**LAB 07 : Supervised learning: Linear Regression and Naïve Bayes using SciKit.**

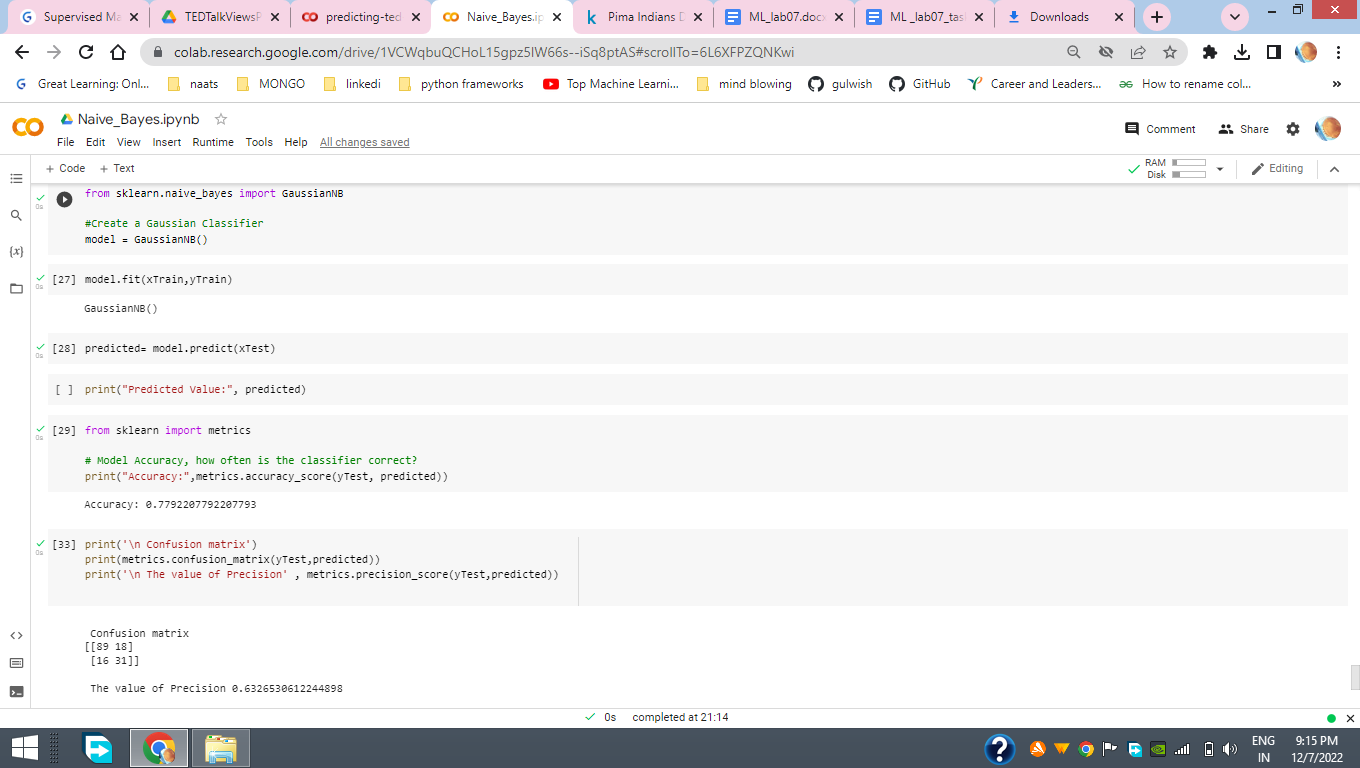
**LAB TASKS:**

**Perform both algorithms with any other dataset and try to analyze the outputs by changing the parameters.**

**NAIVE\_BAYES**

| **Data.columns**  **Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',**  **'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],**  **dtype='object')** |
| --- |

| **from sklearn.model\_selection import train\_test\_split**  **x= data.drop(['Outcome'],axis =1)**  **y=data['Outcome']**  **xTrain, xTest, yTrain, yTest = train\_test\_split(x, y, test\_size = 0.2, random\_state = 0)** |
| --- |



***LINEAR Regression***

| ***x=data\_df.drop(['Class'], axis = 1)#drop the target variable***  ***y=data\_df['Class']***  ***xtrain, xtest, ytrain, ytest = train\_test\_split(x, y, test\_size = 0.2, random\_state = 42)*** |
| --- |

| ***logisticreg = LogisticRegression()***  ***logisticreg.fit(xtrain, ytrain)*** |
| --- |

| ***y\_pred = logisticreg.predict(xtest)***  ***accuracy= logisticreg.score(xtest,ytest)***  ***print('Accuracy score of the Logistic regression model: ', accuracy\*100,'%')***  ***Accuracy score of the Logistic regression model: 99.85955549313577 %*** |
| --- |

| ***cm = metrics.confusion\_matrix(ytest, y\_pred)***  ***print(cm)***  ***[[56829 35]***  ***[ 45 53]]*** |
| --- |