# 🔍 Decision Tree vs Random Forest – Car Price Range Classification

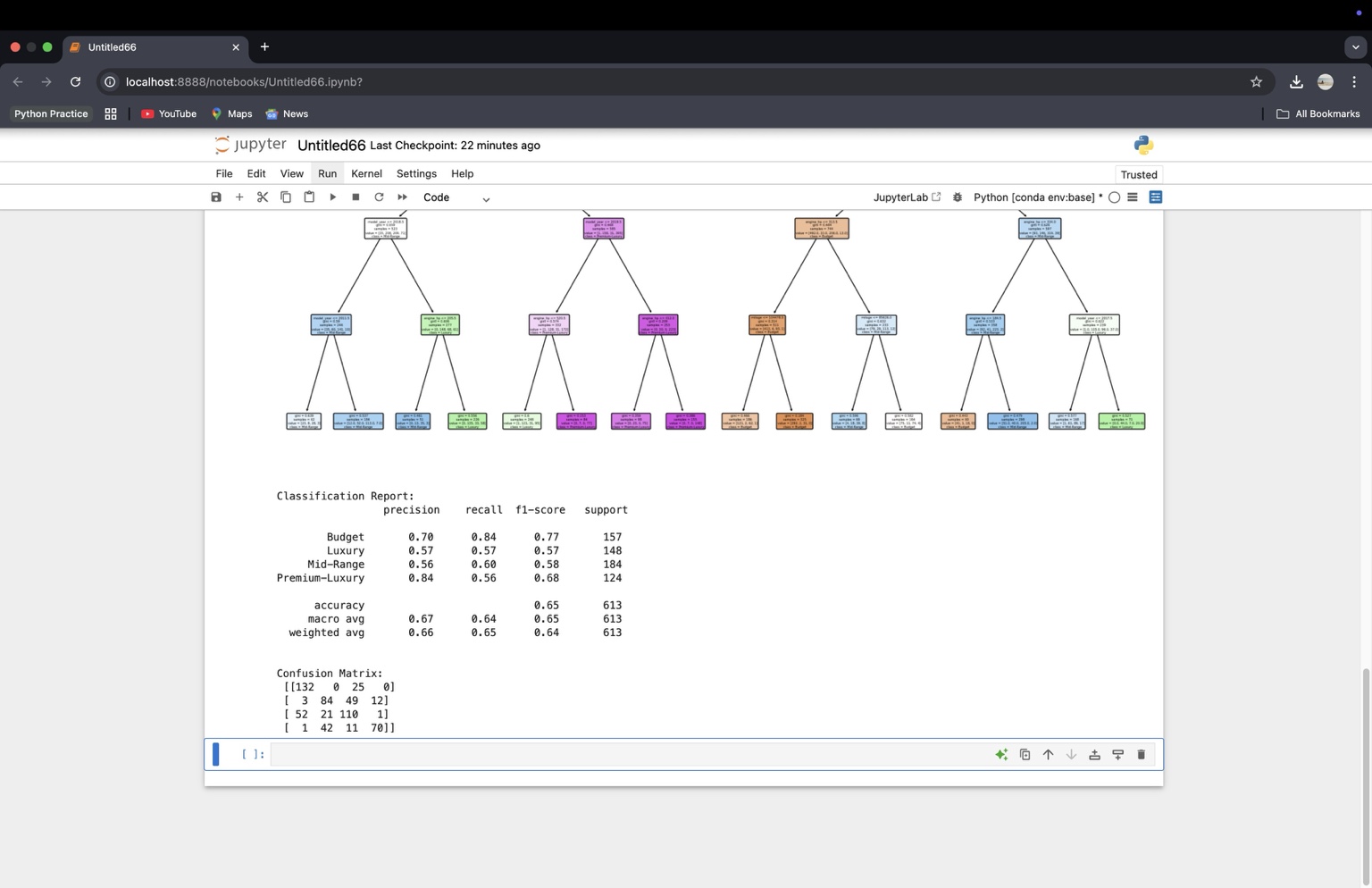
## 1. Model Accuracy

We trained both a Decision Tree and a Random Forest classifier on the same car dataset to predict price ranges.  
  
The accuracies were very close:  
- Decision Tree Accuracy: 65%  
- Random Forest Accuracy: 63%  
This indicates both models performed comparably well, with the Decision Tree having a slight edge.

