МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. Н.Э. Баумана

Кафедра «Системы обработки информации и управления»

|  |  |  |  |
| --- | --- | --- | --- |
|  | Утверждаю  Заведующий кафедрой ИУ-5 |  | Согласовано  научный руководитель |
|  | \_\_\_\_\_\_\_\_\_ В.И.Терехов  "\_\_"\_\_\_\_\_\_\_\_\_\_2024 г. |  | Ю.А. Григорьев\_\_\_\_"\_\_"\_\_\_\_\_\_\_\_\_\_2024 г. |

**Система анализа алгоритмов машинного обучения для решения задач классификации с использованием Pandas**

Техническое задание

(вид документа)

писчая бумага

(вид носителя)

6

(количество листов)

|  |  |
| --- | --- |
| ИСПОЛНИТЕЛЬ: |  |
| Соловьева Александра Михайловна |  |
| "\_\_"\_\_\_\_\_\_\_\_\_ 2024 г. |  |

Москва - 2024

**1. Наименование**

Анализ алгоритмов машинного обучения для решения задач классификации с использованием Pandas на примере предметной области “Статистические данные о занятости и безработице среди населения по возрастным группам”.

**2. Основания для разработки**

Основанием для разработки является задание на выпускную квалификационную работу, подписанное руководителем выпускной работы и утверждённое заведующим кафедрой ИУ5 МГТУ им. Н.Э. Баумана 12 декабря 2023 года.

**3. Назначение разработки**

Разрабатываемая система предназначена для пользователей, желающих получить данные об анализе алгоритмов машинного обучения для решения задач классификации с использованием Pandas на примере предметной области “Статистические данные о занятости и безработице среди населения по возрастным группам”.

Целью является предоставление пользователю результатов анализа алгоритмов машинного обучения для решения задач классификации с использованием Pandas.

**4. Требования к программе или программному изделию**

4.1 Требования к функциональным характеристикам:

4.1.1 Требования к составу выполняемых функций

Система должна выполнять следующие функции:

- загрузку данных в DataFrame Pandas;

- предварительную обработку данных, обработка пропущенных значений, кодирование категориальных признаков и масштабирование числовых признаков;

- разделение данных на обучающий набор и тестовой набор с помощью scikit-learn;

- выбор алгоритмов машинного обучения для классификации, таких как логическая регрессия, метод опорных векторов, случайный лес;

- обучение моделей;

- оценивать точность каждой модели на тестовом наборе;

- выбирать модель с наилучшей производительностью на основе оценки точности;

- вывод результатов анализа, включая графики и метрики для каждой модели.

4.1.2 Требования к входным данным

Входными данными для анализа алгоритмов машинного обучения для решения задач классификации с использованием Pandas должен быть датасет, записанный в файл формата .csv, кодировка UTF-8.

4.1.3 Требования к выходным данным

Выходными данными должны быть аналитические графики и метрики, показывающие результаты анализа алгоритмов машинного обучения для решения задач классификации с использованием Pandas на примере предметной области “Статистические данные о занятости и безработице среди населения по возрастным группам”.

4.2 Требования к надёжности

Система должна надежно и устойчиво функционировать, при вводе некорректных данных выдавать сообщение об ошибке.

4.3 Условия эксплуатации

Пользователь должен быть знаком с основами работы в среде Jupyter Notebook.

4.4 Требования к составу и параметрам технических средств

Компьютер с процессором AMD ryzen 5 c частотой 3.4 ГГц и выше, и ОП не менее 4 ГБ.

4.5 Требования к информационной и программной совместимости

Операционная система Windows 11, приложение Anaconda3, приложение Jupyter Notebook, предустановленные библиотеки в Python: pandas, numpy, sklearn, streamlit.

**5. Требования к программной документации**

Для представления заказчику разрабатываются следующие документы:

1. Техническое задание.

2. Программа и методика испытаний.

3. Расчётно-пояснительная записка.

4. Графический материал по работе в формате презентации.

**6. Технико-экономические показатели**

Требования к данному разделу не предъявляются.

**7. Этапы разработки**

Таблица 1 – Этапы разработки

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Наименование этапов выпускной квалификационной работы** | **Сроки выполнения этапов** | | **Отметка о выполнении** | |
| **план** | **факт** | **Должность** | **подпись, ФИО** |
| 1. 1. | Подготовка документов “Задание на выполнение ВКРБ”, “Календарный план выполнения ВКРБ” | *20.11.2023*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Представление документов “Задание на выполнение ВКРБ”, “Календарный план выполнения ВКРБ” (Смотр № 1) | *03.12.2023*  *Планируемая*  *дата* |  | Нормоконтролер | Кротов Ю.Н. |
|  | Формулирование проблемы, цели и задач работы | *15.01.2024*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Подготовка рабочих материалов | *15.01.2024*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Разработка первой части РПЗ “Постановка задач разработки” | *20.02.2024*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Разработка технического задания | *10.03.2024*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Представление рабочих материалов, технического задания (Смотр № 2) | *20.03.2024*  *Планируемая дата* |  | Нормоконтролер | Кротов Ю.Н. |
|  | Разработка программы и методики испытания | *21.03.2024*  *Планируемая дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Разработка второй части РПЗ «Исследовательская часть» | *10.04.2024*  *Планируемая*  *дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Разработка третьей части РПЗ “Конструкторско-технологическая часть” | *10.04.2024*  *Планируемая дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Защита макета программы  (АСОиУ) | *12.05.2024*  *Планируемая дата* |  | Нормоконтролер | Кротов Ю.Н. |
|  | Разработка заключения, приложений, оформление работы | *23.05.2024*  *Планируемая дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Подготовка доклада и презентации | *25.05.2024*  *Планируемая дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Получение заключения научного руководителя | *28.05.2024*  *Планируемая дата* |  | Руководитель ВКР | Григорьев Ю.А. |
|  | Допуск работы к защите на ГЭК (нормоконтроль) | *30.05.2024*  *Планируемая дата* |  | Нормоконтролер | Кротов Ю.Н. |
|  | Защита работы на ГЭК | *01.06.2024*  *Планируемая дата* |  | Комиссия ГЭК |  |

**8. Порядок контроля и приёмки**

Приём и контроль программного изделия осуществляется в соответствие с документом «Программа и методика испытаний».