

УДК 338.11

Кан Цзинхань

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия

Ю. Ю. Кочинев

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия, e-mail: kochinev@kafedrapik.ru

ПРОГНОЗИРОВАНИЕ АКЦИЙ НА ОСНОВЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ В СОЧЕТАНИИ С НЕЙРОННОЙ СЕТЬЮ С ДОЛГОВРЕМЕННОЙ И КРАТКОВРЕМЕННОЙ ПАМЯТЬЮ

Ключевые слова: прогнозы по акциям, свёрточные нейронные сети, нейронные сети долговременной и кратковременной памяти, машинное обучение, статистические методы.

В данной статье анализируется прогностическая эффективность сверточных нейронных сетей и нейронных сетей с долговременной кратковременной памятью на данных акций Amazon, а также определяются перспективы внедрения методов глубокого обучения в практику прогнозирования акций. Анализируя различные измерения набора данных Amazon и используя архитектуру PyTorch для моделирования сверточных нейронных сетей и нейронных сетей с долговременной кратковременной памятью, выявляются проблемы улучшения их структур. В этом исследовании особое внимание уделяется внутренним структурам сверточных нейронных сетей и нейронных сетей с долговременной кратковременной памятью, а также их преимуществам и недостаткам в прогнозировании акций. Свёрточные нейронные сети (CNN) хорошо справляются с извлечением локальных характеристик, таких как краткосрочные тенденции и характеристики колебаний, из данных о акциях. Однако по-прежнему существуют такие проблемы, как невозможность сосредоточиться на долгосрочных тенденциях и нечувствительность к данным секвенирования. Сети с долговременной краткосрочной памятью (LSTM) хорошо работают с данными временных рядов и могут эффективно фиксировать долгосрочные зависимости в данных об инвентаризации. Однако он нечувствителен к локальным особенностям, и при наличии большого количества шума он будет переобучать обучающие данные, что приведет к недостаточной способности к обобщению. Поэтому в данной статье разрабатывается модель CNN-LSTM, сочетающая преимущества обеих моделей. CNN отвечает за извлечение локальных признаков, в то время как LSTM отвечает за регистрацию долгосрочных тенденций и циклических закономерностей, обеспечивая более полное извлечение признаков и моделирование временных рядов данных. Результаты показывают, что модель CNN-LSTM превосходит модели CNN или LSTM в прогнозировании цен акций и может лучше справляться с нелинейностью и сложными колебаниями рынка.

Kang Jinghan

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

Yu. Yu. Kochinev

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia,
e-mail: kochinev@kafedrapik.ru

STOCK PREDICTION BASED ON CONVOLUTIONAL NEURAL NETWORK COMBINED WITH LONG-TERM AND SHORT-TERM MEMORY NEURAL NETWORK

Keywords: Stock forecasts, convolutional neural networks, long-term and short-term memory neural networks, machine learning, statistical methods.

This paper analyzes the predictive performance of convolutional neural networks and long short-term memory (LSTM) neural networks on Amazon stock data and identifies prospects for implementing deep learning methods in stock forecasting. By analyzing various dimensions of the Amazon dataset and using the PyTorch architecture to model convolutional neural networks and LSTM neural networks, challenges in improving their structures are identified. This study specifically focuses on the internal structures of convolutional neural networks and LSTM neural networks, as well as their advantages and disadvantages in stock forecasting. Convolutional neural networks (CNN) perform well in extracting local features, such as short-term trends and fluctuation characteristics, from stock data. However, there are still challenges such as an inability to focus on long-term trends and insensitivity to sequencing data. Long short-term memory

(LSTM) networks perform well on time series data and can effectively capture long-term dependencies in inventory data. However, it is insensitive to local features, and in the presence of a large amount of noise, it will overfit the training data, resulting in insufficient generalization. Therefore, this paper develops a CNN-LSTM model that combines the advantages of both models. The CNN is responsible for extracting local features, while the LSTM is responsible for capturing long-term trends and cyclical patterns, enabling more comprehensive feature extraction and time-series data modeling. The results show that the CNN-LSTM model outperforms either the CNN or LSTM models in stock price forecasting and is better able to handle nonlinearities and complex market fluctuations.

Введение

Фондовый рынок – один из важнейших сегментов финансового рынка и основной канал корпоративного финансирования и инвестиций [1]. Инвесторы участвуют в управлении компанией, приобретая акции, надеясь получить доход от торговли акциями. Однако на цены фондового рынка влияет множество факторов, включая макроэкономические показатели, государственную политику, тенденции развития отрасли и глобальные события. Сложность и изменчивость этих факторов делают фондовый рынок полон неопределенности и риска. Для лучшего понимания рыночных тенденций и колебаний цен точное прогнозирование цен акций стало важным направлением финансовых исследований.

