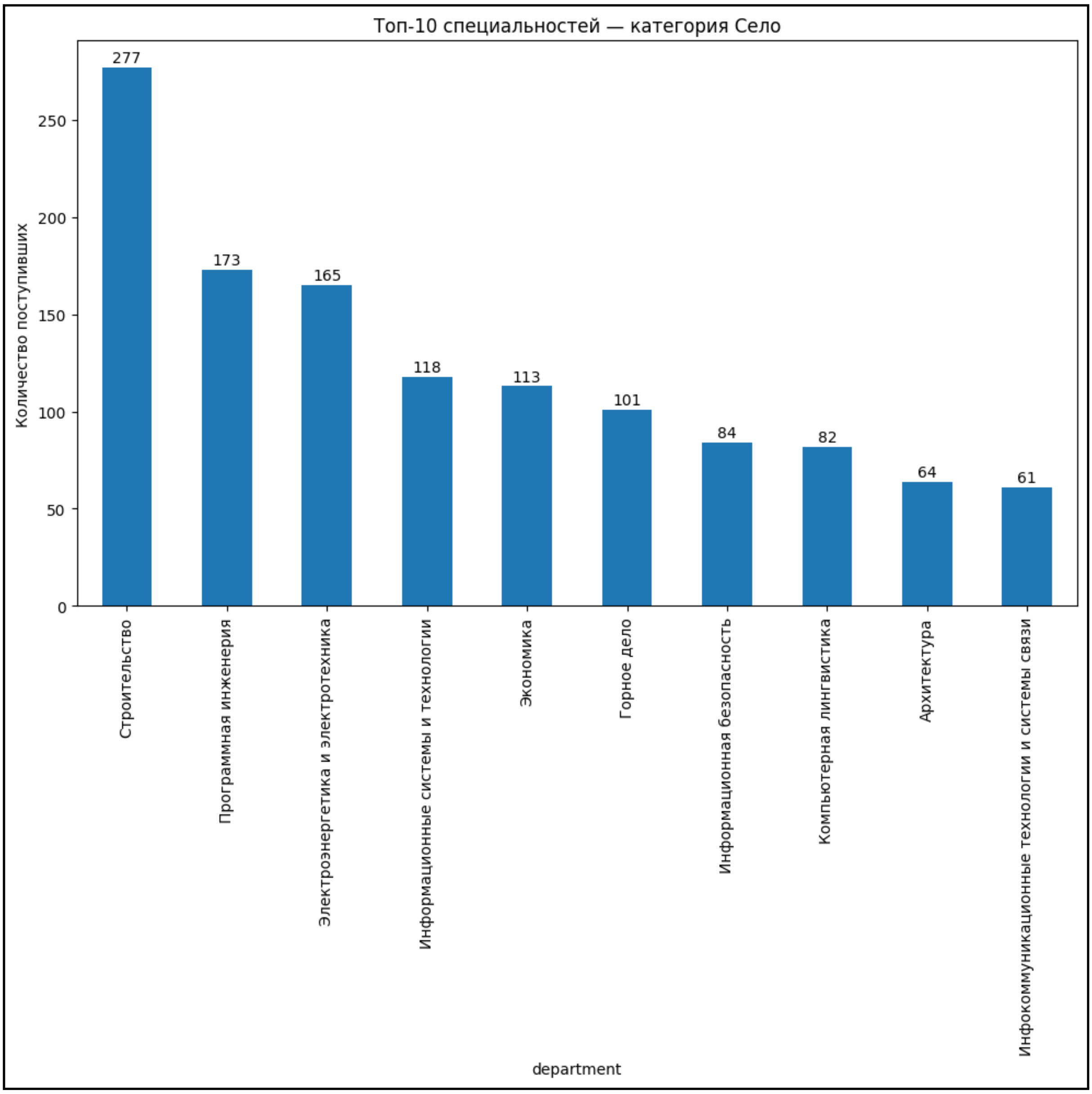

Рисунок 3.13 – Топ специальностей в категории: Село за 2024 и 2025 годы

системы и технологии – 59. Менее выраженный интерес отмечается к программам горное дело (56), архитектура (52), дизайн архитектурной среды (36) и логистика (32). Заметная доля IT-направлений свидетельствует о постепенном смещении интересов абитуриентов в сторону цифровых компетенций и современных технологий.

Для категории Село распределение поступающих демонстрирует наиболее выраженную ориентацию на инженерно-технические специальности.

Лидером является направление «Строительство», на которое поступили 277 абитуриентов, что значительно превышает показатели других категорий. Далее следуют «Программная инженерия» (173 человека) и «Электроэнергетика и электротехника» (165 человек), подтверждая высокий интерес сельских абитуриентов к техническим и цифровым профессиям. Существенную долю также занимают «Информационные системы и технологии» (118 человек) и «Экономика» (113 человек). Значимым остается интерес к прикладным направлениям – «Горное дело» (101 абитуриент), «Информационная безопасность» (84), «Компьютерная лингвистика» (82) и «Архитектура» (64). Сравнительно высокий показатель у специальности «Инфокоммуникационные технологии и системы связи» (61 абитуриент) также подчеркивает тенденцию к выбору направлений, связанных с цифровой инфраструктурой. Это распределение отражает устойчивый спрос сельских абитуриентов на специальности, обеспечивающие доступ к стабильным техническим и инженерным профессиям.

Таблица 3.1 – Выбор специальностей абитуриентами в разрезе категорий населенных пунктов

	Бишкек	Высокогорье	Малые города	Село
ПИ	+	+	+	+
Строительный	+	+	+	+
ИБ	+	-	+	+
Логистика	+	+	+	+
ИСТ	+	+	+	+
Электроэнергетика и электротехника	+	+	+	+
Экономика	+	+	+	+

Дизайн архитектурной среды	+	-	+	-
ИТСС	+	-	-	+
Архитектура	+	-	+	+
Горное дело	-	+	+	+
Прикладная информатика	-	+	-	-
Комп. линг.	-	+	-	+
Техн. транс. процессов	-	+	-	-

Таблица 3.1 отражает распределение выбора специальностей абитуриентами в зависимости от категории населенного пункта. Видно, что технические направления такие как «Программная Инженерия», «Строительство» и «Электроэнергетика и электротехника» – стабильно выбираются во всех регионах. Вместе с тем некоторые специальности имеют выраженную региональную специфику: например, «Горное дело» чаще выбирают абитуриенты из высокогорья и сельской местности, тогда как «Дизайн архитектурной среды» заметно реже представлен в этих категориях. Таким образом, таблица демонстрирует как общее ядро популярных направлений, так и отличия, обусловленные социально-географическим контекстом.

Результаты экспериментов

Чтобы оценить эффективность предложенных архитектур, был проведен сравнительный эксперимент на реальном наборе данных абитуриентов. Задача состояла в том, чтобы предсказать, будет ли студент зачислен на стипендию (бюджет) или на самофинансирование (контракт), используя, как числовые оценки за экзамен, так и категориальные признаки.

Числовые признаки были стандартизированы, а категориальные признаки были закодированы с помощью *One-Hot* энкодера. Модель *MBGN* сравнивалась с традиционной базовой нейронной сетью, реализованной с помощью *Keras*, а также с предложенной *Keras Adaptive Activation*. Обе модели были обучены с использованием бинарной перекрестной энтропии потерь и оценены на основе потерь при валидации, точности и *AUC* (арена под *ROC*-кривой).

