



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) – Отделение экспериментальной физики

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема работы
<i>Использование методов машинного обучения для составления оптимального портфеля ценных бумаг</i>

УДК 004.85.032.26:336.763

Обучающийся

Группа	ФИО	Подпись	Дата
0В91	Галлямов Андрей Ильмирович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Крицкий Олег Леонидович	Кандидат ф. – м. наук, доцент		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Кащук Ирина Вадимовна	Кандидат технических наук, доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Антоневич Ольга Алексеевна	Кандидат биологических наук		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП/ОПОП, должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ ИЯТШ	Крицкий Олег Леонидович	Кандидат ф.– м. наук, доцент		

Томск – 2023 г.



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) – Отделение экспериментальной физики

УТВЕРЖДАЮ:

Руководитель ООП/ОПОП

_____ Крицкий О.Л.
(Подпись) (Дата) (ФИО)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

Обучающийся:

Группа	ФИО
0В91	Галлямов Андрей Ильмирович

Тема работы:

<i>Использование методов машинного обучения для составления оптимального портфеля ценных бумаг</i>	
<i>Утверждена приказом директора (дата, номер)</i>	

Срок сдачи обучающимся выполненной работы:

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Котировки цен анализируемых акций за период 01.07.2022 – 31.12.2022 гг. Ежедневные цены закрытия.
Перечень разделов пояснительной записки подлежащих исследованию, проектированию и разработке	<ol style="list-style-type: none"> 1. Сформировать оптимальный портфель методом Марковица. 2. Применить машинное обучение для формирования оптимального портфеля. 3. Найти доходности портфелей за период 01.01.2023 – 31.12.2023. 4. Провести анализ полученных доходностей.
Перечень графического материала	
Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i>	

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Кащук Ирина Вадимовна
Социальная ответственность	Антоневич Ольга Алексеевна

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

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ ИЯТШ	Крицкий Олег Лениодович	Кандидат ф. – м. наук, доцент		

Задание принял к исполнению обучающийся:

Группа	ФИО	Подпись	Дата
0В91	Галлямов Андрей Ильмирович		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСООБЪЕКТИВНОСТЬ
И РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
0В91	Галлямову Андрею Ильмировичу

Школа	Инженерная школа ядерных технологий	Отделение Школа	Экспериментальной физики
Уровень образования	Бакалавриат	Направление/специальность	01.03.02 Прикладная математика и информатика

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

<i>1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	Стоимость материальных ресурсов и специального оборудования определены в соответствии с рыночными ценами г. Томска Тарифные ставки исполнителей определены штатным расписанием НИ ТПУ
<i>2. Нормы и нормативы расходования ресурсов</i>	Норма амортизационных отчислений на специальное оборудование
<i>3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	Отчисления во внебюджетные фонды 30 %

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<i>1. Анализ конкурентных технических решений (НИ)</i>	Расчет конкурентоспособности SWOT-анализ
<i>2. Формирование плана и графика разработки и внедрения (НИ)</i>	Структура работ. Определение трудоемкости. Разработка графика проведения исследования
<i>3. Составление бюджета инженерного проекта (НИ)</i>	Расчет бюджетной стоимости НИ
<i>4. Оценка ресурсной, финансовой, бюджетной эффективности (НИ)</i>	Интегральный финансовый показатель. Интегральный показатель ресурсоэффективности. Интегральный показатель эффективности.

Перечень графического материала

<ol style="list-style-type: none"> 1. Оценка конкурентоспособности ИР 2. Матрица SWOT 3. Диаграмма Ганта 4. Бюджет НИ 5. Основные показатели эффективности НИ
--

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Кащук Ирина Вадимовна	к.т.н, доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0В91	Галлямов Андрей Ильмирович		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа 0В91		ФИО Галлямову Андрею Ильмировичу	
Школа	Инженерная школа ядерных технологий	Отделение (НОЦ)	Экспериментальной физики
Уровень образования	Бакалавриат	Направление/специальность	01.03.02 Прикладная математика и информатика

Тема ВКР:

Использование методов машинного обучения для составления оптимального портфеля ценных бумаг

Исходные данные к разделу «Социальная ответственность»:

<p>Введение</p> <p>– Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения.</p> <p>– Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации</p>	<p><i>Объект исследования:</i> методы машинного обучения для составления оптимального портфеля ценных бумаг</p> <p><i>Область применения:</i> биржевые рынки, банки, любые проекты, связанные с инвестиционной деятельностью</p> <p><i>Рабочая зона:</i> офис</p> <p><i>Размеры помещения:</i> 27,2 м²</p> <p><i>Количество и наименование оборудования рабочей зоны:</i> 1 персональный компьютер</p> <p><i>Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне:</i> алгоритмическая и программная разработка с использованием персонального компьютера</p>
---	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения:</p> <p>– специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства;</p> <p>– организационные мероприятия при компоновке рабочей зоны.</p>	<p>– ГОСТ 12.2.032-78 «Рабочее место при выполнении работ сидя»</p> <p>– СП 2.4.3648-20 "Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи"</p> <p>– Трудовой кодекс Российской Федерации: федер. Закон от 30 дек. 2001 г. №197-ФЗ Раздел 10</p>
<p>2. Производственная безопасность при разработке проектного решения:</p> <p>– Анализ выявленных вредных и опасных производственных факторов</p>	<p>Опасные факторы:</p> <p>– Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий;</p> <p>Вредные факторы:</p> <p>– Отклонение показателей микроклимата;</p> <p>– Недостаточная освещённость рабочей зоны;</p> <p>– Пониженная световая и цветовая контрастность;</p> <p>– Повышенный уровень шума на рабочем месте;</p>

	– Длительное сосредоточенное наблюдение
3. Экологическая безопасность при разработке проектного решения:	Анализ воздействия на литосферу: – Утилизация компьютеров, оргтехники и бумаги; Анализ воздействия на гидросферу: – Производство компьютерной техники;
4. Безопасность в чрезвычайных ситуациях при разработке проектного решения:	– Наиболее типичная ЧС: Пожар;
Дата выдачи задания для раздела по линейному графику	

Задание выдал консультант:

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

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0B91	Галлямов Андрей Ильмирович		

Реферат

Выпускная квалификационная работа содержит 64 страницы, 19 рисунков, 19 таблиц, 32 источника, 3 приложения.

Ключевые слова: машинное обучение, модели формирования оптимальных портфелей, модель Марковица.

Объектом исследования являются ценные бумаги, их котировки за определенный промежуток времени.

Цель работы: Использование методов машинного обучения для составления оптимального портфеля ценных бумаг.

Методы проведения работы: теоретические (изучение литературы, обзор методов и алгоритмов) и практические (написание программного решения, анализ полученных данных и результатов).

Результаты, полученные по итогу данной работы, могут быть использованы в банках и любых проектах, связанных с инвестиционной деятельностью.

Для выполнения работы использовались MS Excel, Jupiter Notebook, MS Word 2021. Язык программирования Python.

Оглавление

Введение.....	10
1. Теоретическая часть.....	12
1.1. Фондовый рынок.....	12
1.2 Портфель ценных бумаг.....	14
1.3 Модель Марковица.....	15
1.4 Алгоритмы машинного обучения.....	18
2. Практическая часть.....	22
3. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	34
3.1. Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения.....	34
3.1.1. Анализ конкурентных технических решений.....	34
3.1.2. SWOT-анализ.....	36
3.2. Планирование научно-исследовательских работ.....	38
3.2.1. Структура работ в рамках научного исследования.....	38
3.2.2. Определение трудоемкости выполнения работ и разработка графика проведения.....	39
3.3 Бюджет научно-технического исследования.....	41
3.3.1. Расчет амортизации специального оборудования.....	42
3.3.2. Основная заработная плата исполнителей темы.....	43
3.3.3. Отчисления во внебюджетные фонды (страховые отчисления).....	45
3.3.4. Накладные расходы.....	45
3.3.5. Бюджет НИР.....	46
3.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.....	46
Выводы по разделу.....	48
4. Социальная ответственность.....	49
Введение.....	49
4.1. Правовые и организационные вопросы обеспечения безопасности.....	49
4.2. Производственная безопасность.....	51
4.2.1. Отклонение показателей микроклимата.....	52
4.2.2. Недостаточная освещенность рабочей зоны.....	53
4.2.3. Повышенная световая и цветовая контрастность.....	54
4.2.4. Повышенный уровень шума на рабочем месте.....	55
4.2.5. Длительное сосредоточенное наблюдение.....	56
4.2.6. Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий.....	56
4.3. Экологическая безопасность.....	57

4.4. Безопасность в чрезвычайных ситуациях.....	59
Заключение по разделу	60
Заключение	61
Список литературы	62
Приложение 1. Программный код для составления оптимального портфеля с помощью алгоритма Случайный лес	66
Приложение 2. Программный код для составления оптимального портфеля с помощью алгоритма LSTM	67
Приложение 3. Программный код для составления оптимального портфеля с помощью алгоритма Дерево решений.....	68

Введение

Сегодня безусловно важно заниматься наукой и любому образованному человеку стоит понимать, куда движется сейчас человечество. И машинное обучение является одним из ключевых моментов в том куда мы идем. Постоянно выходят новости про то, что искусственный интеллект рисует картины в стиле Ван Гога, побеждает чемпионов мира по разным компьютерным играм, заменил рабочее место человека машиной под управлением сложного программного обеспечения. А также говорят, что именно машинное обучение позволит создать полностью функционирующий искусственный интеллект.

Помимо этого, тема инвестиций становится все популярнее с каждым годом, а инвесторам интересны новые способы нахождения лучших стратегий для заработка на фондовом рынке. Инвестирование в ценные бумаги является более рискованным вложением чем банковский вклад, но также этот способ является более доходным. Существует немало методов и стратегий инвестирования в различные активы, одним из наиболее интересных является стратегия разработанная Гарри Марковицем.

Модель Марковица для составления оптимального портфеля была сформулирована в середине двадцатого века, но даже по сей день она является одной из популярнейших методик для формирования инвестиционного портфеля. Модель Марковица строится на предположении, что инвесторы будут максимизировать доходность портфеля при заданном уровне риска. Для решения этой задачи необходимо выбрать такие активы, которые будут иметь наилучшее соотношение между доходностью и риском.

В данной работе будут анализироваться 10 ценных бумаг входящих в индекс ММВБ-10, а именно Газпром, Норильский никель, Лукойл, Магнит, Новатэк, Полюс, Полиметалл, Роснефть, Сбербанк, Яндекс. Эти ценные бумаги являются самыми ликвидными акциями РФ.

Цель работы – использование методов машинного обучения для составления оптимального портфеля ценных бумаг.

Для достижения поставленной цели были сформулированы следующие задачи:

1. Сформировать оптимальный портфель ценных бумаг максимальной доходности и заданного уровня риска при помощи метода Марковица.

2. Написать программное решение для формирования оптимального портфеля, включающее применение разных алгоритмов машинного обучения.

3. Найти доходности у сформированных разными способами портфелей.

4. Проанализировать полученные результаты.

Объектом для исследования являются методы машинного обучения для составления оптимального портфеля ценных бумаг.

1. Теоретическая часть

1.1. Фондовый рынок

Рыночная экономика состоит из множества рынков, среди которых можно выделить финансовый рынок. Финансовый рынок, в свою очередь, включает в себя различные сегменты, одним из которых является рынок ценных бумаг, также известный как фондовый рынок. На данном рынке осуществляется привлечение капитала путем выпуска и обращения ценных бумаг, а также их перераспределение. Участники рынка могут купить или продать ценные бумаги, что позволяет формировать и изменять их цену. Рынок ценных бумаг является важным инструментом для привлечения инвестиций и позволяет устанавливать справедливые цены на финансовые инструменты.

Фондовый рынок необходим для привлечения средств инвесторов, а также для финансирования и расширения хозяйственной деятельности компаний. Помимо этого, он нужен для предоставления инвесторам информации о состоянии экономики и ориентиров для размещения своих инвестиций. Эта информация отражается в курсовой стоимости ценных бумаг, которая может расти или падать в зависимости от перспектив компаний и состояния экономики в целом. Например, повышение цены акций какой-либо компании может давать неверные сигналы, и инвесторы могут ошибочно оценивать перспективы компаний и экономики. Тем не менее, в целом фондовый рынок дает хорошие ориентиры для принятия инвестиционных решений и является важным инструментом для определения состояния экономики в целом.

Ситуация на фондовом рынке может предвещать будущее состояние экономики. Это происходит потому, что инвесторы постоянно ищут и анализируют информацию о компаниях, чтобы предвидеть будущие результаты деятельности и принимать инвестиционные решения. Однако стоит отметить, что на фондовом рынке могут возникать различные факторы, включая эмоциональные реакции инвесторов или влияние внешних событий, которые могут приводить к краткосрочным колебаниям цен на ценные бумаги и не всегда отражать реальное состояние экономики в целом. Тем не менее, в долгосрочной перспективе фондовый рынок

является важным индикатором экономического развития и может давать полезную информацию для инвесторов и других участников рынка.

Состояние фондового рынка играет важную роль для экономического развития, поскольку крах рынка может вызвать застой в экономике. Это связано с тем, что уменьшение стоимости ценных бумаг делает вкладчиков беднее. Кроме того, падение курсовой стоимости уменьшает возможности предприятий для аккумуляции необходимых им средств за счет выпуска новых бумаг. Таким образом, фондовый рынок представляет собой важный индикатор экономического развития, который может давать полезную информацию для инвесторов и других участников рынка. В данной работе мы будем рассматривать акции как один из видов ценных бумаг, которые торгуются на фондовом рынке.

Акция – это эмиссионная ценная бумага, закрепляющая право ее владельца на получение части прибыли акционерного общества в виде дивидендов, на участие в управлении акционерным обществом и на часть имущества, остающегося после его ликвидации. На рынке цена акции определяется в результате взаимодействия спроса на акции и их предложения. Характеристикой акции является доход, который она приносит акционеру. Доход по акции может быть представлен в двух формах – в виде прироста курсовой стоимости и в качестве периодических выплат по акции. Во втором случае доход называют дивидендом. Прирост курсовой стоимости акции может составить существенную часть доходов инвестора. Чтобы его реализовать акцию необходимо продать. В противном случае существует опасность, что в следующий момент курс бумаги может упасть. Прирост курсовой стоимости возникает по двум причинам. Во-первых, это возможный спекулятивный подъем на рынке. Во-вторых, это реальный прирост активов предприятия. Получив прибыль, акционерное общество делит ее на две части. Одна часть выплачивается в качестве дивидендов, другая – реинвестируется для поддержания и расширения производства. Реинвестируемая прибыль, приобретающая форму основных и оборотных фондов, реально «наполняет» акцию и ведет, как правило, к росту ее стоимости. В результате, в тенденции цена акции на рынке должна расти. Если инвестор ориентируется прежде всего на прирост курсовой стоимости акции, то он выбирает более рискованную

стратегию, так как цена бумаги может в любой момент и понизиться.

