



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

Школа – Инженерная школа ядерных технологий

Направление подготовки – 01.03.02 Прикладная математика и информатика

ООП/ОПОП – Прикладная математика в инженерии

Отделение школы (НОЦ) – Отделение экспериментальной физики

### ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема работы
<i>Создание экспертной рекомендательной системы для клиентов компании</i>

УДК 005.936:005.334:330.322.2

Обучающийся

Группа	ФИО	Подпись	Дата
0В91	Самородний Кирилл Андреевич		

Руководитель ВКР

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

### КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность»

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

По разделу «Социальная ответственность»

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

### ДОПУСТИТЬ К ЗАЩИТЕ:

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

Томск – 2023 г.



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

Школа – Инженерная школа ядерных технологий

Направление подготовки – 01.03.02 Прикладная математика и информатика

ООП/ОПОП – Прикладная математика в инженерии

Отделение школы (НОЦ) – Отделение экспериментальной физики

УТВЕРЖДАЮ:

Руководитель ООП/ОПОП

\_\_\_\_\_ Крицкий О.Л.  
(Подпись) \_\_\_\_\_ (Дата) \_\_\_\_\_ (ФИО)

**ЗАДАНИЕ  
на выполнение выпускной квалификационной работы**

Обучающийся:

Группа	ФИО
0В91	Самородний Кирилл Андреевич

Тема работы:

<i>Создание экспертной рекомендательной системы для клиентов компании</i>	
<i>Утверждена приказом директора (дата, номер)</i>	

Срок сдачи обучающимся выполненной работы:	
--	--

**ТЕХНИЧЕСКОЕ ЗАДАНИЕ:**

<b>Исходные данные к работе</b>	Таблица, с данными о пользователях, которая содержит идентификатор пользователя, игру, способ взаимодействия с игрой и количество часов взаимодействия.
<b>Перечень разделов пояснительной записки подлежащих исследованию, проектированию и разработке</b>	<ol style="list-style-type: none"> <li>1. Создать рекомендательную систему для отдельных игр.</li> <li>2. Применить метод ближайших соседей для создания рекомендательной системы для пользователей.</li> <li>3. Оценить результаты работы алгоритма.</li> </ol>
<b>Перечень графического материала</b>	
<b>Консультанты по разделам выпускной квалификационной работы</b> <i>(с указанием разделов)</i>	
Раздел	Консультант

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

<b>Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику</b>	
---	--

**Задание выдал руководитель:**

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

**Задание принял к исполнению обучающийся:**

Группа	ФИО	Подпись	Дата
0В91	Самородний Кирилл Андреевич		

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

Студенту:

<b>Группа</b>	<b>ФИО</b>
0B91	Самородному Кириллу Андреевичу

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

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

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

**Перечень вопросов, подлежащих исследованию, проектированию и разработке:**

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

**Перечень графического материала**

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

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

**Задание выдал консультант:**

<b>Должность</b>	<b>ФИО</b>	<b>Ученая степень, звание</b>	<b>Подпись</b>	<b>Дата</b>
Доцент ОСГН ШБИП	Кащук Ирина Вадимовна	К.Т.Н доцент		

**Задание принял к исполнению студент:**

<b>Группа</b>	<b>ФИО</b>	<b>Подпись</b>	<b>Дата</b>
0B91	Самородный Кирилл Андреевич		

## ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

<b>Группа</b> 0В91		<b>ФИО</b> Самородному Кириллу Андреевичу	
<b>Школа</b>	<b>Инженерная школа ядерных технологий</b>	<b>Отделение (НОЦ)</b>	<b>Экспериментальной физики</b>
<b>Уровень образования</b>	Бакалавриат	<b>Направление/специальность</b>	01.03.02 Прикладная математика и информатика

Тема ВКР:

<b>Создание экспертной рекомендательной системы для клиентов компании</b>	
<b>Исходные данные к разделу «Социальная ответственность»:</b>	
<p><b>Введение</b></p> <ul style="list-style-type: none"> <li>– Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения.</li> <li>– Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации</li> </ul>	<p><i>Объект исследования:</i> алгоритм для разработки экспертной рекомендательной системы.  <i>Область применения:</i> областью применения данного алгоритма может выступать любая компания, где применимы рекомендательные системы.  <i>Рабочая зона:</i> офис  <i>Размеры помещения:</i> 27,2 м<sup>2</sup>  <i>Количество и наименование оборудования рабочей зоны:</i> 1 персональный компьютер  <i>Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне:</i> алгоритмическая и программная разработка с использованием персонального компьютера</p>
<b>Перечень вопросов, подлежащих исследованию, проектированию и разработке:</b>	
<p><b>1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения:</b></p> <ul style="list-style-type: none"> <li>– специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства;</li> <li>– организационные мероприятия при компоновке рабочей зоны.</li> </ul>	<ul style="list-style-type: none"> <li>– ГОСТ 12.2.032-78 «Рабочее место при выполнении работ сидя»</li> <li>– СП 2.4.3648-20 "Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи"</li> <li>– Трудовой кодекс Российской Федерации: федер. Закон от 30 дек. 2001 г. №197-ФЗ Раздел 10</li> </ul>
<p><b>2. Производственная безопасность при разработке проектного решения:</b></p> <ul style="list-style-type: none"> <li>– Анализ выявленных вредных и опасных производственных факторов</li> </ul>	<p><b>Опасные факторы:</b></p> <ul style="list-style-type: none"> <li>– Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий;</li> </ul> <p><b>Вредные факторы:</b></p> <ul style="list-style-type: none"> <li>– Отклонение показателей микроклимата;</li> <li>– Недостаточная освещённость рабочей зоны;</li> <li>– Пониженная световая и цветовая контрастность;</li> <li>– Повышенный уровень шума на рабочем месте;</li> <li>– Длительное сосредоточенное наблюдение</li> </ul>
<p><b>3. Экологическая безопасность при разработке проектного решения:</b></p>	<p><b>Анализ воздействия на литосферу:</b></p> <ul style="list-style-type: none"> <li>– Утилизация компьютеров, оргтехники и бумаги;</li> </ul> <p><b>Анализ воздействия на гидросферу:</b></p> <ul style="list-style-type: none"> <li>– Производство компьютерной техники;</li> </ul>

<b>4. Безопасность в чрезвычайных ситуациях при разработке проектного решения:</b>	– Возможность затоплений при авариях в водоснабжении; – Наиболее типичная ЧС: Пожар;
<b>Дата выдачи задания для раздела по линейному графику</b>	

**Задание выдал консультант:**

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

**Задание принял к исполнению студент:**

Группа	ФИО	Подпись	Дата
0В91	Самородний Кирилл Андреевич		

## **Реферат**

Выпускная квалификационная работа выполнена на 65 страницах, содержит 10 рисунков, 19 таблиц, 38 источников и одно приложение.

*Ключевые слова:* Экспертная система, рекомендательная система, нейронные сети.

*Объект исследования:* Разработанная экспертная рекомендательная система.

*Цель работы:* Теоретическое изучение темы и практическая реализация проектирования рекомендательной системы для клиентов компании.

*Методы проведения работы:* Теоретические (изучение литературы, обзор методов и моделей анализа) и практические (применение методов для создания рекомендательной системы).

*В результате исследования:* Разработана рекомендательная система, использующая машинное обучение и нейросети, для взаимодействия с пользователями онлайн-сервиса доступа к компьютерным играм.

## **Определения, обозначения, сокращения и нормативные ссылки**

**В данной работе приведены следующие термины с соответствующими определениями:**

Экспертная система — предиктивная система, включающая в себя знания об определенной слабо структурированной и трудно формализуемой узкой предметной области и способная предлагать и объяснять пользователю разумные решения

Рекомендательные системы — это комплекс сервисов и программ, который анализирует предпочтения пользователей и пытается прогнозировать, что может их заинтересовать.

Python — это интерпретируемый, объектно-ориентированный язык программирования высокого уровня, с помощью которого можно выполнять широкий спектр задач. Его используют для обработки числовой и текстовой информации, создания изображений, работы с базами данных, разработки web-сайтов и приложений с графическим интерфейсом

База данных (БД) — это набор информации, которая хранится упорядоченно в электронном виде.

Искусственный интеллект (ИИ) — это одно из направлений информатики, целью которого является разработка аппаратнопрограммных средств, позволяющих пользователю-непрограммисту ставить и решать свои, традиционно считающиеся интеллектуальными, задачи, общаясь с компьютером на ограниченном подмножестве естественного языка.

Интеллектуальная система — автоматизированная система, основанная на знаниях, или комплекс программных, лингвистических и логико-математических средств для реализации основной задачи — осуществления поддержки деятельности человека и поиска информации в режиме продвинутого диалога на естественном языке



## Оглавление

Введение .....	10
1. Обзор литературы .....	11
2. Теоретическая часть .....	15
2.1 Машинное обучение .....	15
2.2 Анализ данных .....	17
2.3 Факторы влияющие на рекомендации .....	17
3. Практическая часть .....	20
4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение .....	28
4.1 Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения .....	28
4.1.1 Анализ конкурентных технических решений .....	28
4.1.2 SWOT-анализ .....	30
4.2 Планирование научно-исследовательских работ .....	32
4.2.1 Структура работ в рамках научного исследования .....	32
4.2.2 Определение трудоемкости выполнения работ и разработка графика проведения .....	33
4.3 Бюджет научно-технического исследования .....	36
4.4 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования .....	41
Выводы по разделу .....	44
5. Социальная ответственность .....	45
5.1 Правовые и организационные вопросы обеспечения безопасности .....	45
5.2 Производственная безопасность .....	47
5.2.1 Отклонение показателей микроклимата .....	48
5.2.2 Недостаточная освещенность рабочей зоны .....	50
5.2.3 Повышенная световая и цветовая контрастность .....	51
5.2.4 Повышенный уровень шума на рабочем месте .....	51
5.2.5 Длительное сосредоточенное наблюдение .....	52
5.2.6 Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий .....	53
5.3 Экологическая безопасность .....	54
5.4 Безопасность в чрезвычайных ситуациях .....	56
5.4.1 Пожар .....	56
Вывод по разделу .....	57
Заключение .....	59
Список используемых источников .....	60
Приложение 1 .....	63

## **Введение**

В наше время тяжело представить какой-либо сайт или приложение без рекомендательной системы. Рекомендации преследуют нас везде, они подбирают нам видеоролики, музыку, фильмы на основе уже просмотренного или прослушанного, а стоит заговорить о том, что куртка износилась и нужна новая, так на следующий день на сайте появится реклама с интернет-магазином, который предложит несколько интересных вариантов.

Все описанные алгоритмы работают благодаря нейронным сетям.

Целью работы является проектирование экспертной рекомендательной системы для клиентов компании с помощью машинного обучения. Для достижения поставленной цели необходимо решить следующие задачи:

- 1) Собрать данные и классифицировать пользователей.
- 2) Выбрать оптимальный метод экспертной оценки.
- 3) Обучить модель для рекомендательной системы.
- 4) Сравнить полученные результаты с реальными данными.
- 5) Оценить результаты работы

## 1. Обзор литературы

В источнике [2] представлены виды рекомендательных систем и как они помогают бизнесу и людям.

Рекомендации могут быть персонализированными и не персонализированными. В первом случае мы знаем что-то о пользователе и основываемся на его предпочтениях, во втором — используем популярные или сопутствующие товары.

