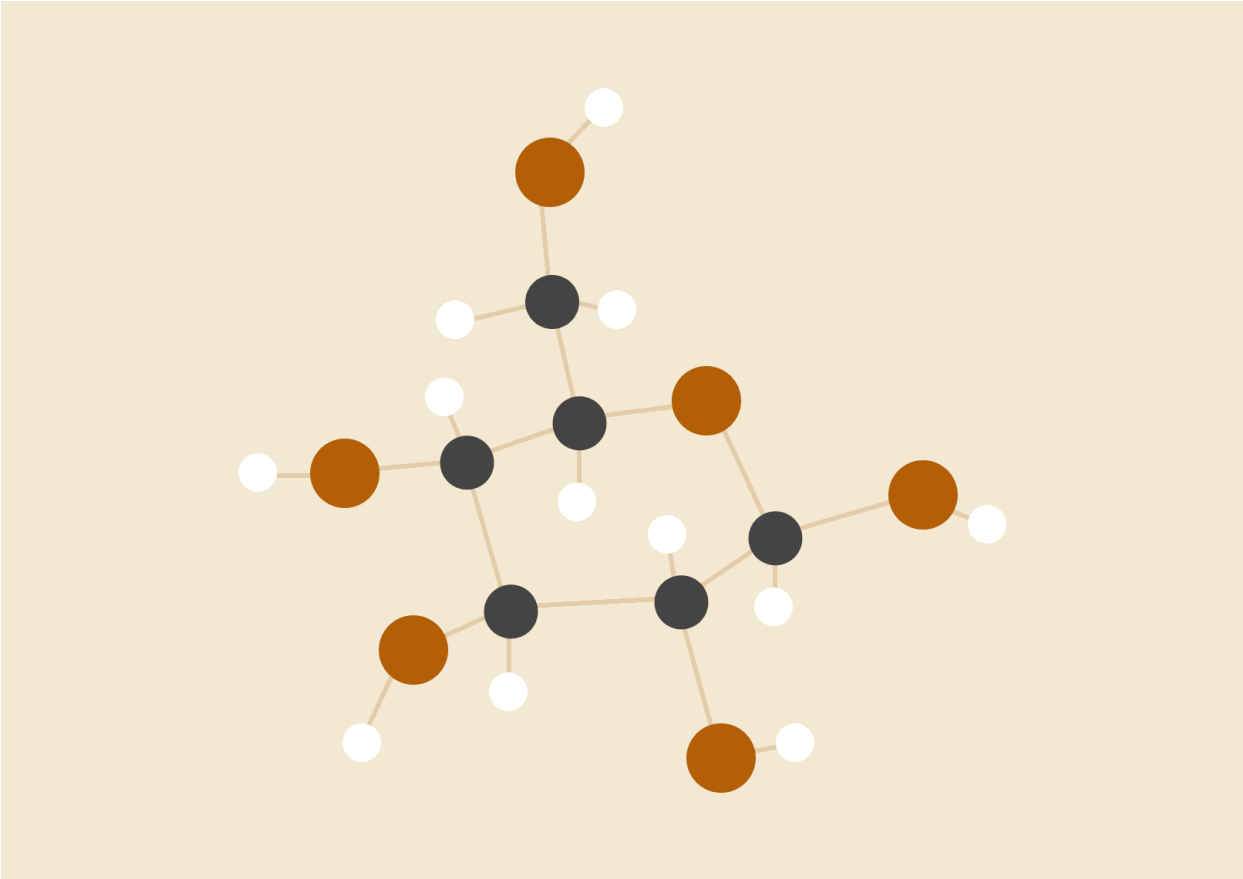
# Machine Learning Lab Report



## Team 7


| | |
|---|---|
| Rahibur Rahman | (211-115-004) |
| Ahad Bin Zaman Jibon | (211-115-009) |
| Sanjib Kumar Singha | (201-115-132) |
| Sree Himel Debonath | (191-115-053) |


## Topic

| Algorithm 1 | Algorithm 2 |
|---|---|
| **K_Nearest_Neighbour** | **Logistic _Regression** |

## INTRODUCTION

Introduction to Loan Prediction with KNN and Logistic Regression:

In the realm of financial decision-making, predicting loan approvals accurately is essential. Two powerful machine learning algorithms, K-Nearest Neighbors (KNN) and Logistic Regression, play a crucial role in this process.loan prediction can be applied to both Logistic Regression and K-Nearest Neighbors (KNN) algorithms. Both algorithms are commonly used for classification tasks, and they can be suitable for predicting whether a loan will be approved or denied based on various features such as credit score, income, and other relevant factors.

K-Nearest Neighbors (KNN):

KNN classifies loan applications based on their similarity to neighboring instances. Considering factors like credit score and income, KNN determines approval by the majority class of the k-nearest neighbors. Its flexibility suits complex decision boundaries.

Logistic Regression:

Logistic Regression assesses the relationship between independent variables and the probability of loan approval. It uses a logistic function to constrain output between 0 and 1, offering insights into the impact of each feature on loan decisions.

