

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

**АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ  
ПРОГНОЗИРОВАНИЕ НА РЫНКЕ ЦЕННЫХ БУМАГ**

Студента 4 курса 412 группы  
направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета

Максимова Олега Алексеевича

Научный руководитель

доцент, к. ф.-м. н., доцент

\_\_\_\_\_

М. Г. Плешаков

Заведующий кафедрой

д. ф.-м. н., доцент

\_\_\_\_\_

С. П. Сидоров

Саратов 2026

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	3
1 ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ .....	5
ЗАКЛЮЧЕНИЕ .....	14

## ВВЕДЕНИЕ

Современный рынок ценных бумаг представляет собой сложную динамическую систему, характеризующуюся высокой волатильностью и значительной степенью неопределённости. В условиях нестабильной макроэкономической среды способность корректно прогнозировать динамику цен финансовых активов приобретает особую значимость как для институциональных, так и для частных инвесторов.

**Актуальность исследования** обусловлена возрастающим интересом к инструментам фондового рынка на фоне расширения доступа к биржевой торговле и цифровизации финансовых сервисов. Рост числа частных инвесторов сопровождается увеличением рыночных рисков, что делает необходимым применение строгих математических методов анализа и прогнозирования. Особую значимость данное исследование приобретает применительно к российскому фондовому рынку, обладающему собственной структурной спецификой и чувствительностью к внешним экономическим факторам.

**Целью работы** является проведение сравнительного анализа эффективности статистических методов и современных алгоритмов машинного обучения при прогнозировании стоимости акций ПАО «Газпром» на различных временных горизонтах.

Для достижения поставленной цели были сформулированы следующие **задачи**:

1. Изучить теоретические основы функционирования рынка ценных бумаг и механизмы формирования биржевых котировок.
2. Рассмотреть математический аппарат и принципы функционирования моделей прогнозирования временных рядов: классической модели ARIMA, метода опорных векторов в задачах регрессии (SVR), ансамблевых алгоритмов Random Forest и XGBoost, а также рекуррентных нейронных сетей типа LSTM.
3. Разработать программную реализацию выбранных моделей на языке Python с использованием библиотек машинного обучения.
4. Провести вычислительные эксперименты по построению прогнозов цен акций ПАО «Газпром» на краткосрочные и среднесрочные периоды (10,

20 и 30 торговых дней).

5. Выполнить сравнительную оценку точности моделей на основе количественных метрик качества (MAPE, RMSE) и определить наиболее эффективный подход для каждого временного горизонта прогнозирования.

**Объектом исследования** является динамика котировок акций ПАО «Газпром» на российском фондовом рынке.

**Предметом исследования** выступают математические модели и алгоритмы прогнозирования финансовых временных рядов, применяемые для оценки будущей стоимости ценных бумаг.

# 1 ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

## Историческое развитие и современное состояние рынка ценных бумаг

В первой главе работы рассматриваются этапы становления и развития фондового рынка как института перераспределения капитала и механизма формирования рыночной стоимости финансовых активов.

Фондовый рынок представляет собой систему экономических и правовых отношений, обеспечивающих выпуск, обращение и оценку ценных бумаг. Его формирование началось в странах Западной Европы в XVII веке с появлением организованной торговли акциями на Амстердамской бирже. Создание вторичного рынка долевых ценных бумаг стало ключевым шагом в развитии механизмов мобилизации капитала и формирования рыночной цены активов.

В XVIII–XIX веках происходит институционализация биржевой торговли и формирование крупнейших мировых площадок — Лондонской и Нью-Йоркской фондовых бирж. XX век характеризуется развитием механизмов регулирования, усилением роли финансовых посредников и переходом к электронной форме торговли. Цифровизация торговых операций существенно повысила скорость обращения информации и усилила роль количественных методов анализа.

Российский рынок ценных бумаг имеет собственную специфику развития. После дореволюционного этапа активного функционирования бирж торговля ценными бумагами была фактически прекращена в советский период. Возрождение рыночной инфраструктуры началось в конце 1980-х годов и сопровождалось формированием биржевых институтов, приватизацией государственных предприятий и развитием системы регулирования.

