

Техническое задание

на курсовой проект по дисциплине «Алгоритмизация и программирование» на тему «Оценка аренды квартиры»

Шифр работы: 30021002

Цель работы: создание модели машинного обучения, предсказывающую стоимость ежемесячной арендной платы за квартиру в Москве и интегрируемой с API Telegram.

Требования к программе:

Бизнес:

1. Проект полного цикла - от сбора данных для обучения, до предсказания целевой переменной.

Функциональные:

1. Команда, запускающая парсинг данных для модели МО.
2. Предобработка данных - преобразование категориальных признаков, используя Dummy-переменные), обработка пропущенных значений, добавление информации о районе (кафе, музеи, фитнес-клубы и т.д.), расстояние до центра города и т.д.
3. Определение координат и района квартиры через API Geopy по адресу, расстояние до ближайшего метро, положение относительно колец Москвы (Садовое, Бульварное, 3 Транспортное, МКАД) и т.д.
4. Получение от пользователя информации о квартире (самостоятельный ввод, ссылка с ЦИАН).
5. Вывод предсказания арендной платы.

6. Визуализация различных зависимостей признаков квартиры (например, как зависят кол-во комнат и район квартиры от арендной платы и наоборот)

7. Возможность получить FAQ, в котором будут описаны алгоритмы, используемые при обучении модели, формально записанные на математическом языке.

Надежность:

1. Проверка входных данных на нужный тип и формат.
2. Запись логов в .txt файл
3. Автоматическая перезагрузка программы в случае ошибки
4. Наличие виртуального окружения, которое позволит запускать бота с любого устройства, соответствующего техническим требованиям.

Требования к составу технических средств:

Рекомендуемые:

Процессор: Intel(R) Core(TM) i7-2630QM CPU @ 2.00GHz

Видеокарта: NVIDIA GeForce GTM 540M

Объем встроенной памяти: 5ГБ

Оперативная память: 6,00 ГБ

Google Chrome - нужную версию смотреть в документации.

Наличие аккаунта в Telegram.

Ожидаемые результаты: рабочая версия программы, документация к программе: техническое задание, итоговый отчет.

Требования к информационной и программной совместимости:

Операционная система: Ubuntu 18.04+, Windows 10 21H2 и выше

Среда программирования: PyCharm Community Edition 2021.3.1.

Язык программирования:

В системе должен быть предустановлен Python 3.9. Для просмотра алгоритма обучения модели машинного обучения желательно наличие Anaconda 3, с предустановленным Jupyter Notebook, также можно воспользоваться Google Colab. *Установка всех необходимых инструментов будет описана в документации.*

Используемые библиотеки:

Pandas - работа с табличными данными.

NumPy - работа с массивами, матрицами.

Aiogram - обращение к API Telegram.

GeoPy - геокодирование.

SeaBorn - визуализация данных

MatPlotLib - визуализация данных

Altair - визуализация данных

Исполнители работы: Рожин А.К.