Прогнозы цен на акции очень важны для инвесторов, бизнеса и развития национальной экономики. Для инвесторов прогнозы цен на акции могут снизить риски, увеличить доходность и помочь разработать более эффективные инвестиционные стратегии [2]. Для предприятий прогнозы цен на акции помогают в принятии стратегических решений и оптимизации бизнеса, повышают интерес инвесторов и повышают привлекательность рынка капитала [3]. Для страны прогнозы цен на акции служат ориентиром для поддержания стабильности финансового рынка, повышения доверия инвесторов и развития корпораций, а также поддержки промышленного планирования и разработки стратегии [4]. Точные прогнозы помогают принимать экономические решения и осуществлять финансовое регулирование, стимулировать экономический рост и занятость, а также содействовать развитию и инновациям в конкретных отраслях.

К обычно используемым моделям в традиционных статистических методах относятся модель ARMA, модель ARIMA, модель GARCH и модель Маркова. В 2015 году Чен [5] предположил, что точность прогнозирования нечетких временных рядов зависит от длины интервала, и предложил модель нечетких временных рядов, основан-

ную на расчете гранулярности блоков и дискретизации энтропии для прогнозирования цен на фондовом рынке, а также проверил результаты прогнозирования с помощью парного двухстороннего t-теста. В 2018 году Эфенди [6] разработал усовершенствованную нечеткую случайную авторегрессионную модель для компенсации волатильности фондовых данных. Результаты показывают, что изменчивость и корректировка спреда являются важными факторами повышения точности модели. В том же году другой ученый Ахмед [7] предложил модель APARCH с переключением Маркова при обобщенном t и обобщенном гиперболическом распределении для полного охвата нечеткой динамики и стилизованных характеристик доходности финансового рынка и генерации прогнозов значений риска. В 2019 году Гуань [8] предложил модель прогнозирования, сочетающую теорию цепей Маркова и теорию нечетких множеств. Эта модель позволяет гибко отражать исторические и текущие взаимосвязи, а также устранять шум благодаря теории нечетких множеств. Она была применена для прогнозирования нескольких фондовых индексов, и результаты превзошли результаты других традиционных моделей. В 2020 году Луо [9] предложил комбинированную линейную и нелинейную модель, объединяющую эмпирическую модовую декомпозицию (EMD), ARIMA и разложение Тейлора для получения более точных результатов прогнозирования [9].

Хотя статистические методы теоретически совершенны, их преимущества заключаются главным образом в обработке линейных данных и требуют строгих базовых предположений. Однако данные о ценах на акции имеют случайные и нелинейные характеристики, что делает традиционные статистические методы неспособными к достижению идеальных результатов прогнозирования. Поэтому необходимо найти новые методы, более подходящие для обработки случайных и нелинейных данных, чтобы повысить точность прогнозирования цен акций.

Будучи особой структурой нейронной сети, серия RNN обладает уникальной конструкцией скрытых состояний, которая позволяет передавать информацию последовательно, эффективно фиксируя последовательные зависимости и динамические закономерности в ценах акций. Эта особенность даёт серии RNN огромное преимущество в области прогнозирования цен акций [10]. LSTM эффективно решает проблему ослабления или усиления градиента во время обратного распространения RNN путем введения сложного механизма стробирования, включая входные вентили, вентили забывания и выходные вентили [11].

Из-за нелинейности и нестабильности данных по акциям одна модель прогнозирования больше не может точно предсказывать цены акций, поэтому в данной статье рассматривается применение комбинированной модели для прогнозирования акций.

Целью данного исследования – использовать сверточную нейронную сеть (CNN) и нейронную сеть с долговременной краткосрочной памятью (LSTM) для прогнозирования тенденций акций на основе набора данных Amazon. В данной статье представлены целевые усовершенствования в двух ключевых областях. Во-первых, на архитектурном уровне локально-глобальная связь реализуется в виде обучаемого механизма с двумя путями. В ветви CNN используются слои свертки с расширенными свертками и каналным вниманием, что явно усиливает чувствительность к внутридневным микроструктурам с высокой волатильностью. Ветвь LSTM включает обучаемый шлюз развязки тренда перед блоком памяти. Сначала он разлагает исходную последовательность на быстрые и медленные латентные состояния, взвешенные по параметрам полураспада, а затем объединяет их с выходом CNN посредством перекрестного внимания. Это предотвращает доминирование градиентов долгосрочными трендами. Во-вторых, на этапе обработки данных мы используем подход перекрывающейся выборки, взвешенной по волатильности, используя обратную величину внутридневной амплитуды в качестве веса выборки. В рамках скользящего окна плотность выборки динамически корректируется, чтобы модель могла сосредоточиться на интервалах с низким соотношением сигнал/шум и высоким информационным содержанием во время

обучения. Это значительно снижает риск переобучения в LSTM.

