

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

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

Кафедра теории функций и стохастического анализа

**ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ
ПРОГНОЗИРОВАНИЯ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

Студентки 2 курса 248 группы

направления 09.04.03 — Прикладная информатика

механико-математического факультета

Хайрова Рената Равилевна

Научный руководитель

доцент, к. ф.-м. н., доцент

А. В. Шаталина

Заведующий кафедрой

д. ф.-м. н., доцент

С. П. Сидоров

Саратов 2023

Во **введение** сформулирована цель и основные задачи исследования, обосновывается актуальность и значимость исследования, а также практическое применение результатов работы.

Актуальность использования нейронных сетей возрастает в условиях современных мировых тенденций по мере увеличения спроса на анализ и обработку большого объема данных. Нейронные сети успешно используются в самых различных областях - бизнесе, экономике, медицине, геологии, физике и т.д., и активно используются на практике везде, где нужно решать задачи прогнозирования, классификации и управления. Такие характеристики нейронных методов, как возможность нелинейного моделирования и сравнительная простота реализации, часто делают их незаменимыми при решении сложнейших многомерных задач. Там, где линейная аппроксимация неудовлетворительна и линейные модели работают плохо, а таких задач достаточно много, основным инструментом становятся нейронные методы. Кроме того, нейронные сети хорошо подходят для обработки большого числа переменных, что является барьером в случае моделирования линейных зависимостей.

Целью магистерской работы является:

1. Изучение различных известных методов нейронных сетей и модификация наиболее подходящего метода для прогнозирования будущих значений основанного на нейронных сетях;
2. Изучение различных классических методов прогнозирования и модификация наиболее подходящего метода для прогнозирования будущих значений основанного на машинном обучении;
3. Выполнение предсказания значения переменных на основе данных методов и формирование вывода в пользу одного или другого метода.

Объектом исследования является нейронные сети и их применение к прогнозированию значений.

Предметом исследования является рекуррентная нейронная сеть LSTM и ее применение в прогнозировании стоимости акций компании AMD.

Работа прошла апробацию на различных конференциях, в частности, на ежегодной студенческой конференции «Актуальные проблемы математики и механики», которую проводил механико-математический факультет СГУ в апреле 2023 года, в секции «Анализ данных», в XI Международной

молодежной научно-практической конференции «Математическое и компьютерное моделирование в экономике, страховании и управлении рисками», ноябрь 2022 года.

Основное содержание работы

В **первом** разделе рассмотрены общие аспекты разработки прогнозирующих систем: основные определения и методы прогнозирования, модели временных последовательностей, критерии производительности прогнозирующих систем и другие общие вопросы, касающиеся разработки прогнозирующих систем.

Прогноз и цели его использования

Прогнозирование - это предсказание будущих событий и один из ключевых этапов в управлении принятием решений. Целью прогнозирования является уменьшение риска при принятии решений. Прогноз обычно получается ошибочным, но ошибка зависит от используемой прогнозирующей системы. Чем больше данных будет на входе прогнозирования, тем точнее получится результат анализа, что непосредственно оказывает влияние на принятие и результат решений управления.

Важно заметить, что снизить среднюю ошибку прогнозирования ниже определенного уровня невозможно, вне зависимости от того насколько сложен примененный метод. Прогнозирование никогда не сможет полностью свести риск при принятии решений к 0, поэтому необходимо явно определять неточность прогноза. Обычно, принимаемое решение определяется результатами прогноза (при этом предполагается, что прогноз правильный) с учетом возможной ошибки прогнозирования.

Критерии производительности

Прогнозирующая система должна выполнять две основные функции: генерацию прогноза и управление прогнозом. Генерация прогноза включает получение данных для уточнения модели прогнозирования, проведение прогнозирования, учет мнения экспертов и предоставление результатов прогноза пользователю. Управление прогнозом включает в себя наблюдение процесса прогнозирования для определения неконтролируемых условий и поиск возможности для улучшения производительности прогнозирования.

Во **втором** разделе рассмотрены наиболее известные виды нейронных

сетей, их модели и принцип работы.

Нейронная сеть (также искусственная нейронная сеть, ИНС) - математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей - сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы.

- С точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п.
- С точки зрения математики, обучение нейронных сетей - это многопараметрическая задача нелинейной оптимизации.
- С точки зрения кибернетики, нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники.
- С точки зрения развития вычислительной техники и программирования, нейронная сеть - способ решения проблемы эффективного параллелизма.
- С точки зрения искусственного интеллекта, ИНС является основой философского течения коннекционизма и основным направлением в структурном подходе по изучению возможности построения (моделирования) естественного интеллекта с помощью компьютерных алгоритмов.

