Prediction of Loan Status at PT. ABC using Machine Learning Algorithms

NAME : SHINTA AMALIA

BATCH : RMT-034

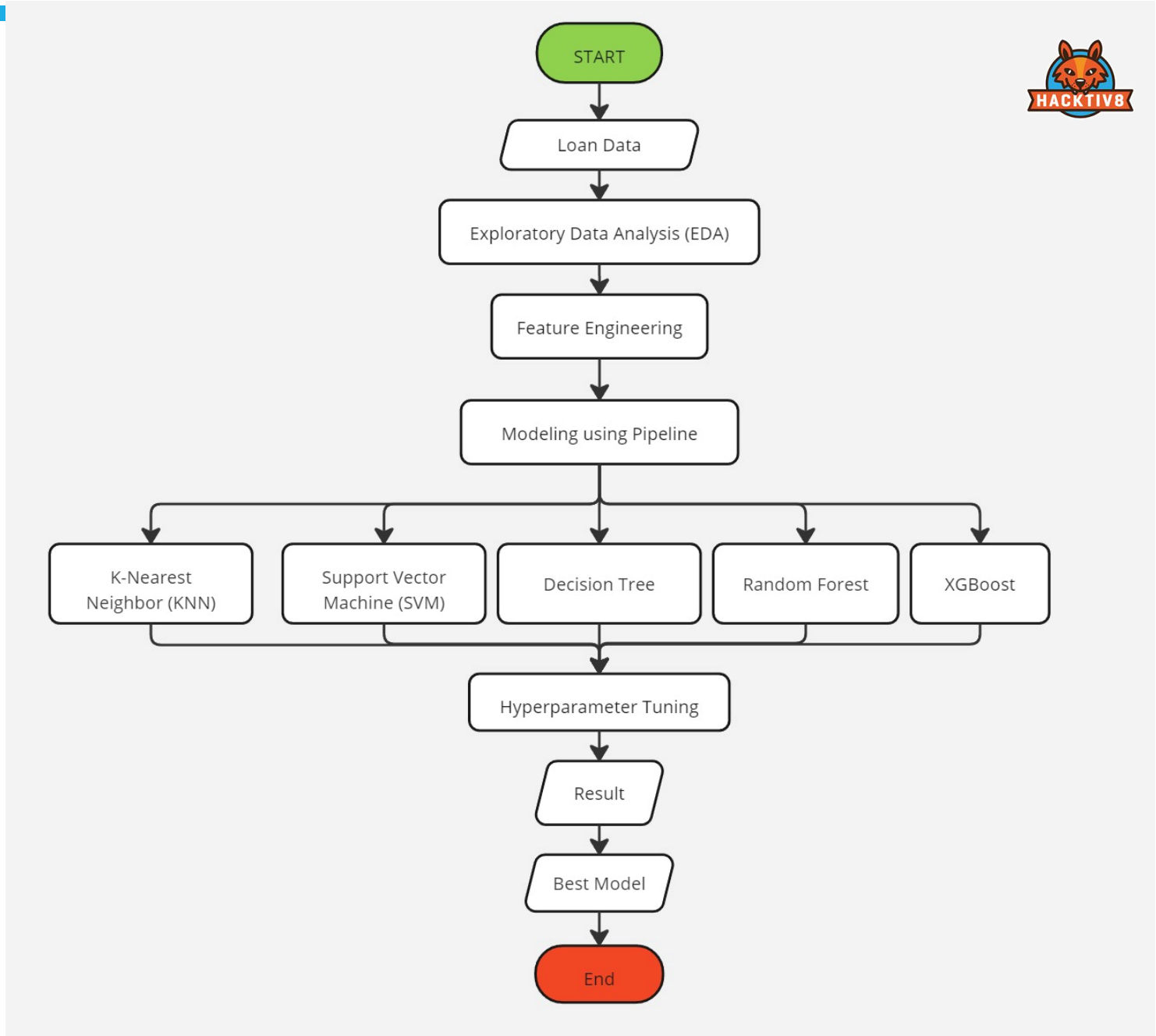
PROBLEM STATEMENT

I used to create a model to study patterns in predicting whether a loan application is approved or not. The model will also be saved so that it can be used to predict new data.

