

Московский Авиационный Институт
(Национальный Исследовательский Университет)

Отчет по лабораторной работе
по курсу "Искусственный интеллект"

Студент: Спиридонов К.А.

Группа: М80-307Б-19

Преподаватель: Ахмед Самир

Москва, 2022

Постановка задачи

В данной лабораторной работе, вы выступаете в роли предприимчивого начинающего стартапера в области машинного обучения. Вы заинтересовались этим направлением и хотите предложить миру что-то новое и при этом неплохо заработать. От вас требуется определить задачу которую вы хотите решить и найти под нее соответствующие данные. Так как вы не очень богаты, вам предстоит руками проанализировать данные, визуализировать зависимости, построить новые признаки и сказать хватит ли вам этих данных, и если не хватит найти еще. Вы готовитесь представить отчет ваши партнерам и спонсорам, от которых зависит дальнейшая ваша судьба. Поэтому тщательно работайте:) И главное, день промедления и вас опередит ваш конкурент, да и сплагиченная работа отразится на репутации

Выбранный датасет:

Mobile Price Classification (<https://www.kaggle.com/iabhishekofficial/mobile-price-classification>)

Решаемая задача

Предсказание признака Price.

Описание входных данных

- *id* – id
- *battery_power* – Емкость батареи в МАЧ
- *blue* – Наличие bluetooth
- *clock_speed* - скорость, с которой микропроцессор выполняет инструкции
- *dual_sim* – Поддержка 2 симкарт
- *fc* – Разрешение фронтальной камеры в мегапикселях
- *four_g* – Наличие 4g
- *int_memory* – Емкость памяти в гб
- *m_dep* – Глубина в см
- *mobile_wt* – Вес в гр
- *n_cores* – Количество ядер
- *pc* – Разрешение главной камеры в мегапикселях
- *ram* – ОЗУ в мб
- *sc_h* – Высота экрана в см
- *sc_w* – Ширина экрана в см
- *talk_time* – Время в режиме разговора
- *three_g* – Наличие 3g
- *touch_screen* – Наличие тачскрина
- *wifi* – Наличие wifi
- *price_range* – К какой ценовой категории относится (0, 1, 2 ,3)

Анализ данных

Типы признаков

- Все признаки количественные
- Бинарные: *blue*, *dual_sim*, *four_g*, *three_g*, *touch_screen*, *wifi*
- Исследуемое значение: *price_range*

Размер

Строк: 2000

Столбцов: 21

Распределение признаков с числовыми полями



Решаемая задача

Классифицировать price_range.

Проблемы

В анализе данных проблем не возникло.

Вывод

В ходе лабораторной работы был проанализирован датасет Mobile Price Classification. Показано, как распределение признака, который предстоит исследовать, так и его зависимость от других признаков. Я считаю, что данных данных достаточно.