Другим источником дохода акционера является дивиденд. Если инвестор менее склонен к риску, ему следует остановиться на акциях, по которым регулярно выплачиваются дивиденды, хотя в этом случае их цена может расти и не очень быстро. Большая надежность такой стратегии состоит в том, что инвестор реально получает доход уже в период владения акцией. Даже если в будущем курсовая стоимость упадет, то все равно он уже реализовал часть дохода [1].

1.2 Портфель ценных бумаг

При формировании портфеля ценных бумаг инвестор может вложить свои средства в различные виды активов, такие как акции, облигации, срочные контракты и недвижимость. Главной целью формирования портфеля является достижение требуемого уровня ожидаемой доходности при более низком уровне ожидаемого риска. Это достигается путем диверсификации портфеля по составу инструментов, то есть распределения средств инвестора между различными активами, а также тщательного подбора финансовых инструментов.

В теории и практике управления портфелем существуют два основных подхода: традиционный и современный. Традиционный подход основывается на фундаментальном и техническом анализе, при котором аналитик оценивает финансовые показатели компаний и рынков, а также их технические характеристики, такие как цены и объемы торгов. Современный подход, в свою очередь, использует математические модели и алгоритмы для принятия инвестиционных решений, учитывая множество факторов, включая макроэкономические данные и глобальные тенденции рынка. Оба подхода имеют свои преимущества и недостатки, и выбор конкретного подхода зависит от инвестора и его инвестиционных целей.

Поэтому инвестор при формировании портфеля строит свой инвестиционный выбор на основе ожидаемых значений доходности и риска. Данные величины оцениваются, в первую очередь, на основе статистической информации за предыдущие периоды времени. Кроме того, при формировании портфеля необходимо учитывать такие факторы, как срок инвестирования, ликвидность активов, инфляция и налоги. Эти факторы могут существенно влиять на конечный результат

инвестирования. Чтобы минимизировать риски и получить максимальную доходность, инвестор должен тщательно анализировать каждый актив, входящий в его портфель, и периодически корректировать его состав в соответствии с изменением рыночных условий и своих инвестиционных целей.

1.3 Модель Марковица

Современная теория портфельного инвестирования была заложена в статьях Гарри Марковица (1952 год), работах Вильяма Шарпа (1964 год) и Джона Литнера (1965 год), и была основана на понятиях систематического (рыночного) и несистематического рисков ценной бумаги [2].

В своих работах Г. Марковиц предложил математическую модель для оптимального выбора активов, обеспечивающих максимальную доходность при минимальном риске. Эта модель, называемая моделью Марковица или моделью оптимального портфеля, позволяет инвесторам определить оптимальный состав портфеля на основе его ожидаемой доходности и риска, а также корреляции между различными активами. С помощью этой модели инвесторы могут выбирать портфели, которые наилучшим образом соответствуют их инвестиционным целям и рискам. За свои работы в области портфельного менеджмента и оптимизации инвестиций Г. Марковиц был удостоен Нобелевской премии по экономике в 1990 году.

Модель портфельного анализа Марковица основана на таких предположениях как:

- Рынок состоит из конечного числа абсолютно ликвидных активов, которые подразумеваются бесконечно делимыми. Доходности рискованных активов являются нормально распределенными случайными величинами, имеющими конечные моменты первого (математическое ожидание) и второго (дисперсия) порядка.
- Индивидуальные предпочтения инвестора задаются функцией полезности от двух аргументов: ожидаемой доходности, измеряемой математическим ожиданием, и риска, оцениваемого дисперсией. Сравнение портфелей осуществляется на основе только двух критериев.
- Инвестор не склонен к риску, т.е. из двух портфелей с одинаковой ожидаемой доходностью он предпочтет портфель с меньшим риском. В то же время

из двух портфелей с одинаковым риском инвестор выберет портфель с большей ожидаемой доходностью.

- Налоги и транзакционные издержки равны нулю [3].

Сегодня современное финансовое инвестирование непосредственно связано с формированием инвестиционного портфеля. Одними из актуальных задач финансового вложения становится анализ и прогнозирование ожидаемой прибыли и рисков [4].

Основу портфельного инвестирования составляет распределение инвестиционных средств между различными группами активов, так как невозможно предвидеть одновременное выполнение двух условий: высокую надежность и максимальную доходность. В зависимости от целей и задач, инвестор анализирует состояние рынка и причины, влияющие на цену актива. Процесс формирования портфеля подразумевает задание наиболее подходящей структуры портфеля для данного вида активов, так же определяется процентное соотношение между финансовыми инструментами. Портфельное инвестирование позволяет улучшить условия вложения капитала, придав совокупности криптовалют такие инвестиционные качества, которые невозможно достигнуть для отдельно взятого финансового инструмента [5].

Таким образом, портфель представляет собой совокупность финансовых инструментов, объединенных вместе для реализации целей инвестора, увеличения прибыли и уменьшения рыночного риска.

Основой портфеля Марковица является принцип предпочтения инвестора, согласно которому при равных условиях инвесторы предпочитают портфели с наименьшим риском.

В модели Марковица формируются стандартные портфели, то есть портфели, состоящие только из купленных активов. Из вышесказанного вытекает первое условие: все бумаги должны иметь положительные доли, то есть

$$A_i > 0, \quad (1)$$

Вторым условием является наличие суммы долей ценных активов, равной 1:

$$\sum_{i=1}^n A_i = 1 \quad (2)$$

где A_i - доли ценных бумаг в портфеле.

Согласно теории Марковица, показателем доходности является его математическое ожидание, а мера риска рассчитывается через стандартное отклонение [6].

Доходность портфеля (ожидаемая доходность) составляется из суммы доходностей, заявленных в портфеле акций с выбранными весовыми коэффициентами. Так как, при прочих равных условиях инвесторы стремятся увеличить доходность и минимизировать риски, то необходимо максимизировать целевую функцию, которая будет иметь вид:

$$\sum_{i=1}^n M_i \cdot A_i \rightarrow \max \quad (3)$$

где M_i - доходность i -й акции.

Доходность портфеля эквивалентно ожидаемой доходности μ_x

$$\mu_x = E[r_x(t)] = \sum_{i=1}^n E[r_i(t)] \cdot x_i = \sum_{i=1}^n \mu_i \cdot x_i \quad (4)$$

где $r_i(t)$ - доходность актива.

Риск портфеля эквивалентен волатильности портфеля σ_x

$$\sigma_x = \text{var}(r_x(t)) = \text{var}\left(\sum_{i=1}^n r_i \cdot x_i\right) = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(r_i(t), r_j(t)) \cdot x_i \cdot x_j \quad (5)$$

где $\sigma_{i,j} = \text{cov}(r_i(t), r_j(t)) = \rho_{ij} \sigma_i \sigma_j$ – ковариация n активов.

$\rho_{ij} = \text{corr}(r_i(t), r_j(t))$ – корреляция n активов.

Согласно модели Марковица, для заданного уровня риска существует оптимальный портфель с наибольшей доходностью. Будем считать значение d_p уровнем максимального риска для инвестора. Кроме среднеквадратичного отклонения в портфеле Марковица учитывается матрица корреляции доходностей. В итоге риск портфеля представляется формулой:

$$\sum_{i=1}^n A_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=i+1}^{n+1} A_i A_j \rho_{ij} \sigma_i \sigma_j < \sigma_p^2 \quad (6)$$

где A_i – доходности i -ых акций, r_{ij} – корреляция между доходностями.

Доходность можно описать, как наклон средней линии, построенной на графике цен активов, а уровень риска теория Марковица описывает как амплитуду колебаний реальной цены по отношению к уровню доходности. Кроме того, Марковиц отмечал, что чем больше амплитуда колебаний, тем менее предсказуемо поведение цены [6].

Можно рассматривать задачу поиска оптимального портфеля с двух различных сторон: обеспечение минимального риска, при котором будет обеспечен определенный уровень дохода, и получение максимальной доходности при заданном уровне риска.

Во избежание избыточного риска инвесторы будут стремиться минимизировать стандартное отклонение доходности путем диверсификации капитала между различными объектами вложений [7]. В этом случае происходит уменьшение риска, поскольку стандартное отклонение доходности портфеля будет меньше, чем средневзвешенные стандартные отклонения активов, которые составляют этот портфель.

Таким образом, экономико-математическая модель задачи формирования оптимального по Марковицу портфеля, примет следующий вид:

$$\begin{aligned} E(X_{\Pi}) &\rightarrow \max \\ X_{\Pi} &= \sum_{k=1}^n \alpha_k \cdot x_k \\ \sum_{k=1}^n \alpha_k &= 1 \\ \sigma_{\Pi}^2 &< \sigma^2 \end{aligned} \quad (7)$$

где X_{Π} – формируемый портфель, x_k – активы, входящие в портфель, α_k – доли акций в портфеле, σ_{Π}^2 – риск портфеля, σ^2 – заданный уровень риска.

1.4 Алгоритмы машинного обучения

В данной работе используются 3 алгоритма, а именно Случайный лес, LSTM (Long Short-Term Memory), Дерево решений.

Случайный лес (Random Forest) — это метод машинного обучения, который сочетает в себе принципы ансамблевого обучения и деревьев решений. Он был предложен Лео Брейманом и Адель Катлер в 2001 году. Основная идея случайного леса заключается в создании ансамбля независимых деревьев решений и объединении их прогнозов для получения более точных результатов [8].

Случайный лес строит несколько деревьев решений на основе случайных подвыборок из обучающего набора данных. Для каждого дерева выбирается случайное подмножество признаков из всех доступных признаков. Это позволяет случайному лесу устойчиво работать с большим количеством признаков и предотвращает переобучение модели [8].

Каждое дерево в случайном лесу обучается независимо от остальных деревьев. Обучение происходит путем разбиения обучающего набора данных на подмножества и применения алгоритма дерева решений к каждому подмножеству [9].

Преимущества случайного леса заключаются в его способности обрабатывать большие объемы данных с большим числом признаков, а также в его устойчивости к переобучению. Случайный лес также может оценить важность признаков, исходя из их вклада в улучшение качества модели [9].

LSTM (Long Short-Term Memory) является методом машинного обучения, относящимся к классу рекуррентных нейронных сетей (RNN), который разработан для работы с последовательными данными. LSTM изначально представлен в работе "Long Short-Term Memory" Хохрайтером и Шмидхубером в 1997 году [10].

Основной элемент LSTM - это нейронная ячейка (cell), которая состоит из нескольких компонентов: входного (input) и выходного (output) затворов, забывающего (forget) затвора и обновляющего (update) затвора. Каждый из этих затворов контролирует поток информации внутри ячейки и определяет, какая информация будет сохранена или забыта. Они работают вместе, чтобы контролировать поток информации внутри сети и сохранять важные контекстные зависимости в данных [11].

Преимущество LSTM заключается в его способности эффективно обрабатывать длинные последовательности и сохранять информацию о важных событиях или зависимостях на больших временных промежутках [12].

Дерево решений (Decision Tree) — это метод машинного обучения, который строит модель прогнозирования в виде древовидной структуры решений. В каждом узле дерева происходит разбиение данных на основе значимых признаков, чтобы минимизировать неопределенность и улучшить прогнозы. Дерево решений можно использовать для решения задач классификации или регрессии [13].

В основе дерева решений лежит принцип разбиения данных на основе информационной энтропии. Информационная энтропия представляет собой меру неопределенности набора данных. Цель дерева решений состоит в том, чтобы разбить данные таким образом, чтобы каждое разбиение уменьшало неопределенность и увеличивало чистоту (или однородность) подмножеств данных [14].

Процесс разбиения данных продолжается до достижения определенного условия остановки, такого как достижение максимальной глубины дерева или достижение минимального количества наблюдений в узле. В каждом листовом узле дерева содержится прогнозное значение для задачи классификации или регрессии [15].

Формулы для расчета матрицы ошибок, точности (accuracy), точности (precision), полноты (recall) и F1-меры:

Пусть у нас есть два класса: положительный (P) и отрицательный (N).

Матрица ошибок:

- True Positive (TP) - количество правильно предсказанных положительных примеров.
- False Positive (FP) - количество неправильно предсказанных положительных примеров.
- False Negative (FN) - количество неправильно предсказанных отрицательных примеров.
- True Negative (TN) - количество правильно предсказанных отрицательных примеров.

Accuracy (точность):

- $Accuracy = (TP + TN) / (TP + FP + FN + TN)$

Precision (точность):

- $Precision = TP / (TP + FP)$

Recall (полнота):

- $Recall = TP / (TP + FN)$

F1-мера:

- $F1 = 2 * (Precision * Recall) / (Precision + Recall)$

Коэффициенты MAE (Mean Absolute Error) и MSE (Mean Squared Error) являются метриками, используемыми для оценки качества прогнозных моделей или регрессионных моделей.

MAE измеряет среднее абсолютное отклонение прогнозов от фактических значений. Для вычисления MAE необходимо выполнить следующие шаги:

- Для каждого наблюдения в выборке вычислить абсолютное значение разницы между фактическим значением и прогнозом.
- Полученные абсолютные значения сложить вместе.
- Разделить сумму на общее количество наблюдений в выборке.

Формула для вычисления MAE выглядит следующим образом:

$$MAE = \frac{1}{n} \cdot \sum |Y_i - Y| \quad (8)$$

где Y_i – фактическое значение, Y – прогнозное значение, n – количество наблюдений.

MSE измеряет среднее квадратичное отклонение прогнозов от фактических значений. Для вычисления MSE необходимо выполнить следующие шаги:

- Для каждого наблюдения в выборке вычислить квадрат разницы между фактическим значением и прогнозом.
- Полученные квадраты сложить вместе.
- Разделить сумму на общее количество наблюдений в выборке.

