

Научная статья
УДК 339.13.01
doi: 10.17586/2713-1874-2023-4-92-101

ПРОГНОЗИРОВАНИЕ КОТИРОВОК АКЦИЙ КИТАЙСКОЙ КОМПАНИИ NIO С ИСПОЛЬЗОВАНИЕМ ДВУНАПРАВЛЕННЫХ LSTM НЕЙРОННЫХ СЕТЕЙ

*Роман Сергеевич Кузнецов¹, Татьяна Гельцевна Тумарова²,
Сергей Викторович Пилипенко³*

^{1,2,3}Санкт-Петербургский государственный экономический университет, Санкт-Петербург, Россия

¹romak2921@mail.ru, <https://orcid.org/0000-0002-7589-526X>

²tumarova@finec.ru, <https://orcid.org/0009-0001-2710-0648>

³martelx27@yahoo.ca, <https://orcid.org/0009-0009-5145-0961>

Язык статьи – русский

Аннотация: Стремительное развитие технологий в сфере искусственного интеллекта и появление новых архитектур нейронных сетей оказывает влияние на компании, связанные с фондовым рынком. Различные нейронные сети могут с высокой точностью прогнозировать тренд и значения котировок акций. Способность грамотно подбирать необходимые нейронные сети и их параметры становится важным конкурентным преимуществом для компаний, функционирующих на фондовом рынке. В исследовании отражена эффективность использования двунаправленных LSTM нейронных сетей с дополнительным механизмом внимания, а также протестированы разные количества предыдущих дней, учитываемых при прогнозе будущего значения.

Ключевые слова: bilstm, котировки, нейронные сети, прогнозирование акций

Ссылка для цитирования: Кузнецов Р. С., Тумарова Т. Г., Пилипенко С. В. Прогнозирование котировок акций китайской компании NIO с использованием двунаправленных LSTM нейронных сетей // Экономика. Право. Инновации. 2023. № 4. С. 92–101. <http://dx.doi.org/10.17586/2713-1874-2023-4-92-101>.

FORECASTING STOCK QUOTES OF THE CHINESE COMPANY NIO USING BIDIRECTIONAL LSTM NEURAL NETWORKS

Roman S. Kuznetsov¹, Tatiana G. Tumarova², Sergey V. Pilipenko³

^{1,2,3}St. Petersburg State University of Economics, St. Petersburg, Russia

¹romak2921@mail.ru, <https://orcid.org/0000-0002-7589-526X>

²tumarova@finec.ru, <https://orcid.org/0009-0001-2710-0648>

³martelx27@yahoo.ca, <https://orcid.org/0009-0009-5145-0961>

Article in Russian

Abstract: The rapid development of technologies in the field of artificial intelligence and the emergence of new neural network architectures has an impact on companies associated with the stock market. Various neural networks can predict the trend and values of stock quotes with high accuracy. The ability to correctly select the necessary neural networks and their parameters becomes an important competitive advantage for companies operating in the stock market. The study reflects the effectiveness of using bidirectional LSTM neural networks with an additional attention mechanism, and also tested different numbers of previous days taken into account when predicting future values.

Keywords: bilstm, neural networks, stock forecasting, quotes

For citation: Kuznetsov R. S., Tumarova T. G., Pilipenko S. V. Forecasting Stock Quotes of the Chinese Company Nio Using Bidirectional LSTM Neural Networks (In Russ.). *Ekonomika. Pravo. Innovacii*. 2023. No. 4. pp. 92–101. (In Russ.). <http://dx.doi.org/10.17586/2713-1874-2023-4-92-101>.

Введение. Ежедневные колебания фондового рынка связаны с потоком больших объемов денежных средств, каждый день миллиарды рублей перетекают между участниками рынка. В истории развития мировой

экономики наблюдался ряд экстремальных случаев, когда колебания фондового рынка оказывали влияние на финансовые системы стран. Примером такого случая является финансовый кризис 2008 года. Поиск

эффективных методов прогнозирования котировок акций на фондовой бирже является актуальной задачей, которая в последние годы получила дополнительный стимул к решению за счет развития и повышения доступности технологий искусственного интеллекта.

Развитием способов улучшения точности прогнозирования занимаются компании из разных секторов экономики: IT-компании, специализирующиеся на искусственном интеллекте, фондовые биржи, консалтинговые и аналитические агентства, частные игроки на рынке. Все больше компаний финансового сектора переходят к механизмам принятия решений, основанным не на человеческой оценке, а на алгоритмах машинного обучения.

