

## Problem Description

Prabhat Doongarwal ( popular as “pd” ) has started his own mobile company. He wants to give tough fight to big companies like Apple,Samsung etc.

He does not know how to estimate price of mobiles his company creates. In this competitive mobile phone market you cannot simply assume things. To solve this problem he collects sales data of mobile phones of various companies.

Prabhat wants to find out some relation between features of a mobile phone(eg:- RAM, Internal Memory etc) and its selling price. But he is not so good at Machine Learning. So he needs your help to solve this problem.

In this problem you do not have to predict actual price but a price range indicating how high the price is.

## Data Description

The data has been divided into 2 groups :

- Training Set (train.csv)
- Test Set (test.csv)

The **training set** should be used to build your machine learning models. For the training set we provide the outcome( also known as target or class label) for each mobile sales data point.

The **test set** should be used to see how well your model performs on unseen data. For the test set we do not provide the outcome(target variable) for each mobile sales data point. For each data point in the test set use the model you trained to predict price range.

The training set contains 2000 data points. Each data point has 20 features and one target variable.

The test set contains 1000 data points. Each data point has 20 features.

## Features Description

- 1) **pc** : Primary Camera mega pixels
- 2) **fc** : Front Camera mega pixels
- 3) **sc\_h** : Screen Height of mobile in cm
- 4) **sc\_w** : Screen Width of mobile in cm
- 5) **m\_dep** : Mobile Depth in cm
- 6) **px\_width** : Pixel Resolution Width
- 7) **px\_height** : Pixel Resolution Height
- 8) **ram** : Random Access Memory in Mega Bytes
- 9) **int\_memory** : Internal Memory in Giga Bytes
- 10) **four\_g** : Has 4G or not
- 11) **three\_g** : Has 3G or not
- 12) **dual\_sim** : Has dual sim support or not
- 13) **battery\_power** : Total energy a battery can store in one time measured in mAh
- 14) **touch\_screen** : Has touch screen or not
- 15) **clock\_speed** : speed at which microprocessor executes instructions
- 16) **n\_cores** : Number of cores of processor
- 17) **wifi** : Has wifi or not
- 18) **blue** : Has bluetooth or not
- 19) **mobile\_wt** : Weight of mobile phone
- 20) **talk\_time** : longest time that a single battery charge will last when you are constantly talking on the phone
- 21) **price\_range** : This is the target variable with value of 0(low cost), 1(medium cost), 2(high cost) and 3(very high cost).

## Goal

Your task is to classify each data point of test.csv into either of 4 classes(0,1,2,3).

## Metric

Your score will be percentage of mobile phones in test.csv that your machine learning model correctly classifies. This metric is known as “accuracy”.

## Submission File Format

You have to submit a CSV file with exactly 1000 entries plus a header row. This csv file must contain output of your model on each point in test set(test.csv). Also you have to submit your source code file.

You can see the sample submission file to see in what format you have to submit your solution.