

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

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.03.04 «Программная инженерия»
Отделение информационных технологий

БАКАЛАВРСКАЯ РАБОТА

Тема работы
Разработка системы диагностики и предсказания уровня заболеваемости COVID-19 в странах мира

УДК 004.65:004.451:004.774:616.9

Студент

Группа	ФИО	Подпись	Дата
8K81	Семенова Ольга Юрьевна		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР ТПУ	Марухина Ольга Владимировна	к.т.н., доцент		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОСГН	Гасанов Магеррам Али оглы	д.э.н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент ООД ШБИП	Мезенцева Ирина Леонидовна			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Чердынцев Евгений Сергеевич	к.т.н.		

Томск – 2022 г.

Планируемые результаты обучения по ООП

Код компетенции	Наименование компетенции
Общекультурные (универсальные) компетенции	
У-1	Способна осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач
У-2	Способна определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений
У-3	Способна эффективно работать индивидуально и в качестве члена группы, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
У-4	Способна использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности
У-5	Способна воспринимать межкультурное разнообразие общества в социально-историческом, этическом и философском контекстах
У-6	Способна управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни
У-7	Способна поддерживать должный уровень физической подготовленности для обеспечения полноценной социальной и профессиональной деятельности;
У-8	Способна создавать и поддерживать безопасные условия жизнедеятельности, в том числе при возникновении чрезвычайных ситуаций
Профессиональные компетенции	
П-1	Применять базовые и специальные естественнонаучные и математические знания в области информатики и вычислительной техники, достаточные для комплексной инженерной деятельности.
П-2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
П-3	Ставить и решать задачи комплексного анализа, связанные с созданием аппаратнопрограммных средств информационных и автоматизированных систем, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
П-4	Разрабатывать программные и аппаратные средства (системы, устройства, блоки, программы, базы данных и т.п.) в соответствии с техническим заданием и с использованием средств автоматизации проектирования.
П-5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретацию полученных данных, в области создания аппаратных и программных средств информационных и автоматизированных систем.
П-6	Внедрять, эксплуатировать и обслуживать современные программно-аппаратные комплексы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.

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

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.03.04 «Программная инженерия»
Отделение информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП
_____ Чердынцев Е.С. _
(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Бакалаврской работы

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8K81	Семеновой Ольге Юрьевне

Тема работы:

Разработка системы диагностики и предсказания уровня заболеваемости COVID-19 в странах мира

Утверждена приказом директора (дата, номер)	№119-50/с от 29.04.2022 г.
---	----------------------------

Срок сдачи студентом выполненной работы:	06.06.2022 г.
--	---------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Объектом проектирования в данной работе является веб-приложение для визуализации уровня заболеваемости COVID-19 в странах мира.
---------------------------------	---

Перечень подлежащих исследованию, проектированию и разработке вопросов	1. Обзор предметной области 2. Выбор средств разработки 3. Анализ и подготовка исходных данных для обработки 4. Разработка системы 5. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение 6. Социальная ответственность
Перечень графического материала	1. Фрагменты кода Python 2. Рисунки, демонстрирующие результаты 3. Матрица SWOT-анализа
Консультанты по разделам выпускной квалификационной работы	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Гасанов Магеррам Али оглы
Социальная ответственность	Мезенцева Ирина Леонидовна

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	15.02.2022
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Марухина О.В.	к.т.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8K81	Семенова Ольга Юрьевна		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
8K81	Семенова Ольга Юрьевна

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.04 Программная инженерия

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

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад руководителя – 45000 руб. Оклад разработчика – 21000 руб.
2. Нормы и нормативы расходования ресурсов	Дополнительной заработной платы 15%; Накладные расходы 15%; Районный коэффициент 1,3.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Тариф отчислений во внебюджетные фонды 7,6%

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	Определение потенциального потребителя результатов исследования. Анализ конкурентных технических решений. SWOT-анализ разработанной стратегии
2. Планирование и формирование бюджета научных исследований	Определение структуры работы. Расчет трудоемкости выполнения работ. Подсчет бюджета исследования
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	Рассчитать показатели финансовой эффективности, ресурсоэффективности и эффективности исполнения

Перечень графического материала:

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT
3. График проведения и бюджет НИ
4. Оценка ресурсной, финансовой и экономической эффективности НИ

Дата выдачи задания для раздела по линейному графику	15.02.2022
--	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОСТН	Гасанов Магеррам Али оглы	Д.Э.Н.		15.02.2022

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8K81	Семенова Ольга Юрьевна		15.02.2022

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа		ФИО	
8К81		Семенова Ольга Юрьевна	
Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.04 Программная инженерия

Тема ВКР:

Разработка системы диагностики и предсказания уровня заболеваемости COVID-19 в странах мира	
Исходные данные к разделу «Социальная ответственность»:	
Введение – Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения. – Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации	Объект исследования – данные о коронавирусной инфекции в странах мира Область применения – система диагностики и предсказания уровня заболеваемости. Рабочая зона: <u>офис</u> Размеры помещения – 2*6 Количество и наименование оборудования рабочей зоны – 1 ноутбук Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне – проектирование и разработка
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения: – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны.	ГОСТ 12.2.032-78. Рабочее место при выполнении работ сидя. Рациональная организация труда в течение рабочего времени предусмотрена Трудовым Кодексом РФ ФЗ-197. ОСТ 21889-76. Система “Человек-машина”. Кресло человека-оператора. ТК РФ. СанПиН 1.2.3685-21. Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.
2. Производственная безопасность при разработке проектного решения: – Анализ выявленных вредных и опасных производственных факторов	Вредные факторы: -отсутствие или недостаток необходимого естественного и искусственного освещения; -зрительное напряжение; -повышенный уровень электромагнитного излучения; -отклонение показателей микроклимата; Опасные факторы: -повышенное значение напряжения в электрической цепи. Требуемые средства коллективной и индивидуальной защиты от выявленных факторов: защитные ограждения; устройства автоматического контроля и сигнализации; знаки безопасности; устройства вентиляции и кондиционирования воздуха.

3. Экологическая безопасность при разработке проектного решения	Воздействие на литосферу: бытовой мусор, ртутное загрязнение. Воздействие на гидросферу: бытовой мусор, ртутное загрязнение. Воздействие на атмосферу: ртутное загрязнение.
4. Безопасность в чрезвычайных ситуациях при разработке проектного решения	Возможные ЧС: пожар, наводнение, землетрясение, удар молнией, террористический акт. Наиболее типичная ЧС: пожар.

Дата выдачи задания для раздела по линейному графику	15.02.2022
---	-------------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Мезенцева Ирина Леонидовна			15.02.2022

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8K81	Семенова Ольга Юрьевна		15.02.2022

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

Инженерная школа информационных технологий и робототехники
Направление подготовки 09.03.04 «Программная инженерия»
Уровень образования бакалавриат
Отделение информационных технологий

Период выполнения осенний / весенний семестр 2021/2022 учебного года

Форма представления работы:

Бакалаврская работы

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

**КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы**

Срок сдачи студентом выполненной работы:

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
24.02.2021	Основная часть	75
20.05.2021	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	15
25.05.2021	Социальная ответственность	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Марухина Ольга Владимировна	к.т.н., доцент		

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Чердынцев Евгений Сергеевич	к.т.н.		

Реферат

Выпускная квалификационная работа содержит 75 страниц, 23 рисунка, 20 таблиц.

Ключевые слова: анализ данных, подготовка данных, машинное обучение, предсказательная модель.

Цель работы – разработка системы диагностики и предсказания уровня заболеваемости COVID-19 в странах мира.

В результате работы была реализована системы диагностики и предсказания уровня заболеваемости COVID-19 в странах мира.

В первом разделе представлено описание предметной области и анализ проблематики, а также поставлены задачи исследования.

Во втором разделе приводится подробное описание подготовки исходных данных для обработки.

В третьем разделе описывается разрабатываемая система и необходимые компоненты для реализации.

Пятый раздел представляет собой выполненное задание по разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение».

В шестом разделе содержится выполненное задание по разделу «Социальная ответственность»

Определения, обозначения, сокращения, нормативные ссылки

Библиотека – сборник подпрограмм или объектов, используемых для разработки программного обеспечения.

Выброс (англ. outlier) - в статистике результат измерения, выделяющийся из общей выборки.

Датасет – От англ. Dataset – набор данных.

Датафрейм (от англ. DataFrame) – объект библиотеки pandas, позволяющий работать с табличными данными.

ИС – информационная система, предназначена для хранения, поиска и обработки информации.

Big Data («большие данные») - обозначение структурированных и неструктурированных данных огромных объёмов и значительного многообразия, эффективно обрабатываемых горизонтально масштабируемыми программными инструментами.

CSV (от англ. Comma-Separated Values - значения, разделённые запятыми) - текстовый формат, предназначенный для представления табличных данных. Строка таблицы соответствует строке текста, которая содержит одно или несколько полей, разделённых запятыми.

Python – высокоуровневый язык программирования общего назначения.

Оглавление

Реферат	9
Определения, обозначения, сокращения, нормативные ссылки	10
Введение	14
Глава 1. Исследование предметной области	16
1.1 Описание предметной области	16
1.2 Анализ проблемы	17
1.3 Задачи исследования	18
1.4 Вывод по главе	20
Глава 2. Подготовка исходных данных для обработки	21
2.1. Методика подготовки исходных данных для построения моделей классификации	21
2.2. Выбор технологий	21
2.3. Предобработка данных	23
2.4. Получение сводной информации и признаках	26
2.5. Обработка выбросов	29
2.6. Обработка пропусков	30
2.7. Удаление дубликатов	32
2.8. Рассмотрение таблицы корреляции	32
2.9. Вывод по главе	33
Глава 3. Разработка системы предсказания	34
3.1. Метод k-средних	34
3.2. Обучение	35
3.3. Создание системы	36
3.4. Вывод по главе	37
Глава 4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	38
4.1 Потенциальные потребители результатов исследования	38
4.2 Анализ конкурентных технических решений	38
4.3 SWOT-анализ	40
	11

4.4 Планирование работ по научно-техническому исследованию	42
4.4.1 Структура работ в рамках научного исследования	42
4.4.2 Определение трудоемкости выполнения работ	43
4.4.3 Разработка графика проведения научного исследования	44
4.5 Бюджет научно-технического исследования (НТИ)	46
4.5.1. Расчет материальных затрат НТИ	46
4.5.2. Расчет затрат на специальное оборудование для научных работ	46
4.5.3 Основная заработная плата исполнителя темы	47
4.5.4 Расчет дополнительной заработной платы	48
4.5.5 Отчисления во внебюджетные фонды	49
4.5.6 Накладные расходы	49
4.5.7 Формирование бюджета затрат научно-исследовательского проекта	50
4.6 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	50
Вывод по главе	52
Глава 6. Социальная ответственность	54
Введение	54
6.1. Правовые и организационные вопросы обеспечения безопасности	55
6.1.1. Правовые нормы трудового законодательства	55
6.1.2. Эргономические требования к правильному расположению и компоновке рабочей зоны	56
6.2. Производственная безопасность	57
6.2.1. Отсутствие или недостаток необходимого естественного и искусственного освещения	57
6.2.2. Аномальные микроклиматические параметры воздушной среды на местонахождении рабочего	58
6.2.3. Нервно-психические перегрузки, связанные с монотонностью труда	59
6.2.4. Нервно-психические перегрузки, связанные с перенапряжением зрительных анализаторов	59

6.2.5. Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает рабочий	60
6.3. Экологическая безопасность	61
6.4. Безопасность в чрезвычайных ситуациях	62
Вывод по разделу	63
Заключение	65
Список литературы	66
Приложение А	68

Введение

11 марта 2020 года Всемирная организация здравоохранения (ВОЗ) объявила о пандемии нового коронавируса (Covid-19), который в декабре 2019 был впервые обнаружен в китайском мегаполисе Ухань. Вирус начал стремительно распространяться по всему миру, охватывая с каждым днем все больше и больше людей.

Вирус и по сей день распространяется с высокой скоростью, в связи с тем, что он передается воздушно-капельным путем. Чем больше зараженных вирусом Covid-19 людей, тем больше контактов, следовательно, и больше больных.

Очень важным фактором замедления распространения коронавирусной инфекции является соблюдение защитных мер. Большинство стран мира быстро отреагировало на появление вируса и ввело карантинные меры. Данные методы профилактики включали в себя, не только ограничения на выезд из страны, но и ограничения на передвижения по городу.

Данные защитные меры безусловно повлияли на жизнь людей. Одной из крупных отраслей экономики, пострадавших от введенных ограничений, стал туризм. Люди на какое-то время перестали путешествовать, чтобы обезопасить себя и окружающих от коронавирусной инфекции.

Защитные меры повлияли на процесс распространения вируса, и спустя некоторое время, люди вновь стали посещать разные страны. Но людям все еще стоит необходимо ответственно подходить к выбору мест, которые они хотят посетить, так как не во всех странах распространение вируса замедлилось. Анализ статистики коронавирусной инфекции занимает много времени, а также не может дать каких-либо прогнозов.

Таким образом, **целью данной работы** является обеспечение пользователей системой, визуализирующей предсказание уровня заболеваемости COVID-19 в странах мира.