Материал и методы исследования

Amazon – одна из крупнейших и самых влиятельных технологических компаний в мире. Собранные в этой статье данные содержат историческую информацию о ценах акций Amazon за 28 лет, с сентября 1997 года по сентябрь 2025 года. Данные получены непосредственно от Yahoo Finance, надежного поставщика данных о финансовых рынках [12]. Этот набор данных относится к более длительному временному ряду, который позволяет нам анализировать рыночные показатели Amazon на протяжении многих лет, наблюдать долгосрочные тенденции и определять ключевые события в истории компании.

В данной статье данные обрабатываются со следующих четырех различных точек зрения:

Обработка пропущенных значений относится к методам работы с пропущенными значениями при анализе данных [13]. В данной статье рассматривается вопрос удаления пропущенных значений из строк/столбцов и заполнения их статистическими данными, такими как среднее значение, медиана, мода и т. д., или оценки пропущенных значений на основе значений из других выборок.

Дедупликация включает в себя выявление и удаление дублирующихся или избыточных записей в наборе данных [14]. Это важный этап очистки и предварительной обработки данных, гарантирующий, что анализ будет проводиться на основе уникальных и точных данных. Дубликаты могут возникать из-за человеческих ошибок, ошибок ввода данных или несоответствий в источнике данных.

Работа с выбросами подразумевает выявление и обработку экстремальных значений в наборе данных, которые существенно отличаются от остальных данных [15]. Выбросы могут существенно повлиять на результаты анализа данных и, при отсутствии должной обработки, исказить результаты. В данной статье используется метод удаления выбросов.

Нормализация данных подразумевает стандартизацию данных в наборы данных с одинаковыми измерениями и относительными размерами [16]. Это помогает предотвратить чрезмерное влияние отдельных характеристик данных на модель.

Таблица 1

Данные по акциям Amazon за 1997–2025 гг.

Показатель	Описания
Дата	Дата записи (ГГГГ-ММ-ДД).
Открыть	Цена открытия акции.
Высокий	Самая высокая цена акции за день.
Низкий	Самая низкая цена акции за день.
Закрыть	Цена закрытия акции.
Отрегулируйте закрытие	Скорректированная цена закрытия.
Объём	Количество торгуемых акций.

Таблица 2

Описательный статистический анализ данных по акциям Amazon

	Открыть	Высокий	Низкий	Закрыть	Объём
Минимум	0.070313	0.072396	0.065625	0.069792	9744000.0
Среднее	62.355853	63.052616	61.594109	62.341676	136790145.935971
Стандарт	44,131184	44.647214	43,572817	44.125563	134223705.314077
25%	2.127875	2.173625	2.098375	2.135375	61778500.0
50%	9.34	9.4795	9.218	9.33725	98998000.0
75%	82,033247	83.25325	80.867874	81.993626	154017000.0
Макс	239.020004	242.520004	238.029999	242.059998	2086584000.0

Благодаря этой серии систематических операций по очистке данная статья может преобразовать исходные, грубые данные в чистый, последовательный и пригодный для аналитического использования набор данных.

Цель обработки данных – повысить ценность информации и упростить принятие решений. Важную роль в этом играют автоматизированные решения для обработки данных, использующие технологии искусственного интеллекта и машинного обучения.

Информация в исходном наборе данных приведена в таблице 1.

Набор данных по акциям Amazon содержит в общей сложности 7133 фрагмента последовательностей данных, каждый из которых содержит 7 элементов информации, перечисленных в таблице 1.

При прогнозировании акций описательный статистический анализ является не только базовым этапом предварительной обработки данных, но и ключевым условием для понимания поведения рынка и построения надежных моделей прогнозирования.

Проводя количественный анализ центральной тенденции, дисперсии и характера распределения ключевых показателей, таких как исторические цены акций, объемы

торгов и доходности, исследователи могут уловить внутренние закономерности и аномальные характеристики данных.

В частности, характеристики колебаний объёма торгов могут отражать стабильность потока капитала. Этот анализ не только служит основой для очистки данных и разработки признаков для последующего моделирования временных рядов, но и позволяет избежать переобучения или неверной оценки модели из-за игнорирования существенных характеристик данных, выявляя исторические циклы колебаний и границы риска.

Как показано в таблице 2, среднее значение цены акций Amazon значительно превышает медиану, а стандартное отклонение составляет почти 70% от среднего значения, что наглядно демонстрирует крайне смещенное вправо, высоковолатильное и длиннохвостое распределение доходности. Это указывает на то, что акции фактически торгуются в низком диапазоне в большинство торговых дней, а отдельные скачки значительно повышают среднее значение. Большой разрыв между 25% и 75% процентиллями дополнительно подтверждает ненормальный характер ценового ряда. Он также показывает, что стратегия простой скользя-

щей средней склонна генерировать ложные сигналы на низких уровнях цен.

Описательная статистика не только количественно определяет характеристики данных, характеризующих высокий рост, высокий уровень шума и ловушки лишения прав, но и закладывает прочную основу для последующей разработки прогностических моделей.

