**Features in variable x as below:**

'Age','DistanceFromHome', 'Education','EnvironmentSatisfaction', 'JobInvolvement', 'JobLevel', 'JobSatisfaction', 'MonthlyIncome', 'NumCompaniesWorked', 'PercentSalaryHike', 'PerformanceRating', 'RelationshipSatisfaction','TotalWorkingYears', 'WorkLifeBalance', 'YearsAtCompany', 'YearsInCurrentRole', 'YearsSinceLastPromotion', 'YearsWithCurrManager','BusinessTravel\_Travel\_Frequently', 'BusinessTravel\_Travel\_Rarely','EducationField\_Life Sciences', 'EducationField\_Marketing', 'EducationField\_Medical', 'EducationField\_Other', 'EducationField\_Technical Degree', 'Gender\_Male', 'JobRole\_Human Resources', 'JobRole\_Laboratory Technician', 'JobRole\_Manager', 'JobRole\_Manufacturing Director', 'JobRole\_Research Director', 'JobRole\_Research Scientist', 'JobRole\_Sales Executive', 'JobRole\_Sales Representative', 'MaritalStatus\_Married', 'MaritalStatus\_Single', 'OverTime\_Yes'

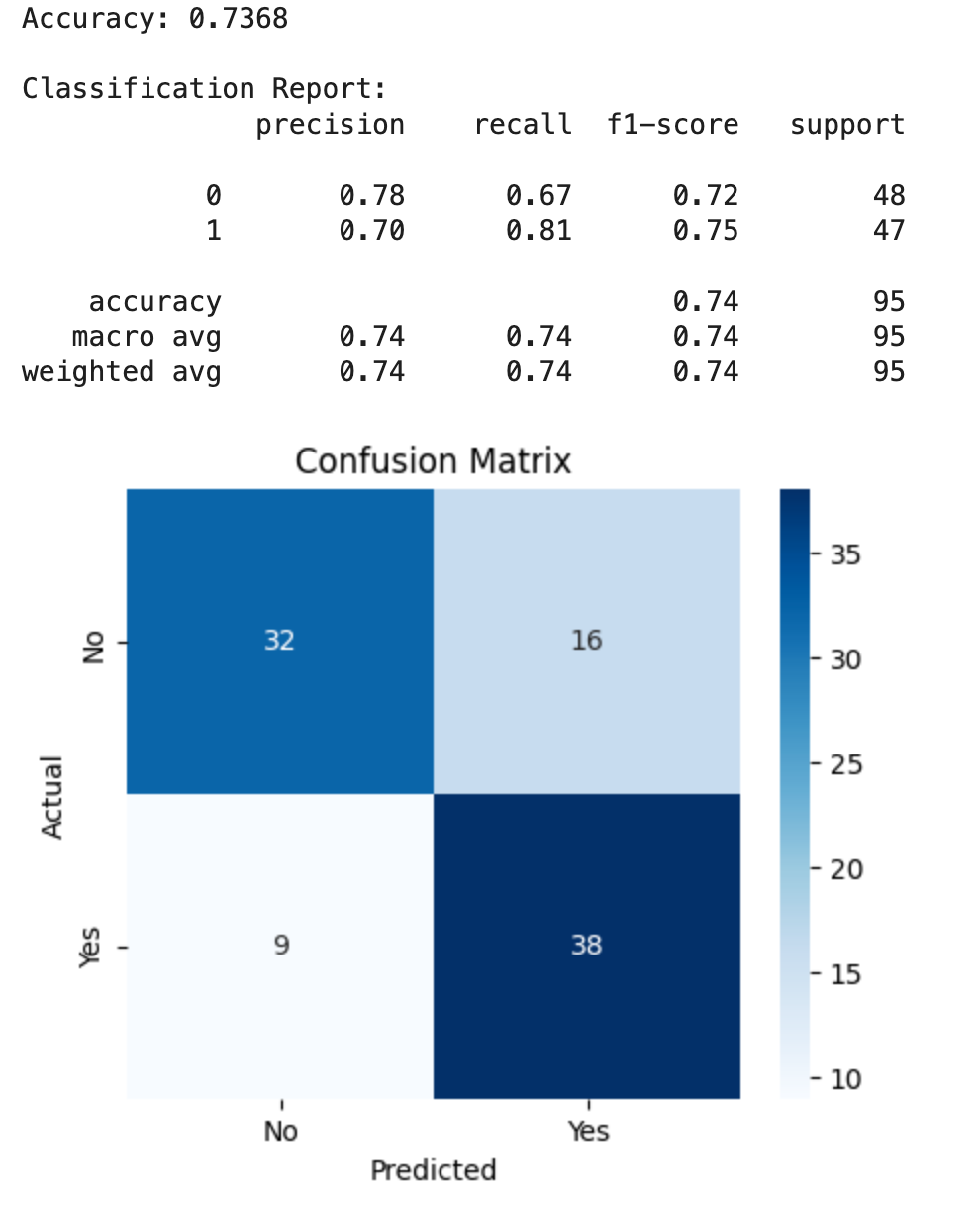
**Target Variable:**

Attrition

**APPROACH:**

* Cleaned data by looking for null, missing values and unknown values like Nan
* Identified the presence of any ordinal and categorical variables
* Performed one-hot encoding to aid the model for better data handling
* Performed Exploratory Data Analysis (EDA) to identify the correlations and key factors influencing the attrition rate (Target)
* Selected key features related to the target variable for model building
* Checked for data balance and under sampling technique was executed to improve the accuracy of the ML model
* Chose appropriate ML algorithm using lazy predict
* Built Logistic Regression Model and trained using training data set
* Evaluated the model using test data set
* Saved the model using pickle
* Loaded model in Streamlit and designed dashboard that enables user (HR) to enter the values of the features and predict the Attrition Rate
* Deployment steps and instructions to enter data and accuracy of the model has been shared below
* Inference and Suggestions to HR (by predicting Attrition using the ML model) has been shared in a separate file

**Model Evaluation:**



**Streamlit App Deployment Instructions:**

* Open a new notebook in VS code or in Google colab
* Upload the given log\_reg.pkl file in the same folder where the notebook is saved
* Run the cells to Install Streamlit, app.py and pyngrok in the notebook
* Run the ngrok authentication cell next
* Use or click the link given as Ngrok tunnel; Streamlit dashboard appears with the features
* Enter the values in the respective features as per the instruction given below and click Predict

**Instructions to enter values in the features:**

**Age:**

Enter the age in number of years

**DistanceFromHome:**

Enter the value of distance in Miles

**Education:**

Enter values between 1 to 4