Для достижения данной цели поставлены следующие **задачи**:

1. Анализ предметной области.
2. Подготовка данных для анализа.
3. Выделение значимых атрибутов и добавление новых.
4. Предобработка данных
5. Выбор алгоритма машинного обучения
6. Обучение моделей
7. Разработка системы, визуализирующей исследование

В результате ожидается система, позволяющая туристам, просматривать страны, которые они желают посетить. Данная система дает возможность путешествовать с минимальными рисками, вызванными коронавирусной пандемией. Для пользования системой пользователям не нужны никакие дополнительные навыки, так как итоговый веб-сайт будет представлять собой карту мира, при нажатии на какую-либо страну будет показываться информация по заражениям на данный момент, а также в виде светофора прогноз уровня заболеваемости.

Глава 1. Исследование предметной области

1.1 Описание предметной области

Big Data (англ. «большие данные») — массивы данных большого объёма и значительного многообразия, анализируемые с помощью специальных компьютерных технологий [1]. Важную роль играет не столько объем данных, сколько возможности анализа этих массивов. Исследование больших данных позволяет выявить скрытые закономерности, оценка которых в дальнейшем может привести к неожиданным результатам.

Анализ Big Data позволяет оптимизировать многие сферы жизни. Так, например, в технологии больших данных пользуются спросом в области страхования. Сегодня около 30% страховых компаний уже используют Big Data и аналитику для того, чтобы персонализировать общение с клиентами и предугадывать их поведение [2]. Оценка рисков — одно из главных преимуществ больших данных. Например, при расчете стоимости страхования автомобиля сейчас учитывается возраст водителей данного автомобиля, их стаж, количество аварий, а также от характеристики самой машины, такие как мощность, год выпуска и другие. Проанализировать все эти факторы и позволяет анализ больших данных. Большой популярностью технологии больших данных пользуются и в торговле через Интернет. Использование анализа данных позволяет составлять портрет покупателя и предлагать необходимые ему товары. Многие компании используют данные сторонних источников, также анализируют данные просмотренных товаров или уже приобретенных и на основании этих данных рекомендуют подходящие конкретному пользователю предложения. В отрасли финансов анализ Big Data также пользуется спросом у банков. Они используют данные технологии в качестве оценки кредитоспособности лица, так называемом кредитном скоринге.

В медицине технологии больших данных используют для прогнозирования побочных эффектов компонентов, для выявления каких-

либо заболеваний на ранних стадиях, а также для того, чтобы дать прогноз по возможным заболеваниям. Коронавирусная инфекция Covid-19 привлекла внимание к большим данным. Анализ информации о симптомах заболевших, их количестве, местоположении и другом позволит не только отследить распространение вируса, но и замедлить скорость распространения, а также спланировать защитные меры по борьбе с пандемией.

Анализ данных по коронавирусу может помочь не только в отрасли медицины, но и в отрасли туризма. Аналитику таких данных можно использовать в туристическом бизнесе, показывая туристам уровень заболеваемости в той или иной стране.

1.2 Анализ проблемы

В сложившейся в мире обстановке люди сильно переживают за свое здоровье и могут отказываться путешествовать как прежде в связи с опасением заразиться вирусом Covid-19. Помимо заражения вирусом, люди во время пребывания в какой-либо стране могут столкнуться с неожиданной вспышкой коронавирусной инфекции, в связи с чем остаться там на неопределенный срок, пока не снимут карантинные меры.

Количество международных туристических поездок в 2021 году остается на 72% меньше, чем в допандемийном 2019 году, а доходы отрасли составляют лишь 54% от уровня двухлетней давности [3].

Коронавирусная инфекция очень быстро распространяется, особенно если не предпринять какие-либо защитные меры. В связи с этим люди очень рискуют, планируя куда-либо путешествия: на период покупки билетов в ту или иную страну, уровень заболеваемости может быть не так высок, но ближе к поездке распространение вируса может начать расти, в связи с чем вводятся карантинные меры, а также отменяют авиаперевозки.

Существуют некоторые веб-сайты, предоставляющие данные о количестве зараженных, выздоровевших и умерших от коронавирусной инфекции в каждой стране. Некоторые такие сайты, например *coronavirus-*

monitor.info, визуализируют количество зараженных коронавирусной инфекции за весь период пандемии. Данная информация не позволяет пользователям сделать какие-либо выводы о нынешней обстановке в стране. Также есть веб-сайты, например news.mail.ru/coronavirus/stat/, представляющие информацию в виде таблиц или графиков, но для просмотра таких диаграмм требуется время, чтобы просмотреть разные страны и сравнить статистику. Пример просмотра статистики о количестве зараженных и умерших за 3 месяца представлен на рисунке 1.

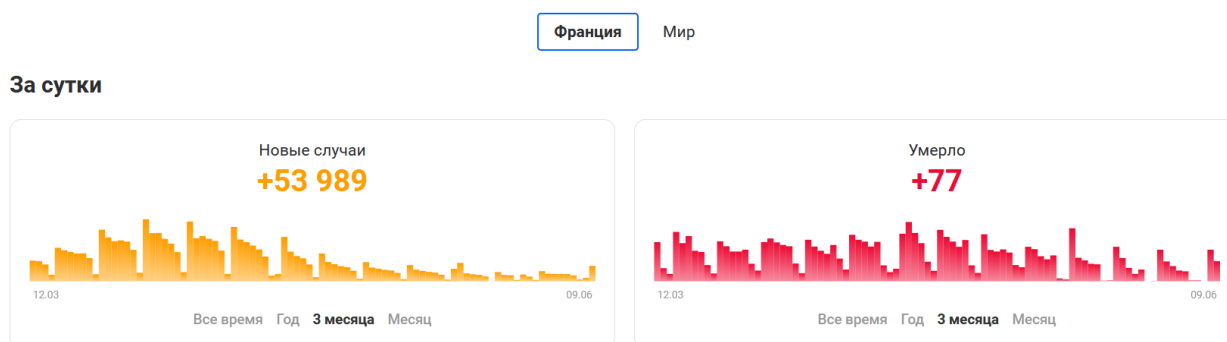


Рисунок 1 – Визуализация статистики распространения вируса Covid-19

Из-за неинформативных веб-сайтов, а также отсутствия какого-либо прогноза уровня заболеваемости люди могут отказываться от поездок или же путешествовать, рискуя собственной безопасностью.

Таким образом, проблема заключается в том, что люди мало информированы об уровне заболеваемости Covid-19. Для решения данной проблемы следует разработать систему, которая будет визуализировать необходимую статистику и прогноз уровня заболеваемости COVID-19 на карте мира.

1.3 Задачи исследования

Для классификации уровней заболеваемости коронавирусной инфекции необходимо сделать подготовку данных и произвести разведочный анализ. Данные для анализа, используемые в данной работе, находятся в открытом доступе и содержат информацию по ежедневной статистике вируса COVID-19

в странах мира с 31 декабря 2019 года. Датасет содержит 192552 записи и 67 характеристики. Описание всех характеристик приведено в приложении А.

Для дальнейшего анализа данных были выделены значимые атрибуты, которые оказывают наибольшее влияние на определение уровня заболеваемости. Проанализировав все атрибуты, было оставлено 26 признака, описание которых находится в таблице 1.

Таблица 1 – Значимых атрибуты датасета

№	Название	Перевод
1	iso_code	Код страны
2	continent	Континент
3	location	Страна
4	date	Дата
5	total_cases	Всего случаев
6	new_cases	Новые случаи
7	total_deaths	Всего смертей
8	new_deaths	Новые смерти
9	reproduction_rate	Скорость воспроизведения
10	new_tests	Новые тесты
11	total_tests	Всего тестов
12	positive_rate	Положительные тесты
13	stringency_index	Индекс строгости
14	population	Население
15	population_density	Плотность населения
16	median_age	Средний возраст
17	aged_65_older	Старше 65
18	aged_70_older	Старше 70
19	gdp_per_capita	ВВП на душу населения
20	extreme_poverty	Крайняя бедность

21	female_smokers	Курильщики женщины
22	male_smokers	Курильщики мужчины
23	handwashing_facilities	Устройства для мытья рук
24	hospital_beds_per_thousand	Больничных коек на тысячу
25	life_expectancy	Продолжительность жизни
26	human_development_index	Индекс развития человека

Также в качестве аргумента был добавлен атрибут, в котором будет содержаться анализ распространения вируса с помощью коэффициента распространения инфекции R_t . R_t - коэффициент распространения коронавируса — это показатель, определяющий среднее количество людей, которых инфицирует один больной до его изоляции. Он рассчитывается на основе данных по приросту новых случаев за последние 8 суток [4]. Коэффициент рассчитывается по формуле 1.

$$R_t = \frac{\text{число зарегистрированных заболеваний за последние 4 дня}}{\text{число зарегистрированных заболеваний за предыдущие 4 дня}} \quad (1)$$

На основе данного атрибута и будет решаться задача классификации стран по уровню заболеваемости.

1.4 Вывод по главе

В данной главе была рассмотрена предметная область, а также сделан анализ проблемы. В результате разрабатываемая в рамках данной работы система будет единственным сервисом, наглядно визуализирующим уровень заболеваемости коронавирусной инфекцией какой-либо страны в сравнении с другими странами. Помимо описания предметной области и изучения уже существующих сервисов со статистикой по вирусу COVID-19,

сформулирована задача исследования, визуализировать которое и будет веб-сайт.

Глава 2. Подготовка исходных данных для обработки

2.1. Методика подготовки исходных данных для построения моделей классификации

Большие данные – это огромные массивы информации, в таком количестве записей могут содержаться ошибочные значения, пропуски или же повторяющиеся строки. Для того чтобы прогноз был более точным, необходимо качественно обработать «сырые» данные, то есть заметить все ошибки в данных, а главное правильно исправить их.

В связи с этим для хороших итоговых результатов необходимо произвести несколько подготовительных этапов:

- проверить данные на дублирующие строки;
- проверить данные на выбросы;
- проверить наличие отсутствующих данных.

Если дублирующие строки можно просто удалить, то для исправления неточностей, таких как выбросы или отсутствующие данные, используют различные методы. Так, например, если выбросов или отсутствующих значений меньше половины, то следует заменить такие значения либо средним значением, либо медианой, либо ближайшим соседом, в противном же случае необходимо удалить данные записи или атрибуты из дальнейшего анализа.

2.2. Выбор технологий

Для анализа огромных массивов информации в последнее время чаще используется такой язык программирования, как Python. Python имеет простой синтаксис, а также множество открытых библиотек.

В данной работе в качестве среды разработки использовалась такая среда разработки, как PyCharm. PyCharm – это самая интеллектуальная Python IDE с полным набором средств для эффективной разработки на языке Python.

Также были в основе анализа данных использовались такие библиотеки, как Pandas, matplotlib и scikit-learn.

2.2.1. Библиотека Pandas

Библиотека Pandas была создана и используется для работы со структурированными в виде таблиц данными. Библиотека является открытой в связи с этим продвинутой и быстрорастущей.

Данная библиотека содержит множество встроенных методов группировки, комбинирования и фильтрации данных, методов анализа временных рядов. Помимо этого, библиотека умеет извлекать и осуществлять операции с данными из различных источников, среди них базы данных SQL, файлы CSV, Excel и JSON.

2.2.2. Библиотека Matplotlib

Matplotlib — это библиотека на языке Python для визуализации данных. В ней можно построить двумерные (плоские) и трехмерные графики [5]. С помощью данной библиотеке можно не только построить разные виды графиков, но и задать любой элемент диаграммы, например название осей, самой диаграммы, отметки на осях и другое.

Данная библиотека низкоуровневая и лежит в основе других библиотек, таким образом, для более быстрых построений графиков можно использовать высокоуровневые библиотеки, но для детализации и прорисовки диаграмм следует использовать Matplotlib.

2.2.3. Библиотека scikit-learn

Scikit-learn популярная библиотека для машинного обучения. Она содержит в себе большой набор алгоритмов, используемых для решения задач машинного обучения, таких как кластеризация и классификация или же регрессия.

В данной библиотеки реализованы методы разбиения массива данных на тестовый и обучающий, вычисление основных метрик над наборами данных, проведение Кросс-валидация.

2.3. Предобработка данных

Датасет с уже отобранными атрибутами состоит из 26 столбца-характеристиками и 192552 записями, увеличивающимися с каждым днем (рисунок 2). Данные взяты на 9 июня 2022 года.