## MATERIALS

1. **PC,Internet connection**
2. **Jupyter Notebook to write code**
3. **Dataset (.CSV format)**
4. **language:Python (version 3)**

## PROCEDURE

1. We select Algorithm _1: K_Nearest_Neighbour
2. We select Dataset : **loan_data_set.csv**
3. Then we open jupyter notebook and give the Title name :

**Loan prediction using _K Nearest Neighbors**

4. Then we started Implementing code and training,testing and Predicting
5. Total Process from building model to predicting

    **Importing dependencies**

    **Reading dataset**

    **Setting the value for dependent and independent variables**

    **Checking for missing values**

    **Imputation of missing values**

    **Implementing the KNN Mode**

    **Feature scaling**

    **Fitting Mode**

    **Prediction on the test set**

    **Evaluating the Model**

    **Getting Graphical Representation By Implementing ROC_AUC curve**

    **Testing the knn model by giving inputs manually**

6. We select Algorithm _2: Logistic_Regression
7. We select Dataset : loan_data_set.csv
8. Then we open jupyter notebook and give the Title name :

    **Loan_prediction Using Logistic_Regression**

9. Total Process from building model to predicting

    **Importing dependencies**

    **Reading dataset**

    **Setting the value for dependent and independent variables**

    **Checking for missing values**

    **Imputation of missing values**

**Exploratory data analysis**

**Model Building**

**Converting some object data type to int**

**Setting the value for dependent and independent variables**

**Splitting the dataset into train and test set**

**Fitting the Logistic regression model**

**Prediction on the test set**

**Evaluating the Model**

**Getting Graphical Representation By Implementing ROC_AUC curve**

**Testing the logistic regression model by giving inputs manually**

### DATASET FEATURES

Our dataset has those below features :

**Loan_ID,Gender, Married, Dependents,Education,Self_Employed,ApplicantIncome,CoapplicantIncome**

**LoanAmount,Loan_Amount_Term,Credit_History,Property_Area,Loan_Status**

## RESULTS

By predicting the dataset using two different algorithms we got two difference accuracy results

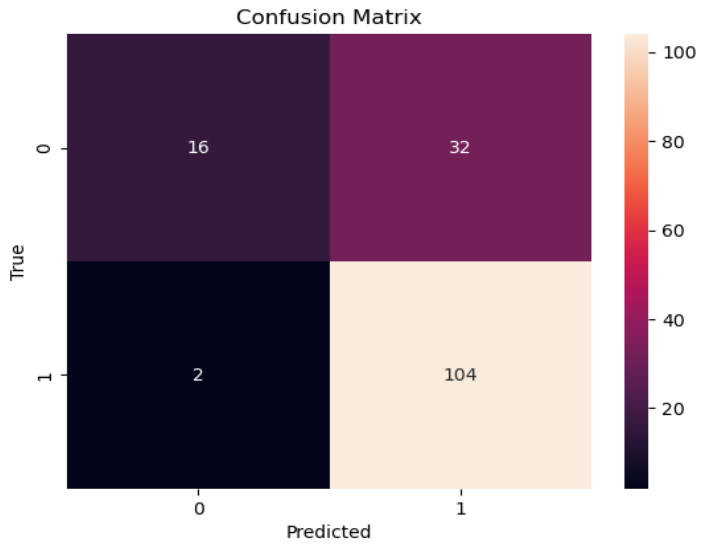
From **K_Nearest_Neighbour** we get the accuracy **74%** or **0.74**

And from **Logistic_Regression** we get the accuracy **78%** or **0.78**

After that we also tested the model by giving it manual input.

**Logistic_Regression**

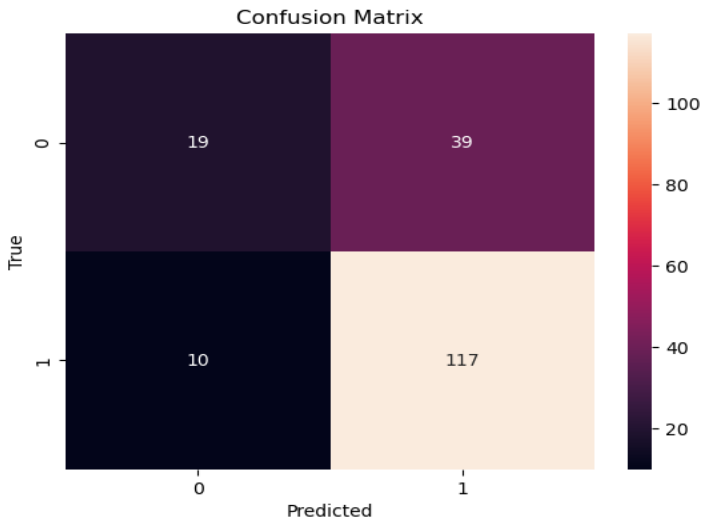
| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.89 | 0.33 | 0.48 | 48 |
| 1 | 0.76 | 0.98 | 0.86 | 106 |
| **accuracy** | | | **0.78** | 154 |
| macro avg | 0.83 | 0.66 | 0.67 | 154 |
| weighted avg | 0.80 | 0.78 | 0.74 | 154 |




