

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ

ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

**РАЗРАБОТКА ПРОГРАММНОГО ПРОДУКТА ДЛЯ
ПРОГНОЗИРОВАНИЯ ФОНДОВОГО РЫНКА**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 451 группы

направления 38.03.05 — Бизнес-информатика

механико-математического факультета

Пузаркина Дениса Вячеславовича

Научный руководитель

к. э. н., доцент

А. Р. Файзлиев

Заведующий кафедрой

д. ф.-м. н., доцент

С. П. Сидоров

Саратов 2024

ВВЕДЕНИЕ

Прогнозирование движения стоимости ценных бумаг на фондовом рынке является важнейшей задачей для участников рынка. Но стоит отметить, что поставленная задача не является тривиальной. Фондовый рынок остается сложной и динамической средой, что делает прогнозирование поведения цен затруднительным.

Тем не менее, несмотря на эту непредсказуемость, существуют различные индикаторы и методы, которые могут помочь выявить определенные паттерны и закономерности в динамике цен. Эти индикаторы могут служить основой для разработки эффективных торговых стратегий.

В настоящее время в сфере технического анализа фондового рынка все большей популярностью пользуется такой метод, как машинное обучение. Алгоритмы машинного обучения позволяют обнаруживать закономерности и паттерны в больших объемах данных, которые человек не в состоянии увидеть.

Однако стоит отметить, что машинное обучение имеет свои ограничения. Эффективность алгоритмов машинного обучения зависит от качества и объема обучающих данных. Если данные недостаточно полны, результаты могут быть искажены.

Таким образом, машинное обучение представляет собой перспективный инструмент анализа динамики фондового рынка. Его использование может помочь участникам рынка принимать решения и улучшить свои торговые результаты. Однако необходимо помнить о его ограничениях и подходить к использованию данного инструмента с должной ответственностью, не полагаясь исключительно на его результаты.

Целью этой работы является разработка программного продукта на языке программирования Python позволяющего прогнозировать движения цен акций и генерировать сигналы о входе в сделку. в основе решения задач лежат применения технического анализа фондового рынка и машинного обучения с использованием нейронных сетей.

Для достижения поставленной цели потребуется выполнение следующих задач:

- разработка моделей рекуррентной нейронной сети LSTM;
- тестирование моделей в торговле;
- сравнение методов.

Выполнение этих задач позволит создать решение для прогнозирования динамики цен на фондовом рынке и демонстрации торговых операций с использованием машинного обучения.

Данная выпускная квалификационная работа состоит из следующих частей: введения, где описаны актуальность работы, цель, и задачи, теоретической части, где описываются понятие фондового рынка, анализ и методы машинного обучения, практической части, где разрабатывается программный продукт для прогнозирования фондового рынка, заключительной части, где подводится итог данной работы.

Основное содержание работы

Введение содержит информацию об актуальности данной работы, её целях и задачах, которые планируется решить в ходе исследования.

Первый раздел рассказывает о теоретических основах прогнозирования динамики цен на фондовом рынке. Здесь рассматривается гипотеза эффективного рынка, приводится определение технического анализа, описываются различные методы и подходы к прогнозированию цен на акции.

Второй раздел посвящен применению методов машинного обучения для прогнозирования цен акций. В нем описываются возможности использования машинного обучения для анализа данных фондового рынка, рассматриваются существующие подходы к применению машинного обучения для прогнозирования фондового рынка.

В третьем разделе описана разработка и тестирование моделей машинного обучения, а именно рекуррентной нейронной сети, основанной на исторических данных, и рекуррентной нейронной сети, основанной на техническом индикаторе MACD.

Для обучения, тестирования и торговли выбрана ценная бумага «SVET» эмитента «Светофор» с таймфреймом в один час. Такой таймфрейм обусловлен тем, что ежечасные данные менее подвержены влиянию внешних факторов, следовательно они будут более предсказуемы.

Исторические данные получены с биржи «МОЕХ» с помощью библиотеки «arimoeх».

Ежечасный таймфрейм представлен в период с 1 января 2023 года по 19 апреля 2024 года. Первое значение закрытия цены равняется 17,6 рублей, последнее значение закрытия цены равняется 28,2 рубля. На рынке за последние месяцы наблюдается нисходящий тренд. Визуализация движения цены представлены на рисунке 1



Рисунок 1 – график временного ряда цен тикера "SVET"

На вход модель получает только цену закрытия, поэтому в данных необходимо оставить только её.

Модели берут определенную последовательность значений временного ряда и основе этой последовательности предсказывать следующее значение. Данная модель берет 10 предыдущих значений. Приведем данные в трехмерный формат методом скользящего окна, где первое измерение - это количество последовательностей, второе - размер последовательности, третье - количество признаков.

Получили выборку объясняющей переменной X . Она имеет 3411 последовательностей, в которых значатся 10 наблюдений и один признак. Также получили выборку объясняемой переменной y . Данная выборка имеет 3411 последовательности, в которой имеется один признак.