В данной статье набор данных делится на обучающий набор, проверочный набор и тестовый набор. Обучающий набор используется для обучения модели. Модель изучает правила сопоставления входных и выходных данных, корректируя свои внутренние параметры, веса и смещения в нейронной сети, изучая образцы и метки в обучающем наборе. Валидационный набор используется для оценки модели во время обучения и настройки её гиперпараметров. Гиперпараметры – это параметры, заданные до начала обучения, такие как скорость обучения, количество слоёв сети и сила регуляризации. Валидационный набор позволяет оценить эффективность модели при различных настройках гиперпараметров без необходимости проведения финального тестирования, тем самым выбирая оптимальный набор гиперпараметров. Тестовый набор используется для окончательной однократной оценки обобщающей способности модели. После завершения обучения модели

и настройки параметров мы оцениваем её на тестовом наборе. Эта оценка считается объективной оценкой эффективности модели на ранее не исследованных данных реального мира.

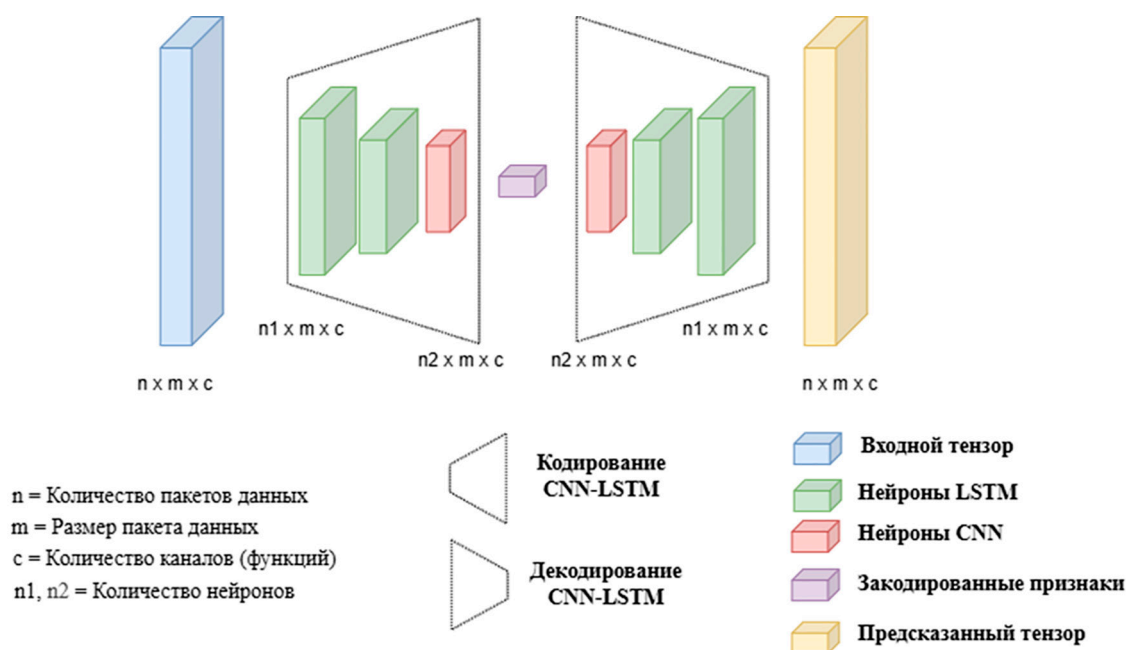
Мы делим набор данных в соотношении 7:1:2, где 70% – для обучения, 10% – для проверки и 20% – для тестирования.

Краеугольным камнем эффективности модели является большой объём обучающих данных. Соотношение 70% гарантирует, что у модели достаточно выборки для изучения сложных закономерностей и закономерностей в данных, что позволяет избежать недообучения, вызванного недостатком данных.

Проверочный набор обеспечивает надёжную и стабильную оценку производительности для различных настроек гиперпараметров.

Поскольку результаты оценки на тестовом наборе являются окончательным заключением, мы надеемся, что они будут максимально точными и надёжными. Более крупный тестовый набор (20%) означает меньшую дисперсию оценки, что может лучше отражать обобщающую способность модели в реальном мире.

Нормализованное среднее значение и стандартное отклонение рассчитываются только для обучающего набора, а затем та же статистика используется для стандартизации проверочного и тестового наборов.



Архитектура CNN-LSTM

Такой подход напрямую блокирует будущую информацию из модели. Если μ и σ рассчитываются с использованием полного набора данных, колебания и экстремальные значения за период тестирования, и даже в будущем, будут переданы на этап обучения. Модель будет видеть будущее распределение заранее во время бэкестинга, что приведет к завышенной эффективности.

При использовании статистики обучающего набора проверочные и тестовые наборы принимаются пассивно, и независимо от того, насколько сильно они колеблются, они не будут загрязнять обучающее распределение, тем самым полностью отсекая канал утечки информации.