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 доля верных ответов (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 3.14

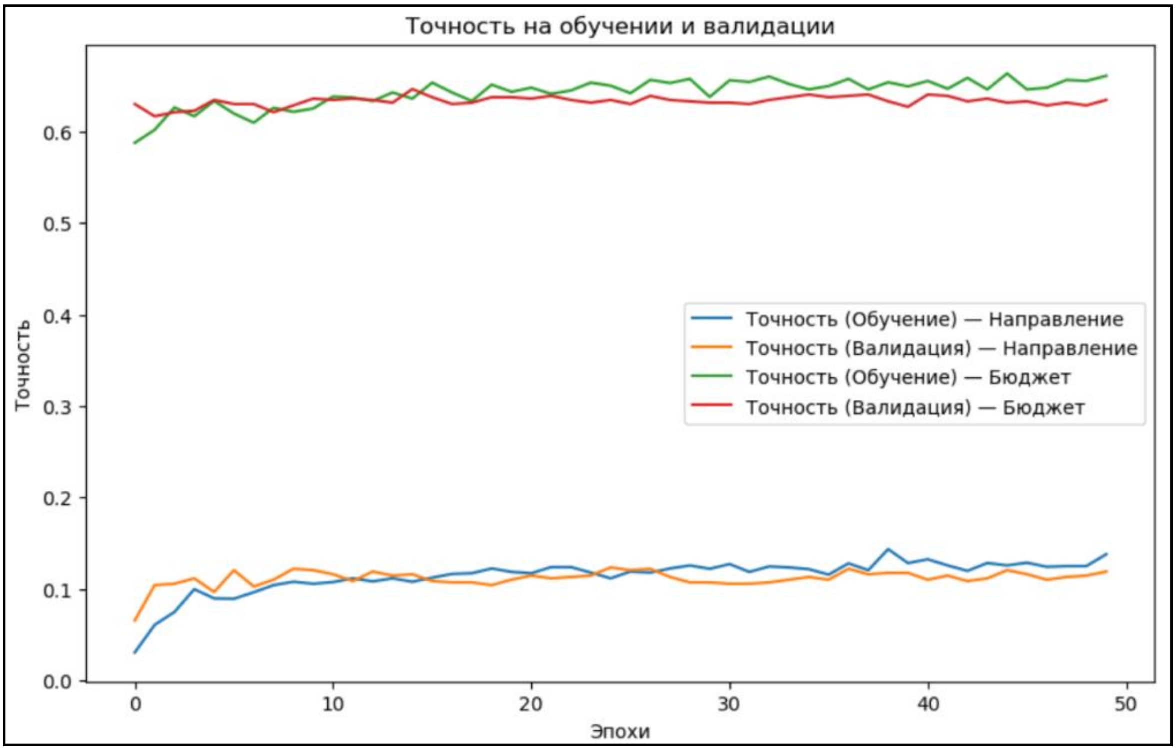

Рисунок 3.14 – График точности на обучении и валидации

Результаты

Графики зависимости точности и потерь от количества эпох показывают, 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.

На графике ошибки (рис. 3.15) наблюдается тенденция к снижению, однако полное приближение к нулевому значению не достигается. Это свидетельствует о том, что доступного объема данных недостаточно для полного обучения модели и минимизации ошибки до приемлемого уровня.

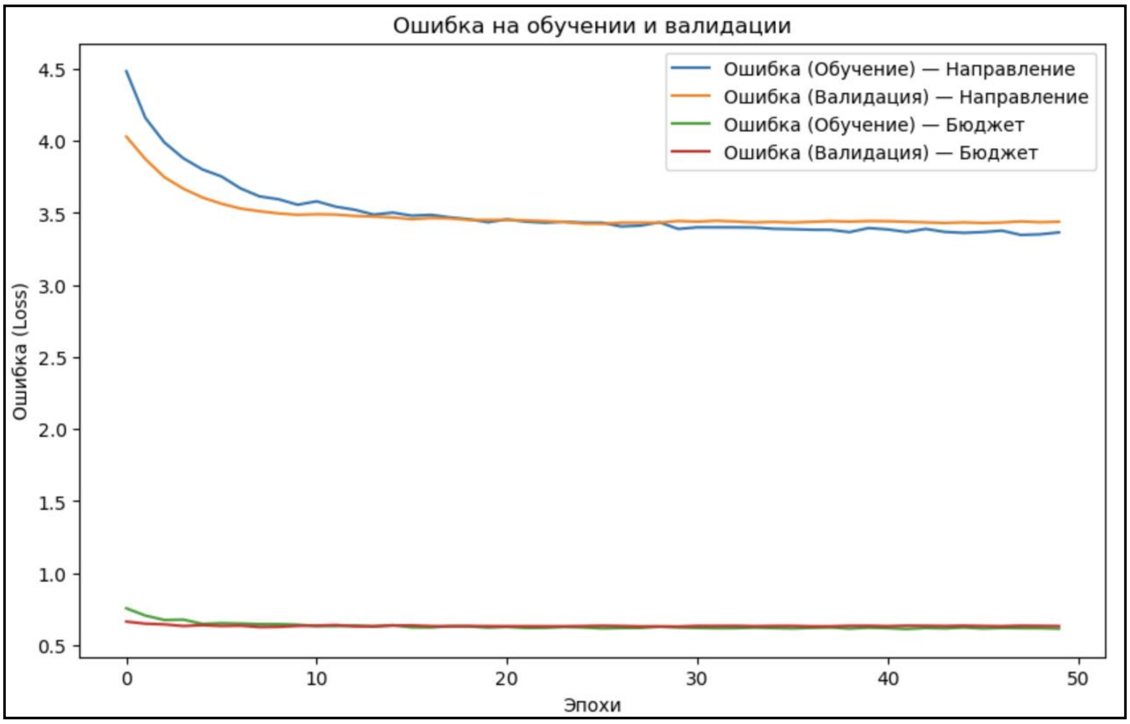

Рисунок 3.15 – График ошибки на обучении и валидации

Результаты эксперимента

Анализ экспериментальных данных показал, что внедрение адаптивной функции активации способствует повышению точности классификации направления подготовки по сравнению с базовой моделью, использующей функцию ReLU:

- Базовая модель (ReLU): 10.70%
- Модель с Adaptive Activation: 11.89%

Абсолютное улучшение составило 1.19 процентных пункта, а относительное – около 11% по сравнению с исходной моделью. Полученные результаты показывают, что адаптивная функция активации обладает потенциалом к более точному моделированию сложных взаимосвязей в данных, благодаря динамической подстройке активации нейронных слоев под характерные шаблоны входной информации.

Однако, несмотря на полученное улучшение прогноза, абсолютное значение точности остается на низком уровне, что указывает на сложность задачи классификации и ограниченность текущего объема данных. Это говорит о необходимости дальнейшего совершенствования модели, в том числе за счет расширения датасета и повышения его репрезентативности.

Адаптивная функция активации влияет на повышение точности модели за счет комбинирования двух различных функций – сигмоиды и гиперболического тангенса. Это позволяет каждому нейрону самостоятельно выбирать, какая функция активации лучше подходит для текущего набора входных данных. Для разных задач можно комбинировать подобным образом и другие функции активации.

Преимущества *AdaptiveActivation* функции. Активационная функция с адаптивными параметрами обладает преимуществвами по сравнению с традиционными функциями активации. Главным преимуществом является **гибкость**: нейроны могут адаптироваться к различным типам входных данных, и этим повышать универсальность и точность модели.

Уменьшение недостатков традиционных функций активации:

- Функция *ReLU* нередко приводит к появлению «мёртвых нейронов», когда отрицательные значения полностью обнуляются, из-за чего такие нейроны фактически престают обучаться
- Сигмоидальная и tanh-активации подвержены ослаблению градиентов, что особенно сильно проявляются в глубоких нейронных архитектурах
- Адаптивная функция активации позволяет комбинировать преимущества этих подходов и существенно смягчает их ограничения, обеспечивая более устойчивое и надёжное распространение градиента по сети.

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 увидеть, что *AdaptiveActivation* помогла модели лучше различать классы за счет более гибкой трансформации входных признаков. Это свойство особенно полезно в задачах, где классические функции активации могут быть недостаточно выразительными.

