

LB/TH/33/2025
TH5920

**FRAUD DETECTION IN FINANCIAL
TRANSACTIONS USING LSTM AND XAI WITH
ONTOLOGICAL VALIDATION**

Sahani Nithya Rambukpitiya

228837G

Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa

Sri Lanka

July 2025

FRAUD DETECTION IN FINANCIAL TRANSACTIONS USING LSTM AND XAI WITH ONTOLOGICAL VALIDATION

Sahani Nithya Rambukpitiya

228837G

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa

Sri Lanka

July 2025

DECLARATION

I declare that this is my own work, and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

.....
Signature of the Student:

01/07/2025
.....
Date:

The candidate above has carried out research for the Master's thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof A.T.P Silva

.....
Signature of the Supervisor:

01/07/2025
.....
Date:

DEDICATION

This thesis is dedicated to my family and friends, whose unwavering support, encouragement, and belief in me made this journey possible. I also extend my heartfelt dedication to my supervisor, Professor A.T.P Silva, for their invaluable guidance, and insightful advice during this research and to all the lecturers of the Department of Computational Mathematics, Faculty of Information Technology, University of Moratuwa, for their continuous inspiration and support. Finally, I dedicate this work to everyone who continues to inspire me to pursue knowledge with passion and perseverance.

ACKNOWLEDGEMENT

I would like to sincerely thank my supervisor, Professor A.T.P Silva, for the constant guidance, encouragement, and insightful advice provided throughout this research. Their support was instrumental in helping me navigate the challenges and stay focused on my goals.

I am also truly grateful to all the lecturers of the Department of Computational Mathematics, Faculty of Information Technology, University of Moratuwa, for their valuable knowledge and support, which laid a strong foundation for this work.

My deepest appreciation goes to my family, whose endless love, patience, and belief in me have been the pillars that carried me through this journey. I am also thankful to my friends for standing by me, offering encouragement, and sharing in both the difficult and joyful moments.

Finally, I extend my gratitude to everyone who, in one way or another, contributed to the completion of this thesis. Your inspiration and kindness have made all the difference.

ABSTRACT

The escalation of digital financial services has increased the occurrence of credit card fraud, demanding the development of advanced, trustworthy fraud detection systems. Traditional and simple rule-based and classical machine learning techniques, while useful, often fail to detect complex, evolving fraud patterns and lack sufficient interpretability for high-stakes financial decision-making. This research proposes a novel, integrated approach combining Long Short-Term Memory networks, Explainable AI techniques, and ontology-based semantic validation to address these challenges. The Long Short-Term Memory model captures sequential transaction behaviors effectively, identifying anomalies indicative of fraud. To overcome the inherent black-box nature of deep learning models, Local Interpretable Model-agnostic Explanations is employed, providing transaction-level interpretability and enabling stakeholders to understand the factors behind each prediction. Further, an ontology is developed to embed domain-specific knowledge, offering semantic validation of the model outputs. This ensures that predictions not only rely on learned patterns but also align with predefined risk rules and expert knowledge. Experimental results demonstrate high accuracy and recall rates, confirming the model's effectiveness in detecting fraudulent activities, while the ontology layer enhances trust and reliability. This hybrid framework thus advances fraud detection by combining predictive power, explainability, and semantic validation, contributing to the development of more secure and transparent financial systems.

Keywords: Credit Card Fraud Detection, Long Short-Term Memory Network, Explainable AI, Ontology-Based Validation, Interpretability

TABLE OF CONTENTS

DECLARATION.....	i
DEDICATION	ii
ACKNOWLEDGEMENT	iii
ABSTRACT.....	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	ix
LIST OF TABLES	x
LIST OF ABBREVIATIONS.....	xi
CHAPTER 1	1
INTRODUCTION	1
1.1 Prolegomena	1
1.2 Background and Motivations	3
1.3 Problem in Brief.....	5
1.4 Objectives	5
1.5 Hypothesis.....	6
1.6 A Novel Approach to Credit Card Fraud Detection using LSTM and XAI With Ontological Validation.....	6
1.7 System Requirements.....	7
1.8 Structure of the thesis.....	7
1.9 Summary	8
CHAPTER 2	9
DEVELOPMENTS IN CREDIT CARD FRAUD DETECTION RESEARCH.....	9
2.1 Introduction.....	9
2.2 Gestation of the Area of Credit Card Fraud Detection	9
2.2.1 Early Landscape of Credit Card Fraud.....	9
2.2.2 Manual and Rule-based Approaches.....	10
2.2.3 Limitations of Early Approaches	10
2.3 Evolution of the Area of Credit Card Fraud Detection	10
2.3.1 Shift to Machine Learning	11
2.3.2 Rise of Deep Learning Models	13
2.3.3 Challenges Faced in Model Performance	14
2.4 Latest Development and Future Trends	15

2.4.1 Explainable AI (XAI).....	15
2.4.2 Integration of Knowledge-based Systems.....	17
2.4.3 Trend Toward Responsible AI.....	18
2.5 Research Findings.....	19
2.6 Research Challenges in the Area of Credit Card Fraud Detection.....	21
2.7 Problem Definition.....	23
2.8 Definition of Objectives.....	24
2.9 Summary.....	24
CHAPTER 3	25
TECHNOLOGY	25
3.1 Introduction.....	25
3.2 Long Short-Term Memory Framework (LSTM).....	25
3.2.1 Data Input and Preprocessing	25
3.2.2 LSTM Model Architecture.....	27
3.2.3 Model Training and Evaluation	31
3.3 Explainable AI (XAI)	32
3.3.1 LIME.....	32
3.3.2 Integration with LSTM Predictions	34
3.3.3 Visualization and Interpretation.....	35
3.4 Ontological Validation Framework	35
3.4.1 Ontology Overview.....	35
3.4.2 Ontology Modeling.....	36
3.4.3 Reasoning and Rule Layer	36
3.4.4 Integration with Model Output	36
3.5 Technology Stack.....	37
3.6 Summary.....	38
CHAPTER 4	39
PROPOSED HYBRID APPROACH.....	39
4.1 Introduction.....	39
4.2 Hypothesis.....	39
4.3 Input.....	39
4.4 Output	39
4.5 Process	40
4.5.1 Data Acquisition and Preprocessing Module.....	40

