**Flight Ticket Prediction Using Gradient Boosting Regressor Compared With Linear Regression**

**ABSTRACT:**

The purpose of this project is to predict airfare for ticket bookings using the Gradient Boosting. Regression device learning is the set of rules as opposed to a brand new linear regression. Materials and Methods: New Linear Regression Algorithm (with a sample size of ten) and Gradient Boosted Regression (with a sample size of ten). These algorithms are calculated on this picture using a total of 20 examples for the algorithm, and there are two firms that are used to calculate them. The size of the sample was determined to be 10, and it was compared with a group using a G Power value of 80%. Results: Values achieved in terms of accuracy are decided by Gradient Boosting regression (82.5%), as opposed to new Linear Regression (62.5%).8%. This is because Gradient Boosting regression is more accurate than new linear regression. In a test with one tail, the statistically significant difference between the new linear regression algorithm and the Gradient Boosting Regressor was found to be 0.00. This result was obtained With in Significance level of p 0.05. Conclusion: After going through all of the methods, it has been determined that the airfare forecast is more accurate than the brand new linear regression. This was revealed after going through all of the procedures.

**Existing Systems:**

In many existing flight ticket prediction systems, linear regression is commonly employed. Linear regression assumes a linear relationship between the input features (e.g., departure time, destination, airline) and the target variable (flight ticket price). It is a simple and interpretable model often used in initial modeling stages.

**Advantages:**

Simplicity: Linear regression is a straightforward and easy-to-understand algorithm.

Computational Efficiency: Training and predicting with linear regression are computationally less demanding compared to more complex algorithms.

Interpretability: The model provides a clear interpretation of the relationship between input features and the predicted ticket prices.

**Disadvantages:**

Linearity Assumption: Linear regression assumes a linear relationship between the input features and the target variable, which might not always reflect the true underlying patterns in flight ticket pricing.

Limited Complexity: Linear regression may struggle to capture intricate non-linear relationships in the data.

**Algorithms Used:**

Linear Regression is commonly used in existing systems for flight ticket prediction.

**Proposed Systems:**

The proposed system suggests using a Gradient Boosting Regressor for flight ticket price prediction. Gradient boosting is an ensemble learning technique that builds a series of weak learners (usually decision trees) and combines them to create a strong predictive model. It is capable of capturing complex non-linear relationships in the data, potentially leading to improved accuracy compared to linear regression.

**Advantages:**

High Predictive Accuracy: Gradient boosting can capture complex relationships in the data, making it potentially more accurate than linear regression.

Robust to Outliers: The ensemble nature of gradient boosting can handle outliers better than linear regression.

Feature Importance: Provides insights into the importance of different features in predicting flight ticket prices.

**Disadvantages:**

Computational Complexity: Training a gradient boosting model can be computationally expensive, especially when dealing with a large dataset or a high number of trees.

Hyperparameter Tuning: Finding the optimal set of hyperparameters for gradient boosting requires thorough experimentation and may be time-consuming.

Interpretability: Gradient boosting models are often considered less interpretable compared to linear regression due to their complex ensemble structure.

**Algorithms Used:**

Gradient Boosting Regressor is proposed for flight ticket prediction in the new system.

**SYSTEM SPECIFICATION:**

**HARDWARE REQUIREMENTS:**

* **System :** Intel i7
* **Hard Disk :** 1 TB.
* **Monitor** : 14’ Colour Monitor.
* **Mouse :** Optical Mouse.
* **Ram :** 8GB.

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 10.
* **Coding Language :** Python.
* **Front-End :** Html. CSS
* **Designing :** Html, CSS, javascript.
* **Data Base :** SQLite.

**REFERENCE:**

1. Zhao, Zhichao, Jinguo You, Guoyu Gan, Xiaowu Li, and Jiaman Ding. 2021. “Civil Airline Fare Prediction with a MultiAttribute Dual-Stage Attention Mechanism.” Applied Intelligence(Dordrecht, Netherlands), August, 1–16

2. Tavana, Madjid, Nadia Nedjah, and Reda Alhajj. 2020. Emerging Trends in Intelligent and Interactive Systems and Applications: Proceedings of the 5th International Conference on Intelligent, Interactive Systems and Applications (IISA2020). Springer

3. Harrington, Peter. 2012. Machine Learning in Action. Simon and Schuster.

4. Yu, Jiangni. 2021. “A New Way of Airline Traffic Prediction Based on GCN-LSTM.” Frontiers in Neurorobotics 15 (December): 661037

5. Rafique, Muhammad, Aleem Dad Khan Tareen, Adil Aslim Mir, Malik Sajjad Ahmed Nadeem, Khawaja M. Asim, and Kimberlee Jane Kearfott. 2020. “Delegated Regressor, A Robust Approach for Automated Anomaly Detection in the Soil Radon Time Series Data.” Scientific Reports 10 (1): 3004.