Ключевыми этапами современного развития стали создание Московской межбанковской валютной биржи, преодоление кризиса 1998 года и объединение торговых площадок в Московскую биржу в 2011 году. В последние годы российский рынок функционирует в условиях санкционного давления и геополитической нестабильности, что усиливает волатильность финансовых инструментов и повышает значимость методов количественного анализа.

Рост числа частных инвесторов и увеличение объёмов торгов подтверждают актуальность разработки математических моделей прогнозирования

динамики цен, что обосновывает выбор темы исследования.

### **Классификация ценных бумаг**

В работе рассматривается классификация финансовых инструментов по уровню инвестиционного риска, что имеет принципиальное значение при анализе динамики рыночных активов.

Ценная бумага представляет собой финансовый актив, удостоверяющий имущественные права её владельца. В рамках математического моделирования ценная бумага характеризуется такими количественными параметрами, как цена, доходность, волатильность и ликвидность.

В соответствии с теорией портфельных инвестиций (Г. Марковиц) и моделью САРМ (У. Шарп) выделяется понятие безрисковой ставки доходности, относительно которой оценивается премия за риск. На основании данного подхода все ценные бумаги условно подразделяются на безрисковые и рискованные.

**Безрисковые ценные бумаги** характеризуются минимальной вероятностью дефолта и относительно низкой волатильностью. К ним относятся:

- *Государственные облигации* (в Российской Федерации — ОФЗ), представляющие собой долговые обязательства государства;
- *Муниципальные облигации*, выпускаемые органами региональной и местной власти;
- *Банковские депозиты с государственной системой страхования вкладов*, обеспечивающие возврат средств в пределах установленного лимита;
- *Облигации государственных корпораций*, эмитируемые организациями с существенным государственным участием.

Следует отметить, что термин «безрисковые» носит условный характер, поскольку данные инструменты подвержены инфляционному, процентному и валютному рискам.

**Рисковые ценные бумаги** предполагают наличие рыночного риска и неопределённости доходности. К данной категории относятся:

- *Акции* — долевые ценные бумаги, предоставляющие право на участие в управлении компанией и получение дивидендов (обыкновенные и при-

- вилегированные);
- *Корпоративные облигации* — долговые обязательства коммерческих организаций;
  - *Производные финансовые инструменты* (фьючерсы, опционы, свопы), стоимость которых зависит от базового актива;
  - *Паи инвестиционных фондов и ETF*, представляющие инструменты коллективного инвестирования.

Стоимость рискованных активов определяется совокупностью макроэкономических, отраслевых и поведенческих факторов, что приводит к высокой изменчивости их рыночной цены.

В рамках настоящего исследования объектом моделирования выступают акции публичной компании как наиболее чувствительный к рыночной информации финансовый инструмент, что делает их репрезентативными для анализа методов прогнозирования временных рядов.

## **Математические методы и модели прогнозирования динамики фондового рынка**

В работе рассматриваются три основных подхода к анализу динамики цен активов: фундаментальный анализ, технический анализ и методы математического моделирования.

Фундаментальный анализ основывается на изучении макроэкономических показателей и финансовой отчетности эмитента. Технический анализ использует статистические закономерности, выявляемые в динамике цен и объёмов торгов. В настоящем исследовании основной акцент сделан на методах математического моделирования и машинного обучения, позволяющих выявлять скрытые зависимости во временных рядах.

Применение количественных методов связано с гипотезой эффективного рынка (EMH), согласно которой текущая цена отражает доступную информацию. Сравнение линейных и нелинейных моделей позволяет оценить степень предсказуемости временного ряда.

### **Модель ARIMAX**

Модель  $ARIMAX(p, d, q)$  учитывает как внутреннюю динамику временного ряда, так и влияние внешних факторов. В общем виде она записывается

следующим образом:

$$\Phi_p(L)(1-L)^d X_t = \sum_{i=1}^k \beta_i (1-L)^d Y_{i,t} + \Theta_q(L)\varepsilon_t \quad (1)$$

где  $L$  — оператор лага ( $L^k X_t = X_{t-k}$ ),  $\Phi_p(L)$  и  $\Theta_q(L)$  — полиномы авторегрессии и скользящего среднего,  $\varepsilon_t$  — белый шум.