Выделяют четыре вида рекомендательных систем:

**Контентная фильтрация.** Контентная фильтрация основывается на потребленном контенте.

**Коллаборативная фильтрация.** В коллаборативной фильтрации не происходит сопоставления интересов пользователя с объектами, а происходит поиск похожих друг на друга пользователей. Если два клиента компании дают одинаковую оценку для одного элемента или ведут себя одинаково, значит они будут оценивать другие элементы таким же образом. То есть похожим пользователям нравятся одинаковые вещи.

**Фильтрация, основанная на знаниях.** В системе, основанной на знаниях рекомендации строятся на основе экспертного мнения. В отличие от рекомендаций в предыдущих системах они чаще всего не персонализированные.

**Гибридные рекомендательные системы.** В гибридной системе используется сразу несколько рассмотренных выше методов. Существует множество различных типов гибридных систем, но будем рассматривать два из них: Аугментация, где выходные данные одного метода используются как входные данные для другого и каскадирование, где одна рекомендательная система уточняет данные для другой.

В источнике [11] рассматриваются основные метрики, которые нам понадобятся для расчета:

Самый распространенный пример меры расстояния — это евклидово расстояние:

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (1.1)$$

где  $n$  - количество измерений (атрибутов), а  $x_k$  и  $y_k$  -  $k$ -е атрибуты объектов данных  $x$  и  $y$ .

Введем обозначения: индекс  $i$  всегда будет обозначать пользователей (всего пользователей будет  $N, i=1..N$ ); индекс  $a$  – предметы (игры), которые мы рекомендуем (всего  $M, a=1..M$ ); когда пользователь  $i$  оценивает предмет  $a$ , он производит отклик (response, rating)  $r_{i,a}$ , этот отклик – случайная величина. Наша задача – прогнозировать оценки  $r_{i,a}$ , зная признаки  $x_i$  и  $x_a$  для всех элементов базы и зная некоторые уже расставленные в базе  $r_{i',a'}$ . Прогноз будем обозначать через  $\hat{r}_{i,a}$ .

Метод ближайших соседей: рассчитывает расстояние между пользователями и будем рекомендовать то, что нравится вашим соседям.

Коэффициент корреляции (коэффициент Пирсона):

$$\omega_{i,j} = \frac{\sum_a (r_{i,a} - \bar{r}_a)(r_{j,a} - \bar{r}_a)}{\sqrt{\sum_a (r_{i,a} - \bar{r}_a)^2} \sqrt{\sum_a (r_{j,a} - \bar{r}_a)^2}} \quad (1.2)$$

Где  $\bar{r}_a$ , средний рейтинг продукта  $a$ , среди всех пользователей.

Косинус угла между векторами рейтингов:

$$\omega_{i,j} = \frac{\sum_a r_{i,a} r_{j,a}}{\sqrt{\sum_a r_{i,a}^2} \sqrt{\sum_a r_{j,a}^2}} \quad (1.3)$$

Простейший способ построить прогноз нового рейтинга  $\hat{r}_{i,a}$  – сумма рейтингов других пользователей, взвешенная их похожестью на пользователя  $i$ :

$$\hat{r}_{i,a} = \bar{r}_a + \frac{\sum_j (r_{j,a} - \bar{r}_j) \omega_{i,j}}{\sum_j |\omega_{i,j}|} \quad (1.4)$$

А чтобы не суммировать по всем пользователям, можно ограничиться ближайшими соседями:

$$\hat{r}_{i,a} = \bar{r}_a + \frac{\sum_{j \in km(i)} (r_{j,a} - \bar{r}_j) \omega_{i,j}}{\sum_{j \in km(i)} |\omega_{i,j}|} \quad (1.5)$$

Для хранения данных о пользователе используется база данных.

В источниках [7] и [8] описан принцип работы и виды баз данных.

Базы данных позволяют обрабатывать, хранить и структурировать намного большие объёмы информации, чем таблицы.

Удалённый доступ и система запросов позволяет множеству людей одновременно использовать базы данных. С электронными таблицами тоже можно работать онлайн всей командой, но системы управления базами данных делают этот процесс организованнее, быстрее и безопаснее.

Объём информации в базах данных может быть огромным и не влиять на скорость работы. А в Google Таблицах уже после нескольких сотен строк или тысяч символов страница будет загружаться очень медленно.

В нашей работе будем использовать реляционную базу данных, основные определения представлены в [8].

Отношение находится в 1НФ, если все его атрибуты являются простыми, все используемые домены должны содержать только скалярные значения. Не должно быть повторений строк в таблице.

Отношение находится во 2НФ, если оно находится в 1НФ и каждый не ключевой атрибут неприводимо зависит от Первичного Ключа.

Отношение находится в 3НФ, когда находится во 2НФ и каждый не ключевой атрибут нетранзитивно зависит от первичного ключа. Проще говоря, второе правило требует выносить все не ключевые поля, содержимое которых может относиться к нескольким записям таблицы в отдельные таблицы.

В источнике [3] подробно описано то, как взаимодействовать с базами данных, как переносить базу данных в python, а также описаны все преимущества данного языка программирования.

Принципы работы основных библиотек для работы с базами данных и массивами взяты из [9] и [10].

В источниках [4] и [6] рассказывается о том, как работают рекомендательные системы в различных компаниях и приведены примеры работы рекомендательных систем на языке python с помощью машинного обучения.

В источнике [1] возьмем классификацию интеллектуальных систем по решаемым задачам:

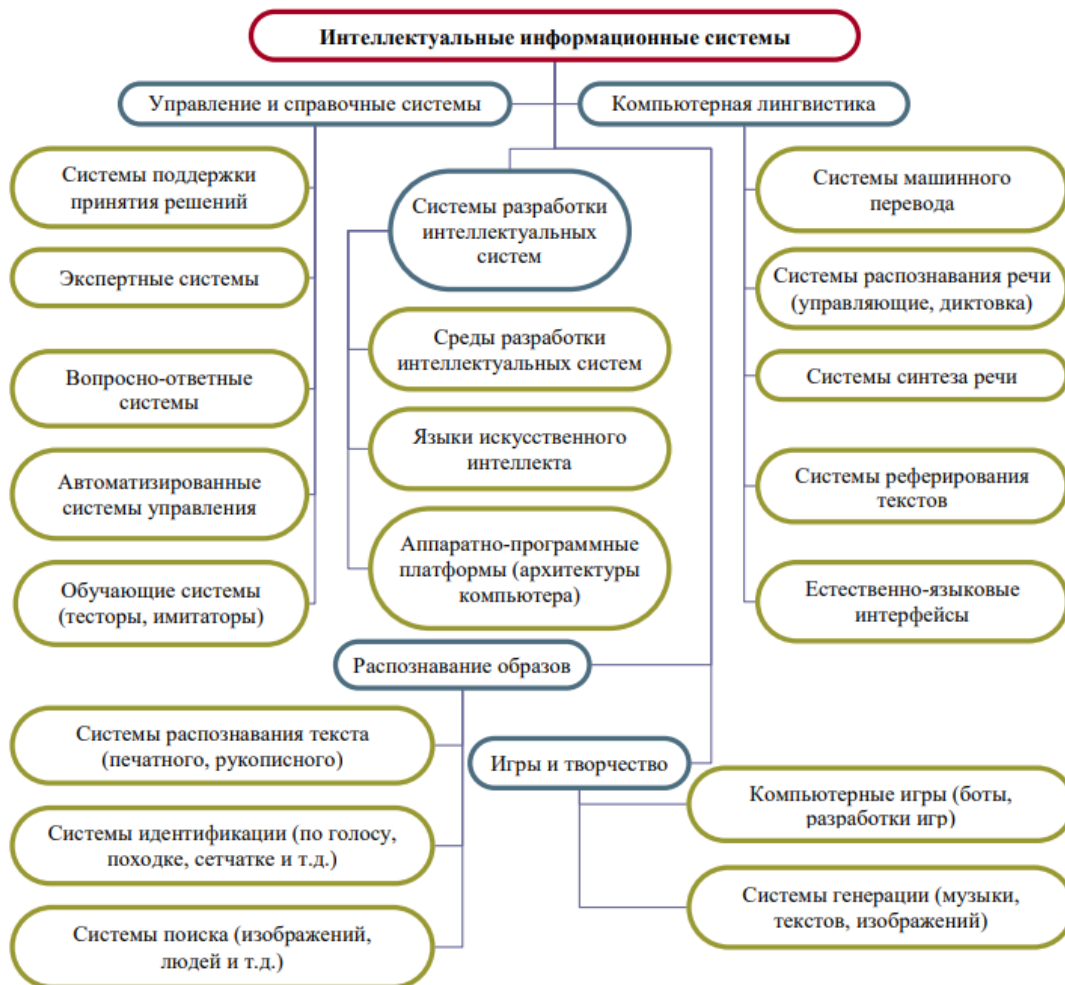


Рисунок 1. Классификация интеллектуальных систем по решаемым задачам [1].

## 2. Теоретическая часть

### 2.1 Машинное обучение

Машинное обучение (ML) — класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме [12].

Составляющие машинного обучения:

1. Данные. Например, если мы хотим спрогнозировать погоду, необходима сводка погоды за последние несколько лет (чем больше, тем лучше). Чем качественнее данные, тем эффективнее будет работать программа. Каким бы совершенным ни был алгоритм работы, если качество данных не очень, результат будет соответствующим [12].

2. Признаки. Набор свойств, характеристик или признаков, которые описывают создаваемую модель. Если мы говорим о погоде, то это температура, скорость ветра, время года. Правильно подобранные признаки — залог успешного обучения [12].

3. Алгоритм. Каждую задачу можно решить разными способами. Для разных целей можно подобрать разные алгоритмы [12].

Методы машинного обучения:

Классическое обучение чаще всего для ИИ используется классическое обучение. Это простые алгоритмы, основанные на закономерностях в данных. Есть 2 типа классического обучения:

Обучение с учителем. Мы обучаем машину на реальных примерах. Допустим, мы хотим научить ее отличать яблоки от груш. Загружаем в программу данные и говорим ей, что на одних картинках изображены яблоки, а на других — груши. Машина должна найти общие признаки и выстроить связи без учителя. Этот метод используется, когда нет возможности предоставить роботу размеченные данные. Программа должна сама найти общие признаки и классифицировать полученные

данные. Такой подход часто используют в таргетированной рекламе, когда действия или предпочтения пользователя нельзя заранее классифицировать [12].

Обучение с подкреплением. Это более сложный вид обучения. ИИ нужно не просто анализировать данные, а действовать самостоятельно в реальной среде. Обучение похоже на игру: за правильно принятое решение машина получает балл, за ошибки — баллы вычитаются.

Рассмотрим на примере игры «Змейка». На поле есть объект, до которого должна добраться змейка. Она не знает, какой путь самый эффективный. Ей известно только расстояние до объекта. Методом проб и ошибок змейка находит оптимальный вариант движения и анализирует ситуации, которые ведут к проигрышу. Этот способ используют для обучения роботов-пылесосов или самоуправляемых автомобилей [12].

Глубокое обучение нейросетей (Deep learning). Глубоким это обучение называется потому, что структура искусственных нейронных сетей состоит из нескольких слоев, которые взаимодействуют между собой и создают сложный процесс анализа данных [12].

