Capstone proposal

MOBILE PRICE PREDICTION

DOMAIN BACKGROUND

- *Mobile price prediction is the concept of predict the price of the mobile based on its qualities and features*
- *The qualities of mobile include its RAM,ROM,battery life, camera pixels, dual sim mobile weight....etc*
- *We should identify the price of the mobile based on these feature . We should know how these features effect the price of the mobile*
- *My intention is to help the shop owners to give correct rate to the mobile. By this they get the price for very genuine rate and they will sell very genuinely to the customers*

- *https://www.kaggle.com/iabhishekofficial/mobile-price-classification*
- The problem should be solved because the customers and shop owners should also understand the price of the mobile .

● PROBLEM STATEMENT

- *The aim of this project was to estimate the price of the mobile by its features such as ram,internal memory...... Etc . We Donot have to find the actual price but a range price should be predicted .This means the price may not be accurate but approximate price of the mobile should be .We can do this by using supervised learning techniques*

Data Sets and Inputs

- The data set will be in 2 data sets one is train and another one is test data set in this data set we have many features . The data set is taken from kaggle .
  https://www.kaggle.com/iabhishekofficial/mobile-price-classification
- The data set contain 20 columns these columns describe the features of the mobile. These are used to predict the price of the mobile
- The data set should be used because we need to predict the prices of the mobile using our past records and we should have knowledge on features of the mobile to predict its price
- In train data set we have the features except the target variable On this we apply our learning algorithm and predict its price
- On test we will know which algorithm have better result based on That we will apply good algorithm
- There are many inputs in the problem these inputs are the features of the mobile the first feature was battery life based on the batteey life of the the price increases because the battery

life is very important feature to mobile and colour blue it not so important because many people will se the features of the mobile than the colour of the mobile clock speed it also effect the price of the mobile if clock speed is more tje price will be more . Dual sim many phones should have dual sim slot then the mobile will have high price. Fc means front camera it the very important feature which describe the price of the mobile many people look for a good camera . Four-g feature is also a features now a days many phones come with 4g features . The memory is also an important feature decide the rate The customers mainly look at internal memory of the phone. M-depth it tje depth of the mobile it is not an important feature which will effect the price of the mobile. Mobile_weight the weight of the mobile is some what a considerable matter because some people donot prefer more weight. N_cores it is an impotant feature because it also indicates the speed of the mobile . Pc is a cammera feature which effects the price of the mobile. Px is also an important feature that effect the price of an mobile px-height it doesnot effect the price of the mobile. Px-width this feature also doesnot effect the price of the mobile. Ram is the most important and it is the first feature which effects the price of the mobile Sc_h this feature also effect the price of the mobile . Talk time is a considerable feature that effect the price of the mobole . 3g is also feature that effect the price mostly . Touch screen many people prefer to have touch screen mobile so it one of the important feature that effect the price of the mobile. Wifi also effect the price of the mobile

Solution

- **My** solution is to apply the supervised learning algorithms to the problem . I will use the regression i.e decision tree algorithm to this model based on the features it is better to classify the problem by this model

Bench Mark model

- **The** bench mark model will be linear regression   and decision tree performs well with the training data but it will over fit the data which results the less performance in test data set. So linear regression model will perform well in this model

Evaluation metrics

- **I have** used the score() on test data by this i have evaluated the performance of the solution model and the bench mark model the decision over fits the data so it has less score than linear regression

Project design

Step 1:

- Firstly i have imported the data from The kaggle website
- I loaded the training,testing set

Step 2:

- I have created the statistics to view the features that effect the price of the mobile
- I have used the decision trees and Used score() function for it.

Step 3:

- Now i have used the bench mark model linear regression
- Then i have compared the score for the decision tree and bench mark model