## 2. Class-wise Performance (F1-Score)

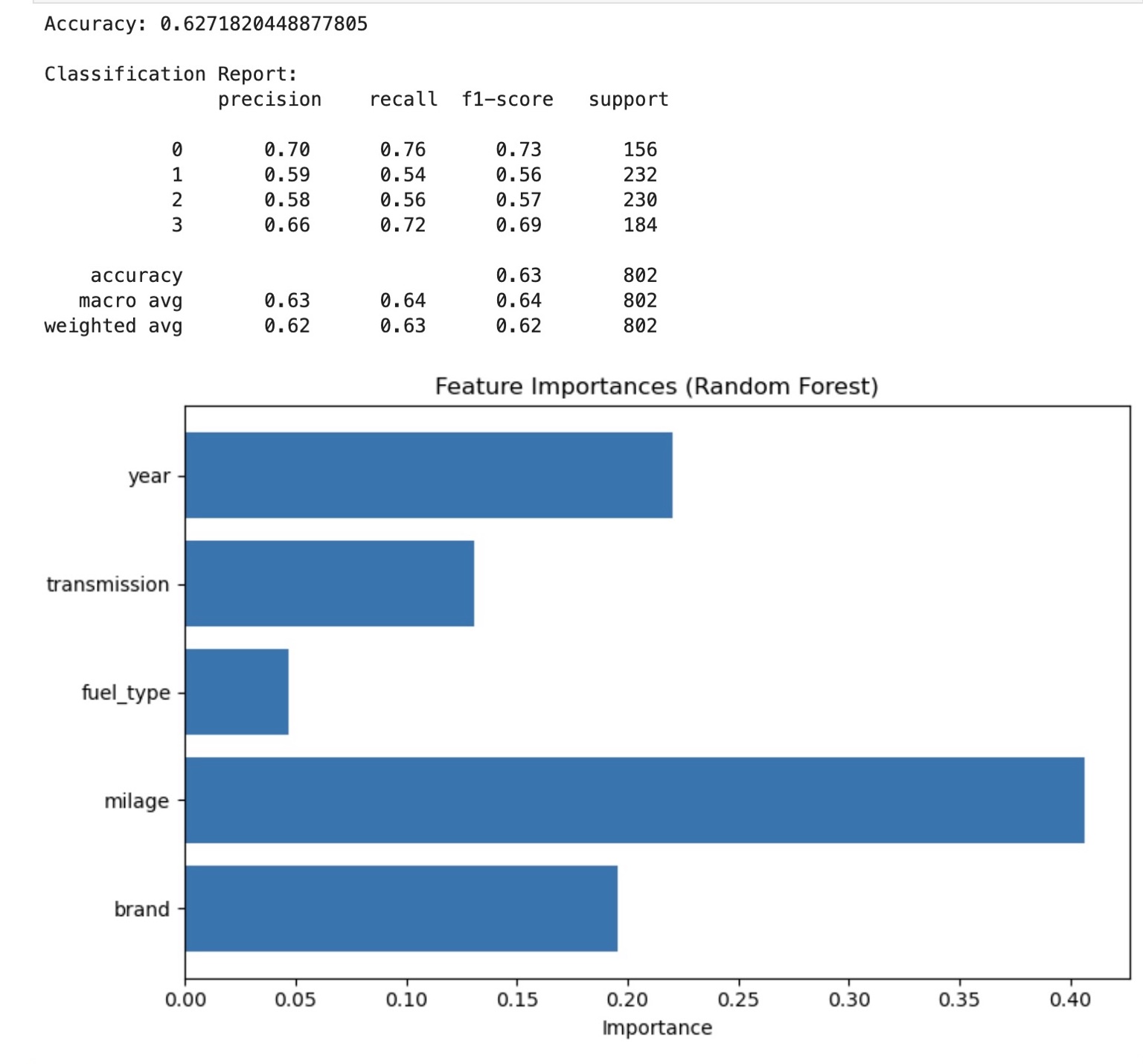
We compare the F1-scores (harmonic mean of precision and recall) for each class:  
  
| Class | Decision Tree | Random Forest |  
|------------------|----------------|----------------|  
| Budget (0) | 0.77 | 0.73 |  
| Luxury (1) | 0.57 | 0.56 |  
| Mid-Range (2) | 0.58 | 0.57 |  
| Premium-Luxury (3) | 0.68 | 0.69 |  
  
🔍 \*Insight:\* Both models struggle most with the Mid-Range and Luxury categories, likely due to overlapping features.

## 3. Confusion Matrix – Decision Tree



This shows how well the Decision Tree classified each price category. For example:  
- 132 Budget cars were correctly classified.  
- Some Luxury cars were misclassified as Mid-Range or Premium.  
- Budget and Premium-Luxury were predicted most accurately.

## 4. Random Forest: Classification & Feature Importance



The Random Forest classification report shows performance across all classes, with an overall accuracy of ~63%.  
  
The feature importance chart reveals:  
- `milage` was the most significant predictor of price.  
- `year` and `brand` were also important.  
- `fuel\_type` had the least influence on the outcome.

## 5. Final Comparison Summary

| Factor | Decision Tree | Random Forest |  
|--------------------|----------------------------|----------------------------|  
| Accuracy | ✅ Slightly higher | ❌ Slightly lower |  
| Interpretability | ✅ Very high (visual tree) | ❌ Lower (black-box) |  
| Overfitting Risk | ❌ Higher | ✅ Lower due to averaging |  
| Stability | ❌ Sensitive to changes | ✅ More stable predictions |  
| Feature Importance | ❌ Not direct | ✅ Clearly available |

🧠 \*Conclusion:\* The Decision Tree model provided slightly better accuracy in this case, and is easier to interpret. However, Random Forest is generally more robust and reliable across varied datasets. For production-grade applications, Random Forest is usually the preferred choice due to its ability to reduce overfitting.