

ПРОГНОЗИРОВАНИЕ КРЕДИТНЫХ СПРЕДОВ КОРПОРАТИВНЫХ ОБЛИГАЦИЙ С ИСПОЛЬЗОВАНИЕМ АНСАМБЛЕЙ ГРАДИЕНТНОГО БУСТИНГАDOI: <https://doi.org/10.24891/elhfng>EDN: <https://elibrary.ru/elhfng>**Константин Владимирович КРИНИЧАНСКИЙ**

ответственный автор, доктор экономических наук, профессор кафедры финансовых рынков и финансового инжиниринга; главный научный сотрудник Института финансовых исследований финансового факультета, Финансовый университет при Правительстве РФ, Москва, Российская Федерация
e-mail: kkrin@ya.ru
ORCID: 0000-0002-1225-7263
SPIN: 3366-2850

Артем Алексеевич КОБЗЕВ

главный специалист отдела методологии рыночных рисков НКЦ, ПАО «Московская Биржа», Москва, Российская Федерация
e-mail: artem.kobzev.2001@mail.ru
ORCID: отсутствует
SPIN: отсутствует

Александра Юрьевна ПИСКАРЁВА

студентка, Финансовый университет при Правительстве РФ, Москва, Российская Федерация
e-mail: 221236@edu.fa.ru
ORCID: 0009-0007-9722-1094
SPIN: отсутствует

История статьи:

Рег. № 844/2025
Получена 17.12.2025
Одобрена 12.01.2026
Доступна онлайн
28.05.2026

Специальность: 5.2.6

УДК 336.76

JEL: C45, G12, G17

Ключевые слова:

рынок корпоративных облигаций, кредитный риск, кредитный спред, машинное обучение, градиентный бустинг

Аннотация

Предмет. Алгоритмы градиентного бустинга для прогнозирования кредитных спредов корпоративных облигаций на российском рынке долговых инструментов.

Цели. Разработка и эмпирическая проверка концепции построения прогнозных моделей кредитных спредов.

Методология. Применены методы многомерного статистического анализа, регрессионного моделирования и ансамблевого машинного обучения. В качестве основного инструментария использованы модели градиентного бустинга. Для оценки вклада и значимости предикторов применена методология интерпретации SHapley Additive exPlanations.

Результаты. Установлено, что использование ансамблевых моделей градиентного бустинга обеспечивает значительное повышение точности прогнозирования кредитных спредов по сравнению с традиционными регрессионными подходами. Разработанные модели выявляют сложные нелинейные взаимодействия между факторами различного уровня и демонстрируют устойчивость в периоды рыночных шоков. Предложены структурированные подходы к интеграции микро- и макроэкономических индикаторов в единую модель прогнозирования.

Выводы. Градиентный бустинг, в частности модель CatBoost, эффективен в качестве инструмента анализа и прогнозирования кредитных спредов. Обоснована применимость метода в задачах финансовой аналитики, риск-

менеджмента и поддержки инвестиционных решений. К ограничивающим факторам следует отнести зависимость качества прогнозов от полноты и достоверности исходных данных, а также необходимость регулярной актуализации моделей. Возможно внедрение предложений в системы автоматизированного принятия инвестиционных решений, кредитного скоринга и стресс-тестирования, направленных на повышение прозрачности и устойчивости российского рынка корпоративного долга.

© Издательский дом ФИНАНСЫ и КРЕДИТ, 2025

Для цитирования: Криничанский К.В., Кобзев А.А., Пискарьева А.Ю. Прогнозирование кредитных спредов корпоративных облигаций с использованием ансамблей градиентного бустинга // Финансы и кредит. – 2026. – № 5. – С. 124 – 141. DOI: 10.24891/elhfng EDN: ELHFNG

Введение

Разработка надежных и точных методов прогнозирования кредитных спредов корпоративных облигаций – часто с использованием кредитных спредов (в том числе G-спредов)¹ – становится важной задачей современной финансовой аналитики. Такие прогнозы имеют решающее значение для управления рисками, оптимизации портфеля, управления активами и пассивами, а также планирования ликвидности в финансовых учреждениях. Более точные ожидания относительно будущих изменений спредов позволяют инвесторам, эмитентам и регулирующим органам лучше предвидеть изменения рыночных условий и кредитоспособности эмитентов, заблаговременно реагировать на изменения в оценке активов и ликвидности, а также разрабатывать более сложные и адаптивные инвестиционные стратегии². На практике кредитные спреды также служат индикатором раннего предупреждения о переоценке рисков, инструментом для выявления недооцененных ценных бумаг и основой для построения стратегий хеджирования на сегментированных или неликвидных рынках облигаций.

Усиление конкуренции среди участников рынка облигаций и растущая сложность глобальных рынков капитала повысили спрос на модели прогнозирования, которые являются точными, гибкими и способными адаптироваться к структурным изменениям. Прогнозирование кредитных спредов осложняется многочисленными источниками неопределенности, включая макроэкономические колебания, фундаментальные факторы, специфичные для каждого эмитента, меняющиеся режимы регулирования. Реакции рынка на информационные шоки часто демонстрируют нелинейность, структурные разрывы и волны настроений инвесторов – особенности, которые подрывают эффективность традиционных линейных моделей и требуют более совершенных аналитических инструментов [1].

Академическая литература по анализу кредитных спредов обычно следует двум взаимодополняющим направлениям. Макроэкономические эмпирические исследования [2, 3] подчеркивают важность наблюдения за такими факторами риска, как инфляционные ожидания, движения процентных ставок, волатильность фондового рынка, финансовая стабильность. Микроэкономические исследования фокусируются на характеристиках эмитентов,

¹ G-спред рассчитывают как разность доходности рискованной облигации и аналогичной по сроку погашения (дюрация) доходностью безрисковой (государственной) облигации. Таким образом, G-спред воплощает собой плату за кредитный риск, связанный с эмитентом.

