

Московский авиационный институт
(национальный исследовательский университет)

Отчёт по лабораторной работе №0
по курсу «Искусственный интеллект и машинное обучение»

Студент Шавандрин Ф.М.
Группа М8О-308Б-19
Дата 06.05.2022

Москва, 2022

Цель работы: анализ данных для решения задач машинного обучения.

Описание

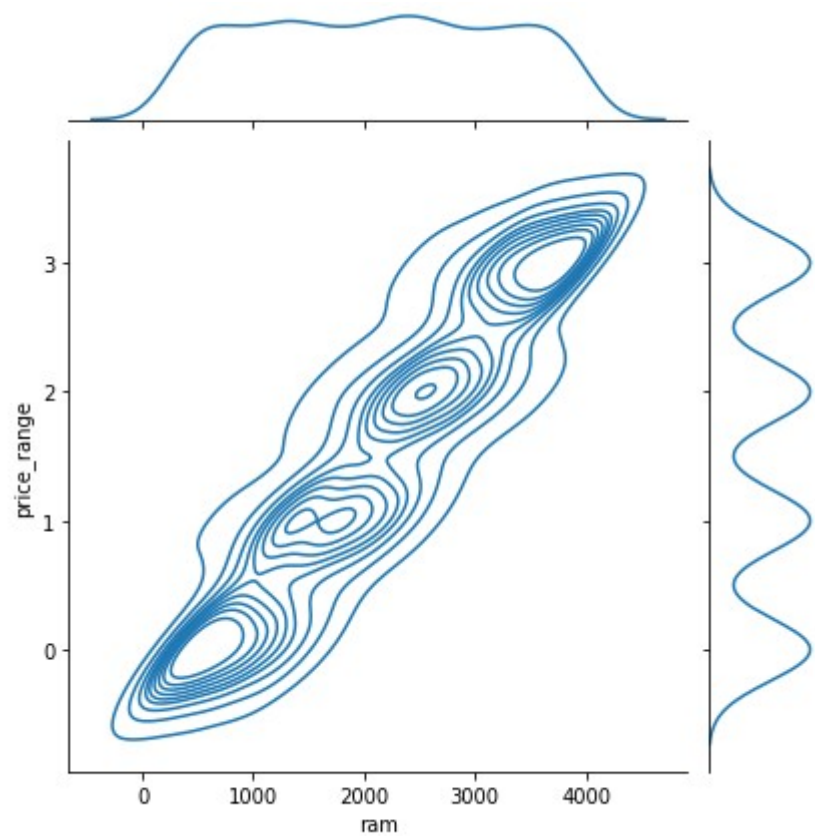
В качестве основной задачи я выбрал задачу классификации цены мобильного телефона на основе его характеристик. Представим, что мы наладили производство телефонов. Нам остаётся понять, по какой цене их стоит продавать, чтобы у нас была прибыль.

Вот некоторые основные **статистические данные**:

	battery_power	blue	clock_speed	dual_sim	fc	four_g	int_memory	m_dep	mobile_wt	n_cores	...	px_height	
count	2000.000000	2000.0000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	2000.000000	...	2000.000000	20
mean	1238.518500	0.4950	1.522250	0.509500	4.309500	0.521500	32.046500	0.501750	140.249000	4.520500	...	645.108000	12
std	439.418206	0.5001	0.816004	0.500035	4.341444	0.499662	18.145715	0.288416	35.399655	2.287837	...	443.780811	4
min	501.000000	0.0000	0.500000	0.000000	0.000000	0.000000	2.000000	0.100000	80.000000	1.000000	...	0.000000	5
25%	851.750000	0.0000	0.700000	0.000000	1.000000	0.000000	16.000000	0.200000	109.000000	3.000000	...	282.750000	8
50%	1226.000000	0.0000	1.500000	1.000000	3.000000	1.000000	32.000000	0.500000	141.000000	4.000000	...	564.000000	12
75%	1615.250000	1.0000	2.200000	1.000000	7.000000	1.000000	48.000000	0.800000	170.000000	7.000000	...	947.250000	16
max	1998.000000	1.0000	3.000000	1.000000	19.000000	1.000000	64.000000	1.000000	200.000000	8.000000	...	1960.000000	19