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	iso_code	192552 non-null	object
1	continent	181361 non-null	object
2	location	192552 non-null	object
3	date	192552 non-null	object
4	total_cases	184979 non-null	float64
5	new_cases	184750 non-null	float64
6	total_deaths	166483 non-null	float64
7	new_deaths	166479 non-null	float64
8	reproduction_rate	143384 non-null	float64
9	total_tests	78116 non-null	float64
10	new_tests	74374 non-null	float64
11	positive_rate	94068 non-null	float64
12	stringency_index	149981 non-null	float64
13	population	191382 non-null	float64
14	population_density	171580 non-null	float64
15	median_age	159024 non-null	float64
16	aged_65_older	157339 non-null	float64
17	aged_70_older	158190 non-null	float64
18	gdp_per_capita	158179 non-null	float64
19	extreme_poverty	103250 non-null	float64
20	female_smokers	119998 non-null	float64
21	male_smokers	118353 non-null	float64
22	handwashing_facilities	77952 non-null	float64
23	hospital_beds_per_thousand	140769 non-null	float64
24	life_expectancy	180075 non-null	float64
25	human_development_index	154566 non-null	float64

Рисунок 2 – Информация о датасете

Для дальнейшего анализа данных необходимо понимать характеристики данных:

1. iso_code – Код страны;
2. continent – Континент;

3. location – Страна;
4. date – Дата;
5. total_cases – Всего подтвержденных случаев COVID-19;
6. new_cases – Новые подтвержденные случаи COVID-19;
7. total_deaths – Всего смертей от COVID-19;
8. new_deaths – Новые смерти от COVID-19;
9. reproduction_rate – Оценка эффективной скорости размножения COVID-19 в режиме реального времени;
- 10.new_tests – Новые тесты на COVID-19;
- 11.total_tests – Всего тестов на COVID-19;
- 12.positive_rate – Доля положительных тестов на COVID-19;
- 13.stringency_index – Индекс строгости ответных мер правительства: составной показатель;
- 14.population – Население;
- 15.population_density – Количество людей на площадь земли;
- 16.median_age – Средний возраст населения;
- 17.aged_65_older – Доля населения в возрасте 65 лет и старше;
- 18.aged_70_older – Доля населения в возрасте 70 лет и старше;
- 19.gdp_per_capita – ВВП на душу населения;
- 20.extreme_poverty – Доля населения, живущего в условиях крайней нищеты;
- 21.female_smokers – Доля курящих женщин;
- 22.male_smokers – Доля курящих мужчин;
- 23.handwashing_facilities – Доля населения, имеющего базовые приспособления для мытья рук в помещениях;
- 24.hospital_beds_per_thousand – Больничные койки на тысячу человек;
- 25.life_expectancy – Ожидаемая продолжительность жизни при рождении;

26.human_development_index – Составной индекс, измеряющий средние достижения по трем основным параметрам человеческого развития: долгая и здоровая жизнь, знания и достойный уровень жизни.

На рисунках 3-5 представлены первые 5 записей датасета.

	iso_code	continent	location	date	total_cases	new_cases	\
0	AFG	Asia	Afghanistan	2020-02-24	5.0	5.0	
1	AFG	Asia	Afghanistan	2020-02-25	5.0	0.0	
2	AFG	Asia	Afghanistan	2020-02-26	5.0	0.0	
3	AFG	Asia	Afghanistan	2020-02-27	5.0	0.0	
4	AFG	Asia	Afghanistan	2020-02-28	5.0	0.0	

	total_deaths	new_deaths	reproduction_rate	total_tests	new_tests	\
0	NaN	NaN	NaN	NaN	NaN	
1	NaN	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	

Рисунок 3 – Первые пять записей набора данных (часть 1)

	positive_rate	stringency_index	population	population_density	\
0	NaN	8.33	39835428.0	54.422	
1	NaN	8.33	39835428.0	54.422	
2	NaN	8.33	39835428.0	54.422	
3	NaN	8.33	39835428.0	54.422	
4	NaN	8.33	39835428.0	54.422	

	median_age	aged_65_older	aged_70_older	gdp_per_capita	extreme_poverty	\
0	18.6	2.581	1.337	1803.987	NaN	
1	18.6	2.581	1.337	1803.987	NaN	
2	18.6	2.581	1.337	1803.987	NaN	
3	18.6	2.581	1.337	1803.987	NaN	
4	18.6	2.581	1.337	1803.987	NaN	

Рисунок 4 – Первые пять записей набора данных (часть 2)

	female_smokers	male_smokers	handwashing_facilities	\
0	NaN	NaN	37.746	
1	NaN	NaN	37.746	
2	NaN	NaN	37.746	
3	NaN	NaN	37.746	
4	NaN	NaN	37.746	

	hospital_beds_per_thousand	life_expectancy	human_development_index
0	0.5	64.83	0.511
1	0.5	64.83	0.511
2	0.5	64.83	0.511
3	0.5	64.83	0.511
4	0.5	64.83	0.511

Рисунок 5 – Первые пять записей набора данных (часть 3)

2.4. Получение сводной информации и признаков

Уже на рисунках выше видно наличие в датасете отсутствующих значений. Получим сводную информацию обо всех признаках набора. Для этого разделим все атрибуты на количественные и категориальные (рисунок 6).

```
\# Разделение признаков на категориальные и количественные
categorical_columns = [c for c in df.columns if
df[c].dtype.name == 'object']
numerical_columns = [c for c in df.columns if
df[c].dtype.name != 'object']
print('categorical and numerical columns respectively')
print(categorical_columns)
print(numerical_columns)
```

Рисунок 6 – Разделение признаков на категориальные и количественные

В результате деления можно заметить, что категориальных признака всего 4, остальные- количественные (рисунок 7).

```
categorical and numerical columns respectively
['iso_code', 'continent', 'location', 'date']
['total_cases', 'new_cases', 'total_deaths', 'new_deaths', 'reproduction_rate',
```

Рисунок 7 – Результаты деления признаков на категориальные и
количественные

Описание количественных признаков содержит информацию об их:

- количестве (count);
- среднем значении (mean);
- стандартном отклонении (std);
- минимальном значении (min);
- медиане (50%);
- значении первой (25%) и третьей (75%) квартилей;
- максимальном значении (max).

На рисунках 8-10 представлена сводная информация для количественных признаков.

	total_cases	new_cases	total_deaths	new_deaths	\
count	1.849790e+05	1.847500e+05	1.664820e+05	166477.000000	
mean	3.337571e+06	1.231062e+04	6.513123e+04	157.081074	
std	2.096520e+07	8.703508e+04	3.392375e+05	787.961828	
min	1.000000e+00	0.000000e+00	1.000000e+00	0.000000	
25%	2.731000e+03	0.000000e+00	9.500000e+01	0.000000	
50%	3.619600e+04	7.000000e+01	9.280000e+02	1.000000	
75%	3.834990e+05	1.021000e+03	8.508000e+03	17.000000	
max	5.336620e+08	4.079892e+06	6.304835e+06	18152.000000	

	reproduction_rate	total_tests	new_tests	positive_rate	\
count	143384.000000	7.811600e+04	7.437400e+04	94068.000000	
mean	0.969716	2.058192e+07	6.701997e+04	0.098994	
std	0.372054	8.291270e+07	2.489342e+05	0.118279	
min	-0.070000	0.000000e+00	1.000000e+00	0.000000	
25%	0.780000	3.549080e+05	2.224000e+03	0.017300	
50%	0.980000	2.006404e+06	8.670000e+03	0.055550	
75%	1.160000	9.846145e+06	3.629875e+04	0.138900	
max	6.130000	9.214000e+09	3.585563e+07	1.000000	

Рисунок 8 – Сводная информация для количественных признаков (часть 1)

	stringency_index	population	population_density	median_age \
count	149981.000000	1.913820e+05	171580.000000	159024.000000
mean	52.733394	1.431532e+08	458.225201	30.653493
std	20.883846	6.950890e+08	2112.475358	9.069964
min	0.000000	4.700000e+01	0.137000	15.100000
25%	37.960000	8.960050e+05	37.312000	22.300000
50%	52.780000	7.552800e+06	88.125000	30.600000
75%	68.980000	3.277620e+07	214.243000	39.100000
max	100.000000	7.874966e+09	20546.766000	48.200000

	aged_65_older	aged_70_older	gdp_per_capita	extreme_poverty \
count	157339.000000	158190.000000	158179.000000	103250.000000
mean	8.832719	5.571715	19634.971689	13.603878
std	6.139328	4.170114	20575.558361	20.025444
min	1.144000	0.526000	661.240000	0.100000
25%	3.526000	2.063000	4449.898000	0.600000
50%	6.704000	4.209000	12951.839000	2.200000
75%	14.312000	8.678000	27936.896000	21.200000
max	27.049000	18.493000	116935.600000	77.600000

Рисунок 9 – Сводная информация для количественных признаков (часть 2)

	female_smokers	male_smokers	handwashing_facilities \
count	119998.000000	118353.000000	77952.000000
mean	10.643339	32.792625	50.919882
std	10.580023	13.530387	31.855789
min	0.100000	7.700000	1.188000
25%	1.900000	21.600000	20.859000
50%	6.300000	31.400000	49.839000
75%	19.300000	41.300000	83.241000
max	44.000000	78.100000	100.000000

	hospital_beds_per_thousand	life_expectancy	human_development_index
count	140769.000000	180075.000000	154566.000000
mean	3.068597	73.649553	0.725199
std	2.520137	7.449282	0.149729
min	0.100000	53.280000	0.394000
25%	1.300000	69.500000	0.602000
50%	2.400000	75.050000	0.743000
75%	4.000000	79.070000	0.845000
max	13.800000	86.750000	0.957000

Рисунок 10 – Сводная информация для количественных признаков (часть 3)

На рисунке 11 представлена сводная информация для категориальных признаков, содержащая количество заполненных ячеек, уникальных значений,

часто встречающееся значение и количество записей, в которых встречается частое значение.

	iso_code	continent	location	date
count	192552	181361	192552	192552
unique	244	6	244	890
top	MEX	Africa	Mexico	2021-08-30
freq	890	44643	890	243

Рисунок 11 – Сводная информация для категориальных признаков

2.5. Обработка выбросов

Значения, которые сильно выбиваются из общего набора данных, называются выбросами. Для построения более точной модели прогнозирования, необходимо обработать такие аномальные значения.

Для обработки данных выбросами будем считать значения, которые не укладываются в заданные квартили. Рассмотрим диаграмму размаха: здесь представлена медиана, интерквартильный размах, а также минимальное и максимальное значение (рисунок 12).

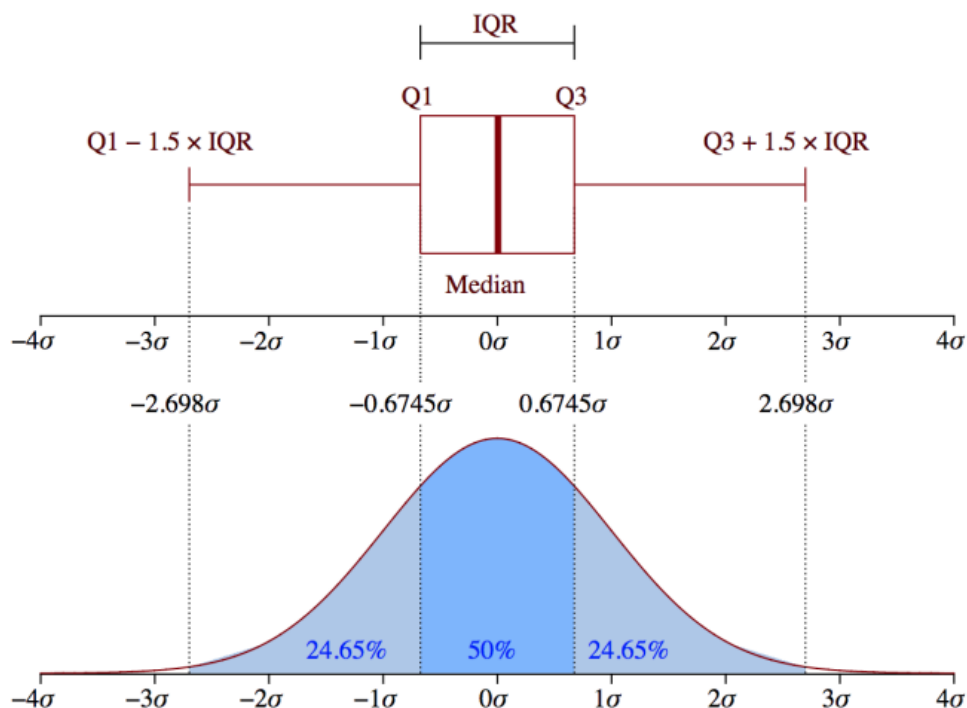


Рисунок 12 – Диаграмма размахов

Значение будет считаться нормальным, если будет лежать внутри интервала от $Q1 - 1.5 * IQR$ до $Q3 + 1.5 * IQR$, остальные значения – выбросы, которые следует либо удалить, либо заменить медианным или средним значением. В данной работе выбросы были заменены на медианное значение.

2.6. Обработка пропусков

Как уже было видно на рисунках выше, данные содержат пропуски (значения NaN), которые также следует обработать. На рисунке 13 представлено количество пропущенных значений в каждом столбце-характеристике.

iso_code	0
continent	11191
location	0
date	0
total_cases	7573
new_cases	7802
total_deaths	26070
new_deaths	26075
reproduction_rate	49168
total_tests	114436
new_tests	118178
positive_rate	98484
stringency_index	42537
population	1170
population_density	20972
median_age	33528
aged_65_older	35213
aged_70_older	34362
gdp_per_capita	34373
extreme_poverty	89302
female_smokers	72554
male_smokers	74199
handwashing_facilities	114600
hospital_beds_per_thousand	51783
life_expectancy	12477
human_development_index	37986
dtype:	int64

Рисунок 13 – Количество пропущенных значений

При большом количестве пропущенных значений строки или столбцы удаляются, в остальных случаях, пропущенные значения заменяются на медианные или средние значения, или же на соседние. В данной работе некоторые признаки, такие как подтвержденные случаи заражения вирусом (total_cases), нет смысла заменять на средние значения, так как статистика ежедневная, это может привести к выбросам, поэтому следует заменить такие пропущенные значения на соседние. Соседние значения в данном наборе данных определяют количество подтвержденных случаев заражения на следующий или предыдущий день, что чаще всего является максимально приближенным к пропущенному числу.

Другие же признаки, следует заменить на медианные значения, так как, в сравнении со средними, они более устойчивы к выбросам.