(3411, 10, 1)

(3411, 1)

При разделении выборки на обучающую и тестовую получим следующий результат:

Training Data shape

(3410, 10, 1)

(3410, 1)

Testing Data shape

(1, 10, 1)

(1, 1)

Для Построении архитектуры модели использовалась библиотека keras.

Имеем следующую топологию сети:

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 10, 10)	480
lstm_1 (LSTM)	(None, 20)	2480
dense (Dense)	(None, 1)	21

Total params: 2,981

Trainable params: 2,981

Non-trainable params: 0

Первыми двумя слоями нейронной сети являются два рекуррентных слоя LSTM по 10 и 20 нейронов соответственно. На выходе стоит полносвязный слой с одним нейроном.

Различные варианты нейронных сетей были протестированы и обучены методом проб и ошибок. В результате предложенная выше архитектура при 125-ти эпохах показала наилучшую точность.

Сопоставление прогнозов с фактическими данными представлено на рисунке 2.

Более детальное представление последних тиков представлены на рисунке 3.

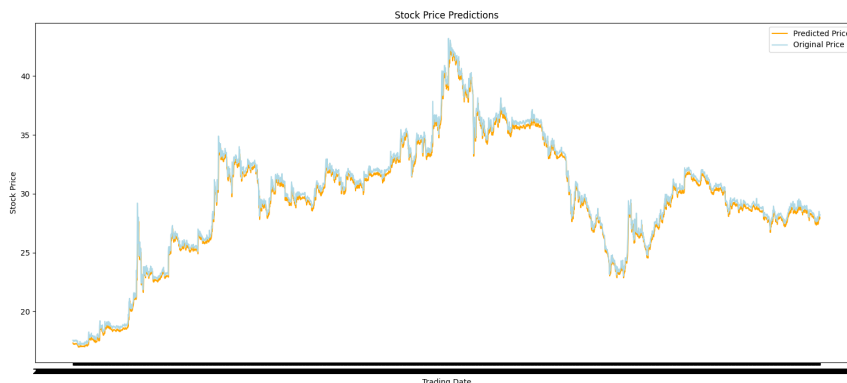


Рисунок 2 – Визуализация прогноза временного ряда цен тикера "SVET"

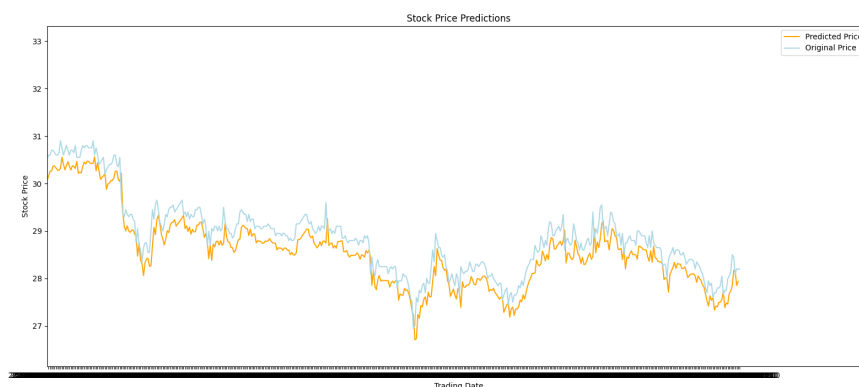


Рисунок 3 – Детальная визуализация прогноза временного ряда цен тикера "SVET"

Потери составили 0,02 процента, оценка RMSE равна 10 процентов, Модель угадывает общее направление тренда, но график прогнозов имеет запаздывание. Фактический прогноз тестовых данных:

Predicted Prices

```
[[27.93793]]
```

Original Prices

```
[[28.2]]
```

Точность прогноза составила 99,07 процентов.

Протестируем модель в торговле. Суть стратегии заключается в том, что если прогнозируемая цена на конец часа оказывается выше текущей, это будет сигналом к входу в сделку. Если предсказанная цена остается выше при входе в сделку, следует ожидать следующего изменения цены. Когда прогнозируемая цена становится ниже текущей, это сигнализирует о необходимости выхода из сделки.

Совершенные сделки представлены на рисунке 4, где зеленой сплошной линией отмечена покупка, а красной пунктирной - продажа.