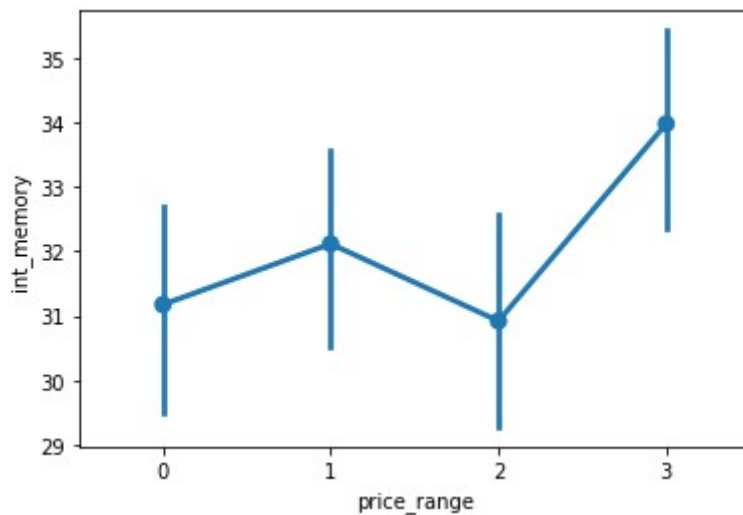
Анализ и визуализации данных.

Посмотрим вероятное значение ценовой категории для данных значений RAM:



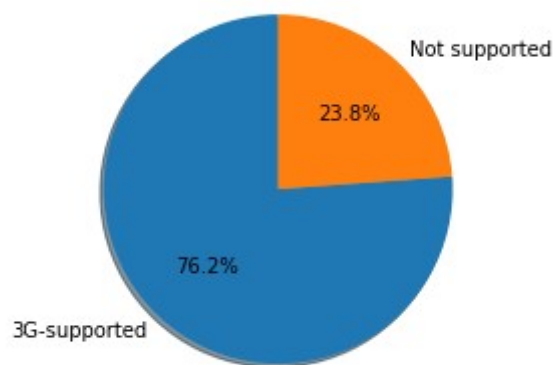
Можно заметить, что примерно в диапазоне значений RAM от 800 до 900, телефон к относится к 1 ценовой категории (или 0 на графике). Аналогично, и для других диапазонов RAM можно сделать похожие выводы.

Внутренняя память телефона и ценовая категория:

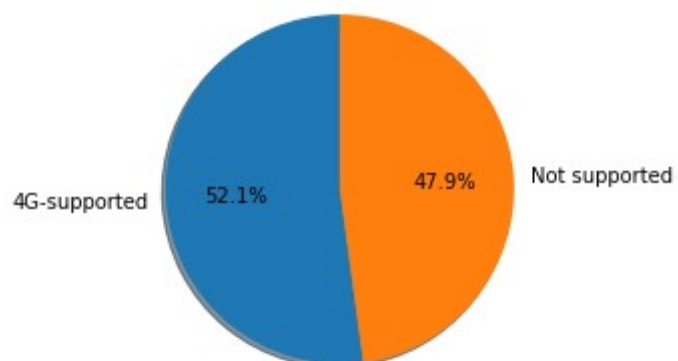


По данному графику мы не можем явно сказать, зависит ли ценовая категория от внутренней памяти телефона.

