

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное образовательное учреждение высшего
 образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа ядерных технологий
 Направление подготовки 01.03.02 Прикладная математика и информатика
 ООП/ОПОП Прикладная математика в инженерии
 Отделение школы (НОЦ) экспериментальной физики

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема работы
Классификация изображений с помощью сверточных нейронных сетей

УДК 004.932.032.26

Обучающийся

Группа	ФИО	Подпись	Дата
0В92	Макаревич Дарья Васильевна		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Семенов Михаил Евгеньевич	к.ф.-м.н.		

Со-руководитель ВКР (по разделу «Концепция стартап-проекта»)

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Горюнова Наталия Николаевна	к.э.н.		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Антоневич Ольга Алексеевна	к.б.н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП/ОПОП, должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Крицкий Олег Леонидович	к.ф.-м.н.		

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа ядерных технологий
Направление подготовки 01.03.02 «Прикладная математика и информатика»
Отделение экспериментальной физики

УТВЕРЖДАЮ:

Руководитель ООП

_____ Крицкий

О.Л.

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Бакалаврской работы (стартап)

Студенту:

Группа	ФИО
0В92	Макаревич Дарье Васильевне

Тема работы:

Классификация изображений с помощью сверточных нейронных сетей	
Утверждена приказом директора (дата, номер)	06.02.2023, №37-46/с

Срок сдачи студентом выполненной работы:	31.05.2023
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	<i>Временные ряды котировок ликвидных компаний (цены закрытия, выборка не менее 252 наблюдений)</i>
Перечень подлежащих исследованию, проектированию и разработке вопросов	<ol style="list-style-type: none"> 1. Провести анализ методов отображения временных рядов в графические изображения 2. Провести анализ методов конструирования признаков 3. Выбор метода прогнозирования многомерного временного ряда с использованием признаков графического изображения 4. Разработать алгоритм и программную реализацию для обучения нейронной сети 5. Провести тестирование и подготовить рекомендации по использованию проекта.
Перечень графического материала	<ol style="list-style-type: none"> 1. Графики исходных данных и рекуррентных диаграмм 2. Архитектура нейронной сети 3. Матрицы ошибок 4. Графики метрик оценки качества решения задачи классификации
Консультанты по разделам выпускной квалификационной работы <i>(если необходимо, с указанием разделов)</i>	
Раздел	Консультант
Концепция стартап-проекта	Горюнова Наталия Николаевна
Социальная ответственность	Антоневич Ольга Алексеевна

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	01.03.2023
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семенов Михаил Евгеньевич	к. ф.-м. н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0В92	Макаревич Дарья Васильевна		

**ЗАДАНИЕ К РАЗДЕЛУ
«КОНЦЕПЦИЯ СТАРТАП-ПРОЕКТА»**

Обучающемуся:

Группа	ФИО
0В92	Макаревич Дарья Васильевна

Школа	Инженерная школа ядерных технологий	Отделение школы (НОЦ)	Экспериментальной физики
Уровень образования	Бакалавриат	Направление/ООП/ОПОП	01.03.02 Прикладная математика и информатика

Перечень вопросов, подлежащих разработке:	
<i>Проблема конечного потребителя, которую решает продукт, который создается в результате выполнения НИОКР (функциональное назначение, основные потребительские качества)</i>	<i>Сложность предсказания цен акций на финансовых рынках в силу чувствительности используемых методов к нестационарности показателей</i>
<i>Способы защиты интеллектуальной собственности</i>	<i>Регистрация исходного кода, патентование алгоритма</i>
<i>Объем и емкость рынка</i>	<i>Объем рынка СФО на год – 388293,7 руб.</i>
<i>Современное состояние и перспективы отрасли, к которой принадлежит представленный в ВКР продукт</i>	<i>Прогнозируется рост объема рынка</i>
<i>Себестоимость продукта</i>	<i>30 000 руб.</i>
<i>Конкурентные преимущества создаваемого продукта и Сравнение технико-экономических характеристик продукта с отечественными и мировыми аналогами</i>	<i>На основании конкурентных преимуществ</i>
<i>Целевые сегменты потребителей создаваемого продукта</i>	<i>Количественные аналитики, трейдеры</i>
<i>Бизнес-модель проекта, производственный план и план продаж</i>	<i>Модель по А. Остервальду</i>
<i>Стратегия продвижения продукта на рынок</i>	<i>Выступления на конференциях, публикации в тематических журналах</i>
Перечень графического материала:	
<i>При необходимости представить эскизные графические материалы(например, бизнес-модель)</i>	<i>Модель по А. Остервальду, таблица сравнения характеристик</i>

Дата выдачи задания к разделу в соответствии с календарным учебным графиком	
--	--

Задание выдал консультант по разделу «Концепция стартап-проекта» (со-руководитель ВКР):

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Горюнова	К.Э.Н.		

	Наталья Николаевна			
--	-----------------------	--	--	--

Задание принял к исполнению обучающийся:

Группа	ФИО	Подпись	Дата
0В92	Макаревич Дарья Васильевна		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа		ФИО	
0В92		Макаревич Дарья Васильевна	
Школа	Инженерная школа ядерных технологий	Отделение (НОЦ)	Экспериментальной физики
Уровень образования	Бакалавриат	Направление/специальность	01.03.02 Прикладная математика и информатика

Тема ВКР:

Классификация изображений с помощью сверточных нейронных сетей	
Исходные данные к разделу «Социальная ответственность»:	
<p>Введение</p> <ul style="list-style-type: none"> – Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения. – Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации 	<p><i>Объект исследования: алгоритм распознавания объектов на изображении</i> <i>Область применения: автоматизация распознавания паттернов на изображении</i> <i>Рабочая зона: офис</i> <i>Размеры помещения: 6 м²</i> <i>Количество и наименование оборудования рабочей зоны: 1 персональный компьютер</i> <i>Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне: написание программного кода</i></p>
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p>1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения:</p> <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	<p>СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда"; Трудовой кодекс Российской Федерации от 30 декабря 2001 г. N 197-ФЗ (ТК РФ); ГОСТ 12.2.032-78 Рабочее место при выполнении работ сидя</p>
<p>2. Производственная безопасность при разработке проектного решения:</p> <ul style="list-style-type: none"> – Анализ выявленных вредных и опасных производственных факторов 	<p>Вредные факторы: Отклонение показателей микроклимата; Недостаточная освещённость рабочей зоны; Повышенный уровень шума на рабочем месте; Повышенная запыленность воздуха рабочей зоны; Опасные факторы: Опасность поражения электрическим током;</p>
<p>3. Экологическая безопасность при разработке проектного решения</p>	<p>Воздействие на селитебную зону: утилизация компьютеров и их составляющих Воздействие на литосферу: образование отходов во время утилизации компьютерной техники и ее составляющих</p>

	Воздействие на гидросферу: изготовление компьютерной техники Воздействие на атмосферу: выбросы углекислого газа при производстве компьютеров и его комплектующих
4. Безопасность в чрезвычайных ситуациях при разработке проектного решения	Возможные ЧС: землетрясение, пожар, затопление Наиболее типичная ЧС: пожар
Дата выдачи задания для раздела по линейному графику	

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Антоневич Ольга Алексеевна	к.б.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0B92	Макаревич Дарья Васильевна		

Реферат

Выпускная квалификационная работа содержит 61 страницу, 15 рисунков, 10 таблиц, 27 источников, 2 приложения. Ключевые слова: бинарная классификация, временные ряды, цепи Маркова, рекуррентные диаграммы, Грамианоугольные угловые поля.

Актуальность данной работы: разработка инновационных методов выявления закономерностей в графических изображениях временных рядов, классификация момента роста или падения на конец периода многомерных временных рядов с использованием модели машинного обучения.

В данной выпускной квалификационной работе (ВКР) изучается использование сверточных нейронных сетей для бинарной классификации изображений.

Объектом ВКР является распознавание объектов на изображении.

Предмет ВКР: выявление паттернов и классификация изображений, используя сверточные нейронные сети.

Целью работы является разработка математической модели и реализация программного кода для классификации объектов изображения.

В результате работы была разработана сверточная нейронная сеть, классифицирующая рост или падение, заданного периода времени.

Структура ВКР состоит из введения, анализа принципов работы сверточных нейронных сетей, анализа методов отображения временных рядов и конструирования признаков, алгоритма нейронной сети, анализа качества модели, стартап-части и социальной ответственности.

Исходными данными являются котировки компаний ZM, BIDU, TSLA, AMZN, MSFT, INTC, NVDA, BAC, PARA, NOC, а именно временные ряды, на основе которых были построены рекуррентные диаграммы, диаграммы переходных вероятностей цепей Маркова и грамианоугольные угловые поля. Для каждой компании бралось несколько периодов: с 31.05. 2014 по 30.01.2015, с 31.05. 2015 по 30.01.2016, с 31.05. 2016 по 30.01.2017, с 31.05. 2018 по 30.01.2019, с 31.05. 2019 по 30.01.2019, с 31.05. 2020 по 30.01.2021, с

31.05. 2021 по 30.01.2022, с 31.05. 2022 по 30.01.2023, из каждого периода бралось по 100 значений.

Для разработки программного модуля был использован язык программирования Python. Модель машинного обучения была написана на основе фреймворка Keras.

Оглавление

Введение.....	13
1. Теоретическая часть.....	15
1.1. Анализ методов распознавания объектов.....	15
1.2. Анализ принципа работы искусственной нейронной сети.....	15
1.3. Анализ методов обучения нейронных сетей.....	17
1.4. Описание рекуррентных диаграмм.....	17
1.5. Описание фрактальных временных рядов.....	18
1.6. Описание работы сверточной нейронной сети.....	19
2. Анализ методов отображения временных рядов в графические изображения.....	20
2.1. Грамианоугольное угловое поле.....	20
2.2. Цепи Маркова и переходные вероятности.....	23
2.3. Рекуррентные диаграммы.....	25
2.4. Анализ методов конструирования признаков.....	28
3. Разработка алгоритма и программная реализация для обучения нейронной сети.....	29
4. Тестирование и подготовка рекомендаций по использованию проекта...	31
5. Концепция стартап-проекта.....	38
5.1. Описание продукта как результата НИР.....	38
5.2. Интеллектуальная собственность.....	39
5.3. Объем и ёмкость рынка.....	39
5.4. Анализ современного состояния и перспектив развития отрасли.....	41
5.5. Планируемая стоимость продукта.....	42
5.6. Конкурентные преимущества создаваемого продукта и Сравнение технико-экономических характеристик продукта с отечественными и	

мировыми аналогами.....	43
5.7. Целевые сегменты потребителей.....	44
5.8. Бизнес-модель проекта, производственный план и план продаж.....	45
5.9. Стратегия продвижения продукта на рынок.....	46
6. Социальная ответственность.....	46
6.1. Правовые и организационные вопросы обеспечения безопасности....	47
6.2. Производственная безопасность.....	48
6.2.1. Отклонение показателей микроклимата.....	49
6.2.2. Недостаточная освещенность рабочей зоны.....	50
6.2.3. Повышенный уровень шума на рабочем месте.....	51
6.2.4. Опасность поражения электрическим током.....	53
Заключение.....	57
Список используемых источников.....	58
Приложения 1.....	62
Приложение 2.....	65

Определения, обозначения, сокращения

НС – нейронная сеть.

ПК – персональный компьютер.

Python – высокоуровневый язык программирования общего назначения.

Keras – фреймворк машинного обучения для языка Python.

Тренировочная выборка – набор данных, по которому производится настройка или оптимизация модели машинного обучения.

Тестовая выборка – набор данных, который полностью независим от тренировочной выборки, предназначенный для оценки качества работы модели.

Валидационная выборка – набор данных, по которому осуществляется выбор наилучшей модели из множества моделей, построенных по тренировочной выборке.

Аккуратность (Accuracy) – метрика, отражающая количество прогнозов, которые модель сделала правильно.

Полнота (Recall) – метрика, отражающая долю объектов положительного класса из всех объектов положительного класса нашел алгоритм.

Точность (Precision) – метрика, отражающая долю объектов, названных классификатором положительными и при этом действительно являющимися положительными

Матрица ошибок – это матрица размером $N \times N$, используемая для оценки эффективности модели классификации, где N - количество целевых классов.

Матрица ошибок сравнивает фактические целевые значения с предсказанными моделью машинного обучения.

Оценка F1 (F-мера) – это метрика, используемая для измерения точности модели классификации и являющаяся средним гармоническим значением precision и recall.

Введение

Финансовый рынок очень обширен. В его состав входит большое количество участников. Например, его участниками являются инвесторы - люди, продающие и покупающие акции различных компаний, эмитенты - компании, которые как раз-таки выпускают эти акции. Инвестор может действовать через посредников, то есть через брокеров или дилеров. Чтобы продать акции по более выгодной цене, необходимо знать, в какой момент времени лучше всего продавать акции. Для этого существует большое количество методов, начиная от обычного анализа новостей и заканчивая применением технологий машинного обучения.

Конечно, никто никогда не сможет со стопроцентной вероятностью сказать, в какой момент акции компании достигнут своего пика в цене, чтобы продать их с максимальной выгодой, так как очень на них влияет большое количество факторов, которые очень сложно предсказать. Но если условно отбросить какие-либо факторы, то математически можно предугадать рост или падение цен с большей вероятностью.

