

Capstone Project

Mobile Price Range Prediction Technical Documentation

Table of Content:-

1. Abstract
2. Introduction
3. Problem Statement
4. Data Description
5. Exploratory Data Analysis
6. Modelling
7. Evaluation
8. Feature Importance
9. Conclusion

Abstract:-

The goal of this Mobile Price Range Prediction project is to build a predictive model that accurately predicts the price range of a mobile phone based on its features. The dataset used in this project contains information on 21 features of various mobile phones such as their processor, display, battery, camera, RAM, etc. The dataset also includes the actual prices of each mobile phone. A machine learning model is then trained on the dataset to predict the price of a mobile phone given its features. The model is evaluated on its accuracy, precision, recall, and other metrics. The results of the model are then compared to other models to determine which one gives the best performance.

Introduction:-

This project aims to introduce a model that can accurately predict the price range of a mobile phone. The model will consider a variety of factors, such as the brand, model, processor, RAM, storage, display size, camera, battery life, and other features. By analysing these features and applying a machine learning algorithm, we can accurately predict the price range of a mobile phone.

The model will be developed using a dataset of prices of existing mobile phones and their features. The dataset will be used to train the model and test its accuracy. The model will be evaluated based on its ability to accurately predict the price range of a mobile phone. The model will be implemented on a web application, allowing users to enter the features of their desired mobile phone, and receive a prediction of its price range.

The model can be used to assist users in making informed decisions when purchasing a mobile phone. It can also help vendors optimise their pricing strategy by providing them with accurate pricing data for their products. Furthermore, this model can be improved over time to accommodate new phones and their features.

Overall, this project aims to develop a machine learning model that can accurately predict the price range of a mobile phone. This model will be implemented on a web application, allowing users to make informed decisions when purchasing a mobile phone.

Problem Statement:-

There are many things we consider before buying a mobile as we use our mobile for various purposes like connecting with our family & Office Colleagues, playing games, taking photos to keep our memory alive. So, these specifications such as RAM, internal memory, Wi-Fi, 3G/4G connectivity etc. play an important role to buy a mobile. To analyse this important factor from time to time and come up with the best set of specifications and price ranges so that people will buy the mobile. Hence through the various ML modules we will help the company to estimate the price of mobiles according to features so the maximum amount of sell will be possible.

Data Description:-

Before performing any operation on the dataset, it is important to understand the data. After loading the data, we observed the dataset by checking a few of the first and last rows. We checked the shape of the dataset and there are 2000 rows and 21 features columns in our dataset.

Let's understand the features present in our dataset.

Battery_power - Total energy a battery can store in one time measured in mAh

Blue - Has bluetooth or not

Clock_speed - speed at which microprocessor executes instructions

Dual_sim - Has dual sim support or not

Fc - Front Camera mega pixels

Four_g - Has 4G or not

Int_memory - Internal Memory in Gigabytes

M_dep - Mobile Depth in cm

Mobile_wt - Weight of mobile phone

N_cores - Number of cores of processor

Pc - Primary Camera mega pixels

Px_height - Pixel Resolution Height

Px_width - Pixel Resolution Width

Ram - Random Access Memory in Mega

Touch_screen - Has touch screen or not

Wifi - Has wifi or not

Sc_h - Screen Height of mobile in cm

Sc_w - Screen Width of mobile in cm

Talk_time - longest time that a single battery charge will last over a call

Three_g - Has 3G or not

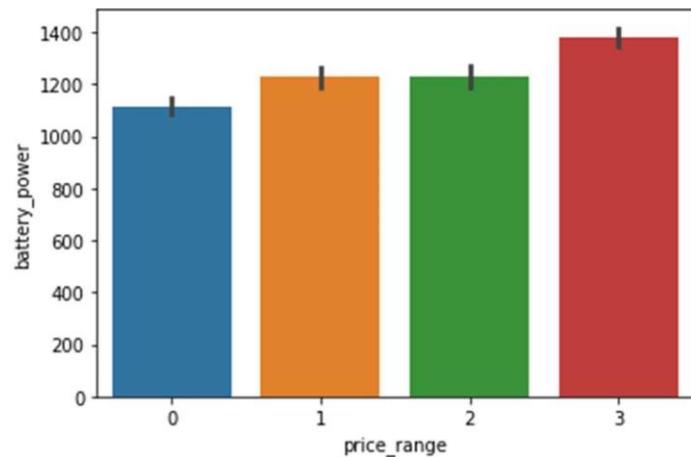
Wifi - Has wifi or not

Price_range - This is the target variable with value of 0(low cost), 1(medium cost),2(high cost) and 3(very high cost)