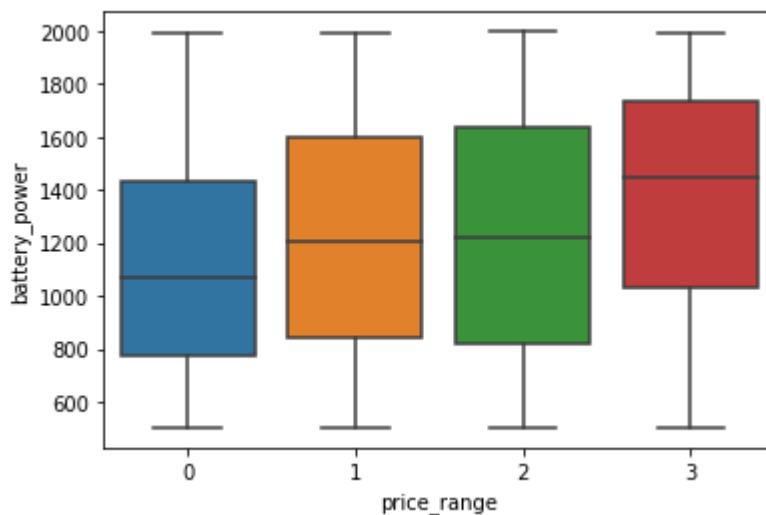
Телефоны, которые поддерживают 3G:



Телефоны, поддерживающие 4G:

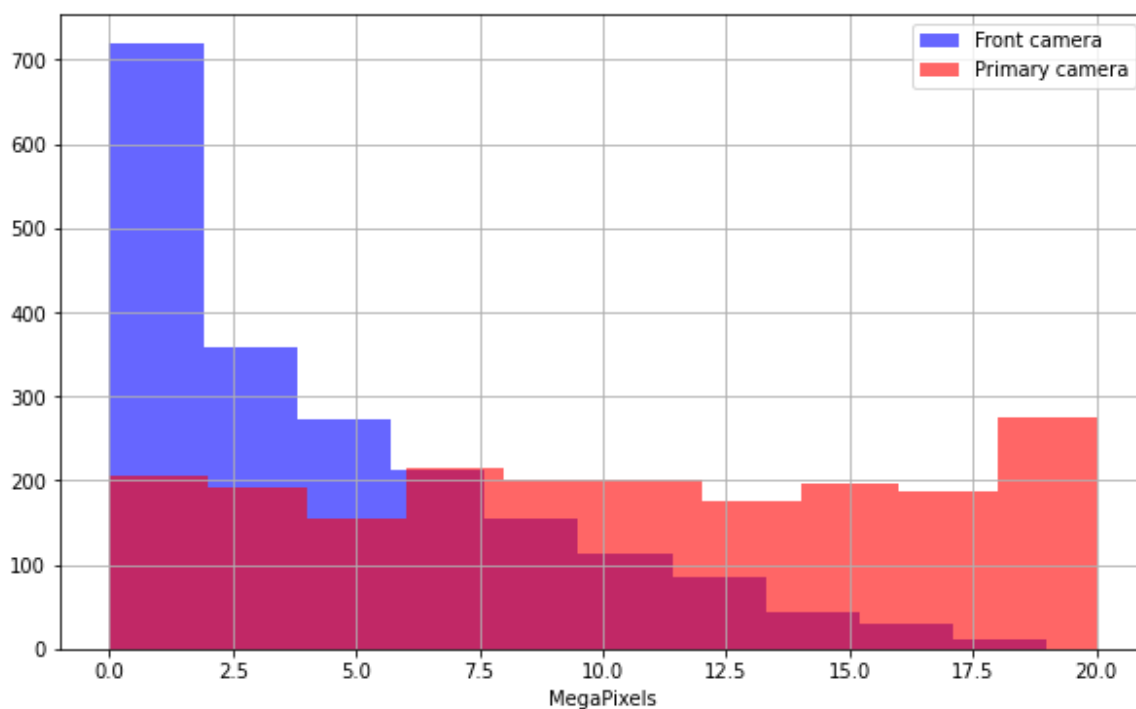


Заряд батареи и ценовая категория:



