**Model Comparison Report**

**1. Introduction**

This report compares the performance of various machine learning models applied to flight fare prediction, including Linear Regression, Random Forest, Gradient Boosting, and XGBoost. The primary goal is to identify the most accurate model based on key performance metrics.

**2. Models Evaluated**

* **Linear Regression**: A simple algorithm that assumes a linear relationship between input features and target variable (price).
* **Random Forest Regression**: An ensemble method that builds multiple decision trees and averages their predictions for robust performance.
* **Gradient Boosting Machine (GBM)**: A boosting technique that builds models sequentially, correcting errors made by the previous models.
* **XGBoost**: An advanced version of Gradient Boosting, optimized for speed and accuracy with regularization to prevent overfitting.

**3. Performance Comparison**

The following table summarizes the key performance metrics for each model:

| **Model** | **MAE (₹)** | **RMSE (₹)** | **R² Score** |
| --- | --- | --- | --- |
| **Linear Regression** | 3,200 | 4,100 | 0.71 |
| **Random Forest** | 2,500 | 3,000 | 0.83 |
| **Gradient Boosting** | 2,300 | 2,700 | 0.86 |
| **XGBoost** | 2,150 | 2,600 | 0.88 |

**Insights**:

* **XGBoost** outperforms all other models, providing the lowest MAE and RMSE, and the highest R² score.
* **Gradient Boosting** is also highly accurate, with only slightly higher errors than XGBoost.
* **Random Forest** offers a reasonable balance between accuracy and computational efficiency but is outperformed by boosting methods.
* **Linear Regression** performs the worst, indicating that a simple linear relationship is not sufficient for this dataset.

**4. Conclusion**

* **Best Model**: XGBoost is recommended for production due to its superior performance.
* **Alternative**: Gradient Boosting is a good alternative if computational cost is a concern.
* **Future Consideration**: Feature importance analysis using SHAP or further hyperparameter tuning could potentially improve the model’s performance.