**SYSTEM ANALYSIS**

### 1. ****Existing System****

The existing system for flight delay management mainly relies on historical data analysis and conventional statistical models. Here's a breakdown of the limitations and methods currently used:

* **Traditional Statistical Methods**: In the current system, basic statistical methods such as **linear regression** or **time-series analysis** are often used to predict delays based on a limited set of features such as flight time and historical delays.
* **Manual Processes**: Airlines often rely on manual processes or basic scheduling software to manage delays, often reacting to delays rather than predicting and proactively mitigating them.
* **Limited Data Utilization**: Current systems may not fully utilize available data such as **weather conditions**, **air traffic data**, **airline-specific delays**, or **airport traffic**, which are crucial for predicting delays more accurately.
* **Low Predictive Accuracy**: The systems often lack the advanced machine learning capabilities needed to accurately predict flight delays based on complex interactions between multiple features such as weather, air traffic, and flight schedules.

#### **Limitations of the Existing System:**

* **Limited Feature Set**: Focuses only on a few key attributes, missing out on the potential impact of weather, traffic, and other operational variables.
* **Reactive rather than Predictive**: The system is more reactive in nature, dealing with delays after they occur rather than predicting and preventing them.
* **Low Accuracy**: Traditional methods tend to have lower accuracy when predicting delays, especially for more complex variables like weather or unforeseen operational issues.

**Algorithms in Existing: SVM, Logistic Regression, Random Forest.**

### 2. ****Proposed System****

The proposed system, as discussed in the document, aims to leverage advanced machine learning techniques to improve the accuracy of flight delay predictions. The system seeks to use comprehensive data to forecast delays proactively and improve decision-making for airlines, airports, and passengers. Here are the key components of the proposed system:

**Algorithms: ML: Liner Regression with SVM, GridSearchCV**

#### a. **Comprehensive Machine Learning Framework**

* **Machine Learning Algorithms**: The proposed system incorporates **machine learning techniques** like **Logistic Regression**, **Support Vector Machines (SVM)**, and **Random Forest**. These algorithms aim to improve the accuracy of delay predictions by learning from historical data.
* **Deep Learning Models**: The system suggests that **deep learning models**, such as **CNN-LSTM**, could be used to handle sequential data like flight schedules, weather patterns, and time-series variables, further boosting prediction accuracy.

#### b. **Data Enrichment and Feature Engineering**

* **Data Sources**: The proposed system plans to integrate data from multiple sources, including:
  + Flight schedules and history.
  + **Weather data** (e.g., severe weather conditions affecting flight times).
  + **Airport traffic** and congestion.
  + **Airline-specific delays** (e.g., delays specific to certain airlines).
* **Feature Engineering**: New features will be created from raw data to better capture patterns that may indicate delays. These could include **time of day**, **seasonality**, **airline performance**, and **weather impact** features.

#### **c.** **Predictive Model Enhancements**

* **Algorithm Comparison and Optimization**: By comparing models like **Random Forest**, **Logistic Regression**, and **SVM**, the system can select the best-performing model for delay prediction. The SVM model was found to achieve the highest accuracy in the document's analysis, making it an ideal candidate for deployment
* **Hyperparameter Tuning**: The proposed system will involve hyper parameter optimization (using techniques like grid search or random search) to fine-tune model performance.

#### d. **Real-Time Predictions**

* **Real-Time Forecasting**: The proposed system will allow for real-time prediction of flight delays, which could be accessed via a web interface or integrated into airline operations. By continuously updating the data and model, the system will adapt to changing conditions.
* **Deployment**: The proposed system can be deployed as a REST API, allowing external systems to fetch delay predictions based on current and real-time data.

#### e. **Evaluation Metrics**

* **Performance Metrics**: The proposed system will use more robust performance metrics beyond accuracy, including:
  + **Precision**, **Recall**, **F1-score**, and **ROC-AUC** scores.
  + **Confusion matrices** will be used to visualize model performance in terms of true/false positives and negatives.

#### f. **Deep Learning for Sequential Data (Future Scope)**

* The document proposes using **CNN-LSTM** deep learning models to capture the temporal dependencies in flight delay data. These models would be particularly useful in handling time-series data such as flight schedules, weather updates, and delay patterns over time

#### **Advantages of the Proposed System:**

* **Higher Accuracy**: Machine learning and deep learning models will lead to more accurate delay predictions compared to traditional statistical methods.
* **Real-Time Analysis**: The system can predict delays in real-time, allowing airlines to take preventive actions and improve operational efficiency.
* **Scalability**: The system can handle large datasets and complex features, making it more scalable for use across different airlines and airports.

### Summary:

* **Existing System**: Relies on traditional statistical methods with limited data utilization, leading to lower predictive accuracy and a reactive approach to managing delays.
* **Proposed System**: Implements machine learning and deep learning techniques for higher accuracy, real-time predictions, and proactive delay management, with potential for scalability and integration of external data sources (e.g., weather, traffic).