Как показано на рисунке 1, исходная последовательность сначала проходит через несколько одномерных сверточных и ReLU-слоев, что позволяет сети эффективно изучить собственные технические характеристики. Слой пулинга сглаживает высокочастотный шум, а также выполняет понижающую дискретизацию, значительно сокращая временной шаг. В результате получается более чистая, высокоуровневая последовательность с извлеченными локальными признаками.

Иерархия и настройки гиперпараметров для ветвей CNN показаны ниже:

Входное окно: 5-мерный вектор $T = 60$ торговых дней.

Период прогнозирования: $\tau = 1$.

(1) Conv1D-1. вход=5, выход=64, ядро=3, расширение=1, причинное заполнение=2, затем проходит через слой BN и активируется ReLU, и, наконец, Dropout устанавливается на 0,2.

(2) Conv1D-2. вход = 64, выход = 128, ядро = 3, расширение = 2, причинный отступ = 4, затем проходит через слой пакетной нормализации (BN) и активируется ReLU, в конечном итоге устанавливая Dropout равным 0,2.

(3) Conv1D-3. вход = 128, выход = 256, ядро = 3, расширение = 4, причинный отступ = 8, затем проходит через слой пакетной нормализации (BN) и активируется ReLU, устанавливая Dropout равным 0,2.

(4) Глобальный слой MaxPooling1D сглаживает данные последовательности в 256-мерный вектор признаков фиксированной длины, содержащий наиболее важную информацию для каждого канала.

(5) Плотный слой принимает 256-мерный вектор, полученный на предыдущем

этапе, и подаёт его на вход полносвязного слоя. Это преобразует 256 входных единиц в 64 новых представления признаков более высокого уровня. Затем используется функция активации ReLU для введения нелинейности, что позволяет сети обучаться и моделировать более сложные паттерны данных.

(6) Выходной слой выдаёт непрерывное числовое значение для прогнозирования определённой величины. Кроме того, он использует сигмоидальную функцию активации, подходящую для бинарной классификации, для вывода значения вероятности, используемого для определения направления тренда или состояния.

Регуляризация: спад веса L2 $1e-4$, импульс BatchNorm 0,1.

Оптимизатор: AdamW, $lr = 1e-3$, косинусный отжиг $T_{max}=50$, партия=256.

Ранняя остановка: проверено, что потеря прекращается, если она не уменьшается в течение 10 последовательных эпох, максимум 200 эпох.

Иерархия и настройки гиперпараметров для ветвей LSTM показаны ниже:

Входное окно: 5-мерный вектор $T = 60$ торговых дней.

Период прогнозирования: $\tau = 1$.

(1) LSTM1. Вход=5, скрытые слои = 128, количество слоев = 1, форма входного тензора – размер пакета, длина последовательности – 5. Нет выпадений между слоями.

(2) LSTM2. Вход = 128, скрытые слои = 64, количество слоев = 1, форма входного тензора = 5, размер партии, длина последовательности = 5. Нет выпадений между слоями.

(3) Выходной слой выдаёт непрерывное числовое значение для прогнозирования определённой величины. Кроме того, он использует сигмоидальную функцию активации, подходящую для бинарной классификации, для вывода значения вероятности, используемого для определения направления тренда или состояния.

Регуляризация: повторное выпадение = 0,15, только LSTM-1. Уменьшение веса L2 $1e-5$.

Оптимизатор: Адам, $lr = 5e-4$, ReduceLROnPlateau (фактор=0,5, терпение=5), партия=256.

Ранняя остановка: проверено, что потеря прекращается, если она не уменьшается в течение 10 последовательных эпох, максимум 200 эпох.

$$\hat{y} = \sigma_{out} \left(W_{out} \cdot \text{LSTM} \left(\text{Pool} \left(\text{ReLU} \left(X * W_{CNN} \right) \right) \right) + b_{out} \right). \quad (1)$$

Как показано в формуле (1), $*$ представляет одномерную свертку, Pool представляет объединение с понижением дискретизации, LSTM(\cdot) представляет конечное скрытое состояние, а σ_{out} относится к соответствующей выходной активации.

$$X = [x_1, x_2, x_3, \dots, x_T] \in R^{TxT}, x_T \in R^f. \quad (2)$$

Как показано в формуле (2), длина одной выборки равна T , а каждый временной шаг содержит f исходных признаков.

$$q_{-t}^{(l,j)} = \max((0 \leq \tau < w_p) h(t \cdot s_p + \tau)^{(l,j)}, \quad (3)$$

$$Q = [q_1, q_2, \dots, q_T] \in R^{(T \times d)}. \quad (4)$$

Как показано в формулах (3) и (4), длина выходных данных еще больше сокращается, и после нескольких слоев стекирования сверток-пулинга окончательно получается Q , где d – число ядер свертки в последнем слое, а $T \ll T$.

$$z_t = \tanh(W(q_z, q_t) + b_z) \in R^{d_z}. \quad (5)$$

Как показано в формуле (5), Q напрямую поступает в LSTM-сеть во временных интервалах без необходимости дополнительных обучаемых параметров. В этой статье мы напрямую устанавливаем $z_t = q_t$, выполняем полносвязную операцию и снова сжимаем её.

$$v = h^T \in R^h. \quad (6)$$

Как показано в формуле (6), последнее скрытое состояние используется в качестве представления всей выборки. h – размерность скрытого слоя LSTM.

