Flight Take Off Data Predication

The data set used for analysis contains data about flights leaving from JFK airport between one year from November 2019 to December 2020. It includes 28820 lines of individual flight information with 23 columns. Other researchers have used this data set to conduct delay prediction using the numerical value of the variable “TAXI\_OUT”. Make a prediction model for flight delays was conducted on the variable “DEP\_DELAY” by binary classification.

| **Attribute Name** | **Description** | **Type** |
| --- | --- | --- |
| **MONTH** | Month | Integer |
| **DAY\_OF\_MONTH** | Date of flight | Integer |
| **DAY\_OF\_WEEK** | Day of the week | Integer |
| **OP\_UNIQUE\_CARRIER** | Carrier code that represents the carrier company | Object |
| **TAIL\_NUM** | Air flight number | Object |
| **DEST** | Destination | Object |
| **DEP\_DELAY** | Departure delay of the flight | Integer |
| **CRS\_ELAPSED\_TIME** | Scheduled journey time of the flight | Integer |
| **DISTANCE** | Distance of the flight | Integer |
| **CRS\_DEP\_M** | Scheduled departure time | Integer |
| **DEP\_TIME\_M** | Actual departure time | Integer |
| **CRS\_ARR\_M** | Scheduled arrival time | Integer |
| **Temperature** | Temperature | Integer |
| **Dew Point** | Dew Point | Object |
| **Humidity** | Humidity | Integer |
| **Wind** | Wind direction | Object |
| **Wind Speed** | Wind speed | Integer |
| **Wind Gust** | Wind gust | Integer |
| **Pressure** | Pressure | Floating Point |
| **Condition** | Condition of the climate | Object |
| **sch\_dep** | Number of flights scheduled for departure | Integer |
| **sch\_arr** | Number of flights scheduled for arrival | Integer |
| **TAXI\_OUT** | Taxi-out time | Integer |