Exploratory Data Analysis:-

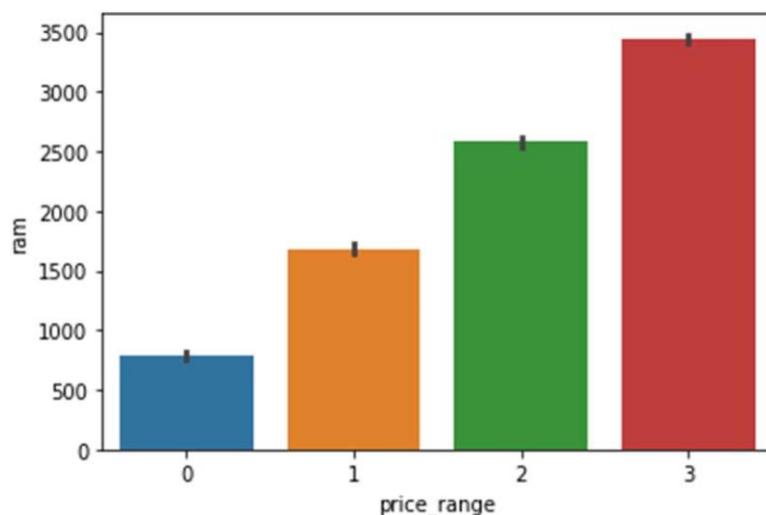
1. Analysis

1. Battery - price range



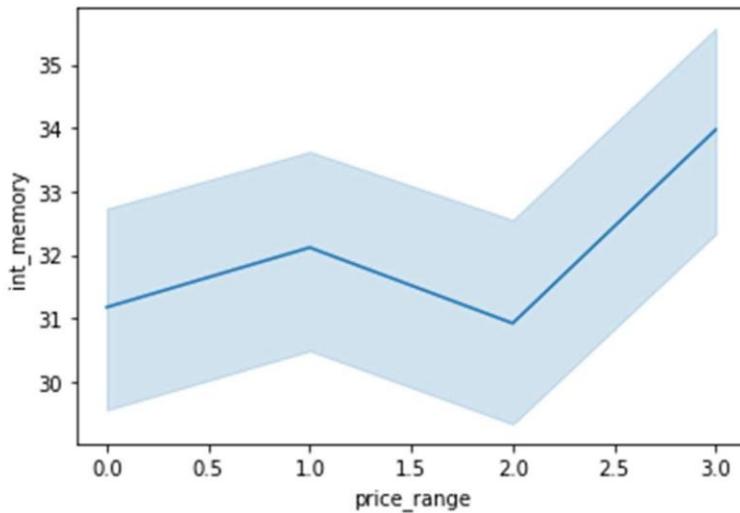
mid-range and high mid-range phones offer good battery performance.

2. Ram - Price Range



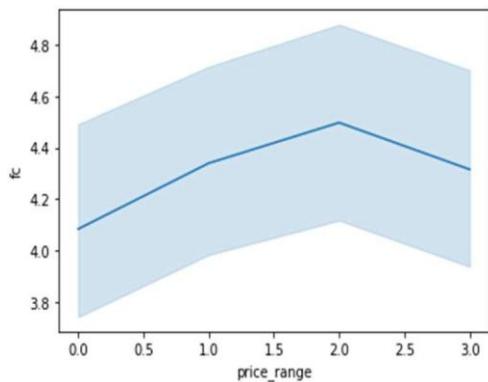
mostly only very high price range smartphones provide good ram