Формула для вычисления MSE выглядит следующим образом:

$$MAE = \frac{1}{n} \cdot \sum (Y_i - Y)^2 \quad (9)$$

Оба коэффициента MAE и MSE являются положительными значениями. Чем меньше значение MAE и MSE, тем лучше модель справляется с прогнозированием или регрессией.

Масштабирование данных — это процесс приведения значений переменных к определенному диапазону. Целью масштабирования является создание стандартизованного диапазона значений, что позволяет сравнивать и анализировать переменные с разными единицами измерения.

Формула для масштабирования данных с использованием метода MinMaxScaler выглядит следующим образом:

$$x_{sc} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (10)$$

x — исходное значение, $\min(x)$ — минимальное значение в наборе данных, $\max(x)$ — максимальное значение в наборе данных, x_{sc} — масштабированное значение.

Формула для обратного масштабирования с использованием метода MinMaxScaler следующая:

$$x = x_{sc} \cdot (\max(x) - \min(x)) + \min(x) \quad (11)$$

2. Практическая часть

Для выполнения задания необходимо сформировать портфель из 10 акций и взять данные их котировок за период от 01.07.2022 по 31.12.2022. В данном периоде было 130 дней, когда на фондовой бирже осуществлялись торги ценными бумагами. Загрузим данные цен акций на момент закрытия биржи в Excel документ.

	A	B	C	D	E	F	G	H	I	J	K
1	Дата	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс
2	01.07.2022	192,50	17100,00	3944,00	4155,00	1060,20	8464,00	372,10	375,15	129,91	1596,00
3	04.07.2022	186,25	17024,00	3949,50	4198,00	1025,00	7997,50	355,10	371,60	131,43	1660,00
4	05.07.2022	197,30	16700,00	4015,00	4314,00	970,20	8100,00	360,10	371,50	133,50	1701,40
5	06.07.2022	195,42	16452,00	3962,00	4306,00	1009,60	8020,00	354,40	369,65	134,00	1655,00
6	07.07.2022	197,92	16550,00	3956,00	4420,00	998,00	8225,00	364,00	364,95	133,09	1711,00
7	08.07.2022	198,00	16720,00	3911,00	4439,00	993,00	8376,00	363,40	351,20	133,30	1704,00
8	11.07.2022	188,90	16270,00	3848,50	4385,00	968,60	8484,00	362,20	343,80	126,80	1678,00
9	12.07.2022	191,40	15922,00	3752,00	4300,00	942,00	8320,00	341,40	333,50	129,00	1660,60
10	13.07.2022	186,00	15200,00	3645,50	4255,50	916,20	8171,00	317,10	321,00	126,00	1600,00
11	14.07.2022	184,55	15302,00	3608,00	4228,00	926,60	8182,00	310,00	315,45	125,90	1613,60

Рис. 1. Котировки цен акций для формирования портфеля

Далее найдем относительные приращения активов по формуле $r_i = \frac{S_t - S_{t-1}}{S_{t-1}}$, где

$t > 0$, а также годовые доходности (среднедневные умноженные на 365).

	A	B	C	D	E	F	G	H	I	J
1	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс
2	Годовые доходности									
3	-0,255511394	-0,22612353	0,160995445	0,251004039	0,130526382	-0,08077718	0,413513772	0,008535612	0,345026762	0,564613692
4	Относительные приращения									
5	-0,0325	-0,0044	0,0014	0,0103	-0,0332	-0,0551	-0,0457	-0,0095	0,0117	0,0401
6	0,0593	-0,0190	0,0166	0,0276	-0,0535	0,0128	0,0141	-0,0003	0,0157	0,0249
7	-0,0095	-0,0149	-0,0132	-0,0019	0,0406	-0,0099	-0,0158	-0,0050	0,0037	-0,0273
8	0,0128	0,0060	-0,0015	0,0265	-0,0115	0,0256	0,0271	-0,0127	-0,0068	0,0338
9	0,0004	0,0103	-0,0114	0,0043	-0,0050	0,0184	-0,0016	-0,0377	0,0016	-0,0041
10	-0,0460	-0,0269	-0,0160	-0,0122	-0,0246	0,0129	-0,0033	-0,0211	-0,0488	-0,0153
11	0,0132	-0,0214	-0,0251	-0,0194	-0,0275	-0,0193	-0,0574	-0,0300	0,0174	-0,0104
12	-0,0282	-0,0453	-0,0284	-0,0103	-0,0274	-0,0179	-0,0712	-0,0375	-0,0233	-0,0365
13	-0,0078	0,0067	-0,0103	-0,0065	0,0114	0,0013	-0,0224	-0,0173	-0,0008	0,0085
14	0,0166	0,0195	0,0280	0,0360	0,0285	0,0236	0,0016	0,0477	0,0238	0,0155

Рис. 2. Относительные приращения и годовые доходности акций

После этого используем средства Excel, а именно анализ данных с помощью которого найдем оценку дневной матрицы ковариаций.

	A	B	C	D	E	F	G	H	I	J
1	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс
2	Годовые доходности									
3	-0,255511394	-0,22612353	0,160995445	0,251004039	0,130526382	-0,08077718	0,413513772	0,008535612	0,345026762	0,564613692
4	Относительные приращения									
5	-0,0325	-0,0044	0,0014	0,0103	-0,0332	-0,0551	-0,0457	-0,0095	0,0117	0,0401
6	0,0593	-0,0190	0,0166	0,0276	-0,0535	0,0128	0,0141	-0,0003	0,0157	0,0249
7	-0,0095	-0,0149	-0,0132	-0,0019	0,0406	-0,0099	-0,0158	-0,0050	0,0037	-0,0273
8	0,0128	0,0060	-0,0015	0,0265	-0,0115	0,0256	0,0271	-0,0127	-0,0068	0,0338
9	0,0004	0,0103	-0,0114	0,0043	-0,0050	0,0184	-0,0016	-0,0377	0,0016	-0,0041
10	-0,0460	-0,0269	-0,0160	-0,0122	-0,0246	0,0129	-0,0033	-0,0211	-0,0488	-0,0153
11	0,0132	-0,0214	-0,0251	-0,0194	-0,0275	-0,0193	-0,0574	-0,0300	0,0174	-0,0104
12	-0,0282	-0,0453	-0,0284	-0,0103	-0,0274	-0,0179	-0,0712	-0,0375	-0,0233	-0,0365
13	-0,0078	0,0067	-0,0103	-0,0065	0,0114	0,0013	-0,0224	-0,0173	-0,0008	0,0085
14	0,0166	0,0195	0,0280	0,0360	0,0285	0,0236	0,0016	0,0477	0,0238	0,0155
15	-0,0049	-0,0190	0,0166	0,0276	-0,0535	0,0128	0,0141	-0,0003	0,0157	0,0249
16	0,0151	-0,0149	-0,0132	-0,0019	0,0406	-0,0099	-0,0158	-0,0050	0,0037	-0,0273
17	0,0200	0,0060	-0,0015	0,0265	-0,0115	0,0256	0,0271	-0,0127	-0,0068	0,0338
18	-0,0192	0,0103	-0,0114	0,0043	-0,0050	0,0184	-0,0016	-0,0377	0,0016	-0,0041

Рис. 3. Построение матрицы ковариаций с помощью анализа данных

	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс
Газпром	0,00123606	0,000228792	0,000227554	0,000210466	0,000188298	0,000342321	0,000299064	0,000322626	0,000373987	0,000364533
НорНикель	0,000228792	0,000552777	0,000305962	0,000274312	0,000384481	0,000433129	0,000547399	0,000449073	0,000407385	0,000361662
Лукойл	0,000227554	0,000305962	0,000564312	0,000209881	0,00030113	0,000261421	0,00032063	0,000348195	0,000314738	0,00030871
Магнит	0,000210466	0,000274312	0,000209881	0,000450013	0,0002728	0,000322981	0,000417738	0,000361214	0,000302237	0,000352549
Новатэк	0,000188298	0,000384481	0,00030113	0,0002728	0,000649647	0,0003783	0,000630684	0,000468328	0,000415607	0,000490544
Полюс	0,000342321	0,000433129	0,000261421	0,000322981	0,0003783	0,000951139	0,000912431	0,000498636	0,000400416	0,000348917
Полиметалл	0,000299064	0,000547399	0,00032063	0,000417738	0,000630684	0,000912431	0,002193981	0,000629605	0,000510135	0,000646748
Роснефть	0,000322626	0,000449073	0,000348195	0,000361214	0,000468328	0,000498636	0,000629605	0,000693848	0,000500425	0,000447867
Сбербанк	0,000373987	0,000407385	0,000314738	0,000302237	0,000415607	0,000400416	0,000510135	0,000500425	0,000602321	0,000475259
Яндекс	0,000364533	0,000361662	0,00030871	0,000352549	0,000490544	0,000348917	0,000646748	0,000447867	0,000475259	0,01023894

Рис. 4. Дневная матрица ковариации

Теперь необходимо получить оценку годовой матрицы ковариаций, для чего необходимо умножить каждое значение матрицы дневной ковариации на 247 (количество рабочих дней в 2022 году).

	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс
Газпром	0,30531	$=N5*247$	0,05621	0,05199	0,04651	0,08455	0,07387	0,07969	0,09237	0,09004
НорНикель	0,05651	0,13654	0,07557	0,06776	0,09497	0,10698	0,13521	0,11092	0,10062	0,08933
Лукойл	0,05621	0,07557	0,13939	0,05184	0,07438	0,06457	0,07920	0,08600	0,07774	0,07625
Магнит	0,05199	0,06776	0,05184	0,11115	0,06738	0,07978	0,10318	0,08922	0,07465	0,08708
Новатэк	0,04651	0,09497	0,07438	0,06738	0,16046	0,09344	0,15578	0,11568	0,10265	0,12116
Полюс	0,08455	0,10698	0,06457	0,07978	0,09344	0,23493	0,22537	0,12316	0,09890	0,08618
Полиметалл	0,07387	0,13521	0,07920	0,10318	0,15578	0,22537	0,54191	0,15551	0,12600	0,15975
Роснефть	0,07969	0,11092	0,08600	0,08922	0,11568	0,12316	0,15551	0,17138	0,12360	0,11062
Сбербанк	0,09237	0,10062	0,07774	0,07465	0,10265	0,09890	0,12600	0,12360	0,14877	0,11739
Яндекс	0,09004	0,08933	0,07625	0,08708	0,12116	0,08618	0,15975	0,11062	0,11739	0,25290

Рис. 5. Годовая матрица ковариации

Следующей задачей является нахождение весовых коэффициентов. Для этого необходимо ввести начальные формулы, чтобы применить к полученным ранее данным инструмент Поиск решения.

	A	B	C	D	E	F	G	H	I	J	K	L
1		Годовые доходности										
2		-0,255511394	-0,22612353	0,16099545	0,25100404	0,13052638	-0,08077718	0,41351377	0,00853561	0,34502676	0,56461369	
3												
4		Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс	
5	Газпром	0,305306899	0,05651151	0,05620586	0,05198501	0,0465095	0,0845532	0,07386869	0,07968858	0,09237482	0,09003966	
6	НорНикель	0,056511506	0,13653582	0,07557263	0,06775502	0,09496682	0,10698292	0,13520767	0,11092109	0,10062409	0,08933046	
7	Лукойл	0,056205864	0,07557263	0,13938513	0,05184059	0,07437906	0,06457108	0,07919551	0,08600413	0,07774038	0,07625143	
8	Магнит	0,051985011	0,06775502	0,05184059	0,11115315	0,06738152	0,07977634	0,10318124	0,08921974	0,07465248	0,08707953	
9	Новатэк	0,046509504	0,09496682	0,07437906	0,06738152	0,16046271	0,09344014	0,15577892	0,11567698	0,10265488	0,12116431	
10	Полюс	0,0845532	0,10698292	0,06457108	0,07977634	0,09344014	0,23493129	0,2253704	0,12316301	0,09890283	0,08618242	
11	Полиметалл	0,073868693	0,13520767	0,07919551	0,10318124	0,15577892	0,2253704	0,54191331	0,15551244	0,12600336	0,15974678	
12	Роснефть	0,079688579	0,11092109	0,08600413	0,08921974	0,11567698	0,12316301	0,15551244	0,17138049	0,12360486	0,11062311	
13	Сбербанк	0,092374819	0,10062409	0,07774038	0,07465248	0,10265488	0,09890283	0,12600336	0,12360486	0,1487733	0,11738905	
14	Яндекс	0,090039664	0,08933046	0,07625143	0,08707953	0,12116431	0,08618242	0,15974678	0,11062311	0,11738905	0,25290193	
15												
16		Весовые коэффициенты портфеля										
17		Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс	1
18		0	0	0	0	0	0	0	0	0	0	0
19												
20	Доходность портфеля	0,0000	$=\text{СУММПРОИЗВ}(B18:K18;B2:K2)$									$=\text{СУММ}(B18:K18)$
21			$=\text{КОРЕНЬ}(\text{МУМНОЖ}(B18:K18;\text{МУМНОЖ}(B5:K14;\text{ТРАНСП}(B18:K18))))$									
22												
23	Волатильность	0,00	<=	0,3								

Рис. 6. Исходные данные для нахождения долей активов

Зададим максимальный риск, на который готов пойти инвестор на уровне 0.3 и найдем оптимальные доли актива для портфеля.

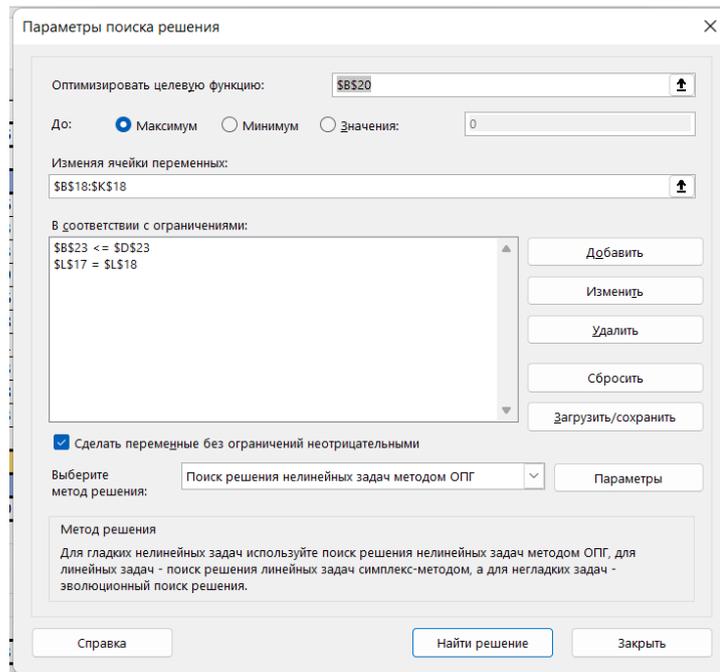


Рис. 7. Окно поиска решения с необходимыми условиями

Весовые коэффициенты портфеля											
Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полус	Полиметалл	Роснефть	Сбербанк	Яндекс		
0,00	0,00	0,24	0,45	0,00	0,00	0,00	0,00	0,18	0,13		1
Доходность портфеля											
0,2866											
Волатильность											
0,30		<=		0,3							

Рис. 8. Портфель полученный методом Марковица

Теперь приступим к задаче формирования оптимального портфеля при помощи машинного обучения.