THE DATA

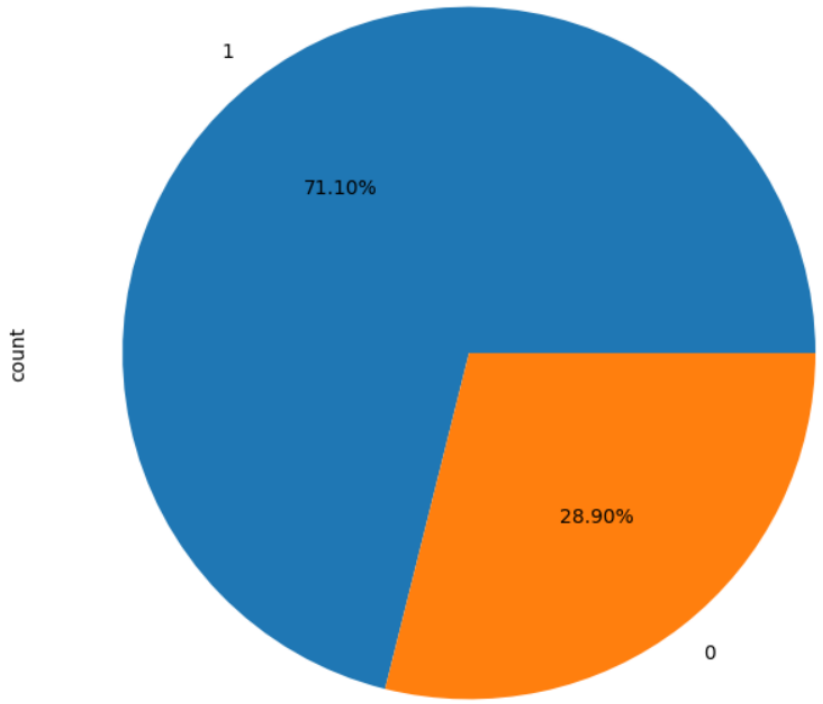
Column	Description
Loan_ID	A unique loan ID
Gender	Either male or female
Married	Married(yes) or Not Married(No)
Dependents	Number of persons depending on the client
Education	Applicant Education(Graduate or Not Graduate)
Self_Employed	Self-employed (Yes/No)
ApplicantIncome	Applicant income
CoapplicantIncome	Co-applicant income
LoanAmount	Loan amount in thousands
Loan_Amount_Term	Terms of the loan in months
Credit_History	Credit history meets guidelines
Property_Area	Applicants are living either Urban, Semi-Urban or Rural
Loan_Status	Loan approved (Approved "1" or Not Approved "0")

MODELING FLOWCHART



EXPLORATORY DATA ANALYSIS (EDA)

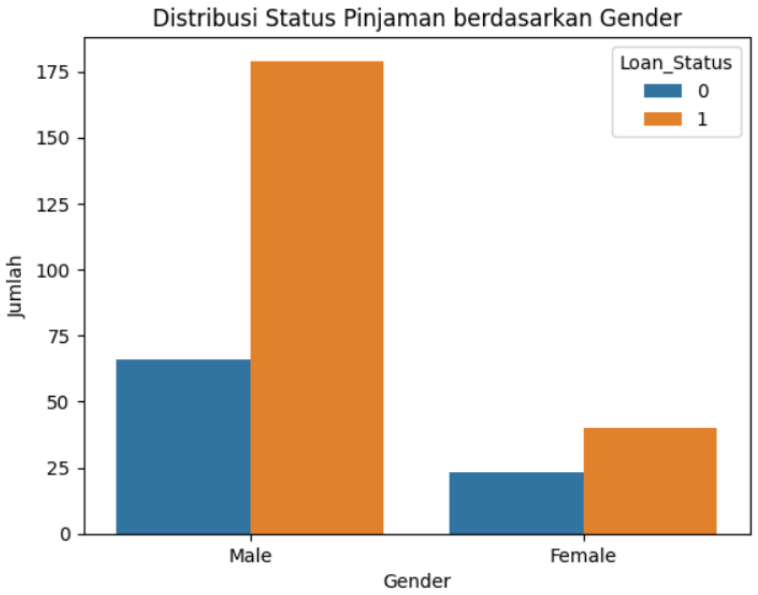
A. Target Variabel (Loan_Status)



