Major Project

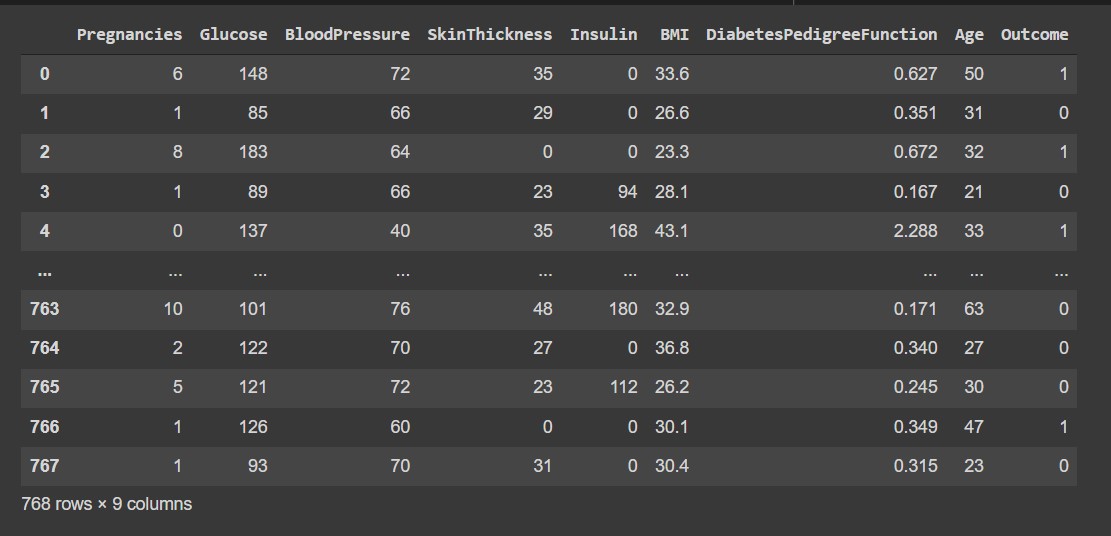
# All the patients are female

import pandas as pd

df=pd.read\_csv('https://raw.githubusercontent.com/2k0v11/Dataset/main/diabetes.csv.csv')

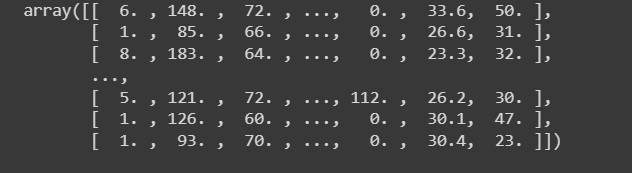
df

# Here the output will be outcome and input will be Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DPF, Age



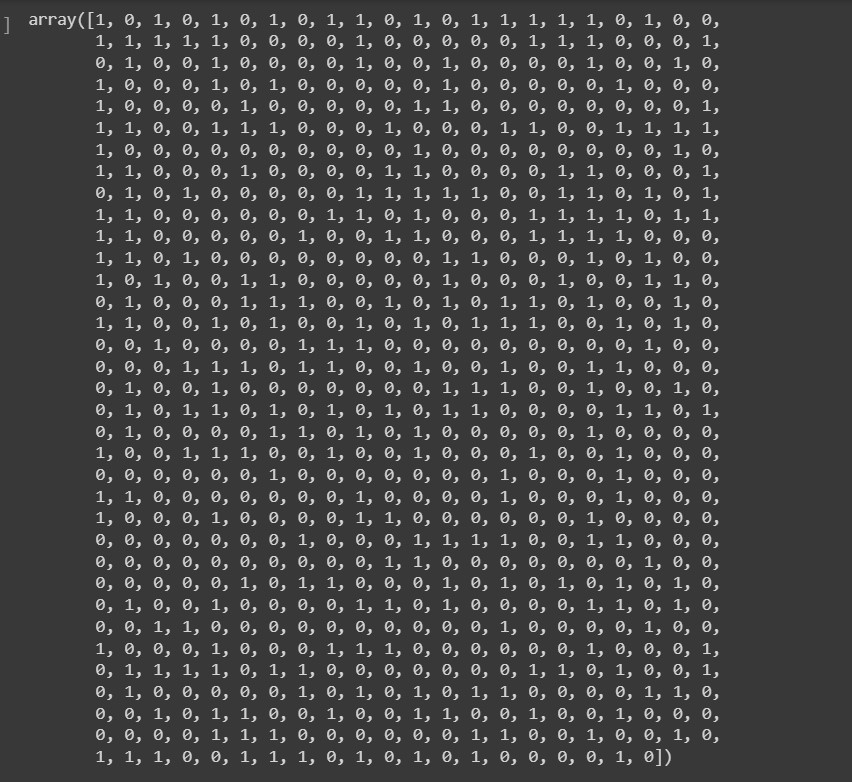
x=df.iloc[:,[0,1,2,3,4,5,7]].values

x



y=df.iloc[:,-1].values

y



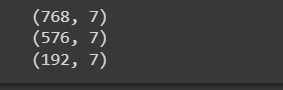
from sklearn.model\_selection import train\_test\_split

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,random\_state=0)

print(x.shape)

print(x\_train.shape)

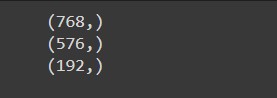
print(x\_test.shape)



print(y.shape)

print(y\_train.shape)

print(y\_test.shape)



from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()

x\_train = scaler.fit\_transform(x\_train)

x\_test = scaler.fit\_transform(x\_test)

from sklearn.linear\_model import LogisticRegression

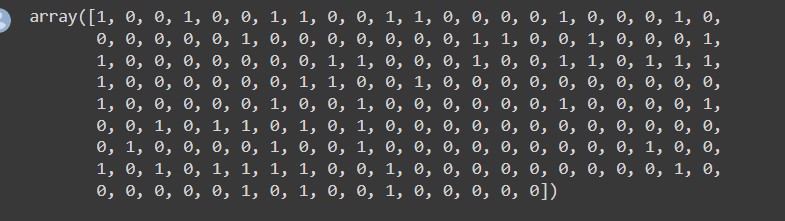
model=LogisticRegression()

model.fit(x\_train,y\_train)

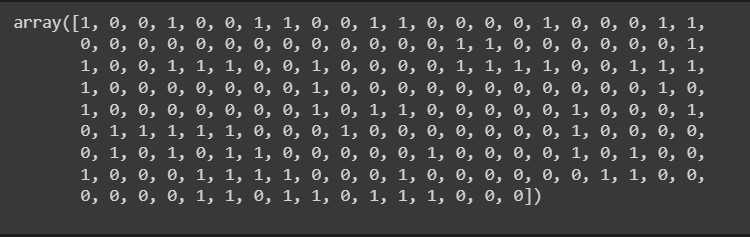


ypred=model.predict(x\_test)

ypred



y\_test



from sklearn.metrics import accuracy\_score

accuracy\_score(ypred,y\_test)\*100



model.predict([[1,200,180,24,200,41,19]]



* Link for the group project performed on google collab: <https://colab.research.google.com/drive/1fBNhddEdNcqb6BF52IQSGMe-dHMbYw0e?usp=sharing>

**Submitted By : Kavya Kartik (kavyakartik6@gmail.com)**