Искусственный интеллект – это моделирование процессов человеческого интеллекта электронно-вычислительными устройствами. ИИ применяется в качестве экспертных систем, средств обработки естественного языка, компьютерного зрения и прогнозирования финансовых и иных показателей. Реальная ценность ИИ обусловлена его способностью облегчить и ускорить процессы анализа данных и принятия управленческих решений в различных отраслях. В рамках ИИ существует его отдельная отрасль – машинное обучение, которое направлено на использование информации и алгоритмов для имитации процесса обучения человека электронно-вычислительными машинами. Финансовый сектор, включая институциональных инвесторов, играет существенную роль в развитии искусственного интеллекта, так как инвестирует в развитие данной технологии с целью формирования эффективных методов анализа фондового и других рынков.

Стремительное развитие интеллектуальных систем приводит к росту количества и форматов биржевых и внебиржевых данных, а также росту спроса на автоматические системы сбора, обработки и прогностической аналитики фондового рынка. Различные системы прогностической аналитики фондового рынка создаются как научными сотрудниками и обучающимися высших учебных заведений, так и сотрудниками высокотехнологичных компаний. В обоих случаях основная цель исследователей – это разработка

наиболее эффективного средства прогнозирования котировок фондового рынка с использованием оптимальных алгоритмов машинного обучения.

В рамках исследования авторами используется двунаправленная LSTM-нейронная сеть с дополнительным механизмом внимания. LSTM (Long Short-Term Memory) архитектура используется в случаях, когда необходимо что-то запоминать при обучении и выявлять долгосрочные зависимости. Двунаправленная LSTM (BiLSTM) представляет собой две сети LSTM с противоположными направлениями, что позволяет обучать модель не только от ввода к выводу, но и в обратном направлении. Механизм внимания позволяет обучить нейронную сеть «концентрироваться» на том участке данных, который обрабатывается и анализируется в конкретный момент времени.

Обзор литературы. Повышение доступности вычислительных мощностей позволяет современным ученым экспериментировать с различными комбинациями слоев и алгоритмов машинного обучения в процессе разработки механизмов прогнозирования фондового рынка. В работе Цзилинь Чжана и Лиши Йе используется CNN-BiLSTM нейронная сеть с механизмом внимания для прогнозирования индекса CSI 300. В качестве результата авторы отмечают, что двунаправленные LSTM сети с механизмом внимания показывают наилучший результат в сравнении с LSTM и CNN-LSTM сетями [1]. Для прогнозирования индекса CSI 300, а также индекса Шанхайской фондовой биржи (SSE Composite) BiLSTM нейронную сеть также используют Ю Чен, Руйсинь Фанг и коллеги в своей научно-исследовательской работе, в которой авторы отмечают CNN-BiLSTM-ECA сеть как наиболее эффективную [2]. Для обучения нейронных сетей используются не только значения котировок на фондовом рынке и показатели технических индикаторов, но и текстовая информация, содержащая комментарии из социальных сетей или новостные статьи. Томас Хатта в своем исследовании обучает двунаправленную LSTM нейронную сеть не только на данных об изменениях котировок, но и учитывает комментарии трейдеров в социальных сетях [3]. Такой подход

позволяет прогнозировать котировки с учетом настроений рыночных участников.

В изученной литературе различаются не только используемые комбинации сети используются для прогнозирования изменения котировок акций, торгуемых на индийском фондовом рынке [4]. Авторы отмечают эффективность использования двунаправленных нейронных сетей при прогнозировании котировок. Но также отмечают, что необходим дополнительный анализ отрасли, в которой функционирует анализируемая компания. Чон Гван, Миген Чо и коллеги в рамках своей научно-исследовательской работы используют Long Short-Term Memory нейронные сети для прогнозирования акций корейских компаний KOSDAQ и KOSPI [5].

Для повышения эффективности торговли прогнозирование котировок акций с использованием нейронных сетей комбинируется с различными методами анализа и оценки рисков при инвестировании на финансовых рынках. В исследовании Ронила Баруа и Анила Шармы обученные двунаправленные LSTM нейронные сети используются совместно с методом Блэка-Литтермана, разработанным Фишером Блэком и Робертом Литтерманом в 1990 году в компании Goldman Sachs. Авторы отмечают, что собранные портфели значительно превосходят по доходности эталонные модели, даже с учетом транзакционных издержек [6].