Loan_Status	Count
1	219
0	89

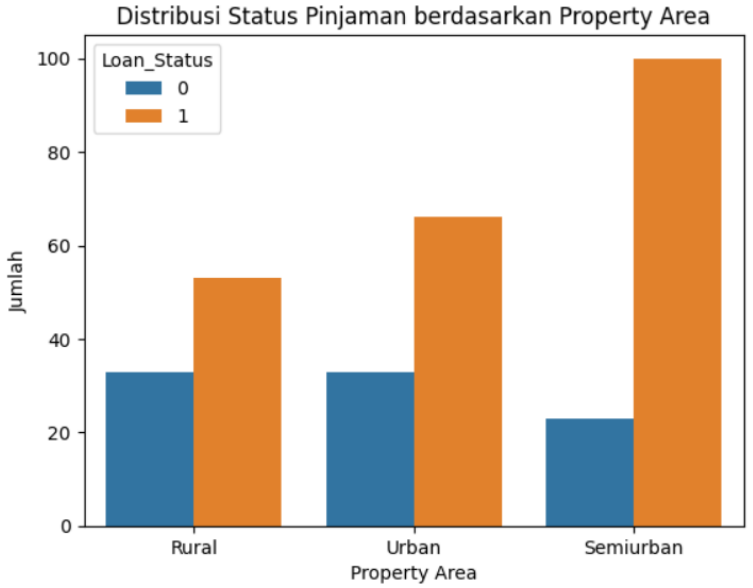
Based on the calculations and pie chart, it was found that the **Loan_Status "1" or approved is more prevalent** at 71.10% compared to "0" or not approved at 28.90%. The detailed numbers are 219 for 1 and 89 for 0.

B. Gender



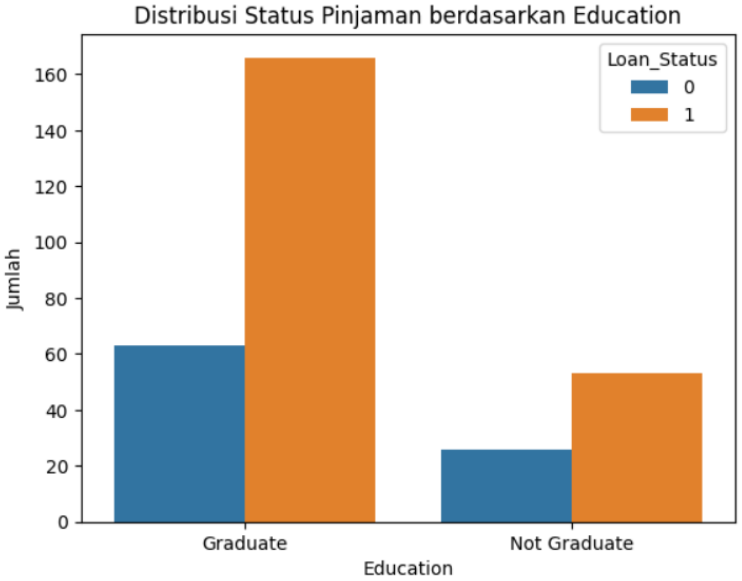
Loans are applied for more by males compared to females.

