

SPRINT – 1 PROJECT DOCUMENT

Date	18 November 2022
Team ID	PNT2022TMID29701
Project Name	Flight Delay Prediction Using Machine Learning

DEVELOPMENT PHASE:

SPRINT-1:

Outline:

- Downloading the dataset
- Pre-processing the dataset
 - Importing libraries
 - Data analysing
 - Handling the missing values if any
 - Data visualisation – Scatterplot
 - Dropping unnecessary if any
 - Encoding – OnehotEncoder
 - Splitting the dataset –
Dependent and Independent

Required Libraries:

- Pandas - Data Pre-processing
- Numpy - Data Pre-processing,
Analysis
- Matplotlib - Visualization
- Seaborn - Visualization
- Isnull, drop - Balancing Data
- Sklearn - Model Building

Software/Tool:

- Anaconda- Jupyter Notebook
- Used Language Python

OVERVIEW:

Downloading the dataset:

Dataset is downloaded from Project workspace in Guided Project as flightdata.csv

Columns Description:

1. Dest - Destination Airport.
2. Crs_dep_time and crs_arr_time - planned departure and arrival time.
3. Crs_elapsed_time - estimated travel time as per plan.
4. Arr_time and dep_time - actual arrival and departure time.
5. Actual_elapsed_time - actual travelled time

Pre-processing the dataset:

Reference : <https://drive.google.com/file/d/1sRKoJ7fPIHOifPe4I-NBtIw4oDfhts0O/view?usp=sharing>