**PYTHON CODING**

**PR: EREQUISITES**

Installed Python Packages :

1. Scikit-learn

2. Num

3. Pandas

4. Matplotlib

5. Flask

# Import libraries

import pandas as pd

import pickle

from sklearn.preprocessing import LabelEncoder

from sklearn.ensemble import RandomForestClassifier

from sklearn.model\_selection import train\_test\_split # Import train\_test\_split function

# Import dataset

df = pd.read\_csv('Data/Processed\_data15.csv')

# Label Encoding

le\_carrier = LabelEncoder()

df['carrier'] = le\_carrier.fit\_transform(df['carrier'])

le\_dest = LabelEncoder()

df['dest'] = le\_dest.fit\_transform(df['dest'])

le\_origin = LabelEncoder()

df['origin'] = le\_origin.fit\_transform(df['origin'])

# Converting Pandas DataFrame into a Numpy array

X = df.iloc[:, 0:6].values # from column(years) to column(distance)

y = df['delayed'].values

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.25, random\_state=18) # 70% training and 30% test

# For 75% Train and 25% test use Random state =809

#Create a Random Forest Classifier

clf = RandomForestClassifier(random\_state=18)

clf.fit(X\_train, y\_train)

# Saving model to disk

pickle.dump(clf,open('model.pkl','wb'))

model = pickle.load(open('model.pkl','rb'))