² Криничанский К.В. Технологическая трансформация рынков долговых ценных бумаг и ее эффекты // Финансы и кредит. 2023. Т. 29. № 2. С. 244–261. DOI: 10.24891/fc.29.2.244 EDN: QLCWKH

среди которых кредитные рейтинги, уровень закредитованности, прибыль, ликвидность, корпоративное управление, демонстрируя, что неоднородность на уровне фирмы играет существенную роль в межфирменных изменениях спредов. Взаимодействие между макроэкономическими факторами и детерминантами на уровне фирм в сочетании с изменяющейся во времени чувствительностью рынка к каждому из них подчеркивает нетривиальный характер прогнозирования кредитных спредов.

Быстрое развитие российского рынка корпоративных облигаций стимулировало эмпирические исследования, специфичные для его институциональной среды. Результаты исследований показывают, что динамика спредов зависит не только от надежности эмитента и отраслевой принадлежности, но и от поведения внутреннего рынка акций [4]. Если опираться на международные исследования [5], можно предположить, что специфические для эмитента или экономики страны факторы (тип облигаций, категория эмитента, характеристики организатора, ключевые макрофинансовые переменные) объясняют лишь незначительную долю вариации спредов, не превышающую 20%, что подчеркивает необходимость разработки моделей, способных учитывать нелинейные взаимодействия и многомерные пространства предикторов.

За последние полтора десятилетия кредитные спреды на российском рынке облигаций демонстрировали выраженную цикличность. Рынок пережил резкое расширение спредов во время финансового кризиса 2014–2015 гг. и снова в 2022 г., когда спреды по эталонным облигациям превысили 800–1 200 базисных пунктов. Периоды финансовой стабилизации, умеренной денежно-кредитной политики и улучшения условий ликвидности привели к устойчивому сужению спредов, особенно в период с 2016 по 2020 г., когда спреды по облигациям инвестиционного класса в среднем составляли примерно 250–350 б.п. – все еще выше сопоставимых уровней на развитых рынках, но заметно ниже, чем в периоды кризиса. После волатильности начала 2022 г. спреды значительно сузились: по состоянию на середину 2025 г. рейтинговые эмитенты торгуются в диапазоне 100–150 б.п., в то время как спреды эмитентов второго и третьего эшелона составляют приблизительно 313 и 675 б.п. соответственно.

Спреды на рынке облигаций инвестиционного уровня в США колеблются в диапазоне 50–200 б.п., что отражает более высокую ликвидность рынка, большую диверсификацию, доверие к отчетности эмитентов и деятельности рейтинговых агентств. По данным ICE³, на июль 2025 г., скорректированный с учетом опционов спред для корпоративных облигаций США инвестиционного класса составляет примерно 83 б.п.

Российские аналоги значительно выше: индексы Cbonds показывают около 417,7 б.п. для общего индекса IFX-Cbonds и 399,9 б.п. для высококачественных эмитентов. Выраженная волатильность кредитных спредов на российском рынке отражает повышенные геополитические риски, более частые макроэкономические потрясения и менее зрелую микроструктуру рынка.

В период с 2022 по 2025 г. российский рынок корпоративного долга вырос более чем вдвое, превысив 30 трлн руб. Такой стремительный рост привел к беспрецедентному увеличению числа эмитентов, объемов выпуска, участия инвесторов и ликвидности рынка. Значительные колебания спредов во время стрессовых ситуаций демонстрируют ограничения факторных линейных регрессий, основанных на небольших наборах предикторов. Такие

³ Ice Data Indices, LLC. ICE BofA US Corporate Index Option-Adjusted Spread. FRED, Federal Reserve Bank of St. Louis, 2026. URL: <https://fred.stlouisfed.org/series/BAMLC0A0CM>

модели часто показывают хорошее соответствие внутри выборки, но плохую точность прогнозирования вне выборки, особенно при наличии структурных разрывов или нелинейных зависимостей [6].

С учетом высокой волатильности российских G-спредов (в стрессовые периоды кредитные спреды высокодоходного сегмента временно вырастали с около 400 б.п. до почти 3 000 б.п.) существует необходимость в более сложных методах прогнозирования. Методы машинного обучения, в частности ансамбли градиентного бустинга, хорошо подходят для этой среды, поскольку они могут учитывать нелинейности, сложные взаимодействия и многомерные данные, обеспечивая при этом высокую точность прогнозирования и устойчивость к шуму. Для институциональных инвесторов, работающих в условиях повышенной неопределенности, такие методы могут улучшить стратегическое управление кредитным и ликвидным риском, предоставляя более надежные прогнозы в широком диапазоне рыночных условий [7].

Обзор литературы

Актуальность использования методов машинного обучения (МО) и искусственного интеллекта (ИИ) для моделирования и прогнозирования кредитных спредов отражается в росте числа авторов которой обсуждают результаты такого использования. К.В. Самохвалова показывает, что российский рынок корпоративных облигаций демонстрирует значительно большую дисперсию, сдвиги режимов и структурные разрывы в сравнении с рынком США, где кредитные спреды часто следуют закономерностям, которые можно достаточно точно описать с помощью традиционных эконометрических моделей [8]. На формирующихся рынках институциональные и частные инвесторы сталкиваются с повышенной неопределенностью и возрастающей потребностью в количественных инструментах, способных учитывать нелинейные зависимости, разнообразные взаимодействия факторов и специфическую структуру российского рынка облигаций.

Эмпирические исследования подчеркивают, что ансамблевые подходы машинного обучения, особенно алгоритмы градиентного бустинга, превосходят метод наименьших квадратов (МНК), метод главных компонент (Principal Component Analysis, PCA), метод частичных наименьших квадратов (Partial Least Squares Regression (PLS)) и другие традиционные методы регрессии в прогнозировании кредитных спредов [2, 9]. Ансамблевые модели повышают точность прогнозирования в многомерных условиях, учитывают мультиколлинеарность, более эффективно обрабатывают пропущенные или зашумленные данные и автоматически учитывают взаимодействие между макроэкономическими, рыночными и специфическими для эмитента переменными. Эти свойства особенно важны на российском рынке, где резкие переходы между режимами риска происходят чаще, а чувствительность спредов к различным группам факторов может быстро меняться в периоды стресса.