Показатели эффективности

В таблице ниже приведены результаты проверки обеих моделей:

Модель	Окончательная потеря в валидации	Наилучшая точность *Val*	*Best Val AUC*
MBGN	0.3759	83.90%	0.9091
Базовая	0.6301	64.34%	-

Таблица 3.2 – Сравнение производительности MBGN с базовой моделью на валидационном множестве

Как показано в таблице 3.2, модель *MBGN* значительно превзошла базовую модель по всем доступным метрикам. Она достигла потери при окончательной проверке 0.3759, лучшей точности 83.89 % и AUC 0.9091, что свидетельствует о высокой дискриминационной эффективности. Для сравнения, базовая модель достигла лишь 64.34 % точности и имела постоянно более высокие потери.

7.3 Динамика обучения

На рисунке 3.16 показаны потери при обучении и проверке по эпохам для обеих моделей. Модель *MBGN* демонстрирует более быструю сходимость и меньшую перестройку по сравнению с базовой моделью.

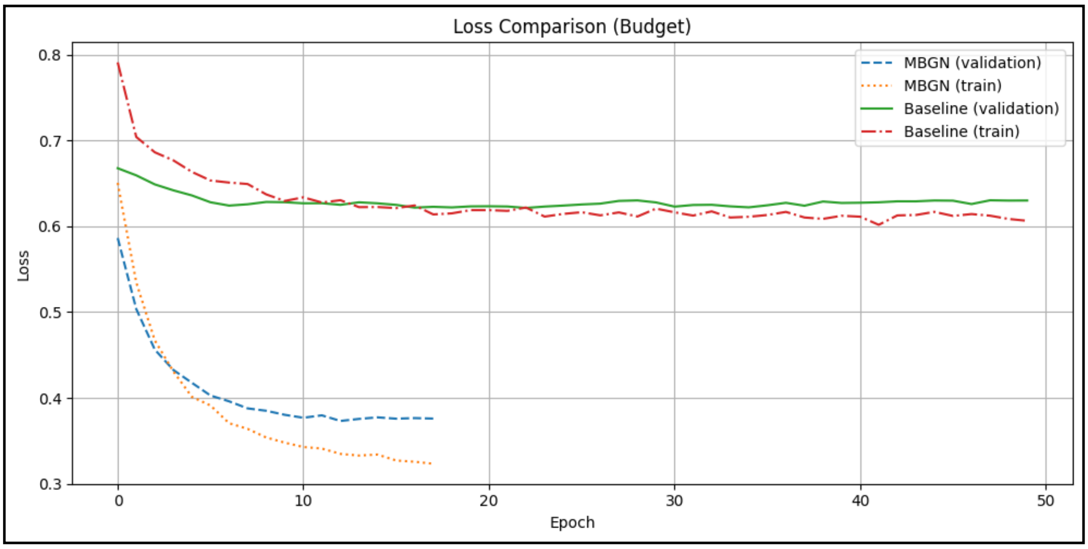

Рисунок 3.16 – Сравнение потерь между MBGN и базовой моделью во время обучения.

На рисунке 3.17 показано сравнение потерь с *Adaptive Activation*. Можно отметить чёткий контраст между моделью *MBGN* и моделью с *Adaptive Activation*, поведение которой близко к результатам *Baseline*.

В то же время модель с Adaptive Activation, аналогично Baseline, остаётся на значительно более высоком уровне потерь (~0.62–0.63) и после первых эпох не демонстрирует заметного улучшения. Несмотря на то, что по точности классификации специальностей Adaptive Activation даёт небольшой прирост (около 11% относительно Baseline), в задаче предсказания формы поступления её кривые практически повторяют динамику базовой модели.

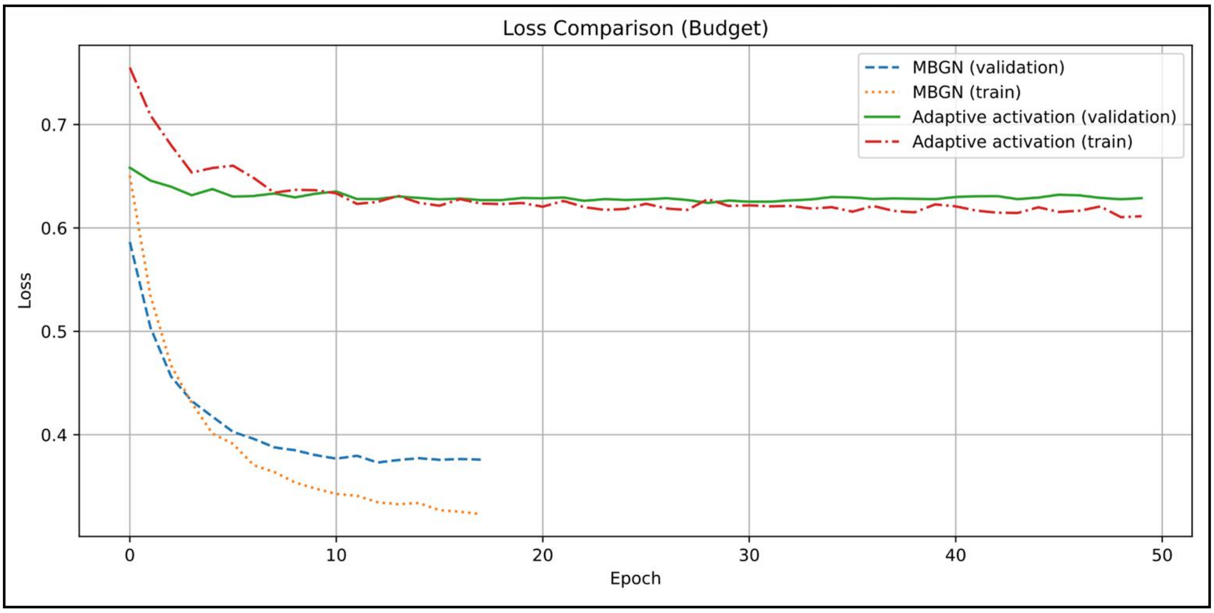

Рисунок 3.17 – Сравнение потерь между MBGN и Adaptive Activation во время обучения.

7.4 Матрица несоответствий

Матрица несоответствий для MBGN, представленная на рисунке 3.18, демонстрирует ее сильную способность разделять случаи стипендии и контракта, с высоким количеством истинно положительных и истинно отрицательных результатов.

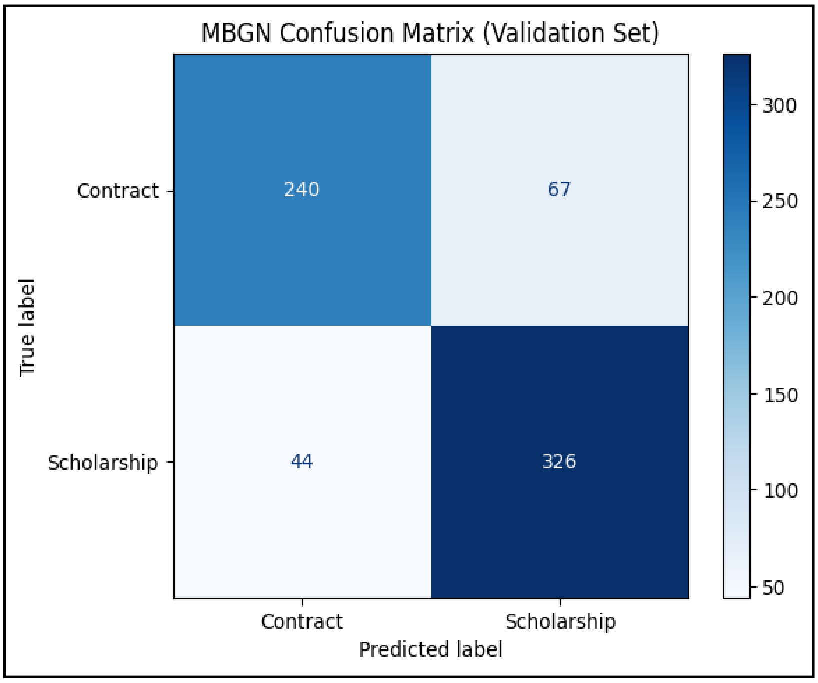

Рисунок 3.18 – Матрица несоответствий для MBGN на валидационном множестве.

