**INSTRUCTIONS**

**Before run a code**

**R**

**Install packages and load:**

* ggplot2
* corrgram
* DMwR
* caret
* randomForest
* unbalanced
* C50
* dummies
* e1071
* Information
* MASS
* rpart
* gbm
* ROSE
* sampling
* DataCombine
* inTreesfastDummies
* psych
* dplyr
* plyr
* data.table
* GGally
* tidyr

You can run the code on RStudio

**Before run a code**

**Python**

**Import packages:**

* import numpy as np
* import pandas as pd
* import matplotlib.pyplot as plt
* %matplotlib inline
* import seaborn as sns
* from matplotlib import pyplot
* sns.set()
* from sklearn.model\_selection import cross\_val\_score
* from sklearn.model\_selection import train\_test\_split
* from sklearn.ensemble import RandomForestClassifier
* from sklearn.metrics import accuracy\_score
* from sklearn.metrics import classification\_report
* from sklearn.preprocessing import StandardScaler
* from sklearn.pipeline import make\_pipeline
* from sklearn import svm
* from sklearn.preprocessing import scale
* from sklearn.model\_selection import GridSearchCV
* from sklearn.linear\_model import LogisticRegression
* from sklearn.metrics import precision\_recall\_curve
* from sklearn.metrics import auc
* from sklearn.metrics import roc\_curve
* from sklearn.metrics import roc\_auc\_score
* from sklearn.decomposition import PCA
* from sklearn.ensemble import GradientBoostingClassifier

You can run the code on Jupyter Notebook