Параллельно в литературе все чаще указывается на потенциал искусственных нейронных сетей, включая рекуррентные архитектуры, такие как сети с долгой краткосрочной памятью (Long Short-Term Memory, LSTM) и совсем недавно модели на основе трансформеров для генерации стабильных и точных прогнозов кредитных спредов в условиях, характеризующихся структурной волатильностью и нелинейными процессами генерации данных [9]. Такие модели способны изучать сложные временные зависимости, которые могут быть недостаточно точно отражены параметрическими структурами временных рядов. Хотя нейронные сети требуют больших обучающих наборов данных и могут страдать от ограниче-

ний интерпретируемости, их гибкость делает их перспективным направлением для будущих исследований в области моделирования рисков корпоративных облигаций.

Методы прогнозирования на рынке облигаций

Методы прогнозирования доходности облигаций и кредитных спредов можно условно разделить на несколько классов, каждый из которых охватывает различные аспекты процесса генерации данных и предполагает компромисс между интерпретируемостью, вычислительной сложностью и эффективностью прогнозирования (табл. 1).

1. Классические линейные и факторные модели.

Традиционные эконометрические подходы включают линейные регрессионные модели, факторные регрессии, например PCR и PLS, поперечные или панельные регрессии, а также упрощенные структурные модели. Обобщенная формула линейной модели прогнозирования кредитного спреда может быть выражена следующим образом:

$$G_t = \beta_0 + \sum_{i=1}^n \beta_i X_{i,t} + \varepsilon_t, \quad (1)$$

где G_t – наблюдаемое значение G-спреда в момент времени t ;

$X_{i,t}$ – i -й предиктор в тот же момент времени, включающий рыночные факторы (например, доходность безрисковых облигаций), микроэкономические характеристики эмитента (например, кредитный рейтинг, финансовые коэффициенты), макроэкономические индикаторы (уровень инфляции, ключевая ставка);

β_i – коэффициенты регрессии, характеризующие интенсивность и направление влияния факторов;

ε_t – ошибка, которая учитывает нелинейные эффекты и прочие источники вариации.

Для повышения гибкости модели могут быть включены члены взаимодействия или полиномиальные разложения, хотя это может привести к снижению интерпретируемости и увеличению риска переобучения.

Эмпирические исследования неизменно показывают, что, хотя линейные модели хорошо работают в стабильных условиях, точность их прогнозирования ухудшается, когда базовые данные демонстрируют структурные разрывы, асимметричные реакции на шоки или нелинейные компромиссы между риском и доходностью. Например, обобщенные компонентные модели обычно уступают альтернативам максимального правдоподобия в прогнозировании разброса вне выборки.

2. Стохастические модели и модели временных рядов.

Стохастические модели явно учитывают временную зависимость и включают ARIMA, векторные авторегрессии (VAR), модели семейства GARCH и связанные с ними представления пространства состояний [10, 11]. Стандартная спецификация авторегрессионной интегрированной скользящей средней (ARIMA) может быть записана следующим образом:

$$\phi(B)(G_t - \mu) = \theta(B)\varepsilon_t, \quad (2)$$

где B – оператор лага;

$\phi(B)$ – полином авторегрессии порядка p ;

$\theta(B)$ – полином скользящего среднего порядка q ;

μ – константа;

ε_t – белый шум с нулевым математическим ожиданием и постоянной дисперсией.

Расширения, такие как ARIMAX, включают экзогенные макроэкономические или рыночные переменные. Несмотря на их полезность в моделировании устойчивости, возврата к среднему значению и кластеризации волатильности, стохастические модели требуют жестких предположений относительно стационарности и могут показывать плохие результаты при структурных изменениях режима или нелинейной динамике – условиях, распространенных на развивающихся и геополитически уязвимых рынках облигаций [12].

3. Нелинейные подходы к машинному обучению.

В литературе быстро растет число работ, в которых нелинейные методы машинного обучения применяются для прогнозирования кредитных спредов. К ним относятся нейронные сети, деревья решений, случайные леса и ансамбли градиентного бустинга [13]. Градиентный бустинг, являющийся предметом настоящего исследования, итеративно строит аддитивную модель:

$$\hat{y}_i^{(m)} = \hat{y}_i^{(m-1)} + v \cdot h_m(x_i), \quad (3)$$

где $\hat{y}_i^{(m)}$ – прогноз на m -ом шаге для объекта i ;

$h_m(\cdot)$ – дерево решений, обученное на ошибках предыдущей модели;

v – скорость обучения, которая регулирует величину шага для обновления прогноза.

Бустинг минимизирует заданную функцию потерь шаг за шагом, позволяя модели адаптивно корректировать систематические ошибки прогнозирования [14]. Такие улучшения, как графики скорости обучения или адаптивное сжатие на основе неопределенности, дополнительно повышают устойчивость и снижают риск переобучения.

В литературе отмечается, что методы машинного обучения, включая градиентный бустинг (XGBoost, LightGBM и CatBoost), случайные леса и нейронные сети LSTM, неизменно превосходят классические модели на коротких и на длительных горизонтах прогнозирования [15]. В частности, ансамблевые древовидные методы предлагают ряд преимуществ: высокую производительность при работе с гетерогенными наборами данных, устойчивость к шуму, эффективную обработку категориальных переменных и возможность учета нелинейных и интерактивных эффектов без явного задания модели.

По нашему видению, CatBoost хорошо подходит для моделирования корпоративных облигаций, поскольку он изначально обрабатывает категориальные характеристики эмитентов (отрасль, тип эмитента, класс инструмента), снижает систематическую ошибку прогнозирования за счет упорядоченного бустинга и демонстрирует высокую стабильность при эмпирических выборках, включающих несколько эмитентов с различным качеством раскрытия информации. С учетом ограниченной доступности длинных исторических рядов на российском рынке, градиентный бустинг представляет собой сбалансированную и эмпирически подтвержденную стратегию, сочетающую в себе точность прогнозирования, методологическую гибкость и практическую интерпретируемость.

Градиентно-бустинговые деревья решений (GBDT) представляют собой мощную структуру ансамблевого обучения, в которой слабые базовые классификаторы – как правило, неглубо-

кие деревья решений – обучаются последовательно, так что каждое новое дерево корректирует остаточные ошибки ансамбля. В отличие от классических алгоритмов бустинга, таких как AdaBoost, методы GBDT допускают произвольные дифференцируемые функции потерь, что делает их подходящими как для задач регрессии, так и для задач классификации в широком диапазоне финансовых приложений [16]. Путем итеративной подгонки деревьев к отрицательным градиентам (антиградиентам) функции потерь бустинг создает гибкое, непараметрическое приближение к сложным функциональным зависимостям [17].