3. Memory - Price Range



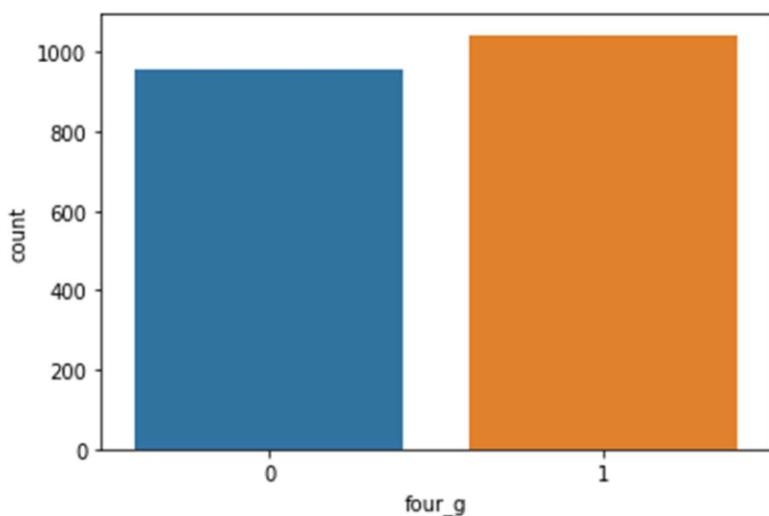
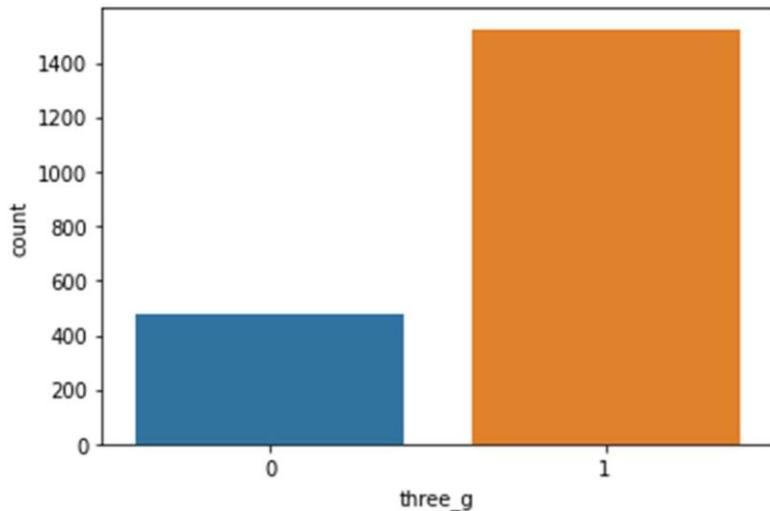
mid range and very high range offer good memory but very high range offer high memory

4. Front mega pixel - price range.

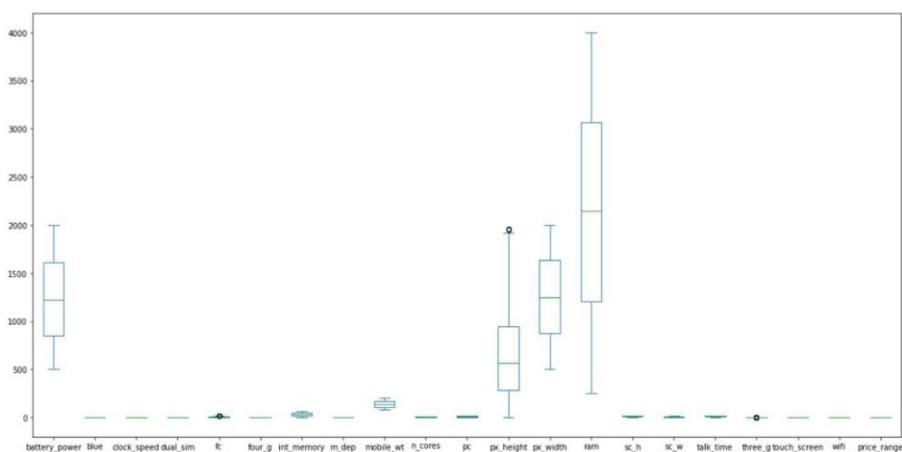


expensive phones and mid range phone offer same specs in terms of front mega pixels camera but high mid range phones offer good front camera

