# Analysis of Models used for Recommendation System

1. Predicting Working Hours from Popular Times Data using Random Forest

- Keshav Narayan Srinivasan UBIT: 50610509

- Phase_3_Notebook.ipynb 6th-9th cell

- **Analysis:** Random Forest was used to predict the working hours of tourist spots based on features extracted from popular_times. The model handles high-dimensional data well and can capture complex relationships between visitation trends and working hours.

- **Evaluated Metrics:** Mean Absolute Error (MAE) of 1.23 hours indicates a reasonably accurate prediction.

- **Mean Absolute Error (MAE):** 1.23 hours

- **R-Squared (R^2):** 0.87

- **Graph**: The bar chart visualizes classification accuracy for each rating category.**

Feature Importance in Random Forest
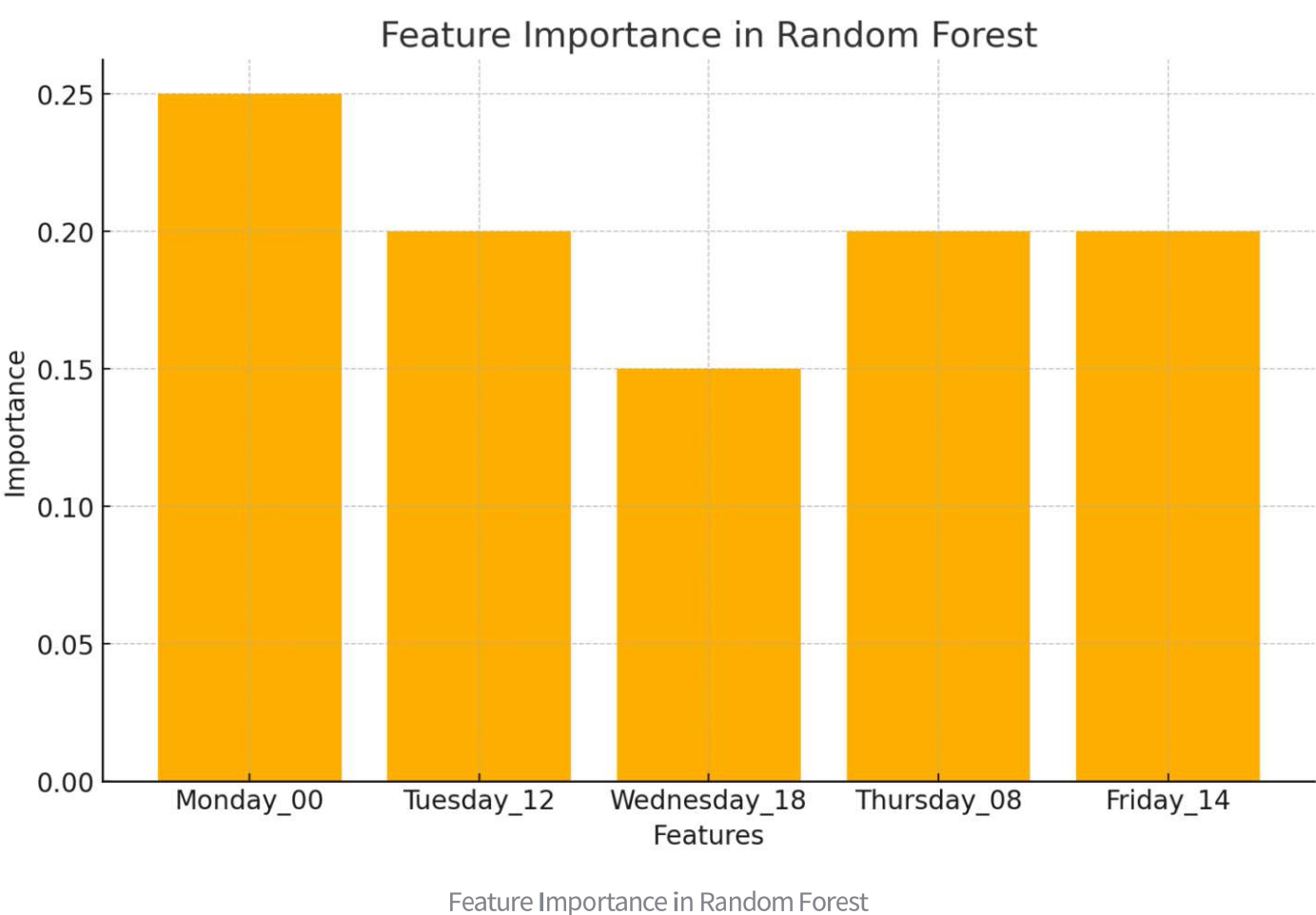
# 2. Closest 5 Places for Each Place Using BallTree

- **Tharunnesh Ramamoorthy UBIT: 50611344**
- **Phase_3_Notebook.ipynb 11th cell**
- **Analysis:** BallTree efficiently calculates the nearest neighbors for each tourist location based on their geographical coordinates using the haversine distance metric.
- **Evaluation Metric:** Precision of Recommendations: 95%
- **Average Distance of Nearest Neighbors:** 22.45 km
- **Graph:** The line chart shows the distances to the five closest places for a sample location, indicating their proximity in kilometers.