Современные реализации градиентного бустинга, включая XGBoost, LightGBM [15] и CatBoost [18], обеспечивают значительную вычислительную эффективность за счет распараллеливания, разбиения на основе гистограмм и расширенной регуляризации. Например, XGBoost оптимизирует последовательность аддитивных деревьев решений:

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i), \quad (4)$$

где f_k – k -е дерево, минимизирующее регуляризованную целевую функцию.

Эти алгоритмы особенно хорошо подходят для задач финансового прогнозирования благодаря своей способности моделировать нелинейности, обрабатывать гетерогенные типы данных и сохранять устойчивость в условиях шума.

Методы GBDT также предоставляют доступ к методам интерпретируемости моделей, таким как оценка важности признаков и значения SHAP, которые играют решающую роль для обеспечения прозрачности в управлении рисками и отчетности перед регулирующими органами.

Создание набора данных, настройка и оценка модели

Эмпирический набор данных состоял из более чем 20 объясняющих переменных, сгруппированных в три факторные категории:

- 1) характеристики эмитента: кредитный рейтинг, объем выпуска, дюрация и структура сроков погашения, отраслевая классификация;
- 2) рыночные переменные: индексы ликвидности облигаций, объемы торгов, уровни индекса IFX-Cbonds, спреды между ценой покупки и продажи (при наличии);
- 3) макроэкономические показатели: ключевая ставка, инфляция, рост ВВП, цены на нефть марки Brent, валютный курс (руб. к долл. США).

Сбор данных проводился ежедневно. 70% наблюдений использовались для обучения модели, а 30% – для тестирования. С учетом временной зависимости, характерной для финансовых временных рядов, оценка модели основывалась на перекрестной проверке временных рядов (скользящее окно / расширяющееся окно), которая обеспечивает более реалистичные оценки производительности по сравнению со случайным перемешиванием.

Для настройки гиперпараметров модели (глубина деревьев, коэффициент обучения, регуляризация) использовалась библиотека Optuna, которая применяет байесовский поиск и обрезку для эффективного исследования многомерных пространств параметров. При необходимости применялись процедуры масштабирования признаков (стандартизация, нормализация) для предотвращения нестабильности градиента в процессе оптимизации.

Эффективность оценивалась с использованием нескольких стандартных показателей:

- R^2 (коэффициент детерминации) – степень объяснения дисперсии целевой переменной;
- MSE (mean squared error) – среднеквадратичная ошибка;
- RMSE – корень из MSE, интерпретируемый в базисных пунктах G-спреда.

Сравнительный анализ также включал линейную регрессию и алгоритм случайного леса, а в некоторых экспериментах – ансамбль Voting Regressor (включавший Ridge, SVM и случайные леса).

Результаты и оценка эффективности

Модель градиентного бустинга продемонстрировала высокую точность прогнозирования на нескольких горизонтах. Для прогноза на один день модель CatBoost показала наиболее точные результаты благодаря усовершенствованной обработке категориальных признаков и регуляризации, которая предотвращает переобучение.

Для целевой переменной на 30 дней модель достигла $R^2 \approx 0,88$, демонстрируя, что нелинейные зависимости между факторами эмитента, рынка и макроэкономическими факторами в совокупности обеспечивают существенную объяснительную силу. Прогнозы оставались стабильными на разных временных горизонтах, что свидетельствует о надежности выявленных взаимосвязей даже в условиях высокой волатильности.

Тем не менее улучшить обучение возможно при помощи подбора параметров; критерий информативности – параметр, используемый для построения дерева, максимальная глубина дерева, минимальное количество объектов для вершины, чтобы разбить данные лучшим образом, минимальное количество объектов в листе дерева, число признаков используемых для предиката в каждой вершине дерева. После подбора лучших гиперпараметров изменилась скорость обучения. Эти результаты согласуются с ранее полученными данными о том, что методы машинного обучения, особенно ансамблевые модели, превосходят линейную регрессию в прогнозировании кредитного спреда в условиях нелинейного режима и высокого уровня шума [5].

Интерпретируемость и анализ SHAP

Для количественной оценки вклада каждого предиктора были вычислены значения SHAP (SHapley Additive exPlanations):

$$\phi_i = \frac{\sum_{S \subseteq N \setminus \{i\}} |S|!(|N|-|S|-1)!}{|N|N!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)], \quad (5)$$

где N – полный набор признаков;

S – подмножество признаков без i -го,

$f_S(\cdot)$ – прогноз модели, обученной на подмножестве признаков S .

Такой подход обеспечивает локальную и глобальную интерпретируемость, позволяя точно определить местоположение прогнозируемых движений спредов.

Обученная модель CatBoost продемонстрировала высокую прогностическую точность, опережающую традиционные алгоритмы. Ключевые результаты:

- $R^2 = 0,78$, что говорит о высокой доле объясненной дисперсии спреда;
- $RMSE \approx 12,5$ б.п., интерпретируемый как средняя ошибка прогноза в терминах G-спреда;
- $MAE < 10$ б.п., что указывает на стабильную ошибку предсказаний.

Анализ важности признаков с использованием SHAP показал, что:

- кредитный рейтинг эмитента объясняет до 43,5% вариации G-спреда;
- отраслевая принадлежность и объем выпуска занимают вторую и третью позиции по значимости;
- макроэкономические показатели (ключевая ставка, инфляция) внесли вклад в размере около 10,5%, что согласуется с наблюдением о том, что кредитные спреды в России более чувствительны к особенностям отдельных эмитентов, чем к общим макроэкономическим тенденциям.

Визуализация ошибок валидации подтвердила стабильную сходимость примерно после 100–200 итераций бустинга.

Результаты исследования подчеркивают способность CatBoost учитывать неоднородность эмитентов, условия ликвидности и нелинейные взаимодействия факторов – критически важные факторы, которые часто упускаются из виду классическими эконометрическими моделями.