Матрица несоответствий (рисунок 3.18) демонстрирует, что модель MBGN достигла высоких и сбалансированных результатов классификации. Из 677 образцов для проверки были правильно классифицированы 240 претендентов на контракт и 326 претендентов на стипендию, что привело к общей точности в 83,6 %. Точность составляет 82,9 %, это означает, что большинство предсказаний стипендий были верными, в то время как recall составляет 88,1 %, что указывает на низкий процент пропущенных случаев стипендий. Показатель F1, равный 85,4 %, подчеркивает сильное гармоническое среднее между точностью и отзывом.

7.5. Сравнительный анализ моделей на задаче прогнозирования формы обучения

Для прогнозирования формы обучения (*budget/contract*) были проведены серии вычислительных экспериментов, направленных на оценку эффективности предложенной комбинированной нейросетевой архитектуры *MBGN*. Для оценки качества использовалась итоговая доля верных ответов (*accuracy*) на тестовой выборке. В качестве базовых методов сравнения были выбраны классические алгоритмы машинного обучения: опорные векторы (*SVC*), метод ближайших соседей (*KNN*) и дерево решений (*Decision Tree*).

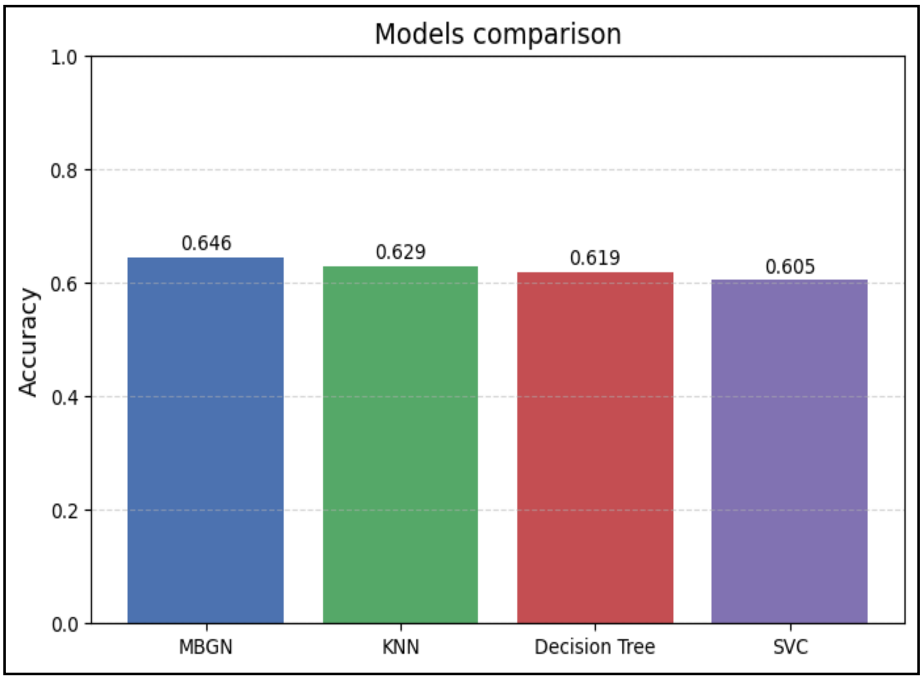

Рисунок 3.19 – Доля верных ответов на задаче прогнозирования формы обучения (*budget/contract*)

Результаты сравнения, представленные на рисунке 3.19 показывают существенные различия в качестве классификации между исследуемыми моделями.

Полученные значения *accuracy* составили:

1. Комбинированная модель (*Adaptive NN* → *Softmax* → *MBGN*): 0.6459
2. KNN: 0.6287
3. Decision Tree: 0.6186
4. SVC: 0.6048

Вывод. Комбинированная модель, включающая адаптивную функцию активации, *softmax-module* и *MBGN*, показала наилучшую точность (около 64.6%). Данный результат подтверждает гипотезу о том, что глубинные архитектуры способны эффективно моделировать сложные нелинейные зависимости между признаками абитуриентов, возникающие в процессе выбора формы обучения. Включение дополнительного softmax-модуля, ответственного за промежуточную классификацию направления подготовки, позволило повысить устойчивость модели за счет учета скрытой семантики специальностей. Использование *MBGN*-гейтинга обеспечило динамическое усиление значимых признаков и подавление шумовых, что также способствовало итоговому росту точности. Ее преимущество перед классическими моделями составляет порядка *2.7–4* процентных пунктов, что является статистически и практически значимым. Полученные результаты подтверждают целесообразность использования нейросетевых архитектур с динамическими механизмами активации и гейтинга для решения задач данного класса.

ГЛАВА 4. РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ РЕКОМЕНДАЦИЙ

На основе разработанной модели и функции активации была реализована информационная система поддержки поступления абитуриентов, которая позволяет использовать входные данные (баллы ОРТ и доп. категориальные признаки) для прогнозирования вероятного распределения абитуриентов по специальностям и формам обучения (бюджет/контракт).

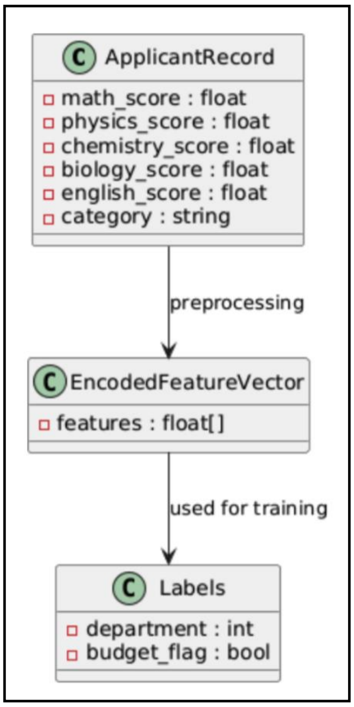

Рисунок 4.1 – Диаграмма классов реализованной информационной системы рекомендаций

Диаграмма классов на рисунке 4.1, демонстрирует логику взаимодействия компонентов и поток данных от ввода данных (Applicant Record) (баллы ОРТ и район проживания) и до формирования рекомендаций (*Labels*). Такая структура обеспечивает модульность и упрощает интеграцию дополнительных функций в будущем.

Разработанная система рекомендаций абитуриентам построена как веб-приложение, состоящее из двух взаимодействующих подсистем:

1. **Kotlin/Vert.x-сервер** - отвечает за веб-интерфейс, прием запросов от пользователя и формирование HTML-страниц с результатами.
2. **Python/Flask-API** - инкапсулирует модель машинного обучения и вычисляет рекомендации по специальностям.

Процесс функционирования разработанной веб-системы рекомендаций включает в себя последовательное взаимодействие двух серверных компонентов.

Диаграмма последовательности на рисунке 4.2, показывает логику обработки данных от момента ввода пользователем результатов ОРТ для получения рекомендаций. Пользователь отправляет запрос с данными в систему, которая направляет их API, а затем в модель для выполнения этапа предобработки - масштабирования и кодирования признаков. После этого модель возвращает системе подготовленный вектор данных. Получив информацию, система формирует предсказание и возвращает результат.