Существуют специалисты такие как, например, количественные аналитики, которые с помощью статистических методов и моделей анализа данных определяют тренды.

Целью выпускной квалификационной работой является разработка математической модели и реализация программного кода для классификации объектов изображения. Для ее достижения необходимо выполнить ряд задачи, такие как:

1. Провести анализ методов отображения временных рядов в графические изображения
2. Провести анализ методов конструирования признаков
3. Выбор метода прогнозирования многомерного временного ряда с использованием признаков графического изображения
4. Разработать алгоритм и программную реализацию для обучения нейронной сети

5. Провести тестирование и подготовить рекомендации по использованию проекта.

1. Теоретическая часть

1.1. Анализ методов распознавания объектов

В данной выпускной квалификационной работе необходимо реализовать алгоритм, способный относить изображения к определенному классу. Для того, чтобы классифицировать изображение, необходимо определить его местоположение на картинке, а далее определить его к определенному классу.

Существует немалое количество методов распознавания объектов на изображении. Например, поиск и анализ особых точек, наложение светофильтров, сравнение с эталоном - заданным шаблоном.

Одним из самых известных методов распознавания объектов является использование машинного обучения. Машинное обучение (ML) — это использование математических моделей данных, которые помогают компьютеру обучаться без непосредственных инструкций. Оно считается одной из форм искусственного интеллекта (ИИ) [1]. Далее речь пойдет о нейронных сетях - подразделе машинного обучения, а если быть конкретнее о сверточных нейронных сетях.

Говоря о достоинствах, можно отметить то, что они эффективно справляются с большим набором данных, имеют относительно высокую скорость выполнения, а также работоспособны к сдвигам и поворотам. Из минусов можно отметить - это то, что необходимо экспериментальным путем вычислять ее количество слоев, сложность обучения, а также вероятность переобучения модели.

В конечном итоге, были рассмотрены основные методы распознавания объектов на изображении.

1.2. Анализ принципа работы искусственной нейронной сети

Искусственная нейронная сеть (ИНС) — математическая модель, а также её программное 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 [2]. Искусственные нейронные сети способны решать огромное количество задачи, такие как, например, классификация, которая будет рассматриваться более подробно далее, регрессионный анализ, компьютерное зрение, кластеризации. Сама нейронная сеть представляет собой некий направленный граф, узлами которого являются нейроны, а дуги - связи с весовыми коэффициентами, все это представлено на рисунке 1.1, где первый слой - это входной слой, второй слой является скрытым слоем, а третий слой - это выходной слой.

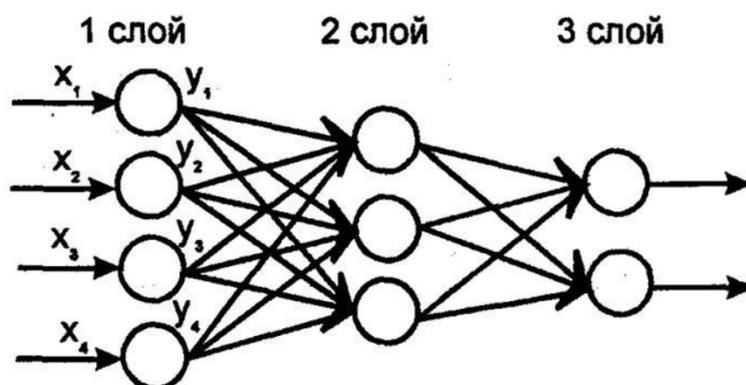


Рисунок 1.1 – Структура ИНС

Искусственная нейронная сеть не обязательно имеет три слоя, их может быть и больше. Если слоёв больше, то такие сети уже называются многослойными. Нейрон — это вычислительная единица, которая получает информацию, производит над ней простые вычисления и передает ее дальше другим нейронам [3]. Он имеет сигналы на входе и активационную единицу на выходе. Его вывод получается с помощью функции активации, с помощью которой обработка данных может производиться не только линейно. Самые популярные активационные функции:

- Сигмоида (Sigmoid)

$$f(x) = \frac{1}{1+e^{-x}}$$

- Гиперболический тангенс(Hyperbolic tangent)

$$f(x) = \frac{e^{ax} - e^{-ax}}{e^{ax} + e^{-ax}}$$

- Линейная(linear)

$$f(x) = kx$$

1.3. Анализ методов обучения нейронных сетей

Обучение нейронной сети — это процесс, в котором параметры нейронной сети настраиваются посредством моделирования среды, в которую эта сеть встроена [4]. Обучение нейронной сети может происходить с учителем или без учителя - это два разных алгоритма обучения. Если говорить о первом алгоритме, то он предполагает, что нейронной сети будет подаваться определенный набор данных, на основе которого нейронная сеть будет обучаться. На вход она получает данные, после чего он проходит обработку, а именно происходит вычисление выходного сигнала, далее полученные значения сравниваются со значениями, которые являются требованиями к выходу сети, после чего вычисляется ошибка, на основе которой корректируются веса. Когда ошибка становится достаточно малой, тогда обучение можно останавливать. Веса могут регулироваться как случайным образом — стохастическим методом, так и детерминированным, т.е. регулируются только конкретные веса.

Если говорить про вторую разновидность обучения нейронных сетей - алгоритмы без учителя, то здесь можно сказать о том, что у таких алгоритмов отсутствуют очевидные указания к его действиям, имеется лишь оценка всех его действий.

1.4. Описание рекуррентных диаграмм

В 1987 г. был предложен способ отображения m -мерной фазовой траектории состояний $x(t)$ на двумерную квадратную двоичную матрицу размером $N \times N$, в которой единицы, выступающие черными точками,

являются повторением состояния при некотором времени i в некоторое другое время j .

Рекуррентный анализ основывается на изображении фазовых портретов для анализа нелинейных динамических систем. Диаграмма рекуррентности показывает, когда временной ряд

$X = \{x(t_0), \dots, x(t_i), \dots, x(t_j), \dots\}$ повторно возвращается в ранее наблюдаемое состояние. Стоит отметить, что изображениями тоже можно описывать векторы. Для построения диаграммы рекуррентности используется матрица R с элементами:

$$R_{i,j}(\epsilon) = \Theta(\epsilon - \|x(t_i) - x(t_j)\|) = \{0, 1\}, \quad i, j = 1 \dots K, \quad (1)$$

где $\epsilon > 0$ – предопределенный порог,

$$\Theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}, \quad (2)$$

- ступенчатая функция Хевисайда

В конечном итоге, элемент матрицы R принимает значение 1 в том случае, если $x(t_i)$ и $x(t_j)$ отличаются на значение, которое меньше ϵ . Заметим, что квадратная матрица R может быть использована как матрица смежности для построения графа.

Цель рекуррентных диаграмм — анализ зрительных траекторий фазовых портретов. На основе рекуррентных диаграмм далее будут вычисляться паттерны, на основе которых будет обучаться сверточная нейронная сеть, создание которой будет описано в следующей главе.

1.5. Описание фрактальных временных рядов

Существует понятие фрактальной геометрии, которая асимметрична евклидовой геометрии, которая является более сложным понятием. Если последняя геометрия основывается на гладкости и идеальности, то фрактальная же, в свою очередь, берет за основу грубости и асимметрию. С помощью нее человечество может изображать окружающий мир более реалистично. Фрактальное время представляет собой очень

интересную концепцию. Все события могут сочетать в себе случайности и детерменизм, что почти может не укладываться в голове. Фрактальная размерность предполагает, что размерность объекта не обязательно может являться целым числом как в евклидовой геометрии, т.е. она описывает то, как объект заполняет пространство. Говоря о временном ряде, фрактальная размерность показывает то, насколько временной ряд является изрезанным. Увидеть это можно по следующей формуле:

$$N \cdot (2 \cdot r)^d = 1, \quad (3)$$

где N - число окружностей

r - радиус

d - фрактальная размерность

Отсюда можно, соответственно, вычислить фрактальную размерность (4), выразив ее из формулы (3) и преобразовав:

$$d = \frac{\log(N)}{\log(\frac{1}{2 \cdot r})}, \quad (4)$$

1.6. Описание работы сверточной нейронной сети

Идея архитектуры основывается на биологической зрительной системе. Первым слоем данной нейронной сети является сверточным. Он реализует операцию свертки. Ядром свертки является матрица, в которой элементами выступают весовые коэффициенты. В зависимости от типа изображения зависит количество матриц. Если изображение одно, то подается одна матрица, если цветное, то три. Ядро свертки в результате показывает число признаков, которые будут объединяться в один. В конечном итоге, после операции свертки образуется карта признаков, у которой выделены какие-то определенные признаки. В слоях свертки может быть ни одно ядро свертки. Чем их больше, тем больше информации об изображении можно получить.

Математически операцию свертки можно представить следующим образом:

$$g = f \cdot h, \quad (5)$$

где f - фрагмент изображения

h - матрица ядра

g - соответствующий элемент карты признаков

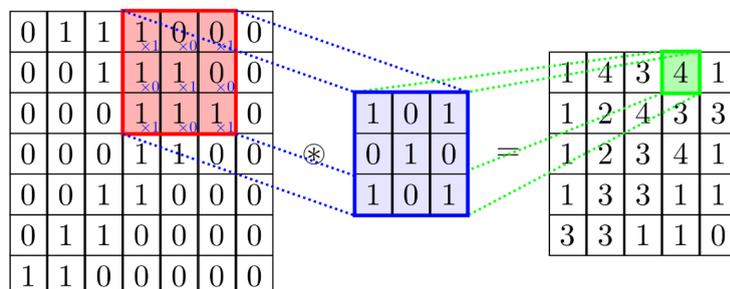


Рисунок 1.2 - Операция свертки

После сверточного слоя идет слой подвыборки. Заменяя некоторые области всего одним пикселем данный слой делает размерности карты признаков меньше. В итоге остаются только доминирующие признаки, вследствие чего уменьшается вычислительная нагрузка компьютера. Учитываются размер шага и размер ядра самой подвыборки.

Далее может идти слой свертки, затем опять слой подвыборки и т.д. Последним слоем выступают полносвязные слои, с помощью которых классифицируется изображение.

В данном разделе был выполнен анализ методов распознавания объектов, способов обучения нейронных сетей, разобраны рекуррентные диаграммы, которые будут в дальнейшем использоваться для обучения нейронных сетей, и принципы работы сверточных нейронных сетей.

2. Анализ методов отображения временных рядов в графические изображения

2.1. Грамианугольное угловое поле

Это один из способов отображения временных рядов. Его суть заключается в том, что он создает матрицу временных корреляций для каждого значения. Если вкратце, то его суть заключается в том, что в

диапазоне— $1 \leq a \leq b \leq 1$, далее, взяв \arccos , идет вычисление полярных координат масштабированных временным рядом, а затем уже считается \cos суммы углов для поля суммирования углов Грамиана (GASF) или синус разности углов для грамианского углов Разностное поле (GADF).

$$x_i = a + (b - a) \times \frac{x_i - \min(x)}{\max(x) - \min(x)}, \forall i \in \{1 \dots, n\}, (6)$$

$$\phi_i = \arccos(x'_i), \forall i \in \{1 \dots, n\}, (7)$$

$$GASF_{i,j} = \cos(\phi_i + \phi_j), \forall i, j \in \{1 \dots, n\}, (8)$$

$$GADF_{i,j} = \sin(\phi_i + \phi_j), \forall i, j \in \{1 \dots, n\}, (9)$$

Было построено Грамианоугольные угловые поля на языке Python, для начала рассмотрим для одномерного ряда, в качестве примера приведена компания ZM, в период с 31.05.2022 по 24.08.2022 (100 значений):

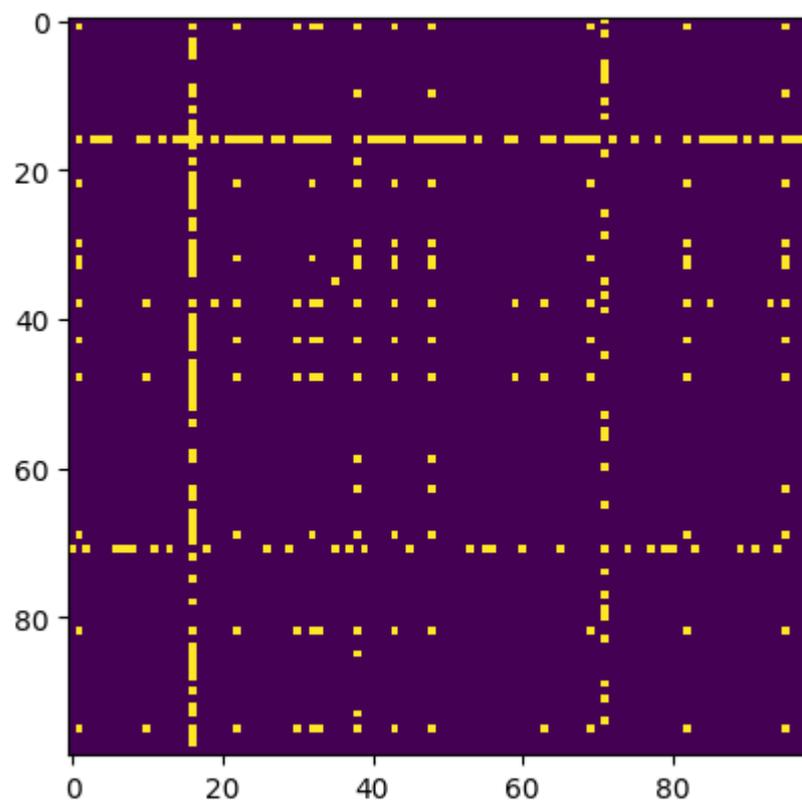


Рисунок 2.1 - Графианоугольное угловое поле для одномерного ряда: по осям отложены порядковые номера, соответствующие значениям одномерного временного ряда

Наличие ярко выраженных пиков на диаграмме может указывать на наличие выделяющихся угловых особенностей на поверхности. Отсутствие выраженных структур может указывать на более равномерное распределение угловых характеристик. Как можно заметить, на диаграмме присутствуют небольшие пики.