Logistic_Regression

**K_Nearest_Neighbour**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.66 | 0.33 | 0.44 | 58 |
| 1 | 0.75 | 0.92 | 0.83 | 127 |
| accuracy | | | 0.74 | 185 |
| macro avg | 0.70 | 0.62 | 0.63 | 185 |
| weighted avg | 0.72 | 0.74 | 0.70 | 185 |

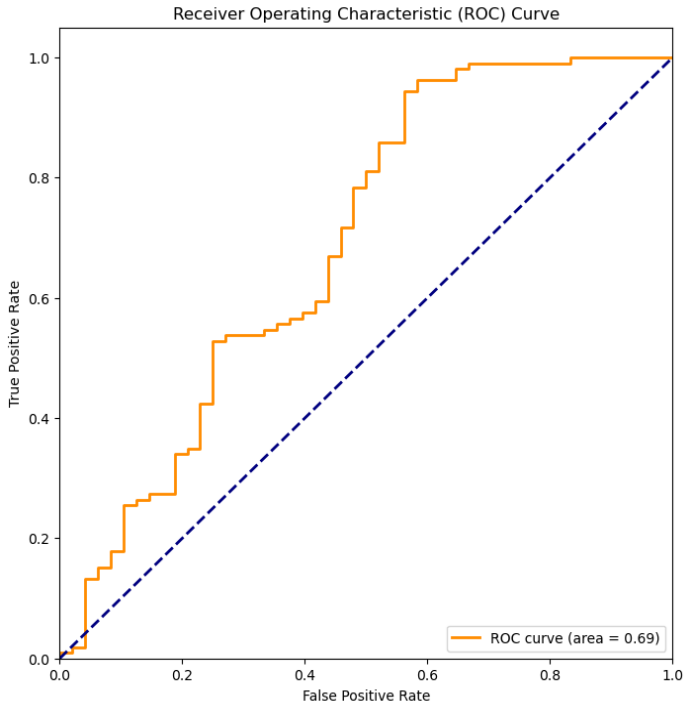



K_Nearest_Neighbour

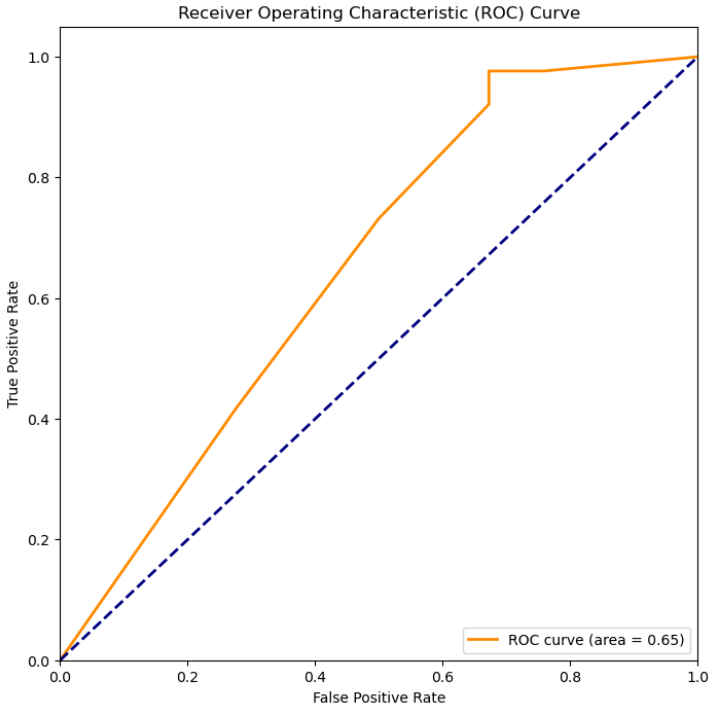
Above The confusion matrix we can see the total number of true value and predicted value for both algorithms where for Logistic_Regression TP(True positive)=16,TN(True Negative)=104,FP(False positive)=2,FN(False negative)=32,

For K_Nearest_neighbour TP=19,TN=117,FP=10,FN=39

**Getting Graphical Representation By Implementing ROC_AUC curve**




Logistic Regression




K_Nearest_Neighbour

## CONCLUSION

In conclusion, upon applying Logistic Regression and K-Nearest Neighbors (KNN) algorithms to the same dataset, it is evident that Logistic Regression outperforms KNN, yielding superior prediction accuracy.That means Logistic gives better results than KNN.

## REFERENCES


1. https://en.wikipedia.org/wiki/Confusion_matrix (Confusion Matrix)
2. https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/#:~:text=AUC%20ROC%20stands%20for%20%E2%80%9CArea,summary%20of%20the%20ROC%20curve. (Roc_Auc curve)
3. https://www.pythonkitchen.com/logistic-regression-vs-k-nearest-neighbors-in-machine-learning/ (logistic vs KNN)