Сравнение с линейной регрессией показало ее чувствительность к мультиколлинеарности и неспособность улавливать сложные нелинейные паттерны, особенно в периоды высокой рыночной волатильности. CatBoost же демонстрировал устойчивость к выбросам, неполным данным и автокоррелированным наблюдениям. В *табл. 2* представлено сравнение бустинга с альтернативными методами.

Рекомендации по использованию результатов

Надежные прогнозы кредитных спредов играют решающую роль в управлении портфельными рисками, соблюдении нормативных требований, планировании ликвидности и тактическом или стратегическом распределении активов. Изложим соображения относительно применения рассмотренных методов прогноза.

Для институциональных инвесторов.

1. Оптимизация портфеля и выбор облигаций: прогнозируемые спреды помогают выявлять недооцененные облигации с чрезмерно высокими премиями за риск, что позволяет разрабатывать стратегии, направленные на повышение ожидаемой доходности на единицу риска.
2. Хеджирование кредитных спредов: точные прогнозы спредов позволяют своевременно использовать производные инструменты или альтернативные инструменты с фиксированным доходом для снижения рисков в периоды повышенной волатильности.
3. Соблюдение нормативных требований: банки могут интегрировать прогнозы G-спреда во внутренние модели достаточности капитала и системы раннего предупреждения для прогнозирования ухудшения кредитного качества.

Для управляющих активами и инвестиционных фондов.

1. Улучшение коэффициента Шарпа за счет динамической перебалансировки на основе ожидаемых движений спреда.
2. Анализ сценариев с использованием имитационных моделей «от прошлого к будущему», включающих прогнозы, основанные на машинном обучении.

Для розничных инвесторов.

1. Более обоснованное принятие решений на основе сравнения рыночных спредов с расчетными справедливыми спредами, полученными с помощью модели.
2. Избегание приобретения завышенных по цене или ухудшающихся по кредитному качеству инструментов.

Для включения этих прогнозов в институциональную инфраструктуру компаниям следует:

- внедрять автоматизированные каналы передачи данных по облигациям, макроэкономическим показателям и рыночным данным;
- обновлять прогнозы модели (ежедневно или в течение дня);
- разрабатывать панели мониторинга для визуализации прогнозов спредов, вклада SHAP, сигналов ликвидности и оповещений о трендах;
- интегрировать прогнозы в процессы оценки кредитоспособности в качестве «рыночного фактора риска».

Модульность алгоритмов бустинга позволяет настраивать функции потерь (MAE, Huber, log-cosh) под конкретные сценарии использования, такие как анализ чувствительности к риску в хвостовой части распределения или прогнозирование в стрессовых сценариях.

Заключение

Данное исследование посвящено одной из центральных задач современной финансовой аналитики: точному прогнозированию кредитных спредов на российском рынке облигаций, где волатильность, зависимость от режима и структурная неоднородность ослабляют эффективность традиционных эконометрических моделей. С помощью алгоритмов градиентного бустинга – в частности, CatBoost, XGBoost и LightGBM – исследование демонстрирует, что современные методы машинного обучения могут успешно моделировать нелинейные взаимодействия между переменными на уровне эмитента, рынка и макроэкономическими переменными.

На базе использования обширных ежедневных данных за 2018–2025 гг. модель CatBoost достигла высокой точности прогнозирования ($R^2 = 0,78$, $MAE < 10$ б.п.) и превзошла классические регрессии и базовые модели деревьев решений. Анализ интерпретируемости на основе SHAP показал, что характеристики качества эмитента доминируют в структуре кредитных спредов в России, в то время как макроэкономические факторы, хотя и значимые, играют второстепенную роль.

Полученные результаты подтверждают гипотезу о том, что ансамблевые модели машинного обучения обладают существенными преимуществами перед линейными моделями на нестабильных развивающихся рынках. Они также закладывают методологическую основу

для интеграции моделей на основе ИИ в управление рисками, принятие инвестиционных решений и финансовое регулирование.

Касаясь вопроса перспектив будущих исследований, укажем на следующие:

- расширение набора функций за счет использования обработки естественного языка при анализе информации, раскрываемой эмитентами, и оценок новостных сообщений;
- сочетание градиентного бустинга со структурными или упрощенными моделями кредитного риска;
- разработка гибридных ансамблевых архитектур на основе нейронных сетей для улучшения прогнозирования на долгосрочный период.

В целом исследование способствует модернизации и трансформации аналитических инструментов в российской финансовой системе, повышает прозрачность, адаптивность и устойчивость рынков облигаций.

Таблица 1

Классификация методов прогнозирования доходности облигаций и кредитных спредов

Table 1

Classification of methods for forecasting bond yields and credit spreads

Метод	Класс	Основные преимущества
OLS-регрессия, PCR, PLS	Линейные регрессии	Простота, интерпретируемость
VAR, ARIMA, GARCH	Стохастические модели	Учет автокорреляции, волатильности
Нейросети (MLP, LSTM)	Машинное обучение	Выявление сложных нелинейных связей
Случайный лес (Random Forest)	Ансамблевые деревья	Стабильность, автоматический отбор признаков
Градиентный бустинг (GBDT)	Ансамблевые деревья	Устойчивость к шуму, интерпретируемость

Источник: авторская разработка

Source: Authoring

Таблица 2
Обзор архитектуры нейронных сетей

Table 2
Overview of neural network architectures

Архитектура	Описание	Ограничения
Рекуррентные нейронные сети (RNN) (LSTM и GRU)	Демонстрируют высокую эффективность при работе с последовательными финансовыми данными, где критично учитывать автокорреляции, сезонность и нелинейные эффекты, способны запоминать долгосрочные зависимости в структуре спредов	Требуют значительных объемов обучающей выборки, чувствительны к параметризации и имеют ограниченную интерпретируемость*
Глубокие нейросетевые модели с механизмом внимания (attention-based models)	Обеспечивают высокую точность прогноза за счет способности выделять ключевые элементы временной последовательности	Использование в прикладных задачах ограничивается высокой вычислительной сложностью и необходимостью специализированной экспертизы**
Многомерные GARCH-модели и другие продвинутые эконометрические конструкции	Позволяют учитывать условную волатильность и корреляционные зависимости между доходностями различных активов	Чувствительны к спецификации модели, сложности параметрической оценки и ограниченной масштабируемости***