Теперь рассмотрим для многомерного ряда. Использовались три временных ряда компаний BIDU, ZM и COIN в период с 31.05.2022 по 24.08.2022 (100 значений):

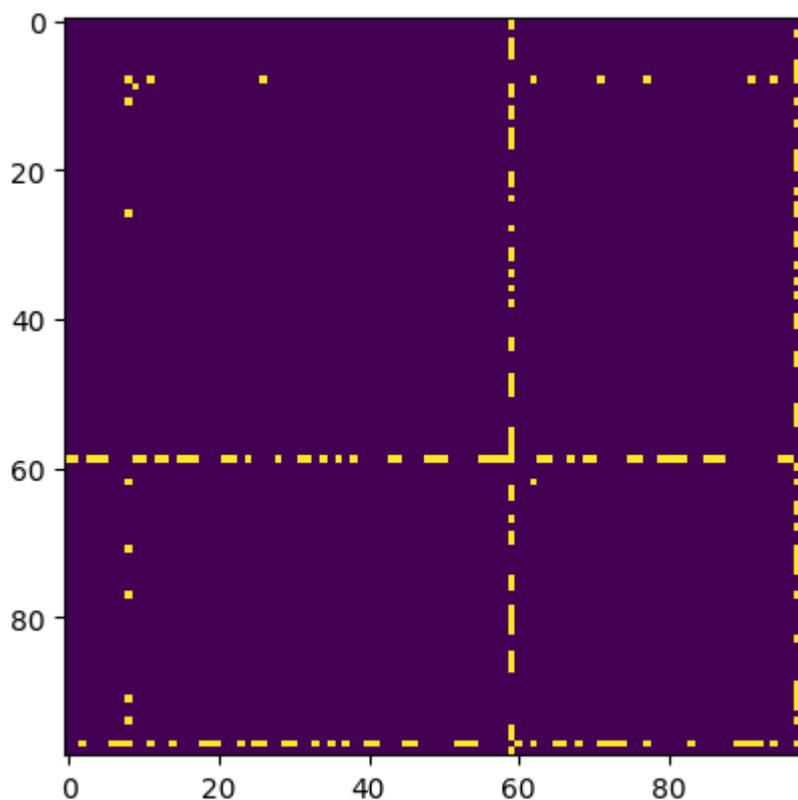


Рисунок 2.2 - Грамианугольное угловое поле для многомерного ряда: по осям отложены порядковые номера, соответствующие значениям многомерного временного ряда

Можно заметить, что пиков стало еще меньше, их практически нет, что свидетельствует об отсутствии выраженных паттернов, поэтому дальнейшее рассмотрение многомерных рядов грамианугольного поля в рамках данной работы не будет рассмотрено.

2.2. Цепи Маркова и переходные вероятности

Дискретизирует временные ряды в бункеры. Затем он вычисляет матрицу перехода Маркова дискретизированного временного ряда. Наконец, он распределяет матрицу перехода на поле, чтобы уменьшить потерю временной информации.

Диаграммы переходных вероятностей цепей Маркова представляют собой графическое изображение модели Маркова - математической модели,

которая описывает случайные процессы, в которых вероятность перехода из одного состояния в другое зависит только от текущего состояния (10).

$$P = [P_{ij}], \quad (10)$$

где:

P_{ij} - это вероятность перехода из состояния i в состояние j .

Каждый столбец матрицы должен иметь сумму элементов, равную 1.

Рассмотрим для начала диаграмму переходных вероятностей и цепей Маркова для одномерного ряда компании ZM, в период с 31.05.2022 по 24.08.2022 (100 значений):

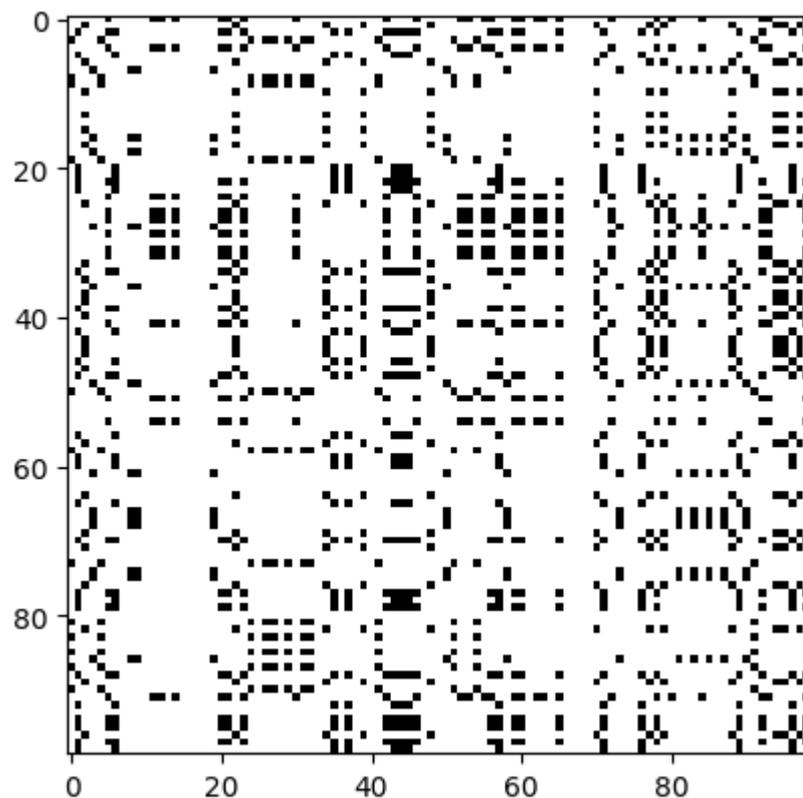


Рисунок 2.3 - Цепи Маркова и переходные вероятности для одномерного ряда: по осям отложены порядковые номера, соответствующие значениям одномерного временного ряда

Точки изображают всплески, наличие высокой вероятности перехода из одного состояния в другое.

Также была построена диаграмма переходных вероятностей цепей Маркова для многомерного ряда трех компаний BIDU, ZM и COIN в период с 31.05.2022 по 24.08.2022 (100 значений):

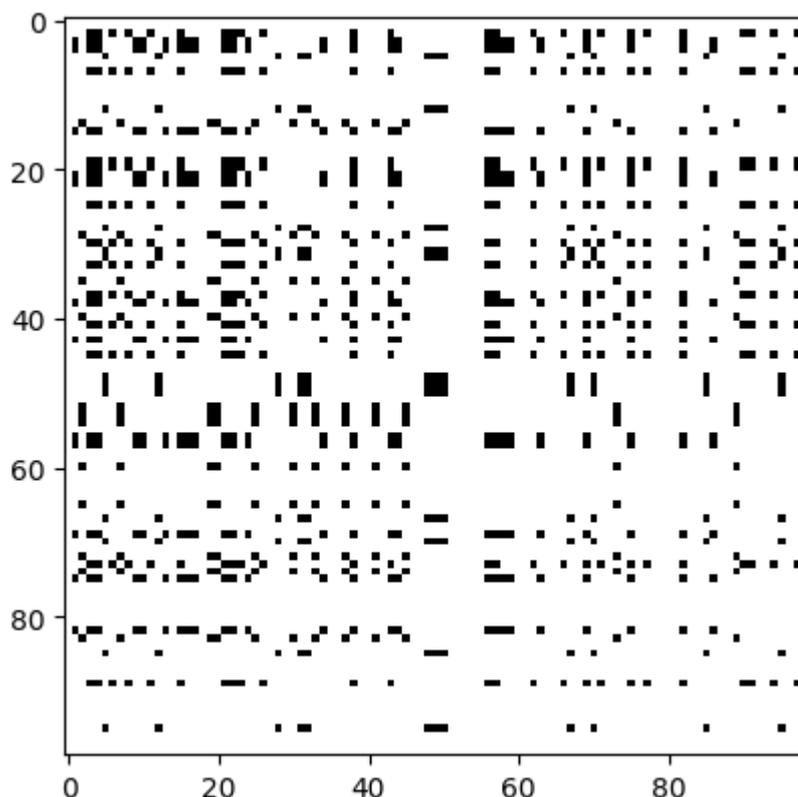


Рисунок 2.4 - Цепи Маркова и переходные вероятности для многомерного ряда: по осям отложены порядковые номера, соответствующие значениям многомерного временного ряда

Можно заметить, что, используя многомерность, количество находимых паттернов уменьшается. Было принято решение не использовать диаграммы цепей Маркова и переходных вероятностей, построенных с помощью многомерных рядов.

2.3. Рекуррентные диаграммы

Если говорить о рекуррентных диаграммах, то весь процесс заключается в том, что идет извлечение траекторий из временных рядов и вычисляет попарные расстояния между этими траекториями. Траектории определены как:

$$x_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}), \forall i \in \{1 \dots, n - (m - 1)\tau\}, (11)$$

где m - из траекторий и τ . График повторения, обозначаемый R , является бинаризованная попарная матрица расстояний между траекториями

$$R_{i,j} = \Theta(\varepsilon - \|x'_i - x'_j\|), \forall i, j \in \{1 \dots, n - (m - 1)\tau\}, (12)$$

Для выбора порога можно использовать различные стратегии, например, заданный поплавок или квантиль расстояний.

Было построена Рекуррентная диаграмма на языке Python в качестве примера приведена компания ZM, в период с 31.05.2022 по 24.08.2022 (100 значений):

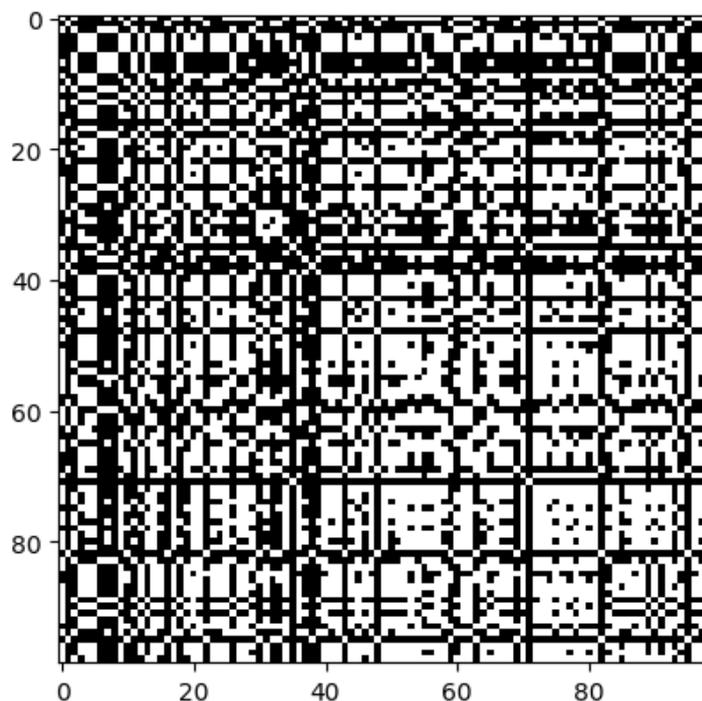


Рисунок 2.5 - Рекуррентная диаграмма для одномерного ряда: по осям отложены порядковые номера, соответствующие значениям одномерного временного ряда

Можно увидеть, что количество взаимосвязей в разы больше, каждая черная точка означает повторение состояния.

Теперь построим для многомерного ряда трех компаний VIDU, ZM и COIN в период с 31.05.2022 по 24.08.2022 (100 значений):



Рисунок 2.6 - Рекуррентная диаграмма многомерного временного ряда: по осям отложены порядковые номера, соответствующие значениям многомерного временного ряда

Цена на котировки либо растет, либо падает, так таковой стабильности обычно не бывает. Выше уже было сказано, что рекуррентные диаграммы представляют собой расстояние от i -ого элемента до j -ого. Если расстояние чересчур велико, то тогда значение не окрасится в черный цвет, иными словами прерывающаяся последовательность черных точек говорит нам о большой разнице в значениях (рост или падение цен). Сразу можно увидеть, что Рис. 2.5 не подходит, т.к. значения в основном повторяются и отследить рост или падение не получится. Колебание цен можно заметить в момент частой смены “черных” и “белых” точек, как это показано на Рис. 2.5.

Учитывание многомерности - процесс сложный, т.к. из всех построенных графиков было выявлено, что зависимостей между выбранными компаниями нет, поэтому в сеть не будут подаваться диаграммы, построенные для многомерных рядов.

2.4. Анализ методов конструирования признаков

Существует большое количество методов конструирования признаков для облегчения работы разработчиков нейронной сети. Каждому классу необходимо присваивать специальную метку, чтобы сеть понимала, к какому классу ее необходимо классифицировать. Выделение признаков помогает модели не переобучиться, а также улучшает ее производительность. Далее будут рассмотрены несколько популярных методов.