Запрашиваем еще раз проверку на пропущенные значения, чтобы убедиться в отсутствии пропущенных значений (рисунок 14).

location	0
date	0
total_cases	0
new_cases	0
total_deaths	0
new_deaths	0
reproduction_rate	0
total_tests	0
new_tests	0
positive_rate	0
stringency_index	0
population	0
population_density	0
median_age	0
aged_65_older	0
aged_70_older	0
gdp_per_capita	0
extreme_poverty	0
female_smokers	0
male_smokers	0
handwashing_facilities	0
hospital_beds_per_thousand	0
life_expectancy	0
human_development_index	0

Рисунок 14 – Количество пропущенных значений

2.7. Удаление дубликатов

Удаление дубликатов также является важной частью обработки данных. На рисунке 15 представлен метод обнаружения и удаления строк-дубликатов.

```
print('Deleting duplicates')
df = df.drop_duplicates()
print(df.shape)
print('\r')
```

Рисунок 15 – Удаление дубликатов

В результате удаления количество строк не изменилось, следовательно этот набор данных не содержал повторяющихся строк.

2.8. Рассмотрение таблицы корреляции

Последним этапом предобработки является проверка данных на мультиколлинеарность. Для этого построим таблицу корреляционную матрицу, с помощью которой можно установить корреляцию признаков.

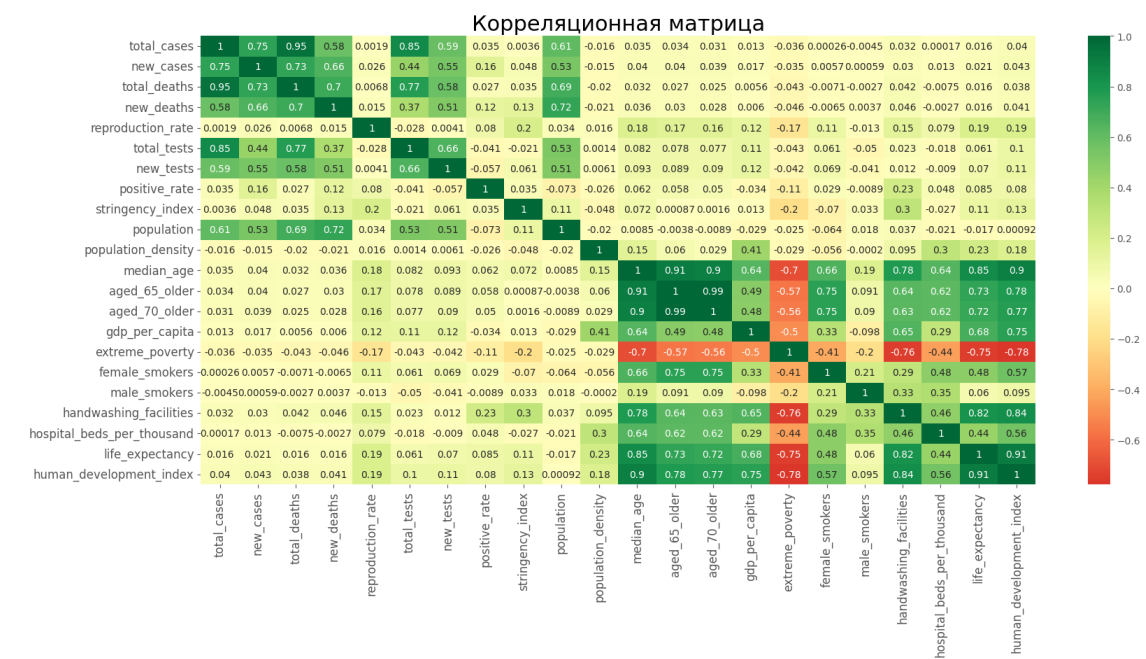


Рисунок 16– Корреляционная матрица

Таким образом, если коэффициент корреляции больше или равен 0,75, то признаки линейно зависимы друг от друга. В дальнейшем такие признаки

могут сказать на точности прогнозировании результатов, в связи с чем необходимо удалить некоторые признаки.

Вновь построим корреляционную матрицу, чтобы убедиться, что все зависимые признаки были удалены (рисунок 17).

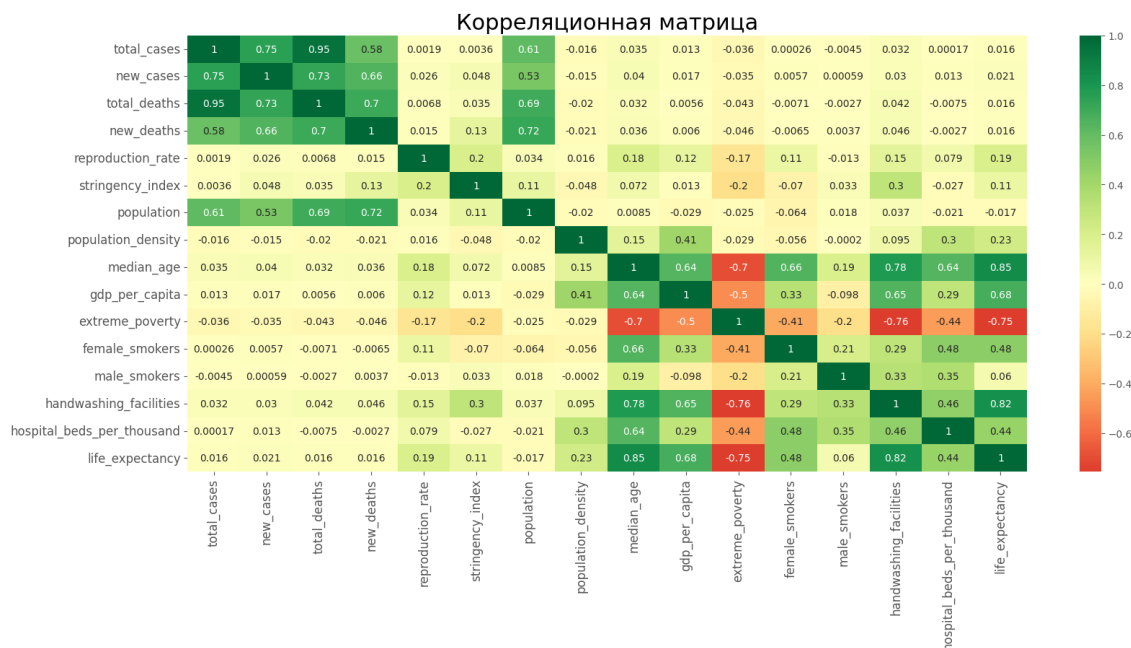


Рисунок 17– Корреляционная матрица

2.9. Вывод по главе

В данной главе была проведена обработка «сырых» данных, в результате чего сформировался набор данных для дальнейшего прогнозного анализа. В ходе данного раздела были определены и удалены выбросы, восстановлены пропущенные значения, удалены дубликаты строк, а также проведена проверка на коллинеарные признаки.

После проведения такой предобработки данных, можно приступить к дальнейшему анализу.

Глава 3. Разработка системы предсказания

3.1. Метод k-средних

В данной работе в качестве модели машинного обучения была выбрана модель машинного обучения без учителя - метод k-средних (k-Means Clustering). Метод K-средних – это метод кластерного анализа, целью которого является разделение m наблюдений на k кластеров, при этом каждое наблюдение относится к тому кластеру, к центру (центроиду) которого оно ближе всего [6].

Данный метод был выбран в связи с тем, что нам заранее известно количество кластеров, на которые необходимо разбить данные. В данном случае необходимо разделить данные на 3 группы. Данные группы будут характеризовать уровень опасности посещения той или иной страны: «зеленый» - посещение безопасно, «желтый» - средний уровень опасности и «красный» - посещение опасно.

Алгоритм k-средних заключается в том, что выбирает лучшие точки данных K , чтобы сформировать центр кластера. После этого он повторяет 2 этапа для каждой точки: присваивает точку данных ближайшему центру кластеров, создает новый центр, взяв среднее значение всех точек данных из этого кластера [7].

Так как алгоритм k-средних является методом машинного обучения без учителя нет необходимости разбивать историческую выборку на обучающую и тестовую.

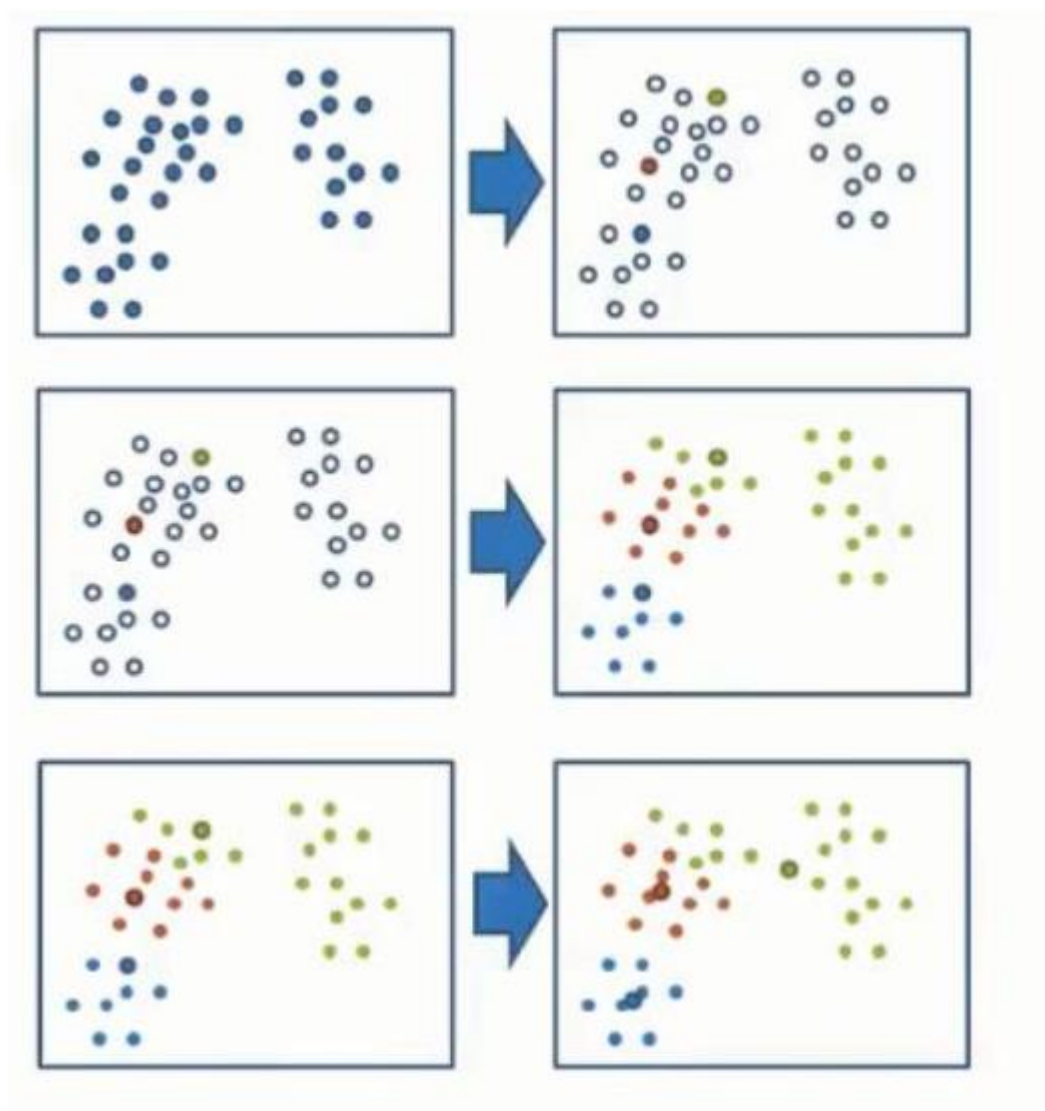


Рисунок 18– Метод k-средних

3.2. Обучение

В качестве метода машинного обучения используется алгоритм k-средних.

Обучим модель с помощью библиотеки scikit-learn

```
from sklearn.cluster import KMeans  
model = KMeans(n_clusters=3)  
model.fit(df[0])
```

Рисунок 19 – Обучение

Далее визуализируем точность предсказаний модели (рисунок 20).

```
f, (ax1, ax2) = plt.subplots(1, 2, sharey=True, figsize=(10, 6))
ax1.set_title('Наши предсказания')
ax1.scatter(df[0][:, 0], df[0][:, 1], c=model.labels_)
ax2.set_title('Реальные значения')
ax2.scatter(df[0][:, 0], df[0][:, 1], c=df[1]);
plt.show()
```

Рисунок 20 – Визуализация точности модели

На рисунке 21 представлен результат точности предсказаний.