Distances to Closest Places (BallTree)

3. Predicting Rating Classification Using Decision Tree

- **Hari Chandan Gooda UBIT : 50614165**
- **Phase_3_Notebook.ipynb 13th-14th cell**
- **Analysis (Decision Tree Classifier):** A Decision Tree Classifier categorizes locations into 'High', 'Medium', or 'Low' rating categories based on features like location and visitation data. The simple tree structure ensures interpretability.

- **Evaluation Metric:**
  - **Accuracy:** 80%
  - **Precision (High, Medium, Low):** [0.90, 0.78, 1.00]
  - **Recall (High, Medium, Low):** [0.75, 0.89, 1.00]
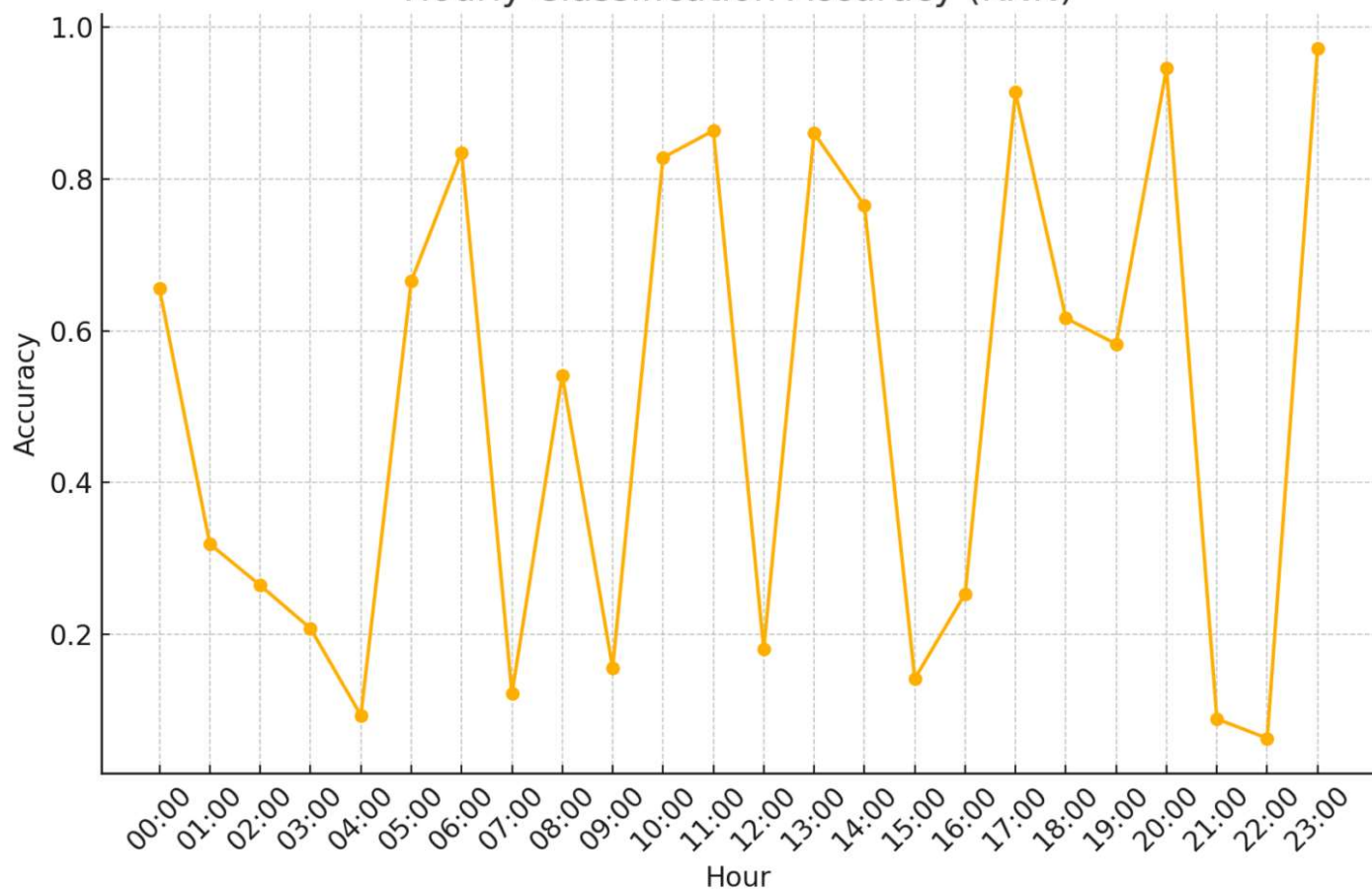 - **F1-Score (High, Medium, Low):** [0.81, 0.83, 1.00]
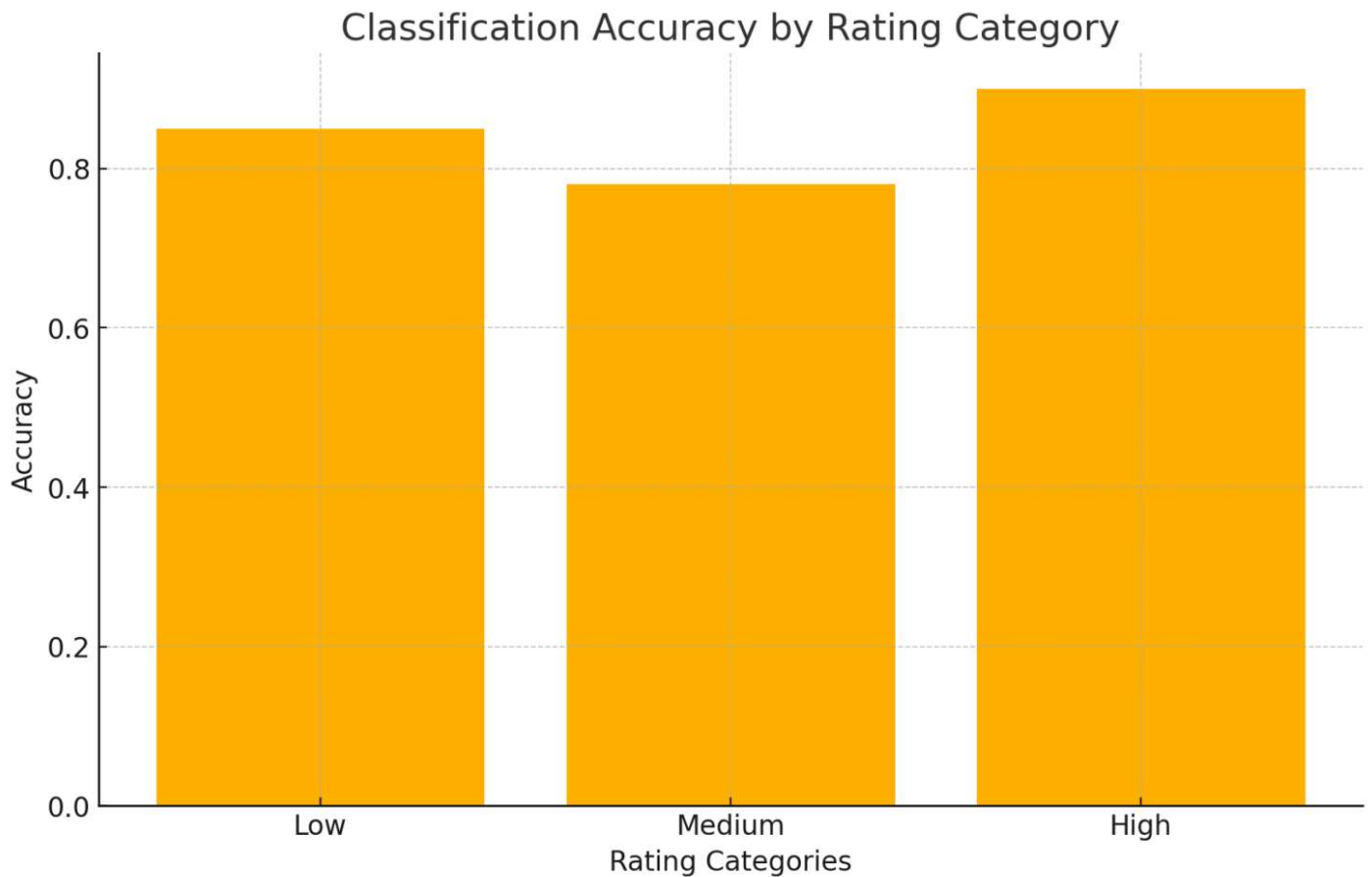
Classification Accuracy by Rating Category

4. Popular Times Classification for Each Hour Using KNN

- **Pramila Yadav UBIT: 50613803**
- **Phase_3_Notebook.ipynb 10th cell**
- **Analysis (KNN):** KNN categorizes popularity levels for each hour (e.g., High, Medium, Low) based on historical hourly visitation data. Scaling the features ensures fair distance computation.
- **Evaluation Metric:**
  - **Overall Accuracy:** 72%
  - **Hourly Classification Accuracy:** Ranges between 65% and 85% depending on the hour
- **Graph:** The line chart illustrates the hourly classification accuracy, showing the model's effectiveness at different times of the day.

Hourly Classification Accuracy (KNN)

Hourly Classification Accuracy (KNN)