Затем укороченная последовательность поступает в LSTM-сеть. На этом этапе LSTM-сети достаточно сосредоточиться только на относительно плавной эволюции латентного состояния со средними и низкими частотами, и проблема исчезающего градиента значительно смягчается.

Таким образом, CNN отвечает за восприятие локальной формы, а LSTM – за запоминание долгосрочного контекста. Они соединены последовательно, образуя взаимодополняющую функцию, а не позволяют каждой из сторон в отдельности нести всю временную шкалу.

В реальных рыночных условиях 80% ежеминутных колебаний представляют собой случайный шум. Объединение данных в CNN – это естественный фильтр нижних частот, что делает его более гибким, чем

ручные скользящие средние. Более того, традиционный количественный анализ требует ручного построения более 20 индикаторов, таких как MACD и RSI. Ядра свертки CNN представляют собой локальные паттерны, что снижает необходимость в априорных допущениях. LSTM работает с последовательностями с пониженной дискретизацией, отображая 60 шагов понижения дискретизации на исходные 240 одномоментных линий, не беспокоясь о взрывном росте градиентов.

Таким образом, свёрточный блок в данной работе действует как локальное увеличительное стекло, сначала извлекая действительно полезную микротекстуру из рыночных данных высокого разрешения. Затем блок LSTM действует как низкоскоростной видеомаягнитофон, воспроизводя долгосрочную историю на уже увеличенных ключевых кадрах.

CNN-LSTM не просто объединяет две сети. Вместо этого используется двухступенчатый конвейер локальной абстракции и моделирования временных рядов. Это позволяет преодолеть недальновидность CNN и одновременно снизить неоднозначность памяти LSTM. Таким образом, это практичное решение, сочетающее точность и надёжность для многомасштабных задач прогнозирования временных рядов с высоким уровнем шума и низким отношением сигнал/шум, таких как прогнозирование фондового рынка.

Результаты исследования и их обсуждение

Мы разработали эксперименты на CNN, LSTM и CNN в сочетании с LSTM и выбрали пять показателей для оценки предсказательной способности модели.

Таблица 3

Оценка эффективности различных моделей на основе данных по акциям Amazon

	MAE	MSE	RMSE	MAPE	R^2
CNN	21.4285	461.8109	37.4009	59%	0.4584
LSTM	16.4103	282.4047	23.7053	27%	0.5635
CNN-LSTM	6.7808	72.8485	8.5351	13%	0.9475

Чтобы полностью изучить прогностические возможности различных моделей и представить различия в модельных прогнозах, в данной статье используются следующие оценочные индикаторы.

Формула расчета средней абсолютной погрешности приведена в формуле (7).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (7)$$

Расчет среднеквадратической ошибки показан в формуле (8).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (8)$$

Расчет среднеквадратической ошибки показан в формуле (9).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (9)$$

Расчет средней абсолютной процентной погрешности показан в формуле (10).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (10)$$

Расчет коэффициента решения показан в формуле (11).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2}. \quad (11)$$

где y_i используется для представления истинного значения, \hat{y}_i – прогнозируемое значение, n – количество выборок, а \bar{y} – среднее значение.

Среди этих метрик оценки MAE измеряет, насколько близки прогнозы модели к истинным значениям. MSE и RMSE измеряют степень вариации между прогнозируемыми и истинными значениями. MAPE оценивает среднее отклонение между прогнозируемыми и истинными значениями.

Чем ближе прогнозируемое значение к истинному, тем меньше соответствующее значение показателя; и наоборот, при увеличении разрыва значение показателя увеличивается. Следовательно, чем меньше значение вышеуказанного показателя, тем точнее прогноз модели. Кроме того, чем ближе R^2 к 1, тем лучше соответствие модели. Когда $R^2=1$, это означает, что прогнозируемое значение полностью соответствует истинному.

Как показано в таблице 3, CNN имеет только локальное поле зрения и воспринимает тренды как шум. Прогнозируемые значения часто колеблются в пределах $\pm 35\%$ от текущей цены, а среднеквадратичное отклонение (СКО) значительно увеличивается из-за большого отклонения. Прогнозируемое значение отклоняется от истинного в среднем на 60%, что хуже случайного блуждания и не может быть реализовано в транзакциях. Модель может объяснить только 46% колебаний, а оставшиеся 54% рассматриваются как белый шум и недообучение.