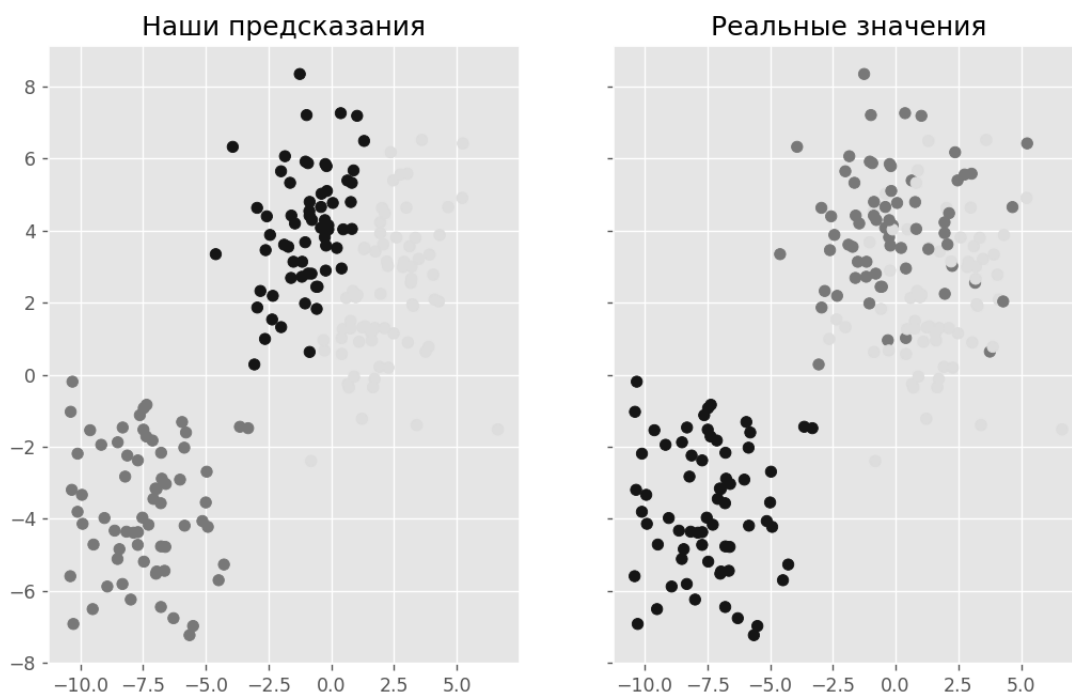


Рисунок 21 – результат точности предсказаний

3.3. Создание системы

Результатом работы должна являться система, доступная и интуитивно понятная всем пользователям. Для этого необходимо разработать веб-сайт с картой мира, на котором будет отображаться информация об эпидемиологической обстановке.

Для реализации данной задачи была сделана карта мира (рисунок 22).



Рисунок 22 – карта мира для визуализации уровня опасности

Для итогового результата на карту необходимо добавить результаты предсказаний (рисунок 23).

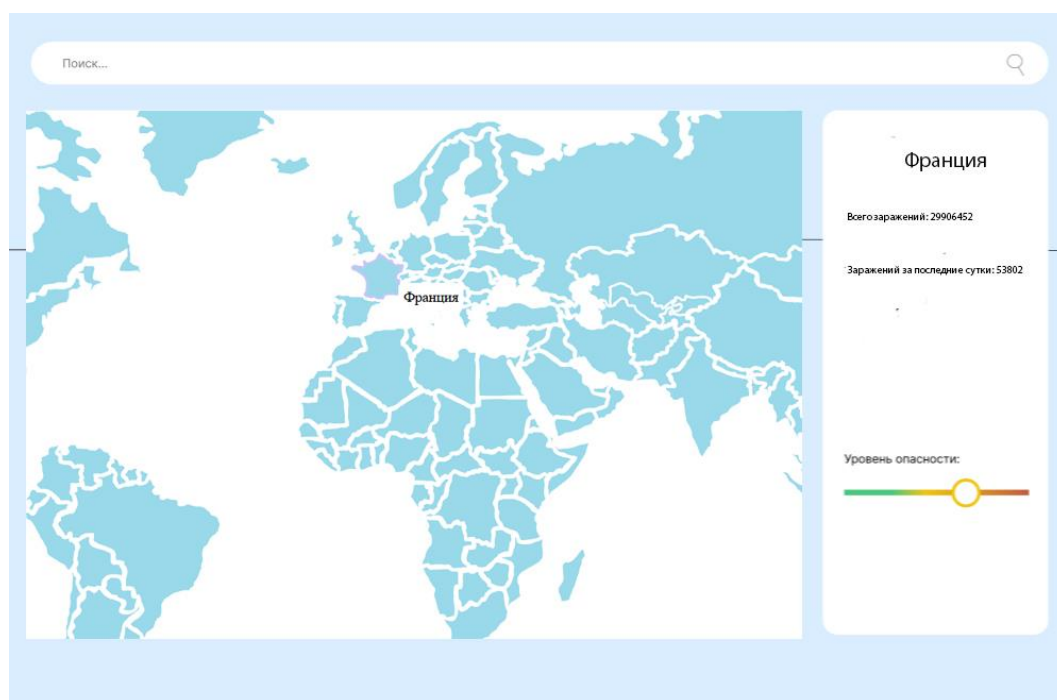


Рисунок 23 – готовая система диагностики и предсказаний

3.4. Вывод по главе

В данной главе была выбрана и применена модель машинного обучения. После обучения был разработан веб-сайт, содержащий карту миру, на которой отображается уровень опасности заражения COVID-19.

Глава 4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Разработка НИ производится группой работников, состоящей из двух человек – руководителя и студента.

Данная выпускная квалификационная работа заключается в исследовании и предсказании распространения вируса COVID-19 в разных странах мира, а также проектировании и разработке веб-приложения на их основе. Соответственно, задача раздела – экономически обосновать разработку, определить и рассчитать трудовые и денежные затраты на ее создание.

4.1 Потенциальные потребители результатов исследования

Целевым рынком для данной разработки является рынок туризма, осуществляющий продажу билетов, туров в различные страны мира.

Целевой аудиторией могут являться:

1. туристические агентства
2. авиакомпании
3. туристические операторы.

Предполагается, что вышеперечисленные организации будут предлагать свои продукты, используя данное исследование, с большей точностью распространения вируса COVID-19 в той или иной стране.

4.2 Анализ конкурентных технических решений

В ходе выполнения работы были выявлены и проанализированы конкуренты разрабатываемой программной системы. В качестве конкурента 1 была взята система «coronavirusstat», которая отображает статистику активных, выздоровевших, а также умерших на сегодняшний день случаев каждой страны. На основе данных делается прогноз на ближайшие десять дней.

В качестве конкурента 2 была взята система «coronavirus-monitor» визуализации статистики заражения коронавирусной инфекцией. Она представляет собой мировую карту, отображающую количество заражений, смертей и выздоровлений за все время с начала пандемии.

Анализ конкурентных технических решений с позиции ресурсоэффективности и ресурсосбережения позволяет провести оценку сравнительной эффективности научной разработки и определить направления для ее будущего повышения.

По результатам анализа была составлена таблица 2 «Оценочная карта для сравнения конкурентных технических решений».

Таблица 2 – Оценочная карта для сравнения конкурентных технических решений

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б _ф	Б _{к1}	Б _{к2}	К _ф	К _{к1}	К _{к2}
1	2	3	4	5	6	7	8
Интуитивно-понятный интерфейс	0,15	5	3	4	0,75	0,45	0,6
Красивая визуализация данных	0,15	4	3	4	0,6	0,45	0,6
Удобство эксплуатации	0,2	5	4	3	1	0,8	0,6
Прогнозируемость данных	0,2	5	4	3	1	0,8	0,6
Предполагаемый срок эксплуатации	0,1	5	5	5	0,5	0,5	0,5
Срок выхода на рынок	0,1	3	5	5	0,3	0,5	0,5
Объем трудозатрат	0,1	3	4	4	0,3	0,4	0,4
Итого	1	30	28	28	4,45	3,9	3,8

Анализ конкурентных технических решений определяется по формуле:

$$K = \sum B_i \times B_i \quad (4.1)$$

где K – конкурентоспособность вида;

B_i– вес критерия (в долях единицы);

B_i – балл i-го показателя.

Анализ конкурентоспособности показал, что продукт опережает

технические и экономические показатели эффективности. Основными преимущественными факторами являются: интуитивно-понятный интерфейс, удобство эксплуатации и прогнозируемость данных.

4.3 SWOT-анализ

Произведем также в данном разделе SWOT – анализ, позволяющий оценить факторы и явления, способствующие или препятствующие продвижению проекта на рынок. SWOT-анализ применяют для исследования внешней и внутренней среды проекта. Метод помогает понять, какие действия необходимо предпринять для максимизации возможностей проекта и нейтрализации слабых сторон и угроз.

Целью использования SWOT-анализа является определение возможной эффективности и прогнозирование направлений будущего развития разрабатываемого решения.

На первом этапе SWOT анализа в таблице 3 были описаны сильные и слабые стороны проекта, выявлены возможности и угрозы реализации проекта.

Таблица 3 – Матрица SWOT анализа

Сильные стороны	Возможности во внешней среде
С1. Интуитивно-понятный интерфейс; С2. Использование математической точности при прогнозе распространения коронавирусной инфекции; С3. Востребованность подобных систем на рынке туризма.	В1. Возможность продажи туристическим организациям; В2. Легкая интеграция нового функционала; В3. Получение нового опыта разработки веб-приложений.
Слабые стороны	Угрозы внешней среды
Сл1. Отсутствие рекламной компании; Сл2. Отсутствие опыта проведения подобного рода исследований; Сл3. Значительные временные и интеллектуальные затраты.	У1. Появление аналогов; У2. Отсутствие спроса.

Второй этап состоит в выявлении соответствия сильных и слабых сторон проекта внешним условиям окружающей среды. Для этого необходимо построить интерактивную матрицу проекта, которая позволит оценить связи факторов между собой. Интерактивная матрица рассматриваемого продукта представлена в таблицах 4 и 5.

Таблица 4 - Интерактивная матрица сильных и слабых сторон и возможностей

Возможности проекта	Сильные стороны				Слабые стороны		
		C1	C2	C3	Сл1	Сл2	Сл3
	B1	0	0	+	+	+	-
	B2	+	+	+	-	+	+
	B3	+	+	-	-	+	0

Таблица 5 - Интерактивная матрица сильных сторон и слабых сторон и угроз

Угрозы проекта	Сильные стороны				Слабые стороны		
		C1	C2	C3	Сл1	Сл2	Сл3
	У1	+	-	+	-	-	-
	У2	-	-	+	+	-	+

Корреляцию возможностей и угроз с сильными и слабыми сторонами можно записать в следующей форме:

- B2B3C1C2; B1B2C3;
- B1Сл1; B1B2B3Сл2; B2Сл3;
- У1C1; У1У2C3;
- У2Сл1; У2Сл3.

Такой метод записи помогает выявить группы факторов, которые имеют единую природу.

Наконец, в рамках третьего этапа составляется итоговая матрица SWOT-анализа, представленная в таблице 6.

Таблица 6 - Итоговая матрица SWOT-анализа

	Сильные стороны: С1. Интуитивно-понятный интерфейс; С2. Использование математической точности при прогнозе распространения коронавирусной инфекции; С3. Востребованность подобных систем на рынке туризма.	Слабые стороны: Сл1. Отсутствие рекламной компании; Сл2. Отсутствие опыта проведения подобного рода исследований; Сл3. Значительные временные и интеллектуальные затраты.
Возможности: В1. Возможность продажи туристическим организациям; В2. Легкая интеграция нового функционала; В3. Получение нового опыта разработки веб-приложений.	Актуальный и простой в эксплуатации продукт с прогнозируемыми данными повышает шанс на продажу туристическим организациям.	Легкая интеграция нового функционала повысят уровень конкурентоспособности продукта на рынке, что повысит спрос к нему.
Угрозы: У1. Появление аналогов; У2. Отсутствие спроса.	Появление аналогов, отвечающих интересам той же целевой аудитории, напрямую влияет на востребованность данного продукта. С ростом количества аналогов спрос на конкретную систему становится меньше.	Самой большой угрозой для проекта является отсутствие спроса на данный продукт.

4.4 Планирование работ по научно-техническому исследованию

4.4.1 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.

Перечень этапов и работ, распределение исполнителей по данным видам работ приведен в таблице 7.

Таблица 7– Перечень этапов, работ и распределение исполнителей

№	Этапы работы	Исполнители работы
1	Выбор научного руководителя бакалаврской работы	Семенова О.Ю.
2	Составление и утверждение темы бакалаврской работы	Марухина О.В. Семенова О.Ю.
3	Постановка целей и задач	Марухина О.В. Семенова О.Ю.
4	Разработка календарного плана	Марухина О.В. Семенова О.Ю.
5	Подбор и изучение материалов по тематике	Семенова О.Ю.
6	Проведение анализа предметной области	Семенова О.Ю.
7	Проведение анализа данных	Семенова О.Ю.
8	Проектирование	Семенова О.Ю.
9	Разработка	Марухина О.В. Семенова О.Ю.
10	Согласование выполненной работы с научным руководителем	Марухина О.В. Семенова О.Ю.
11	Выполнение разделов финансовый менеджмент, социальная ответственность	Семенова О.Ю.
12	Подведение итогов, оформление работы	Марухина О.В. Семенова О.Ю.

4.4.2 Определение трудоемкости выполнения работ

Трудовые затраты в большинстве случаях образуют основную часть стоимости разработки, поэтому важным моментом является определение трудоемкости работ каждого из участников научного исследования.

