Mobile Price Classification Dataset Report

Introduction:

The Mobile Price Classification dataset is designed to predict the price range of mobile phones based on various features. This dataset is useful for classification tasks and can help in understanding the factors that influence the pricing of mobile phones.

Dataset Overview:

Number of Rows: 2000

Number of Columns: 21

Features:

The dataset includes the following features:

battery_power: Total energy a battery can store in one time measured in mAh

blue: Whether the phone has Bluetooth (0: No, 1: Yes)

clock_speed: Speed at which microprocessor executes instructions

dual_sim: Whether the phone supports dual SIM (0: No, 1: Yes)

fc: Front camera megapixels

four_g: Whether the phone supports 4G (0: No, 1: Yes)

int_memory: Internal memory in GB

m_dep: Mobile depth in cm

mobile_wt: Weight of the mobile phone

n_cores: Number of cores of the processor

pc: Primary camera megapixels

px_height: Pixel resolution height

px_width: Pixel resolution width

ram: Random Access Memory in MB

sc_h: Screen height of the mobile in cm

sc_w: Screen width of the mobile in cm

talk_time: Maximum time in hours a battery will last

three_g: Whether the phone supports 3G (0: No, 1: Yes)

touch_screen: Whether the phone has a touch screen (0: No, 1: Yes)

wifi: Whether the phone has WiFi (0: No, 1: Yes)

price_range: Target variable with values (0: Low Cost, 1: Medium Cost, 2: High Cost)

Data Exploration:

First 5 Rows: Displaying the initial entries to understand the structure and initial values.

Last 5 Rows: Displaying the final entries to ensure data consistency.

Shape of the Dataset: The dataset contains 2000 rows and 21 columns.

Descriptive Statistics:

Provides summary statistics for numerical features, including mean, standard deviation, minimum, and maximum values.

Data Cleaning

Missing Values: The dataset does not contain any missing values, ensuring completeness.

Null Values: No null values are present in the dataset.

Data Transformation

Column Renaming: Columns are renamed for better readability (e.g., blue to bluetooth).

Binary Columns Transformation: Columns like bluetooth, dual_sim, four_g, three_g, touch_screen, and wifi are t ransformed to categorical values (0: No, 1: Yes).

Price Range Transformation: The price_range column is mapped to categorical values (0: Low Cost, 1: Medium Cost, 2: High Cost).

Data Visualization:

Correlation Analysis: Identifying the highest correlated features with the price_range.

3G vs Sale Price: Bar plot showing the relationship between 3G support and price range.

Bluetooth Support vs Price Range: Count plot showing the distribution of Bluetooth support across different price ranges.

Pixel Resolution vs Price Range: Scatter plot showing the relationship between pixel resolution height and width with price range.

Screen Dimensions vs Price Range: Scatter plot showing the relationship between screen height and width with price range.

Conclusion:

The Mobile Price Classification dataset provides a comprehensive set of features that can be used to predict the price range of mobile

phones. Through data exploration, cleaning, transformation, and visualization, we can gain valuable insights into the factors that influence mobile phone pricing. This analysis can be further extended to build predictive models and enhance our understanding of mobile phone market dynamics.