LSTM-метод обладает памятью и может использовать вчерашнюю цену закрытия в качестве ориентира, что сужает погрешность. Однако он по-прежнему подвержен высокочастотным сбоям, приводящим к появлению хвоста в остаточном распределении. MAPE достигла порогового значения, которое составляет менее 30%, однако внутрисуточный диапазон по-прежнему часто меняет направление, и реальная торговля часто останавливает потери. Он увеличился до 56%, но все еще был ниже практического проходного балла 0,6.

В модели CNN-LSTM CNN использует локальные признаки для подавления 1-минутного шума. LSTM работает только с чистыми последовательностями, уменьшая как смещение, так и дисперсию. Среднеквадратичное отклонение (RMSE) приближается к среднему значению ошибки (MAE), а эффект «хвоста» исчезает. MAPE уже на 15% ниже эмпирической линии, что означает, что 1-минутная K-линия имеет ошибку всего на 0,13% и имеет высокую толерантность к проскальзыванию.



Рис. 2. Прогноз цен закрытия торгов Amazon (тестовый набор)

Объяснительная способность R^2 достигает 95%, что практически точно предсказывает тренд и среднечастотный диапазон. В остаточной части остаётся только микро-структурный шум, который и является той частью, которая не подлежит торговле.

Как показано на рисунке 2, разработанная нами модель CNN-LSTM может точно предсказать цену закрытия Amazon. Эта модель использует свёрточную нейронную сеть для локального восприятия и понижения разрешения исходной многомерной последовательности. Используя механизмы распределения весов и объединения, она сжимает временной шаг на порядок, сохраняя при этом дискриминантные мелкозернистые признаки, что значительно снижает пространственно-временную сложность последующей сети долго-кратковременной памяти.

Часть LSTM использует преимущества стробированной памяти на очищенных коротких последовательностях, фиксируя долгосрочные динамические зависимости и достигая дополнительного моделирования от локального до глобального. Сквозное совместное обучение позволяет градиентам напрямую достигать передней части, устраняя необходимость в ручной сегментации кадров и проектировании признаков, а также автоматически обучая оптимальному пространственно-временному представлению. Эта структура естественным образом устойчива к изменениям

длины входных данных, а свёрточный слой можно предварительно обучить и перенести на сценарии с небольшими выборками, принимая во внимание как эффективность, так и обобщение.

LSTM хорошо подходит для фиксации долгосрочных зависимостей в данных последовательностей, тогда как CNN хорошо подходит для извлечения локальных особенностей данных изображений. Объединив преимущества обоих подходов, мы можем позволить модели учитывать как временную, так и пространственную информацию данных, сократить параметры и снизить риск переобучения, тем самым обеспечивая более точные прогнозы, лучшую производительность и более высокую эффективность обучения.

Выводы

В данном исследовании акции Amazon используются в качестве образца для систематического сравнения предсказательной эффективности CNN, LSTM и их структур слияния в реальных рыночных сценариях. На основе этого предлагается и верифицируется сквозная, тонко настраиваемая гибридная структура CNN-LSTM.

Экспериментальные результаты согласуются с разложением ошибок, указывая на то, что механизм локального рецептивного поля CNN может надёжно улавливать микрохарактеристики, такие как внутрисуточные колебания и внезапные разрывы событий,

но он ограничен фиксированной длиной ядра и отсутствием памяти, что затрудняет экстраполяцию долгосрочных тенденций. LSTM имеет значительное преимущество в памяти для циклических компонентов и поворотных моментов макроэкономической политики благодаря своей рекуррентной структуре с управляемыми воротами, но она очень склонна к переобучению в условиях высокого уровня шума и медленно реагирует на локальные мутации. Новая парадигма использования CNN в качестве извлекателя признаков на входе и LSTM в качестве декодера последовательностей может улучшить различные показатели оценки, сохраняя при этом общее количество параметров неизменным.

Подводя итог, можно сказать, что гибридная модель CNN-LSTM представляет собой новую парадигму прогнозирования цен акций, которая обеспечивает баланс

между микро- и макропамятью. Её архитектура, основанная на локальном свёрточном извлечении и последовательном моделировании памяти, подходит для прогнозирования акций, предоставляя компаниям и инвесторам точную справочную информацию.

Хотя эта модель значительно превосходит чистые базовые показатели CNN или LSTM на ежедневных данных Amazon, эксперимент охватывает только одну американскую акцию, одну частоту и период изменений американских акций. Его устойчивость на различных рынках, таких как акции класса А, акции Гонконга и криптоактивы, при разной частоте выборки или в течение более длительных периодов бэкестинга пока не подтверждена. Поэтому для подтверждения его обобщающей способности необходимы дополнительные испытания на различных классах активов, различных таймфреймах и различных рынках.