Основная часть. Двунаправленная LSTM сеть с механизмом внимания. Сеть LSTM была разработана в 1997 году Хохрейтером и Шмидхубером как разновидность рекуррентной нейронной сети, пригодной для обработки и прогнозирования событий с большими интервалами и задержками во

параметров и архитектур нейронных сетей, но и региональные рынки, по акциям которых осуществляется прогноз. В работе Гириджара Маджи и коллег двунаправленные нейронные временных рядах. Нейронная сеть с двунаправленной краткосрочной памятью (BiLSTM) является оптимизацией обычной LSTM, которая способна учитывать прошлую и будущую информацию за счет объединения обычного и обратного слоев Long Short-Term Memory.

Механизм внимания позволяет значительно повысить производительность нейронных сетей. Он назначает разные веса в соответствии с различными признаками, то есть назначать большие веса критической информации для повышения эффективности обработки информации за счет дифференцированного присвоения веса и решать проблему потери информации, вызванную длинными последовательностями LSTM.

Большинство существующих методов направлены на разработку более сложных модулей внимания для достижения лучшей производительности, что неизбежно увеличивает сложность и вычислительную нагрузку на модель. В исследовании авторы используют модуль с эффективным канальным вниманием (ECA – Effective Channel Attention). ECA может не только генерировать веса для каждого канала, но и анализировать корреляцию между различными каналами. Для данных временного ряда большие веса будут назначены для ключевых признаков, а меньшие – для нерелевантных. Таким образом, ECA фокусируется на полезной информации, что повышает чувствительность сети к основным признакам. Уравнения слоя внимания представлены ниже [7].

$$e_t = \tanh(w_a[x_1, \dots, x_T] + b) \quad (1)$$

$$a_T = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)} \quad (2)$$

где e_t – результат вычисления весов для разных признаков;

\tanh – функция активации;

w_a – матрица весов механизма внимания;

$[x_1, \dots, x_T]$ – входные данные слоя внимания;

b – отклонение внимания.

a_t – конечные веса;

e_t – результат вычисления весов для разных признаков.

В рамках исследования на программном уровне нейронная сеть и все процессы, связанные с анализом и прогнозированием, реализованы на языке программирования python

версии 3.9. Для построения слоев нейронной сети используются библиотеки tensorflow (версия 2.11.0) и keras (версия 2.11.0), разработанные специально для работы с алгоритмами искусственного интеллекта. В качестве

основного обработчика данных выступает библиотека pandas (версия 1.5.2), а сбор данных осуществляется с помощью библиотеки requests (версия 2.31.0). Процесс прогнозирования отражен на рисунке 1.