Выбор структуры модели осуществляется на основе информационного критерия Шварца:

$$BIC = \ln(n)k - 2\ln(\hat{L}) \quad (2)$$

Модель предполагает линейность зависимости, стационарность процесса и отсутствие мультиколлинеарности экзогенных факторов.

### Метод опорных векторов (SVR)

Для обучающей выборки  $\{(y_i, \mathbf{x}_i)\}_{i=1}^n$  задача SVR формулируется как задача минимизации:

$$\frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \rightarrow \min_{\mathbf{w}, b, \xi, \xi^*} \quad (3)$$

при ограничениях:

$$\begin{cases} y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b \leq \varepsilon + \xi_i \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (4)$$

Для учета нелинейности используется ядерная функция, в частности радиально-базисное ядро:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (5)$$

Алгоритм обеспечивает структурную минимизацию риска и устойчивость к шуму.

## Метод случайного леса (Random Forest)

Random Forest представляет собой ансамбль решающих деревьев, построенных на бутстрап-выборках. В задачах регрессии разбиение узлов осуществляется по критерию минимизации дисперсии:

$$Q(X, j, s) = \frac{1}{|X|} \left( \sum_{x_i \in X} (y_i - \bar{y})^2 - \sum_{x_i \in X_L} (y_i - \bar{y}_L)^2 - \sum_{x_i \in X_R} (y_i - \bar{y}_R)^2 \right) \quad (6)$$

Итоговый прогноз формируется усреднением:

$$\hat{y}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N b_i(\mathbf{x}, \theta_i) \quad (7)$$

Метод снижает дисперсию модели и повышает устойчивость к шумам.

## Градиентный бустинг (XGBoost)

XGBoost реализует последовательное построение ансамбля с регуляризацией.

Целевая функция имеет вид:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (8)$$

Оптимальный вес листа определяется как:

$$w_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (9)$$

где  $g_i$  и  $h_i$  — градиент и гессиан функции потерь. Метод направлен на последовательное уменьшение смещения модели.

## Рекуррентные нейронные сети LSTM

Архитектура LSTM предназначена для моделирования временных зависимостей. Работа ячейки описывается системой уравнений:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (10)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad \tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (11)$$

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (12)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad h_t = o_t \tanh(c_t) \quad (13)$$

Обучение осуществляется методом обратного распространения ошибки во времени (ВРТТ). Модель способна выявлять долгосрочные зависимости во временных рядах.

## Практическая часть

В практической части работы проведено сравнительный анализ моделей прогнозирования финансовых временных рядов на примере акций ПАО «Газпром».

В качестве целевой переменной использовалась доходность акций, а не цена. Такой выбор обусловлен нестационарностью ценовых временных рядов, тогда как доходности в большинстве случаев обладают свойством стационарности или близки к нему. Приведение данных к стационарному виду соответствует требованиям эконометрики временных рядов и позволяет снизить риск переобучения моделей.

Доходность рассчитывалась как относительное изменение цены:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (14)$$

Модели обучались на исторических значениях доходности. Для оценки точности прогнозирования использовалась метрика средней абсолютной процентной ошибки (MAPE):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|. \quad (15)$$

В качестве базовой модели сравнения применялся наивный прогноз, предполагающий сохранение последнего наблюдаемого значения.

Несмотря на то, что модели обучались на доходностях, итоговый прогноз формировался в терминах цены. Для этого использовалась рекурсивная схема восстановления ценового ряда. Если  $P_0$  — последняя известная фактическая цена, а  $r_t$  — спрогнозированная доходность, то восстановление цены осуществлялось по формуле:

$$P_t = P_{t-1}(1 + r_t). \quad (16)$$

Сравнение моделей проводилось на горизонтах 10, 20 и 30 торговых дней. Итоговые значения MAPE представлены в таблице 1.