Библиографический список

1. Асайын Е.Ж., Барлыбаев А., Шарипбай А. Методы глубокого обучения в прогнозировании фондового рынка: обзор и будущие направления // Вестник науки. 2024. №11 (80). URL: <https://cyberleninka.ru/article/n/metody-glubokogo-obucheniya-v-prognozirovanii-fondovogo-rynka-obzor-i-buduschie-napravleniya> (дата обращения: 10.09.2025).
2. Глебова А.Г., Ковалева А.А. Прогнозирование волатильности российского биржевого рынка акций в условиях международных экономических санкций // Финансы: теория и практика. 2024. № 1. URL: <https://cyberleninka.ru/article/n/prognozirovanie-volatilnosti-rossiyskogo-birzhevogo-rynka-aktsiy-v-usloviyah-mezhdunarodnyh-ekonomicheskikh-sanktsiy> (дата обращения: 10.09.2025).
3. Теньковская Л.И. Прогнозирование котировок акций ПАО «НК «Лукойл» на основе полиномиальных моделей тенденции // Векторы благополучия: экономика и социум. 2024. Т. 52. №4. С. 168-185. DOI: 10.18799/26584956/2024/4/1924.
4. Теньковская Л.И. Моделирование котировок акций компании в условиях цикличности экономики // Journal of new economy. 2024. № 3. URL: <https://cyberleninka.ru/article/n/modelirovanie-kotirovok-aktsiy-kompanii-v-usloviyah-tsiklichnosti-ekonomiki> (дата обращения: 10.09.2025).
5. Chen Y., Lin W., Wang J.Z. A Dual-Attention-Based Stock Price Trend Prediction Model With Dual Features // IEEE Access. 2019. Vol. 7. P. 148047-148058. DOI: 10.1109/ACCESS.2019.2946223.
6. Efendi Riswan, Nureize Arbaiy, Mustafa Mat Deris. A new procedure in stock market forecasting based on fuzzy random auto-regression time series model // Information Sciences. 2018. № 441. № 113-132. DOI: 10.1016/j.ins.2018.02.016.
7. Ben Saïda Ahmed et al. Value-at-risk under market shifts through highly flexible models // Journal of Forecasting. 2018. № 37.8. № 790-804. DOI: 10.1002/for.2503.
8. Guan H., Jie H., Guan S. et al. A novel fuzzy-Markov forecasting model for stock fluctuation time series // Evol. Intel. 2020. № 13. С. 133–145. DOI: 10.1007/s12065-019-00328-0.
9. Luo Zhidan et al. A hybrid model for financial time-series forecasting based on mixed methodologies // Expert Systems. 2021. № 38.2. P. e12633. DOI: 10.1111/exsy.12633.
10. Pawar K., Jalem R.S., Tiwari V. Stock Market Price Prediction Using LSTM RNN. In: Rathore V., Worring M., Mishra D., Joshi A., Maheshwari S. eds. Emerging Trends in Expert Applications and Security. Advances in Intelligent Systems and Computing. 2019. Vol. 841. Springer, Singapore. DOI: 10.1007/978-981-13-2285-3_58.
11. Гарафутдинов Р.В., Шевченко М.В. Формирование портфеля акций с использованием нейронных сетей архитектуры LSTM и фрактального анализа // Цифровые модели и решения. 2025. № 2. URL: <https://>

cyberleninka.ru/article/n/formirovanie-portfelya-aksiy-s-ispolzovaniem-neyronnyh-setey-arhitektury-lstm-i-fraktalnogo-analiza (дата обращения: 10.09.2025).

12. Источники данных по акциям Amazon. [Электронный ресурс]. Режим доступа. URL: <https://finance.yahoo.com/quote/AMZN/history/?guccounter=2> (дата обращения: 02.09.2025).

13. Копырин А.С., Макарова И.Л. Алгоритм препроцессинга и унификации временных рядов на основе машинного обучения для структурирования данных // Программные системы и вычислительные методы. 2020. № 3. URL: <https://cyberleninka.ru/article/n/algoritm-preprotssinga-i-unifikatsii-vremennyh-ryadov-na-osnove-mashinnogo-obucheniya-dlya-strukturirovaniya-dannyh> (дата обращения: 10.09.2025).

14. Власенко А.В., Жданов А.А. Предварительная обработка временных рядов производственных систем // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. 2018. № 4 (231). URL: <https://cyberleninka.ru/article/n/predvaritelnaya-obrabotka-vremennyh-ryadov-proizvodstvennyh-sistem> (дата обращения: 10.09.2025).

15. Скляр А.Я. Анализ временных рядов и выявление процессов с размытой периодичностью // Кибернетика и программирование. 2018. № 6. С. 56-64. DOI: 10.25136/2644-5522.2018.6.27069. URL: https://nbpublish.com/library_read_article.php?id=27069 (дата обращения: 12.09.2025).

16. Гололобов Н.В. Статистические алгоритмы в задачах выявления аномалий в наборах данных // Актуальные вопросы современной науки: Сборник трудов по материалам VIII Всероссийского конкурса научно-исследовательских работ. 2022. 15 марта. URL: https://perviy-vestnik.ru/wp-content/uploads/2022/05/2022-E-198-03_22.pdf#page=79 (дата обращения: 15.09.2025).