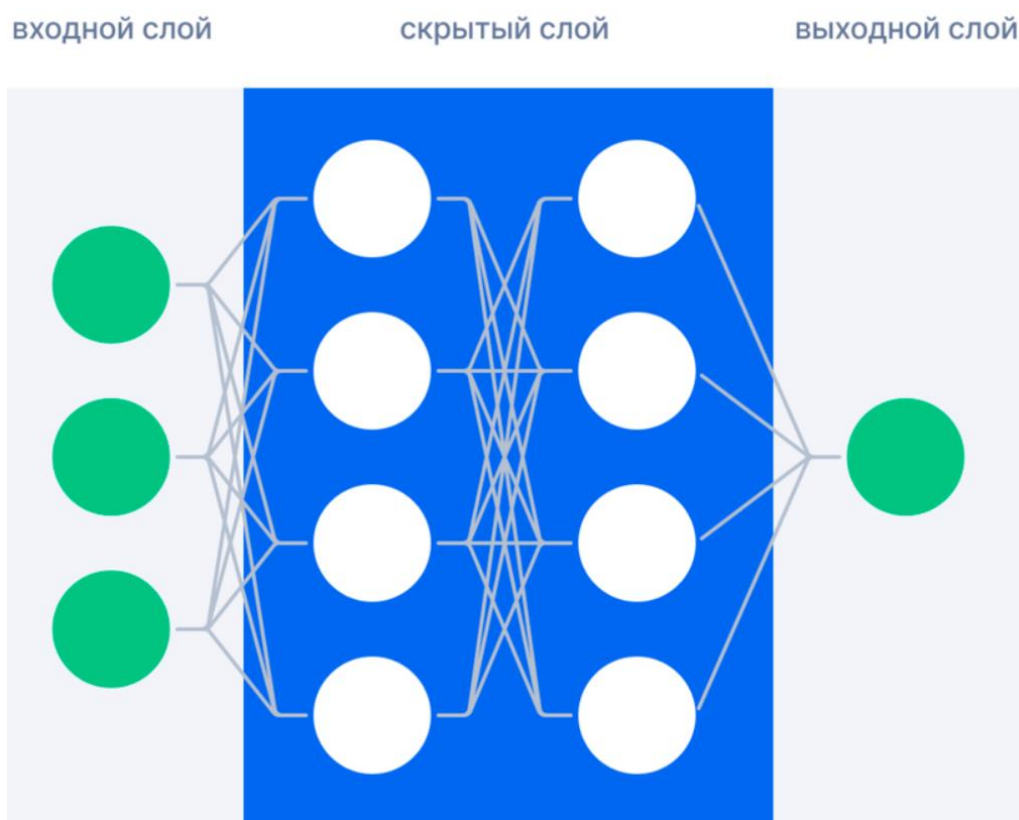


Рис 2. Классическая нейронная сеть прямого распространения.



Входной слой получает различные формы данных, которые сеть обрабатывает. Далее данные проходят через один или несколько скрытых слоев, где они преобразуются в форму, которую может использовать выходной слой. А выходной слой нейронной сети представляет собой место, где сеть отвечает на данные, которые ей были переданы и обработаны.

## **2.2 Анализ данных**

Анализ данных и использование статистических методов для определения предпочтений и поведения пользователей являются важной частью разработки рекомендательных систем. Для этого можно использовать различные методы статистического анализа данных, такие как множественная регрессия, факторный анализ, кластерный анализ и другие [11].

Одним из ключевых задач в анализе данных является выделение признаков, которые могут быть использованы для определения предпочтений и поведения пользователей. Эти признаки могут включать в себя информацию о предыдущих покупках, оценках, просмотрах или других действиях пользователей на сайте. Для анализа этих данных можно использовать различные методы, такие как анализ корреляции, анализ главных компонент, анализ временных рядов и другие. Например, можно использовать методы кластерного анализа для выявления групп пользователей с похожими интересами, что может помочь в персонализации рекомендаций [13].

Для оценки качества рекомендательных систем можно использовать различные метрики, такие как точность, покрытие, разнообразие и удовлетворенность пользователей. Например, точность может быть измерена с помощью метрик, таких как средняя абсолютная ошибка (MAE) или средняя квадратичная ошибка (MSE) [11].

## **2.3 Факторы влияющие на рекомендации**

Качество рекомендаций в рекомендательных системах может быть существенно улучшено за счет учета различных факторов, которые могут влиять на предпочтения и поведение пользователей. Некоторые из этих факторов включают в себя контекстные условия, социальные связи, пользовательские

интересы и предпочтения, а также другие факторы, такие как временной фактор, сезонность и многие другие.

Контекстные условия могут включать в себя местоположение, время суток, устройство и другие параметры, которые могут влиять на предпочтения пользователя. Например, пользователь может искать рестораны в определенном районе в определенное время суток, и рекомендации должны учитывать эти факторы для того, чтобы быть наиболее релевантными [14].

Социальные связи также могут оказывать влияние на предпочтения и поведение пользователей. Например, пользователи могут оставлять отзывы и оценки на продукты, а также обмениваться мнениями и рекомендациями с друзьями в социальных сетях. Эти связи могут быть использованы для создания персонализированных рекомендаций, учитывающих предпочтения пользователей и их социальные связи [15].

Пользовательские интересы и предпочтения являются ключевыми факторами, влияющими на рекомендации. Рекомендательные системы могут использовать информацию о предыдущих действиях пользователя, таких как покупки, просмотры и оценки, для определения его предпочтений. Также могут быть использованы методы анализа текстовых данных, таких как отзывы и комментарии, для определения интересов и предпочтений пользователей [16].

Другие факторы, которые могут влиять на рекомендации, включают в себя временной фактор, сезонность, наличие акций и скидок, а также другие параметры, которые могут быть важными для конкретной категории продуктов или услуг.

Разберем критерии, влияющие на оценку рекомендаций:

Оценка качества рекомендательных систем является важной задачей, чтобы понимать, насколько хорошо система выполняет свою задачу и как можно ее улучшить. Существует несколько методов оценки качества рекомендаций, включая точность, покрытие, разнообразие и удовлетворенность пользователей.

Точность— это мера того, насколько хорошо рекомендательная система прогнозирует, что понравится пользователю. Точность может быть измерена с помощью метрик, таких как средняя абсолютная ошибка (MAE), корень из

среднеквадратичной ошибки (RMSE), коэффициент корреляции Пирсона и многие другие [17].

$$MAE = \frac{1}{N} \sum_{i,j=1}^N |y_i - y_j|, \text{ где} \quad (2.3.1)$$

где  $N$  — число примеров обучающей выборки,

$y_i$  — целевое значение  $i$ -го примера,

$y_j$  — предсказанное моделью значение.

$$MSE = \frac{1}{N} \sum_{i,j=1}^N (y_i - y_j)^2 \text{ где} \quad (2.3.2)$$

где  $N$  — число примеров обучающей выборки,

$y_i$  — целевое значение  $i$ -го примера,

$y_j$  — предсказанное моделью значение.

Покрытие — это мера того, как много элементов из всего множества элементов было рекомендовано пользователю. Чем выше покрытие, тем больше вероятность, что пользователь найдет что-то, что ему понравится. Покрытие может быть измерено как доля уникальных элементов, которые были рекомендованы пользователю, от общего числа элементов [18].

$$K = \frac{N_{unique}}{N}, \text{ где} \quad (2.3.3)$$

$N_{unique}$  — уникальные рекомендованные элементы,

$N$  — число элементов.

Разнообразие — это мера того, насколько разнообразны рекомендации, которые предлагаются пользователю. Разнообразие может быть измерено с помощью метрик, таких как среднее расстояние между элементами, которые были рекомендованы пользователю [19].

Удовлетворенность пользователей — это мера того, насколько пользователи довольны рекомендациями, которые были им предложены. Удовлетворенность может быть измерена с помощью опросов, анкетирований и других методов сбора обратной связи от пользователей [20].

### 3. Практическая часть

Программная обработка данных была проведена за семь основных этапов

1. Выгрузили данные из базы данных в среду программирования Google Colaboratory.

2. Подготовили данные: проверили данные на отсутствие пропущенных значений, дубликатов или некорректных значений.

3. Разделили данные на обучающую и тестовую выборки: разделили данные на две части - обучающую и тестовую выборки. Обучающая выборка использовалась для обучения модели, а тестовая выборка - для оценки качества модели.

4. Выбрали метод машинного обучения: выбрали подходящий метод машинного обучения для рекомендательной системы.

5. Обучили модель.

6. Оценили качество модели: оценили метрики, такие как средняя абсолютная ошибка (MAE) и средняя квадратичная ошибка (MSE) с помощью формул (2.3.1) и (2.3.2).

7. Использовали модель для рекомендаций: с помощью метода ближайших соседей, нашли похожих пользователей и предлагаем им игры, в которые они еще не играли, но которые понравились похожим пользователям с помощью формулы (1.5).

Для начала посмотрим, как выглядят рекомендации в самом магазине steam:

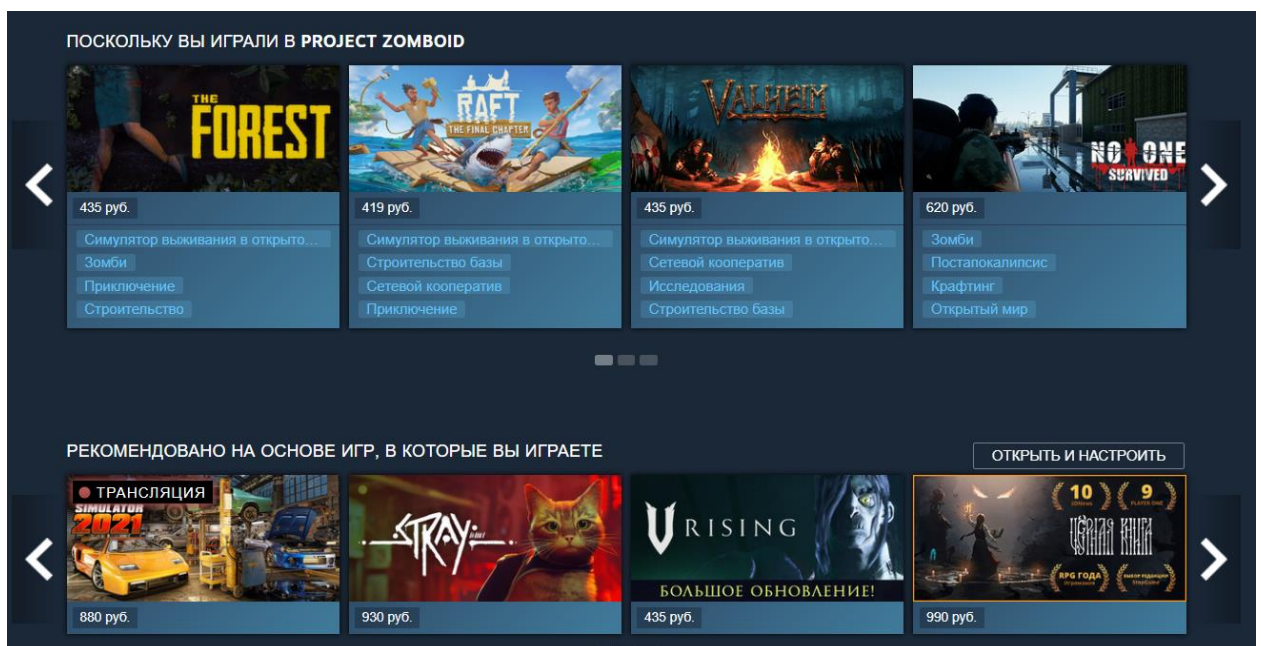


Рис 3. Начальные данные.