Таблица 1 – Сравнение моделей по метрике MAPE (%)

Модель	10 дней	20 дней	30 дней
Naive (базлайн)	1.80	1.95	2.25
<b>XGBoost</b>	<b>1.46</b>	2.33	2.07
Random Forest	1.69	2.44	2.12
SVR	<b>1.01</b>	3.42	3.16
LSTM	1.41	2.47	<b>2.00</b>
ARIMAX	1.07	<b>2.15</b>	7.65

Анализ полученных результатов показывает различную динамику поведения моделей в зависимости от горизонта прогнозирования.

На краткосрочном горизонте (10 дней) наилучший результат продемонстрировала модель SVR (1.01%), что свидетельствует о её способности эффективно улавливать локальные паттерны доходности и превосходить наивный базлайн (1.8%).

Однако при увеличении горизонта прогнозирования ошибка модели SVR существенно возрастает (до 3.16–3.42% на 20–30 днях), что указывает на накопление ошибки при рекурсивном построении прогноза и консервативный характер предсказаний.

Модели XGBoost и Random Forest демонстрируют наиболее устойчивые результаты на всех горизонтах. В частности, ошибка XGBoost возрастает умеренно при переходе от 10 к 30 дням (1.46% → 2.07%), оставаясь ниже показателя наивного прогноза (2.20%), что свидетельствует о высокой стабильности ансамблевого подхода и его способности обобщать нелинейные зависимости.

Модель LSTM показывает сопоставимую с XGBoost стабильность (1.41% → 2.00%), подтверждая эффективность рекуррентных архитектур для работы с последовательными финансовыми данными, хотя на коротких горизонтах она незначительно уступает градиентному бустингу.

Классическая линейная модель ARIMAX демонстрирует приемлемую точность на коротких интервалах (1.07–2.15%), однако на горизонте 30 дней ошибка резко возрастает до 7.65%, что подтверждает ограниченность линейных зависимостей при моделировании нестационарных финансовых рядов с меняющейся волатильностью.

Наивный прогноз остаётся конкурентоспособным ориентиром, особенно на длинных горизонтах, что согласуется со слабой формой гипотезы эффективного рынка и подчёркивает фундаментальную сложность задачи прогнозирования.

Проведённый анализ показывает, что при выборе модели необходимо учитывать не только минимальное значение ошибки на отдельном горизонте, но и устойчивость результатов при увеличении периода прогнозирования. С этой точки зрения наиболее сбалансированные и стабильные показатели продемонстрировали ансамблевые методы и рекуррентные сети, прежде всего XGBoost и LSTM.

## ЗАКЛЮЧЕНИЕ

В работе исследована задача прогнозирования динамики акций ПАО «Газпром» с использованием статистических моделей и методов машинного обучения.

В практической части в качестве целевой переменной использовалась доходность, что обусловлено нестационарностью ценовых рядов и требованиями эконометрики временных рядов. Прогнозируемые значения доходности рекурсивно преобразовывались в прогнозные значения цены, что позволило сохранить экономическую интерпретацию результатов и корректно оценить ошибку на разных горизонтах.

Были реализованы модели ARIMAX, SVR, Random Forest, XGBoost и LSTM. Сравнение проводилось на горизонтах 10, 20 и 30 торговых дней по метрике MAPE с использованием наивного прогноза в качестве базового ориентира.

Результаты показали, что минимальное значение ошибки на краткосрочном интервале продемонстрировала модель SVR (1.01%), однако её поведение оказалось нестабильным при увеличении горизонта прогнозирования из-за накопления рекурсивной ошибки. Наиболее устойчивые и сбалансированные показатели на всех интервалах продемонстрировали ансамблевые методы (XGBoost, MAPE 1.46–2.07%) и рекуррентная сеть LSTM (MAPE 1.41–2.00%), которые стабильно превосходили наивный базлайн. Линейная модель ARIMAX уступила нелинейным подходам на длинных горизонтах, что подтверждает необходимость учёта нестационарности рыночных данных.

Поставленная цель работы достигнута: проведён сравнительный анализ моделей прогнозирования, выявлены их сильные и слабые стороны, а также подтверждена целесообразность применения методов машинного обучения при корректной постановке задачи, учёте свойств финансовых временных рядов и использовании наивного прогноза в качестве контрольного ориентира.