Импортируем библиотеки для работы с данными, а также загрузим цены акций. Используем первую строчку в файле, как заголовки, а остальные строчки отделим в переменную `prices`. После этого посчитаем ежедневные доходности для акций и сохраним их в переменную `returns`.

```
import pandas as pd
import numpy as np

# Загрузка данных из CSV файла
data = pd.read_csv('Книга3.csv')

# Получение ценных бумаг
assets = data.columns[0:]

# Получение данных о ценах для акций
prices = data[assets].values

# Расчет ежедневных доходностей
returns = np.diff(prices, axis=0) / prices[:-1]
```

Рис. 9. Загрузка данных в среду для программирования

Далее используем алгоритмы машинного обучения, а именно Случайный лес,

LSTM, Дерево решений, чтобы получить новые данные о ценах активов. Для метода Случайный лес мы сначала разделим данные на обучающую и тестовую выборку, объем которых равен 80% и 20% от общего объема данных соответственно. После этого создается сама модель случайного леса. Затем происходит обучение модели. И в конце прогнозируются доходности для тестовых данных на следующие 180 дней, которые будут использоваться в дальнейшем.

```
# Разделение данных на обучающую и тестовую выборки
train_size = int(0.8 * len(returns))
train_returns = returns[:train_size]
test_returns = returns[train_size:]

# Создание модели Random Forest для прогнозирования доходностей
model = RandomForestRegressor(n_estimators=3, random_state=5)

# Обучение модели на обучающих данных
X_train = train_returns[:-1]
y_train = train_returns[1:]
model.fit(X_train, y_train)

# Прогнозирование доходностей для тестовых данных
X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

# Прогноз на следующие 180 дней
predicted_returns = []
for i in range(180):
    # Применяем модель для текущего дня прогноза
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

# Обновляем текущую доходность для следующего дня
current_returns = np.append(current_returns[1:], predicted_return)
```

Рис. 10. Алгоритм Случайный лес

После этого найдем метрики качества для данного алгоритма:

```
# Вычисляем метрики качества
actual_returns = test_returns[1:] # Фактические доходности
mae = mean_absolute_error(actual_returns, predicted_returns)
mse = mean_squared_error(actual_returns, predicted_returns)

# Выводим метрики качества
print("MAE:", mae)
print("MSE:", mse)
```

Рис. 11. Метрики качества для алгоритма Случайный лес

Получим следующие значения:

```
MAE: 0.017878735303746396
MSE: 0.0005460766094436458
```

Рис. 12. Значения метрик качества для алгоритма Случайный лес

MAE (средняя абсолютная ошибка) составляет примерно 0,0179. Это означает, что средняя абсолютная разница между фактическими и прогнозными доходностями

составляет примерно 0,0179. Тогда годовая равна этому значению, умноженному на корень из количества рабочих дней в году, то есть 0,28.

MSE (среднеквадратичная ошибка) составляет примерно 0,000546. Это означает, что средняя квадратичная разница между фактическими и прогнозными доходностями составляет примерно 0,000546. Годовая равна 0,0086.

Для алгоритма LSTM начальные данные масштабируются для нормализации и улучшения сходимости модели. Затем происходит деление на обучающую и тестовую выборку. После этого создается модель LSTM для прогнозирования доходностей. К модели добавляется слой с 64 нейронами. Также в модели используется среднеквадратичная ошибка в качестве функции потерь. Затем происходит обучение модели. И в конце прогнозируются доходности для тестовых данных на следующие 180 дней, которые будут использоваться в дальнейшем, а также обратное масштабирование.

```
# Масштабирование доходностей
scaler = MinMaxScaler()
returns_scaled = scaler.fit_transform(returns)

# Разделение данных на обучающую и тестовую выборки
train_size = int(0.8 * len(returns_scaled))
train_returns = returns_scaled[:train_size]
test_returns = returns_scaled[train_size:]

# Создание модели LSTM для прогнозирования доходностей
model = Sequential()
model.add(LSTM(64, input_shape=(1, len(assets))))
model.add(Dense(len(assets)))
model.compile(loss='mean_squared_error', optimizer='adam')

# Обучение модели на обучающих данных
X_train = np.expand_dims(train_returns[:-1], axis=1) # Изменено
y_train = train_returns[1:]
model.fit(X_train, y_train, epochs=20, batch_size=1)

# Прогнозирование доходностей для тестовых данных
X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

# Прогноз на следующие 180 дней
predicted_returns = []
for i in range(180):
    # Применяем модель для текущего дня прогноза
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

    # Обновляем текущую доходность для следующего дня
    current_returns = np.append(current_returns[1:], predicted_return)

# Обратное масштабирование прогнозируемых доходностей
predicted_returns = scaler.inverse_transform(predicted_returns)
```

Рис. 13. Алгоритм LSTM

MAE: 0.015870130907311705
MSE: 0.0003755117647903608

Рис. 14. Метрики качества для алгоритма LSTM

Для алгоритма «Дерево решений» создаются обучающая и тестовая выборка, затем создается сама модель, где используется параметр `fit`, который выполняет оптимизацию параметров на основе данных обучения. После этого происходит прогнозирование доходностей на следующие 180 дней.

```
# Разделение данных на обучающую и тестовую выборки
train_size = int(0.8 * len(returns))
train_returns = returns[:train_size]
test_returns = returns[train_size:]

# Создание модели дерева принятия решений для прогнозирования доходностей
model = DecisionTreeRegressor()

# Обучение модели на обучающих данных
model.fit(train_returns[:-1], train_returns[1:])

# Прогнозирование доходностей для тестовых данных
X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

# Прогноз на следующие 180 дней
predicted_returns = []
for i in range(180):
    # Применяем модель для текущего дня прогноза
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

# Обновляем текущую доходность для следующего дня
current_returns = np.append(current_returns[1:], predicted_return)
```

Рис. 15. Алгоритм «Дерево решений»

MAE: 0.014082257376937062
MSE: 0.0003209041038459962

Рис. 16. Метрики качества для алгоритма «Дерево решений»

Отметим, что метрики качества у каждого алгоритма находятся в пределах нормы.

Теперь напишем программное решение, которое будет выполнять оптимизацию портфеля с использованием модели Марковица. Для начала вычислим доходность и риск портфеля при помощи функции `portfolio_performance`. Затем используем функцию `minimize_risk` для минимизации риска портфеля при оптимизации. Далее поставим условия, что сумма весов активов равна 1, доля актива в портфеле находится в диапазоне от 0 до 1. Далее при помощи итерационного метода SLSQP (Sequential Least Squares Programming), который приближает решение задачи, обновляя значение переменных на каждой итерации оптимизируем портфель. В конце выведем результат.

```

# Функция для расчета риска и доходности портфеля
def portfolio_performance(weights, returns):
    portfolio_returns = np.dot(weights, returns.mean(axis=0))
    portfolio_risk = np.sqrt(np.dot(weights.T, np.dot(np.cov(returns.T), weights)))
    return portfolio_returns, portfolio_risk

# Функция для минимизации риска портфеля
def minimize_risk(weights, returns):
    portfolio_returns, portfolio_risk = portfolio_performance(weights, returns)
    return portfolio_risk

# Определение ограничений и начального приближения для весов активов
constraints = [{'type': 'eq', 'fun': lambda x: np.sum(x) - 1}]
bounds = tuple((0, 1) for _ in assets)
initial_guess = np.array([1 / len(assets) for _ in assets])

# Оптимизация портфеля по Марковицу для минимизации риска
optimized_portfolio = minimize(minimize_risk, initial_guess, args=(predicted_returns,),
                               method='SLSQP', bounds=bounds, constraints=constraints)

# Получение оптимальных весов активов
optimal_weights = optimized_portfolio.x

# Вывод результатов
for asset, weight in zip(assets, optimal_weights):
    print(f"{asset}: {weight}")

```

Рис. 17. Программная реализация метода построения оптимального портфеля

Покажем портфели, которые были получены при помощи составленных программ использующие метод Случайный лес, LSTM, Дерево решений.

	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс	
Случайный лес	0,12	0,00	0,03	0,79	0,00	0,06	0,00	0,00	0,00	0,00	1
LSTM	0,00	0,23	0,13	0,09	0,14	0,22	0,00	0,00	0,19	0,00	1
Дерево решений	0,10	0,00	0,12	0,42	0,13	0,17	0,07	0,00	0,00	0,00	1

Рис. 18. Портфели, составленные программами.

Проанализируем все полученные портфели и сравним их между собой. Для этого загрузим цены активов за период с 01.01.2023 по 31.01.2023. Затем найдем доходность портфеля путем перемножения долей активов на ежедневные приращения этих активов.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Дата	Лукойл	Приращения	Магнит	Приращения	Сбербанк	Приращения	Яндекс	Приращения	Доходность портфеля		Доля в портфеле	Лукойл	Магнит	Сбербанк	Яндекс
2	03.01.2023	4123,50		4360,50		141,78		1832,00		100%			0,24	0,45	0,18	0,13
3	04.01.2023	4116,00	-0,0018	4369,00	0,0019	141,43	-0,0025	1867,00	0,0191	100%						
4	05.01.2023	4087,00	-0,0070	4325,50	-0,0100	141,27	-0,0011	1855,40	-0,0062	100%						
5	06.01.2023	4089,00	0,0005	4319,50	-0,0014	141,40	0,0009	1850,40	-0,0027	99%						
6	09.01.2023	4110,00	0,0051	4330,00	0,0024	142,40	0,0071	1866,80	0,0089	100%						
7	10.01.2023	4079,00	-0,0075	4372,50	0,0098	142,81	0,0029	1870,00	0,0017	100%						
8	11.01.2023	4090,50	0,0028	4458,00	0,0196	149,96	0,0501	1909,00	0,0209	102%						
9	12.01.2023	4078,50	-0,0029	4638,00	0,0404	149,30	-0,0044	1897,80	-0,0059	104%						
10	13.01.2023	4074,00	-0,0011	4620,00	-0,0039	151,69	0,0160	1912,00	0,0075	104%						
11	16.01.2023	4052,00	-0,0054	4674,00	0,0117	153,71	0,0133	1953,80	0,0219	105%						
12	17.01.2023	4025,00	-0,0067	4628,00	-0,0098	150,45	-0,0212	1889,60	-0,0329	104%						
13	18.01.2023	4021,00	-0,0010	4575,50	-0,0113	151,47	0,0068	1906,60	0,0090	103%						
14	19.01.2023	3981,00	-0,0099	4504,50	-0,0155	149,34	-0,0141	1882,40	-0,0127	102%						
15	20.01.2023	3944,00	-0,0093	4575,50	0,0158	151,38	0,0137	1904,80	0,0119	103%						
16	23.01.2023	3972,50	0,0072	4634,50	0,0129	153,50	0,0140	1945,00	0,0211	104%						
17	24.01.2023	3942,50	-0,0076	4586,50	-0,0104	152,66	-0,0055	1919,80	-0,0130	103%						
18	25.01.2023	3923,50	-0,0048	4597,00	0,0023	152,81	0,0010	1914,80	-0,0026	103%						
19	26.01.2023	3911,00	-0,0032	4602,00	0,0011	152,40	-0,0027	1902,80	-0,0063	103%						
20	27.01.2023	3917,00	0,0015	4609,00	0,0015	153,20	0,0052	1912,00	0,0048	103%						
21	30.01.2023	3947,50	0,0078	4653,50	0,0097	153,38	0,0012	1934,00	0,0115	104%						
22	31.01.2023	3955,00	0,0019	4665,00	0,0025	158,07	0,0306	2047,20	0,0585	106%						

Рис. 19. Доходность портфеля полученного методом Марковица.

Для портфелей, полученных программой, сделаем аналогичным образом и по итогу получим следующие доходности:

Случайный лес	LSTM	Дерево решений
100%	100%	100%
100%	100%	100%
99%	99%	99%
99%	99%	99%
99%	99%	99%
100%	99%	99%
102%	100%	101%
105%	101%	102%
105%	101%	103%
106%	102%	104%
105%	101%	103%
104%	101%	102%
103%	100%	101%
104%	100%	102%
105%	101%	103%
104%	101%	103%
104%	101%	103%
105%	101%	104%
105%	102%	104%
106%	104%	106%
106%	104%	106%

Рис. 20. Доходности портфелей, полученные программами

Также для сравнения доходностей приведем результат, который получится, если вложиться в индекс ММВБ-10, который состоит из акций, которые анализируются в данной задаче, примерно в одинаковых долях.

	A	B	C	D
1	Дата	ММВБ-10	Приращения	Доходность
2	03.01.2023	4074,48		1
3	04.01.2023	4073,06	0,000	1,000
4	05.01.2023	4046,96	-0,006	0,993
5	06.01.2023	4046,34	0,000	0,993
6	09.01.2023	4060,98	0,004	0,997
7	10.01.2023	4051,4	-0,002	0,994
8	11.01.2023	4077,14	0,006	1,001
9	12.01.2023	4073,87	-0,001	1,000
10	13.01.2023	4097,51	0,006	1,006
11	16.01.2023	4145,92	0,012	1,018
12	17.01.2023	4098,51	-0,011	1,006
13	18.01.2023	4099,48	0,000	1,006
14	19.01.2023	4045,22	-0,013	0,993
15	20.01.2023	4043,75	0,000	0,993
16	23.01.2023	4074,24	0,008	1,000
17	24.01.2023	4053,03	-0,005	0,995
18	25.01.2023	4050,63	-0,001	0,994
19	26.01.2023	4048,57	-0,001	0,994
20	27.01.2023	4086,35	0,009	1,003
21	30.01.2023	4143,24	0,014	1,017
22	31.01.2023	4190,23	0,011	1,029

Рис. 21. Доходность индекса ММВБ-10.

