

Data

The mobile price classification dataset is a dataset containing various attributes of mobile phones, which can be used to build a machine learning model for predicting price ranges of mobile devices. The dataset includes features such as:

- **Id** : ID.
- **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 megapixels.
- **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 megapixels.
- **px_height** : Pixel Resolution Height.
- **px_width** : Pixel Resolution Width.
- **ram** : Random Access Memory in Megabytes.
- **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.
- **three_g** : has 3g or not.
- **touch_screen** : Has touch screen or not
- **wifi** : has wifi or not.
- **price_range**: This is the target variable with values of 0(low cost), 1(medium cost), 2(high cost) and 3(very high cost).

Objective

The objective of this classification task is to develop a machine learning model that can predict the price range of a mobile phone based on its various attributes. The dataset can be used to train and evaluate the model's performance in accurately classifying mobile prices.

By analyzing this dataset, candidates are required to follow machine learning workflow particularly by:

- Exploring the data and providing descriptive statistics.
- Performing data visualization. (provide comments on each graph)
- Applying necessary data transformations and feature engineering. (provide justification for each step)
- Performing feature selection to select most related features.
- Trying training more than one classification model using “train.csv” and performing hyperparameter tuning.
- Evaluating models on training and testing data by providing classification scores like accuracy, precision, recall ,f1-score, etc.. (provide your comments about results)
- Selecting the best model and saving it as a pkl file.
- **Bonus**: Use flask framework to serve predictions as an API.

Submission Format

Candidates are required to follow the following folder structure:

(Note: This is the basic structure, You can add more files based on your need especially in bonus part)

```
.
└─ Mobile_price_prediction/
    │
    ├── data/
    │   │
    │   ├── train.csv
    │   └─ test.csv
    ├── model/
    │   └─ model.pkl           // your best model
    └─ notebooks/
        └─ price_prediction.ipynb // notebook for your solution
```