**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**

**федеральное государственное автономное образовательное учреждение высшего образования**

**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**

**РУТ(МИИТ)**

**Институт управления и цифровых технологий**

Кафедра «Вычислительные системы, сети и информационная безопасность»

**Отчет По ПРОЕКТНОй ДЕЯТЕЛЬНОСТИ**

**НА тему:**

Направление: 09.03.01 Информатика и вычислительная техника

Профиль:Вычислительные системы и сети

Выполнили:  
студенты группы УВВ-111

Жданов А. И.

Подрез А. П.

Рассадин К. Ю.

Проверил:

Ассистент Желенкова М. Б.

**МОСКВА 2025**

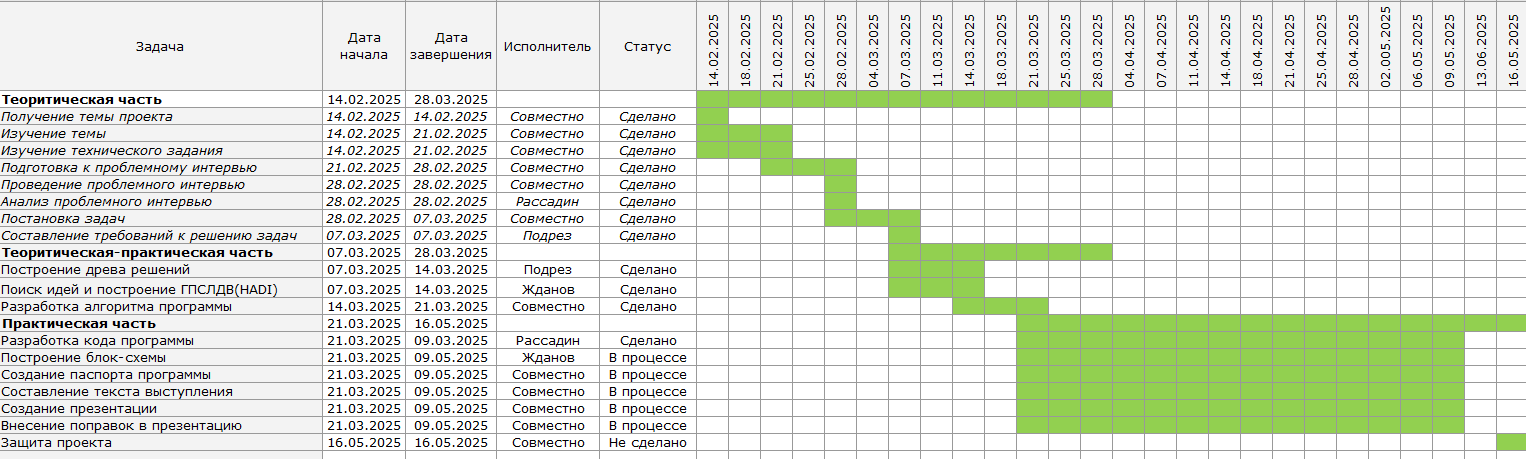
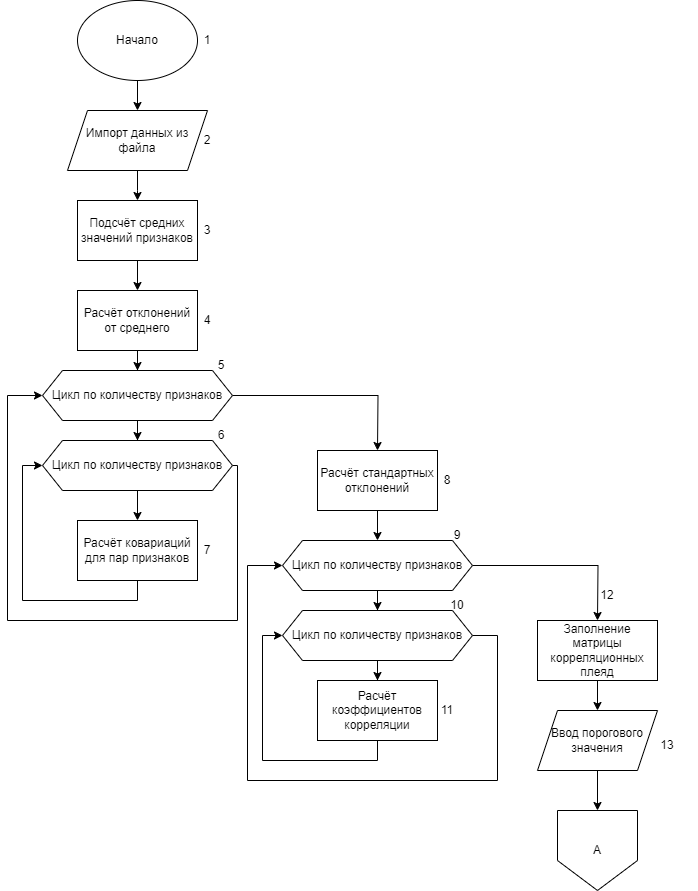