Покажем таблицу долей активов в каждом портфеле:

	Газпром	НорНикель	Лукойл	Магнит	Новатэк	Полюс	Полиметалл	Роснефть	Сбербанк	Яндекс	
Портфель Марковица	0,00	0,00	0,24	0,45	0,00	0,00	0,00	0,00	0,18	0,13	1
Случайный лес	0,12	0,00	0,03	0,79	0,00	0,06	0,00	0,00	0,00	0,00	1
LTSM	0,00	0,23	0,13	0,09	0,14	0,22	0,00	0,00	0,19	0,00	1
Дерево решений	0,10	0,00	0,12	0,42	0,13	0,17	0,07	0,00	0,00	0,00	1
ММВБ-10	0,10	0,09	0,01	0,09	0,01	0,10	0,01	0,01	0,01	0,10	1

Рис. 22. Доли активов в каждом портфеле

Каждый метод дает разные доли для ценных бумаг в оптимальном портфеле, так как они имеют свои собственные алгоритмы и критерии для оптимизации.

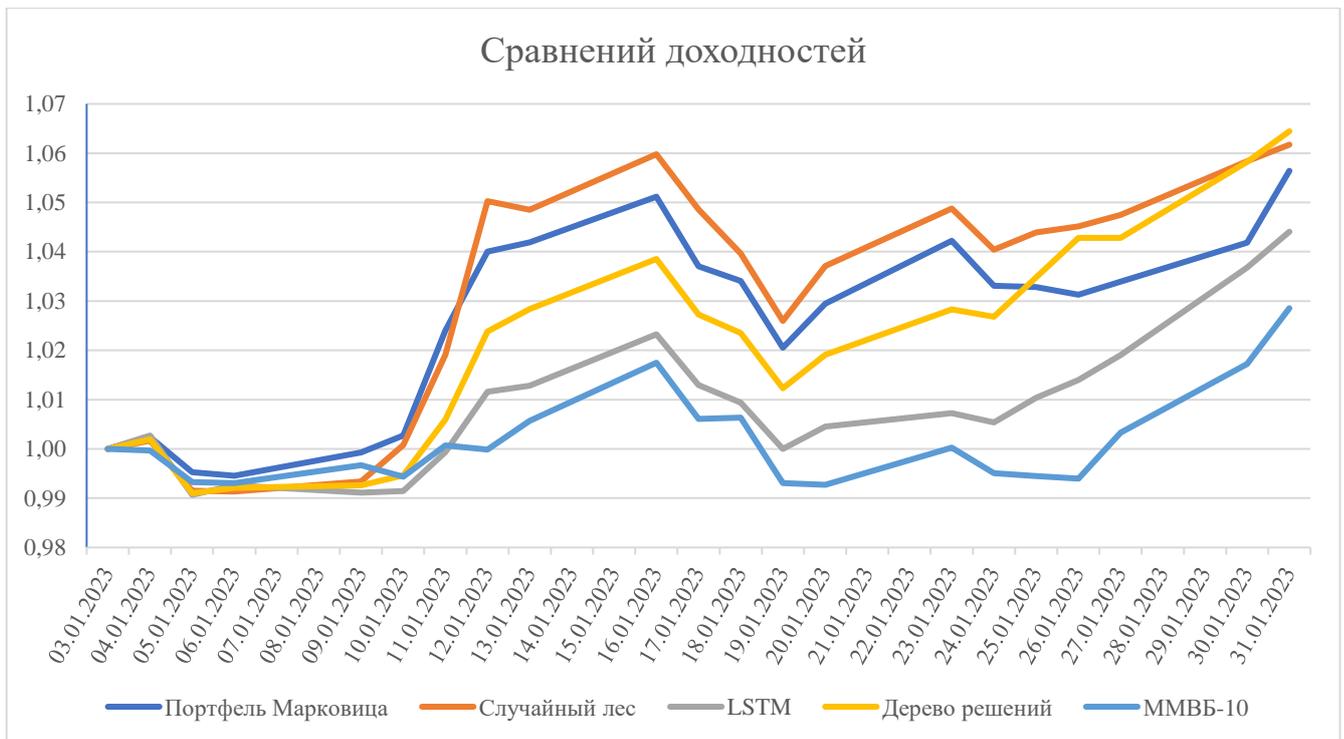


Рис. 23. График доходностей каждого портфеля.

Проведем анализ данных, представленных на рис. 23. Для всех методов динамика цен очень схожа во многих моментах, хоть и доли активов абсолютно везде разные. Тем не менее метод Случайный лес показывает наивысшую доходность почти каждый день, а также имеет достаточно высокую доходность по окончанию месяца, равную 6,4%, или 75% годовых. Также высокую доходность имеют методы Случайный лес и метод Марковица, равные 6,2% и 5,6% соответственно, или 72% и 66% годовых соответственно. Немного меньше, но при этом все еще хорошую доходность в 4,4% в месяц, или 51% годовых показывает метод LSTM. И самый худший результат показывает индекс ММВБ-10 в 2,9%, или 33% годовых, что логично.

Причина, по которой алгоритмы «Дерево решений» и «Случайный лес» показывают наивысшую доходность заключается в том, что они способны выявлять нелинейные зависимости: алгоритмы могут обнаружить нелинейные зависимости между переменными, которые другие алгоритмы могут упустить. Это особенно важно в случае финансовых рынков, где связи между переменными могут быть сложными и нелинейными.

Таким образом можно сделать вывод, что составление оптимального портфеля

ценных бумаг при помощи алгоритмов машинного обучения открывает новые возможности для инвесторов и позволяет учитывать более сложные зависимости и нелинейности в данных. При правильном выборе алгоритма и грамотной настройке параметров можно достичь конкурентноспособной доходности портфеля, превосходящей результаты классического подхода на основе модели Марковица.

3. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Основная цель данного раздела – оценить перспективность развития и планировать финансовую и коммерческую ценность конечного продукта, представленного в рамках исследовательской программы. Коммерческая ценность определяется не только наличием более высоких технических характеристик над конкурентными разработками, но и тем, насколько быстро разработчик сможет ответить на следующие вопросы – будет ли продукт востребован на рынке, какова будет его цена, каков бюджет научного исследования, какое время будет необходимо для продвижения разработанного продукта на рынок.

Данный раздел, предусматривает рассмотрение следующих задач:

- Оценка коммерческого потенциала разработки.
- Планирование научно-исследовательской работы;
- Расчет бюджета научно-исследовательской работы;
- Определение ресурсной, финансовой, бюджетной эффективности исследования.

Цель данной ВКР – Использование методов машинного обучения для составления оптимального портфеля ценных бумаг.

3.1. Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения

3.1.1. Анализ конкурентных технических решений

Т.к. основной целью для любого инвестора является получение прибыли от собственного капитала, то можно выделить три основных способа достижения поставленной цели:

- Банковские вклады, которые рассматриваются, как альтернатива инвестиций в фондовый рынок, но с меньшим риском
- Использование разработанной инвестиционной стратегии, т.е. вложение денег в полученный портфель
- Вложение денег в индекс ММВБ – 10 (в соответствии капитализацией входящих в индекс компаний)

Каждый из выделенных способов имеет свои особенности, которые будут отражены в сводной таблице. Для сравнений трёх способов будут использованы следующие критерии:

- Повышение производительности труда пользователя. Увеличение скорости расчета, возможность работать с большими объемами данных.
- Удобство в эксплуатации.
- Точность вычислений.
- Сложность вычислений
- Доходность портфеля. Доходность портфеля отображает, сколько процентов от капитала может получить инвестора. Таким образом, высокая доходность является наиболее привлекательной для инвестора, но также не стоит и забывать о соотношении потенциальный риск / доходность.
- Диверсифицированность портфеля. Диверсификация портфеля позволяет клиенту быть уверенным быть в том, что его деньги распределены корректно между активами.
- Надежность. Надёжность является немаловажным фактором при выборе способа заработка от капитала, т.к. чем надёжней способ заработка, тем комфортнее инвестору переживать просадки в портфеле.
- Ликвидность. Данный показатель отражает способность активов быть быстро проданными по цене, близкой к рыночной.

Таблица 1 – Оценочная карта конкурентоспособности.

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б _ф	Б _{к1}	Б _{к2}	К _ф	К _{к1}	К _{к2}
Технические критерии оценки ресурсоэффективности							
1. Повышение производительности труда пользователя	0,1	5	5	5	0,5	0,5	0,5
2. Удобство в эксплуатации	0,1	4	3	2	0,4	0,3	0,2
3. Точность вычислений	0,1	5	4	4	0,5	0,4	0,4
4. Сложность вычислений	0,1	4	4	4	0,4	0,4	0,4
Экономические критерии оценки эффективности							
1. Доходность портфеля	0,25	2	4	4	0,5	1	1
2. Диверсифицированность портфеля	0,15	3	4	2	0,45	0,6	0,3
3. Надежность	0,1	5	5	5	0,5	0,5	0,5

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б _ф	Б _{к1}	Б _{к2}	К _ф	К _{к1}	К _{к2}
4. Ликвидность	0,1	5	5	3	0,5	0,5	0,5
Итого	1	33	34	29	3,75	4,2	3,8

Позиция разработки оценивается по каждому показателю экспертным путем по пятибалльной шкале, где 1 – наиболее слабая позиция, а 5 – наиболее сильная. Анализ конкурентных решений определяется по формуле:

$$K = \sum B_i \cdot B_i,$$

где K – конкурентоспособность решения или конкурента, B_i – вес показателя (в долях единицы), B_i – балл i -го показателя.

Таким образом, исходя из полученных данных, инвестиционный портфель, построенный в ходе работы, является наиболее предпочтительный, чем два других механизма заработка.

3.1.2. SWOT-анализ

SWOT-анализ – это сводная таблица, иллюстрирующая взаимосвязь между какими-либо внутренними и внешними факторами. При проведении анализа выявляются слабые и сильные стороны компании или продукта. Таким образом, производители могут объективно оценивать продукцию или услугу в отраслях. Далее проведём SWOT-анализ для портфеля, составленного из акций, входящих в индекс ММВБ - 10.

Таблица 2 – Матрица SWOT-анализа

	Сильные стороны	Слабые стороны
	<p>С1. Высокая степень доходности при заданном уровне риска.</p> <p>С2. Портфель состоит из «голубых фишек», что обеспечивает высокую ликвидность.</p>	<p>Сл1. Необходимость проверки состава индекса каждый квартал.</p> <p>Сл2. Индекс, в основном, ориентирован на «скальперов».</p> <p>Сл3. Для того чтобы вложиться в ММВБ-10</p>

	С3. Инвестор может сам определить для себя допустимый уровень риска.	необходимо приобрести акции всех 10 компаний.
Возможности В1. Повышение инвестиционной активности населения В2. Защита частных инвестиций путем принятия новых законов и нормативно-правовых актов. В3. Повышение финансовой грамотности населения	В1С1. Реализация большого числа финансовых операций, обеспечивающих вложение денежных средств в различные отрасли российской экономики. В1С3. Проведение инвестором нескольких рискованных операций при увеличении доходности. В3С2. Формирование особого слоя населения, отличающегося высокой степенью Финансовой грамотности.	В3Сл1. Возможность последующего изменения процесса вхождения котировок в индекс ММВБ-10. В1Сл3. Реализация изменения количества котировок, входящих в индекс ММВБ-10.
Угрозы У1. Частые финансовые кризисы У2. Отток российского капитала за рубеж. У3. Непредсказуемые скачки валютных курсов.	У1У2С1. Обеспечение защиты инвесторов путем принятия необходимых законов и правовых актов. У3С1С2С3. Обеспечение конкуренции с индексами других стран.	У1Сл1. Обеспечение поддержки известных компаний. У2Сл3. Возможность изменения концепции индекса ММВБ10 и модификации её под интересы инвесторы.

Таким образом, становится очевидно, что для наибольшей эффективности в реализации продуктов данной отрасли, следует:

- Повышать общий уровень финансовой грамотности инвесторов, а также необходимо приводить клиентам объективную оценку российских компаний, чтобы повысить их уровень доверия к ним. Таким образом, совокупный объём инвестиций в Российский фондовый рынок увеличится, соответственно будет расти и объём инвестиций в составленный портфель

- Увеличить уровень участия государства, чтобы обеспечить большую

стабильность портфелю, т.к. защищённые активы становятся более привлекательными для инвесторов.

- Усилить взаимодействие с самими компаниями, входящими в состав индекса, чтобы усилить потенциальную доходность, а также обладать нужной информацией для реформирования портфеля.

3.2. Планирование научно-исследовательских работ

3.2.1. Структура работ в рамках научного исследования

Планирование комплекса предполагаемых работ осуществляется в следующем порядке:

- определение структуры работ в рамках научного исследования;
- определение участников каждой работы;
- установление продолжительности работ;
- построение графика проведения научных исследований.

Для разработки алгоритма сформируем рабочую группу, которая будет состоять из разработчика и научного руководителя. На таблице 3 представлен порядок выполнения работ разработки.

Таблица 3 – Перечень этапов, работ и распределение исполнителей

№ раб.	Этапы работ	Должность исполнителя
1	Постановка целей и задачи	Научный руководитель
2	Разработка ТЗ	Разработчик, научный руководитель
3	Изучение материала по заданной теме	Разработчик, научный руководитель
4	Разработка плана выполнения работы	Разработчик
5	Обсуждение и подбор литературы	Разработчик, научный руководитель
6	Разработка программы	Разработчик
7	Тестирование разработанной программы	Разработчик
8	Оформление РПЗ (Расчетно-пояснительной записки)	Разработчик
9	Оформления необходимых графических материалов	Разработчик, научный руководитель
10	Анализ полученных результатов	Разработчик

3.2.2. Определение трудоемкости выполнения работ и разработка графика проведения

При проведении научных исследований основную часть стоимости разработки составляют трудовые затраты, поэтому определение трудоемкости проводимых работ является важным этапом составления сметы.

Для определения ожидаемого (среднего) значения трудоемкости использована следующая формула:

$$t_{\text{ож}i} = \frac{3t_{\text{min}i} + 2t_{\text{max}i}}{5}, \quad (3.1)$$

где $t_{\text{ож}i}$ – ожидаемая трудоемкость выполнения i -ой работы, человеко-дни;

$t_{\text{min}i}$ – минимально возможная трудоемкость выполнения заданной i -ой работы, человеко-дни;

$t_{\text{max}i}$ – максимально возможная трудоемкость выполнения заданной i -ой работы, человеко-дни.