4.5.2 Modeling with LSTM	40
4.5.3 Explainable AI Module.....	41
4.5.4 Ontological Validation.....	41
4.5.5 System Flow and Integration	41
4.6 Features	42
4.7 Users	42
4.8 Summary	43
CHAPTER 5	44
DESIGN.....	44
5.1 Introduction.....	44
5.2 Top-Level Architecture.....	44
5.3 Modules.....	45
5.3.1 Data Acquisition and Preprocessing Module.....	45
5.3.2 LSTM-based Fraud Detection Module	47
5.3.3 Explainable AI Module.....	48
5.3.4 Ontology-Based Validation Module	49
5.3.5 Decision Making and Reporting Module.....	50
5.4 Connectors	51
5.5 Summary	51
CHAPTER 6	52
IMPLEMENTATION.....	52
6.1 Introduction.....	52
6.2 Implementation of Modules	52
6.2.1 Data Acquisition and Preprocessing Module.....	52
6.2.2 LSTM-based Fraud Detection Module	54
6.2.3 Explainable AI Module.....	56
6.2.4 Ontology-Based Validation Module	57
6.2.5 Decision Making and Reporting Module.....	60
6.2.6 Interactive Web Application using Streamlit.....	61
6.3 Tools and Technologies	62
6.3.1 Python	62
6.3.2 TensorFlow and Keras	62
6.3.3 Pandas and NumPy	62
6.3.4 LSTM.....	63

6.3.5 LIME.....	63
6.3.6 Matplotlib and Seaborn.....	63
6.3.7 Neo4j.....	63
6.3.8 Cypher Query Language	63
6.3.9 Ontology and Rule-Based Risk Tags	63
6.3.10 Jupyter Notebook	63
6.3.11 OpenPyXL	63
6.3.12 Streamlit.....	64
6.4 Summary	64
CHAPTER 7	65
EVALUATION	65
7.1 Introduction.....	65
7.2 LSTM Evaluation.....	65
7.3 Ontology Evaluation	68
7.4 Evaluation Using Realistic Test Cases with Ontology-Driven Validation	69
7.4.1 Test Case 1: High-Risk Merchant with Blacklist Tag	69
7.4.2 Test Case 2: High Spending and Blacklisted Merchant.....	70
7.4.3 Test Case 3: Legitimate Transaction with No Risk Indicators	71
7.4.4 Test Case 4: High-Risk Customer and Merchant with Multiple Ontology Tags	72
7.4.5 Test Case 5: Legitimate Prediction with Moderate Ontology Warnings	73
7.5 Summary	74
CHAPTER 8	75
CONCLUSION AND FURTHER WORK	75
8.1 Introduction.....	75
8.2 Conclusion	75
8.3 Results Overview	76
8.4 Limitations and Further Work.....	76
8.5 Summary	77
REFERENCES	78
APPENDIX A.....	81
SAMPLE CODES.....	81

LIST OF FIGURES

Figure	Description	Page
Figure 3.1	LSTM Unit Structure	28
Figure 3.2	Stacked LSTM Network	29
Figure 3.3	Full LSTM Network	30
Figure 4.1	Flowchart of the Proposed System	42
Figure 5.1	Top Level Architecture	44
Figure 6.2.1	Dataset Used	52
Figure 6.2.2	Class Distribution of the dataset	53
Figure 6.2.3	LSTM Model Summary	55
Figure 6.2.4	LSTM Model with Parameters	55
Figure 6.2.5	LIME Explanations	57
Figure 6.2.6	Sample Set of Transactions	58
Figure 6.2.7	Sample Transaction Matched to Multiple Rules	59
Figure 6.2.8	Sample Transactions Matched to a Rule	60
Figure 6.2.9	Fraud Transactions Validated by Ontology	60
Figure 6.2.10	Overview of the Web Application	62
Figure 7.1	Training vs Validation Accuracy	65
Figure 7.2	Training vs Validation Loss	66
Figure 7.3	Confusion Matrix - Training	66
Figure 7.4	Confusion Matrix – Test	66
Figure 7.5	ROC Curve	65
Figure 7.6	LIME Predictions for a Test Customer	68
Figure 7.7	Ontology Validated LIME Predictions for a Test Customer	68
Figure 7.8	Test Case 1	69
Figure 7.9	Test Case 2	70
Figure 7.10	Test Case 3	71
Figure 7.11	Test Case 4	72
Figure 7.12	Test Case 5	73

LIST OF TABLES

Figure	Description	Page
Table 2.1	Comparison of Deep Learning Techniques used in Fraud Detection	19

LIST OF ABBREVIATIONS

Abbreviation	Description
AI	Artificial Intelligence
LSTM	Long Short-Term Memory
ML	Machine Learning
RNN	Recurrent Neural Network
DL	Deep Learning
DNN	Deep Neural Network
GRU	Gated Recurrent Unit
XAI	Explainable Artificial Intelligence
SHAP	SHapley Additive exPlanations
LIME	Local Interpretable Model-Agnostic Explanations
GNN	Graph Neural Network
CNN	Convolutional Neural Network
SVM	Support Vector Machine
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
SMOTE	Synthetic Minority Over-sampling Technique