Само конструирование признаков предполагает выполнения нескольких этапов, а именно выбор признаков, их извлечение и преобразование.

Методы конструирования признаков - это способы создания новых признаков из имеющихся данных с целью улучшения качества моделей машинного обучения. Они актуальны тогда, когда имеющиеся признаки не позволяют достигнуть необходимого качества модели или когда количество признаков ограничено.

Одним из наиболее распространенных методов является создание новых признаков на основе существующих. Например, для предсказания стоимости жилья можно создать признак "цена за квадратный метр", разделив цену на площадь квартиры. Это поможет модели учитывать различия в размере и стоимости различных квартир.

Еще одним методом конструирования признаков является использование текстовых признаков. Например, в задаче определения тональности текста можно использовать набор ключевых слов, которые указывают на настроение автора. Эти ключевые слова могут быть извлечены из самого текста либо получены путем анализа отзывов на предмет включения в них слов, характеризующих настроение.

Еще один метод - использование сигнальных признаков. В задачах компьютерного зрения, например, можно использовать признаки, которые выявляют наличие элементов изображения, таких как края, текстуры или специфические образцы.

В рамках данной работы было принято решение не выполнять отдельное конструирование признаков, так как необходимая точность была достигнута.

3. Разработка алгоритма и программная реализация для обучения нейронной сети

Весь алгоритм реализован на языке Python с использованием библиотеки Keras.

Первым делом проводилась подготовка обучающей, тестовой выборок и валидационной выборок. Были созданы специальные функции, строящие рекуррентные диаграммы, цепи Маркова и переходные вероятности и грамианоугольные угловые поля. На вход подавались следующие аргументы: название компании, дата начала периода, месяц начала периода, год начала периода, дата конца периода, месяц конца периода, год конца периода, необязательный аргумент тип данных, означающий в какую папку складывать график (тренировочную, тестовую или валидационную), по умолчанию стояла тренировочная, а также номер графика. Внутри методов производилась загрузка данных с `finance.yahoo`, подготовка данных и само построение графика конкретно интересующего метода. Далее с помощью циклов создавались сами данные.

После всех вышеупомянутых манипуляций, была написана сама нейронная сеть, архитектура которой представлена чуть ниже. Функцией потерь была выбрана стандартная `binary_crossentropy` для бинарной классификации, оптимизатор `adam`.

Алгоритм нейронной сети:

1. Слой свертки, размер ядра 3x3, функция активации ReLU.
2. Слой подвыборки, выбор максимального значения из квадрата 2x2
3. Слой свертки, размер ядра 3x3, функция активации ReLU.

4. Слой подвыборки, выбор максимального значения из квадрата 2x2
5. Слой свертки, размер ядра 3x3, функция активации ReLU.
6. Слой подвыборки, выбор максимального значения из квадрата 2x2
7. Слой преобразования из двумерного в одномерное представление
8. Полносвязный слой, 64 нейрона, функция активации ReLU.
9. Слой Dropout.
10. Выходной слой, 1 нейрон, функция активации sigmoid

Графически представлен следующим образом:

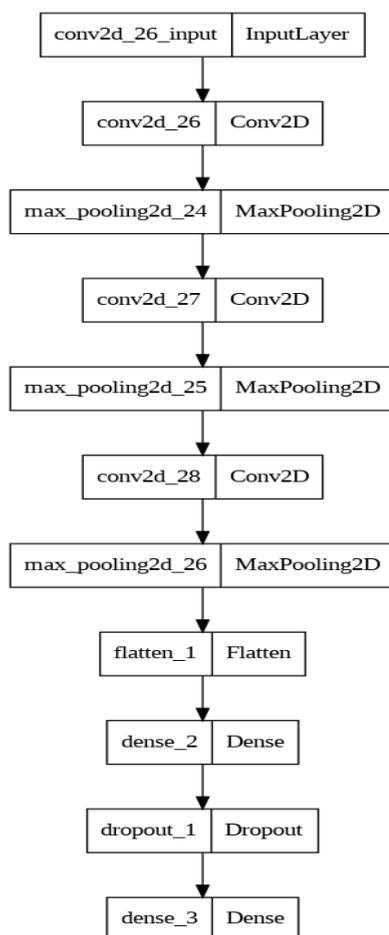


Рисунок 2.7 - Визуальное представление архитектуры нейронной сети

Слои, начиная с первого и заканчивая шестым, необходимы для выделения важных признаков в изображении, остальные слои для самой классификации.

4. Тестирование и подготовка рекомендаций по использованию проекта

Реализовав модель было проведено тестирование. В качестве оценки качества модели была выбрана аккуратность. Получились следующие результаты: для рекуррентных диаграмм аккуратность составляет 85.71%, для цепей Маркова и их переходных вероятностей равняется 77.5%, для грамианоугольного 72.1% углового поля. Использовалась абсолютно идентичная нейронная сеть для каждого метода. Исходя из аккуратности, можно сделать вывод, что рекуррентные диаграммы являются лучшим методом отображения временных рядов в качестве выборок для данной нейронной сети. Однако требуется дополнительно исследовать еще несколько метрик, так как данные являются несбалансированными, то есть количество объектов, находящихся в одном классе, могут отличаться от количества объектов, которые находятся в другом. Результаты не являются окончательными, можно попробовать добиться результатов лучше с более тонкой настройкой.

В рамках данной работы была также построена матрица ошибок:

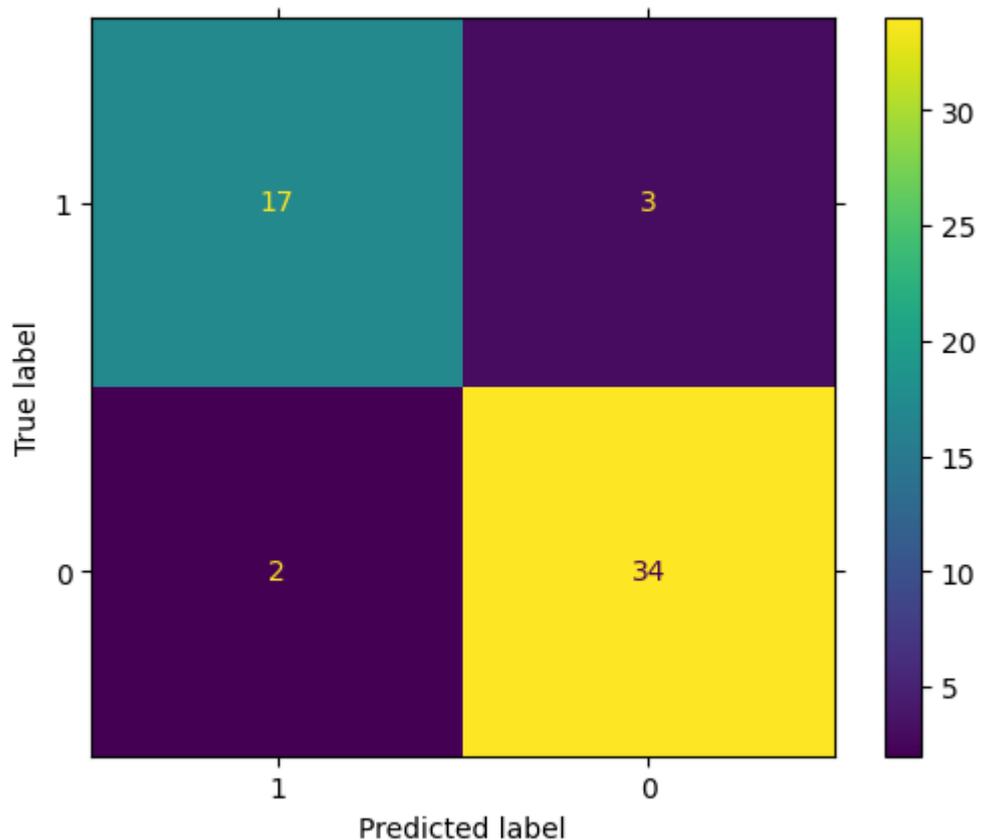


Рисунок 2.8 - Матрица ошибок НС, обученной на рекуррентных диаграммах: по осям расположены истинные и предсказанные классы

При анализе матрицы ошибок необходимо смотреть в первую очередь на главную диагональ, она показывает значения истинно положительных и истинно отрицательных значений. Можно увидеть, что истинно предсказанных изображений, которые классифицировались как рост значений, 17, а истинно отрицательных, то есть моменты падения, 34. Ложно истинных 2, а ложно отрицательных 3 из выборки, размер которой составляет 56 изображений.

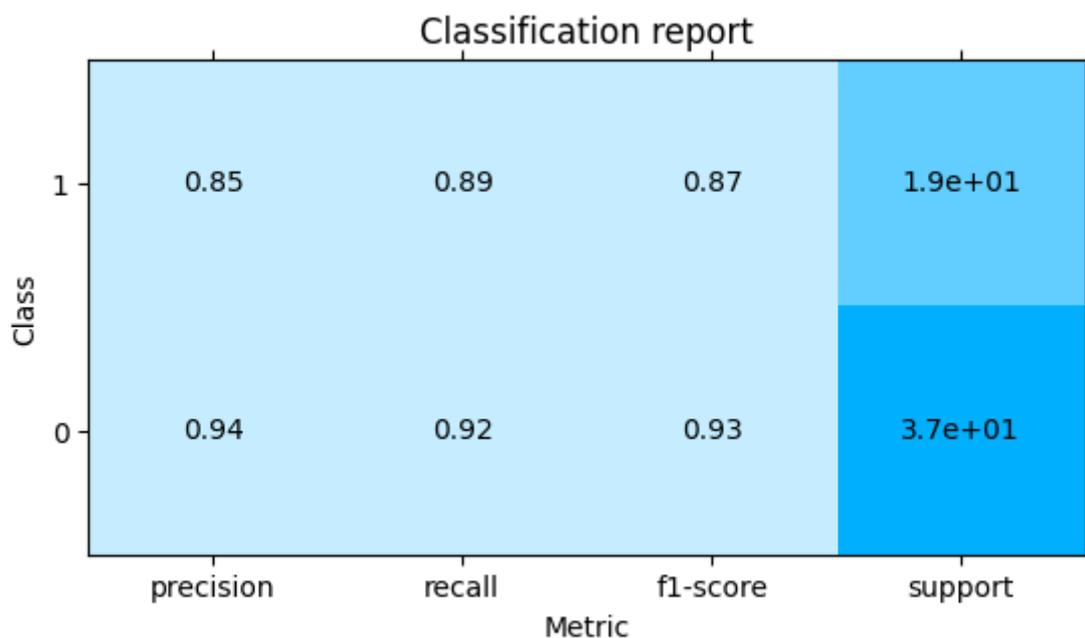


Рисунок 2.9 - График метрик НС, обученной на рекуррентных диаграммах: по осям расположены названия классов и метрик

Precision и recall не зависят от соотношении классов, поэтому такие метрики в нашем случае являются более подходящими, так как количество классов не сбалансировано. Precision показывает предсказанную долю истинных объектов, которые действительно являлись истинными. Recall, в свою очередь, показывает долю истинных предсказанных значений, относительно всех объектов. Но стоит обратить также внимание на f1-score. Эта метрика является мерой успешностью классификатора. Её расчет связан с precision и recall, поэтому данная метрика также подходит для несбалансированных данных. С помощью нее можно увидеть такие моменты, например, когда в обучающей выборки проблема с тренировочными данными, из-за чего обученная модель будет предсказывать значения только для одного класса. Чем ближе значение f1-score к единице, тем более точной является модель, и наоборот. Значения 0.87 и 0.93 находятся ближе к единице, что говорит о том, что модель близка к идеальной.