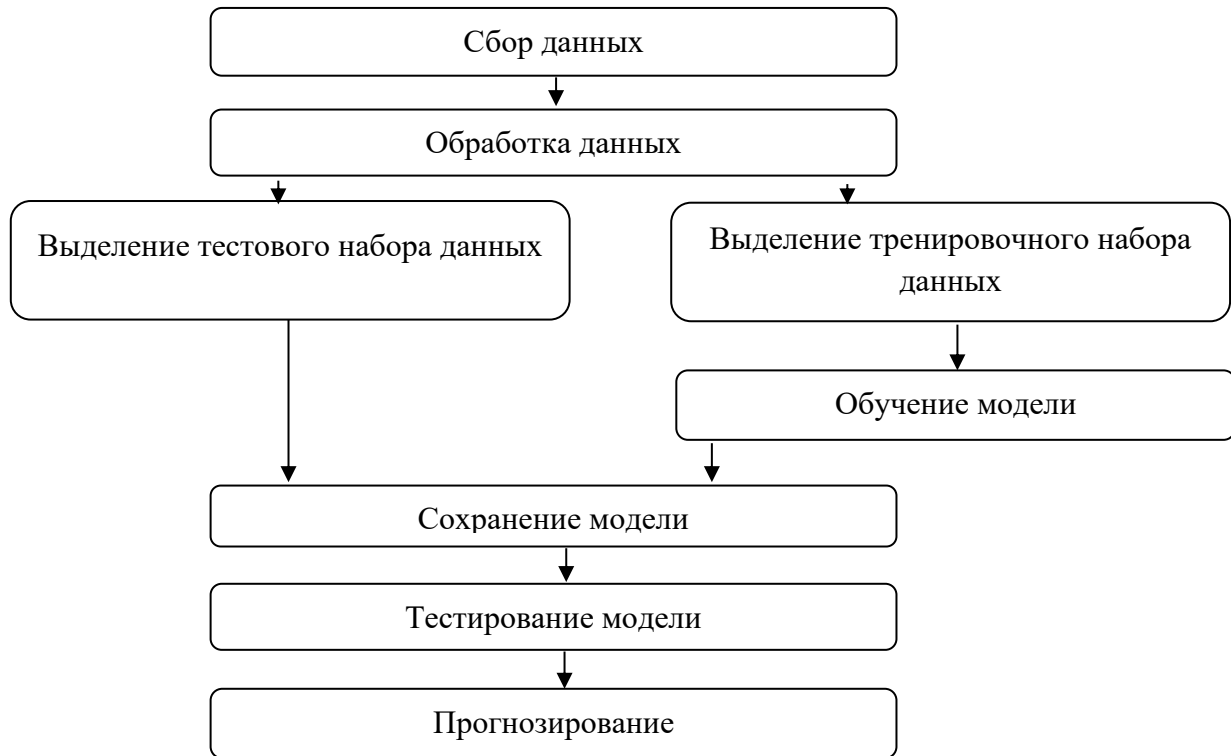


Рисунок 1 – Схема процесса прогнозирования котировок

Источник: составлен авторами

В качестве источника данных выступает веб-сайт Yahoo Finance – международный провайдер финансовой информации, собирающий данные об изменениях котировок по мировым акциям, новости в разрезе компаний и различных финансовых секторов, показатели финансовых отчетностей компаний и иные метрики. В качестве тренировочного и тестового набора данных используются котировки акций компании NIO Inc. Компания основана в 2014 году в Шанхае и специализируется на производстве электромобилей. Компания NIO отличается постоянными технологическими прорывами и инновациями, такими как передовые в отрасли технологии замены аккумуляторов, Battery as Service (BaaS), а также запатентованные технологии автономного вождения [8]. Данные, используемые для обучения и тестирования модели, содержат название компании, дневные значе-

ния котировок открытия, закрытия, дневные максимум и минимум (см. Таблицу 1).

Собранный средствами python датасет содержит данные 1273 торговых дней за период 12.09.2018 – 02.10.2023 и включает в себя следующие столбцы:

- дата торгов на бирже, к которой относятся соответствующие значения котировок;
- тикер – название компании, по которой собраны данные;
- цена открытия – значение котировки акции на выбранную дату в момент открытия биржи;
- дневной максимум – максимальное значение котировки за один торговый день;
- дневной минимум – минимальное значение котировки за один торговый день;
- цена закрытия – значение котировки акции на выбранную дату в момент закрытия биржи.

**Пример набора данных по акциям NIO Inc.,
используемых для обучения и тестирования модели**

Источник: составлен авторами на основе [9]

Дата	Тикер	Цена открытия	Дневной максимум	Дневной минимум	Цена закрытия
12.09.2018	Nio Inc	6	6,93	5,35	6,6
13.09.2018	Nio Inc	6,62	12,69	6,52	11,6
14.09.2018	Nio Inc	12,66	13,8	9,22	9,9
17.09.2018	Nio Inc	9,61	9,75	8,5	8,5
18.09.2018	Nio Inc	8,73	9,1	7,67	7,68
19.09.2018	Nio Inc	7,88	9,15	7,54	8,5
20.09.2018	Nio Inc	9,07	9,47	8,42	8,78
21.09.2018	Nio Inc	9,05	9,07	8,5	8,59
24.09.2018	Nio Inc	8,32	8,39	7,85	7,87
25.09.2018	Nio Inc	8,06	8,19	7,5	7,58
26.09.2018	Nio Inc	7,59	7,76	7,5	7,5
27.09.2018	Nio Inc	7,55	7,55	7,01	7,01

Для обучения и тестирования использованы данные об изменениях котировок в момент закрытия биржи, содержащиеся в столбце «Цена закрытия». Тренировочный датасет содержит девяносто пять процентов значений котировок или 1210 торговых дней, а тестовый, соответственно, пять процентов или 63 торговых дня.

Модели обучены прогнозировать значение одной котировки на день вперед. Для эксперимента взяты три разных периода, за счет которых происходит прогноз. Первая модель прогнозирует будущее значение за счет анализа тридцати предыдущих котировок, вторая модель – за счет анализа десяти предыдущих котировок, а третья модель – за счет трех. В качестве входных параметров модели используются следующие:

– количество прогнозных значений, равное одному;

– количество используемых для прогноза значений, равное тридцати, десяти и трем для трех разных моделей;

– количество эпох, равное ста сорока.

В качестве основных слоев нейронной сети используются два слоя двунаправленной LSTM нейронной сети и слой внимания. В качестве дополнительных слоев используются dropout слой, позволяющий предотвратить переобучение модели, а также dense слой, используемый для классификации (см. рисунок 2).