C. Property Area



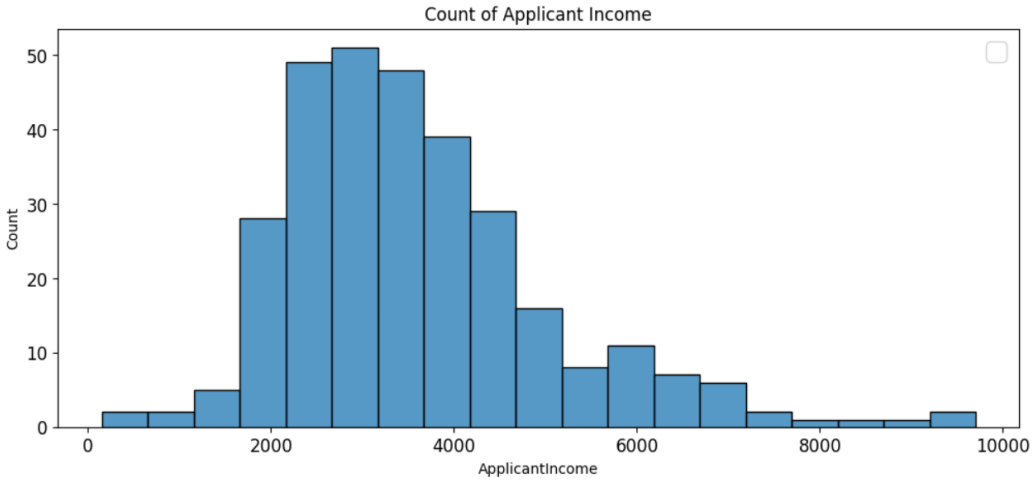
Loans are most frequently applied for in the Semi Urban area followed by Urban and Rural areas.

D. Education



Loans are mostly applied for by applicants with a Graduate.

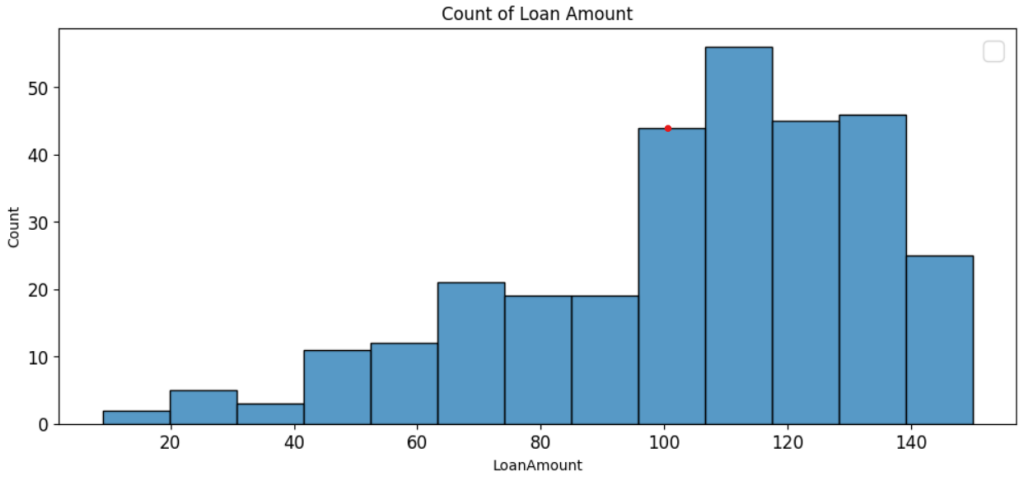
E. Applicant Income



Maximum : 9703 Thousand US Dollar
Minimum : 150 Thousand US Dollar
Average : 3599.13 Thousand US Dollar