Далее представлены метрики, получившиеся в результате работы нейронной сети, обученной на изображениях диаграмм цепей Маркова и их переходных вероятностей:

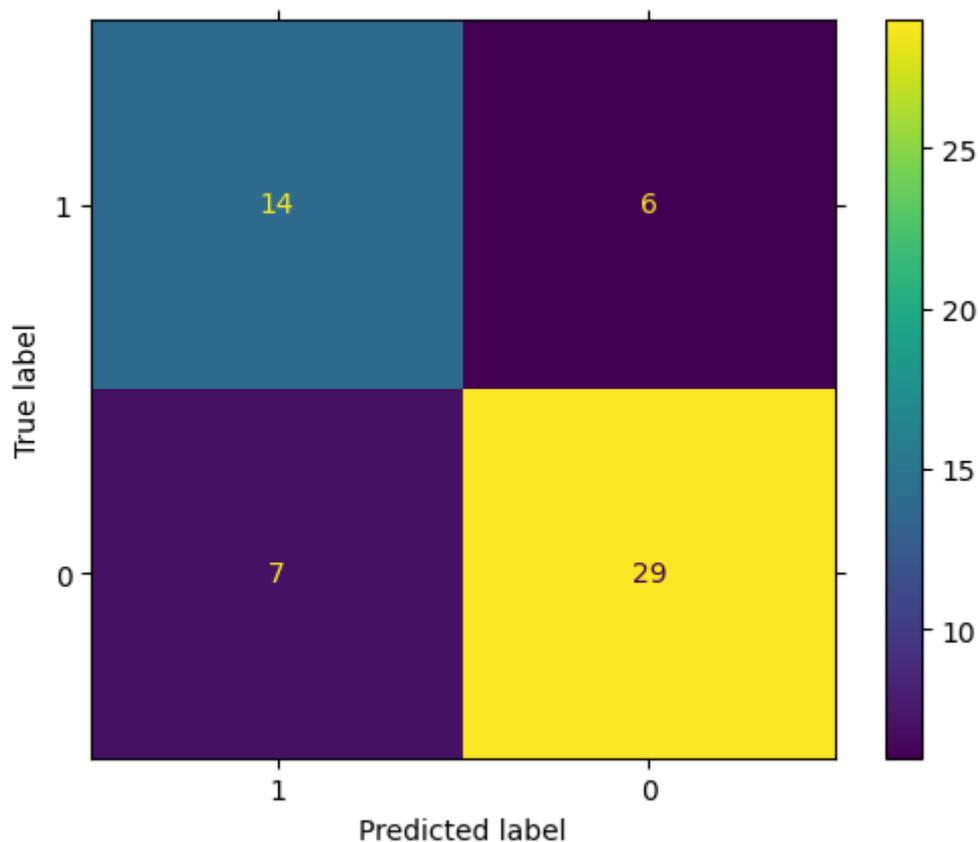


Рисунок 2.10 - Матрица ошибок НС, обученной на диаграммах цепей Маркова и переходных вероятностей: по осям расположены истинные и предсказанные классы

Заметно, что показатели стали чуть хуже, ошибочных классифицированных значений получилось 13 (разница с рекуррентными диаграммами составляет 8 изображений). Казалось бы, что это не такая большая разница, но необходимо взглянуть на полученные метрики качества модели перед тем, как делать окончательные выводы:

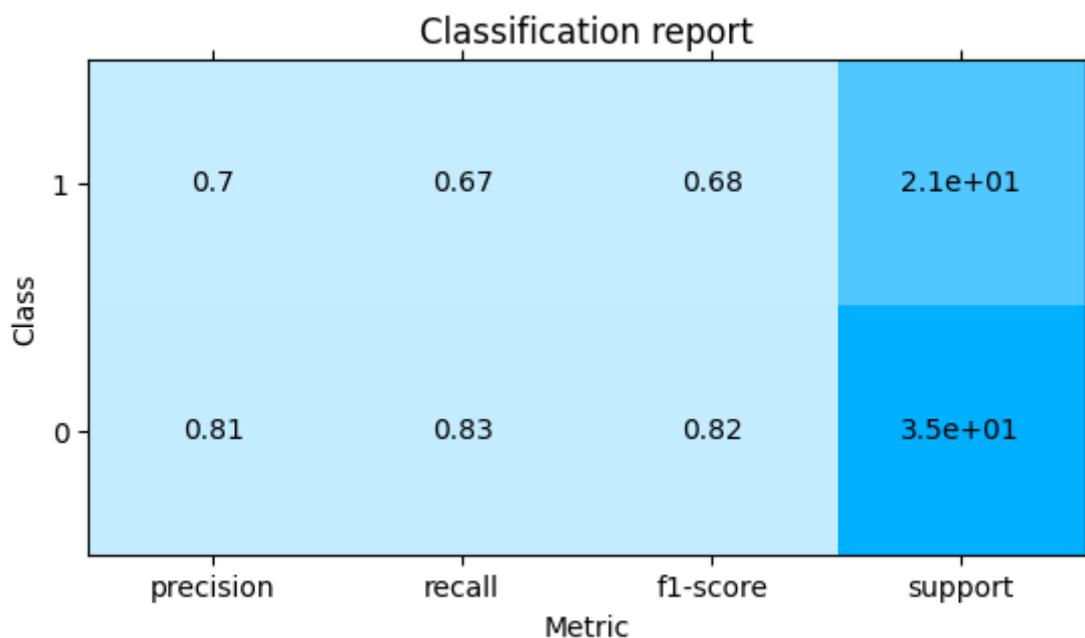


Рисунок 2.11 - График метрик НС, обученной на диаграммах цепей Маркова и переходных вероятностей: по осям расположены названия классов и метрик

F1-score для класса “0” хоть и стала на две сотых больше, но для класса “1” стала гораздо ниже, что свидетельствует о менее качественной классификации изображений в сравнении с классификацией изображений, основанной на рекуррентных диаграммах.

Для грамианогоугольного углового поля получились следующие результаты:

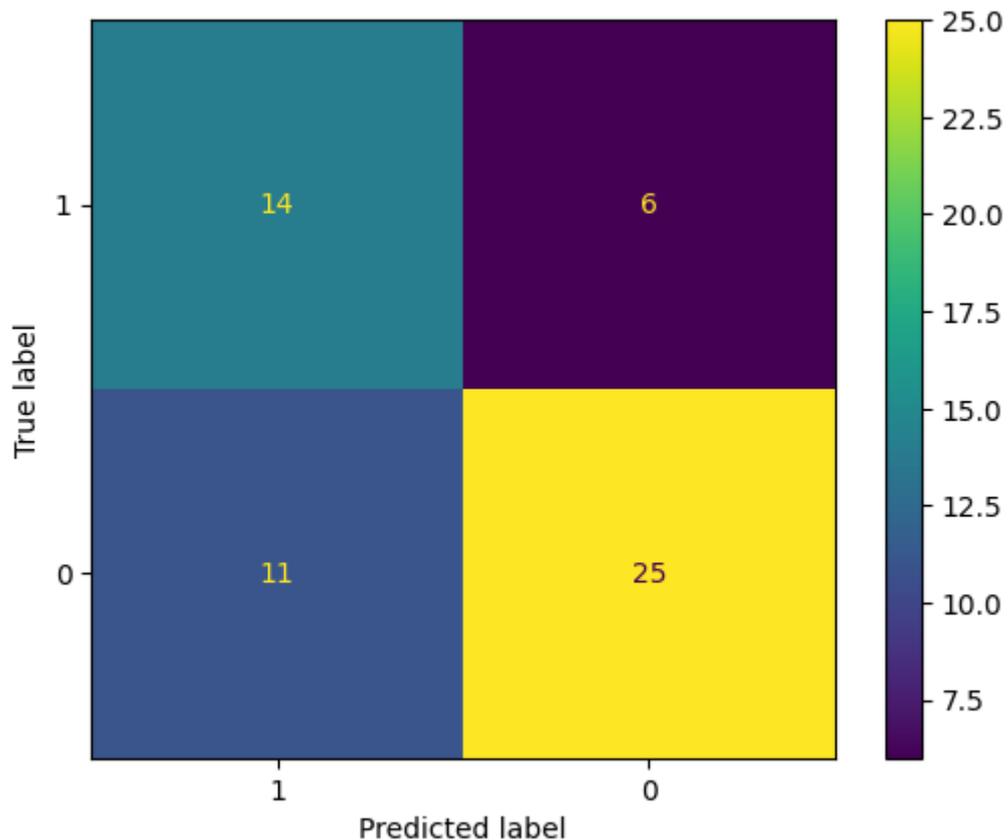


Рисунок 2.12 - Матрица ошибок НС, обученной на трапецианых угловых полях: по осям расположены истинные и предсказанные классы

На Рис. 2.12 можно увидеть, что истинно предсказанных изображений, которые классифицировались как рост значений, 14, а истинно отрицательных, то есть моменты падения, 25. Ложно истинных 11, а ложно отрицательных так же 6 из выборки, размер которой составляет 56 изображений. Если сравнивать с матрицей ошибок нейронной сети, обученной на рекуррентных диаграммах, то можно увидеть, что результаты однозначно хуже. Если суммарно нейронная сеть рекуррентных диаграмм неправильно классифицировала 5 изображений, то нейронная сеть трапецианого поля неправильно классифицировала 17 изображений, то есть почти в 3.5 больше неправильных значений.

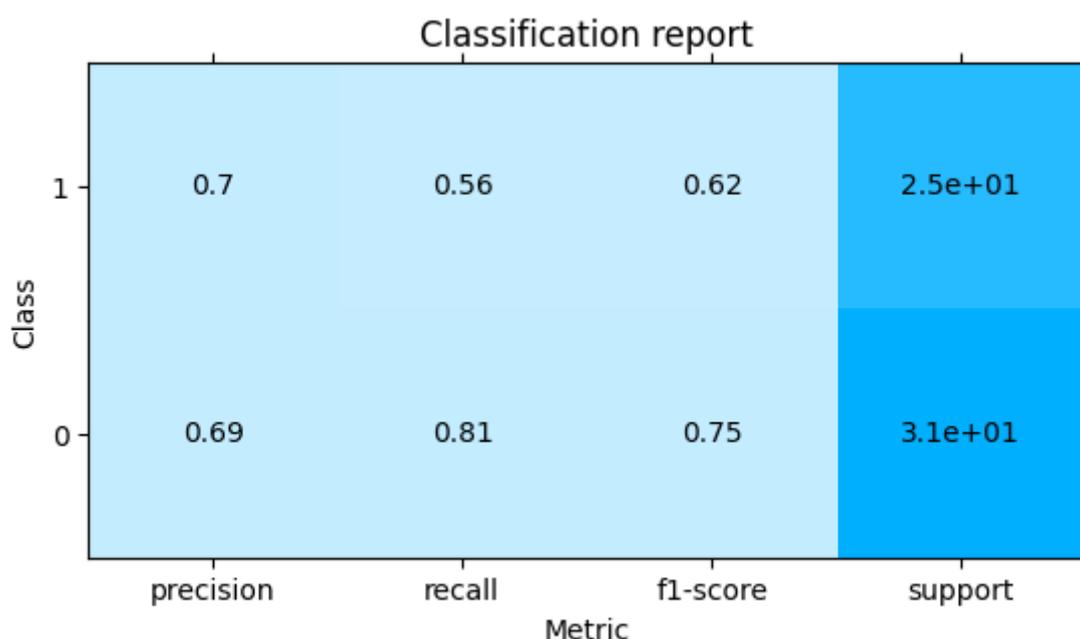


Рисунок 2.13 - График метрик НС, обученной на грамианугольных угловых полях: по осям расположены названия классов и метрик

Глядя на полученные метрики, можно увидеть, что F1-score для класса “1” ближе к 0.5, что говорит о том, что такая модель является скорее базовой и является далека от идеала. F1-score грамианугольного углового поля получились самыми низкими метриками по сравнению с другими методами, что свидетельствует о менее качественной модели. Из полученных результатов можно сделать вывод, что классификация изображений, основанная на рекуррентных диаграммах, является более качественной в рамках данной нейронной сети.

Исходя из полученных результатов, рекомендациями к использованию проекта являются более тонкая настройка проекта, создание еще большего набора данных для обучения и интерфейса для более удобной эксплуатации, также рекомендуется использовать сеть, которая обучена на рекуррентных диаграммах, так как такая сеть оказалась более эффективной, исходя из полученных метрик, описанных ранее.

5. Концепция стартап-проекта

5.1. Описание продукта как результата НИР

На сегодняшний день очень популярна тема предсказания роста и падения акций, некоторые люди хотят инвестировать в них свои деньги, ничего не делая, то есть получать пассивный доход.

В настоящий момент времени существует большое количество различных способов предсказания значений котировок, начиная от элементарного анализа новостей, заканчивая применением машинного обучения. Но стоит задумываться о создании и других способов предсказания.

Предсказание цен акций с помощью сверточной нейронной сети, обученной на рекуррентных диаграммах - новшество. Новые научные методы нужны всегда во всех областях для развития самой науки, сделать что-то новое, это уже является достоинством, но если говорить о данном стартапе, то стоит отметить, важным преимуществом рекуррентной диаграммы является то, что данный метод отображения временного ряда не чувствителен к нестационарности, т.е. не требует предварительной фильтрации, что позволяет ее сразу использовать для анализа. Данный метод может оказаться эффективнее и позволит очень просто предсказывать значения, лишь предварительно обучив сеть на свежих диаграммах. Сами же диаграммы также очень просто составляются, лишь вставив в код название акций и диапазон дат. В рамках данной работы нейронная сеть только находит эти моменты роста и падения, однако ее можно немного доработать, чтобы она сама предсказывала значения

В рамках данной работы написан программный код, который:

1. Составляет рекуррентные диаграммы;
2. Нейронная сеть, которая умеет классифицировать момент роста или падения на конец заданного периода времени

5.2. Интеллектуальная собственность

Объектом интеллектуальной собственности является программа для ЭВМ. Отношения, возникающие в связи с правовой охраной и использованием программ для ЭВМ, регулирует Гражданский кодекс РФ, часть 4, ст. 1261 и ст. 1262. Самостоятельно проанализировав данные на сайте fips.ru, похожих патентов не было выявлено.

Полностью запатентовать программное обеспечение невозможно, т.к. ПО не является изобретением согласно ст. 1350 п. 5 ГК РФ. Разрабатываемая программа состоит из алгоритма и исходного кода. Только алгоритм признается патентоспособным. Исходный код охраняется авторским правом. Ключевое различие заключается в том, что авторское право защищает автора от копирования программы, а патентное право защищает идею программы в целом [17].

Поиски по базе [fips](http://fips.ru) (Федерального института промышленной собственности) с использованием ключевых слов сверточная нейронная сеть, рекуррентная диаграмма показали отсутствие аналогичных действующих патентов.

Таким образом, для защиты интеллектуальной собственности принято решение о регистрации исходного кода (программы для ЭВМ), и патентование алгоритма.