* Петрова Д.А., Трунин П.В Прогнозирование основных финансовых показателей с использованием RNN. URL: <https://ru.readkong.com/page/petrova-d-a-trunin-p-v-prognostirovanie-osnovnyh-9724988>

** Искусство макропрогнозирования: модели, эксперты и «большие данные». Доклад НИУ ВШЭ. М.: НИУ ВШЭ, 2024.128 с.

*** Султанов И.Р. Оценка факторов, влияющих на спред доходности корпоративных облигаций стран БРИКС. URL: <https://aspirantura.hse.ru/data/2020/02/14/1573456014/Султанов.pdf>

Источник: авторская разработка

Source: Authoring

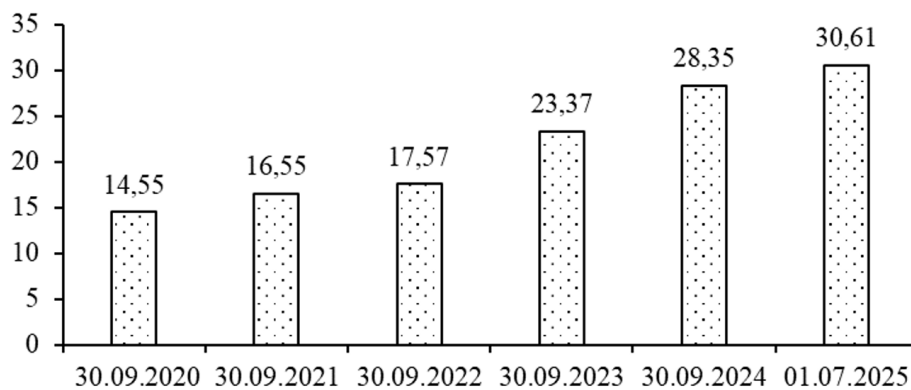
Таблица 3
Сопоставление градиентного бустинга с альтернативами

Table 3
Comparison of gradient boosting with alternatives

Критерий	Линейные модели	Случайный лес	Градиентный бустинг
Учет нелинейности	–	+	+
Интерпретируемость	+	+(умеренная)	+(при использовании SHAP)
Устойчивость к шуму	–	+	+(с регуляризацией)
Точность на табличных данных	+	+	++
Возможность применения на российском рынке	+	+	++

Источник: авторская разработка по [9, 10]

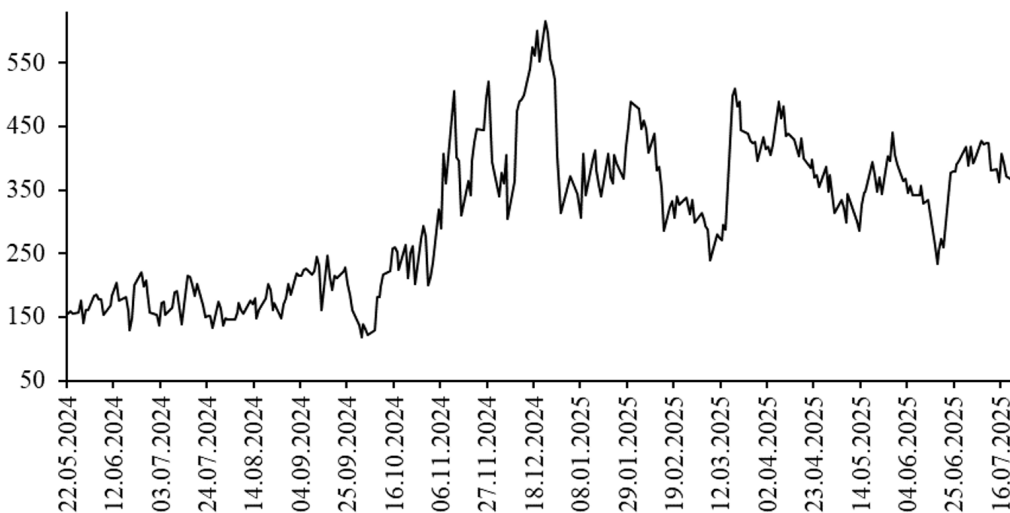
Source: Authoring, based on [9, 10]

Рисунок 1**Динамика объема российского рынка корпоративных облигаций, трлн руб.****Figure 1****Dynamics of the Russian corporate bond market volume, trillion RUB**

Источник: авторская разработка по данным Никонов М., Галиева Г. Запас прочности истощается: обзор долгового рынка 2024–202.

URL: https://rusbonds.ru/rb-docs/analytics/RA_expert_1730276051.pdf

Source: Authoring, based on data from Nikonov M., Galiyeva G. Safety margin is depleting: debt market overview 2024–202. URL: https://rusbonds.ru/rb-docs/analytics/RA_expert_1730276051.pdf

Рисунок 2**Динамика индекса IFX-Cbonds G-Spread****Figure 2****Dynamics of the IFX-Cbonds G-Spread index**

