

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:

4.1.1.1 импорт библиотек

4.1.1.2 загрузка данных в DataFrame Pandas

- 4.1.1.3 показ основных характеристик датасета, удаление лишних столбцов
- 4.1.1.4 замена названия колонок, обработка пропущенных значений, проверка дубликатов и замета типов данных
- 4.1.1.5 кодирование категориальных признаков, масштабирование данных, и корреляционный анализ данных;
- 4.1.1.6 разделение данных на обучающий набор и тестовой набор;
- 4.1.1.7 выбор алгоритмов машинного обучения для решения задачи классификации
- 4.1.1.8 обучение модели
- 4.1.1.9 оценивать качество каждой модели
- 4.1.1.10 выбор модели с наилучшей производительностью
- 4.1.1.11 вывод результатов анализа.

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 данные, 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 надёжно 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 с 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, seaborn, matplotlib, scikit-learn.

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.	Подготовка документов “Задание на выполнение ВКРБ”, “Календарный план выполнения ВКРБ”	<u>20.11.2023</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
2.	Представление документов “Задание на выполнение ВКРБ”, “Календарный план выполнения ВКРБ” (Смотр № 1)	<u>03.12.2023</u> Планируемая дата		Нормоконтролер	Кротов Ю.Н.
3.	Формулирование проблемы, цели и задач работы	<u>15.01.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.

4.	Подготовка рабочих материалов	<u>15.01.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
5.	Разработка первой части РПЗ “Постановка задач разработки”	<u>20.02.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
6.	Разработка технического задания	<u>10.03.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
7.	Представление рабочих материалов, технического задания (Смотр № 2)	<u>20.03.2024</u> Планируемая дата		Нормоконтролер	Кротов Ю.Н.
8.	Разработка программы и методики испытания	<u>21.03.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
9.	Разработка второй части РПЗ «Исследовательская часть»	<u>10.04.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
10.	Разработка третьей части РПЗ “Конструкторско-технологическая часть”	<u>10.04.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
11.	Защита макета программы (АСОиУ)	<u>12.05.2024</u> Планируемая дата		Нормоконтролер	Кротов Ю.Н.
12.	Разработка заключения, приложений, оформление работы	<u>23.05.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
13.	Подготовка доклада и презентации	<u>25.05.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
14.	Получение заключения научного руководителя	<u>28.05.2024</u> Планируемая дата		Руководитель ВКР	Григорьев Ю.А.
15.	Допуск работы к защите на ГЭК (нормоконтроль)	<u>30.05.2024</u> Планируемая дата		Нормоконтролер	Кротов Ю.Н.
16.	Защита работы на ГЭК	<u>01.06.2024</u> Планируемая дата		Комиссия ГЭК	

8. Порядок контроля и приёмки

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