Трудоемкость выполнения научного исследования оценивается экспертным путем в человеко-днях и носит вероятностный характер, который зависит от множества трудно учитываемых факторов. Для определения ожидаемого (среднего) значения трудоемкости $t_{ожі}$ используется следующая формула:

$$t_{ожі} = \frac{3t_{\min i} + 2t_{\max i}}{5}, \quad (4.2)$$

где $t_{ожі}$ – ожидаемая трудоемкость выполнения i -ой работы чел.-дн.;

$t_{\min i}$ – минимально возможная трудоемкость выполнения заданной i -ой работы, чел.-дн.;

$t_{\max i}$ – максимально возможная трудоемкость выполнения заданной i -ой работы, чел.-дн.;

Исходя из ожидаемой трудоемкости работ, определяется продолжительность каждой работы в рабочих днях T_p , учитывающая параллельность выполнения работ по нескольким исполнителями.

$$T_{pi} = \frac{t_{ожі}}{Ч_i}, \quad (4.3)$$

где T_{pi} – продолжительность одной работы, раб.дн.;

$t_{ожі}$ – ожидаемая трудоемкость выполнения одной работы, чел.-дн.;

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

4.4.3 Разработка графика проведения научного исследования

Наиболее удобным и наглядным представлением проведения научных работ является построение ленточного графика в форме диаграммы Ганта.

Диаграмма Ганта – горизонтальный ленточный график, на котором работы по теме представляются протяженными во времени отрезками, характеризующимися датами начала и окончания выполнения данных работ.

Для удобства построение графика, длительность каждого из этапов работ из рабочих дней следует перевести в календарные дни. Для этого необходимо воспользоваться следующей формулой:

$$T_{ki} = T_{pi} \cdot k_{\text{кал}}, \quad (4.4)$$

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

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

$k_{\text{кал}}$ – коэффициент календарности.

Коэффициент календарности определяется по следующей формуле:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - (T_{\text{вых}} + T_{\text{пр}})}, \quad (4.5)$$

где $T_{\text{кал}}$ – количество календарных дней в году;

$T_{\text{вых}}$ – количество выходных дней в году;

$T_{\text{пр}}$ – количество праздничных дней в году.

Расчет коэффициента календарности:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - (T_{\text{вых}} + T_{\text{пр}})} = \frac{365}{365 - 118} = 1,48$$

Таблица 8 – Временные показатели проведения научного исследования

Наименование работы	Исполнители работы	Трудоемкость работ, чел-дни			Длительность работ, дни	
		$t_{\min}$	$t_{\max}$	$t_{\text{ож}}$	$T_{\text{р}}$	$T_{\text{к}}$
Выбор научного руководителя ВКР	Семенова О.Ю.	1	2	1,4	1	2
Составление и утверждение темы бакалаврской работы	Марухина О.В.	2	4	2,8	3	4
	Семенова О.Ю.	2	4	2,8	3	4
Постановка целей и задач	Марухина О.В.	5	7	5,8	6	7
	Семенова О.Ю.	5	7	5,8	6	7
Разработка календарного плана	Марухина О.В.	1	2	1,4	1	2
	Семенова О.Ю.	1	2	1,4	1	2
Подбор и изучение материалов по тематике	Семенова О.Ю.	7	14	9,8	10	12
Проведение анализа предметной области	Семенова О.Ю.	7	14	9,8	10	12
Проведение анализа данных	Семенова О.Ю.	25	45	35	35	44
Проектирование	Семенова О.Ю.	14	21	16,8	17	21

Разработка	Семенова О.Ю.	25	45	35	35	44
Согласование выполненной работы с научным руководителем	Марухина О.В.	3	7	4,6	5	6
	Семенова О.Ю.	3	7	4,6	5	6
Выполнение разделов финансовый менеджмент, социальная ответственность	Семенова О.Ю.	5	7	5,8	6	7
Подведение итогов, оформление работы	Марухина О.В.	2	7	4	4	5
	Семенова О.Ю.	2	7	4	4	5

4.5 Бюджет научно-технического исследования (НТИ)

4.5.1. Расчет материальных затрат НТИ

При расчете бюджета следует учитывать специфику продукта, так как работа представляет собой проектирование и разработку информационной системы, материальные затраты будут равны нулю.

4.5.2. Расчет затрат на специальное оборудование для научных работ

В специальное оборудование входит оборудование для рабочего места, которое является необходимым для реализации программного обеспечения. Затраты на специальное оборудование приведены в таблице 9.

Таблица 9 – Расчет бюджета затрат на приобретение спецоборудования для научных работ

Наименование оборудования	Количество единиц оборудования	Цена за 1 ед. оборудования	Затраты, руб.
Ноутбук	1	60000	60000

Подставка для ноутбука	1	2000	2000
Компьютерная мышшь	1	1000	1000
Итого:			63000

Срок полезного использования офисных машин (код 330.28.23.23) составляет от 2 до 3 лет. Для вычисления амортизации данный срок можно принять за 3 года.

Норма амортизации вычисляется по следующей формуле:

$$A_n = \frac{100\%}{3} = 33,33\% \quad (4.6)$$

Годовые амортизационные вычисления составляют:

$$A_r = S * \frac{A_n}{100\%} = 63\,000 * 0,33 = 20\,790 \text{ рублей} \quad (4.7)$$

Ежемесячные амортизационные отчисления составят:

$$A_m = \frac{A_r}{12} = \frac{20790}{12} = 1732,5 \text{ рублей} \quad (4.8)$$

Итого за весь период выполнения исследовательской работы, с учётом того, что его продолжительность равна приблизительно 6 месяцам, амортизация равна:

$$A = A_m * 6 = 1732,5 * 6 = 10\,395 \text{ рублей} \quad (4.9)$$

4.5.3 Основная заработная плата исполнителя темы

Основная зарплата – это та часть зарплата, которая выплачивается работнику гарантированно, не зависимо от результатов труда. Расчёт баланса рабочего времени приведён в таблице 10.

Таблица 10 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Студент
Календарное число дней	365	365
Количество нерабочих дней - выходные дни - праздничные дни	66	66

Потери рабочего времени		
- отпуск	56	80
- невыходы по болезни	0	0
Действительный годовой фонд рабочего времени	243	219

Исходя из всех найденных показателей можно составить таблицу расчета основной заработной платы. Зарплата студента принимается за 21 000 рублей в месяц, а зарплата преподавателя за 45 000 рублей.

Таблица 11 – Расчет основной заработной платы

Исполнители	$Z_{тс}$ руб	кпр	кд	кр	Z_m , руб	$Z_{дн}$, руб	T_p	$Z_{осн}$
Семенова О.Ю.	16 153	-	-	1,3	21 000	700	133	93 100
Марухина О.В.	34 615	-	-	1,3	45 000	1 500	19	28 500
Итого:								121 600

Общий размер основной заработной платы составил 121 600 рублей.

4.5.4 Расчет дополнительной заработной платы

Дополнительная заработная плата учитывает величину предусмотренных Трудовым кодексом РФ доплат за отклонение от нормальных условий труда, а также выплат, связанных с обеспечением гарантий и компенсаций (при исполнении государственных и общественных обязанностей, при совмещении работы с обучением, при предоставлении ежегодного оплачиваемого отпуска и т.д.).

Расчет дополнительной заработной платы рассчитывается по формуле:

$$Z_{\text{доп}} = k_{\text{доп}} \cdot Z_{\text{осн}}, \quad (4.10)$$

где $k_{\text{доп}}$ – коэффициент дополнительной заработной платы, принятый на стадии проектирования за 0,15.

Таблица 12 – Расчет дополнительной заработной платы

Исполнители	З _{осн} , руб.	k _{доп}	З _{осн} , руб.
Марухина О.В.	28 500	0,15	4 275
Семенова О.Ю.	93 100		13 965
Итого			18 240

4.5.5 Отчисления во внебюджетные фонды

В данной статье расходов отражаются обязательные отчисления по установленным законодательством Российской Федерации нормам органам государственного социального страхования (ФСС), пенсионного фонда (ПФ) и медицинского страхования (ФФОМС) от затрат на оплату труда работников.

Величина отчислений во внебюджетные фонды определяется исходя из формулы:

$$З_{внеб} = k_{внеб} \cdot (З_{осн} + З_{доп}) \quad (4.11)$$

где $k_{внеб}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

В соответствии с Федеральным законом от 24.07.2009 №212-ФЗ установлен размер страховых взносов равный 30,2%.

Отчисления во внебюджетные фонды представлены в таблице 13.

Таблица 13 – Отчисления во внебюджетные фонды

Исполнитель	Основная заработная плата, руб.	Дополнительная заработная плата, руб.
Марухина О.В.	28 500	4 275
Семенова О.Ю.	93 100	13 965
Коэффициент отчислений во внебюджетные фонды	0,302	
Итого	42 231,68	

4.5.6 Накладные расходы

Накладные расходы учитывают прочие затраты организации, не попавшие в предыдущие статьи расходов. Их величина определяется по формуле:

$$З_{накл} = (З_{м} + З_{ам} + З_{осн} + З_{доп} + З_{внеб}) \cdot k_{нр} \quad (4.12)$$

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

Величину коэффициента накладных расходов можно взять в размере 16%.

Накладные расходы составили:

$$З_{\text{накл}} = (10395 + 121600 + 18240 + 42231,68) \cdot 0,16 = 30\,794,67 \text{ руб.}$$

4.5.7 Формирование бюджета затрат научно-исследовательского проекта

Рассчитанная величина затрат научно–исследовательской работы является основой для формирования бюджета затрат проекта. Определение бюджета затрат на научно–исследовательский проект приведено в таблице 14.

Таблица 14 –Расчет бюджета затрат НТИ

Наименование статьи	Сумма, руб	Примечание
Амортизация основных средств	10 395	Пункт 4.5.2
Затраты на основную заработную плату	121 600	Пункт 4.5.3
Затраты на дополнительную заработную плату	18 240	Пункт 4.5.4
Страховые взносы	42 231,68	Пункт 4.5.5
Накладные расходы	30 794,67	Пункт 4.5.6
Общий бюджет	223 261,35	Сумма всех пунктов

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

Определение эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Его нахождение связано с определением двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования определяется как:

$$I_{\text{фин.р}}^{\text{исп.}i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}} \quad (4.13)$$

где $I_{\text{фин.р}}^{\text{исп.}i}$ – интегральный финансовый показатель разработки;

Φ_{pi} – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения научно-исследовательского проекта.

Интегральный показатель ресурсоэффективности вариантов исполнения объекта исследования можно определить следующим образом:

$$I_{pi} = \sum_{i=1}^n a_i \times b_i \quad (4.14)$$

где I_{pi} – интегральный показатель ресурсоэффективности для i -го варианта исполнения разработки;

a_i – весовой коэффициент i -го варианта исполнения разработки;

b_i^a, b_i^p – бальная оценка i -го варианта исполнения разработки, устанавливается экспертным путем по выбранной шкале оценивания;

n – число параметров сравнения.

Таблица 15 – Сравнительная оценка характеристик вариантов исполнения проекта

Объект исследования Критерии	Весовой коэффициент параметра	Исп. 1	Исп. 2
1. Способствует росту производительности труда пользователя	0,1	4	4
2. Удобство в эксплуатации (соответствует требованиям потребителей)	0,2	5	4
3. Помехоустойчивость	0,15	4	4
4. Энергосбережение	0,15	5	4
5. Надежность	0,2	5	3
6. Материалоемкость	0,2	5	4
Итого	1	4,75	3,8

Интегральный показатель эффективности вариантов исполнения разработки ($I_{испi}$) определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп1} = \frac{I_{р-исп1}}{I_{фин.р}^{исп1}} = \frac{4,75}{1} = 4,75$$

$$I_{исп2} = \frac{I_{р-исп2}}{I_{фин.р}^{исп2}} = \frac{3,8}{1} = 3,8;$$

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволит определить сравнительную эффективность проекта и выбрать наиболее целесообразный вариант из предложенных.

Сравнительная эффективность проекта ($\mathcal{E}_{ср}$):

$$\mathcal{E}_{ср} = \frac{I_{исп2}}{I_{исп1}} \quad (4.15)$$

Таблица 16 – Сравнительная эффективность разработки

№	Показатели	Исп.1	Исп.2
1	Интегральный финансовый показатель разработки	1	1
2	Интегральный показатель ресурсоэффективности разработки	4,75	3,8
3	Интегральный показатель эффективности	4,75	3,8
4	Сравнительная эффективность вариантов исполнения	1	0,8

Сравнив значения интегральных показателей эффективности, можно сделать вывод, что реализация технологии в первом исполнении является более эффективным вариантом решения задачи, поставленной в данной работе с позиции финансовой и ресурсной эффективности.

Вывод по главе

В ходе выполнения раздела финансового менеджмента проведен анализ финансово-экономических аспектов разработки программной системы. Составлен перечень проводимых работ, их исполнителей и продолжительность выполнения этапов работ, составлен линейный график.

Основываясь на результатах проведенного анализа, разработка системы является конкурентоспособной и перспективной.

Глава 6. Социальная ответственность

Введение

В рамках выпускной квалификационной работы была создана система диагностики и предсказания уровня заболеваемости COVID-19 в разных странах мира. Система предназначена для людей, желающих посещать различные страны, но беспокоившиеся об уровне распространения коронавирусной инфекции. Пользователи системы могут выбрать любую страну и посмотреть уровень опасности посещения данной страны в виде светофора.

В связи с распространяющимся вирусом люди столкнулись с рядом ограничений, в том числе и на ограничение в путешествиях. Туризм – сфера экономики, сильно пострадавшая из-за пандемии. С течением времени люди начали соблюдать меры безопасности, вследствие чего многим странам удалось замедлить распространение вируса COVID-19. Туристы могут снова начать посещать многие места, но следует ответственно подходить к выбору страны, чтобы обезопасить себя и окружающих. Изучение статистики по распространению коронавирусной инфекции и ее анализ приводит к дополнительным затратам времени. Для решения данной проблемы была сформирована идея создания системы, которая совместит в себе анализ данных распространения вируса и его визуализацию.