Источник: авторская разработка по данным IFX-Cbonds G-Spread. Cbonds.ru

URL: <https://cbonds.ru/indexes/IFX-Cbonds-G-Spread/>

Source: Authoring, based on data from IFX-Cbonds G-Spread. Cbonds.ru

URL: <https://cbonds.ru/indexes/IFX-Cbonds-G-Spread/>

Список литературы

1. Козлов С.В., Седенков С.А. Анализ LSTM и GRU моделей для построения прогнозов временных рядов // *International Journal of Open Information Technologies*. 2024. Т. 12. № 7. С. 43–50. EDN: DMKFKO
2. Карминский А.М., Дьячкова Н.Ф. Исследование взаимосвязи кредитных циклов с изменениями кредитных рейтингов // *Журнал Новой экономической ассоциации*. 2020. № 4. С. 138–161. DOI: 10.31737/2221-2264-2020-48-4-6 EDN: LVYGXP
3. Косников С.Н., Демченко Т.С., Молчанов Ф.А. Экономические циклы и их влияние на финансовые рынки // *Естественно-гуманитарные исследования*. 2025. № 2. С. 588–593. EDN: ERHWUU
4. Ерофеева Т.М. Прогнозирование спреда доходности на российском долговом рынке // *Вестник Московского университета. Серия 6: Экономика*. 2021. № 6. С. 54–76. DOI: 10.38050/0130010520216.3 EDN: MTFLWJ
5. Gu S., Kelly B., Xiu D. Empirical asset pricing via machine learning. *Review of Financial Studies*, 2020, vol. 33, iss. 5, pp. 2223–2273. DOI: 10.1093/rfs/hhaa009 EDN: DCWZEU
6. Wang S., Chen B. Credit card attrition: An overview of machine learning and deep learning techniques // *Информатика. Экономика. Управление*. 2023. Т. 2. № 4. С. 134–144. DOI: 10.47813/2782-5280-2023-2-4-0134-0144 EDN: YQLKTW
7. Коростелкина И.А., Дедкова Е.Г., Коростелкин М.М., Маматкулова О.И. Применение методов оценки и прогнозирования кредитной деятельности в системе управления рисками в коммерческом банке // *Финансовые рынки и банки*. 2025. № 5. С. 255–262. DOI: 10.24412/2658-3917-2025-5-255-262 EDN: NXQNXU
8. Самохвалова К.В. Особенности развития рынка корпоративных облигаций в современной России // *Международный журнал гуманитарных и естественных наук*. 2020. № 3-2. С. 60–65. DOI: 10.24411/2500-1000-2020-10262 EDN: JFUBNH
9. Machado M.R., Osterrieder J., Chen D. Forecasting Commercial Customers Credit Risk Through Early Warning Signals Data: A Machine Learning based Approach. *SSRN Electronic Journal*, 2024. DOI: 10.2139/ssrn.4754568
10. Галимнуров А.А., Исмагилова А.С. Применение математического моделирования для прогнозирования спреда корпоративных облигаций // *Инженерный вестник Дона*. 2025. № 2. С. 721–737. EDN: PDEGSR
11. Шильцина Е.А. Применение машинного обучения для прогнозирования временных рядов // *Вестник науки*. 2025. Т. 2. № 5. С. 953–960. EDN: GHFRIT
12. Aghayev N.B., Nazarli D.Sh. Hybrid forecasting model of non-scheduled passenger air transportation // *Научный вестник МГТУ ГА*. 2025. Т. 28. № 2. С. 8–21. DOI: 10.26467/2079-0619-2025-28-2-8-21 EDN: GMDLDM
13. Dzhunkeev U. Forecasting inflation in Russia using gradient boosting and neural networks. *Russian Journal of Money and Finance*, 2024, vol. 83, no. 1, pp. 53–76. EDN: IANEFK
14. Berg F., Kölbel J.F., Rigobon R. Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 2022, vol. 26, iss. 6, pp. 1315–1344. DOI: 10.1093/rof/rfac033 EDN: BRYCLT

15. Oukhouya H., Kadiri H., El Himdi Kh., Guerbaz R. Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, and Backtesting XGBoost Models. *Statistics, Optimization and Information Computing*, 2023, vol. 12, iss. 1, pp. 200–209. DOI: 10.19139/soic-2310-5070-1822 EDN: QVDXWM
16. Abbasov R. Revolutionizing risk management in banking: Implementation of AI/ML-based gradient boosting machines (GBM) and random forest models for credit risk management. *International Journal of Research in Finance and Management*, 2023, vol. 6, iss. 1, pp. 441–444. DOI: 10.33545/26175754.2023.v6.i1d.324 EDN: SBKDCJ
17. Konstantinov A.V. Deep gradient boosting for regression problems. *Computing, Telecommunications and Control*, 2021, vol. 14, iss. 3, pp. 7–19. DOI: 10.18721/JCSTCS.14301 EDN: CZBFQN
18. Peykani P., Peymany Foroushany M., Tanasescu C. et al. Evaluation of Cost-Sensitive Learning Models in Forecasting Business Failure of Capital Market Firms. *Mathematics*, 2025, vol. 13, iss. 3. DOI: 10.3390/math13030368 EDN: LDAOKC

Информация о конфликте интересов

Мы, авторы данной статьи, со всей ответственностью заявляем о частичном и полном отсутствии фактического или потенциального конфликта интересов с какой бы то ни было третьей стороной, который может возникнуть вследствие публикации данной статьи. Настоящее заявление относится к проведению научной работы, сбору и обработке данных, написанию и подготовке статьи, принятию решения о публикации рукописи.

FORECASTING CORPORATE BOND CREDIT SPREADS USING GRADIENT BOOSTING ENSEMBLES

DOI: <https://doi.org/10.24891/elhfng>

EDN: <https://elibrary.ru/elhfng>

Konstantin V. KRINICHANSKII

Corresponding author, Financial University under Government of Russian Federation,
Moscow, Russian Federation

e-mail: kkrin@ya.ru

ORCID: 0000-0002-1225-7263

Artem A. KOBZEV

PAO Moscow Exchange, Moscow, Russian Federation

e-mail: artem.kobzev.2001@mail.ru

ORCID: not available

Aleksandra Yu. PISKAREVA

Financial University under Government of Russian Federation, Moscow, Russian Federation

e-mail: 221236@edu.fa.ru

ORCID: 0009-0007-9722-1094

Article history:

Article No. 844/2025

Received 17 Dec 2025

Accepted 12 Jan 2026

Available online

28 May 2026

JEL Classification:

C45, G12, G17

Keywords: corporate bond market, credit risk, credit spread, machine learning, gradient boosting

Abstract

Subject. Gradient boosting algorithms for forecasting credit spreads of corporate bonds in the Russian debt instrument market.

Objectives. To develop and empirically test a concept for building predictive models of credit spreads.

Methods. Methods of multivariate statistical analysis, regression modelling and ensemble machine learning were applied. Gradient boosting models were used as the main toolkit. To assess the contribution and significance of predictors, the SHapley Additive exPlanations (SHAP) interpretation methodology was employed.

Results. It has been established that the use of ensemble gradient boosting models provides a significant improvement in the accuracy of credit spread forecasting compared to traditional regression approaches. The developed models identify complex nonlinear interactions between factors at various levels and demonstrate stability during periods of market shocks. Structured approaches to integrating micro- and macroeconomic indicators into a unified forecasting model have been proposed.