В магазине представлен обширный список рекомендаций, основанный на абсолютно разных факторах. Например, на основе того, в какой игре вы проводите больше всего времени, в какие игры играют друзья, с которыми вы играете чаще всего, на основе списка желаемых игр и так далее.

Работа реализована на языке Python.

Выгружаем данные из базы данных в среду программирования Google Colaboratory, импортируем библиотеки для работы с данными и совершаем их предобработку.

	<b>user_id</b>	<b>игра</b>	<b>тип</b>	<b>часов сыграно</b>
<b>0</b>	151603712	The Elder Scrolls V Skyrim	play	273.0
<b>1</b>	151603712	Fallout 4	purchase	1.0
<b>2</b>	151603712	Fallout 4	play	87.0
<b>3</b>	151603712	Spore	purchase	1.0
<b>4</b>	151603712	Spore	play	14.9
...	...	...	...	...
<b>199994</b>	128470551	Titan Souls	play	1.5
<b>199995</b>	128470551	Grand Theft Auto Vice City	purchase	1.0
<b>199996</b>	128470551	Grand Theft Auto Vice City	play	1.5
<b>199997</b>	128470551	RUSH	purchase	1.0
<b>199998</b>	128470551	RUSH	play	1.4

Рис 4. Исходные данные.(данные компании)

В первой колонке содержится идентификатор пользователя, во второй название игры, в третьей способ взаимодействия с игрой, их представлено 2: Покупка и игра. В четвертой колонке представлено количество часов.

Построим график, чтобы посмотреть на самые популярные игры, которые представлены в наших данных, предоставленных компанией:

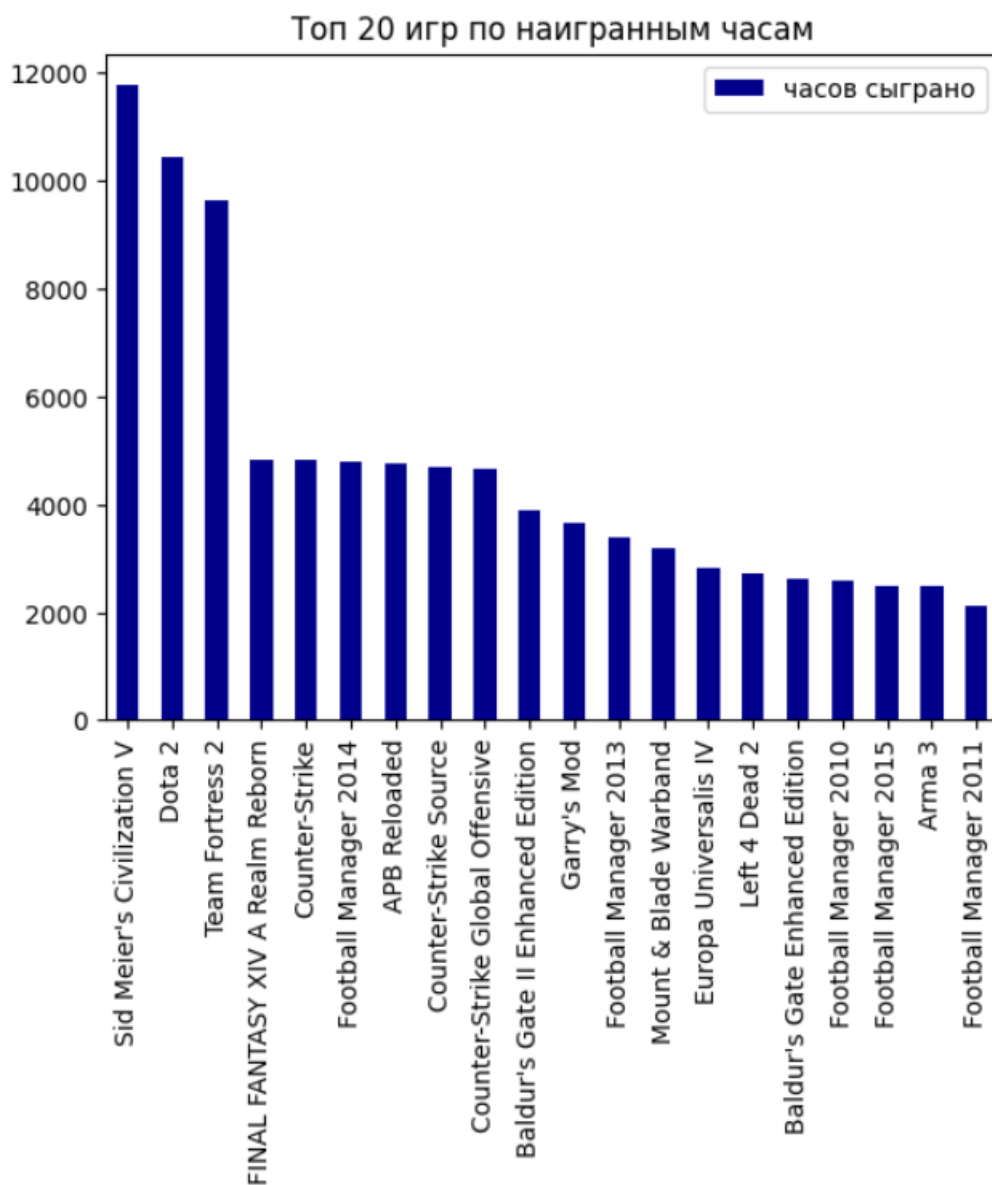


Рис 5. Двадцать популярных игр.

Видно, что большинство игр являются многопользовательскими и либо были популярны когда-либо, либо популярны до сих пор и держат одни из самых высоких показателей онлайн игроков каждый день.

Теперь создадим сводную таблицу, заполним в ней пропущенные значения по формуле минимаксной нормализации  $X' = \frac{X - X_{mean}}{X_{max} - X_{min}}$  [21].

	игра	007 Legends	ORBITALIS	2... 3... KICK IT! (Drop That Beat Like an Ugly Baby)	10 Second Ninja	10,000,000	100% Orange Juice	1000 Amps	12 Labours of Hercules	12 Labours of Hercules II The Cretan Bull	12 Labours of Hercules III Girl Power	rFactor 2	realMyst	realMyst Masterpiece Edition	resident evil 4 / biohazard 4	rymdkapsel	sZone-Online
user_id																	
5250	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
76767	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
86540	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103360	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
144736	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
309554670	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
309626088	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
309812026	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
309824202	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
309903146	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Рис 6. Минимаксная нормализация.

Теперь напишем функцию для рекомендации игр, схожих на заданную игру.

*def gameRec(g):*

*vector = steam\_std[g]*

*vector = steam\_data.corrwith(vector).dropna()*

*gameData = data.groupby('game').agg({'hours\_played': [np.size, np.mean]})*

*gameSim = gameData['hours\_played']['size'] >= 100*

*df = gameData[gameSim].join(pd.DataFrame(vector, columns=['схожесть']))*

*return df.sort\_values(['схожесть'], ascending=False)[:5]*

Функция `gameRec(g)` используется для рекомендации игр на основе их сходства с заданной игрой `g`. В этой функции используется коэффициент корреляции Пирсона для вычисления сходства между заданной игрой и всеми остальными играми в наборе данных. Сначала функция создает вектор признаков `vector` для заданной игры `g` путем центрирования и стандартизации значений ее признаков. Затем функция вычисляет коэффициент корреляции Пирсона между вектором `vector` и каждой строкой матрицы `steam_data`, используя метод `corrwith`. Далее функция создает новый датафрейм `gameData`, который содержит информацию о количестве игроков и среднем времени игры для каждой игры в

наборе данных. Затем функция фильтрует датафрейм `gameData`, оставляя только те игры, которые были сыграны хотя бы 100 часов. Затем функция присоединяет новый столбец `схожесть`, который содержит коэффициент корреляции Пирсона для каждой игры, к датафрейму `gameData`, используя метод `join`. В итоге, функция сортирует датафрейм по убыванию значения столбца `схожесть` и возвращает топ 5 игр, наиболее похожих на заданную игру `g`.

Первой игрой для рекомендаций выберем популярную игру `Left 4 Dead 2`.

игра	(часов сыграно, size)	(часов сыграно, mean)	схожесть
<b>Left 4 Dead 2</b>	1752	19.719007	1.000000
<b>Left 4 Dead</b>	485	18.434845	0.260082
<b>PAYDAY The Heist</b>	479	3.739040	0.178957
<b>Borderlands 2</b>	815	28.339755	0.151266
<b>Killing Floor Mod Defence Alliance 2</b>	241	1.034440	0.144940

Рис 7. Рекомендации к игре `Left 4 Dead 2`.

В результате получил первую часть `Left 4 Dead`, являющуюся приквелом, `Killing Floor Mod Defence Alliance 2`, являющаяся одинаковой по жанру и имеющей схожую аудиторию и две кооперативные игры `PAYDAY The Heist` и `Borderlands 2`

Следующая игра для рекомендаций будет `Fallout 4`

игра	(часов сыграно, size)	(часов сыграно, mean)	схожесть
<b>Fallout 4</b>	335	32.629254	1.00
<b>Alan Wake</b>	218	5.389908	0.35
<b>Darksiders</b>	245	5.506122	0.32
<b>Middle-earth Shadow of Mordor</b>	209	11.657895	0.27
<b>Company of Heroes</b>	241	3.274689	0.27

Рис 8. Рекомендации к игре `Fallout 4`.

Функция выдает схожие по жанру игры, поэтому пользователь может увидеть и заинтересоваться рекомендуемыми играми.



Теперь реализуем с помощью метода ближайших соседей рекомендацию игр для отдельного пользователя

```
from sklearn.neighbors import NearestNeighbors

from sklearn.metrics import mean_absolute_error, mean_squared_error

# Загружаем данные и переименовываем столбцы

df = pd.read_csv('steam-200k.csv')

df.rename(columns = {'151603712': 'user_id', '1.0' : 'hours_played', 'The Elder Scrolls V Skyrim'
: 'game', 'purchase' : 'type'}, inplace = True )

# Создадим матрицу и заполним пропущенные значения

user_game_matrix = pd.pivot_table(df, values='hours_played', index='user_id',
columns='game').fillna(0)

# Инициализируем модель ближайших соседей

k_neighbors = 10

model = NearestNeighbors(metric='cosine', algorithm='brute', n_neighbors=k_neighbors)

# Обучим модель на нашей матрице

model.fit(user_game_matrix)

# Будем выбирать отдельного пользователя по его id для рекомендации

user_id = задаем айди пользователя

# Найдем ближайших соседей для выбранного пользователя

user_index = user_game_matrix.index.get_loc(user_id)

distances, indices = model.kneighbors(user_game_matrix.iloc[user_index, :].values.reshape(1, -
1), n_neighbors=k_neighbors)

# Получаем список игр, которые еще не были играны выбранным пользователем

unplayed_games = user_game_matrix.columns[user_game_matrix.iloc[user_index, :] == 0]

# Создаем список суммарных оценок игр среди ближайших соседей

game_scores = user_game_matrix.iloc[indices.flatten(), :].sum(axis=0)

# Отсортируем игры по суммарным оценкам и выберем топ для рекомендации

top_games = game_scores.sort_values(ascending=False)[:6].index.tolist()
```

```

# Выводим список рекомендованных игр для выбранного пользователя

print("Рекомендуемые игры для пользователя с id", user_id, ":")

for game in top_games:

    if game in unplayed_games:

        print(game)

# Индексы игр, которые были оценены пользователями

played_games_index = np.where(user_game_matrix.loc[user_id] != 0)[0]

# Реальные значения только для оцененных игр

y_true = user_game_matrix.iloc[user_index, played_games_index]

# Предсказанные значения только для оцененных игр

y_pred = user_game_matrix.mean(axis=0).iloc[played_games_index]

# Вычисление средней абсолютной ошибки

mae = mean_absolute_error(y_true, y_pred)

```

Теперь проверим работоспособность алгоритма для пользователей.

```

Рекомендуемые игры для пользователя с id 236735185 :
Hitman Absolution
The Elder Scrolls V Skyrim
Risen 2 - Dark Waters
Ryse Son of Rome
Alan Wake
MAE: 1.81
MSE: 3.27

```

Рис 9. Рекомендации для пользователя 1.

```

Рекомендуемые игры для пользователя с id 30246419 :
Assassin's Creed IV Black Flag
The Witcher Enhanced Edition
Grand Theft Auto V
MAE: 1.68
MSE: 15.89

```

Рис 10. Рекомендации для пользователя 2.