5. How many 3g phones and 4g phones

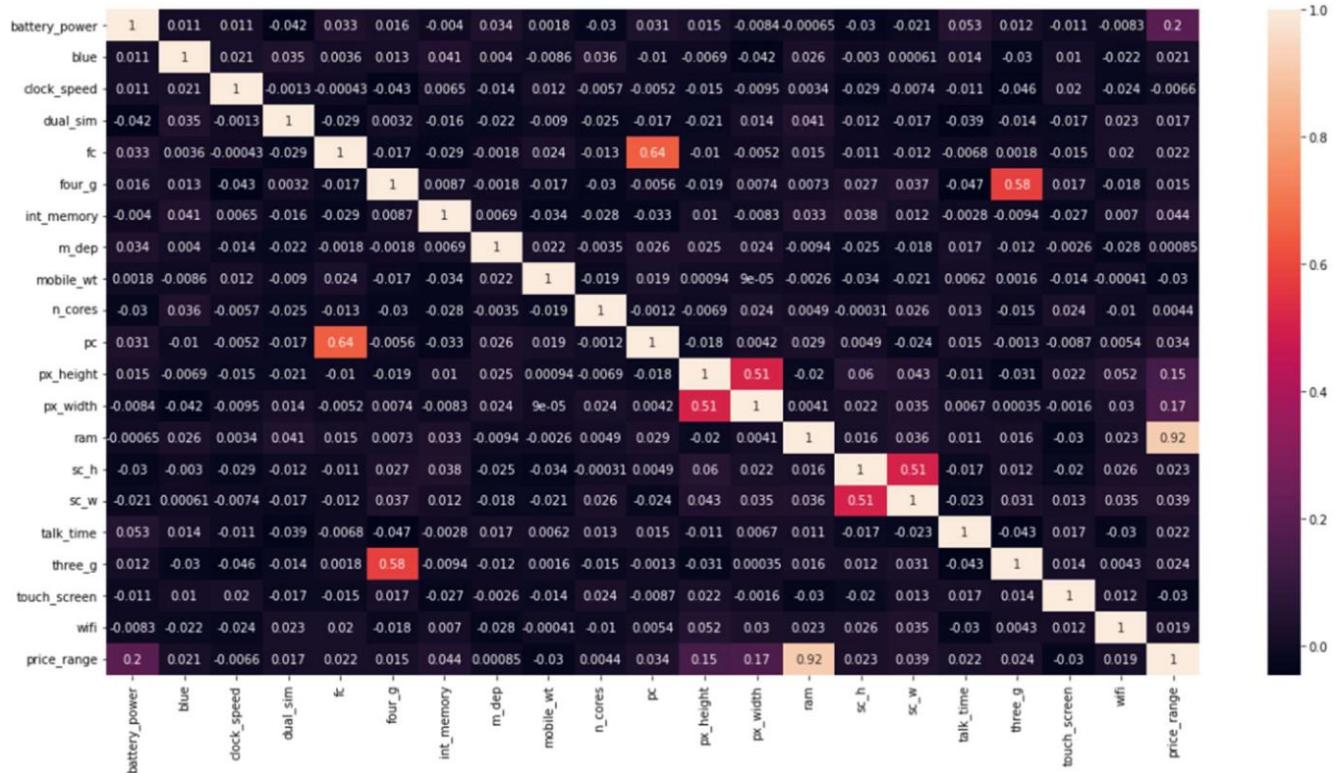


2. Outliers



in this data we can see that there are very few outliers in the data.