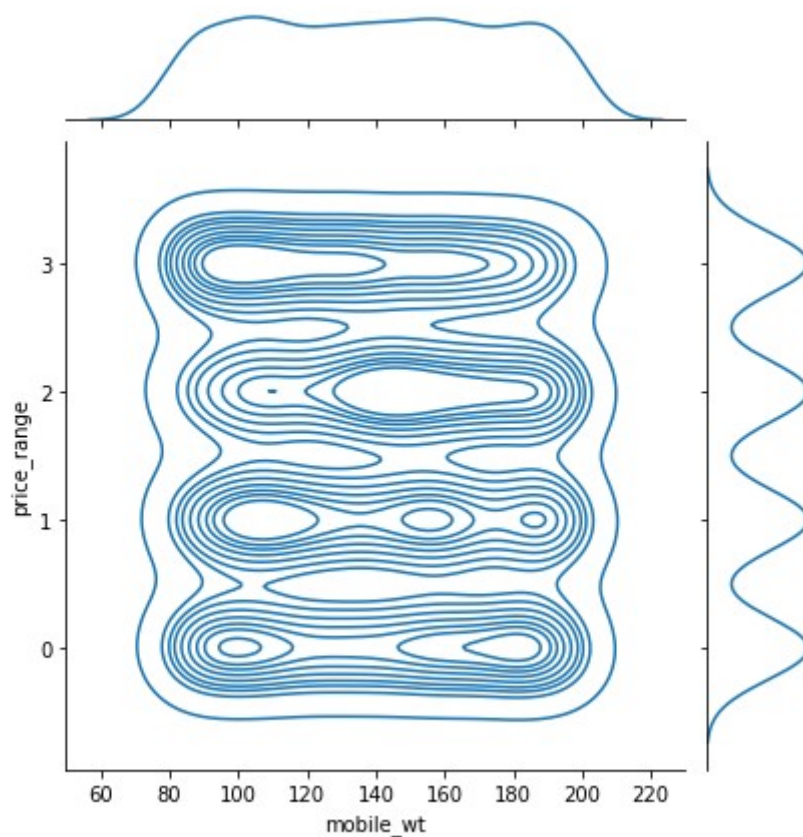
Анализируя данный график, можно увидеть легкую корреляцию заряда батареи и ценовой категории. Чем больше заряд, тем телефон дороже.

Гистограмма для количества телефонов и числа мегапикселей на переднем и задней камере:



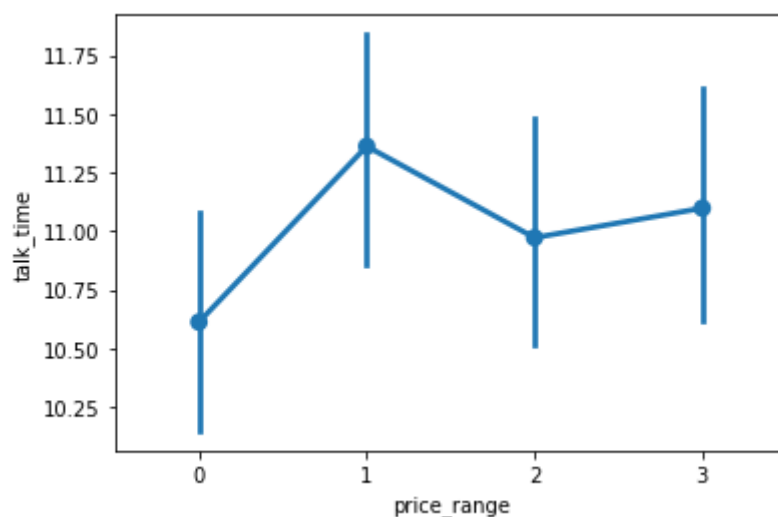
По графику можно сделать вывод о том, что нельзя однозначно сказать, сколько телефонов имеет некоторое число мегапикселей на передней или задней камере.

Вес телефона и ценовая категория:



Здесь мы можем заметить вероятное значение ценовой категории, равное 3 (или 2 на графике) при значении веса телефона от 140 до 160.

Время разговоров и ценовая категория:



По данному графику мы не можем явно сказать, зависит ли ценовая категория от времени разговора.

Выводы: Был проведен анализ данных характеристик различных телефонов для задачи классификации. Была произведена также визуализация данных, которая помогла определить некоторые закономерности между теми или иными признаками. Как итог, мы имеем достаточно информативный набор данных, который можно использовать в задаче классификации.