Таким образом, с помощью метода ближайших соседей удалось составить список рекомендуемых игр для пользователей. Для первого типа пользователей

значения MAE и MSE получились равными 1.81 и 3.27, а для второго типа пользователя 1.68 и 15.89.

## **4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение**

### **Введение**

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

Данный раздел, предусматривает рассмотрение следующих задач:

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

Цель данной ВКР – разработка алгоритма для создания экспертной рекомендательной системы для клиентов компании.

### **4.1 Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения**

#### **4.1.1 Анализ конкурентных технических решений**

На данном этапе выбираются конкурентные технические решения, схожие по функционалу с объектом исследования. Анализ конкурентных технических решений с позиции ресурсоэффективности и ресурсосбережения позволяет провести оценку сравнительной эффективности научной разработки и определить направления для ее будущего повышения.

Конкурентами проекта по созданию рекомендательной системы для магазина игр могут быть другие магазины игр, которые уже имеют подобные системы, а также онлайн-платформы для покупки и скачивания игр, такие как Steam, Epic Games Store, GOG и др. Кроме того, могут быть конкуренты среди компаний-разработчиков рекомендательных систем, которые предлагают свои решения для магазинов и других компаний.

В качестве конкурентов были выбраны:

1. «Epic Games Store» - цифровая платформа для покупки и скачивания видеоигр, запущенная компанией Epic Games в 2018 году. Она предлагает широкий выбор игр от разных издателей и разработчиков, а также эксклюзивные игры, которые недоступны на других платформах. Epic Games Store имеет свою рекомендательную систему, которая основывается на анализе покупок и предпочтений пользователей, а также на отзывах и рейтингах игр.
2. «GOG» - цифровая платформа для покупки и скачивания видеоигр, запущенная компанией CD Projekt в 2008 году. GOG специализируется на продаже классических и ретро игр, а также на играх без цифровой защиты (DRM). Она также предлагает рекомендательную систему на основе анализа предпочтений пользователей и их покупок.

Рассмотрим оценочную карту конкурентоспособности для анализа технических и экономических критериев.

Таблица 4.1 – Оценочная карта конкурентоспособности.

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б <sub>ф</sub>	Б <sub>к1</sub>	Б <sub>к2</sub>	К <sub>ф</sub>	К <sub>к1</sub>	К <sub>к2</sub>
<b>Технические критерии оценки ресурсоэффективности</b>							
1. Функциональность исследования	0,1	4	4	4	0,4	0,4	0,4
2. Надежность системы	0,125	5	5	5	0,625	0,625	0,625
3. Эффективность рекомендаций	0,2	4	4	3	0,8	0,8	0,6
4. Безопасность пользователей	0,1	5	5	5	0,5	0,5	0,5
5. Поддержка и обслуживание системы	0,05	2	5	4	0,1	0,25	0,2
6. Масштабируемость системы	0,075	3	4	4	0,225	0,3	0,3
<b>Экономические критерии оценки эффективности</b>							

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Б <sub>ф</sub>	Б <sub>к1</sub>	Б <sub>к2</sub>	К <sub>ф</sub>	К <sub>к1</sub>	К <sub>к2</sub>
1. Стоимость продукта	0,15	5	2	2	0,75	0,3	0,3
2. Срок эксплуатации	0,1	3	4	4	0,3	0,4	0,4
3. Необходимость в финансировании	0,1	4	2	2	0,4	0,2	0,2
<b>Итого</b>	<b>1</b>	<b>35</b>	<b>35</b>	<b>33</b>	<b>4,1</b>	<b>3,8</b>	<b>3,5</b>

В данной таблице использовались следующие обозначения: Б<sub>ф</sub>-баллы, начисляемые за разработанный алгоритм; Б<sub>к1</sub>- баллы, начисляемые Epic Games Store; Б<sub>к2</sub>- баллы, начисляемые GOG;

Конкурентоспособность оценивается по формуле:  $K_i = VB_i$ , где В – вес критерия. Полученные значения оказались  $K_{ф}=4,1$  ед.,  $K_{к1}=3,8$  ед.,  $K_{к2}=3,5$  ед. Проведенный анализ конкурентных технических решений показал, что исследование является актуальным и перспективным, имеет конкурентоспособность. Стоит отметить, что веса критериев подбирались таким образом, чтобы выделить сильные качества собственного алгоритма.

#### 4.1.2 SWOT-анализ

Данный инструмент предназначен для детальной оценки внутренней и внешней среды проекта – выделяются сильные и слабые стороны внутренней среды проекта, а также угрозы и возможности внешней среды проекта.

Таблица 4.2 – Матрица SWOT-анализа

	Сильные стороны	Слабые стороны
	<p><b>С1.</b>Рекомендательная система позволит улучшить пользовательский опыт покупки игр в магазине.</p> <p><b>С2.</b> Машинное обучение и нейронные сети могут обеспечить высокую точность рекомендаций.</p>	<p><b>Сл1.</b>Необходимость большого объема данных для обучения системы</p> <p><b>Сл2.</b>Возможность искажения рекомендаций из-за недостаточного разнообразия данных.</p>

	<b>С3.</b> Система может помочь магазину увеличить продажи и клиентскую базу.	<b>Сл3.</b> Высокие затраты на разработку и поддержку системы
<p><b>Возможности</b></p> <p><b>В1.</b> Возможность увеличения продаж и доходов за счет более точной рекомендательной системы.</p> <p><b>В2.</b> Возможность привлечения новых клиентов и удержания, существующих за счет улучшения пользовательского опыта</p> <p><b>В3.</b> Возможность использования данных, собранных системой, для анализа потребительского поведения и улучшения ассортимента игр.</p>	<p><b>В2С2.</b> Возможность привлечения новых клиентов и удержания, существующих за счет введения новой рекомендательной системы, основанной на машинном обучении, следовательно улучшения пользовательского опыта.</p> <p><b>В3С1.</b> Полученная система способствует увеличению продаж и доходов.</p>	<p><b>В2В3Сл3.</b> Увеличение команды разработчиков позволит улучшить систему, но компании придется увеличить бюджет на разработку и поддержание данной системы.</p> <p><b>В1В2Сл2.</b> При увеличении клиентской базы и удачных рекомендациях, пропадет возможность искажения данных, так как у нас будет большое разнообразие профилей пользователей</p>
<p><b>Угрозы</b></p> <p><b>У1.</b> Конкуренция от других магазинов игр с аналогичными рекомендательными системами</p> <p><b>У2.</b> Потенциальные проблемы с конфиденциальностью данных клиентов</p> <p><b>У3.</b> Возможность технических проблем и сбоев в работе системы</p>	<p><b>У1С1С2С3.</b> Конкуренция от других магазинов игр будет способствовать улучшению модели.</p>	<p><b>У3С1С2С3</b> Возможность технических проблем и сбоев в работе системы может привести к ухудшению пользовательского опыта.</p>

--	--	--

В результате SWOT-анализа можно сделать вывод, что преимущества данного проекта преобладают над недостатками. А некоторые слабые стороны и угрозы алгоритма способствуют нахождению новых векторов развития и приобретения сильных сторон.

## 4.2 Планирование научно-исследовательских работ

### 4.2.1 Структура работ в рамках научного исследования

Планирование комплекса предполагаемых работ осуществляется в следующем порядке:

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

Для разработки алгоритма сформируем рабочую группу, которая будет состоять из разработчика и научного руководителя. На таблице 4.3 представлен порядок выполнения работ разработки.

Таблица 4.3 – Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб.	Содержание работ	Должность исполнителя
Разработка технического задания	1	Составление и утверждение технического задания, утверждение план-графика	Научный руководитель
	2	Календарное планирование выполнения работ	Разработчик, научный руководитель
Выбор способа решения поставленной задачи	3	Исследование научной литературы	Разработчик, научный руководитель
	4	Выбор оптимальной среды реализации проекта	Разработчик
Разработка алгоритма	5	Построение блок-схемы алгоритма	Разработчик, научный руководитель
	6	Реализация программного кода	Разработчик
	7	Тестирование программы	Разработчик



Обобщение и оценка результатов	8	Анализ полученных результатов	Разработчик
	9	Оценка эффективности и модернизация алгоритма	Разработчик, Научный руководитель
Оформление отчета по НИР (комплекта документации по ОКР)	10	Составление пояснительной записки	Разработчик

#### 4.2.2 Определение трудоемкости выполнения работ и разработка графика проведения

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

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

$$t_{\text{ож}i} = \frac{3t_{\text{min}i} + 2t_{\text{max}i}}{5}, \quad (4.1)$$

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

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

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

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

$$T_{pi} = \frac{t_{\text{ож}i}}{Ч_i}, \quad (4.2)$$

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

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

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

Для перевода длительности каждого этапа из рабочих в календарные дни, необходимо воспользоваться формулой (4.3):

$$T_{ki.инж} = T_{pi} \cdot k_{кал}, \quad (4.3)$$

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

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

$k_{кал}$  – календарный коэффициент.

Календарный коэффициент определяется по формуле:

$$k_{кал.инж} = \frac{T_{кал}}{T_{кал} - T_{вых} - T_{пр}} = \frac{365}{365 - 104 - 14} = 1,48 \quad (4.4)$$

где  $T_{кал}$  – общее количество календарных дней в году;  $T_{вых}$  – общее количество выходных дней в году;  $T_{пр}$  – общее количество праздничных дней в году (2022 год).

Расчеты временных показателей проведения научного исследования обобщены в таблице 4.4.

Таблица 4.4 – Временные показатели проведения научного исследования

Название работы	Трудоёмкость работ						Длительность работ в рабочих днях $T_{pi}$	Длительность работ в календарных днях $T_{ki}$
	$t_{min}$ , чел-дни		$t_{max}$ , чел-дни		$t_{ожс}$ , чел-дни			
	Исп.1	Исп.2	Исп.1	Исп.2	Исп.1	Исп.2		
1	2	3	4	5	6	7	8	9
1. Составление и утверждение технического задания, утверждение плана-графика	-	2	-	5	-	3,2	3,2	4
2. Календарное планирование выполнения работ	2	2	4	4	2,8	2,8	2,8	4

3. Исследование научной литературы	7	3	10	5	8,2	3,8	6	8
4. Выбор оптимальной среды реализации проекта	8	-	15	-	10,8	-	10,8	15
5. Построение блок-схемы алгоритма	1	1	3	3	1,8	1,8	1,8	2
6. Реализация программного кода	15	-	35	-	23	-	23	34
7. Тестирование программы	5	-	30	-	15	-	15	22
8. Анализ полученных результатов	3	-	10	-	5,8	-	5,8	8
9. Оценка эффективности и модернизация алгоритма	3	3	6	6	4,2	4,2	4,2	6
10. Составление пояснительной записки	12	-	17	-	14	-	14	20
<b>Итого:</b>	<b>56</b>	<b>11</b>	<b>130</b>	<b>23</b>	<b>85,6</b>	<b>15,8</b>	<b>86,6</b>	<b>123</b>

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

На основе таблицы составлен календарный план-график выполнения проекта с использованием диаграммы Ганта (таблица 4.5).

