Mobile Phone Price Prediction Report

Objective

Build a system to predict the price range (low/medium/high/very high) of a mobile phone based on its features.

Dataset

The dataset contains various features of mobile phones and their price range. Below is a description of the columns:

Feature	Description
battery_power	Battery Capacity in mAh
blue	Has Bluetooth or not
clock_speed	Processor speed
dual_sim	Has dual sim support or not
fc	Front camera megapixels
four_g	Has 4G or not
int_memory	Internal Memory in GB
m_deep	Mobile depth in cm
mobile_wt	Weight in gm
n_cores	Processor Core Count
pc	Primary Camera megapixels
px_height	Pixel Resolution height
px_width	Pixel Resolution width
ram	Ram in MB
sc_h	Mobile Screen height in cm
sc_w	Mobile Screen width in cm
talk_time	Battery talk time in hours
three_g	Has 3G or not
touch_screen	Has touch screen or not
wifi	Has WiFi or not
price_range	Target: 0=low, 1=medium, 2=high, 3=very high

Approach

- Data loaded and split into features and target (`price_range`).
- Data split into training and test sets (80/20 split).
- Pipeline: Imputation (mean) → Scaling → Random Forest Classifier.
- Hyperparameter tuning using GridSearchCV (5-fold cross-validation).
- Model evaluation on test set.

Results

Best Parameters: - `max_depth`: 10 - `min_samples_split`: 2 - `n_estimators`: 100

Test Accuracy: 0.8925

Classification Report:

Class	Precision	Recall	F1-score	Support
0	0.95	0.95	0.95	100
1	0.83	0.84	0.84	100
2	0.84	0.83	0.83	100
3	0.95	0.95	0.95	100
accuracy			0.89	400
macro avg	0.89	0.89	0.89	400
weighted avg	0.89	0.89	0.89	400

Model saved as: `mobile_price_model.joblib`

Conclusion

The Random Forest model achieved high accuracy (89.25%) in predicting the price range of mobile phones based on their features. The model can be used to assist in pricing decisions for new mobile phones based on their specifications.