EXPERIMENT 12**:** Development of Ensemble model for an Application

**Name:** D V Ravi Sridhar

**Reg No.:** RA1811003010732

**Section:** CSE – D2

# Date: 26/04/2021

**Problem Description :-**

Ensemble model using Random Forest for classification with 100 trees and split points chosen from a random selection of 3 features.

Samples of the training dataset are taken with replacement, but the trees are constructed in a way that reduces the correlation between individual classifiers. Specifically, rather than greedily choosing the best split point in the construction of the tree, only a random subset of features are considered for each split. The example below provides an example of Random Forest for classification with 100 trees and split points chosen from a random selection of 3 features.

***Code :-***

# Random Forest Classification import pandas

from sklearn import model\_selection

from sklearn.ensemble import RandomForestClassifier

url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians- diabetes.data.csv"

names = ['preg', 'plas', 'pres', 'skin', 'test', 'mass', 'pedi', 'age', 'class']

dataframe = pandas.read\_csv(url, names=names) array = dataframe.values

X = array[:,0:8]

Y = array[:,8] seed = 7

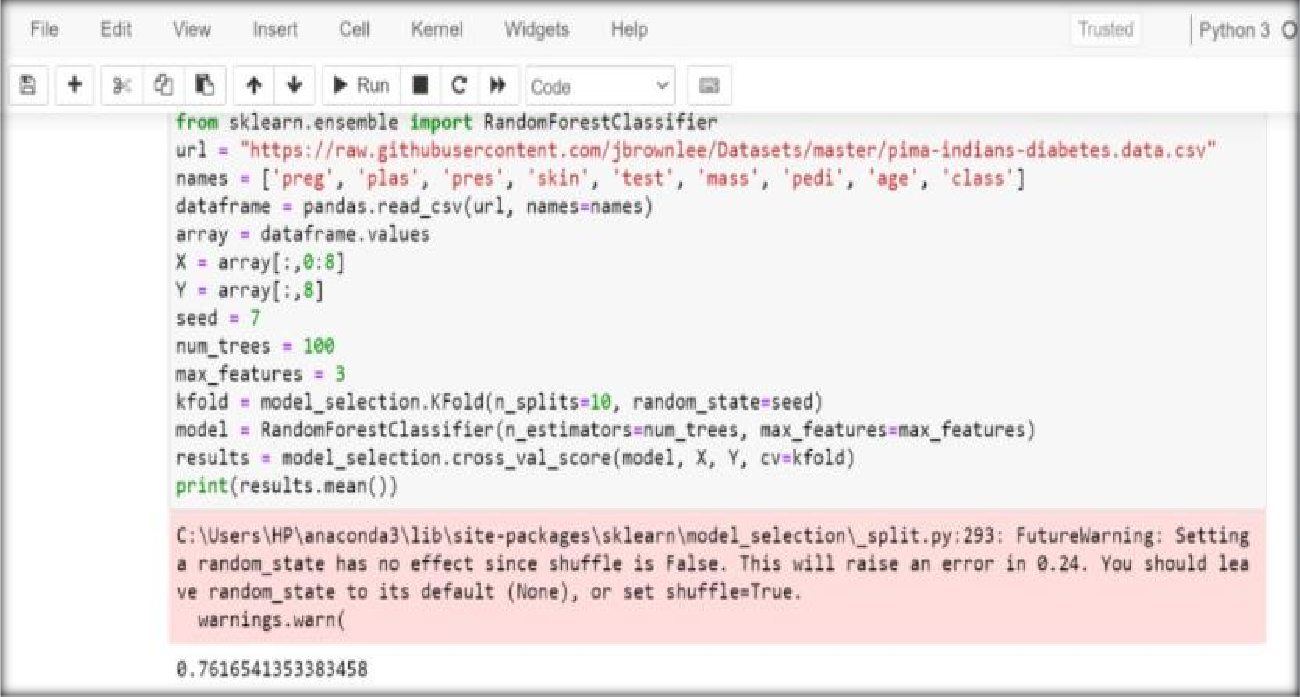
num\_trees = 100

max\_features = 3

kfold = model\_selection.KFold(n\_splits=10, random\_state=seed)

model = RandomForestClassifier(n\_estimators=num\_trees, max\_features=max\_features) results = model\_selection.cross\_val\_score(model, X, Y, cv=kfold) print(results.mean())

**Screenshot from Output :-**



**Result**

The experiment was successfully implemented and executed.