Результаты прогноза трех обученных моделей представлены на графиках ниже (рисунки 3–5).

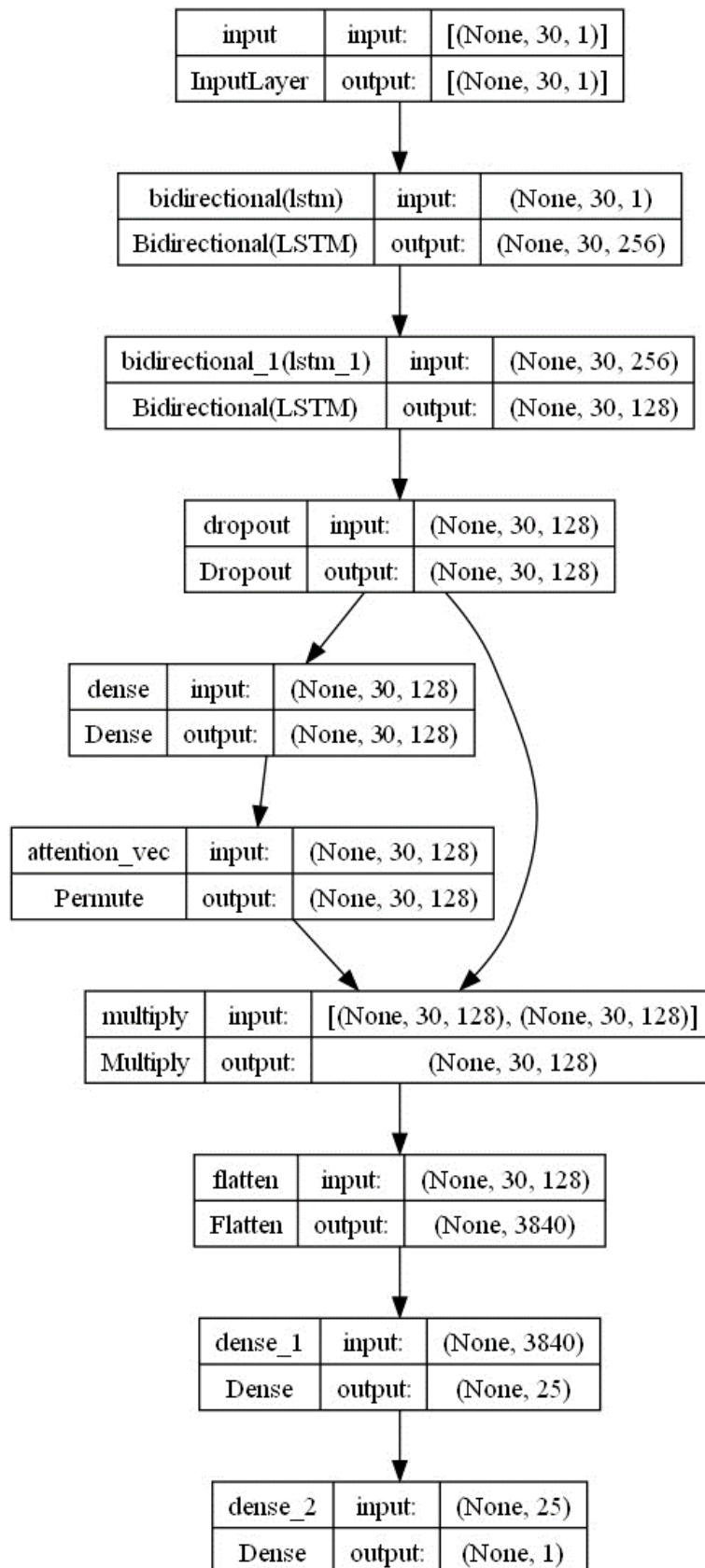


Рисунок 2 – Описание слоев нейронной сети, используемой в исследовании

Источник: составлено авторами



Рисунок 3 – Результаты прогнозирования модели, учитывающей 30 предыдущих значений
Источник: составлено авторами

