Project

Linear Models for Classification

In the previous lab, we learned about classification where the task is to predict the class or category. In this project we'll also work on a classification task.

As always, we are going to approach our problem following a typical Machine Learning workflow.

- Problem formation
- Finding data (Import the dataset)
- Exploring insights in data or EDA
- Data preprocessing
- Choosing and training a model
- Evaluating a model

By being systematic and keeping things organized, it will help you to reproduce some parts of the project or reuse them into other problems.

1. Problem Formulation

Let's say you have an idea of a revolutionary mobile phone and you want to establish a startup, but you know little about the price of the mobile phones. You are interested in learning that!

Fortunately, there is this mobile dataset that you can use to learn about the price ranges of mobiles based on their features such as wifi & bluetooth support etc…

2. Finding the Data

Here are the details of the features. It has 21 features. The target feature is the price range and it has four price ranges: 0(low cost), 1(medium cost), 2(high cost) and 3(very high cost).

- **batter_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 Bytes
- **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 when you are talking
- **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 value of 0(low cost), 1(medium cost), 2(high cost) and 3(very high cost).