Взаимодействие компонентов происходит по следующей схеме:

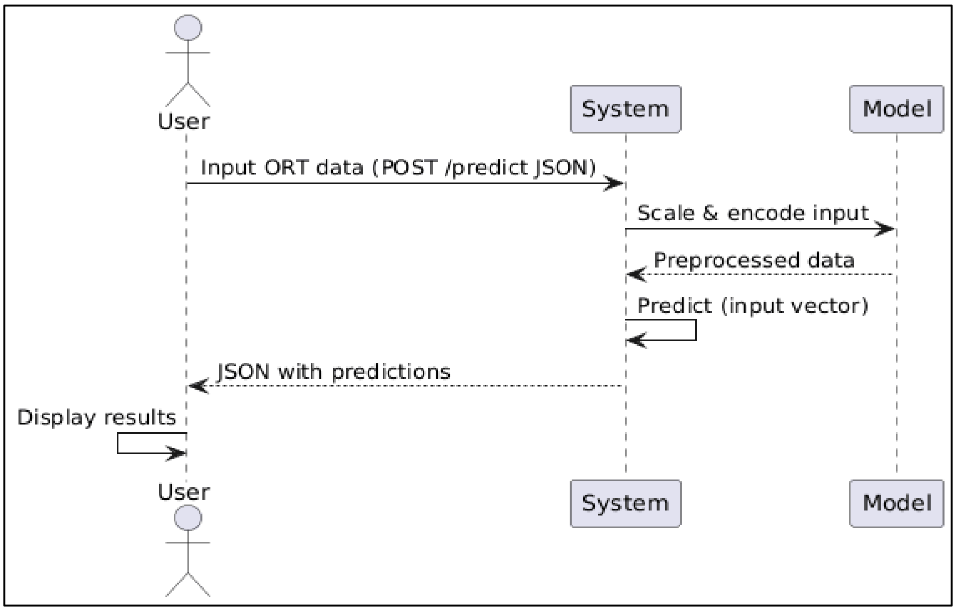

Рисунок 4.2. Диаграмма последовательности реализованной системы рекомендаций.

1. Пользователь через веб-интерфейс Kotlin-сервера вводит исходные данные и отправляет форму.
2. Kotlin-сервер преобразует данные формы в JSON и выполняет HTTP-запрос */predict* к API.
3. API выполняет все операции по подготовке признаков для модели, возвращая JSON-ответ с результатами.
4. Kotlin-сервер на основе ответа формирует страницу с результатами и отправляет её браузеру пользователя.

Для упрощения развертывания предусмотрен процесс автоматического запуска Python-API из Kotlin-приложения: при старте веб-сервера создается отдельный процесс интерпретатора Python, в котором запускается Flask-приложение.

Проектная часть

Веб-подсистема реализована с использованием фреймворка Vert.x для Kotlin.

Основные элементы:

- HTTP-сервер на базе Vert.x, слушающий порт, задаваемый переменным окружением
- Router – компонент маршрутизации, на котором определены основные точки входа
- GET / – маршрут, выдающий стартовую HTML-страницу (indexpage.html) с формой ввода данных абитуриента
- POST /predict – основной маршрут обработки запросов на формирование рекомендаций

Обработка запросов:

Маршрут POST /predict реализует следующую последовательность операций:

1. **Извлечение данных формы**

С помощью метода *formAttributes()* из объекта формы извлекаются значения следующих полей:

- «ort» – баллы за основной тест ОРТ.
- «biology» – баллы за предметный тест по Биологии.
- «chemistry» – баллы за предметный тест по Химии.
- «physics» – баллы за предметный тест по Физике
- «english» – баллы за предметный тест по Английскому языку
- «math» – баллы за предметный тест по Математике

Данные считываются как числовые значения, которые впоследствии используются для построения входного вектора признаков.

2. **Формирование структуры входных данных.** На основе извлеченных значений формируется экземпляр объекта класса *ApplicantRecord*. Этот объект служит оберткой для входных данных модели.
3. **Сериализация в JSON.** Объект InputData сериализуется в формат JSON с помощью библиотеки *kotlinx.serialization*. В результате получается корректное JSON-представление входных данных, которое можно напрямую передавать в REST API модели для предсказания.
4. **Отправка запроса к Python-API.** Для взаимодействия с моделью используется компонент *WebClient*, который формирует *HTTP*-запрос с следующими характеристиками:
   - метод *POST*;
   - адрес *http://localhost:8000/predict*;
   - заголовок Content-Type: application/json;
   - тело – сериализованный JSON с входными данными.

   Запрос выполняется асинхронно, что соответствует реактивной модели Vert.x и позволяет системе в дальнейшем масштабироваться при увеличении количества пользователей.
5. **Получение ответа.** В случае успешного выполнения запроса и получения ответа тело ответа читается, как строка и десериализуется в объект *ApiResponse*.

Формирование страницы результата

После получения ответа в объекте веб-подсистема выполняет:

1. Создание страницы результата
2. Построение строк страницы результата
3. Формирование массивов для визуализации графиков
4. Подстановку и отображение итоговой страницы пользователю

Подход с использованием шаблонов упрощает сопровождение интерфейса и позволяет вносить изменения в представление без изменения серверной логики.

Подсистема прогнозирования реализована на Python + Flask, как простое REST-приложение с одним основным эндпоинтом /Predict.

При запуске сервиса выполняются следующие действия:

- рабочий каталог устанавливается в директорию, где размещен скрипт, что обеспечивает корректный доступ к файлам с моделями
- из файлов, сохраненных на этапе обучения, загружаются:
  - модель многоклассовой классификации направлений *department_model*
  - словарь бинарных моделей для оценки вероятности бюджета по каждому направлению *budget_models*
  - кодировщик меток *label_encoder*, связывающий числовые индексы классов с текстовыми определениями
  - масштабировщик числовых признаков scaler
  - список всех признаков *feature_columns*, включающий как числовые, так и категориальные (*One-Hot*) столбцы.

Использование файла конфигурации признаков *feature_columns* гарантирует, что формат входных данных на этапе применения полностью соответствует формату, использованному при обучении.

При вызове */predict*, как параметр ожидается заранее подготовленный *JSON-объект*, который содержит:

- числовые поля: общий балл ОРТ и баллы по отдельным предметам
- строковое поле: category.

После происходит валидация входных данных, которая при некорректности данных вернет ошибку *400 Bad Request* пользователю с указанием отсутствующего поля.

Из числовых данных создается новый *DataFrame*. Далее для категориального признака выполняется восстановление *One-Hot* представления: для каждого категориального признака устанавливается значение 1 только для столбца, соответствующего переданной категории, остальные принимают значение 0. Для всех признаков из *feature_columns*, не попавших в исходные данные, создаются столбцы с нулевыми значениями.

В результате мы получаем готовый вектор признаков необходимый для работы с моделью.

Числовые признаки преобразуются при помощи обученного мастабировщика *scaler*, это обеспечивает сопоставимость распределения значений с тем, на котором обучалась модель.

Многоклассовая модель *department_model* применяет метод *predict_probability*, в результате чего получается вектор вероятностей для всех возможных направлений. С помощью *label_encoder* числовые индексы классов переводятся в текстовые названия направлений. Полученные тройки (название, индекс, вероятность) сортируются по убыванию вероятности. После, для формирования ответа количество лучших вероятностей сокращается до 10, для простого вывода рекомендации.

Для каждого направления, по индексу возвращается ответ в виде распределения от 0–1, которая предполагает вероятность поступления на бюджетную форму обучения. Если данного ответа не существует возвращается 0.

Такой подход сопутствует дальнейшему обучению модели при накоплении новой статистики.

Для каждого направления формируется запись Recommendation, которая возвращается ответом на запрос.

С точки зрения конечного пользователя процесс работы системы выглядит следующим образом:

- Абитуриент открывает веб-страницу системы и в форме вводит свои результаты ОРТ и предметных тестов, а также выбирает категорию (район проживания).
- По нажатию кнопки отправки формы данные отправляются на маршрут POST /predict Kotlin-сервера.
- Веб-сервер сериализует введённые данные в JSON и передаёт их Python-API прогнозирования.
- Python-API проверяет корректность данных, восстанавливает вектор признаков, масштабирует его и применяет обученные модели:
    - определяет наиболее вероятные направления обучения
    - оценивает вероятность поступления на бюджет по каждому из них.
- Результаты в виде JSON возвращаются в веб-сервер, который формирует HTML-страницу результатов: таблицу с ранжированным списком направлений и вероятностями, а также графические элементы визуализации.
- Абитуриент просматривает рекомендации и при необходимости может изменить исходные данные и выполнить повторный расчёт, используя систему в режиме анализа «что-если».