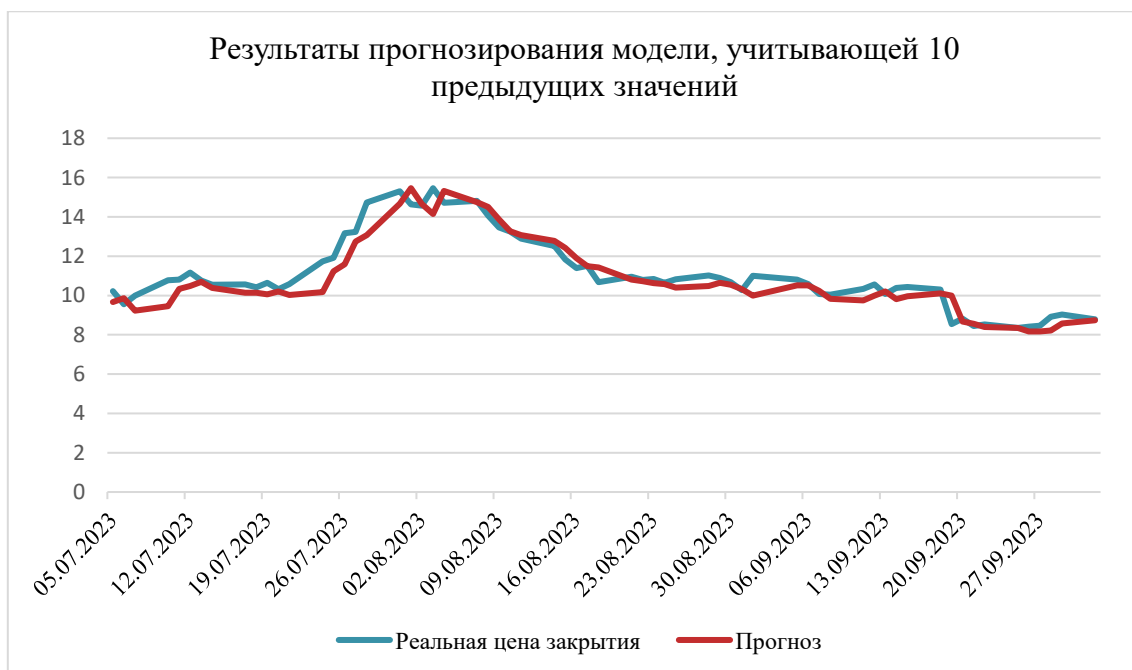


Рисунок 4 – Результаты прогнозирования модели, учитывающей 10 предыдущих значений
Источник: составлено авторами

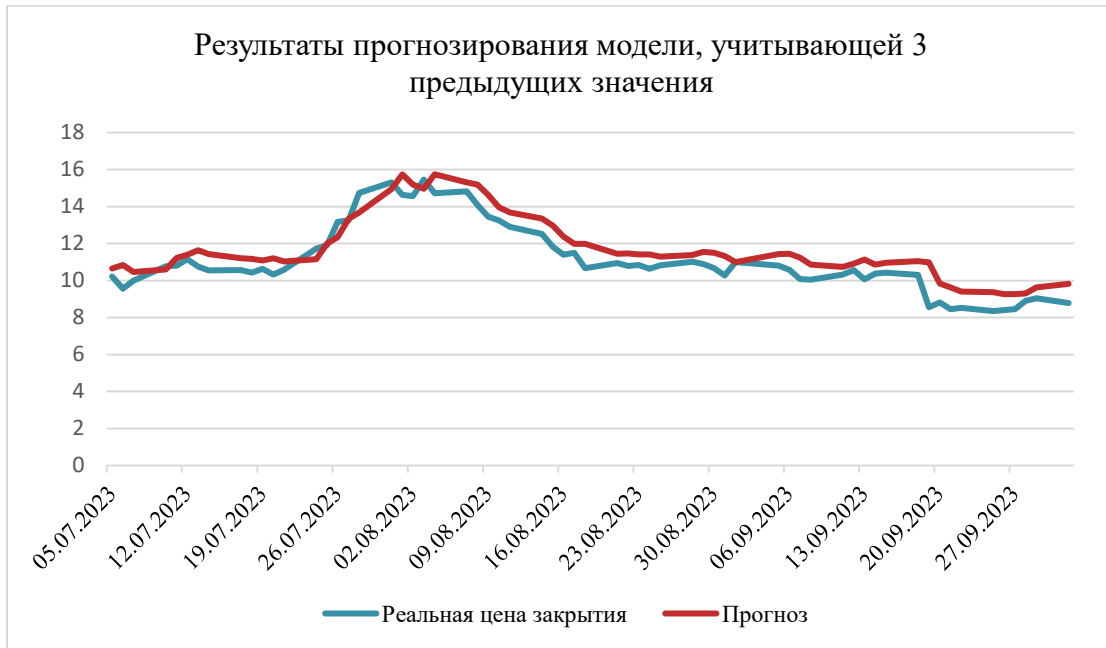


Рисунок 5 – Результаты прогнозирования модели, учитывающей 3 предыдущих значения
 Источник: составлено авторами