Зная величину ожидаемой трудоемкости, можно определить продолжительность каждой i -ой работы в рабочих днях T_{pi} , при этом учитывается параллельность выполнения работ разными исполнителями. Данный расчёт позволяет определить величину заработной платы.

$$T_{pi} = \frac{t_{\text{ож}i}}{Ч_i}, \quad (3.2)$$

где T_{pi} – продолжительность одной работы, рабочие дни;

$t_{\text{ож}i}$ – ожидаемая трудоемкость выполнения одной работы, человеко-дни;

$Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Для перевода длительности каждого этапа из рабочих в календарные дни, необходимо воспользоваться формулой (3.3):

$$T_{ki.инж} = T_{pi} \cdot k_{\text{кал}}, \quad (3.3)$$

где T_{ki} – продолжительность выполнения i -й работы в календарных днях;

T_{pi} – продолжительность выполнения i -й работы в рабочих днях;

$k_{кал}$ – календарный коэффициент.

Календарный коэффициент определяется по формуле:

$$k_{кал.инж} = \frac{T_{кал}}{T_{кал} - T_{вых} - T_{пр}} = \frac{365}{365 - 104 - 14} = 1,48 \quad (3.4)$$

где $T_{кал}$ – общее количество календарных дней в году; $T_{вых}$ – общее количество выходных дней в году; $T_{пр}$ – общее количество праздничных дней в году (2023 год).

Расчеты временных показателей проведения научного исследования обобщены в таблице 4.

Таблица 4 – Временные показатели проведения научного исследования

Название работы	Трудоёмкость работ						Длительность работ в рабочих днях T_{pi}	Длительность работ в календарных днях T_{ki}
	t_{min} , чел-дни		t_{max} , чел-дни		$t_{ожид}$, чел-дни			
	Исп.1	Исп.2	Исп.1	Исп.2	Исп.1	Исп.2		
1	2	3	4	5	6	7	8	9
1. Постановка целей и задачи	-	2	-	5	-	3,2	3,2	4
2. Разработка ТЗ	2	2	4	4	2,8	2,8	2,8	4
3. Изучение материала по заданной теме	7	3	10	5	8,2	3,8	6	8
4. Разработка плана выполнения работы	8	-	15	-	10,8	-	10,8	15
5. Обсуждение и подбор литературы	1	1	3	3	1,8	1,8	1,8	2
6. Разработка программы	15	-	35	-	23	-	23	34
7. Тестирование разработанной программы	5	-	30	-	15	-	15	22
8. Оформление РПЗ(Расчетно-пояснительной записки)	3	-	10	-	5,8	-	5,8	8

9. Оформление необходимых графических материалов	3	3	6	6	4,2	4,2	4,2	6
10. Анализ полученных результатов	12	-	17	-	14	-	14	20
Итого:	56	11	130	23	85,6	15,8	86,6	123

Примечание: Исп. 1 – разработчик, Исп. 2 – научный руководитель.

На основе таблицы составлен календарный план-график выполнения проекта с использованием диаграммы Ганта (таблица 5).

Таблица 5 – Диаграмма Ганта

№	Вид работ	Исп	T _{кi} , кал. дн.	Продолжительность работ													
				февр			март			апр			май				
				1	2	3	1	2	3	1	2	3	1	2	3		
1	Составление и утверждение технического задания, утверждение плана-графика	Исп2	4	■													
2	Календарное планирование выполнения работ	Исп1 Исп2	4	■	■												
3	Обзор научной литературы	Исп1	8		■												
4	Обзор фреймворков среды реализации	Исп1	15		■												
5	Построение блок-схемы алгоритма	Исп1 Исп2	2			■											
6	Реализация кода	Исп1	34				■	■	■								
7	Отладка программы	Исп1	22							■	■	■					
8	Обработка полученных результатов	Исп1	8										■	■			
9	Оценка эффективности, модернизация алгоритма	Исп1 Исп2	6											■	■		
10	Составление пояснительной записки	Исп1	20													■	■

Примечание:

■ – Исп. 1 (разработчик), ■ – Исп. 2 (научный руководитель)

3.3 Бюджет научно-технического исследования

При планировании бюджета научно-технического исследования учитывались все виды расходов, связанных с его выполнением. В этой работе использовать следующую группировку затрат по следующим статьям:

- материальные затраты научно-исследовательской работы (НИР);

- затраты на специальное оборудование для экспериментальных работ;
- основная заработная плата исполнителей темы;
- дополнительная заработная плата исполнителей темы;
- отчисления во внебюджетные фонды (страховые отчисления);
- накладные расходы НИР.

3.3.1. Расчет амортизации специального оборудования

Расчет амортизации проводится следующим образом:

Норма амортизации: рассчитывается по формуле:

$$H_A = \frac{1}{n}, \quad (4.5)$$

где n – срок полезного использования в количестве лет.

Амортизация оборудования рассчитывается по формуле:

$$A = \frac{H_A I}{12} \cdot m, \quad (4.6)$$

где I – итоговая сумма, тыс. руб.; m – время использования, мес.

Данный проект подразумевает научный подход для реализации. По этой причине затраты на сырьё, комплектующие и различные материалы будут отсутствовать. Перейдем сразу к затратам на специальное оборудование.

Таблица 6 - Затраты на программное оборудование

№ п/п	Наименование оборудования	Кол-во единиц оборудования	Цена единицы оборудования, руб.	Общая стоимость оборудования, тыс.руб.
1.	Ноутбук HP 15s-eq2090ur	1	55 434	55 434
2.	Ноутбук Acer Nitro-5	1	50 000	50 000
3.	Операционная система - Windows Pro	1	30 000	60 000

Стоит отметить, что учитываться затраты будут только на амортизационные отчисления, так как у рабочей группы уже имеются ноутбуки. Лицензия операционной системы приобретается на бессрочный период, но так как ресурс ноутбука, на который устанавливается ПО, ограничен (в среднем 8 лет), то перерасчёт

производится на количество времени, затраченное на проект (2,3 месяца). Стоимость одной лицензии – 30 000 рублей.

Норма амортизационных отчисления на приобретение лицензии:

$$H_A = \frac{1}{8} \cdot 100 = 12,5\%$$

Из этого следует, что размер амортизации составит:

$$A_{\text{лиц}} = \frac{30\,000 \cdot 2}{12} \cdot 0,125 \cdot 4,1 = 2562,5 \text{ руб.}$$

Затраты на установку ОС на 2 устройства:

$$A_{\text{устан}} = 5000 \text{ рублей}$$

Также дистрибутив языка программирования и данные анализа хранятся в открытом доступе (тип лицензии).

Рассчитаем амортизацию оборудования линейным способом. Производитель даёт гарантию на модель 5 лет. Соответственно, рассчитаем размер амортизационных отчислений за специальное оборудование (компьютеры) за время разработки алгоритма:

$$A_{\text{лиц}} = \frac{105434}{5 \cdot 12} \cdot 4,1 = 7205 \text{ руб.}$$

Итоговая сумма затрат использование программного оборудования:

$$Z_{\text{оборуд}} = A_{\text{устан}} + A_{\text{лиц}} + A_{\text{лиц}} = 5000 + 2562,5 + 7205 = 14767,5 \text{ руб.}$$

3.3.2. Основная заработная плата исполнителей темы

В данном разделе рассчитывается заработная плата инженера и руководителя, помимо этого необходимо рассчитать расходы по заработной плате, определяемые трудоемкостью проекта и действующей системой оклада.

Основная заработная плата $Z_{\text{осн}}$ одного работника рассчитывается по следующей формуле:

$$Z_{\text{осн}} = Z_{\text{он}} \cdot T_p, \quad (3.7)$$

где $Z_{\text{он}}$ – среднедневная заработная плата, руб.; T_p – продолжительность работ, выполняемых работником, раб.дн. (таблица 4).

Среднедневная заработная плата рассчитывается по формуле:

Для шестидневной рабочей недели (рабочая неделя научного руководителя):

$$Z_{он} = \frac{Z_m \cdot M}{F_{\delta}} = \frac{51285 \cdot 10,3}{246} = 2147,3 \text{ руб.}, \quad (3.8)$$

где Z_m – месячный должностной оклад работника, руб.; F_{δ} – действительный годовой фонд рабочего времени научно-технического персонала, раб. дней; M – количество месяцев работы без отпуска в течение года:

- при отпуске в 28 раб. дня – $M = 11,2$ месяца, 5-дневная рабочая неделя;
- при отпуске в 56 раб. дней – $M = 10,3$ месяца, 6-дневная рабочая неделя.

Для пятидневной рабочей недели (рабочая неделя разработчика):

$$Z_{он} = \frac{Z_m \cdot M}{F_{\delta}} = \frac{33150 \cdot 11,2}{213} = 1743,1 \text{ руб.}. \quad (3.9)$$

Должностной оклад работника за месяц:

- для научного руководителя:

$$Z_m = Z_{мс} \cdot (1 + k_{np} + k_{\delta}) k_p = 26300 \cdot (1 + 0,3 + 0,2) \cdot 1,3 = 51285 \text{ руб.} \quad (3.10)$$

- для разработчика:

$$Z_m = Z_{мс} \cdot (1 + k_{np} + k_{\delta}) k_p = 17000 \cdot (1 + 0,3 + 0,2) \cdot 1,3 = 33150 \text{ руб.}, \quad (3.11)$$

где $Z_{мс}$ – заработная плата, согласно тарифной ставке, руб.; k_{np} – премиальный коэффициент, равен 0,3; k_{δ} – коэффициент доплат и надбавок, равен 0,2; k_p – районный коэффициент, равен 1,3 (для г. Томска).

В таблице 7 представлены показатели рабочего времени, исходя от вида рабочей недели. В таблице 8 представлены расчеты основной заработной платы рабочей группы.

Таблица 7 – Баланс рабочего времени исполнителей

Показатели рабочего времени	Научный руководитель	Разработчик
Календарное число дней	365	365
Количество нерабочих дней	52/14	104/14
- выходные дни		
- праздничные дни		
Потери рабочего времени	48/5	24/10
- отпуск		
- невыходы по болезни		
Действительный годовой фонд рабочего времени	246	213

Таблица 8 – Расчет основной заработной платы исполнителей

Исполнители НИ	$Z_{мс}, руб$	$k_{пр}$	k_{∂}	k_p	$Z_m, руб$	$Z_{он}, руб$	$T_p, раб.дн.$	$Z_{осн}, руб$
Научный руководитель	26300	0,3	0,2	1,3	51285	2147,3	15,8	33927,3
разработчик	17000	0,3	0,2	1,3	33150	1743,1	85,6	149209,4
Итого:								183136,7

Дополнительная заработная плата определяется по формуле:

– для руководителя:

$$Z_{доп} = k_{доп} \cdot Z_{осн} = 0,15 \cdot 33927,3 = 5089,1 руб. \quad (3.12)$$

– для инженера:

$$Z_{доп} = k_{доп} \cdot Z_{осн} = 0,15 \cdot 149209,4 = 22381,4 руб., \quad (3.13)$$

где $k_{доп}$ – коэффициент дополнительной заработной платы (принимается равным 0,15).

3.3.3. Отчисления во внебюджетные фонды (страховые отчисления)

Отчисления во внебюджетные фонды определяется по формуле:

– для научного руководителя:

$$Z_{внеб} = k_{внеб} \cdot (Z_{осн} + Z_{доп}) = 0,3 \cdot 39016 = 11704 руб. \quad (3.14)$$

– для инженера:

$$Z_{внеб} = k_{внеб} \cdot (Z_{осн} + Z_{доп}) = 0,3 \cdot 171590 = 51477 руб. \quad (3.14)$$

где $k_{внеб}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд ОМС и социальное страхование). Общая ставка взносов составляет в 2020 году – 30% (ст. 425, 426 НК РФ).

3.3.4. Накладные расходы

Накладные расходы учитывают прочие затраты: оплата услуг связи, электроэнергии, почтовые и телеграфные расходы и т.д. Их величина определяется по следующей формуле:

$$Z_{накл} = (\text{сумма статей}1 \div 5) \cdot k_{пр}, \quad (3.15)$$

где $k_{пр}$ – коэффициент, учитывающий накладные расходы.

Величина коэффициента накладных взята в размере 16%. Тогда отчисления по накладным расходам составит:

$$Z_{\text{накл}} = (63181 + 183136,7 + 14767,5 + 27\,470,5) \cdot 0,16 = 288\,555,7 \cdot 0,16 = 46\,168,912 \text{ руб.}$$

3.3.5. Бюджет НИР

На основании полученных данных по отдельным статьям затрат составляется калькуляция плановой себестоимости ВКР «Использование методов машинного обучения для составления оптимального портфеля ценных бумаг» по форме, приведенной в таблице 9. В таблице также представлено определение бюджета затрат двух конкурирующих проектов.

Таблица 9 – Группировка затрат по статьям

№	Наименование статьи	Сумма, руб.	Примечание
		Текущий Проект	
1	Материальные затраты НИР	-	-
2	Затраты на специальное оборудование	14767,5	Пункт 3.3.1
3	Затраты по основной заработной плате исполнителей темы	183136,7	Пункт 3.3.2
4	Затраты по дополнительной заработной плате исполнителей темы	27470,5	-
5	Отчисления во внебюджетные фонды	63181	Пункт 3.3.3
6	Накладные расходы	46 168,912	Пункт 3.3.4
Бюджет затрат НИР		334 724,61	

3.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Эффективность проекта будем оценивать при помощи расчёта интегрального показателя эффективности научного исследования, нахождение которого связано с нахождением двух средневзвешанных величин:

- Финансовая эффективность
- Ресурсоэффективность

Т.к. конкурентными решениями в работе установлены индекс ММВБ - 10 и банковские вклады, стоимость вложений в создание которых невозможно узнать или оценить. Тогда эффективность проекта можно установить, используя Интегральный показатель эффективности и Сравнительная эффективность вариантов исполнения. Для конкретики, допустим, что затраты на все три варианта являются одинаковыми, хотя, скорее всего, затраты на работу целой команды на постоянную ребалансировку портфеля, расчёт необходимых процентов по вкладам, превышают затраты, необходимые для реализации НИП.