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Возможность обучения - одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что в случае успешного обучения сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искажённых данных.

Способности нейронной сети к прогнозированию напрямую следуют из её способности к обобщению и выделению скрытых зависимостей между входными и выходными данными. После обучения сеть способна предсказать бу-

дущее значение некой последовательности на основе нескольких предыдущих значений и (или) каких-то существующих в настоящий момент факторов. Например, прогнозирование котировок акций на основе котировок за прошлую неделю может оказаться успешным (а может и не оказаться), тогда как прогнозирование результатов завтрашней лотереи на основе данных за последние 50 лет почти наверняка не даст никаких результатов.

В **третьем** разделе представлены: теоретическая часть описывающая работу рекуррентной нейронной сети - LSTM, общее описание предметной области и программа на языке Python выполняющая прогнозирование цены акции компании AMD.

LSTM — это подвид сверточной нейронной сети со способностью «запоминания» значений наблюдений ранних этапов для использования их в будущем.

Long short-term memory нейронная сеть (долгая краткосрочная память) — это искусственная сверточная нейронная сеть, используемая в области глубокого обучения. В отличие от обычных нейронных сетей с прямой связью, LSTM сети имеют обратные связи. Подобные сети способны обрабатывать не только отдельные единичные данные (например, изображения), но и целые последовательности данных (например, аудиозапись речи или видеозаписи). LSTM нейронные сети хорошо подходят для классификации, обработки и построения прогнозов на основе временных рядов, где взаимосвязанные явления могут происходить с неопределенным временным лагом.

В LSTM сети существуют 3 типа слоев:

1. Слой «забывания» (Forget gate) — на выход подается число от 0 до 1, где 1 обозначает необходимость полного запоминания, а 0 полностью стирает из памяти.

2. Слой «запоминания» (Memory gate) выбирает, какие данные необходимо сохранить. В первую очередь с помощью сигмоидного слоя выбираются значения, которые затем запоминаются.

3. Выходной слой (Output gate) выбирает информацию из каждой «ячейки», в которой произведено запоминание.

Описание предметной области

AMD (Advanced Micro Devices) — американский производитель интегральной микросхемной электроники, один из крупнейших производителей центральных процессоров, графических процессоров и адаптеров (после приобретения ATI Technologies в 2006 году), материнских плат и чипсетов для них также твердотельные накопители под торговой маркой Radeon.

Для прогнозирования значения будущей цены будем использовать показатели с 2015 по 2022 г.:

- Минимальной цены акции за период;
- Максимальной цены акции за период;
- Цены открытия на рынке;
- Цены закрытия на рынке.

Входные данные для анализа и прогноза были взяты с сайта Yahoo Finance, содержащий следующие показатели: «Дата», «Открытие», «Максимум», «Минимум», «Закрытие», «Скорректированное закрытие», «Объем» за каждый открытый день торгов на бирже. На рисунке 1 представлена статистическая характеристика данных за весь период с 2015 по 2022 годы:

	Open	High	Low	Close	Adj Close	Volume
count	1765.000000	1765.000000	1765.000000	1765.000000	1765.000000	1.765000e+03
mean	37.881949	38.644725	37.061813	37.866385	37.866385	5.657274e+07
std	38.974316	39.745377	38.100747	38.927574	38.927574	3.861427e+07
min	1.620000	1.690000	1.610000	1.620000	1.620000	2.606600e+06
25%	9.990000	10.170000	9.820000	10.000000	10.000000	3.266560e+07
50%	20.010000	20.590000	19.469999	19.990000	19.990000	4.860020e+07
75%	57.070000	58.130001	55.630001	56.950001	56.950001	7.294650e+07
max	163.279999	164.460007	156.100006	161.910004	161.910004	3.250584e+08

Рисунок 1 - Статистическая характеристика данных показателей акций компании AMD

Анализ значений акций AMD за период с 2015 по 2022 годы

В рамках работы был выполнен анализ и сравнение показателей минимальной и максимальной цены, а также анализ и сравнение показателей цены открытия и закрытия торгов на бирже в разрезе каждого года с 2015 по 2022 г. на языке программирования Python.

Общий график, демонстрирующий движение значений акций AMD по всем выбранным показателям с 2015 по 2022 годы, представленный на рисунке 2.