Conclusions. Gradient boosting, particularly the CatBoost model, is effective as a tool for analyzing and forecasting credit spreads. The applicability of the method in financial analytics, risk management and supporting investment decisions has been substantiated. Limiting factors include the dependence of forecast quality on the completeness and reliability of the input data, as well as the need for regular model updates. The proposed approaches could be implemented in systems for automated investment decision-making, credit scoring and stress testing, aimed at enhancing the transparency and stability of the Russian corporate debt market.

Please cite this article as: Krnichanskii K.V., Kobzev A.A., Piskareva A.Yu. Forecasting corporate bond credit spreads using gradient boosting ensembles. *Finance and Credit*, 2026, iss. 5, pp. 124–141. DOI: 10.24891/elhfng EDN: ELHFNG

References

1. Kozlov S.V., Sedenkov S.A. [LSTM and GRU model analysis for time series forecasting]. *International Journal of Open Information Technologies*, 2024, vol. 12, no. 7, pp. 43–50. (In Russ.) EDN: DMKFKO
2. Karminsky A.M., Dyachkova N.F. [Empirical study of the relationship between credit cycles and changes in credit ratings]. *Zhurnal Novoi ekonomicheskoi assotsiatsii*, 2020, no. 4, pp. 138–161. (In Russ.) DOI: 10.31737/2221-2264-2020-48-4-6 EDN: LVYGXP
3. Kosnikov S.N., Demchenko T.S., Molchanov F.A. [Economic cycles and their impact on financial markets]. *Estestvenno-gumanitarnye issledovaniya*, 2025, no. 2, pp. 588–593. (In Russ.) EDN: ERHWUU
4. Erofeeva T.M. [Predicting the yield spreads on Russia's debt market]. *Vestnik Moskovskogo universiteta. Seriya 6: Ekonomika*, 2021, no. 6, pp. 54–76. (In Russ.) DOI: 10.38050/0130010520216.3 EDN: MTFLWJ
5. Gu S., Kelly B., Xiu D. Empirical asset pricing via machine learning. *Review of Financial Studies*, 2020, vol. 33, iss. 5, pp. 2223–2273. DOI: 10.1093/rfs/hhaa009 EDN: DCWZEU
6. Wang S., Chen B. Credit card attrition: An overview of machine learning and deep learning techniques. *Informatika. Ekonomika. Upravlenie*, 2023, vol. 2, no. 4, pp. 134–144. (In Russ.) DOI: 10.47813/2782-5280-2023-2-4-0134-0144 EDN: YQLKTW
7. Korostelkina I.A., Dedkova E.G., Korostelkin M.M., Mamatkulova O.I. [Application of methods of assessment and forecasting of credit activity in the risk management system in a commercial bank]. *Finansovye rynki i banki*, 2025, no. 5, pp. 255–262. (In Russ.) DOI: 10.24412/2658-3917-2025-5-255-262 EDN: NXQNXU
8. Samokhvalova K.V. [Peculiarities of the corporate bonds market development in modern Russia]. *Mezhdunarodnyi zhurnal humanitarnykh i estestvennykh nauk*, 2020, no. 3-2, pp. 60–65. (In Russ.) DOI: 10.24411/2500-1000-2020-10262 EDN: JFUBHH
9. Machado M.R., Osterrieder J., Chen D. Forecasting Commercial Customers Credit Risk Through Early Warning Signals Data: A Machine Learning based Approach. *SSRN Electronic Journal*, 2024. DOI: 10.2139/ssrn.4754568
10. Galimnurov A.A., Ismagilova A.S. [The application of mathematical modeling for forecasting corporate bond spreads]. *Inzhenernyi vestnik Dona*, 2025, no. 2, pp. 721–737. (In Russ.) EDN: PDEGSR
11. Shiltsina E.A. [Application of machine learning for time series forecasting]. *Vestnik nauki*, 2025, vol. 2, no. 5, pp. 953–960. (In Russ.) EDN: GHFRIT
12. Aghayev N.B., Nazarli D.Sh. [Hybrid forecasting model of non-scheduled passenger air transportation]. *Nauchnyi vestnik MGTU GA*, 2025, vol. 28, no. 2, pp. 8–21. (In Russ.) DOI: 10.26467/2079-0619-2025-28-2-8-21 EDN: GMDLDM
13. Dzhunkeev U. Forecasting inflation in Russia using gradient boosting and neural networks. *Russian Journal of Money and Finance*, 2024, vol. 83, no. 1, pp. 53–76. EDN: IANEFK

14. Berg F., Kölbel J.F., Rigobon R. Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 2022, vol. 26, iss. 6, pp. 1315–1344. DOI: 10.1093/rof/rfac033 EDN: BRYCLT
15. Oukhouya H., Kadiri H., El Himdi Kh., Guerbaz R. Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, and Backtesting XGBoost Models. *Statistics, Optimization and Information Computing*, 2023, vol. 12, iss. 1, pp. 200–209. DOI: 10.19139/soic-2310-5070-1822 EDN: QVDXWM
16. Abbasov R. Revolutionizing risk management in banking: Implementation of AI/ML-based gradient boosting machines (GBM) and random forest models for credit risk management. *International Journal of Research in Finance and Management*, 2023, vol. 6, iss. 1, pp. 441–444. DOI: 10.33545/26175754.2023.v6.i1d.324 EDN: SBKDCJ
17. Konstantinov A.V. Deep gradient boosting for regression problems. *Computing, Telecommunications and Control*, 2021, vol. 14, iss. 3, pp. 7–19. DOI: 10.18721/JCSTCS.14301 EDN: CZBFQN
18. Peykani P., Peymany Foroushany M., Tanasescu C. et al. Evaluation of Cost-Sensitive Learning Models in Forecasting Business Failure of Capital Market Firms. *Mathematics*, 2025, vol. 13, iss. 3. DOI: 10.3390/math13030368 EDN: LDAOKC

Conflict-of-interest notification

We, the authors of this article, bindingly and explicitly declare of the partial and total lack of actual or potential conflict of interest with any other third party whatsoever, which may arise as a result of the publication of this article. This statement relates to the study, data collection and interpretation, writing and preparation of the article, and the decision to submit the manuscript for publication.