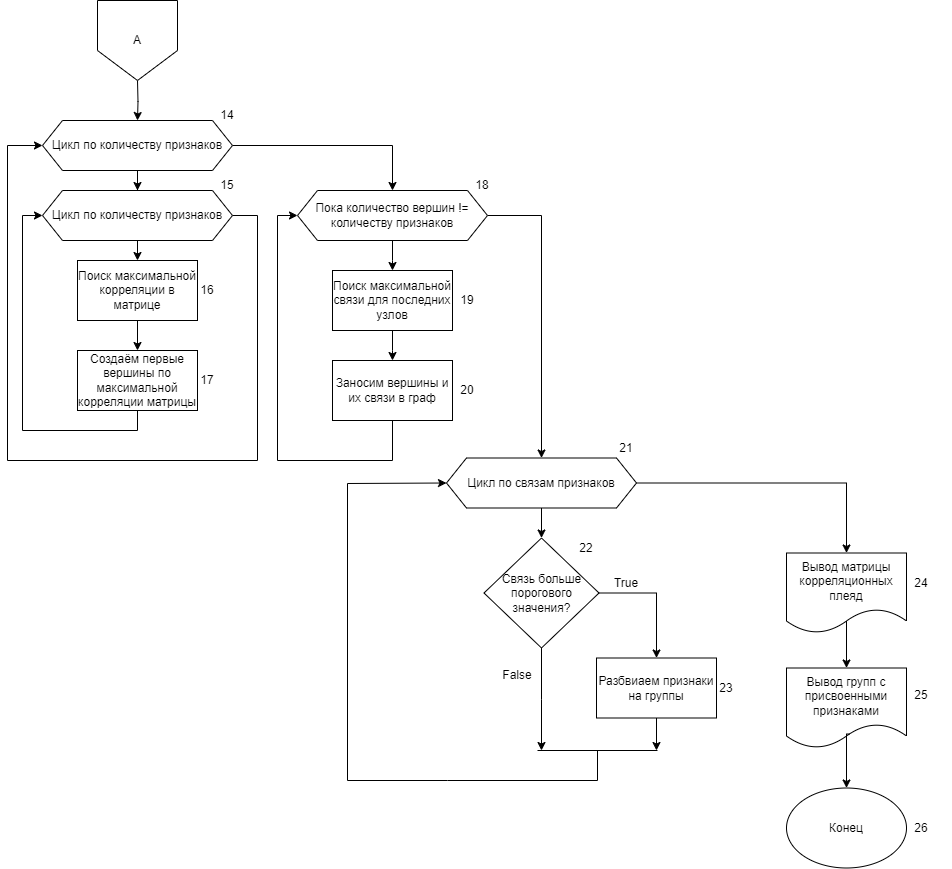
Таблица 1

**Разделы паспорта проекта на стадиях проектной заявки, рабочего паспорта проекта и итогового паспорта проекта**

| **В проектной заявке** | **В рабочем паспорте** | **В итоговом паспорте** |
| --- | --- | --- |
| Определение опасных участков железных дорог | Определение опасных участков железных дорог | Определение опасных участков железных дорог |
| Заказной; учебно-прикладной; предпринимательский; | Заказной; учебно-прикладной; предпринимательский; | Заказной; учебно-прикладной; предпринимательский;; |
|  | Желенкова М.Б. | Желенкова М.Б. |
| С февраля по май 2025 года | -- | -- |
|  | С февраля по май 2025 года | -- |
|  |  | 16.02.2025 – 15.05.2025 |
| Подрез Алексей Павлович; РУТ МИИТ; студент; +79156589334; PodrezAP@yandex.ru | Подрез Алексей Павлович; РУТ МИИТ; студент; +79156589334; PodrezAP@yandex.ru | -- |
| ОАО «РЖД». Сфера – обработка данных. | -- | -- |
| Наш заказчик хочет создать новый алгоритм обработки данных, но не может, потому что ему мешает неэффективный алгоритм обработки данных, а настоящий способ имеет недостатки: человеческий фактор, медленная скорость обработки данных. | -- | -- |
|  | Наш заказчик хочет создать новый алгоритм обработки данных, но не может, потому что ему мешает неэффективный алгоритм обработки данных, а настоящий способ имеет недостатки: человеческий фактор, медленная скорость обработки данных. | Наш заказчик хочет создать новый алгоритм обработки данных, но не может, потому что ему мешает неэффективный алгоритм обработки данных, а настоящий способ имеет недостатки: человеческий фактор, медленная скорость обработки данных. |
| Программное обеспечение | -- | -- |
|  | Решение проблемы является программа, которая практически полностью исключает использование человеческого ресурса, что делает её эффективной. Результатом являются группы отрезков железнодорожных дорог, поделённые по степени опасности | -- |