Таблица 4.5 – Диаграмма Ганта

№	Вид работ	Исп	$T_{ki}$ кал. дн.	Продолжительность работ												
				февр			март			апр			май			
				1	2	3	1	2	3	1	2	3	1	2	3	
1	Составление и утверждение технического задания, утверждение план-графика	Исп2	4	█												
2	Календарное планирование выполнения работ	Исп2 Исп1	4	█												
3	Исследование научной литературы	Исп1	8		█											
4	Выбор оптимальной среды реализации проекта	Исп1	15			█										

№	Вид работ	Исп	$T_{ki}$ кал. дн.	Продолжительность работ											
				февр			март			апр			май		
				1	2	3	1	2	3	1	2	3	1	2	3
5	Построение блок-схемы алгоритма	Исп1 Исп2	2			■									
6	Реализация программного кода	Исп1	34			■	■	■	■						
7	Тестирование программы	Исп1	22						■	■	■	■			
8	Анализ полученных результатов	Исп1	8									■			
9	Оценка эффективности и модернизация алгоритма	Исп2 Исп1	6									■	■		
10	Составление пояснительной записки	Исп1	20											■	■

Примечание:

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

### 4.3 Бюджет научно-технического исследования

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

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

#### 4.3.1 Расчет амортизации специального оборудования

Расчет амортизации проводится следующим образом:

Норма амортизации: рассчитывается по формуле:

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

где  $n$  – срок полезного использования в количестве лет.

Амортизация оборудования рассчитывается по формуле:

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

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

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

Таблица 4.6 - Затраты на программное оборудование

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

Стоит отметить, что учитываться затраты будут только на амортизационные отчисления, так как у рабочей группы уже имеются ноутбуки. Лицензия операционной системы приобретается на бессрочный период, но так как ресурс ноутбука, на который устанавливается ПО, ограничен (в среднем 8 лет), то перерасчёт производится на количество времени, затраченное на проект (2,3 месяца). Стоимость одной лицензии – 30 000 рублей.

Норма амортизационных отчисления на приобретение лицензии:

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

Из этого следует, что размер амортизации составит:

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

Затраты на установку ОС на 2 устройства:

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

Также дистрибутив языка программирования и данные анализа хранятся в открытом доступе.

Рассчитаем амортизацию оборудования линейным способом. Производитель даёт гарантию на модель 5 лет. Соответственно, рассчитаем размер

амортизационных отчислений за специальное оборудование (компьютеры) за время разработки алгоритма:

$$A_{\text{лиц}} = \frac{100000}{5 \cdot 12} \cdot 4,1 = 6833,3 \text{ руб.}$$

Итоговая сумма затрат использование программного оборудования:

$$Z_{\text{оборуд}} = A_{\text{устан}} + A_{\text{лиц}} + A_{\text{лиц}} = 5000 + 2562,5 + 6833,3 = 14395,8 \text{ руб.}$$

#### 4.3.2 Основная заработная плата исполнителей темы

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

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

$$Z_{\text{осн}} = Z_{\text{дн}} \cdot T_p, \quad (4.7)$$

где  $Z_{\text{дн}}$  – среднедневная заработная плата, руб.;  $T_p$  – продолжительность работ, выполняемых работником, раб.дн. (таблица 4.9).

Среднедневная заработная плата рассчитывается по формуле:

Для шестидневной рабочей недели (рабочая неделя научного руководителя):

$$Z_{\text{дн}} = \frac{Z_m \cdot M}{F_o} = \frac{51285 \cdot 10,3}{246} = 2147,3 \text{ руб.}, \quad (4.8)$$

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

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

– при отпуске в 56 раб. дней –  $M = 10,3$  месяца, 6-дневная рабочая неделя.

Для пятидневной рабочей недели (рабочая неделя разработчика):

$$Z_{\text{дн}} = \frac{Z_m \cdot M}{F_o} = \frac{33150 \cdot 11,2}{213} = 1743,1 \text{ руб.} \quad (4.9)$$

Должностной оклад работника за месяц:

– для научного руководителя:

$$Z_{.m} = Z_{mc} \cdot (1 + k_{np} + k_{\partial}) k_p = 26300 \cdot (1 + 0,3 + 0,2) \cdot 1,3 = 51285 \text{ руб.} \quad (4.10)$$

– для разработчика:

$$Z_{.m} = Z_{mc} \cdot (1 + k_{np} + k_{\partial}) k_p = 17000 \cdot (1 + 0,3 + 0,2) \cdot 1,3 = 33150 \text{ руб.}, \quad (4.11)$$

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

В таблице 4.7 представлены показатели рабочего времени, исходя от вида рабочей недели. В таблице 4.8 представлены расчеты основной заработной платы рабочей группы.

Таблица 4.7 – Баланс рабочего времени исполнителей

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

Таблица 4.8 – Расчет основной заработной платы исполнителей

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

Дополнительная заработная плата определяется по формуле:

– для руководителя:

$$Z_{доп} = k_{доп} \cdot Z_{осн} = 0,15 \cdot 33927,3 = 5089,1 \text{ руб.} \quad (4.12)$$

– для инженера:

$$Z_{доп} = k_{доп} \cdot Z_{осн} = 0,15 \cdot 149209,4 = 22381,4 \text{ руб.}, \quad (4.13)$$

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

### 4.3.3 Отчисления во внебюджетные фонды (страховые отчисления)

Отчисления во внебюджетные фонды определяется по формуле:

– для научного руководителя:

$$Z_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0.3 \cdot 39016 = 11704 \text{ руб.} \quad (4.14)$$

– для инженера:

$$Z_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0.3 \cdot 171590 = 51477 \text{ руб.} \quad (4.15)$$

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

### 4.3.4 Накладные расходы

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

$$Z_{\text{накл}} = (Z_{\text{внеб}} + Z_{\text{осн}} + Z_{\text{оборуд}}) \cdot k_{\text{нр}} \quad (4.16)$$

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

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

$$Z_{\text{накл}} = (63181 + 183136,7 + 14767,5) \cdot 0,16 = 252844,2 \cdot 0,16 = 41733,6 \text{ руб.}$$

### 4.3.5 Бюджет НИР

На основании полученных данных по отдельным статьям затрат составляется калькуляция плановой себестоимости ВКР «Создание экспертной рекомендательной системы для клиентов компании» по форме, приведенной в таблице 4.9. В таблице также представлено определение бюджета затрат двух конкурирующих проектов.

Таблица 4.9 – Группировка затрат по статьям

№	Наименование статьи	Сумма, руб.			Примечание
		Текущий Проект	EGS	GOG	



1	Материальные затраты НИР	-	-	-	-
2	Затраты на специальное оборудование	14395,8	100000	100000	Пункт 4.3.1
3	Затраты по основной заработной плате исполнителей темы	183136,7	18420800	18420800	Пункт 4.3.2
4	Затраты по дополнительной заработной плате исполнителей темы	27470,5	2210496	2210496	-
5	Отчисления во внебюджетные фонды	63181	6189388,8	6189388,8	Пункт 4.3.3
6	Накладные расходы	41733,6	4307309	4307309	Пункт 4.3.4
Бюджет затрат НИР		329917,6	31227994	31227994	

#### **4.4 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования**

Для определения эффективности исследования рассчитан интегральный показатель эффективности научного исследования при помощи определения интегральных показателей финансовой эффективности и ресурсоэффективности.

**Интегральный показатель финансовой эффективности** научного исследования получен в процессе оценки бюджета затрат трех конкурирующих проекта. Для этого наибольший интегральный показатель реализации технической задачи принят за базу расчета (как знаменатель), с которым соотносятся финансовые значения по всем вариантам исполнения.

В качестве аналогов данной НИР рассмотрены:

- 1) «Epic Games Store»;
- 2) «GOG».

Интегральный финансовый показатель разработки рассчитывается как:

$$I_{\text{финр}}^{\text{исп.}i} = \frac{\Phi_{\text{pi}}}{\Phi_{\text{max}}}, \quad (4.17)$$

где  $I_{финр}^{исп.i}$  – интегральный финансовый показатель разработки;

$\Phi_{pi}$  – стоимость  $i$ -го варианта исполнения;

$\Phi_{max}$  – максимальная стоимость исполнения.

$\Phi_{текущ.проект} = 329917,6$  руб,  $\Phi_{исп.1} = 31227994$  руб,  $\Phi_{исп.2} = 31227994$  руб.

$$I_{финр}^{тек.пр.} = \frac{\Phi_{тек.пр.}}{\Phi_{max}} = \frac{329917,6}{31227994} = 0.0106$$

$$I_{финр}^{исп.2} = \frac{\Phi_{исп.2}}{\Phi_{max}} = \frac{31227994}{31227994} = 1$$

$$I_{финр}^{исп.3} = \frac{\Phi_{исп.3}}{\Phi_{max}} = \frac{31227994}{31227994} = 1$$

В результате расчета с точки зрения финансовой эффективности текущий проект является более привлекательным и выгодным в несколько раз.

**Интегральный показатель ресурсоэффективности** вариантов выполнения

НИР ( $I_{pi}$ ) определен путем сравнительной оценки их характеристик, распределенных с учетом весового коэффициента каждого параметра (таблица 4.10).

Таблица 4.10 – Сравнительная оценка характеристик вариантов НИР

Объект исследования Критерии	Весовой коэффициент параметра	Текущий проект	Исп.2	Исп.3
1. Надежность системы	0,2	9	9	9
2. Поддержка и обслуживание системы	0,3	8	9	9
3. Эффективность рекомендаций	0,2	8	8	6
4. Безопасность пользователей	0,1	9	9	9
5. Масштабируемость системы	0,2	3	9	9
ИТОГО	1	7,3	8,8	8,4

Расчет интегрального показателя для разрабатываемого проекта:

$$I_{p1} = 0,2 * 9 + 0,3 * 8 + 0,2 * 8 + 0,1 * 9 + 0,2 * 3 = 7,3$$

$$I_{p2} = 0,2 * 9 + 0,3 * 9 + 0,2 * 8 + 0,1 * 9 + 0,2 * 9 = 8,8$$

$$I_{p3} = 0,2 * 9 + 0,3 * 9 + 0,2 * 6 + 0,1 * 9 + 0,2 * 9 = 8,4$$

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

$$I_{исп.i} = \frac{I_{р-исп.i}}{I_{финр}} \quad (20)$$

$$I_{исп.1} = \frac{7,3}{0.0106} = 688,7; I_{исп.2} = \frac{8,8}{1} = 8,8; I_{исп.3} = \frac{8,4}{1} = 8,4$$

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволит определить сравнительную эффективность проекта (таблицы 4.11) и выбрать наиболее целесообразный вариант из предложенных.

