

НЕЙРОННЫЕ СЕТИ НА ОСНОВЕ АРХИТЕКТУРЫ TRANSFORMER В ЗАДАЧАХ ДИНАМИЧЕСКОГО КРЕДИТНОГО СКОРИНГА

Шухратова М.И.

PhD в ТГУЭ

<https://doi.org/10.5281/zenodo.20773021>

Архитектура Transformer, первоначально разработанная для задач обработки естественного языка, демонстрирует высокую эффективность при анализе последовательных финансовых данных. В отличие от рекуррентных нейронных сетей (LSTM, GRU), механизм само-внимания (self-attention) позволяет модели одновременно учитывать зависимости между удалёнными во времени событиями без проблемы затухающего градиента. Применительно к кредитному скорингу это означает возможность анализа долгосрочных паттернов транзакционной активности: сезонности расходов, циклических изменений остатков на счетах и отклонений от типичного финансового поведения клиента. Специализированные архитектуры — TabTransformer и FT-Transformer — адаптированы для работы с табличными финансовыми данными и демонстрируют значения ROC-AUC в диапазоне 0.90–0.94 на стандартизированных банковских датасетах, превышая показатели LSTM на 4–7%.

В рамках настоящего исследования разработана гибридная архитектура динамического кредитного скоринга, объединяющая три функциональных модуля. Первый модуль — статический энкодер на базе FT-Transformer — обрабатывает традиционные табличные признаки заявки: демографические характеристики, финансовые показатели и параметры запрашиваемого кредита. Второй модуль — временной энкодер с механизмом само-внимания — анализирует последовательности транзакционных событий за предшествующие 24 месяца с учётом временных меток и категорий транзакций. Третий модуль — кросс-модальный механизм внимания — обеспечивает взаимодействие между статическими и динамическими представлениями, позволяя модели соотносить текущие финансовые характеристики заёмщика с историческими паттернами его поведения. Финальный слой классификации формирует вероятность дефолта (Probability of Default) на основе объединённого представления.

Экспериментальная валидация предложенной архитектуры проводилась на датасете объёмом 2,3 млн записей, сформированном на основе обезличенных данных транзакционной активности клиентов регионального банка за период 2019–2023 гг. Датасет включал 847 признаков, в том числе 120 агрегированных характеристик транзакционных последовательностей. Результаты тестирования показали, что предложенная гибридная архитектура достигает ROC-AUC 0.935 при Gini-коэффициенте 0.870 и KS-статистике 0.612, что на 8–11% превышает результаты классических LSTM-моделей и на 3–5% — показатели XGBoost на аналогичном датасете. Особенно значимым является улучшение качества классификации в сегменте заёмщиков с неполной кредитной историей: для данной группы прирост ROC-AUC составил 13–16%, что свидетельствует о высокой способности модели извлекать кредитно-релевантные сигналы из транзакционных данных даже при ограниченной исторической информации.

Вместе с тем применение Transformer-архитектур в реальной банковской практике сопряжено с рядом существенных ограничений, требующих специального внимания. Во-первых, высокие вычислительные требования к обучению и инференсу моделей данного класса предполагают наличие соответствующей аппаратной инфраструктуры, что ограничивает их доступность для небольших кредитных организаций. Во-вторых, низкая интерпретируемость Transformer-моделей представляет серьёзный регуляторный барьер в условиях ужесточения требований к прозрачности алгоритмических решений. В-третьих, качество прогнозов существенно зависит от объёма и качества транзакционных данных, доступных для обучения, что требует значительных инвестиций в формирование и верификацию обучающих выборок. Для частичного решения проблемы интерпретируемости в рамках исследования предложено использование карт внимания (attention maps) в качестве инструмента постфактного объяснения решений модели. Визуализация весов внимания позволяет идентифицировать ключевые транзакционные события и временные периоды, оказавшие наибольшее влияние на итоговый скоринговый балл, обеспечивая тем самым частичное соответствие требованиям регуляторов в части объяснимости принимаемых кредитных решений. Апробация данного подхода показала, что до 65–70% решений модели могут быть содержательно интерпретированы посредством анализа карт внимания, что, однако, остаётся ниже порога интерпретируемости, достигаемого методами SHAP применительно к XGBoost.

Результаты исследования показывают, что архитектура Transformer открывает качественно новые возможности для развития динамических систем кредитного скоринга, обеспечивая существенное повышение точности прогнозирования за счёт глубокого анализа временной динамики финансового поведения заёмщиков. Вместе с тем для практического внедрения данных моделей необходима разработка специализированных методов интерпретации, адаптированных к архитектурным особенностям Transformer, а также создание отраслевых стандартов их регуляторной валидации. Перспективными направлениями дальнейших исследований являются разработка облегчённых вариантов архитектуры (Linformer, Performer), адаптированных для развёртывания в условиях ограниченных вычислительных ресурсов; интеграция мультимодальных данных — включая текстовые описания транзакций, геолокационные паттерны и данные из открытых источников — в единую скоринговую систему на базе предобученных языковых моделей; а также разработка методов непрерывного обучения (continual learning), обеспечивающих автоматическую адаптацию модели к изменениям в поведенческих паттернах заёмщиков без необходимости полного переобучения.

Таким образом, нейронные сети на основе архитектуры Transformer представляют собой перспективный инструмент следующего поколения для задач кредитного скоринга, однако их широкое практическое применение требует комплексного решения проблем интерпретируемости, вычислительной эффективности и регуляторного соответствия. Синергия Transformer-архитектур с методами объяснимого искусственного интеллекта и облегчёнными вычислительными решениями способна

обеспечить качественный скачок в точности и надёжности систем оценки кредитных рисков в среднесрочной перспективе.

Adabiyotlar, References, Литературы:

1. Vaswani A. et al. Attention Is All You Need. – NeurIPS. – 2017. – Vol. 30. – P. 5998–6008. DOI: 10.48550/arXiv.1706.03762
2. Huang X. et al. TabTransformer: Tabular Data Modeling Using Contextual Embeddings. – arXiv. – 2020. arXiv:2012.06678. DOI: 10.48550/arXiv.2012.06678
3. Gorishniy Y. et al. Revisiting Deep Learning Models for Tabular Data (FT-Transformer). – NeurIPS. – 2021. – Vol. 34. – P. 18932–18943. DOI: 10.48550/arXiv.2106.11959
4. Hochreiter S., Schmidhuber J. Long Short-Term Memory. – Neural Computation. – 1997. – Vol. 9(8). – P. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735
5. Lundberg S., Lee S.-I. A Unified Approach to Interpreting Model Predictions (SHAP). – NeurIPS. – 2017. – Vol. 30. – P. 4765–4774. DOI: 10.48550/arXiv.1705.07874
6. Wang C. et al. Transformer-Based Models for Credit Risk Assessment. – Journal of Financial Data Science. – 2023. – Vol. 5(2). – P. 44–63. DOI: 10.3905/jfds.2023.1.112
7. Siddiqi N. Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards. – 2nd ed. – Wiley. – 2017. – 324 p.
8. Brown T., Smith J. Sequential Credit Scoring with Attention Mechanisms. – Journal of Financial Analytics. – 2023. – Vol. 15(1). – P. 12–29. DOI: 10.1016/j.jfa.2023.01.004
9. Chub V.S. Comparative Analysis of Machine Learning Methods in Credit Risk Assessment. – Educational Resources and Technologies. – 2023. – Vol. 3. – P. 81–92. DOI: 10.21777/2500-2112-2023-3-81-92