Тогда для любых значений i получим $I_{финр}^{исп.i} = \frac{\Phi_{pi}}{\Phi_{max}} = 1$, так как было допущено что затраты на создание и поддержание портфеля по индексу ММВБ - 10, а также содержание вклада, равны затратам по НИП.

Таблица 10 – Сравнительная оценка характеристик вариантов НИР

Объект исследования Критерии	Весовой коэффициент параметра	Банковский вклад	Разработанный портфель	Индекс ММВБ-10
1. Доходность	0,5	2	4	4
2. Диверсификация риска	0,2	3	5	2
3. Надежность	0,1	5	5	5
4. Ликвидность	0,1	5	5	3
5. Государственное регулирование	0,1	5	5	2
ИТОГО	1	20	24	16

Расчет интегрального показателя для разрабатываемого проекта:

$$I_{p1} = 0,5 * 2 + 0,2 * 3 + 0,1 * 5 + 0,1 * 5 + 0,1 * 5 = 3,1$$

$$I_{p2} = 0,5 * 4 + 0,2 * 5 + 0,1 * 5 + 0,1 * 5 + 0,1 * 5 = 4,5$$

$$I_{p3} = 0,5 * 4 + 0,2 * 2 + 0,1 * 5 + 0,1 * 3 + 0,1 * 2 = 3,4$$

Интегральный показатель эффективности вариантов исполнения разработки вычисляется на основании показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп.i} = \frac{I_{p-исп.i}}{I_{финр}^{исп.i}} \quad (3.16)$$

Очевидно, что, т.к. все знаменатели равны 1 из - за допущений, которые были указаны ранее. Таким образом $I_{исп.i} = I_{pi}$

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволит определить сравнительную эффективность проекта (таблицы 11) и выбрать наиболее целесообразный вариант из предложенных. Сравнительная эффективность проекта (\mathcal{E}_{cp}):

$$\mathcal{E}_{cp} = \frac{I_{исп.1}}{I_{исп.2}}$$

Таблица 11 – Сравнительная эффективность разработки

№ п/п	Показатели	Банковский вклад	Разработанный портфель	Индекс ММВБ-10
1	Интегральный финансовый показатель разработки	1	1	1
2	Интегральный показатель ресурсоэффективности разработки	3,1	4,5	3,4
3	Интегральный показатель эффективности	3,1	4,5	3,4
4	Сравнительная эффективность вариантов исполнения	0,78	1,11	1,42

Проанализировав сравнительную эффективность разработки, можно сделать вывод о том, что данный проект является более эффективным, в сравнении с другими конкурентами.

Выводы по разделу

В ходе выполнения части работы по финансовому менеджменту, ресурсоэффективности и ресурсосбережению был проведен анализ разрабатываемого исследования.

Оценен коммерческий потенциал и перспективность проведения исследования. Полученные результаты говорят о потенциале и перспективности на уровне выше среднего.

Проведено планирование НИР, а именно: определена структура и календарный план работы, трудоемкость и бюджет НТИ. Результаты соответствуют требованиям.

Для оценки затрат на реализацию проекта разработан проектный бюджет, который составляет 334 724,61 руб;

Определена эффективность исследования в разрезах ресурсной, финансовой, бюджетной, социальной и экономической эффективности.

4. Социальная ответственность

Введение

Объектом разработки данной ВКР является программный код, который использует методы машинного обучения для составления оптимального портфеля. Машинное обучение позволяет решать различные задачи при помощи обучения на большом количестве структурированных данных. При помощи алгоритмов и выявлении закономерностей поставленная задача может быть решена.

Проект выполняется на персональном компьютере (ПК), поэтому в данном разделе проводится анализ опасных и вредных факторов при работе с ПК, влияния этих факторов на окружающую среду и мероприятий по её защите.

Предметом исследования является рабочая зона разработчика, включая компьютерный стол, ПК, клавиатуру, компьютерную мышь и стул. Работы выполнялись в компьютерном классе 427А 10 корпуса ТПУ.

4.1. Правовые и организационные вопросы обеспечения безопасности

Специальная оценка условий труда является единым комплексом последовательно осуществляемых мероприятий по идентификации вредных и (или) опасных факторов производственной среды и трудового процесса и оценке уровня их воздействия на работника с учетом отклонения их фактических значений от установленных уполномоченным Правительством Российской Федерации федеральным органом исполнительной власти нормативов (гигиенических нормативов) условий труда и применения средств индивидуальной и коллективной защиты работников [16]. По результатам проведения специальной оценки условий труда устанавливаются классы (подклассы) условий труда. Согласно трудовому кодексу РФ работникам высшей категории (офисные работники) предусматриваются:

1. Понятие рабочего времени. Нормальная продолжительность рабочего времени [ТК РФ Статья 91]: для работников, условия труда на рабочих местах, которых по результатам специальной оценки условий труда отнесены к оптимальным

– нормальная продолжительность рабочего времени не может превышать 40 часов в неделю.

2. Обеспечение нормальных условий труда для выполнения норм выработки [ТК РФ Статья 163];

3. Предоставления работникам предусмотренных ТК РФ выходных и праздничных дней [ТК РФ г. 18], а также оплачиваемые отпуска [ТК РФ г. 19];

4. Установления работникам предусмотренных ТК РФ гарантий и компенсаций [ТК РФ г. 28].

Разработка программного обеспечения происходит за компьютерным столом. Рабочее место разработчика должно быть организовано согласно требованиям, представленным в «Типовая инструкция по охране труда для пользователей персональными электронно-вычислительными машинами (ПЭВМ) в электроэнергетике» [17]:

1) При размещении рабочих мест с ПЭВМ расстояние между рабочими столами с видеомониторами, должно быть не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов - не менее 1,2 м;

2) Рабочие места с ПЭВМ в помещениях с источниками вредных производственных факторов должны размещаться в изолированных кабинках с организованным воздухообменом;

3) Экран видеомонитора должен находиться от глаз пользователя на расстоянии 600-700 мм, но не ближе 500 мм с учетом размеров алфавитно-цифровых знаков и символов;

4) Конструкция рабочего стола должна обеспечивать оптимальное размещение на рабочей поверхности используемого оборудования. При этом допускается использование рабочих столов различных конструкций, отвечающих современным требованиям эргономики;

5) Конструкция рабочего стула (кресла) должна обеспечивать поддержание рациональной рабочей позы при работе на ПЭВМ, позволять изменять позу с целью снижения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления.

При разработке алгоритма для формирования оптимального графика ППП было предоставлено рабочее место, где соблюдены все требования по организации труда с ЭВМ.

4.2. Производственная безопасность

При разработке алгоритма разработчики подвергаются воздействию различных вредных и опасных факторов, которые представлены в таблице 12. В таблице также представлены соответствующие нормативные документы и этапы работ, во время которых разработчики могут столкнуться с их влиянием.

Таблица 12 – 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» [18]
Недостаточная освещённость рабочей зоны	СП 52.13330.2016 «Естественное и искусственное освещение» [19]
Повышенная световая и цветовая контрастность	СП 52.13330.2016 «Естественное и искусственное освещение» [19]
Повышенный уровень шума на рабочем месте	ГОСТ 12.1.003-2014 ССБТ. Шум. Общие требования безопасности [20] СП 51.13330.2011 «Защита от шума. Актуализированная редакция СНиП 23-03-2003» [31]
Длительное сосредоточенное наблюдение	МР 2.2.9.2311 – 07 «Профилактика стрессового состояния работников при различных видах профессиональной деятельности» [21]
Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий	ГОСТ Р 58698-2019 «Защита от поражения электрическим током» [23]

По данной таблице можно сделать вывод, что на разработчиков алгоритма в ходе их деятельности воздействуют только физические и психологические факторы, а химические и биологические факторы отсутствуют.

4.2.1. Отклонение показателей микроклимата

Отклонение показателей микроклимата на рабочем месте от комфортных непосредственно влияет на здоровье работников. Повышение скорости движения воздуха и понижение температуры может привести к переохлаждению организма путем усиления теплообмена и процесса теплоотдачи при испарении пота. Недостаточная влажность в свою очередь ведет к интенсивному испарению влаги со слизистых оболочек. Это может привести к пересыханию, растрескиванию и затем к заражению болезнетворными бактериями. При разработке алгоритма используются персональные компьютеры, которые могут непосредственно влиять на микроклимат путем снижения относительной влажности и повышению температуры в рабочем помещении.

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

Работа, выполняемая командой разработки программного обеспечения, по энергозатратам относится к категории Ia (производится сидя, сопровождается незначительными физическими усилиями). В таблицах 13 и 14 представлены оптимальные и допустимые значения показателей микроклимата на рабочих местах для данной категории.

Таблица 13 – 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 14 – 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

В производственных помещениях, где поддерживать допускаемые нормативные величины локального микроклимата не представляется возможным, необходимо проводить мероприятия по защите работников от возможного перегревания и охлаждения. Это достигается разными способами: использование систем местного кондиционирования воздуха; регламентацией периодов работы в неблагоприятном локальном микроклимате и отдыха в помещении с микроклиматом, нормализующим тепловое состояние; уменьшение длительности рабочей смены и др.

4.2.2. Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны является вредным производственным фактором, приводящим к повышенной утомляемости и снижению работоспособности человека на предприятии. Продолжительная работа в условиях низкой освещенности приводит к ухудшению зрения.

Нормы естественного, искусственного и совместного освещения регламентируются СП 52.13330.2016 «Естественное и искусственное освещение». Разработка программного обеспечения 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 15 представлены требования к освещению рабочего помещения для разряда Б1.

Таблица 15 – 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

Яркий свет в зоне периферийного зрения заметно увеличивает глазное напряжение. Для снижения влияния вредного фактора недостаточной освещенности необходимо, чтобы уровень естественного освещения рабочего пространства приблизительно совпадал с яркостью дисплея. Проблему недостаточной освещенности помещения можно решить при помощи установки дополнительных осветительных приборов, расширения световых проемов.

4.2.3. Повышенная световая и цветовая контрастность

Отклонение светового и цветового контраста на рабочем месте приводит к быстрому утомлению и снижению уровня работоспособности человека на предприятии. Продолжительное воздействие этого вредного фактора приводит к возникновению проблем со зрением. Нормы светового и цветового контраста регламентируются СП 52.13330.2016 «Естественное и искусственное освещение». Для работы за компьютером (категория работ Б1) нормы контраста представлены в таблице 16.

Таблица 16 – Требования к освещению рабочего помещения

Характеристика зрительной работы	Контраст объекта с фоном	Характеристика фона
Высокой точности	Малый	Средний
	Средний	Темный

Для изменения светового и цветового контраста необходимо отрегулировать уровень естественной и искусственной освещенности рабочего помещения или заменить текущее оборудование (мониторы) на более качественные, которые позволят сгладить контраст.

4.2.4. Повышенный уровень шума на рабочем месте

Превышение уровня шума на рабочем месте создает психологический и физический стресс, снижающий производительность, концентрацию, внимание, повышает утомляемость. Повышение уровня шума на рабочем месте возможно из-за фона, создаваемого работой персональных компьютеров, наличия центральной системы вентиляции и кондиционирования воздуха.

Предельно допустимые показатели уровня звука, звукового давления регламентируются СП 51.13330.2011. Защита от шума. Актуализированная редакция СНиП 23-03-2003. Для команды разработчиков программного обеспечения, эти показатели представлены в таблице 17.

Таблица 17 – Предельно допустимые и допустимые уровни звукового давления, уровни звука, эквивалентные и максимальные скорректированные по А уровни звука в помещениях производственных, жилых, общественных зданий и на территории жилой застройки СП 51.13330.2011. Защита от шума. Актуализированная редакция СНиП 23-03-2003 [31].

Назначение помещений или территорий	Для источников постоянного шума										Для источников непостоянного шума		
	Уровни звукового давления дБ, в октавных полосах частот со среднегеометрическими частотами, Гц										Уровень звука скорректированный по А, дБ	Эквивалентный скорректированный по А уровень звука $L_{A,эв}$, дБ	Максимальный скорректированный по А уровень звука $L_{A,макс}$, дБ
	Время суток, ч												
Помещения офисов, рабочие помещения и кабинеты административных зданий, научно-исследовательских организаций	86	31,5											
	71	63											
	61	125											
	54	250											
	49	500											
	45	1000											
	42	2000											
	40	4000											
	38	8000											
	50												
50													
65													

Для снижения уровня шума в производственном помещении можно использовать защитные звукопоглощающие экраны. Для любого оборудования необходимо регулярно проводить техническое обслуживание, так как загрязнение может увеличить производимый шум.

4.2.5. Длительное сосредоточенное наблюдение

При разработке алгоритма необходим контроль процесса формирования популяций и значений целевых функций, который вызывает зрительную и умственную нагрузку на организм человека.

При умственной нагрузке необходима длительность сосредоточенного внимания, выраженная ответственность, плотность сигналов и сообщений в единицу времени по МР 2.2.9.2311 – 07 «Профилактика стрессового состояния работников при различных видах профессиональной деятельности» [21]. Оказывает угнетающее влияние на психическую деятельность ухудшаются функции внимания (объем, концентрация, переключение), памяти (кратковременной и долговременной), восприятия (появляется большое число ошибок).

При зрительной нагрузке необходима высокая координация сенсорных и моторных элементов зрительной системы. Вызывает головную боль, ухудшение зрения, астенопию – патологического состояния, связанного с быстрым переутомлением глаз.

Для устранения накопленной усталости и нагрузки на организм человека необходимо выполнять комплекс физических упражнений на координацию движений, концентрацию внимания, комплекс упражнений на глаз, использовать методику психической саморегуляции.

4.2.6. Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий

Вследствие работы с электрооборудованием и компьютерами возникает вероятность прохождения электрического тока через тело человека. Опасное и вредное воздействия на людей электрического тока проявляются в виде электротравм (судороги, остановка сердца, остановка дыхания, ожоги и др.) и заболеваний. Результат воздействия тока на человека зависит от величины силы тока, его рода и

частоты, продолжительности воздействия и множества других факторов. Причиной поражения электрическим током в условиях лаборатории могут стать случайное прикосновение к токоведущим частям или появление напряжения на металлических частях оборудования.

Под электробезопасностью подразумевается система технических и организационных мероприятий, направленных на защиту людей от опасного воздействия электрического тока, статического электричества и электромагнитного поля. Значения вышеперечисленных факторов регулируются ГОСТ Р 58698-2019.