5.3. Объём и ёмкость рынка

Проведём оценку рынка двумя методами: снизу-вверх и сверху-вниз. Для первого метода возьмём данные из Справочника организаций и предприятий России [15] о количестве финансовых предприятий и бирж, находящихся в СФО – таких компаний 295. Таким образом TAM составляет $295 \cdot 39\,000 = 11\,505\,000$ руб., исходя из стоимости разрабатываемого продукта (39 000 руб.).

Теперь оценим сколько из предприятий, попавших в TAM, имеют достаточно большую выручку и хороший рейтинг среди клиентов (выше 4), чтобы стартап оказался выгодным для них и чтобы понимать, что такой компании действительно можно верить. Исходя из описаний, таких предприятий около 45% (примерно 192 компании), поэтому SAM составляет $295 \cdot 39\,000 \cdot 0.45 = 7\,478\,250$ руб.

Исходя из географии расположения предприятий, примерно с 7.5% из них получится установить сотрудничество в течение года, таким образом SOM составляет $295 \cdot 39\,000 \cdot 0.45 \cdot 0.075 = 388\,293,75$ руб.

Проведём оценку методом сверху-вниз. Согласно анализу проекта Digital объем рынка искусственного интеллекта в 2023 году может превысить 500 млн долларов (39 985 000 000 рублей) [16]. Стоит учесть, что рынок искусственного интеллекта очень обширный. Доля решений, связанных с предсказанием цен акций, по данным наших конкурентов составляет 0.05%. Таким образом, TAM равен $39\,985\,000\,000 \cdot 0.0005 = 19\,992\,500$ руб.

Примерно 10% из TAM географически находятся в СФО, соответственно SAM составляет $39\,985\,000\,000 \cdot 0.0005 \cdot 0.1 = 1\,999\,250$ руб. При этом, три наших конкурента могут прийти и на рынок СФО, таким образом, SOM составляет $1\,999\,250 / (3 + 1) = 499\,812,5$ руб.

По результатам оценки рынка, наиболее оптимистичный прогноз получен методом сверху-вниз и объём рынка на год составляет 8 246 906,25 руб.

Таблица 1. – Сравнение двух методов оценки рынка

Снизу-вверх	Сверху вниз
TAM	

<p>Всего финансовых предприятий и бирж в СФО – 295 $295 * 39\ 000 = 11\ 505\ 000$</p>	<p>Прогнозируемый объём рынка всех решений в России области искусственного интеллекта на 2023 - 39 985 000 000. Приблизительная доля предсказания цен - 0.05%. $39\ 985\ 000\ 000 * 0,0005 = 19\ 992\ 500$</p>
SAM	
<p>Около 45% из предприятий соответствует критериям, чтобы решение было им выгодно. $295 * 0,45 * 39\ 000 = 7\ 748\ 250$</p>	<p>Приблизительная доля предприятий из СФО, готовых приобретать решения, связанных с предсказанием цен- 10%. $39\ 985\ 000\ 000 * 0,0005 * 0,1 = 1\ 999\ 250$</p>
SOM	
<p>Приблизительная доля предприятий, с которыми мы сможем установить сотрудничество за год, исходя из географии - 7.5%. $295 * 0,45 * 39\ 000 * 0,075 = 388\ 293,7$</p>	<p>Конкуренты, которые потенциально могут сделать предоставить аналогичную разработку для данных предприятий - 3 компании + моё, итого 4 компании на одном рынке в одной области. $1\ 999\ 250 / 4 = 499\ 812,5$</p>

В качестве оценки объема рынка используем оценку снизу-вверх, т.к. расчет таким методом имеет менее оптимистичный прогноз (разница составляет 111 518,8 рублей с оценкой сверху-вниз).

5.4. Анализ современного состояния и перспектив развития отрасли

По данным Центробанка (ЦБ) с начала 2019 г. по 2021 г. количество брокерских счетов россиян увеличилось приблизительно в шесть раз — до 14,2 млн, а число счетов по доверительному управлению капиталом поднялось почти в четыре раза — до 597 тыс [13].

В фондовый рынок сейчас инвестирует около трети экономически активного населения. По итогам 2022 года клиентская база брокеров выросла на 36% за счет новичков. Всего на брокерском обслуживании находится 22,9 млн человек. Объем активов граждан составил 6 трлн рублей, из них 5,3 трлн приходилось на ценные бумаги [14].

Существуют две основных методики прогнозирования цен акций. Фундаментальный анализ — в этом случае аналитики оценивают информацию, которая больше относится к компании, чьи акции торгуются на бирже, нежели к самим акциям. Решения о тех или иных действиях на рынке принимаются на основе анализа предыдущей деятельности компании, прогнозах выручки и прибыли и так далее. Вторая методика называется техническим анализом. В данном случае рассматривается поведение цены акций и выявляются его разнообразные паттерны (используется анализ временных рядов).

Машинное и глубокое обучение стали новой эффективной стратегией, которую для увеличения доходов используют многие инвестиционные фонды. Всё чаще аналитики используют именно данный подход.

Таким образом, отрасль прогнозирования цен значений акций имеет перспективы развития.

5.5. Планируемая стоимость продукта

Для определения планируемой стоимости годовой подписки на использование нейронной сети была рассчитана себестоимость по затратному методу ценообразования.

Написание нейронной сети требует привлечения специалиста в области машинного обучения. Стоимость услуг специалиста в области написания нейронной сети оценивается в 50 000 руб./месяц. Стоит еще взять за внимание, что необходимо за каждого сотрудника необходимо платить взносы на обязательное социальное страхование, которое равняется 30.2% (15 100 руб.). Таким образом, стоимость сотрудника с учетом взносов составляет 65 100 руб./месяц.

Разработка программного обеспечения предполагает дистанционную занятость и не предусматривает аренду офисного помещения, а также подразумевает наличие персонального компьютера у разработчика.

Длительность создания программного продукта для одного заказчика оценивается в 2 месяца. Данный срок обусловлен наличием базовой нейронной сети, разработанной автором в результате выполнения ВКР, что позволяет довольно быстро адаптировать нейронную сеть для прогнозирования. Расчёт себестоимости представлен в таблице X.

Таблица 2. – Себестоимость товара

Оборудование, тыс. руб	50, 000
Заработная плата и услуги сторонних организаций, тыс. руб	130, 200
Итого на 6 лицензий, тыс. руб	180, 200
Себестоимость единицы продукции, тыс. руб	30, 000

Примем прибыль от продажи годовой лицензии равной 30% в первый год пользования, тогда планируемая стоимость годовой лицензии составляет 39 000 рублей. Таким образом, выручка от продажи 6 лицензий без учёта налога на добавленную стоимость составит 234 000 рублей.

5.6. Конкурентные преимущества создаваемого продукта и Сравнение технико-экономических характеристик продукта с отечественными и мировыми аналогами

Для проведения конкурентного анализа были выбраны наиболее популярные коммерческие решения по предсказанию движению цен акций. Сведения о продуктах «NeuroTrend Crypto» от Олега Мазуренко, «Прогнозирующая машина» от Евгения Миронова и решении «StocksNeural» от Дмитрий Луковкина и Петра Пермякова были взяты из открытых источников.

В качестве наиболее важных критериев для сравнения были выделены следующие факторы: удобство в эксплуатации, индивидуализация, безопасность и повышение производительности труда пользователя. Удобство в эксплуатации важно для пользователей, которые далеки от программирования, так как иначе им придется потратить время на освоение

инструмента предсказания. Индивидуализация же важна на этапе разработки, потому что позволяет сократить время на адаптацию продукта под новые требования. Безопасность важна, поскольку никто не захочет, чтобы злоумышленники похитили его персональные данные. Повышение производительности труда важно, так как при использовании инструментов для прогнозирования аналитики и трейдеры затрачивают меньше времени для анализа движения цен, что и увеличивает производительность.

Таблица 3. – Конкурентный анализ

Решение	Разрабатываемое решение	Stocks Neural	Прогнозирующая машина	NeuroTrend Crypto
Критерий				
Удобство эксплуатации	-	+	+	+
Индивидуализация	+	+	+	+
Безопасность	+	+	+	+
Повышение производительности труда пользователя	+	+	+	+

Таким образом, согласно выделенным критериям все проекты оказались конкурентоспособными, а также после анализа конкурентов было выявлено, что необходимо будет в будущем разработать решение, предоставляющее удобное эксплуатирование проектом.

5.7. Целевые сегменты потребителей

Целевая аудитория данного стартап проекта - это трейдеры, количественные аналитики. Трейдерам очень важно знать, в какой момент времени лучше всего продавать акции. Чем цена выше, тем лучше, однако скачки в росте значений котировок очень сложно предугадать, поэтому данный подход будет очень им полезен. Количественные аналитики занимаются выявлением закономерностей цен акций, используя различные подходы, включая машинное обучение. Данная разработка может помочь

ускорить процесс.

Также данный стартап-проект может быть полезен обычным людям, которые занимаются инвестированием или хотят им начать заниматься.

Достаточно иметь обычный компьютер, сама программа не требовательна к его характеристикам.

5.8. Бизнес-модель проекта, производственный план и план продаж

Ниже представлено схематическое описание бизнес-модели проекта в виде модели А. Остервальда, представляющую собой схему из 9 блоков, описывающих различные бизнес-процессы проекта.

Таблица 4. – Бизнес-модель по А. Остервальду

Ключевые партнёры	Ключевые виды деятельности	Ценностные предложения	Взаимоотношения с клиентами	Потребительские сегменты
Молодые IT-компании, сотрудники в которых можно привлечь к разработке и модернизации продукта.	Разработка программного кода; Техническая поддержка.	Прогнозирование роста и падения значений цен котировок.	Персональная техническая поддержка.	Количественные аналитики, трейдеры, обычные пользователи.
	Ключевые ресурсы Финансы (з/п сотрудникам);		Каналы сбыта Основной канал сбыта: информационный – участие в форумах, конференциях.	
Структура издержек			Потоки поступления доходов	
Постоянные: заработная плата сотрудников тех. поддержки на начальном этапе реализации ПО.			Продажа годовой лицензии на продукт, продление лицензии.	
Переменные: заработная плата разработчика.				

5.9. Стратегия продвижения продукта на рынок

Целью продвижения проекта на рынок является привлечь не менее 6 заказчиков в первый год работы за счет посещения мероприятий (конференции, форумы и т.д.), которыми интересуются потенциальные клиенты, входящих в целевую аудиторию, а также посредством публикаций в тематических журналах. Приоритетными точками касания в данном случае являются личные встречи, звонки и переписки.

Определим то, как должен выглядеть путь клиента:

1. Увидел статью в журнале/выступление на конференции;
2. Сохранил контакты;
3. Позвонил для уточнения каких-либо деталей/оформления лицензии (или всё сразу);
4. Купил продукт.

Первый пункт действия клиента варьируется в зависимости от того, каким образом потенциальный клиент узнал о продукте, что определяется тем или иным способом продвижения. Третий пункт варьируется в зависимости от самого клиента, его уровня знаний в данной области, т.к. это напрямую влияет на то, будут ли у клиента вопросы. В будущем для сокращения пути клиента можно задуматься о создании сайта с интерфейсом (об интерфейсе уже говорилось ранее). Также стоит отметить, что после покупки продукта взаимодействие с клиентом может продолжиться, если он продлит лицензию на пользование ПО внесением оплаты через год.

6. Социальная ответственность

Объектом исследования данной ВКР является распознавание объектов на изображении, а именно определение роста или падения цен котировок, изображенные на рекуррентных диаграммах, Грамианоульных угловых полях и цепях Маркова и переходных вероятностей. Потенциально данный

алгоритм может помогать предсказывать значения цен на котировки различных компаний.

Работа выполняется на персональном компьютере (ПК), поэтому в данной главе представляется анализ опасных и вредных факторов при работе с ПК, влияние этих факторов на окружающую среду, а также меры по защите окружающей среды.

Предметом исследования выступает рабочая зона программиста, а именно ПК, компьютерный стол, клавиатура, компьютерная мышь и стул.

6.1. Правовые и организационные вопросы обеспечения безопасности

Написание алгоритма нейронной сети осуществляется за компьютерным столом, поэтому рабочее место должно удовлетворять требованиям ГОСТ 12.2.032-78 «Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя» [18] и СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда" [19]. Согласно ГОСТ 12.2.032-78 Конструкцией рабочего места должно быть обеспечено выполнение трудовых операций в пределах зоны досягаемости моторного поля, также должно быть обеспечено оптимальное положение работающего, при работе двумя руками органы управления размещают с таким расчетом, чтобы не было перекрещивания рук. В соответствии СП 2.2.3670-20 на рабочем месте, которое предназначено для работы в положении сидя, производственное оборудование и рабочие столы должны иметь пространство для размещения ног высотой не менее 600 мм, глубиной - не менее 450 мм на уровне колен и 600 мм на уровне стоп, шириной не менее 500 мм.

Согласно ст.100 ТК РФ «Рабочее время» [20] режим рабочего времени устанавливается правилами внутреннего трудового распорядка, которые утверждаются работодателем, учитывая мнение представительного органа работников. Также в данном документе регламентируется продолжительность

рабочего времени, которая не должна быть меньше указанного времени в трудовом договоре, но и не должна превышать 40 часов в неделю. Вдобавок ко всему законодательством предусмотрено установление в течение рабочего дня перерывов для питания и отдыха.

