**SYSTEM ANALYSIS**

**EXISTING SYSTEM:**

Supervised automatic learning algorithms Support Vector Machine and the k-nearest neighbor to predict delays in the arrival of operated flights including the five busiest US airports. The precision achieved was very low with gradient booster as a classifier with a limited data set. Applied machine learning algorithms k-Nearest Neighbors to predict delays on individual flights. Flight schedule data and weather forecasts have been incorporated into the model. Sampling techniques were used to balance the data and it was observed that the accuracy of the classifier trained without sampling was more that of the trained classifier with sampling techniques.

**DISADVANTAGES OF EXISTING SYSTEM:**

* Non-parametric nature do not assume a particular functional form of the response under investigation data.
* The predictability may additionally range because of factors such as the number of origin destination pairs and the forecast horizon.
* The forecasts were based on some key attributes.
* **Algorithm**: Multiple Linear Regression, Support Vector Machine, k-nearest neighbor.

**PROPOSED SYSTEM:**

To predict flight delays to train models, we have collected data accumulated by the Bureau of Transportation, U.S. Statistics of all the domestic flights taken in 2015 was used. The US Bureau of Transport Statistics provides statistics of arrival and departure that includes actual departure time, scheduled departure time, and scheduled elapsed time, wheels-off time, departure delay and taxi-out time per airport. Cancellation and Rerouting by the airport and the airline with the date and time and flight labelling along with airline airborne time are also provided. The data set consists of 31 columns and 20277 and it can grow able by our implementation. By using pandas library we can fill the missing values which is essential for processing data for model.

**ADVANTAGES OF PROPOSED SYSTEM:**

* Supervised learning technique to gather the advantages of having the schedule and real arrival time.
* Algorithms are light computation cost will taken.
* We develop a system that predicts for a delay in flight departure based on certain parameters.

**Algorithm**: Logistic Regression, Decision Tree Regressor, Bayesian Ridge, Random Forest Regressor, Gradient Boosting Regressor