**LOGISTIC RIGRESSION ASSIGNMENT**

**Que 2:-**

**Business Problem =** Output variable y -> Whether the client has subscribed a term deposit or not , Binomial ("yes" or "no")

Name of the File :- bank\_data.csv

Size of File :- 3139 kb

Data :- 45211 Observations and 32 Varaibles

**Exploratory data Analysis** =

* **Outliers: -**  Outliers are only in balance and age variable is present
* **Missing Value: -** Data don’t have Missing Values
* **Normality: -** Data are not normal
* **Output:** - Continues
* **Confusion matrix :**-

0 1

FALSE 39013 3587

TRUE 909 1702

1. **﻿**Measuring Percentage of the client has subscribed a term deposit in output columns

**Deposit** - 25%

**Not Deposit** – 85 %

**Model Building on Train Data =**

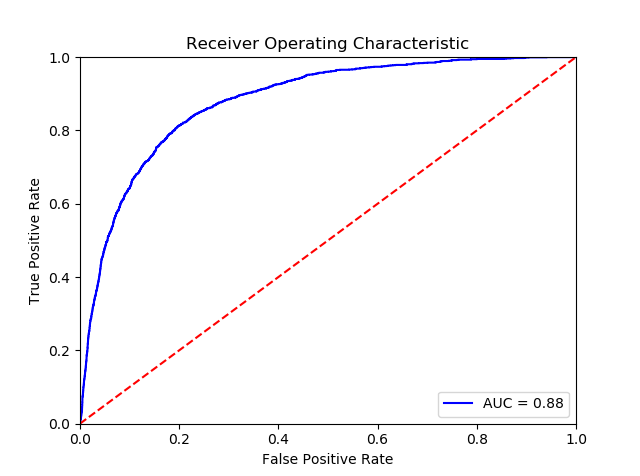
* **Summary: -**

# 

**AIC: -** 15660

* **﻿Accuracy :-** 89%
* **﻿Sensitivity :-** 97%
* **﻿Specificity: -** 32%
* **No observation: - 31647**
* **﻿Df Residuals:** - 31619

**Roc Curve**: -



**Roc Curve**: - 88%

**Python code file**: - Bank Data Assig-M9.py

**Packages used: -**

* ﻿pandas
* numpy
* scipy.stats
* matplotlib.pylab
* pylab
* statsmodels.formula.api
* seaborn as sns
* ﻿sklearn.model\_selection import train\_test\_split
* ﻿statsmodels.api
* Sklearn.metrics import confusion\_matrix,roc\_curve, roc\_auc\_score