Необходимо отметить, что уже при визуальном анализе рисунков выше видно, что прогнозная модель, учитывающая значения котировок десяти предыдущих дней, дает наиболее точные результаты. Для более точного анализа результатов прогнозирования необходимо проанализировать показатели точности прогноза моделей. В качестве показателей авторами используются следующие:

1) Root Mean Square Error (RMSE) – среднеквадратичное отклонение, представляющее собой отклонение остатков (ошибок прогнозирования). Остатки – это показатель того, насколько удалены точки данных от линии регрессии. RMSE показывает концентрацию данных вокруг линии регрессии. Формула расчета отражена ниже (см. уравне-

ние 3), где f – прогнозные значения; o – фактические значения.

2) Mean Square Error (MSE) – среднеквадратичная ошибка, измеряющая величину ошибки в статистических моделях. MSE оценивает среднюю квадратичную разницу между наблюдаемыми и прогнозируемыми значениями. Если модель не содержит ошибок, то значение MSE равно нулю. Формула расчета отражена ниже (см. уравнение 4), где y_i – i -е фактическое значение; \hat{y}_i – i -е прогнозное значение; n – количество наблюдений.

3) Mean Absolute Error (MAE) – средняя абсолютная ошибка, рассчитываемая по формуле, отраженной ниже (см. уравнение 5), где n – количество ошибок; $|x_i - x|$ – абсолютные ошибки.

$$MSE = \sqrt{(f - o)^2} \tag{3}$$

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \tag{4}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x| \tag{5}$$

4) Mean Gamma Deviance (MGD) – среднее гамма-отклонение, эквивалентное отклонению Твиди с параметром силы, равным двум. Показатель рассчитывается с помощью библиотеки языка python – sklearn.

5) Mean Poisson Deviance (MPD) – среднее отклонение Пуассона, эквивалентное отклонению Твиди с параметром силы, равным одному. Показатель рассчитывается с помощью библиотеки языка python – sklearn.

Перечисленные выше показатели широко применяются в различных исследованиях, направленных не только на прогнозирование фондового рынка, но и на прогнозирование других показателей, таких как уровень заболеваемости, распространение экологических загрязнений и многое другое. В рамках исследования авторами были рассчитаны данные показатели для каждой модели (см. таблицу 2).

Таблица 2

Показатели точности обученных моделей

Источник: составлено авторами

Показатель	Модель 1, учитывающая 30 предыдущих значений	Модель 2, учитывающая 10 предыдущих значений	Модель 3, учитывающая 3 предыдущих значения
RMSE	0,722	0,611	0,826
MSE	0,521	0,374	0,682
MAE	0,588	0,454	0,734
MGD	0,004	0,003	0,005
MPD	0,046	0,033	0,061

Анализируя данные, отметим, что наибольшая точность прогнозирования достигается второй моделью, учитывающей десять предыдущих котировок. На втором месте по точности находится модель номер один, учитывающая тридцать предыдущих котировок. Необходимо отметить, что все три модели в целом показали хорошие результаты и могут быть использованы в аналитической деятельности компаний. Модели не всегда могут предсказать точное будущее значение котировки, но могут в целом указать на тренд. Модели могут быть использованы как аналитическими отделами компаний в качестве инструмента прогнозирования и анализа рынка, так и трейдинговыми отделами в качестве вспомогательного инструмента для принятия торговых решений.

Заключение. Под воздействием развития технологий передачи и обработки информации, глобализации мирового новостного пространства и высокочастотной торговли увеличивается скорость изменения котировок на фондовом рынке. В данных условиях растет необходимость и важность разработки эффективных методов прогнозирования котировок фондового рынка. Потребителями

информации о прогнозных значениях выступают разные компании, вовлеченные в широкий спектр видов финансовой деятельности: аналитика фондового рынка, составление инвестиционных портфолио, оценка компаний, представленных на фондовом рынке, разработка торговых роботов и многое другое.