|  |  | Решение проблемы является программа, которая практически полностью исключает использование человеческого ресурса, что делает её эффективной. Результатом являются группы отрезков железнодорожных дорог, поделённые по степени опасности |
| Работники компании РЖД, отдел: Управление анализа и статистики Департамента информатизации (ЦЧУ) | -- | -- |
|  | Работники компании РЖД, отдел: Управление анализа и статистики Департамента информатизации (ЦЧУ) | Работники компании РЖД, отдел: Управление анализа и статистики Департамента информатизации (ЦЧУ) |
| Начальник управления ЦБА. Компания: ОАО “РЖД” | Начальник управления ЦБА. Компания: ОАО “РЖД” | Начальник управления ЦБА. Компания: ОАО “РЖД” |
| Среда разработки кода (PyCharm) | -- | -- |
|  | https://github.com/sntuew/project | https://github.com/sntuew/project |
|  |  |  |
|  | https://github.com/sntuew/project | https://github.com/sntuew/project |
| Разработчики | -- | -- |
|  | Разработчики | Разработчики |
|  | Жданов А. И.  Подрез А. П.  Рассадин К. Ю. | Жданов А. И.  Подрез А. П.  Рассадин К. Ю. |
|  | Время, человеческие, вычислительные | Время, человеческие, вычислительные |
| Презентации, исходные данные (таблица с объектами и их признаками) | -- | -- |
|  | Презентации, исходные данные (таблица с объектами и их признаками) | Презентации, исходные данные (таблица с объектами и их признаками) |
| Аналоги отсутствуют | -- | -- |
|  | Аналоги отсутствуют | Аналоги отсутствуют |
| Желенкова М.Б. | -- | -- |
|  | Желенкова М. Б. | -- |
|  |  | -- |
|  |  | 1. Подготовка данных   Необходимо внести данные в файл формата «csv», разделителем между числами является: «;»   1. Переместить файл с данными в папку с проектом (кодом) 2. Ввести пороговое значение 3. Запустить программу 4. Получить результат программы |
| Дополнительные видеоролики, статьи про различные библиотеки для работы с данным на языке программирования Python | -- | -- |
|  | Гайд по использованию приложения PyCharm | -- |