Таблица 18 – 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;
- размещение опасных для жизни и здоровья человека участков электропроводов и приборов вне зоны досягаемости рукой;
- ограничение напряжения или питание должно осуществляться от безопасного источника питания;
- автоматическое отключение питания (защитное устройство, которое будет отключать систему, питающую электрическое оборудование в случае замыкания)

Защита от поражения электрическим током может осуществляться посредством системы безопасного сверхнизкого напряжения (БСНН) и защитного сверхнизкого напряжения (ЗСНН).

4.3. Экологическая безопасность

Программное обеспечение не оказывает влияния на окружающую среду, так как его разработка и использование происходит при помощи персональных

компьютеров, однако использование самого компьютера может оказывать влияние на окружающую среду.

ПЭВМ состоит из различных деталей, которые имеют различный класс опасности. Ртутные лампы, материнские платы и аккумуляторы содержат свинец, литий, кадмий; материнские платы содержат олово; трансформаторы содержат обмотку из меди, а также железный сердечник; преобразователи и проводка содержат алюминий магнит; корпус состоит из металла и пластика. Класс опасности и предельно-допустимая концентрация указаны в таблице 19 [32].

Таблица 19 – Предельно-допустимые концентрации (ПДК) вредных веществ в воздухе рабочей зоны

Наименование вещества	Формула	Величина ПДК, мг/м ³	Класс опасности
Ртуть	Hg	0,01/0,05	I
Литий и его неорганические соединения		0,02	I
Свинец и его неорганические соединения		-/0,05	I
Олово фторид	FSn	1/0,2	II
Железо	Fe	-/10	IV
Полиэтиленхлорид	[C ₂ H ₃ Cl] _x	6	III
Алюминий магнит	AlMg	-/6	IV

Для оказания наименьшего влияния на окружающую среду, необходимо проводить специальную процедуру утилизации ПЭВМ и оргтехники, при которой разбирается, сортируется и более 90% отправится на вторичную переработку и менее 10% будут отправлены на свалки. При этом она должна соответствовать процедуре утилизации ГОСТ Р 53692-2009 «Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов» [25].

4.4. Безопасность в чрезвычайных ситуациях

Причинами возникновения пожара при работе с ЭВМ может служить короткое замыкание проводки, в том числе в следствии неисправности прибора, сильный перегрев ЭВМ в результате его использования в режиме повышенной нагрузки.

Для предотвращения возникновения пожара, необходимо проводить периодическую своевременную диагностику оборудования и электрической проводки, соблюдать нормы при работе с ЭВМ.

На основании Федерального закона от 22.07.2008 N 123-ФЗ (ред. От 30.04.2021) "Технический регламент о требованиях пожарной безопасности" помещения учебной аудитории оборудованы следующими средствами пожаротушения: огнетушитель ручной углекислотный ОУ-5, пожарный кран с рукавом, также аудитория оборудована системой противопожарной сигнализации [27]. Помещение аудитории категории помещения группы – В4, возможный класс пожара – Е. [27].

При появлении возгорания необходимо сообщить в службу пожарной охраны адрес и место возникновения пожара.

Заключение по разделу

В данном разделе были рассмотрены правовые и организационные вопросы обеспечения безопасности, аспекты производственной и экологической безопасности, а также рассмотрен вопрос безопасности рабочей группы при чрезвычайном происшествии. В результате можно сделать вывод о работнике:

- Имеет нормальную продолжительность рабочего времени;
- Обеспечен нормальными условиями труда для выполнения норм выработки;
- Имеет, предусмотренные ТК РФ, выходные и праздничные дни;
- Имеет категорию тяжести труда - Ia (производится сидя, сопровождается незначительными физическими усилиями), согласно [18];
- Относится к первой группе по электробезопасности [29];

А также о рабочей зоне:

- Рабочая зона имеет оптимальные значения показаний микроклимата согласно таблице 13;
- Фактические показатели освещения совпадают с оптимальными в таблице 15;
- Мониторы ЭВМ имеют оптимальную световую и цветовую контрастность;
- Соблюдаются меры предосторожности для основной защиты от поражения электрическим током;
- Вышедшие из строя ПЭВМ утилизируются согласно ГОСТ Р 53692-2009;
- Помещения учебной аудитории оборудованы положенными средствами пожаротушения;
- Рабочая зона имеет категорию помещения по электробезопасности согласно ПУЭ – без повышенной опасности [28];
- Помещение аудитории категории помещения группы – В4 (пожароопасное);
- Помещение аудитории находится в объекте, относящийся к IV категории по оказанию негативного воздействия на окружающую среду [30].

Заключение

В результате выполнения выпускной квалификационной работы были получены следующие основные результаты:

1. Классическим методом Марковица сформирован оптимальный портфель ценных бумаг с максимальной ожидаемой доходностью и заданным уровнем волатильности 30% годовых;

2. Для сравнения полученного результата было создано три оптимальных рискованных портфеля с помощью нейросетевых алгоритмов "Случайный лес", "сеть с долгосрочно-краткосрочной памятью" (LSTM), "Дерево решений". При этом для построения портфеля использовался итерационный квазиньютоновский метод нелинейной оптимизации с целевой функцией, отличной от классической, выбранной Марковицем;

3. На исторических данных второго полугодия 2022 года нейросетевые алгоритмы "Случайный лес" и "Дерево решений" позволили построить портфели с наблюдаемой в 2023 году доходностью 6,2% и 6,4% в месяц, или 75% и 72% годовых соответственно. Портфель Марковица показал доходность в 5,6% в месяц, или 66% годовых, а нейросетевой алгоритм LSTM сформировал портфель с доходностью в 4,4%, или 52% годовых. Рост индекса ММВБ-10 составил за этот же период 2,9%, или 33% годовых.

Список литературы

1. Буренин А.Н. Управление портфелем ценных бумаг/ А.Н. Буренин. -Москва: Научно-техническое общество имени академика С.И. Вавилова, 2008. - 440 с.
2. Дорохина А.С. "Формирование оптимального по Марковицу портфеля ценных бумаг" г. Томск: Изд-во ТПУ, 2008. — С. 243-244.
3. Мицель А.А. Основы финансовой математики: учебное пособие / А.А. Мицель. - Часть 2. – Томск: Национальный исследовательский Томский политехнический университет, 2016. – 102с.
4. Уильям Ф. Шарп, Гордон Дж. Александер, Джеффри В. Бэйли. Инвестиции. – М.:Инфра-М.-2006-1035 с.
5. Markowitz Harry. Portfolio Selection // Journal of Finance.-1952.-Vol. 7.- № 1. - pp. 71-91.
6. Кравченко И. М, Постников М.М. Криптовалюта: роль в современном мире//Московский Государственный Университет им. М.В. Ломоносова – Москва 2017 г.
7. О. А. Бельснер, О. Л. Крицкий "Оптимизация портфеля финансовых инструментов" г. Томск: Изд-во ТПУ, - С. 2-56
8. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32.
9. Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783-2792.
10. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.
11. Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. Neural Computation, 12(10), 2451-2471.
12. Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 2222-2232.
13. Mitchell, T. (1997). Machine Learning. McGraw-Hill.
14. Quinlan, J.R. (1986). Induction of Decision Trees. Machine Learning, 81-106.

15. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.

16. 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: 12.03.2023);

17. РД 153-34.0-03.298-2001 Типовая инструкция по охране труда для пользователей персональными электронно-вычислительными машинами (ПЭВМ) в электроэнергетике // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200031404> (дата обращения: 12.03.2023);

18. 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: 12.03.2023);

19. 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: 12.03.2023);

20. СН 2.2.4/2.1.8.562-96 «Шум. Общие требования безопасности» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200118606> (дата обращения: 12.03.2023);

21. МР 2.2.9.2311-07 «Состояние здоровья работающих в связи с состоянием производственной среды» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200072234> (дата обращения: 12.03.2023);

22. ГОСТ 12.1.005-88 «Общие санитарно-гигиенические требования к воздуху

рабочей зоны» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200003608> (дата обращения: 12.03.2023);

23. 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: 12.03.2023).

24. Трудовой кодекс (ТК РФ) «Рабочее время»// Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <http://base.garant.ru/12125268/> (дата обращения: 12.03.2023);

25. ГОСТ Р 53692-2009 «Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200081740> (дата обращения: 12.03.2023);

26. СП 12.13130.2009 Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности;

27. Федеральный закон от 22 июля 2008 г. №123-ФЗ «Технический регламент о требованиях пожарной безопасности» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/902111644> (дата обращения: 12.03.2023);

28. «Правила устройства электроустановок» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200030216> (дата обращения: 12.03.2023);

29. Приказ «Об утверждении Правил по охране труда при эксплуатации электроустановок» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573264184> (дата обращения: 12.03.2023);

30. Постановление «Об утверждении критериев отнесения объектов, оказывающих негативное воздействие на окружающую среду, к объектам I, II, III и IV категорий» // Электронный фонд правовой и нормативно-технической

документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573292854> (дата обращения: 12.03.2023);

31. СП 51.13330.2011 «Защита от шума. Актуализированная редакция СНиП 23-03-2003» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200084097> (дата обращения: 12.03.2023);

32. ГН 2.2.5.3532-18 "Предельно допустимые концентрации (ПДК) вредных веществ в воздухе рабочей зоны" // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/557235236> (дата обращения: 12.03.2023);

Приложение 1. Программный код для составления оптимального портфеля с помощью алгоритма Случайный лес

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from scipy.optimize import minimize
from sklearn.ensemble import RandomForestRegressor

data = pd.read_csv('Книга3.csv')
assets = data.columns[0:]
prices = data[assets].values
returns = np.diff(prices, axis=0) / prices[:-1]

train_size = int(0.8 * len(returns))
train_returns = returns[:train_size]
test_returns = returns[train_size:]

model = RandomForestRegressor(n_estimators=3, random_state=5)

X_train = train_returns[:-1]
y_train = train_returns[1:]
model.fit(X_train, y_train)

X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

predicted_returns = []
for i in range(180):
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

    current_returns = np.append(current_returns[1:], predicted_return)

actual_returns = test_returns[1:] # Фактические доходности
mae = mean_absolute_error(actual_returns, predicted_returns)
mse = mean_squared_error(actual_returns, predicted_returns)

print("MAE:", mae)
print("MSE:", mse)

def portfolio_performance(weights, returns):
    portfolio_returns = np.dot(weights, returns.mean(axis=0))
    portfolio_risk = np.sqrt(np.dot(weights.T, np.dot(np.cov(returns.T), weights)))
    return portfolio_returns, portfolio_risk

def minimize_risk(weights, returns):
    portfolio_returns, portfolio_risk = portfolio_performance(weights, returns)
    return portfolio_risk

constraints = ({'type': 'eq', 'fun': lambda x: np.sum(x) - 1})
bounds = tuple((0, 1) for _ in assets)
initial_guess = np.array([1 / len(assets) for _ in assets])

optimized_portfolio = minimize(minimize_risk, initial_guess,
                               args=(predicted_returns,),
                               method='SLSQP', bounds=bounds,
                               constraints=constraints)

optimal_weights = optimized_portfolio.x

for asset, weight in zip(assets, optimal_weights):
    print(f"{asset}: {weight}")
```

Приложение 2. Программный код для составления оптимального портфеля с помощью алгоритма LSTM

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from scipy.optimize import minimize
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeRegressor

data = pd.read_csv('Книга3.csv')

assets = data.columns[0:]

prices = data[assets].values

returns = np.diff(prices, axis=0) / prices[:-1]

train_size = int(0.8 * len(returns))
train_returns = returns[:train_size]
test_returns = returns[train_size:]

model = DecisionTreeRegressor()

model.fit(train_returns[:-1], train_returns[1:])

X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

predicted_returns = []
for i in range(180):
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

    current_returns = np.append(current_returns[1:], predicted_return)

def portfolio_performance(weights, returns):
    portfolio_returns = np.dot(weights, returns.mean(axis=0))
    portfolio_risk = np.sqrt(np.dot(weights.T, np.dot(np.cov(returns.T), weights)))
    return portfolio_returns, portfolio_risk

def minimize_risk(weights, returns):
    portfolio_returns, portfolio_risk = portfolio_performance(weights, returns)
    return portfolio_risk

constraints = ({'type': 'eq', 'fun': lambda x: np.sum(x) - 1})
bounds = tuple((0, 1) for _ in assets)
initial_guess = np.array([1 / len(assets) for _ in assets])

optimized_portfolio = minimize(minimize_risk, initial_guess,
                               args=(predicted_returns,),
                               method='SLSQP', bounds=bounds,
                               constraints=constraints)

optimal_weights = optimized_portfolio.x

for asset, weight in zip(assets, optimal_weights):
    print(f"{asset}: {weight}")
```

Приложение 3. Программный код для составления оптимального портфеля с помощью алгоритма Дерево решений

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from scipy.optimize import minimize
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeRegressor

data = pd.read_csv('Книга3.csv')

assets = data.columns[0:]

prices = data[assets].values

returns = np.diff(prices, axis=0) / prices[:-1]

train_size = int(0.8 * len(returns))
train_returns = returns[:train_size]
test_returns = returns[train_size:]

model = DecisionTreeRegressor()

model.fit(train_returns[:-1], train_returns[1:])

X_test = test_returns[:-1] # Тестовые данные
current_returns = test_returns[-1] # Последняя доступная доходность

predicted_returns = []
for i in range(180):
    predicted_return = model.predict([current_returns])
    predicted_returns.append(predicted_return[0])

    current_returns = np.append(current_returns[1:], predicted_return)

def portfolio_performance(weights, returns):
    portfolio_returns = np.dot(weights, returns.mean(axis=0))
    portfolio_risk = np.sqrt(np.dot(weights.T, np.dot(np.cov(returns.T), weights)))
    return portfolio_returns, portfolio_risk

def minimize_risk(weights, returns):
    portfolio_returns, portfolio_risk = portfolio_performance(weights, returns)
    return portfolio_risk

constraints = ({'type': 'eq', 'fun': lambda x: np.sum(x) - 1})
bounds = tuple((0, 1) for _ in assets)
initial_guess = np.array([1 / len(assets) for _ in assets])

optimized_portfolio = minimize(minimize_risk, initial_guess,
                               args=(predicted_returns,),
                               method='SLSQP', bounds=bounds,
                               constraints=constraints)

optimal_weights = optimized_portfolio.x

for asset, weight in zip(assets, optimal_weights):
    print(f"{asset}: {weight}")
```