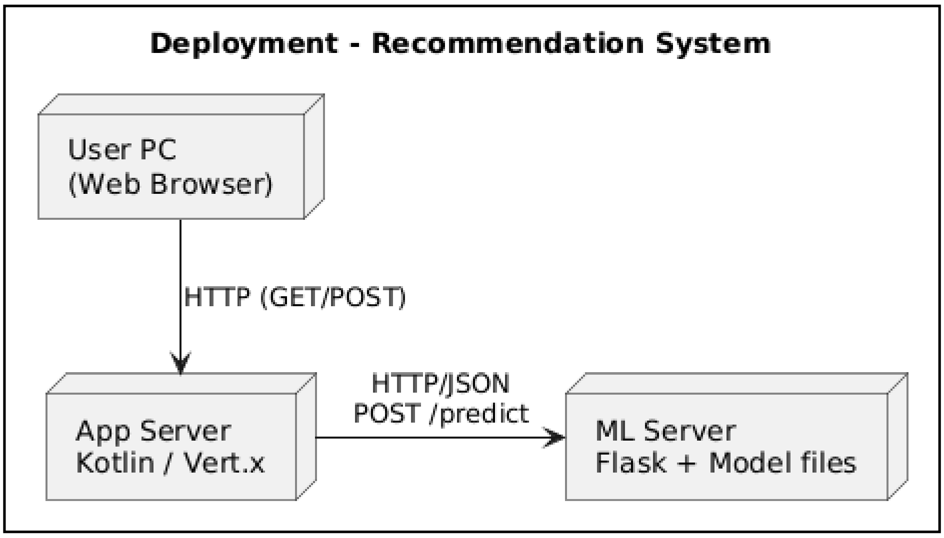

Рисунок 4.3 – Архитектура развертывания системы.

Архитектура развертывания разработанной системы представлена на рисунке 4.3.

На ней показано клиентское рабочее место пользователя (User PC с веб-браузером), сервер приложений (App Server на базе Kotlin / Vert.x), а также отдельный сервер подсистемы машинного обучения (ML Server с Flask и файлами моделей).

Взаимодействие браузера с сервером приложений осуществляется по протоколу HTTP (GET/POST), а обмен данными между сервером приложений и сервером машинного обучения выполняется по HTTP/JSON-запросу /predict.

Интерфейс программы

Форма, представленная на рисунке 4.4, реализует **основной сценарий взаимодействия пользователя с системой** и выполняет функции ввода исходных данных для последующего расчёта рекомендаций.

Основные функциональные элементы интерфейса:

1. Блок ввода результатов ОРТ

- Поле Балл ОРТ (общий результат теста) – ввод общего рейтинга ОРТ.
- Поля Биология, Химия, Физика, Английский, Математика – ввод баллов по предметным тестам.

- Все эти значения используются как числовые признаки в модели. Если предмет не сдавался, пользователь вводит 0.

2. Блок Регион проживания

- Выпадающий список с выбором региона. Выбранный регион передаётся на сервер, как категориальный признак и учитывается при формировании рекомендаций.

3. Кнопка Получить рекомендации

- При нажатии собираются все введённые значения полей.
- Данные отправляются на сервер приложений, далее – в подсистему машинного обучения. После обработки пользователь переходит на страницу с таблицей и графиками рекомендаций.

4. Блок Инструкция по использованию

- Кратко объясняет, как правильно вводить баллы и что делать, если тест по предмету не сдавался. Блок снижает риск ошибок ввода, но не участвует в вычислениях.

5. Блок Объяснение графиков

- Поясняет, что показывают диаграммы на странице результатов (столбчатая и круговая).
- Это помогает правильно интерпретировать вероятности, которые показывает система.

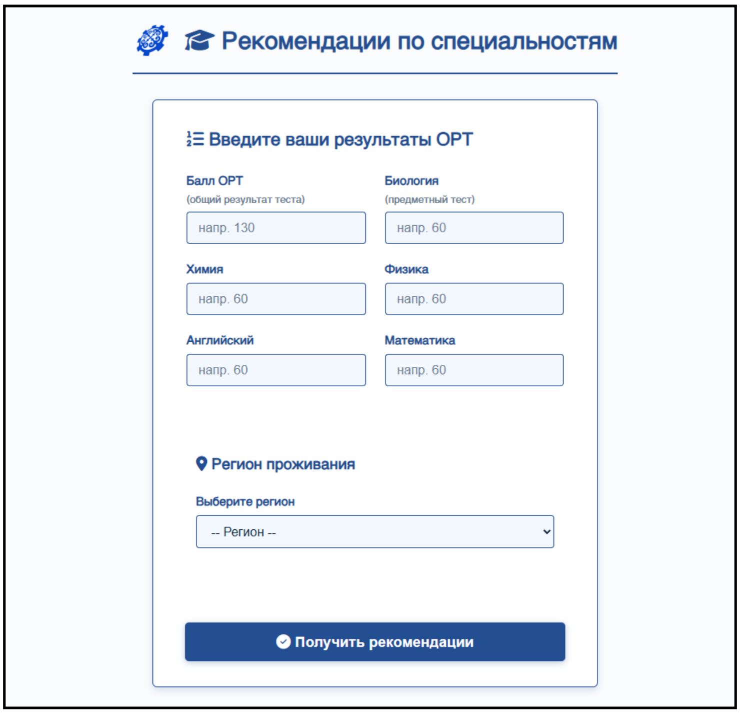

Рисунок 4.4 – Главная страница
реализованной системы рекомендаций.

Страница, представленная на рисунке 4.4, отображает вывод рекомендаций.

1. Таблица рекомендаций
   - В столбце «Специальность» выводится перечень рекомендованных направлений подготовки, отсортированных по убыванию вероятности.
   - В столбце «Вероятность поступления» показывается вероятность того, что абитуриент сможет поступить на данную специальность.
   - Столбец «Вероятность бюджета» показывает предполагаемую вероятность получения бюджетного места по каждой из предложенных специальностей.
2. Столбчатая диаграмма
   - Отражает вероятности поступления по рекомендованным специальностям в процентах.
   - Позволяет быстро сравнить между собой направления по уровню шансов поступления.
3. Круговая диаграмма

- Показывает распределение вероятности бюджета между выбранными специальностями.
- Помогает визуально оценить, на какие направления приходится основной вклад по шансу бюджетного обучения.

4. Кнопка «Вернуться к форме»

- Обеспечивает возврат на экран ввода исходных данных.
- Позволяет пользователю изменить оценки или регион проживания и запустить повторный расчёт рекомендаций.

Рисунок 4.5 – Страница
полученной рекомендации.

Пользователь проверяет результат, задавая входные данные (*баллы ОРТ, район*) и сравнивая полученные рекомендации. Предсказуемое изменение списка специальностей при варьировании входных данных подтверждает корректность работы системы.

Оценка качества рекомендаций

Для проверки качества работы модели использовалась тестовая выборка абитуриентов, не участвовавшая в обучении. Для каждого объекта сравнивалась предсказанная специальность и форма обучения. В качестве метрик использовалась доля верных ответов (*accuracy*) модели, в том случае, когда специальность входила в список направлений.

Таким образом, с точки зрения пользователя результат легко проверить через интерактивный интерфейс, за счет изменения входных данных и получения рекомендации.

Выводы

Проведенное исследование подтвердило эффективность использования нейронной сети для задачи классификации специальностей абитуриентов. По сравнению с базовой моделью, внедрение *AdaptiveActivation* позволило добиться повышения точности предсказаний, что указывает на потенциальную значимость данного подхода в практических приложениях систем поддержки принятия решений.