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

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

Таблица 4.11 – Сравнительная эффективность разработки

№ п/п	Показатели	Текущий проект	Исп.2	Исп.3
1	Интегральный финансовый показатель разработки	0.0106	1	1
2	Интегральный показатель ресурсоэффективности разработки	7,3	8,8	8,4
3	Интегральный показатель эффективности	688,7	8,8	8,4
4	Сравнительная эффективность вариантов исполнения	1	0,0128	0,0122

Таким образом, данный проект является более привлекательным, в сравнении с другими конкурентами, так как данный проект превосходит другие системы в ресурсоэффективности.

## Выводы по разделу

В результате выполнения целей раздела можно сделать следующие выводы:

1. Результатом анализа конкурентных технических решений является выбор одного из вариантов реализации НИР как наиболее подходящего и оптимального по сравнению с другими.
2. В ходе планирования для руководителя и инженера был разработан график реализации этапа работ, который позволяет оценивать и планировать рабочее время исполнителей. Определено следующее: общее количество календарных дней для выполнения работ составляет 123 дней; общее количество дней, в течение которых работал инженер, составляет 100 дней; общее количество дней, в течение которых работал руководитель, составляет 23 дней;
3. Проведен расчет бюджета исследовательской работы = 330 т.р. Основные расходы были направлены на: затраты на специальное оборудование, затраты по основной заработной плате рабочей группы, отчисления во внебюджетные фонды и накладные расходы;  
Для оценки затрат на реализацию проекта разработан проектный бюджет, который составляет 329917,6 руб;
4. Результат оценки эффективности ИР показывает следующие выводы:
  - 1) значение интегрального финансового показателя ИР составляет 0,0106, что является показателем того, что ИР является финансово выгодной по сравнению с аналогами;
  - 2) значение интегрального показателя ресурсоэффективности ИР составляет 7,3, по сравнению с 8,8 и 8,4;
  - 3) значение интегрального показателя эффективности ИР составляет 688,7, по сравнению с 8.8 и 8.4, и является наиболее высоким, что означает, что техническое решение, рассматриваемое в ИР, является наиболее эффективным вариантом исполнения.

## **5. Социальная ответственность**

### **Введение**

Объектом разработки данной ВКР является алгоритм для разработки экспертной рекомендательной системы. Экспертная система — предиктивная система, включающая в себя знания об определенной слабо структурированной и трудно формализуемой узкой предметной области и способная предлагать и объяснять пользователю разумные решения. Рекомендательные системы — это комплекс сервисов и программ, который анализирует предпочтения пользователей и пытается предсказать, что может их заинтересовать. Данный алгоритм будет полезен для любой компании, где применимы рекомендательные системы.

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

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

### **5.1 Правовые и организационные вопросы обеспечения безопасности**

Специальная оценка условий труда является единым комплексом последовательно осуществляемых мероприятий по идентификации вредных и (или) опасных факторов производственной среды и трудового процесса и оценке уровня их воздействия на работника с учетом отклонения их фактических значений от установленных уполномоченным Правительством Российской Федерации федеральным органом исполнительной власти нормативов (гигиенических нормативов) условий труда и применения средств индивидуальной и коллективной защиты работников [22]. По результатам проведения специальной оценки условий труда устанавливаются классы

(подклассы) условий труда. Согласно трудовому кодексу РФ работникам высшей категории (офисные работники) предусматриваются:

1. Понятие рабочего времени. Нормальная продолжительность рабочего времени [ТК РФ Статья 91]: для работников, условия труда на рабочих местах, которых по результатам специальной оценки условий труда отнесены к оптимальным – нормальная продолжительность рабочего времени не может превышать 40 часов в неделю.
2. Обеспечение нормальных условий труда для выполнения норм выработки [ТК РФ Статья 163];
3. Предоставления работникам предусмотренных ТК РФ выходных и праздничных дней [ТК РФ г. 18], а также оплачиваемые отпуска [ТК РФ г. 19];
4. Установления работникам предусмотренных ТК РФ гарантий и компенсаций [ТК РФ г. 28].

Разработка программного обеспечения происходит за компьютерным столом. Рабочее место разработчика должно быть организовано согласно требованиям, представленным в «Типовая инструкция по охране труда для пользователей персональными электронно-вычислительными машинами (ПЭВМ) в электроэнергетике» [23]:

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

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

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

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

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

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

## 5.2 Производственная безопасность

При разработке алгоритма разработчики подвергаются воздействию различных вредных и опасных факторов, которые представлены в таблице 5.1. В таблице также представлены соответствующие нормативные документы и этапы работ, во время которых разработчики могут столкнуться с их влиянием.

Таблица 5.1 – Возможные опасные и вредные факторы

Факторы (ГОСТ 12.0.003-2015)	Нормативные документы
Отклонение показателей микроклимата	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» [25]
Недостаточная освещённость рабочей зоны	СП 52.13330.2016 «Естественное и искусственное освещение» [25]
Повышенная световая и цветовая контрастность	СП 52.13330.2016 «Естественное и искусственное освещение» [25]

Повышенный уровень шума на рабочем месте	ГОСТ 12.1.003-2014 ССБТ. Шум. Общие требования безопасности [26] СП 51.13330.2011 «Защита от шума. Актуализированная редакция СНиП 23-03-2003» [38]
Длительное сосредоточенное наблюдение	МР 2.2.9.2311 – 07 «Профилактика стрессового состояния работников при различных видах профессиональной деятельности» [29]
Производственные факторы, связанные с электрическим током, вызываемым разницей электрических потенциалов, под действие которого попадает работающий	ГОСТ Р 58698-2019 «Защита от поражения электрическим током» [30]

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

### **5.2.1 Отклонение показателей микроклимата**

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

Общие требования к микроклимату производственных помещений регламентируются СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека



факторов среды обитания». Санитарные нормы регулируют оптимальные и допустимые значения показателей в рабочей зоне, соответствующие физиологическим потребностям организма человека, для создания комфортных и безопасных условий труда.

Работа, выполняемая командой разработки программного обеспечения, по энергозатратам относится к категории Ia (производится сидя, сопровождается незначительными физическими усилиями). В таблицах 5.2 и 5.3 представлены оптимальные и допустимые значения показателей микроклимата на рабочих местах для данной категории.

Таблица 5.2 – Оптимальные величины показателей микроклимата на рабочих местах

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

Таблица 5.3 – Допустимые величины показателей микроклимата на рабочих местах

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

В производственных помещениях, где поддерживать допускаемые нормативные величины локального микроклимата не представляется возможным, необходимо проводить мероприятия по защите работников от

возможного перегрева и охлаждения. Это достигается разными способами: использование систем местного кондиционирования воздуха; регламентацией периодов работы в неблагоприятном локальном микроклимате и отдыха в помещении с микроклиматом, нормализующим тепловое состояние; уменьшение длительности рабочей смены и др.

### 5.2.2 Недостаточная освещенность рабочей зоны

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

Нормы естественного, искусственного и совместного освещения регламентируются СП 52.13330.2016 «Естественное и искусственное освещение». Разработка программного обеспечения относится к категории работ высокой точности – Б (наименьший или эквивалентный объект различения 0,30 – 0,50 мм), подразряд 1 (относительная продолжительность зрительной работы при направлении зрения на рабочую поверхность не менее 70%).

В таблице 5.4 представлены требования к освещению рабочего помещения для разряда Б1.

Таблица 5.4 – Требования к освещению рабочего помещения

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

Яркий свет в зоне периферийного зрения заметно увеличивает глазное напряжение. Для снижения влияния вредного фактора недостаточной

освещенности необходимо, чтобы уровень естественного освещения рабочего пространства приблизительно совпадал с яркостью дисплея. Проблему недостаточной освещенности помещения можно решить при помощи установки дополнительных осветительных приборов, расширения световых проемов.

### **5.2.3 Повышенная световая и цветовая контрастность**

Отклонение светового и цветового контраста на рабочем месте приводит к быстрому утомлению и снижению уровня работоспособности человека на предприятии. Продолжительное воздействие этого вредного фактора приводит к возникновению проблем со зрением. Нормы светового и цветового контраста регламентируются СП 52.13330.2016 «Естественное и искусственное освещение». Для работы за компьютером (категория работ Б1) нормы контраста представлены в таблице 5.5.

Таблица 5.5 – Временные допустимые уровни ЭМП, создаваемых ПЭВМ

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

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

### **5.2.4 Повышенный уровень шума на рабочем месте**

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

Предельно допустимые показатели уровня звука, звукового давления регламентируются СП 51.13330.2011. Защита от шума. Актуализированная редакция СНиП 23-03-2003. Для команды разработчиков программного обеспечения, эти показатели представлены в таблице 5.6.

Таблица 5.6 – Предельно допустимые и допустимые уровни звукового давления, уровни звука, эквивалентные и максимальные скорректированные по А уровни звука в помещениях производственных, жилых, общественных зданий и на территории жилой застройки СП 51.13330.2011. Защита от шума. Актуализированная редакция СНиП 23-03-2003 [38].

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

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

### 5.2.5 Длительное сосредоточенное наблюдение

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

При умственной нагрузке необходима длительность сосредоточенного внимания, выраженная ответственность, плотность сигналов и сообщений в единицу времени по МР 2.2.9.2311 – 07 «Профилактика стрессового состояния работников при различных видах профессиональной деятельности» [28]. Оказывает угнетающее влияние на психическую деятельность ухудшаются функции внимания (объем, концентрация, переключение), памяти (кратковременной и долговременной), восприятия (появляется большое число ошибок).

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

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

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

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

лаборатории могут стать случайное прикосновение к токоведущим частям или появление напряжения на металлических частях оборудования.

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

Таблица 5.7 – Пороги напряжения прикосновения для реагирования

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

Меры предосторожности для основной защиты от поражения электрическим током:

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

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

### **5.3 Экологическая безопасность**

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

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

ПЭВМ состоит из различных деталей, которые имеют различный класс опасности. Ртутные лампы, материнские платы и аккумуляторы содержат свинец, литий, кадмий; материнские платы содержат олово; трансформаторы содержат обмотку из меди, а также железный сердечник; преобразователи и проводка содержат алюминий магний; корпус состоит из металла и пластика. Класс опасности и предельно-допустимая концентрация указаны в таблице 8 [39].

Таблица 5.8 – Предельно-допустимые концентрации (ПДК) вредных веществ в воздухе рабочей зоны

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

Для оказания наименьшего влияния на окружающую среду, необходимо проводить специальную процедуру утилизации ПЭВМ и оргтехники, при которой разбирается, сортируется и более 90% отправится на вторичную переработку и менее 10% будут отправлены на свалки. При этом она должна соответствовать процедуре утилизации ГОСТ Р 53692-2009

«Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов» [32].

## **5.4 Безопасность в чрезвычайных ситуациях**

### **5.4.1 Пожар**

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

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

На основании Федерального закона от 22.07.2008 N 123-ФЗ (ред. От 30.04.2021) "Технический регламент о требованиях пожарной безопасности" помещения учебной аудитории оборудованы следующими средствами пожаротушения: огнетушитель ручной углекислотный ОУ-5, пожарный кран с рукавом, также аудитория оборудована системой противопожарной сигнализации [34]. Помещение аудитории категории помещения группы – В4, возможный класс пожара – Е [34].