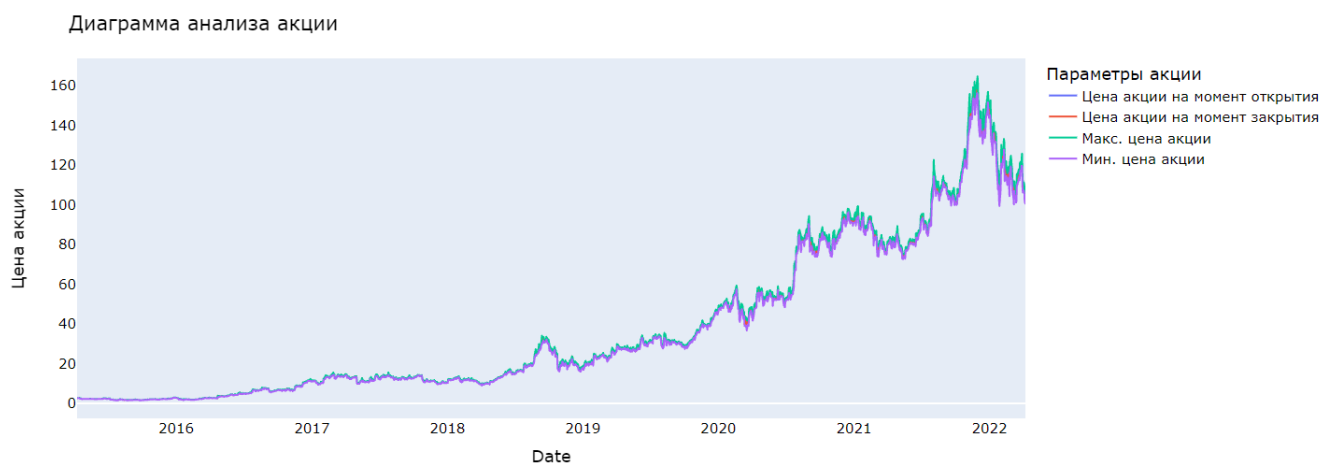


Рисунок 2 - Диаграмма анализа акций компании AMD с 2015 по 2018 годы по всем выбранным показателям

С полным перечнем графиков можно ознакомиться в Приложении А к Выпускной квалификационной работе.

Построение модели LSTM для прогноза цены акций AMD

В ходе работы над LSTM последовательно были выполнены следующие шаги:

- 1) Подготовка и нормализация данных;
- 2) Разделение данных на обучающую и тестовую выборки;
- 3) Построение модели LSTM, определение функции потерь и оптимизатора;
- 4) Обучение модели;
- 5) Анализ модели;

6) Построение графика прогнозируемой цены.

В результате выполнения программы был получен график цен акций на момент закрытия с прогнозом на следующий 30-дневный период. Результат представлен на рисунке 3:



Рисунок 3 - График цен акций AMD на момент закрытия с прогнозом на следующий 30-дневный период

Результатом работы нейронной сети является спрогнозированное значение на 30 дней вперед, а именно, цена акции AMD будет равна 109,2\$.

В четвертом разделе представлены: теоретическая часть описывающая работу модели авторегрессии скользящего среднего - ARIMA и программа на языке Python выполняющая прогнозирование цены акции.

ARIMA — это обобщенная модель ARMA (Autoregressive moving average), которая соединяет в себе AR (Autoregressive) и MA (Moving average) процессы. Модель ARIMA(p, d, q) состоит из следующих частей:

AR - autoregressive. Регрессионная модель, которая использует зависимость между наблюдениями и число интегрированных наблюдений (p);

I - integrated. Для обеспечения стационарности путем взятия разностей (d);

MA - moving average. Подход, анализирующий зависимость между наблюдениями и остатками при применении модели к интегрированным наблюдениям (q).

Комбинация этих двух моделей дает нам модель ARIMA порядка (p, q):
$$x_t = c + \sum^p \varphi_i x_{t-1} + \epsilon_t + \sum^q \theta_i \epsilon_{t-1} .$$

Параметры p и q называются порядком AR и MA, соответственно. Использование ARIMA позволяет делать прогноз на нестационарных данных из-за введения в модель интегрированности. Это достигается с помощью взятия разностей — вычитания уровней временного ряда друг из друга.

Прогнозирование цены акции AMD с помощью модели ARIMA

В ходе работы над моделью ARIMA был выбран аналогичный временной интервал, что для работы над нейронной сетью LSTM.

Последовательно были выполнены следующие шаги:

- 1) Подготовка и нормализация данных;
- 2) Разделение данных на обучающую и тестовую выборки;
- 3) Построение модели ARIMA и графика прогнозируемой цены;
- 5) Анализ модели.