Во время разработки и написания алгоритма было предоставлено рабочее место, где соблюдены все требования по организации труда с ЭВМ.

6.2. Производственная безопасность

Во время написания программного алгоритма программист подвергается воздействию различных вредных и опасных факторов, которые представлены в таблице 5. Также в таблице приведены соответствующие нормативные документы и этапы работ, во время которых данные факторы могут повлиять на программиста.

Таблица 5. – Возможные опасные и вредные факторы

Факторы (ГОСТ 12.0.003-2015)	Нормативные документы
Отклонение показателей микроклимата	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» [21]
Недостаточная освещённость рабочей зоны	СП 52.13330.2016 «Естественное и искусственное освещение» [22]
Повышенный уровень шума на рабочем месте	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» [23]
Опасность поражения электрическим током	ГОСТ Р 58698-2019 «Защита от поражения электрическим током» [24]

Исходя из данной таблицы можно сделать вывод, что на программиста в ходе написания нейронной сети не воздействует химические и биологические факторы, а только физические и психологические факторы.

6.2.1. Отклонение показателей микроклимата

Отклонение показателей микроклимата во время рабочего процесса от нормы оказывает влияние на здоровье работника. Влияние недостаточной влажности выражается в испарении влаги со слизистых оболочек, что, в свою очередь, может вызвать пересыхание, растрескивание, а также заражение микробами. Низкая температура воздуха и повышенная скорость движения воздуха могут вызвать переохлаждение организма путем увеличения интенсивности теплообмена и процесса теплоотдачи во время испарения пота.

Нормы микроклимата производственных помещений регламентируются СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания», они регулируют оптимальные и допустимые значения показателей в рабочей зоне, соответствующие физиологическим потребностям организма человека, для создания комфортных и безопасных условий труда.

В течение написания алгоритма используется ПК, который может влиять на микроклимат путем снижения относительной влажности и повышению температуры в рабочем помещении. Рабочий процесс по энергозатратам относится к категории Ia (производится сидя). Для данной категории в таблицах 6 и 7 приводятся оптимальные и допустимые значения показателей микроклимата на рабочих местах.

Таблица 6. – Оптимальные величины показателей микроклимата на рабочих местах

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	22-24	21-25	60-40	0,1
Тёплый	23-25	22-26	60-40	0,1

Таблица 7. – Допустимые величины показателей микроклимата на рабочих местах

Период года	Температура воздуха, °С		Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с	
	диапазон ниже оптимальных величин	диапазон выше оптимальных величин			для диапазона температур воздуха ниже оптимальных величин, не более	для диапазона температур воздуха выше оптимальных величин, не более
Холодный	20,0 – 21,9	24,1 – 25,0	19,0 – 26,0	15 – 75	0,1	0,1
Тёплый	21,0 – 22,9	25,1 – 28,0	20,0 – 29,0	15 – 75	0,1	0,1

Если невозможно поддерживать допустимые величины локального микроклимата на рабочем месте, то в таком случае необходимо осуществлять мероприятия по защите работников от возможного перегревания и охлаждения. Например, регламентирование рабочего времени и времени, предоставляемого для отдыха, в помещении с микроклиматом, нормализующим тепловое состояние и уменьшение длительности рабочего дня.

6.2.2. Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны является вредным производственным фактором, который приводит к повышению утомляемости и снижению работоспособности человека на предприятии, также длительная работа в условиях низкой освещенности может привести к снижению остроты зрения.

Все нормы естественного, искусственного и совместного освещения регламентируются СП 52.13330.2016 «Естественное и искусственное освещение». Написание программного кода относится к категории работ высокой точности – Б (наименьший или эквивалентный объект различения

0,30 – 0,50 мм), подразряд 1 (относительная продолжительность зрительной работы при направлении зрения на рабочую поверхность не менее 70%).

В таблице 8 представлены требования к освещению рабочего помещения для разряда Б1.

Таблица 8. – Требования к освещению рабочего помещения

Искусственное освещение				Естественное освещение	
Освещенность на рабочей поверхности от системы общего освещения, лк	Цилиндрическая освещенность	Объединенный показатель дискомфорта, не более	Коэффициент пульсации освещенности, Кп, %, не более	Коэффициент естественной освещенности, %, при	
				верхнем или комбинированном	боковом
300	100	21	15	3	1

Повышенное глазное напряжение может быть вызвано ярким светом в зоне периферийного зрения. Для того, чтобы снизить влияние фактора, а именно недостаточную освещенность, необходимо, чтобы уровень естественного освещения рабочего пространства приблизительно совпадал с яркостью дисплея. Проблему недостаточной освещенности помещения можно решить с помощью установки дополнительных осветительных приборов.

6.2.3. Повышенный уровень шума на рабочем месте

Превышенный уровень шума на рабочем месте может создать психологический и физический стресс, который также может снизить производительность работника, его концентрацию и вызвать повышение уровня утомляемости. Данный фактор может быть вызван из-за шума, создаваемого работой ПК и кондиционером для воздуха.

Предельно допустимые показатели уровня звука, звукового давления регламентируются СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека

факторов среды обитания». Для разработчика нейронной сети данные показатели отображены в таблице 9. Интенсивный шум, превышающий значение, равное 80 дБА, при длительном воздействии может привести к полной или частичной потере слуха, так что такой максимальный уровень является критическим для разработчика. Шум приводит к быстрой утомляемости человека, что в свою очередь ведет к производственным ошибкам.

Таблица 9. – Предельно допустимые уровни звукового давления, уровни звука и эквивалентные уровни звука для инженера-программиста

Вид трудовой деятельности, рабочее место	Уровни звукового давления, дБ в октавных полосах со среднегеометрическими частотами, Гц							Уровни звука и эквивалентные уровни звука (дБА)
	31,5	63	125	250	500	1000	2000	
Творческая деятельность, руководящая работа с повышенными требованиями, научная деятельность, конструирование и проектирование, программирование, преподавание и обучение, врачебная деятельность. Рабочие места в помещениях дирекции, проектноконструкторских бюро, расчетчиков, программистов вычислительных машин, в лабораториях для теоретических работ и обработки данных,	86	71	61	54	49	45	42	50

приема больных в здравпунктах								
----------------------------------	--	--	--	--	--	--	--	--

Исходя из табличных значений, берем за норму уровень шума равный 50 дБА. При таком уровне шума всё отчётливо слышно. Для того, чтобы снизить уровень шума на рабочем месте необходимо для всего оборудования регулярно проводить техническое обслуживание, т.к. загрязнение, а именно пыль, может увеличить производимый шум.

6.2.4. Опасность поражения электрическим током

Когда речь идет об электробезопасности, то подразумевается система мероприятий, которая направлена на защиту людей от опасного воздействия электрического тока, статического электричества и электромагнитного поля. Значения вышеперечисленных факторов регулируются ГОСТ Р 58698-2019.

Таблица 10. – Пороги напряжения прикосновения для реагирования

Характер реагирования	Пороги напряжения, В
Реакция испуга	2 (переменный ток)
	8 (постоянный ток)
Мышечная реакция	20 (переменный ток)
	40 (постоянный ток)

Меры предосторожности для основной защиты от поражения электрическим током:

- использование защитных ограждений или оболочек;
- автоматическое отключение питания (защитное устройство, отключающее систему, которая питает электрическое оборудование, в случае возникновения замыкания);
- размещение опасных для жизни и здоровья человека участков электропроводов и приборов вне зоны досягаемости рукой;

- питание должно осуществляться от безопасного источника питания

6.3. Экологическая безопасность

Теперь необходимо проанализировать, какое влияние на окружающую среду оказывает написанная нейронная сеть и возможные источники загрязнения окружающей среды, которые могут возникнуть в ходе написания ВКР.

Сама нейронная сеть не создает вредные выбросы, которые могли бы повлиять на гидросферу, литосферу и атмосферу, повышения радиационного фона не было, но стоит отметить, что сам персональный компьютер оказывает влияние на окружающий воздух, а именно при его нагреве воздух вокруг него становится сухим. Также неправильная утилизация компьютера и его составляющих может привести к загрязнению литосферы, поэтому необходимо извлекать его комплектующие, их сортировать и отправлять на повторное использование для уменьшения вреда окружающей среде.

При соблюдении всех норм во время использования и утилизации компьютерной техники вредное воздействие на окружающую среду будет сокращаться.

Одна из главных проблем при защите окружающей среды в компьютерных лабораториях - утилизация люминесцентных ламп, которые в себе содержат ртуть в дозах от 1 до 70 мг, ядовитое вещество, относящееся к 1-ому классу опасности. Такое количество ядовитого вещества может причинить вред здоровью при повреждении лампы. Хранение и удаление отходов, а именно в данном случае люминесцентных ламп, происходят в соответствии с требованиями Постановления Правительства Российской Федерации от 28.12.2020 № 2314 «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь

причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [25]. Герметически стальной крышкой закрывают наполненную тару с отходами, если есть необходимость, то ее заваривают и передают по договору специализированным предприятиям, которые имеют на их утилизацию.

6.4. Безопасность в чрезвычайных ситуациях

6.4.1. Пожар

Пожар при работе с компьютером может случиться из-за разных причин, например, курение, короткое замыкание, сильный перегрев самого ПК.

Для того, чтобы избежать пожар, необходимо не курить, иметь средства пожаротушения в рабочем состоянии, своевременная диагностика и обслуживание техники. Само здание должно удовлетворять всем требованиям пожарной безопасности, а именно: наличие плана эвакуации, пожарной сигнализации, углекислотных огнетушителей с проверенным клеймом, эвакуационные выходы и, естественно, табличек, которые к ним направляют.

Согласно СП 12.13130.2009 [26] рабочее место разработчика программного обеспечения относится к категории «В», (пожароопасные), так как в данном помещении присутствует пыль, вещества и материалы, которые при взаимодействии с воздухом могут гореть.

В здании, в котором находится рабочее место, присутствуют огнетушители, есть противопожарная сигнализация, датчики дыма, имеются люди, которые ответственны за противопожарную безопасность, а также имеется план эвакуации.

Вывод по разделу

В ходе написания раздела «Социальная ответственность» были рассмотрены основные нормативные акты, регламентирующую безопасную работу за ПК, и проанализированы опасные и вредные факторы, которые могут возникнуть в ходе рабочего процесса. В ходе написания ВКР все меры безопасности были соблюдены, вредное воздействие на окружающую среду

сведены к минимуму, а также описаны меры безопасности для наиболее типичной чрезвычайной ситуации.

В ходе написания раздела «Социальная ответственность» были рассмотрены правовые и организационные вопросы безопасности на рабочем месте, которое должно удовлетворять всем требованиям безопасности. Требуемое освещение обеспечивается за счет нескольких энергосберегающих ламп. Уровень шума находится в допустимом диапазоне. Микроклиматические условия соблюдаются за счет системы отопления в холодное время и проветривания в теплое время. Защита от воздействия электрического тока обеспечивается путем проверки состояния ПК и соблюдения правил безопасности при работе с ним. Рабочее место имеет все необходимые средства для обеспечения пожарной безопасности.

В соответствии ПУЭ, рабочее место относится к 1 категории по электробезопасности: помещение без повышенной опасности. Для работы в офисе персоналу присваивается 1 группа по электробезопасности.

В соответствии СанПиН 1.2.3685-21 работа разработчика относится к категории 1б, так как разработчик проводит большую часть времени за компьютером, передвигается из кабинета в кабинет или по коридору.

В соответствии СП 12.13130.2009 рабочее место разработчика относится к категории «В», (пожароопасные), потому что в данном помещении присутствует вещества и материалы, которые способны при взаимодействии с воздухом гореть.

В соответствии Постановлению от 31 декабря 2020 года №2398 [27] категория объекта, в текущем случае ПЭВМ, оказывающего негативное воздействие на окружающую среду, относится к IV категории.

Вдобавок ко всему были рассмотрены экологическая безопасность, уменьшение вредного влияния на окружающую среду и человека, а также безопасность в чрезвычайных ситуациях и меры по их предупреждению.

Заключение

Проведен анализ методов отображения признаков и конструирования признаков, был выбран метод рекуррентных диаграмм в качестве основного метода отображения признаков, на основе которого шли дальнейшие сравнения. Созданы обучающие, тестовые и валидационные выборки для трех методов отображения признаков, произведена разметка данных, разработан алгоритм сверточной нейронной сети и программно реализован.

На основе матриц ошибок и таблице с метриками качества, в состав которой входили такие метрики как precision, recall и f1-score. Самые лучшие показатели оценки качества модели получились у сети, обученной на рекуррентных диаграммах, а самые худшие у грамианоугольного углового поля. Аккуратность сети для рекуррентных диаграмм составила 85.71%, для цепей Маркова и их переходных вероятностей составила 77.5%, для грамианоугольного 72.1% углового поля. F1-score для рекуррентных диаграмм составила 0.87 и 0.93, для цепей Маркова и их переходных вероятностей 0.68 и 0.82, для грамианоугольного углового поля 0.62 и 0.75.

В качестве рекомендаций было выдвинуто использовать сеть, обученную на рекуррентных диаграммах, также реализация интерфейса для более удобного использования.