Нейронные сети, основанные на архитектуре Long Short-Term Memory, хорошо себя зарекомендовали в качестве инструмента прогнозирования временных рядов. Добавление обратного LSTM-слоя в модель, то есть трансформация LSTM модели в BiLSTM модель позволяет делать более точные прогнозы за счет учета как входной, так и выходной информации при обучении. Использование слоя внимания позволяет обучать модель концентрироваться на более приоритетных показателях и тем самым еще больше увеличивать точность модели. Необходимо отметить, что при разработке моделей необходимо экспериментировать не только с количеством учитываемых предыдущих показателей, но и с количеством слоев, эпох и других параметров нейронной сети. В качестве показателей точности прогноза могут быть использованы

стандартные метрики, такие как RMSE, MSE, MAE, MGD и MPD.

Отобранная в рамках исследования модель может быть также использована в роли бинарного торгового советника. Торговая стратегия, основанная на данной модели, может базироваться на двух сигналах: если прогнозируемое значение выше фактического, то

трейдер открывает новую или удерживает текущую длинную позицию. Если прогнозное значение ниже текущего, то трейдер открывает или удерживает уже существующую короткую позицию. Сигналом к продаже выступает ситуация, когда прогнозное значение ниже фактического, либо наоборот в случае с открытой короткой позицией.

Список источников

1. Jilin Z., Lishi Y., Yongzeng L. Stock Price Prediction Using CNN-BiLSTM-Attention Model // *Mathematics*. 2023. № 11 (9). С. 1–18. (In Eng.).
2. Yu C., Ruixin F., Ting L. Stock Price Forecast Based on CNN-BiLSTM-ECA model // *Hindawi. Scientific programming*. 2021. № 5. С. 1–20. (In Eng.). DOI: 10.1155/2021/2446543.
3. Hatta D., Abida R. Stock Prediction Based on Twitter Sentiment Extraction Using BiLSTM-Attention // *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*. 2022. №. 1. С. 187–198. (In Eng.).
4. Maji G., Sen S. Stock Price Prediction Using LSTM on Indian Share Market // *EPiC series in Computing*. 2019. №. 63. С. 101–110. (In Eng.).
5. Chatterjee I., Gwan J. An NLP and LSTM Based Stock Prediction and Recommender System for KOSDAQ and KOSPI // *Intelligent Human Computer Interaction*. 2021. С. 403–413. (In Eng.). DOI: 10.1007/978-3-030-68449-5_40.
6. Barua R., Sharma A. Dynamic Black Litterman Portfolios with Views Derived via CNN-BiLSTM Predictions // *Finance Research Letters*. 2022. № 49. С. 20–34. (In Eng.). DOI: 10.1016/j.frl.2022.103111.
7. Ming-Che L., Jia-Wei C. Applying Attention-Based BiLSTM and Technical Indicators in the Design and Performance Analysis of Stock Trading Strategies // *Neural Computing and Applications*. 2022. № 34. С. 13267–13279. (In Eng.). DOI:10.1007/s00521-021-06828-4.
8. NIO // Официальный сайт NIO Inc. [Электронный ресурс]. – Режим доступа: <https://ir.nio.com> (In Eng.).

References

1. Jilin Z., Lishi Y., Yongzeng L. Stock Price Prediction Using CNN-BiLSTM-Attention Model. *Mathematics*. 2023. No. 11 (9). pp. 1–18.
2. Yu C., Ruixin F., Ting L. Stock Price Forecast Based on CNN-BiLSTM-ECA model. *Hindawi. Scientific programming*. 2021. No. 5. pp. 1–20. DOI: 10.1155/2021/2446543.
3. Hatta D., Abida R. Stock Prediction Based on Twitter Sentiment Extraction Using BiLSTM-Attention. *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*. 2022. No. 1. pp. 187–198.
4. Maji G., Sen S. Stock Price Prediction Using LSTM on Indian Share Market. *EPiC Series in Computing*. 2019. No. 63. pp. 101–110.
5. Chatterjee I., Gwan J. An NLP and LSTM Based Stock Prediction and Recommender System for KOSDAQ and KOSPI. *Intelligent Human Computer Interaction*. 2021. pp. 403–413. DOI: 10.1007/978-3-030-68449-5_40.
6. Barua R., Sharma A. Dynamic Black Litterman Portfolios with Views Derived via CNN-BiLSTM Predictions. *Finance Research Letters*. 2022. No. 49. pp. 20–34. DOI: 10.1016/j.frl.2022.103111.
7. Ming-Che L., Jia-Wei C. Applying Attention-Based BiLSTM and Technical Indicators in the Design and Performance Analysis of Stock Trading Strategies. *Neural Computing and Applications*. 2022. No. 34. pp. 13267–13279. DOI:10.1007/s00521-021-06828-4.
8. Official website NIO Inc. Available at: <https://ir.nio.com>