

XAI Phase 2

Ziad Moutaz 202201252

April 2025

Contents

1	Models	3
1.1	Logistic Regression	3
1.1.1	Multi Colinearity Assumption	3
1.1.2	Results	3
1.1.3	LIME Interpretation	4
1.2	SVM	4
1.2.1	Results	5
1.2.2	LIME Interpretation	5
2	Conclusion	6

Models

1.1 Logistic Regression

A paper used the logistic regression model with different undersampling ratios. We used the highest achieved accuracy, which was with a 50%:50% ratio. Forward selection (Wrapper method) was used to select the top 10 effective features in the dataset.

1.1.1 Multi Colinearity Assumption

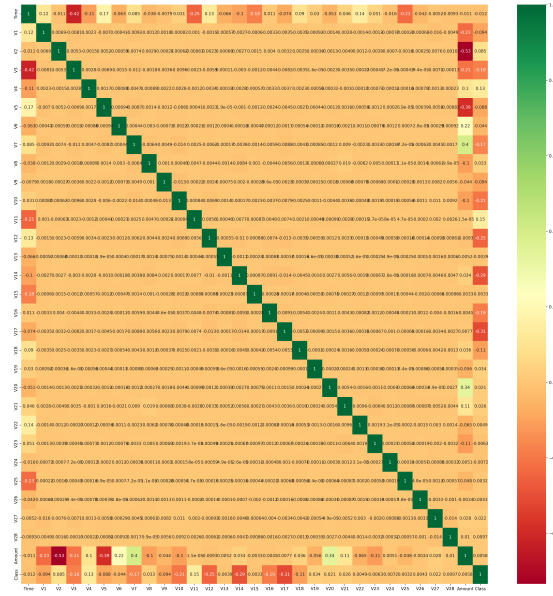


Figure 1: LIME Results for SVM Model

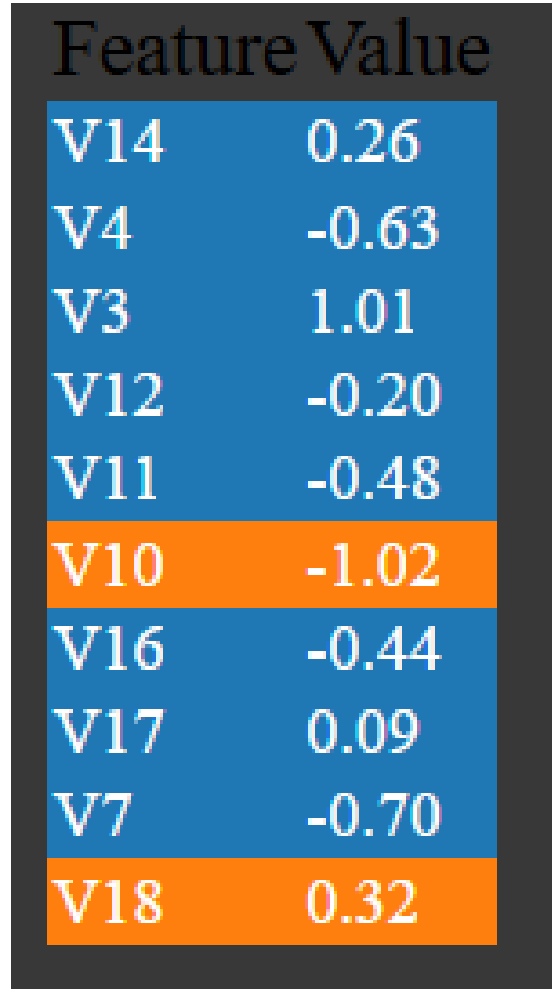
1.1.2 Results

Table 1: Logistic Regression Evaluation Metrics

Metric	Value
Accuracy	94.7%
Precision	0.98
Recall	0.91
F1-score	0.95
Sensitivity	0.91
AUC Score	0.95

1.1.3 LIME Interpretation

LIME (Local Interpretable Model-agnostic Explanations) provided an interpretation of the Logistic Regression model. LIME provided valuable information about how each input feature influences the prediction results. Model decisions become more trusted because of their explainable characteristics at a local level.



Feature Value	
V14	0.26
V4	-0.63
V3	1.01
V12	-0.20
V11	-0.48
V10	-1.02
V16	-0.44
V17	0.09
V7	-0.70
V18	0.32

Figure 2: LIME Results for Logistic Regression Model

1.2 SVM

The authors applied various models in their analysis but determined that Support Vector Machine with RBF kernel delivered most effective results. A soft margin value of C equal to 0.1 enabled the model to accept a few misclassifi-

cations while seeking an overall more generalized decision boundary. Forward selection conducted a Wrapper method selection of the top 10 effective features in the available dataset.

1.2.1 Results

Table 2: SVM Evaluation Metrics

Metric	Value
Accuracy	91.5%
Precision	1.0
Recall	0.84
F1-score	0.91
False Positive Rate (FPR)	1.0
AUC Score	0.92

1.2.2 LIME Interpretation

The SVM model received interpretation through application of the LIME approach. Local explanations reveal the SVM model decision boundary alongside showing how features actively contribute to prediction results thus making this model more understandable.

Feature Value	
V14	0.26
V4	-0.63
V3	1.01
V17	0.09
V12	-0.20
V11	-0.48
V16	-0.44
V18	0.32
V10	-1.02
V7	-0.70

Figure 3: LIME Results for SVM Model

2 Conclusion

This report introduces two predictive models composed of Logistic Regression along with Support Vector Machine (SVM), followed by LIME application for model interpretation. Research results demonstrated that SVM reached 91.5% while Logistic Regression showed 94.7% accuracy. Through the LIME technique we gained significant insights that helped understand how each model operates along with its important features in decision-making. Minimal transparency issues exist in essential domains like healthcare and finance when model verification becomes possible.