Список используемых источников

1. Что такое машинное обучение? // Azure URL:
<https://azure.microsoft.com/ru-ru/resources/cloud-computing-dictionary/what-is-machine-learning-platform> (дата обращения: 23.03.2023).
2. Искусственная нейронная сеть // VI Международный конкурс научно-исследовательских и творческих работ учащихся Старт в науке URL: <https://school-science.ru/6/4/37944> (дата обращения: 26.03.2023).
3. Нейросеть и Искусственный интеллект // Всемирный Клубъ-Музей-Лекторий "Маски, Лики, Фигуры и энергоартефакты мира" URL:
<https://mif-mira.ru/akademicheskie-sredy/post/nejroset-i-iskusstvennyj-intellect> (дата обращения: 28.03.2023).
4. Обучение нейронной сети // NEURONUS.com URL:
<https://neuronus.com/theory/nn/238-obucheniya-nejronnoi-seti.html> (дата обращения: 28.03.2023).
5. Imaging time series // pyts URL:
<https://pyts.readthedocs.io/en/stable/modules/image.html#gramian-angular-field> (дата обращения: 05.03.2023).
6. jason887 /
Using-Deep-Learning-Neural-Networks-and-Candlestick-Chart-Representation-to-Predict-Stock-Market // github URL:
<https://github.com/jason887/Using-Deep-Learning-Neural-Networks-and-Candlestick-Chart-Representation-to-Predict-Stock-Market> (дата обращения: 17.03.2023).
7. HOW TO GET CONFUSION MATRIX AND ROC CURVE WHEN USING FLOW_FROM_DIRECTORY? // AppsLoveWorld URL:
<https://www.appsloveworld.com/machine-learning/32/how-to-get-confusion-matrix-and-roc-curve-when-using-flow-from-directory> (дата обращения: 02.05.2023).

8. Choosing Colormaps in Matplotlib // matplotlib URL:
<https://matplotlib.org/stable/tutorials/colors/colormaps.html> (дата обращения: 12.04.2023).
9. tf.keras.preprocessing.image.ImageDataGenerator // TensorFlow URL:
https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator#flow_from_directory (дата обращения: 19.04.2023).
10. Матрица ошибок (Confusion Matrix) // TensorFlow URL:
<https://www.helenkapatsa.ru/matritsa-oshibok/> (дата обращения: 20.05.2023).
11. Расшифровка матрицы путаницы // MachineLearningMastery.ru URL:
<https://machinelearningmastery.ru/decoding-the-confusion-matrix-bb4801decbb/> (дата обращения: 21.05.2023).
12. Гафаров Ф.М., Галимянов А.Ф. ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ И ИХ ПРИЛОЖЕНИЯ. - УДК 004.032.26 изд. - Казань: Издательство Казанского университета, 2018. - 121 с.
13. ЦБ РФ: россияне всё больше интересуются торговлей ценными бумагами // Открытый журнал URL:
<https://journal.open-broker.ru/novosti/rossiyane-vse-bolshe-interesuyutsya-torgovlej-cennymi-bumagami/> (дата обращения: 27.05.2023).
14. Более трети экономически активного населения имеет брокерские счета: итоги 2022 года // Банк России URL:
<https://cbr.ru/press/event/?id=14602> (дата обращения: 27.05.2023).
15. Справочник организаций и предприятий России [Электронный ресурс]. – Режим доступа: <https://sfo.spr.ru/>, свободный (дата обращения 27.05.2023)
16. Тренды развития искусственного интеллекта в России и темпы его роста в России и мире максимально сблизились // Специальный проект DIGITAL URL:
<https://rg.ru/2023/04/17/trendy-razvitiia-iskusstvennogo-intellekta-i-tempy-ego-rosta-v-rossii-i-mire-maksimalno-sblizilis.html> (дата обращения:

27.05.2023).

17. Мухаметгалиева К. А. Патентование программного обеспечения / К. А. Мухаметгалиева // Правовая защита, экономика и управление интеллектуальной собственностью : материалы всероссийской научно-практической конференции, Екатеринбург, 21 апреля 2015 г. — Екатеринбург : [УрФУ], 2015. — Т. 1 — С. 84-91.
18. ГОСТ 12.2.032-78 Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя. Общие эргономические требования // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200003913> (дата обращения: 16.05.2023).
19. СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда" // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573230583> (дата обращения: 16.05.2023).
20. Трудовой кодекс (ТК РФ) «Рабочее время» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <http://base.garant.ru/12125268/> (дата обращения: 16.05.2023).
21. СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2021. – Режим доступа: <https://docs.cntd.ru/document/573500115> (дата обращения: 19.05.2023).
22. СП 52.13330.2016 «Естественное и искусственное освещение» // Электронный фонд правовой и нормативно-технической документации

- [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/456054197> (дата обращения: 16.05.2023).
23. СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573500115> (дата обращения: 19.05.2023).
24. ГОСТ Р 58698-2019 «Защита от поражения электрическим током» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200170001> (дата обращения: 16.05.2023).
25. Постановление Правительства Российской Федерации от 28.12.2020 № 2314 «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573275660> (дата обращения: 19.05.2023).
26. СП 12.13130.2009 «Системы противопожарной защиты. Эвакуационные пути и выходы» [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200071156> (дата обращения: 19.05.2023).
27. Постановлению от 31 декабря 2020 года №2398 «Об утверждении критериев отнесения объектов, оказывающих негативное воздействие на окружающую среду, к объектам I, II, III и IV категорий» [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573292854> (дата обращения: 19.05.2023).

Приложения 1

Код методов, использованных для создания тестовой, валидационной и обучающих выборок

```
import yfinance as yahooFinance
from pyts.datasets import load_gunpoint
from mpl_toolkits.axes_grid1 import ImageGrid
from pyts.image import MarkovTransitionField, GramianAngularField, RecurrencePlot
import requests
import datetime
import matplotlib.pyplot as plt
import numpy as np
import apimoeX
import pandas as pd
from scipy.signal import argrelextrema
from scipy.spatial.distance import pdist, squareform
# Check the root folder
!!ls /content/drive/
# Path should be visually same as in drive.google.com
path_to_train = '/content/drive/MyDrive/TRAIN_DATA'
path_to_test = '/content/drive/MyDrive/TEST_DATA'
path_train_Markov = '/content/drive/MyDrive/Markov/TRAIN_DATA'
path_test_Markov = '/content/drive/MyDrive/Markov/TEST_DATA'
path_train_Gramian = '/content/drive/MyDrive/Gramian/TRAIN_DATA'
path_test_Gramian = '/content/drive/MyDrive/Gramian/TEST_DATA'
d_start = 31
m_start = 5
y_start = [2013,2014,2015,2016,2017,2018,2019,2020,2021,2022]
d_end = 30
m_end = 1
y_end = [2014,2015,2016,2017,2018,2019,2020,2021,2022,2023]
def get_class_name(start_price, last_price):
    if start_price < last_price:
        return 1
    else:
        return 0
def Markov_plot(ticker, DAY_START, MONTH_START, YEAR_START,
```

```

DAY_END, MONTH_END, YEAR_END, number, type_data = 'TRAIN_DATA'):
# startDate , as per our convenience we can modify
startDate = datetime.datetime(YEAR_START, MONTH_START, DAY_START)
transformer_markov = MarkovTransitionField()
# endDate , as per our convenience we can modify
endDate = datetime.datetime(YEAR_END, MONTH_END, DAY_END)
GetInformation = yahooFinance.Ticker(f"{ticker}")
# pass the parameters as the taken dates for start and end
dataS = GetInformation.history(start=startDate,
end=endDate)
dfS = pd.DataFrame(dataS)
prices = dfS['Close'][:100].to_numpy()
normalized_prices = np.log(prices[1:]/prices[:-1])
Markov = transformer_markov.transform([normalized_prices])
plt.imshow(X=np.int64(Markov[0] > 0.249), cmap='binary')
NAME = get_class_name(start_price=prices[0], last_price=prices[-1])
if NAME == 0:
plt.savefig(f'/content/drive/MyDrive/Markov/{type_data}/0/{NAME}. {number}.png')
else:
plt.savefig(f'/content/drive/MyDrive/Markov/{type_data}/1/{NAME}. {number}.png')
def Gramian_plot(ticker, DAY_START, MONTH_START, YEAR_START,
DAY_END, MONTH_END, YEAR_END, number, type_data = 'TRAIN_DATA'):
# startDate , as per our convenience we can modify
startDate = datetime.datetime(YEAR_START, MONTH_START, DAY_START)
transformer_gramian = GramianAngularField()
# endDate , as per our convenience we can modify
endDate = datetime.datetime(YEAR_END, MONTH_END, DAY_END)
GetInformation = yahooFinance.Ticker(f"{ticker}")
# pass the parameters as the taken dates for start and end
dataS = GetInformation.history(start=startDate,
end=endDate)
dfS = pd.DataFrame(dataS)
prices = dfS['Close'][:100].to_numpy()
normalized_prices = np.log(prices[1:]/prices[:-1])
Gramian = transformer_gramian.transform([normalized_prices])
plt.imshow(X=np.int64(Gramian[0] > 0.03))
NAME = get_class_name(start_price=prices[0], last_price=prices[-1])
if NAME == 0:

```

```

plt.savefig(f'/content/drive/MyDrive/Gramian/{type_data}/0/{NAME}.{number}.png')
else:
plt.savefig(f'/content/drive/MyDrive/Gramian/{type_data}/1/{NAME}.{number}.png')
def recur_plot(ticker, DAY_START, MONTH_START, YEAR_START,
DAY_END, MONTH_END, YEAR_END, number, type_data = 'TRAIN_DATA'):
# startDate , as per our convenience we can modify
startDate = datetime.datetime(YEAR_START, MONTH_START, DAY_START)
# endDate , as per our convenience we can modify
endDate = datetime.datetime(YEAR_END, MONTH_END, DAY_END)
GetInformation = yahooFinance.Ticker(f'{ticker}')
# pass the parameters as the taken dates for start and end
dataS = GetInformation.history(start=startDate,
end=endDate)
dfS = pd.DataFrame(dataS)
prices = dfS['Close'][:100].to_numpy()
normalized_prices = np.log(prices[1:]/prices[:-1])
REC_PLOT = RecurrencePlot().transform([normalized_prices])
NAME = get_class_name(start_price=prices[0], last_price=prices[-1])
plt.imshow(X=np.int64(REC_PLOT[0] > 0.03),cmap='binary')
if NAME == 0:
plt.savefig(f'/content/drive/MyDrive/{type_data}/0/{NAME}.{number}.png')
else:
plt.savefig(f'/content/drive/MyDrive/{type_data}/1/{NAME}.{number}.png')

```

Приложение 2

Код сверточной нейронной сети

```
from keras.models import Sequential
import scipy
import matplotlib.pyplot as plt
from sklearn_evaluation import plot
import numpy as np
from keras.models import Model
from keras.utils.vis_utils import plot_model
import dataset
from keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.layers import Conv2D, MaxPooling2D
from tensorflow.keras.layers import Dense, Flatten, Dropout, Activation, Input
from sklearn.metrics import classification_report, confusion_matrix, ConfusionMatrixDisplay
source_train_dir = '/content/drive/MyDrive/TRAIN_DATA'
source_test_dir = '/content/drive/MyDrive/TEST_DATA'
source_valid_dir = '/content/drive/MyDrive/VALID_DATA'
test_dir = 'test' # Каталог с набором данных для тестирования
train_dir = 'train' # Каталог с данными для обучения
val_dir = 'valid' # Каталог с данными для проверки
dir_name = 'data'
model_outh = '/content/drive/MyDrive/NETWORKS/plot_model/model_plot.png'
nb_images = 64
img_width, img_height = 416, 416
input_shape = (img_width, img_height, 3)
epochs = 30
batch_size = 20
batch_size_val = 10
batch_size_test = 8
nb_train_samples = 148
nb_validation_samples = 60
nb_test_samples = 56
model = Sequential()
model.add(Conv2D(32, (3, 3), activation='relu', input_shape=input_shape))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(32, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
```

```

model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(64, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(1, activation='sigmoid'))
plot_model(model, to_file=model_puth, show_layer_names=True)
train_generator = datagen.flow_from_directory(
source_train_dir,
shuffle=True,
target_size=(img_width, img_height),
batch_size=batch_size,
class_mode='binary')
val_generator = datagen.flow_from_directory(
source_valid_dir,
shuffle=True,
target_size=(img_width, img_height),
batch_size=batch_size_val,
class_mode='binary')
test_generator = datagen.flow_from_directory(
source_test_dir,
shuffle=False,
target_size=(img_width, img_height),
batch_size=batch_size_test,
class_mode='binary')
model.compile(loss='binary_crossentropy',
optimizer='adam',
metrics=['accuracy'])
model.fit_generator(
train_generator,
steps_per_epoch=nb_train_samples // batch_size,
epochs=epochs,
validation_data=val_generator,
validation_steps=nb_validation_samples // batch_size_val)
scores = model.evaluate_generator(test_generator, nb_test_samples // batch_size)
print("Аккуратность на тестовых данных: %.2f%%" % (scores[1]*100))
Y_pred = model.predict_generator(test_generator, nb_test_samples // batch_size_test)
y_pred = np.round(Y_pred)

```

```
cnf_matrix = confusion_matrix(test_generator.classes, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cnf_matrix, display_labels=['1', '0'])
disp.plot()
plt.show()
target_names = ['1', '0']
print(classification_report(test_generator.classes, y_pred, target_names=target_names))
plot.ClassificationReport.from_raw_data(y_pred, test_generator.classes, target_names=target_names)
```