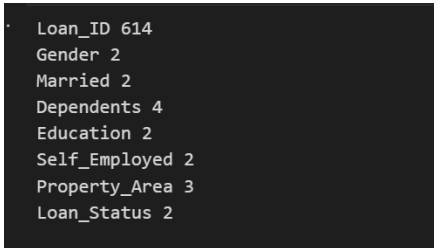
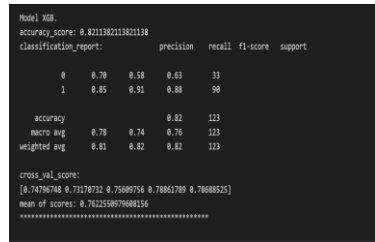


Project Development Phase Model Performance Test

Date	10 November 2022
Team ID	PNT2022TMID06904
Project Name	SMART LENDER - APPLICANT CREDIBILITY PREDICTION FOR LOAN APPROVAL
Maximum Marks	10 Marks

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S. No.	Parameter	Values	Screenshot
1.	Metrics	<p>Classification Model:</p> <p>Accuracy Score- Xgboost Model Testing Accuracy 0.905135135135135 Xgboost Model Training Accuracy 0.905135135135135</p> <p>Classification Report -</p> <pre> Model XGB. accuracy_score: 0.8211382113821138 classification_report: precision recall f1-score support 0 0.70 0.58 0.63 33 1 0.85 0.91 0.88 90 accuracy 0.82 123 macro avg 0.78 0.74 0.76 123 weighted avg 0.81 0.82 0.82 123 cross_val_score: [0.74796748 0.73170732 0.75609756 0.78861789 0.78688525] mean of scores: 0.7622550979608156 ***** </pre>	 <pre> Loan_ID 614 Gender 2 Married 2 Dependents 4 Education 2 Self_Employed 2 Property_Area 3 Loan_Status 2 </pre> <pre> ('XGB', XGBClassifier(learning_rate=0.1, n_estimators=100, max_depth=6, min_child_weight = 1, gamma=0., subsample=0.8, scale_pos_weight=1, random_state=27)), 1 </pre> <pre> for model_name, model in models: print('\nModel No.' + model_name) full_pipeline = Pipeline(steps=[(preprocessor, preprocessor), (model, model) </pre>  <pre> Model XGB. accuracy_score: 0.8211382113821138 classification_report: precision recall f1-score support 0 0.70 0.58 0.63 33 1 0.85 0.91 0.88 90 accuracy 0.82 123 macro avg 0.78 0.74 0.76 123 weighted avg 0.81 0.82 0.82 123 cross_val_score: [0.74796748 0.73170732 0.75609756 0.78861789 0.78688525] mean of scores: 0.7622550979608156 ***** </pre>

2.

Tune the Model

Hyperparameter Tuning

- No tuning is performed as we have got 91% accuracy

Parameters used-

n_estimators=5000,max_depth=80,max_features='log2'

Validation Method

- In-sample validation

```
df = pd.read_csv("C:\\Users\\mural\\OneDrive\\Desktop\\Surya_project-main\\Surya_project-main\\Loan_Approval.csv")
print(df.info())
df.head()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 614 entries, 0 to 613
Data columns (total 13 columns):
 #   Column                Non-Null Count  Dtype  ---
 0   Loan_ID               614 non-null    object
 1   Gender                601 non-null    object
```

Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term
LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	
LP001003	Male	Yes	1	Graduate	No	4593	1500.0	120.0	
LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	
LP001006	Male	Yes	0	Not Graduate	No	2583	2350.0	120.0	
LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	