Целью реализации проекта является обеспечение устойчивых конкурентных преимуществ на рынке, уменьшение временных затрат и повышение качества работы за счёт привлечения современных информационных технологий для автоматизации бизнес-процессов.

Данное мобильное приложение разрабатывалось во время обучения в ТПУ, ИШИТР, ОИТ студенткой в офисе размером 2*6 м. В процессе проектирования и разработки использовался 1 ноутбук.

6.1. Правовые и организационные вопросы обеспечения безопасности

6.1.1. Правовые нормы трудового законодательства

Трудовой кодекс Российской Федерации [8] определяет трудовые отношения между работниками и работодателями и имеет приоритетное значение перед другими принятыми федеральными законами, связанными с трудовыми отношениями.

Создание благоприятных условий труда, установление государственных гарантий трудовых прав и свобод граждан, а также защита прав и интересов работников и работодателей являются главными целями трудового законодательства.

Специфика данной области по разработке программного обеспечения такова, что для полноценной работы достаточно иметь лишь персональный компьютер и стабильный доступ в сеть Интернет. Но также немаловажными факторами производительности сотрудника являются удобство его рабочего места, режим рабочего дня, атмосфера в коллективе и т.д.

Основные положения:

– В статье 108 Трудового Кодекса РФ «Перерывы для отдыха и питания» сказано, что работнику в течение рабочего дня должен быть предоставлен перерыв продолжительностью не более двух часов и не менее 30 минут, который в рабочее время не включается.

– Согласно статье 91 Трудового Кодекса РФ «Понятие рабочего времени. Нормальная продолжительность рабочего времени» нормальная продолжительность рабочего времени не может превышать 40 часов в неделю.

– Статья 111 Трудового Кодекса РФ «Выходные дни»: всем работникам предоставляются выходные дни (еженедельный непрерывный отдых). При пятидневной рабочей неделе работникам предоставляются два выходных дня в неделю, при шестидневной рабочей неделе - один выходной день.

6.1.2. Эргономические требования к правильному расположению и компоновке рабочей зоны

При разработке системы сотрудник выполняет свою работу преимущественно сидя за своим рабочим местом, которое должно быть организовано с учетом требований ГОСТ 12.2.032-78 «Система стандартов безопасности труда. Рабочее место при выполнении работ сидя. Общие эргономические требования» [9]. Согласно указанным требованиям, должны быть соблюдены следующие условия:

1. Конструкцией рабочего места должно быть обеспечено выполнение трудовых операций в пределах зоны досягаемости моторного поля.
2. Высота рабочей поверхности при организации рабочего места для работы данной специфики, выполняемой женщинами, - 630 мм, а высота сидения – 400 мм.
3. При работе двумя руками органы управления размещают с таким расчетом, чтобы не было перекрещивания рук.
4. Очень часто используемые средства отображения информации, требующие точного и быстрого считывания показаний, следует располагать в вертикальной плоскости под углом $\pm 15^\circ$ от нормальной линии взгляда и в горизонтальной плоскости под углом $\pm 15^\circ$ от сагиттальной плоскости.

Организация правильного освещения рабочего места является немаловажным, что регулирует СП 52.13330.2016 «Естественное и искусственное освещение» [10].

При выполнении выпускной квалификационной работы правовых и организационных нарушений по указанным требованиям не было выявлено, рабочее место было оборудовано согласно всем нормам и правилам.

6.2. Производственная безопасность

Появление вредных и опасных факторов производства может быть спровоцировано условиями труда и устройством, с помощью которого осуществляется разработка системы.

Согласно «ГОСТ 12.0.003-2015 Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы. Классификация», при выполнении работ на персональном компьютере воздействуют факторы, представленные в таблице 17 [11].

Таблица 17 – Возможные опасные и вредные производственные факторы на рабочем месте разработчика мобильных приложений

Факторы (ГОСТ 12.0.003-2015)	Нормативные документы
Отсутствие или недостаток необходимого естественного и искусственного освещения	СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95*
Аномальные микроклиматические параметры воздушной среды на местонахождении рабочего	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания»
Нервно-психические перегрузки, связанные с монотонностью труда	Трудовой кодекс Российской Федерации от 30.12.2001 №197-ФЗ (ред. от 25.02.2022)
Нервно-психические нагрузки, связанные с перенапряжением зрительных анализаторов	СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания
Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает рабочий	ГОСТ 12.1.038-82 Система стандартов безопасности труда (ССБТ). Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов

6.2.1. Отсутствие или недостаток необходимого естественного и искусственного освещения

Недостаточная освещенность приводит к понижению работоспособности, а также может вызвать проблемы со здоровьем, а именно может повлиять на качество зрения работников.

Согласно СП 52.13330.2016 [10], зрительную работу разработчика программного обеспечения можно характеризовать как работу разряда Б – высокой точности (наименьший эквивалентный размер объекта различения

составляет 0,3-0,5 мм), подразряда 1 (относительная продолжительность зрительной работы при направлении зрения на рабочую поверхность не менее 70%). В таблице 18 представлены требования к освещению рабочего помещения для указанного разряда.

Таблица 18 – Требования к освещению рабочего помещения для разряда Б1

Искусственное освещение				Естественное освещение	
Освещенность на рабочей поверхности от системы общего освещения, лк	Цилиндрическая освещенность, лк	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

Искусственное освещение в помещениях для эксплуатации персонального компьютера должно осуществляться системой общего равномерного освещения. В случаях преимущественной работы с персональным компьютером следует применять системы комбинированного освещения. Окна в помещениях, где эксплуатируется вычислительная техника, преимущественно должны быть ориентированы на север и северо-восток.

Если следовать приведенным мерам, то удастся минимизировать вероятность нарушения зрительной функции.

6.2.2. Аномальные микроклиматические параметры воздушной среды на местонахождении рабочего

Данный фактор можно отнести к вредным, так как его воздействие на организм человека проявляется в виде быстрой утомляемости, потери работоспособности сотрудника.

Работа, выполняемая разработчиком, выполняется в помещениях категории 2, так как является умственным трудом и выполняется в положении сидя с минимальными физическими напряжениями. В таблице 19 представлены оптимальные и допустимые значения показателей

микроклимата на рабочих местах для данной категории согласно СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» пункт 29 [12].

Таблица 19 – Оптимальные величины показателей микроклимата

Период года	Температура воздуха, С ⁰		Результирующая температура, С ⁰		Относительная влажность, %		Скорость движения воздуха, м/с	
	Оптимальная	Допустимая	Оптимальная	Допустимая	Оптимальная	Допустимая	Оптимальная, не более	Допустимая, не более
Холодный	19-21	18-23	18-20	17-22	45-30	60-30	0,2	0,3
Теплый	23-25	18-28	22-24	19-27	60-30	65-30	0,15	0,25

6.2.3. Нервно-психические перегрузки, связанные с монотонностью труда

Разработка информационной системы тесно связана с малоподвижной, однотипной деятельностью, что является показателем монотонного труда.

Для снижения уровня монотонности работы разработчика программного обеспечения следует принимать следующие меры:

- во время рабочего дня делать регулярные перерывы;
- вне рабочего времени стоит проводить время на свежем воздухе;
- стараться организовывать свое рабочее время таким образом, чтобы не заниматься одной задачей на протяжении длительного времени.

6.2.4. Нервно-психические перегрузки, связанные с перенапряжением зрительных анализаторов

Работа на ПК сопровождается постоянным и значительным напряжением функций зрительного анализатора. Одной из основных особенностей является иной принцип чтения информации, чем при обычном чтении. Допустимые уровни ультрафиолетового излучения, создаваемые изделиями, в соответствии с СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» указаны в таблице 20 [12].

Таблица 20 – Допустимые уровни ультрафиолетового излучения

Вид изделий	Спектральный диапазон длин волн, нм	Допустимая интенсивность облучения, Вт/м ²
Экраны телевизоров, видеомониторов, осциллографов измерительных и других приборов, средств отображения информации с визуальным контролем	Свыше 315 до 400	Не более 0,1
	Свыше 280 до 315	Не более 0,0001
	От 200 до 280	Не допускается

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

В деятельности организации широко используется электричество для питания компьютерной техники, которая может являться источником опасности. Поражение электрическим током может произойти при прикосновении к токоведущим частям, находящимся под напряжением, на которых остался заряд или появилось напряжение. Электрический ток оказывает на человека термическое, электролитическое, биологическое и механическое воздействия, что может привести к травмам, таким как ожоги, или гибели.

ГОСТ 12.1.038-82 ССБТ. «Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов» устанавливает ПДЗ напряжений прикосновения и токов, предназначенные для проектирования способов и средств защиты людей [14]. Так, для переменного тока частотой 50 Гц допустимое значение напряжения прикосновения составляет 2 В, а силы тока – 0,3 мА, для тока частотой 400 Гц, соответственно – 2 В и 0,4 мА, для постоянного тока – 8 В и 1 мА.

Мерами защиты от воздействия электрического тока являются оградительные устройства, устройства автоматического контроля и сигнализации, изолирующие устройства и покрытия, устройства защитного

заземления, устройства автоматического отключения, предохранительные устройства.

Пользование мерами защиты, приведенными выше, позволяют избежать поражения электрическим током.

6.3. Экологическая безопасность

Разработка информационной системы несет минимальное воздействие на литосферу и по большей части сводится к обычному бытовому мусору и отходам жизнедеятельности человека.

В процессе работы за своим рабочим местом образуется бытовой мусор: искусственные источники освещения, канцелярский мусор, пищевые отходы и т.д. Каждая категория мусора должна быть утилизирована в соответствии в определенным классом опасности или переработан, чтобы не оказывать негативное влияние на состояние литосферы.

Вышедшие из строя ПЭВМ и сопутствующая оргтехника относится к IV классу опасности и подлежит специальной утилизации. В жидкокристаллических мониторах с диагональю 26 дюймов содержится от 2 до 4 U-образных ламп массой по 13 г, которые содержат ртуть в количестве 62,14 мг/кг, что превышает норму допустимого содержания в почве в 30 раз. Неутилизированные компьютеры и оргтехника вызывают ртутное загрязнение литосферы и гидросферы, а при сжигании – и атмосферы. Для оказания наименьшего влияния на окружающую среду, необходимо проводить специальную процедуру утилизации ПЭВМ и оргтехники, при которой 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» [15].

При выполнении выпускной квалификационной работы техника не выходила из строя и не нуждалась в утилизации, поэтому данный источник

загрязнения не воздействовал на окружающую среду и является лишь предполагаемым.

6.4. Безопасность в чрезвычайных ситуациях

Выпускная квалификационная работа студента по проектированию и разработке системы для диагностики и предсказания уровня заболеваемости COVID 19 в странах мира выполнялась в офисе. Перечень возможных чрезвычайных ситуаций для данного типа помещения представлен ниже:

- природные катастрофы (наводнения, цунами, ураган и т.д.);
- геологические воздействия (землетрясения, оползни, обвалы, провалы территории и т.д.);
- техногенные аварии (пожар).

Одной из наиболее возможных чрезвычайных ситуаций в помещениях такого типа является пожар. Пожаром называется неконтролируемое горение во времени и пространстве, наносящее материальный ущерб и создающее угрозу жизни и здоровью людей.

Причинами возгорания при работе с компьютером могут быть:

- токи короткого замыкания;
- неисправность устройства компьютера или электросетей;
- небрежность оператора при работе с компьютером;
- воспламенение ПК из-за перегрузки.

В связи с этим, согласно ГОСТ 12.1.004-91 «Система стандартов безопасности труда. Пожарная безопасность. Общие требования», при работе с компьютером необходимо соблюдать следующие нормы пожарной безопасности [16]:

- Пожарная безопасность объекта должна обеспечиваться системами предотвращения пожара и противопожарной защиты, в том числе организационно-техническими мероприятиями.

– Для предохранения сети от перегрузок запрещается одновременно подключать к сети количество потребителей, превышающих допустимую нагрузку

– Объекты должны иметь системы пожарной безопасности, направленные на предотвращение воздействия на людей опасных факторов пожара, в том числе их вторичных проявлений, на требуемом уровне.

– Работы за компьютером проводить только при исправном состоянии оборудования, электропроводки

Если все же не удалось предотвратить пожар, то каждый сотрудник должен:

– незамедлительно сообщить об это в пожарную охрану;

– принять меры по эвакуации людей, каких-либо материальных ценностей согласно плану эвакуации;

– отключить электроэнергию, приступить к тушению пожара первичными средствами пожаротушения.

На основании Федерального закона от 22.07.2008 N 123-ФЗ (ред. от 30.04.2021) "Технический регламент о требованиях пожарной безопасности" наиболее вероятно при разработке информационной системы возникновение пожара горючих веществ и материалов электроустановок, находящихся под напряжением – класс Е. Первичными средствами пожаротушения данного класса пожара будут являться огнетушители, маркировка которых содержит букву Е.