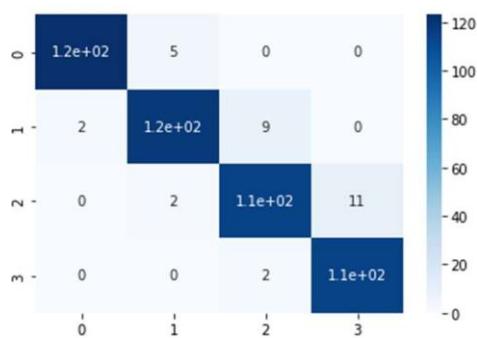
3. Correlation Analysis



after analysing heat map, there no independent variable are highly correlated except ram

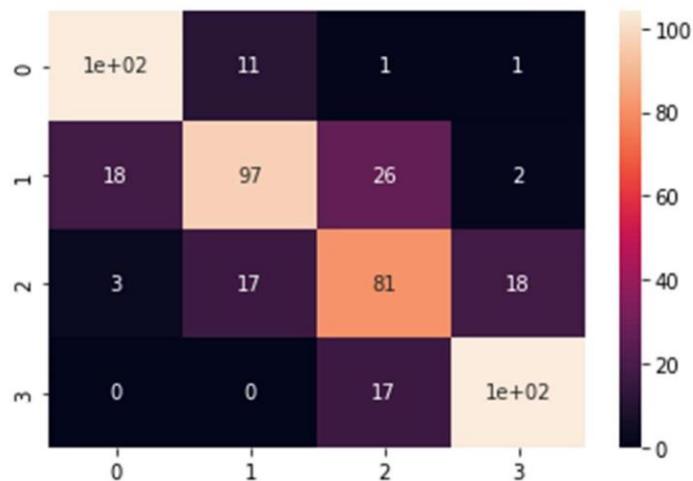
Modelling :-

A :- K-NEAREST NEIGHBORS:



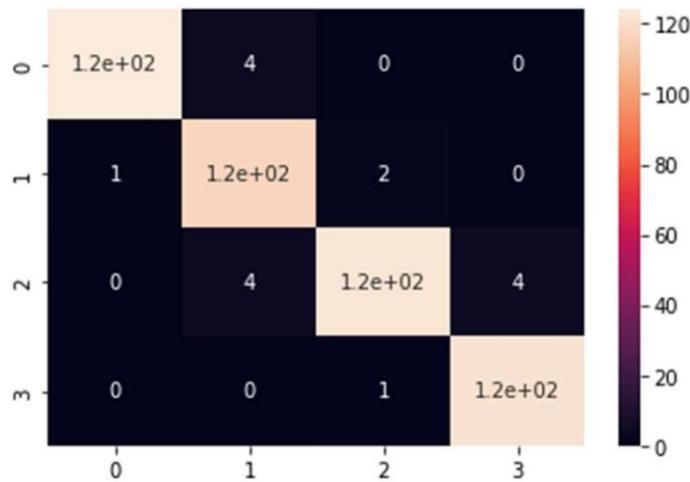
	precision	recall	f1-score	support
0	0.98	0.96	0.97	128
1	0.94	0.91	0.93	129
2	0.91	0.90	0.90	127
3	0.91	0.98	0.95	116
accuracy			0.94	500
macro avg	0.94	0.94	0.94	500
weighted avg	0.94	0.94	0.94	500

B :- DECISION TREE:



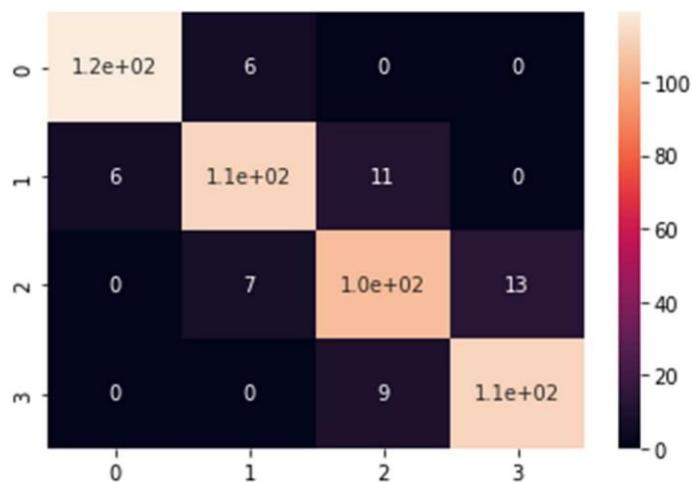
	precision	recall	f1-score	support
0	0.83	0.89	0.86	117
1	0.78	0.68	0.72	143
2	0.65	0.68	0.66	119
3	0.83	0.86	0.85	121
accuracy			0.77	500
macro avg	0.77	0.78	0.77	500
weighted avg	0.77	0.77	0.77	500