В результате выполнения программы была построена модель с прогнозом на следующий 30-дневный период. Результат работы кода представлен на рисунке 4:

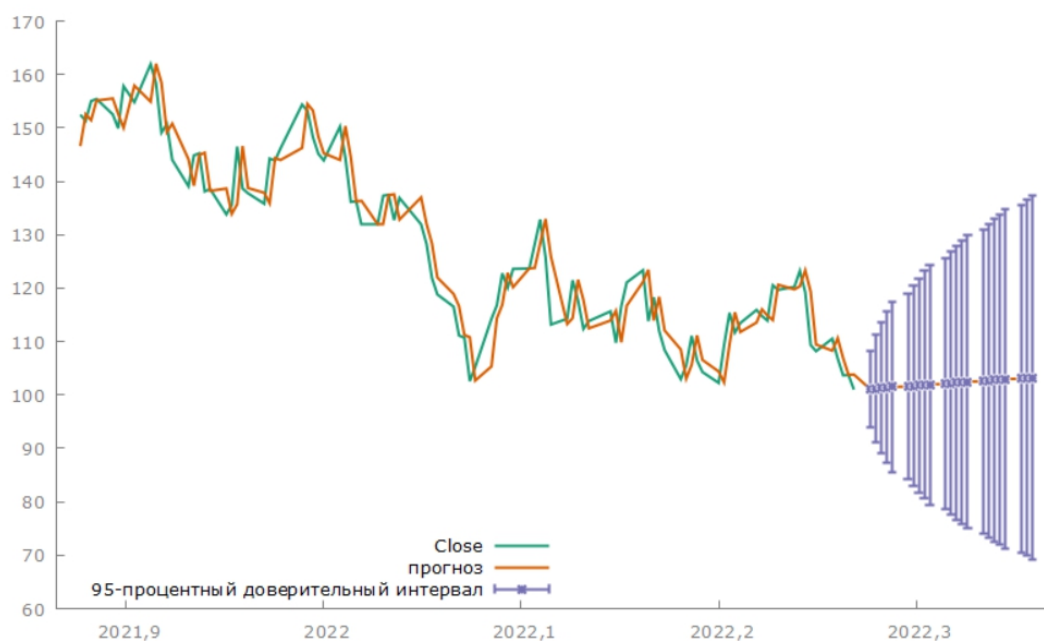


Рисунок 4 - График цен акций AMD на момент закрытия с прогнозом на следующий 30-дневный период

Результатом работы регрессионной модели ARIMA является диапазон спрогнозированных значений на 30 дней вперед, и конкретное значение, а именно, цена акции AMD будет равна 114\$.

В заключении приведены результаты магистерской работы.

В ходе выполнения работы был выполнен обзор основных теоретических положений, составляющих математическую основу нейронных сетей и регрессионного анализа. Был выполнен поиск, проведена подготовка и предварительный анализ статистических данных для акции компании AMD. На основе этих данных выполнено моделирование нейронной сети LSTM и построение модели ARIMA, прогнозирующей значение акции на 30 дней вперед.

Также были освоены основы программирования на языке Python и разработана на нем программа, выполняющую процедуру прогнозирования нейронной сети LSTM.

В процессе работы также было необходимо реализовать программу, позволяющую вычислять прогнозное значение цены акции для модели машинного обучения ARIMA.

Анализ выполненных исследований показал, что на финансовых временных рядах модель LSTM показала себя лучше, чем модель ARIMA:

- с помощью рекуррентной нейронной сети LSTM был получен точный результат будущего прогноза цены акции компании AMD через 30 дней, который равен 109\$;

- с помощью модели ARIMA получен диапазон значений возможного движения цены акции компании AMD на ближайшие 30 дней. Значение, спрогнозированное моделью равно 114\$. Но предложенный диапазон достаточно велик, что не позволяет сделать однозначные выводы по изменению цены акции и, соответственно, определить выгоду от покупки/продажи.

Также значения корня из среднеквадратичной ошибки RMSE моделей равны 4,5 и 5,6, соответственно. Применение модели LSTM позволяет уменьшить значение RMSE на 22,5% по сравнению с моделью ARIMA

Результаты данной работы имеют практическую ценность в области применений рекуррентных нейронных сетей и могут быть хорошим инструментом для практического применения в области торговли на бирже. Однако, важно учитывать много факторов перед применением подобных инструментов и основная задача, определяющая результат прогнозирования - это подбор параметров для будущей нейронной сети.