The highest education level attained by the employee (e.g., "1" = Below College, "2" = College(UG), “3” = Masters (PG), “4” = Phd & Others.

**EnvironmentSatisfaction:**

Enter values between 1 to 4

Satisfaction with the work environment (e.g., "1" = Low, "2" = Medium, "3" = High, "4" = Very High)

**JobInvolvement:**

Enter the value between 1 to 3

The level of involvement the employee has in their job (e.g., "1" = Low, "2" = Medium, "3" = High

**JobLevel:**

Enter the value starting from 1

The job level of the employee within the company (e.g., "1" = Entry Level, "2" = Mid-Level, etc., based on hierarchy in the company)

**JobSatisfaction:**

Enter the value between 1 to 4

Satisfaction with the job (e.g., "1" = Low, "2" = Medium, "3" = High, "4" = Very High)

**MonthlyIncome:**

Enter the monthly salary of the employee in INR

**NumCompaniesWorked:**

Enter the value starting from 0

**PercentSalaryHike:**

Enter the value between 1 to 100 indicating the percentage of hike

**PerformanceRating:**

Enter the value between 1 to 4

The employee’s performance rating (e.g., "1" = Poor, "2" = Average, "3" = Good, “4” = Excellent)

**RelationshipSatisfaction:**

Enter the value between 1 to 4

Satisfaction with Manager at work (e.g., "1" = Low, "2" = Medium, "3" = High, "4" = Very High)

**TotalWorkingYears:**

Enter the value starting from 0

**WorkLifeBalance:**

Enter the value between 1 to 4

Employee’s work-life balance satisfaction (e.g., "1" = Low, "2" = Medium, "3" = High, "4" = Very High)

**YearsAtCompany:**

Enter the value starting from 0

**YearsInCurrentRole:**

Enter the value starting from 0

**YearsSinceLastPromotion:**

Enter the value starting from 0

**YearsWithCurrManager:**

Enter the value starting from 0

**BusinessTravel\_Travel\_Frequently:**

Select the value Yes or No

**BusinessTravel\_Travel\_Rarely:**

Select the value Yes or No

**EducationField\_Life Sciences:**

Select the value Yes or No

**EducationField\_Marketing:**

Select the value Yes or No

**EducationField\_Medical:**

Select the value Yes or No

**EducationField\_Other:**

Select the value Yes or No

**EducationField\_Technical Degree:**

Select the value Yes or No

**Gender\_Male:**

Select the value Yes or No

**JobRole\_Human Resources:**

Select the value Yes or No

**JobRole\_Laboratory Technician:**

Select the value Yes or No

**JobRole\_Manager:**

Select the value Yes or No

**JobRole\_Manufacturing Director:**

Select the value Yes or No

**JobRole\_Research Director:**

Select the value Yes or No

**JobRole\_Research Scientist:**

Select the value Yes or No

**JobRole\_Sales Executive:**

Select the value Yes or No

**JobRole\_Sales Representative:**

Select the value Yes or No

**MaritalStatus\_Married:**

Select the value Yes or No

**MaritalStatus\_Single:**

Select the value Yes or No

**OverTime\_Yes:**

Select the value Yes or No