Loan_Status	Max	Min	Avg
1	9703	645	3630.70
0	7660	150	3521.43

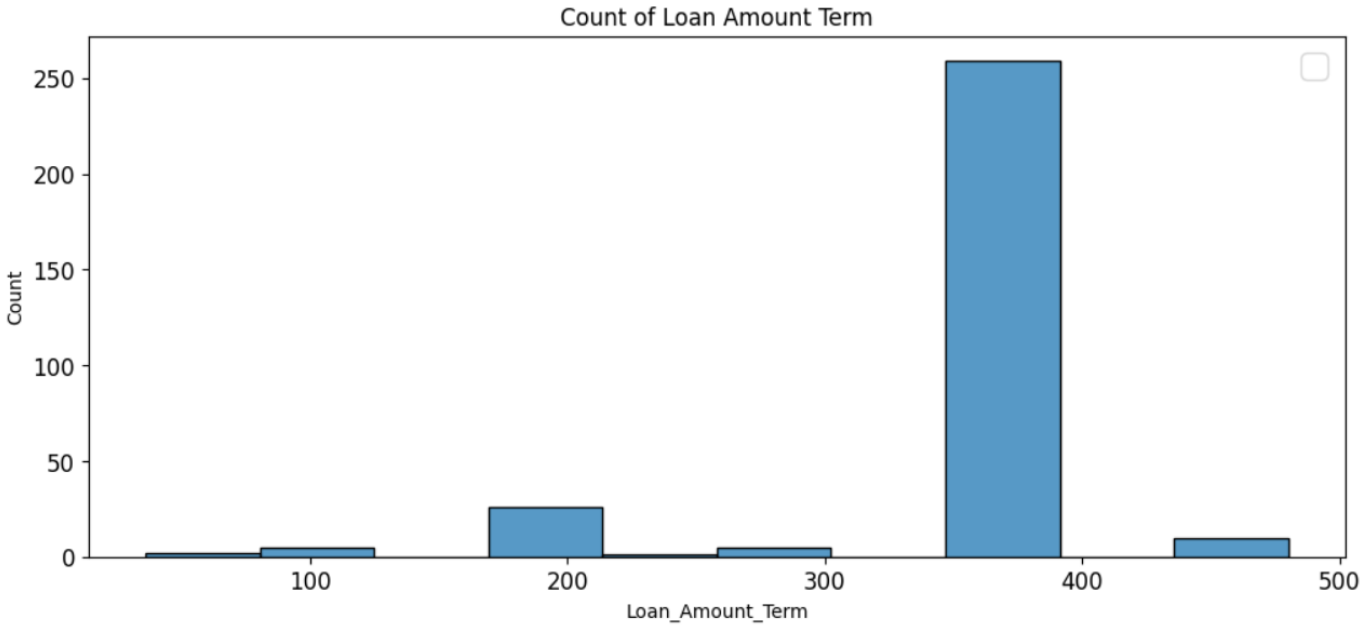
F. Loan Amount



Maximum: 150.0 Thousand US Dollar
Minimum: 9.0 Thousand US Dollar
Average : 104.62 Thousand US Dollar