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



### **Вывод по разделу**

В данном разделе были рассмотрены правовые и организационные вопросы обеспечения безопасности, аспекты производственной и экологической безопасности. В результате можно сделать вывод о работнике:

- Имеет нормальную продолжительность рабочего времени;
- Обеспечен нормальными условиями труда для выполнения норм выработки;
- Имеет, предусмотренные ТК РФ, выходные и праздничные дни;
- Имеет категорию тяжести труда - Ia (производится сидя, сопровождается незначительными физическими усилиями), согласно [25];
- Относится к первой группе по электробезопасности [36];

Также о рабочей зоне:

- Рабочая зона имеет оптимальные значения показаний микроклимата согласно таблице 5.2;
- Фактические показатели освещения совпадают с оптимальными в таблице 5.4;
- Мониторы ЭВМ имеют оптимальную световую и цветовую контрастность;
- Соблюдаются меры предосторожности для основной защиты от поражения электрическим током;
- Вышедшие из строя ПЭВМ утилизируются согласно ГОСТ Р 53692-2009;
- Помещения учебной аудитории оборудованы положенными средствами пожаротушения;
- Рабочая зона имеет категорию помещения по электробезопасности согласно ПУЭ – без повышенной опасности [35];
- Помещение аудитории категории помещения группы – В4 (пожароопасное);

- Помещение аудитории находится в объекте, относящийся к IV категории по оказанию негативного воздействия на окружающую среду [37].

## Заключение

В результате выполнения выпускной квалификационной работы можно сделать основные выводы:

1. Построена рекомендательная система для клиентов компании, предоставляющей доступ к компьютерным играм. Система основана на кластеризации данных профилей существующих игроков по методу единичной связи по принципу ближнего соседа;
2. все типы пользователей были разделены на два основных класса: “Заинтересованные игроки” и “безразличные к рекомендациям”. Для первого типа пользователя значения MAE и MSE получились равными 1.81 и 3.27 соответственно, а для второго типа пользователя — 1.68 и 15.89, что позволяет их хорошо различать относительно друг друга;
3. в результате применения рекомендательной системы в компании ожидается рост выручки на 16%(до 1,44 млн руб)

## Список используемых источников

1. Интеллектуальные системы / А.В. Остроух. – Красноярск: Научно-инновационный центр, 2015. – 110 с.
2. Рекомендательные системы: что это и как работает алгоритм рекомендаций [Электронный ресурс] // mindbox.ru 2023. URL: <https://mindbox.ru/academy/education/rekomendatelnye-sistemy/> (дата обращения: 10.01.2023).
3. Прохоренок Н.А. Python 3. Самое необходимое / Н.А. Прохоренок, В.А. Дронов. — СПб.: БХВ-Петербург, 2016. — 464 с.
4. Рекомендательные системы и машинное обучение [Электронный ресурс] // ittransition URL: <https://www.ittransition.com/machine-learning/recommendation-systems> (дата обращения 12.01.2023)
5. Классификация рекомендательных систем [Электронный ресурс] // URL: [http://itmu.vsuet.ru/Posobija/Predstavlenie\\_znan/htm/2\\_t.html](http://itmu.vsuet.ru/Posobija/Predstavlenie_znan/htm/2_t.html) (дата обращения 10.01.2023)
6. Recommendation systems explained [Электронный ресурс] // towardsdatascience.com URL: <https://towardsdatascience.com/recommendation-systems-explained-a42fc60591ed> (дата обращения 13.01.2023)
7. Как работают базы данных [Электронный ресурс] // practicum.yandex URL: [practicum.yandex.ru/blog/chto-takoe-bazy-dannyh/](http://practicum.yandex.ru/blog/chto-takoe-bazy-dannyh/) (дата обращения 13.01.2023)
8. Нормализация отношений в базе данных [Электронный ресурс] // habr URL: <https://habr.com/ru/post/254773/> (дата обращения 13.01.2023)
9. Библиотека Numpy [Электронный ресурс] pythonworld.ru URL: <https://pythonworld.ru/numpy/1.html> (Дата обращения: 15.12.2020).
10. Библиотека Pandas [Электронный ресурс] khashtamov.com URL: <https://khashtamov.com/ru/pandas-introduction/> (Дата обращения: 17.12.2020).
11. Francesco Ricci, Lior Rokach Recommender systems handbook. / Lior Rokach, Bracha Shapira, Paul B. Kantor — Canada: 2011. — 512 с.
12. Машинное обучение [Электронный ресурс] // URL: [https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5\\_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5\\_\(Machine\\_Learning\)](https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_(Machine_Learning)) (дата обращения 5.05.2023)
13. Building Recommender Systems with Machine Learning [Электронный ресурс] // github URL: <https://github.com/PacktPublishing/Building-Recommender-Systems-with-Machine-Learning-and-AI> (дата обращения 14.04.2023)
14. Adomavicius, G., & Tuzhilin, A. (2015). Context-aware recommender systems. Recommender Systems Handbook.
15. Burke, R. (2012). Social recommenders. Recommender systems handbook.

16. Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. Recommender systems handbook.
17. Herlocker, J. L., Konstan, J. A., & Riedl, J. (2004). Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS).
18. Zhou, T., Kuscsik, Z., Liu, J. G., Medo, M., Wakeling, J. R., & Zhang, Y. C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. Proceedings of the National Academy of Sciences, 107.
19. Adomavicius, G., & Kwon, Y. (2012). New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems, 27(2).
20. Pu, P., Chen, L., & Hu, R. (2012). User satisfaction with recommender systems: A review and future research directions. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6).
21. Нормализация данных [Электронный ресурс] // URL: <https://wiki.loginom.ru/articles/data-normalization.html> (дата обращения 25.04.2023)
22. ГОСТ 12.2.032-78 Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя. Общие эргономические требования // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200003913> (дата обращения: 18.04.2023);
23. РД 153-34.0-03.298-2001 Типовая инструкция по охране труда для пользователей персональными электронно-вычислительными машинами (ПЭВМ) в электроэнергетике // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200031404> (дата обращения: 18.04.2023);
24. СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2021. – Режим доступа: <https://docs.cntd.ru/document/573500115> (дата обращения: 18.04.2023);
25. СП 52.13330.2016 «Естественное и искусственное освещение» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/456054197> (дата обращения: 20.04.2023);
26. СН 2.2.4/2.1.8.562-96 «Шум. Общие требования безопасности» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200118606> (дата обращения: 20.04.2023);
27. МР 2.2.9.2311-07 «Состояние здоровья работающих в связи с состоянием производственной среды» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200072234> (дата обращения: 20.04.2023);

28. ГОСТ 12.1.005-88 «Общие санитарно-гигиенические требования к воздуху рабочей зоны» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200003608> (дата обращения: 20.04.2023);

29. ГОСТ Р 58698-2019 «Защита от поражения электрическим током» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200170001> (дата обращения: 20.04.2023).

30. Трудовой кодекс (ТК РФ) «Рабочее время» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <http://base.garant.ru/12125268/> (дата обращения: 20.04.2023);

31. ГОСТ Р 53692-2009 «Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200081740> (дата обращения: 21.04.2023);

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

33. Федеральный закон от 22 июля 2008 г. №123-ФЗ «Технический регламент о требованиях пожарной безопасности» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/902111644> (дата обращения: 21.04.2023);

34. «Правила устройства электроустановок» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200030216> (дата обращения: 21.04.2023);

35. Приказ «Об утверждении Правил по охране труда при эксплуатации электроустановок» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573264184> (дата обращения: 21.04.2023);

36. Постановление «Об утверждении критериев отнесения объектов, оказывающих негативное воздействие на окружающую среду, к объектам I, II, III и IV категорий» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573292854> (дата обращения: 21.04.2023);

37. СП 51.13330.2011 «Защита от шума. Актуализированная редакция СНиП 23-03-2003» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200084097> (дата обращения: 21.04.2023);

38. ГН 2.2.5.3532-18 "Предельно допустимые концентрации (ПДК) вредных веществ в воздухе рабочей зоны" // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/557235236> (дата обращения: 21.04.2023);

## Приложение 1

```
import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import math

from scipy import stats

import statistics as st

data = pd.read_csv('/content/sample_data/steam-200k.csv')

data

data.rename(columns = {'151603712': 'user_id', '1.0' : 'часов сыграно', 'The Elder Scrolls V Skyrim' :
'игра', 'purchase' : 'тип'}, inplace = True )

data.drop(data.columns[[-1]], axis = 1)

chislo = data['user_id'].unique()

print(len(chislo))

data.describe()

data.info()

top_games = data.groupby('игра')['часов сыграно'].max().reset_index()

top_games = top_games.sort_values(by='часов сыграно', ascending = False)

top_games_graph = top_games.head(20).plot.bar(x = 'игра', y = 'часов сыграно',
title = "Топ 20 игр по наигранным часам",
color = "darkblue")

plt.show()

steam_data = data.pivot_table(columns='игра', index='user_id', values='часов сыграно', fill_value=0)

def center(row):
    our_row = (row - row.mean()) / (row.max() - row.min())
    return our_row

steam_std = steam_data.apply(center)

steam_data

def gameRec(g):
    vector = steam_std[g]
```

```

vector = steam_data.corrwith(vector).dropna()

gameData = data.groupby('игра').agg({'часов сыграно': [np.size, np.mean]})
gameSim = gameData['часов сыграно']['size'] >= 100
df = gameData[gameSim].join(pd.DataFrame(vector, columns=['схожесть']))
df['схожесть'] = df['схожесть'].round(2)
return df.sort_values(['схожесть'], ascending=False)[:5]

gameRec('Left 4 Dead 2')
gameRec("Fallout 4")

from sklearn.neighbors import NearestNeighbors
from sklearn.metrics import mean_absolute_error, mean_squared_error

# Создадим матрицу и заполним пропущенные значения
user_game_matrix = pd.pivot_table(df, values='часов сыграно', index='user_id',
columns='игра').fillna(0)

# Инициализируем модель ближайших соседей
k_neighbors = 10
model = NearestNeighbors(metric='cosine', algorithm='brute', n_neighbors=k_neighbors)

# Обучим модель на нашей матрице
model.fit(user_game_matrix)

# Будем выбирать отдельного пользователя по его id для рекомендации
user_id = 30246419

# Найдем ближайших соседей для выбранного пользователя
user_index = user_game_matrix.index.get_loc(user_id)
distances, indices = model.kneighbors(user_game_matrix.iloc[user_index, :].values.reshape(1, -1),
n_neighbors=k_neighbors)

# Получаем список игр, которые еще не были играны выбранным пользователем
unplayed_games = user_game_matrix.columns[user_game_matrix.iloc[user_index, :] == 0]

# Создаем список суммарных оценок игр среди ближайших соседей

```



```

game_scores = user_game_matrix.iloc[indices.flatten(), :].sum(axis=0)

# Отсортируем игры по суммарным оценкам и выберем топ для рекомендации
top_games = game_scores.sort_values(ascending=False)[:6].index.tolist()

# Выводим список рекомендованных игр для выбранного пользователя
print("Рекомендуемые игры для пользователя с id", user_id, ":")
for game in top_games:
    if game in unplayed_games:
        print(game)

from sklearn.metrics import mean_absolute_error, mean_squared_error

# Индексы игр, которые были оценены пользователями
played_games_index = np.where(user_game_matrix.loc[user_id] != 0)[0]

# Реальные значения только для оцененных игр
y_true = user_game_matrix.iloc[user_index, played_games_index]

# Предсказанные значения только для оцененных игр
y_pred = user_game_matrix.mean(axis=0).iloc[played_games_index]

# Вычисление средней абсолютной ошибки
mae = mean_absolute_error(y_true, y_pred)

# Вычисление средней квадратичной ошибки
mse = mean_squared_error(y_true, y_pred)
print('MAE:', mae.round(2))
print('MSE:', mse.round(2))

```