Блок-схема 1



Блок-схема 2

Блок-схема 1

import numpy as np  
  
file\_path = "test.csv"  
data = np.genfromtxt(file\_path,delimiter=";").astype(int)  
avg = np.mean(data, axis = 0)  
dev = data - avg  
n = data.shape[0]  
cov\_matrix = np.zeros((data.shape[1], data.shape[1]))  
for i in range(data.shape[1]):  
 for j in range(data.shape[1]):  
 cov = np.sum(dev[:, i] \* dev[:, j]) / (n - 1)  
 cov\_matrix[i, j] = cov  
std\_dev = np.sqrt(np.diag(cov\_matrix))  
corr\_matrix = np.zeros((data.shape[1], data.shape[1]))  
for i in range(data.shape[1]):  
 for j in range(data.shape[1]):  
 if std\_dev[i] != 0 and std\_dev[j] != 0:  
 corr = cov\_matrix[i, j] / (std\_dev[i] \* std\_dev[j])  
 else:  
 corr = 0  
 corr\_matrix[i, j] = corr  
corr\_matrix = np.round(corr\_matrix,3)  
mod\_corr\_matrix = corr\_matrix.copy()  
np.fill\_diagonal(mod\_corr\_matrix, -1)  
mod\_corr\_matrix[mod\_corr\_matrix == 1] = 0.99  
np.fill\_diagonal(mod\_corr\_matrix, 1)  
corr\_matrix = mod\_corr\_matrix  
R = float(input("Введите пороговое значение = "))  
N = corr\_matrix.shape[0]  
graph = {}  
remaining\_indices = list(range(N))  
max\_val = -1  
max\_i, max\_j = -1, -1  
for i in range(N):  
 for j in range(N):  
 if i != j and corr\_matrix[i, j] > max\_val:  
 max\_val = corr\_matrix[i, j]  
 max\_i, max\_j = i, j  
graph[max\_i] = [(max\_j, max\_val)]  
graph[max\_j] = [(max\_i, max\_val)]  
remaining\_indices.remove(max\_i)  
remaining\_indices.remove(max\_j)  
last\_nodes = [max\_i, max\_j]  
while len(graph) < N:  
 max\_val = -1  
 max\_k = -1  
 parent\_node = -1  
 for node in last\_nodes:  
 for k in remaining\_indices:  
 if corr\_matrix[node, k] > max\_val:  
 max\_val = corr\_matrix[node, k]  
 max\_k = k  
 parent\_node = node  
 if max\_k == -1:  
 break  
 if parent\_node in graph:  
 graph[parent\_node].append((max\_k, max\_val))  
 else:  
 graph[parent\_node] = [(max\_k, max\_val)]  
 if max\_k in graph:  
 graph[max\_k].append((parent\_node, max\_val))  
 else:  
 graph[max\_k] = [(parent\_node, max\_val)]  
 remaining\_indices.remove(max\_k)  
 last\_nodes = [parent\_node, max\_k]  
visited = set()  
groups = []  
for node in graph:  
 if node not in visited:  
 group = []  
 stack = [node]  
 while stack:  
 current\_node = stack.pop()  
 if current\_node not in visited:  
 visited.add(current\_node)  
 group.append(current\_node)  
 for neighbor, weight in graph.get(current\_node, []):  
 if weight >= R and neighbor not in visited:  
 stack.append(neighbor)  
 if group:  
 groups.append(group)  
all\_nodes = set(graph.keys())  
for node in all\_nodes - visited:  
 groups.append([node])  
print("Корреляционная матрица:")  
print(corr\_matrix)  
print("\nГруппы признаков с порогом R =", R)  
for i, group in enumerate(groups, 1):  
 print(f"Группа {i}: {sorted(group)}")