Loan_Status	Max	Min	Avg
1	150	17	105.62
0	150	9	102.18

G. Loan Amount Term



Maximum : 480.0 Months
Minimum : 36.0 Months
Average : 341 Months

Loan_Status	Max	Min	Avg
1	480	60	341
0	480	36	342

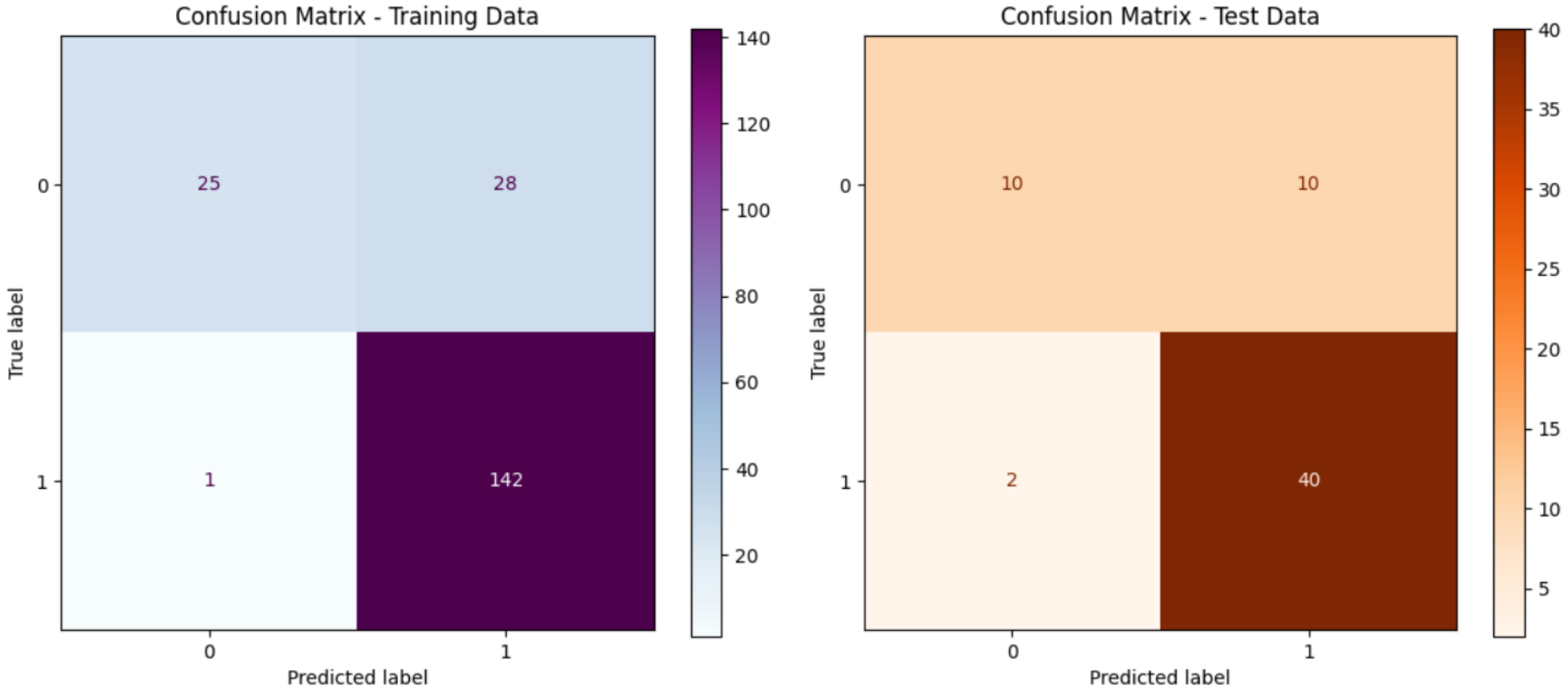
RESULT AND DISCUSSION

In the evaluation model, the resulting evaluation matrix is Accuracy Score, Precision Score, Recall Score, and F1 Score. However, due to the case study regarding the prediction of 'Loan_Status' whether it is approved or not, the **Precission Score** is the main evaluation matrix.

	Precision	KNN	SVM	Decision Tree	Random Forest	XGBoost
Without Hyperparameter Tuning	Train	0.83735	0.83529	1.00000	1.00000	1.00000
	Test	0.72222	0.80000	0.73913	0.81250	0.80000
	Mean (Cross Validation)	0.78186	0.83559	0.86298	0.85297	0.85518
	Std (Cross Validation)	0.04174	0.01928	0.03306	0.02888	0.03123
With Hyperparameter Tuning	Train	0.80814	0.83529	0.94483	0.90446	1.00000
	Test	0.76471	0.80000	0.76744	0.79167	0.80435
	Mean (Cross Validation)	0.80135	0.83559	0.87125	0.84598	0.85982
	Std (Cross Validation)	0.03064	0.01928	0.02862	0.01998	0.02544

Based on the comparison results, I chose **SVM as the best model** because the model is classified as good fit with the STD results at the lowest precision score among other models.

CONFUSSION MATRIX OF SVM MODEL



BUSINESS INSIGHT

- The tested SVM model can be immediately implemented into the company's loan approval system to support the decision-making process.
- The company needs to offer a variety of loan products with flexible amounts and terms to meet the diverse needs of customers.
- In addition to income, the company needs to consider other factors such as job stability, credit history, and the amount of assets in the credit assessment process.

CONCLUSION

To achieve optimal prediction results for loan status, whether approved or not, the SVM model is the best model to use. This is reflected in the precision score produced, which is 83% on the training set and 81% on the test set, with a difference of 3% < 5%, indicating that the resulting model is classified as a good fit. Thus, the tested SVM model can be immediately implemented into the company's loan approval system to support decision-making.

THANK YOU