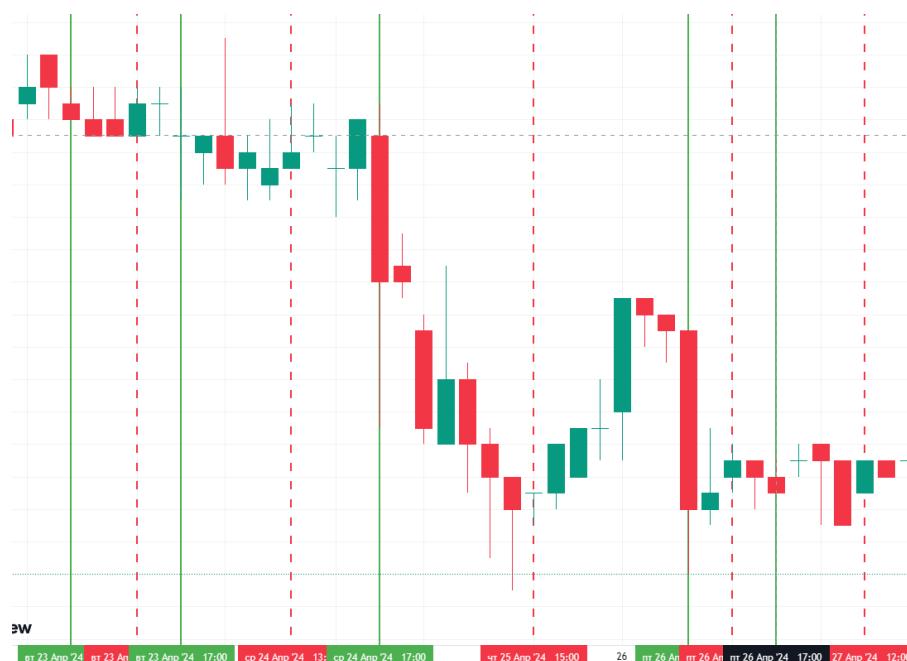


Рисунок 4 – Совершенные сделки с помощью модели рекуррентной нейронной сети LSTM на основе исторических данных цены

Итого сумма прибыли за неделю составляет -0,2 рубля.

Данные для второй модели представлены в период с 1 января 2023 года по 12 апреля 2024 года. Первое значение закрытия цены равняется 17,6 рублей, последнее значение закрытия цены равняется 28,3 рубля. На рынке за последние месяцы наблюдается нисходящий тренд.

Вторая модель основана на техническом индикаторе MACD. Алгоритм прогнозирует движение линии MACD. Входом в сделку будет являться разность значений MACD и сигнальной линии. Если данное значение меняет свой знак, то это будет являться сигналом к покупке или продаже. Для этого необходимо вычислим функцию MACD и Signal и в следующей колонке

«MACD_Signal_diff» записать разность значений.

В этом случае, если значение разности колонки отрицательное, то запишем 0 в колонку «signal», если положительное, то 1. На вход модель будет получать цену закрытия и значение сигнала.

Построим архитектуру нейронной сети, обучим и протестируем ее.

```
Model: "sequential"
```

```
-----  
Layer (type)                Output Shape                Param #  
-----  
lstm (LSTM)                  (None, 16, 48)             9984  
  
lstm_1 (LSTM)                 (None, 96)                  55680  
  
dense (Dense)                 (None, 1)                   97
```

```
=====  
Total params: 65,761  
Trainable params: 65,761  
Non-trainable params: 0  
-----
```

Потери составили 9 процентов, оценка RMSE равна 5,8 процентов, модель в целом верно прогнозирует сигналы, но прогнозируемый сигнал отстает от реальных данных на один тик. Фактический прогноз тестовых данных:

```
y_test  
[1. 1. 1. 1. 1.]
```

```
pred  
[[0.9435729 ]  
 [0.9550271 ]  
 [0.94751924]  
 [0.9362049 ]  
 [0.92984635]]
```


Протестируем модель в торговле. Стратегия будет заключаться в сравнении текущего сигнала и прогнозируемого. Если сигнал будет меняться с нуля на единицу и при этом MACD будет меньше нуля - это будет считаться входом в сделку. Если сигнал будет меняться с единицы на ноль и при этом MACD будет больше нуля - это выход из сделки.

Итого сумма прибыли составляет -0,3 рубля. За период тестирования совершилась одна сделка. Цена покупки равняется 28,65 рублей, цена продажи — 28,35 рубля.

Исходя из полученных данных, построенные модели обладают способностью выявлять тренды и прогнозировать цены с точностью, достигающей 99 процентов. Тем не менее, обе модели демонстрируют определенное запаздывание, что ограничивает их возможности по получению более выгодных точек входа.

Первая стратегия ориентирована на открытие коротких позиций с целью получения быстрой прибыли, в то время как вторая стратегия нацелена на более длительные позиции.

Для получения прибыли возможно использование иных методов генерирования сигналов и других более научно обоснованных стратегий.

Для повышения точности прогнозирования рекомендуется дополнительно использовать различные переменные и технические индикаторы.

Например, можно добавить в набор данных значения открытия и закрытия, максимум и минимум цены.

Также можно использовать технические индикаторы, такие как скользящие средние и индекс относительной силы (RSI)

Интеграция этих дополнительных переменных и индикаторов может помочь моделям более точно прогнозировать движение цен и предоставлять более выгодные точки входа для принятия торговых решений.

Основные результаты

В данной работе выполнены следующие задачи:

- разработаны модели рекуррентной нейронной сети LSTM на языке программирования Python, которые позволяют прогнозировать движения цен акций и генерировать сигналы о покупке или продаже;

- были исследованы различные архитектуры и параметры моделей LSTM для достижения оптимальной производительности;
- в основе решения задач легли применение технического анализа фондового рынка и машинного обучения с использованием нейронных сетей
- решения были протестированы на исторических данных, а также в реальных торговых условиях;
- были проанализированы результаты тестирования и торговли и сделаны выводы об эффективности моделей

В целом, разработанные модели и методы прогнозирования могут быть использованы для принятия торговых решений.