Вывод по разделу

В результате выполнения работы по разделу ВКР «Социальная ответственность» были изучены правовые и организационные вопросы обеспечения безопасности, правовые нормы трудового законодательства и требования к организации рабочей зоны, в которой проводилась разработка информационной системы. Была рассмотрена производственная безопасность, в т.ч. проанализированы вредные и опасные факторы, которые могут

возникнуть при разработке и эксплуатации проектируемого решения, и предложены мероприятия по защите разработчика от их действия. Проанализирован характер воздействия проектируемого решения на окружающую среду и возможные чрезвычайные ситуаций, которые могут возникнуть при разработке и эксплуатации мобильного приложения.

Были выявлены следующие показатели:

- категория помещения по электробезопасности согласно ПУЭ – 1, нет условий для возникновения повышенной или особой опасности;

- сами разработчики относятся к первой группе по электробезопасности согласно Правил по охране труда при эксплуатации электроустановок, однако должностные лица, осуществляющие контроль и надзор за соблюдением требований безопасности при эксплуатации электроустановок в офисном помещении, должны иметь группу по электробезопасности не ниже IV;

- 2 категория помещений по СанПиН 1.2.3685-21 "Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания";

- категория помещений по взрывопожарной и пожарной опасности согласно СП 12.13130.2009 «Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности» – В4, так как удельная пожарная нагрузка составила примерно 43 МДж/м².

- вышедшие из строя ПЭВМ и сопутствующая оргтехника относится к IV классу опасности и подлежит специальной утилизации.

Заключение

В результате выпускной квалификационной работы был исследован набор данных, содержащий ежедневные данные статистики по коронавирусной инфекции.

Визуализация статистики позволяет наглядно видеть эпидемиологическую обстановку в каждой стране. На ее основе можно принимать решения о посещении той или иной страны.

Результатом работы являются:

1. Анализ и предобработка данных, состоящая из:
 - обнаружения и устранения пропущенных значений,
 - обнаружения и устранения выбросов,
 - устранения зависимых признаков,
 - удаления дубликатов строк.
2. Проанализированный метод k-средних
3. Реализован веб-сайт, визуализирующий уровни опасности заражения COVID-19 разных стран мира.

Данная платформа была разработана для туристов, желающих посетить какую-либо страну, с целью улучшения и автоматизации анализа статистики заражений вирусом COVID-19 стран мира.

На этапе финансового анализа были выявлены конкурентные черты разработки собственного решения, бюджет и сроки реализации.

На этапе анализа данных социальной ответственности было отмечено отсутствие нарушений при выполнении выпускной квалификационной работы по различным аспектам в области безопасности

Список литературы

1. Big Data [Электронный ресурс]. 2022. URL: <https://secretmag.ru/enciklopediya/chto-takoe-big-data-obyasnyаем-prostymi-slovami.htm> (дата обращения: 04.06.2022)
2. Страховые компании и Big Data [Электронный ресурс]. 2022. URL: https://www.cnews.ru/news/top/strahovye_kompanii_vstayut_na_put_big_data (дата обращения: 04.06.2022)
3. Ассоциация туроператоров [Электронный ресурс]. 2022. URL: <https://www.atorus.ru/news/press-centre/new/58410.html> (дата обращения: 04.06.2022)
4. Коэффициент распространения коронавируса (Rt) [Электронный ресурс]. 2022. URL: <https://gogov.ru/articles/covid-rt> (дата обращения: 04.06.2022)
5. Matplotlib [Электронный ресурс]. 2022. URL: <https://blog.skillfactory.ru/glossary/matplotlib/> (дата обращения: 04.06.2022)
6. Алгоритм К-средних [Электронный ресурс]. 2022. URL: <https://axd.semestr.ru/upr/average.php> (дата обращения: 04.06.2022)
7. Модели машинного обучения [Электронный ресурс]. 2022. URL: <https://vc.ru/ml/126354-modeli-mashinnogo-obucheniya-obyasnyаем-pyatiletnemu-rebenku> (дата обращения: 04.06.2022)
8. Трудовой кодекс Российской Федерации // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2022. URL: <https://docs.cntd.ru/document/901807664> (дата обращения: 27.05.2022)
9. 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]. 2022. URL: <https://docs.cntd.ru/document/1200003913> (дата обращения: 17.05.2022)

10. СП 52.13330.2016. Естественное и искусственное освещение // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2022. URL: <https://docs.cntd.ru/document/456054197> (дата обращения: 27.05.2022)
11. ГОСТ 12.0.003-2015. Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы. Классификация. – М.: Стандартинформ, 2019. – URL: <https://docs.cntd.ru/document/1200136071> (дата обращения: 27.05.2022)
12. СанПиН 1.2.3685-21. Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания. – М., 2021. – URL: <https://docs.cntd.ru/document/573500115> (дата обращения: 27.05.2022)
13. ГОСТ 12.1.019-2017. Система стандартов безопасности труда (ССБТ). Электробезопасность. Общие требования и номенклатура видов защиты. – URL: <https://docs.cntd.ru/document/1200161238> (дата обращения: 27.05.2022)
14. ГОСТ 12.1.038-82 ССБТ. «Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2022. URL: <https://docs.cntd.ru/document/5200313> (дата обращения: 27.05.2022)
15. ГОСТ Р 53692-2009 Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2022. URL: <https://docs.cntd.ru/document/1200081740> (дата обращения: 27.05.2022)
16. ГОСТ 12.1.004-91 «Система стандартов безопасности труда. Пожарная безопасность. Общие требования // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2022. URL: <https://docs.cntd.ru/document/9051953> (дата обращения: 27.05.2022)

Приложение А

№	Название	Перевод	Значение
1	iso_code	Код страны	
2	continent	Континент	
3	location	Страна	
4	date	Дата	
5	total_cases	Всего случаев	Всего подтвержденных случаев COVID-19
6	new_cases	Новые случаи	Новые подтвержденные случаи COVID-19
7	new_cases_smoothed	Новые случаи (среднее)	Новые подтвержденные случаи COVID-19 (среднее за 7 дней)
8	total_deaths	Всего смертей	Всего смертей от COVID-19
9	new_deaths	Новые смерти	Новые смерти от COVID-19
10	new_deaths_smoothed	Новые смерти (среднее)	Новые смерти от COVID-19 (среднее за 7 дней)
11	total_cases_per_million	Всего случаев на миллион	Всего подтвержденных случаев COVID-19 на миллион человек
12	new_cases_per_million	Новые случаи на миллион	Новые подтвержденные случаи COVID-19 на миллион человек
13	new_cases_smoothed	Новые случаи на миллион	Новые подтвержденные случаи COVID-19 на миллион человек (среднее за 7 дней)

	_per_million	миллион (среднее)	
14	total_deaths_per_million	Всего смертей на миллион	Всего смертей от COVID-19 на миллион человек
15	new_deaths_per_million	Новые смерти на миллион	Новые смерти от COVID-19 на миллион человек
16	new_deaths_smoothed_per_million	Новые смерти на миллион (среднее)	Новые смерти от COVID-19 на миллион человек (среднее за 7 дней)
17	reproduction_rate	Скорость воспроизведения	Оценка эффективной скорости размножения COVID-19 в режиме реального времени
18	icu_patients	ОРИТ пациенты	Количество пациентов с COVID-19 в отделениях интенсивной терапии в данный день
19	icu_patients_per_million	ОРИТ пациенты на миллион	Количество пациентов с COVID-19 в отделениях интенсивной терапии в данный день на миллион человек
20	hospital_patients	Пациенты в больнице	Количество пациентов с COVID-19 в больнице в данный день
21	hospital_patients_per_million	Пациенты в больнице на миллион	Количество пациентов с COVID-19 в больнице в данный день на миллион
22	weekly_icu_admissions	ОРИТ пациенты в неделю	Количество пациентов с COVID-19 в отделениях интенсивной терапии за данную неделю

23	weekly_icu_admissions_per_million	ОРИТ пациенты в неделю на миллион	Количество пациентов с COVID-19 в отделении интенсивной терапии за данную неделю на миллион
24	weekly_hospital_admissions	Пациенты в больнице в неделю	Количество пациентов с COVID-19 в больнице за данную неделю
25	weekly_hospital_admissions_per_million	Пациенты в больнице в неделю на миллион	Количество пациентов с COVID-19 в больнице за данную неделю на миллион человек
26	new_tests	Новые тесты	Новые тесты на COVID-19 (расчитаны только на последующие дни)
27	total_tests	Всего тестов	Всего тестов на COVID-19
28	total_tests_per_thousand	Всего тестов на тысячу	Всего тестов на COVID-19 на тысячу человек
29	new_tests_per_thousand	Новые тесты на тысячу	Новые тесты на COVID-19 на тысячу человек
30	new_tests_smoothed	Новые тесты (среднее)	Новые тесты на COVID-19 (среднее за 7 дней)
31	new_tests_smoothed_per_thousand	Новые тесты на тысячу (среднее)	Новые тесты на COVID-19 на тысячу человек (среднее за 7 дней)

32	positive_rate	Положительные тесты	Доля положительных тестов на COVID-19, указанная как скользящее среднее значение за 7 дней (обратное значение text_per_case)
33	tests_per_case	Тестов на случай	Тесты, проведенные на каждый новый подтвержденный случай COVID-19, представленный как среднее значение за 7 дней (обратное значение positive_rate)
34	tests_units	Тестовые единицы	Единицы, используемые местоположением для сообщения данных тестирования
35	total_vaccinations	Всего прививок	Общее количество введенных доз вакцины против COVID-19
36	people_vaccinated	Привитые люди	Общее число людей, получивших хотя бы одну дозу вакцины
37	people_fully_vaccinated	Полностью вакцинированные люди	Общее количество людей, получивших все дозы вакцины от COVID-19, предусмотренных протоколом вакцинации
38	total_boosters	Всего бустеров	Общее количество введенных доз вакцины против COVID-19, введенных сверх количества, предусмотренного протоколом вакцинации
39	new_vaccinations	Новые прививки	Новые введенные вакцины (рассчитаны только за последующие дни)
40	new_vaccinations_smoothed	Новые прививки (среднее)	Новые введенные вакцины (среднее за 7 дней)
41	total_vaccinations_per_hundred	Всего вакцин на сотню	Общее количество доз вакцины против COVID-19, введенных на 100 человек

42	people_vaccinated_per_hundred	Привитые люди на сотню	Общее количество людей, получивших хотя бы одну дозу вакцины против COVID-19, на сотню
43	people_fully_vaccinated_per_hundred	Люди, полностью вакцинированные, на сотню	Общее количество людей, получивших все дозы, предусмотренные протоколом первичной вакцинации, на сотню человек
44	total_boosters_per_hundred	Всего бустеров на сотню	Общее количество доз вакцины против COVID-19, введенных сверх количества, на сотню человек
45	new_vaccinations_smoothed_per_million	Новые прививки на сотню (среднее)	Новые введенные вакцины на сотню (среднее за 7 дней)
46	new_people_vaccinated_smoothed	Люди, получившие первую дозу (среднее)	Ежедневное количество людей, получивших первую дозу вакцины (среднее за 7 дней)
47	new_people_vaccinated_smoothed_per_hundred	Люди, получившие первую дозу на сотню (среднее)	Ежедневное количество людей, получивших первую дозу вакцины (среднее за 7 дней) на 100 человек

48	stringency_index	Индекс строгости	Индекс строгости ответных мер правительства: составной показатель
49	population	Население	
50	population_density	Плотность населения	Количество людей на площадь земли
51	median_age	Средний возраст	Средний возраст населения
52	aged_65_older	Старше 65	Доля населения в возрасте 65 лет и старше
53	aged_70_older	Старше 70	Доля населения в возрасте 70 лет и старше
54	gdp_per_capita	ВВП на душу населения	ВВП на душу населения
55	extreme_poverty	Крайняя бедность	Доля населения, живущего в условиях крайней нищеты
56	cardiovascular_death_rate	Смертность от сердечно-сосудистых заболеваний	Смертность от сердечно-сосудистых заболеваний в 2017 году
57	diabetes_prevalence	Распространенность диабета	Распространенность диабета
58	female_smokers	Курильщик и женщины	Доля курящих женщин
59	male_smokers	Курильщик и мужчины	Доля курящих мужчин

60	handwashing_facilities	Устройства для мытья рук	Доля населения, имеющего базовые приспособления для мытья рук в помещениях
61	hospital_beds_per_thousand	Больничных коек на тысячу	Больничные койки на тысячу человек
62	life_expectancy	Продолжительность жизни	Ожидаемая продолжительность жизни при рождении
63	human_development_index	Индекс развития человека	Составной индекс, измеряющий средние достижения по трем основным параметрам человеческого развития: долгая и здоровая жизнь, знания и достойный уровень жизни
64	excess_mortality_cumulative_absolute	Абсолютная избыточная совокупная смертность	Совокупная разница между зарегистрированным числом смертей с 1 января 2020 г. и прогнозируемым числом смертей за тот же период на основе предыдущих лет
65	excess_mortality_cumulative	Избыточная совокупная смертность	Процентная разница между совокупным числом смертей с 1 января 2020 г. и совокупным прогнозируемым числом смертей за тот же период на основе предыдущих лет
66	excess_mortality	Избыточная смертность	Процентная разница между зарегистрированным количеством еженедельных или ежемесячных смертей в 2020–2021 гг. и прогнозируемым количеством смертей за тот же период на основе предыдущих лет

67	excess_mortality_cumulative_per_million	Избыточная совокупная смертность на миллион	Совокупная разница между зарегистрированным числом смертей с 1 января 2020 года и прогнозируемым числом смертей за тот же период на основе предыдущих лет на миллион человек
----	---	---	--