Детальное изучение экспериментальных результатов показало, что использование адаптивной функции активации позволяет модели более гибко и контекстно осмысленно преобразовать выходные признаки. Благодаря этому сеть лучше выделяет значимые особенности данных, что способствует повышению качества разделения классов. В отличие от традиционных активации, таких как ReLU или sigmoid, адаптивные функции способны менять свое поведение в зависимости от структуры обучающегося пространства. Это особенно важно при работе с неоднородными или слабо организованными данными, которые типичны для задач профессиональной ориентации.

Несмотря на то что абсолютный прирост точности оказался незначительным, он является статистически значимым, что подтверждает целесообразность внедрения адаптивных механизмов в архитектуру моделей машинного обучения. Это подкрепляет гипотезу о перспективности адаптивных функций активации и открывает новые направления для дальнейших исследований. Среди них — сравнительный анализ различных типов адаптивных активационных функций, их интеграция с методами регуляризации, внимания (attention mechanisms), обучения с подкреплением, а также оценка эффективности в условиях потоковой обработки данных и на разнообразных наборах.

Особое внимание заслуживает концепция нейрона с многоразветвлённой структурой и динамическими управляющими элементами (Multi-Branch Gated Neuron), примененная в архитектуре модели. Такая структура позволяет параллельно обрабатывать входную информацию в различных представлениях, что значительно повышает гибкость и адаптивность нейросети. Управляющие (гейтинговые) механизмы позволяют динамически регулировать вклад каждой ветви, обеспечивая контекстно-зависимое поведение модели.

Сочетание мульти-ветвевой структуры и динамического стробирования расширяет потенциал нейросетевых решений как в плане качества предсказаний, так и эффективности использования ресурсов. Это особенно актуально в условиях ограниченных вычислительных мощностей, когда важно оптимизировать затраты на обучение и инференс.

Важным направлением дальнейшего развития предложенного подхода является его адаптация к обработке больших объёмов данных. Внедрение современных архитектур глубокого обучения, таких как трансформеры или сверточные нейронные сети с расширенными возможностями агрегации признаков, может существенно повысить практическую применимость метода. Это особенно актуально в условиях постоянно растущих требований к скорости и точности обработки информации в реальном времени.

В перспективе дальнейшие исследования могут быть направлены на:

- оптимизацию структуры адаптивных функций активации с использованием эволюционных алгоритмов и нейроэволюции;
- внедрение методов самообучения и трансферного обучения для повышения устойчивости модели к шумам и неполным данным;
- разработку гибридных систем, сочетающих преимущества нейросетевых подходов и традиционных алгоритмов классификации;
- интеграцию предложенной архитектуры в промышленные рекомендательные системы, цифровые платформы профориентации и адаптивные образовательные траектории.

Таким образом, полученные результаты не только демонстрируют эффективность предложенного метода, но и формируют прочную основу для последующих теоретических и прикладных разработок в области интеллектуальных систем классификации и предсказательного моделирования.

Adaptive Activation может использоваться в:

- **Рекомендательных системах и глубоких классификаторах**: В таких задачах, как рекомендации различных траекторий развитий или в медицинских предсказаниях.
- **Обучении с шумными выборками**: За счет более гибкой трансформации признаков Adaptive Activation может быть менее подвержен шуму и переобучению. Что важно при работе с данными объем которых ограничен.
- **Многозадачном обучении**: В архитектурах с несколькими выходами (например, как предсказание специальностей и одновременное прогнозирование формы обучения) адаптивные функции будут более эффективным подходом для подстройки слоев на разные подзадачи.

MBGN может использоваться в:

- **Обработке естественного языка**: В таких задачах, как перевод, обобщение или анализ настроения, MBGN помогает моделям динамически выбирать наиболее подходящие пути обработки.
- **Компьютерном зрении**: В визуальных задачах MBGN позволяет одновременно улавливать, как глобальные, так и локальные особенности.
- **Унифицированных моделях ИИ**: Структуры MBGN перспективны для крупномасштабных систем ИИ общего назначения, которые должны работать с различными типами данных и контекстами.
- **Обучении с ограниченными данными (low-resource learning):** за счёт более эффективного использования признаков MBGN способен лучше адаптироваться к малым выборкам, чем традиционные архитектуры.

В результате работы были решены следующие задачи:

1. Выполнен анализ существующих методов классификации и прогнозирования в образовательной аналитике, а также выявлены их основные ограничения

2. Проведен анализ исходных данных и установлены взаимосвязи между характеристиками абитуриентов и выбранными направлениями подготовки
3. Разработана и исследована адаптивная функция активации (Adaptive Activation Function), направленная на повышение точности и устойчивости нейросетевых моделей.
4. Разработана архитектура Multi-Branch Gated Neuron, обеспечивающая учет сложных нелинейных зависимостей между результатами ОРТ и выбором специальности.
5. Проведены экспериментальные количественные исследования предложенных архитектур на реальных и синтетических данных, выполнено сравнение с классическими моделями.
6. Разработано программное обеспечение для практического применения предложенных методов в автоматизированных системах поддержки приемной кампанией.

REFERENCES

1. Agarap, A. F.: Deep learning using Rectified Linear Units (ReLU). arXiv preprint arXiv:1803.08375. https://arxiv.org/abs/1803.08375 (2018).
2. Alyahyan, E., Düştegör, D.: 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 (2020).
3. Apicella, A., Donnarumma, F., Isgrò, F., Prevete, R.: A survey on modern trainable activation functions. Neural Networks 138, 14–32. https://doi.org/10.1016/j.neunet.2021.01.026 (2021).
4. Bozieva, A. M., Tseyeva, F. M., Khatukhova, D. V.: Application of machine learning methods for assessing the performance of higher education institutions [in Russian]. (2024).
5. Belotsky, E. A., Suetin, A. V.: Building a recommender system for selecting higher education institutions for applicants [in Russian]. (2024).
6. Even-Dar, E., Mannor, S., Mansour, Y.: Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. Journal of Machine Learning Research 7, 1079–1105. (2006).
7. Gavrilenkov, V. I., Demakov, V. A.: Machine learning in education. Volga Academy of Education and Arts named after St. Alexius, Metropolitan of Moscow, Tolyatti, Russia [in Russian]. (2024).
8. Jabeen, H., Baig, A. R.: Review of classification using genetic programming. International Journal of Engineering Science and Technology 2(2), 94–103. https://www.researchgate.net/publication/275022403_Review_of_classification_using_genetic_programming (2010).
9. Lee, K., Yang, J., Lee, H., Hwang, J. Y.: Stochastic adaptive activation function. arXiv preprint arXiv:2210.11672. https://arxiv.org/abs/2210.11672 (2022).
10. Lin, X., Zhong, G., Chen, K., Li, Q., Huang, K.: Attention-Augmented Machine Memory. Cognitive Computation 13(3), 751–760. https://doi.org/10.1007/s12559-021-09854-5 (2021).

11. Lops, P., de Gemmis, M., Semeraro, G.: Content-based recommender systems: State of the art and trends. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P. B. (eds.) Recommender Systems Handbook, 1st edn., pp. 73–105. Springer. https://doi.org/10.1007/978-0-387-85820-3_3 (2011).
12. Ma, C., Wu, J., Si, C., Tan, K. C.: Scaling supervised local learning with augmented auxiliary networks. arXiv preprint arXiv:2402.17318. https://arxiv.org/abs/2402.17318 (2024).
13. Molina, M., Blasco, G.: 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 (2003).
14. Romero, C., Ventura, S.: 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://dx.doi.org/10.1109/TSMCC.2010.2053532 (2010).
15. Sakho, A., Malherbe, E., Scornet, E.: 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 (2024).
16. Sebastiani, F.: Machine learning in automated text categorization. ACM Computing Surveys 34(1), 1–47. https://doi.org/10.1145/505282.505283 (2002).
17. Miftahul, J. M., Sumi, K., Mohammad, S. A.: 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 (2020).
18. Misra, D.: Mish: A self-regularized non-monotonic activation function. arXiv preprint arXiv:1908.08681. https://arxiv.org/abs/1908.08681 (2020).
19. Kustitskaya, T.A., Noskov, M.V., Weinstein, Y.V.: Predicting Learning Success: Research Problems and Challenges. Informatics and Education 1, 4-15. http://dx.doi.org/10.1504/IJLT.2008.019376 (2023).
20. Mityanina, A.V., Gazha, K.V.: Recommender system «Abiturient» for university admission committee. (2024).