C :- LOGISTIC REGRESSION:



	precision	recall	f1-score	support
0	0.99	0.97	0.98	128
1	0.94	0.97	0.96	120
2	0.98	0.94	0.96	130
3	0.97	0.99	0.98	122
accuracy			0.97	500
macro avg	0.97	0.97	0.97	500
weighted avg	0.97	0.97	0.97	500

D :-RANDOM FOREST:-

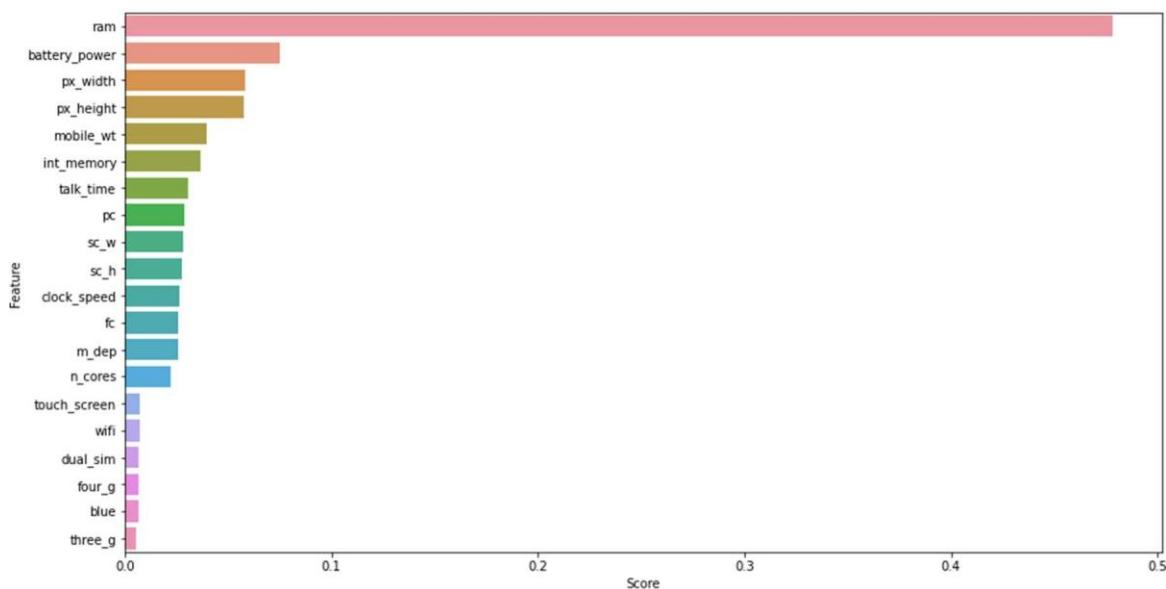


	precision	recall	f1-score	support
0	0.95	0.95	0.95	125
1	0.90	0.87	0.88	129
2	0.84	0.84	0.84	125
3	0.90	0.93	0.91	121
accuracy			0.90	500
macro avg	0.90	0.90	0.90	500
weighted avg	0.90	0.90	0.90	500

Evaluation

	Model	accuracy
0	LogisticRegression classifier	96.8
1	KNN classifier	93.8
2	RandomForest classifier	89.6
3	DecisionTree classifier	77.2

Feature Importance



Conclusion

1. From EDA we can see that mid-range and high mid-range phones offer good battery performance and expensive phones and mid-range phone offer same specs in terms of front mega pixels camera but high mid-range phones offer good front camera
2. There are some top features like Ram, battery_power, px_width, px_height, mobile_wt.
3. Form all the above experiments we can conclude that logistic regression and, KNN we got the best results