Loan Recovery Prediction

Decision Tree Classification

Decision Tree Classification

Complexity penalty	Splits	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy
0.000	875	6670	1668	2084	0.889	0.902

Data Split

Train: 6670	Validation: 1668	Test: 2084	Total: 10422

Confusion Matrix

		Predicted	
		DEFAULT	NO DEFAULT
Observed	DEFAULT	1088	125
	NO DEFAULT	80	791

Model Performance Metrics

	DEFAULT	NO DEFAULT	Average / Total
Support	1213	871	2084
Accuracy	0.902	0.902	0.902
Precision (Positive Predictive Value)	0.932	0.864	0.903
Recall (True Positive Rate)	0.897	0.908	0.902
False Positive Rate	0.092	0.103	0.097
False Discovery Rate	0.068	0.136	0.102
F1 Score	0.914	0.885	0.902
Matthews Correlation Coefficient			NaN
Area Under Curve (AUC)	0.903	0.903	0.903
Negative Predictive Value	0.864	0.932	0.898
True Negative Rate	0.908	0.897	0.903
False Negative Rate	0.103	0.092	0.097
False Omission Rate	0.136	0.068	0.102
Threat Score	3.818	2.397	3.107
Statistical Parity	0.560	0.440	1.000

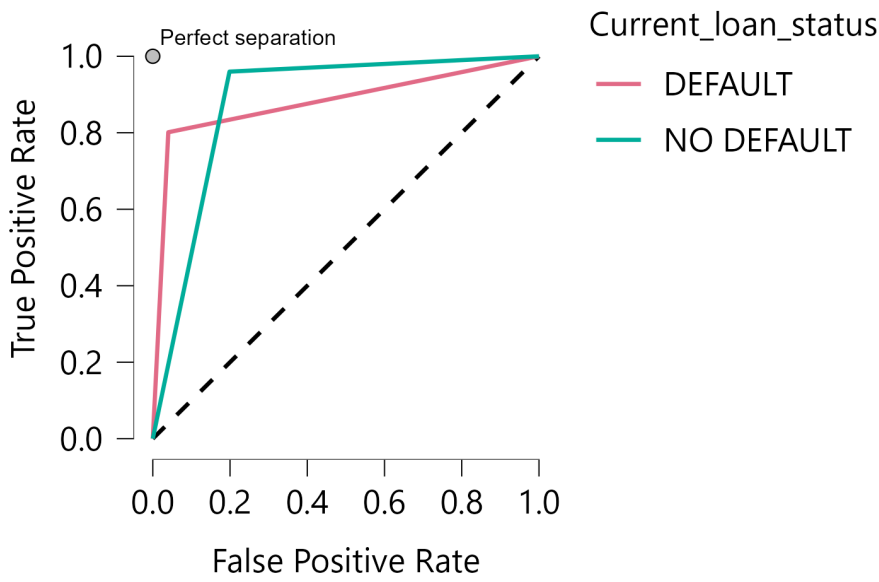*Note.* All metrics are calculated for every class against all other classes.
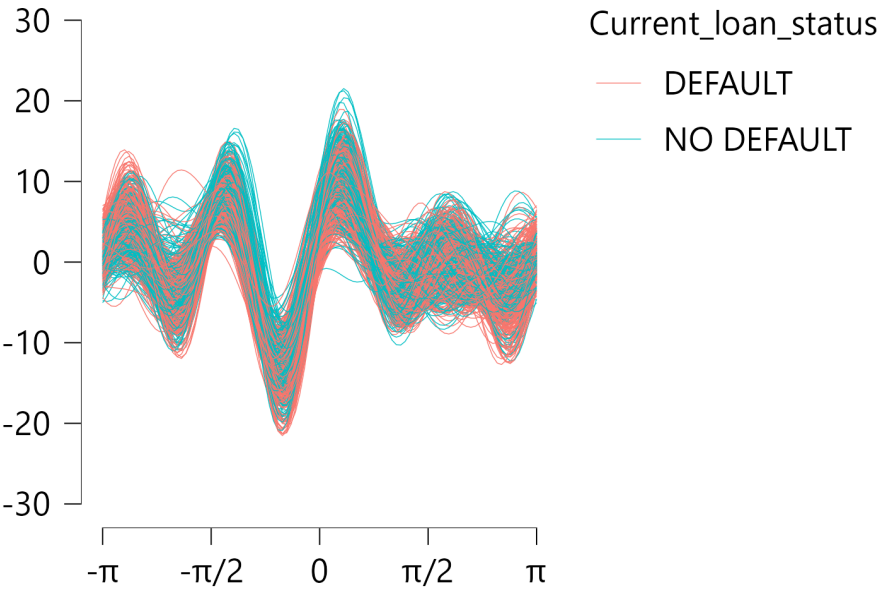
Feature Importance Metrics

	Relative Importance	Mean dropout loss
customer_age	18.780	0.093
cred_hist_length	18.285	0.180
historical_default	13.468	0.279
loan_int_rate	13.285	0.130
loan_grade	9.482	0.161
customer_income	7.558	0.157
home_ownership	4.870	0.119
loan_amnt	4.819	0.126
employment_duration	4.600	0.089
loan_intent	3.806	0.122
term_years	1.049	0.070

Note. Mean dropout loss is based on 50 permutations.

ROC Curves Plot

Andrews Curves Plot

Decision Tree Plot