21. Molchanov, A.N., Burmistrov, A.V., Grishunov, S.S., Chukhraev, I.V.: Some issues of data collection and processing in the information system of a higher educational institution. [in Russian]. https://cyberleninka.ru/article/n/nekotorye-voprosy-sbora-i-obrabotki-dannyh-v-informatsionnoy-sisteme-vysshego-uchebnogo-zavedeniya/pdf (2024).
22. Dubey, S. R., Singh, S. K., Chaudhuri, B. B.: Activation functions in deep learning: A comprehensive survey and benchmark. https://doi.org/10.1016/j.neucom.2022.06.111 (2022).
23. Chen, S., Li, Z.: Intent-Enhanced Data Augmentation for Sequential Recommendation. State Key Laboratory of Complex & Critical Software Environment, Beihang University, Beijing 100191, China. https://doi.org/10.48550/arXiv.2410.08583 (2025).
24. Kostiuk, S.: Using Adaptive Activation Functions in Pre-Trained Artificial Neural Network Models. https://www.researchgate.net/publication/375758835_Using_Adaptive_Activation_Functions_in_Pre-Trained_Artificial_Neural_Network_Models (2024).
25. Harby, F., Thaljaoui, A., Nayab, D., Aladhadh, S., El Khediri, S., Khan, R. U.: Data Augmentation and Random Multi-Model Deep Learning for Data Classification. https://doi.org/10.32604/cmc.2023.029420 (2024).
26. Siemens, G., Baker, R. S. J. d.: Learning analytics and educational data mining: Towards communication and collaboration. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK'12), pp. 252–254. https://doi.org/10.1145/2330601.2330661 (2012).
27. Lin, Y., Chen, H., Xia, W., Lin, F., Wang, Z., Liu, Y.: A Comprehensive Survey on Deep Learning Techniques in Educational Data Mining. arXiv preprint arXiv:2309.04761. https://doi.org/10.48550/arXiv.2309.04761 (2024).
28. Küçük, E., Cincil, F., Karal, Y.: A systematic review of the ethical use of artificial intelligence (AI) in education. https://doi.org/10.21203/rs.3.rs-4370610/v1 (2024).
29. Misra, D.Mish: A self-regularized non-monotonic activation function. https://doi.org/10.48550/arXiv.2109.14545 (2020).

Акт о внедрении результатов научно-исследовательской работы в КГТУ им И. Раззакова

«УТВЕРЖДАЮ»
Ректор КГТУ им И. Раззакова
М. К. Чыныбаев
2025 г.

АКТ

**о внедрении результатов научно - исследовательской работы
«Разработка системы принятия решений (СППР) для набора абитуриентов ВУЗа в условиях неопределенности»**

Результаты научно-исследовательской работы «Разработка системы принятия решений (СППР) для набора абитуриентов ВУЗа в условиях неопределенности », автором которой является Болотбек уулу Н., внедрены в учебный процесс кафедры «Программное Обеспечение Компьютерных Систем» (ПОКС) Кыргызского Государственного Технического Университета им. И. Раззакова.

В рамках внедрения используются следующие научные результаты:

- Адаптивная функция активации, сочетающая результаты базовых функций (например, сигмоиды и гиперболического тангенса) с самообучаемыми коэффициентами что улучшает качество предсказаний.
- Многоветвевой нейронный блок с механизмом gated-управления (MBGN), позволяющий параллельно обрабатывать признаки по нескольким вычислительным ветвям с взвешиванием их вклада, что увеличивает выразительную способность моделей при анализе данных.
- Автоматизированные data-pipelines сбора, очистки, обогащения и валидации данных абитуриентов, обеспечивающие подготовку для последующего использования в моделях машинного обучения и системах поддержки принятия решений.

Внедренные результаты применяются в учебном процессе при изучении дисциплин, связанных с искусственным интеллектом, анализом данных и разработкой интеллектуальных систем поддержки принятия решений, способствуя повышению практической подготовки студентов в области современных методов машинного обучения.

Заведующий кафедрой ПОКС
доктор физ.-мат. наук, профессор

Салиев А. Б. ____________________

Акт о внедрении результатов научно-исследовательской работы в Международном Университете в Центральной Азии

«УТВЕРЖДАЮ»
Вице-Президент по стратегическому развитию и академическим вопросам МУЦА
_______________ А. М. Джумабаев
«10 ноября» 2025 г.

АКТ

о внедрении результатов научно - исследовательской работы
«Разработка системы принятия решений (СППР) для набора абитуриентов ВУЗа в условиях неопределенности»

Результаты научно-исследовательской работы «Разработка системы принятия решений (СППР) для набора абитуриентов ВУЗа в условиях неопределенности», автором которой является Болотбек уулу Н., внедрены в учебный процесс программы «Информационные Технологи» Международного университета в Центральной Азии.

В рамках внедрения используются следующие научные результаты:

- Адаптивная функция активации, сочетающая результаты базовых функций (например, сигмоиды и гиперболического тангенса) с самообучаемыми коэффициентами что улучшает качество предсказаний.
- Многоветвевой нейронный блок с механизмом gated-управления (MBGN), позволяющий параллельно обрабатывать признаки по нескольким вычислительным ветвям с взвешиванием их вклада, что увеличивает выразительную способность моделей при анализе данных.
- Автоматизированные data-pipelines сбора, очистки, обогащения и валидации данных абитуриентов, обеспечивающие подготовку для последующего использования в моделях машинного обучения и системах поддержки принятия решений.

Внедренные результаты применяются в учебном процессе при изучении дисциплин, связанных с искусственным интеллектом, анализом данных и разработкой интеллектуальных систем поддержки принятия решений, способствуя повышению практической подготовки студентов в области современных методов машинного обучения.

Глава направления IT
Вейс П. Л. _______________

Сертификат об обучении в период мобильности. Università Di Camerino. Italia

UNICAM Università di Camerino 1336

Università di Camerino

Erasmus + KA171/2022
Project n. 2022-1-IT02-KA171-HED-000073309
Academic Year 2023-2024
Certificate of Attendance

This is to certify that

Mr. **BOLOTBEK Nursultan** (CF BLTNSL97A12Z256Q)

coming from the **KYRGYZ TECHNICAL UNIVERSITY NAMED AFTER ISKHAAK RAZZAKOV (KSTU)** has performed his mobility period for study at the host Institution UNIVERSITY OF CAMERINO – SCHOOL OF SCIENCE AND TECNOLOGY (Computer Science Department - Via Madonna delle Carceri, 7 - 62032 Camerino (MC)

Erasmus code: I CAMERIN01

He started his mobility persio at our Institution on **September 26th, 2023**

And completed his mobility period on **February 20th, 2024**

Place and date	**Camerino, 20/02/2024**
Name and status of the signatory	**Prof. Luciano Barboni** **Rector Delegate for Extra-UE Countries Programs**
Signature	
Host institution Stamp	UNIVERSITAS CAMERINENSIS MCCCXXXVI MDCCXXVII MDCCLIII

Amministrazione

C.F.81001910439
protocollo@pec.unicam.it
www.unicam.it

Area Servizi agli Studenti

Ufficio Relazioni Internazionali

Polo degli Studenti "Franco Biraschi"
Via Gentile III da Varano, 26
62032 **Camerino** (MC) Italy
tel + 39 0737 404619
erasmus@unicam.it
relazioni.internazionali@unicam.it
international.unicam.it