



**ADVANCING  
RANSOMWARE DETECTION SYSTEM USING  
MACHINE LEARNING**

G.A.A.I.S De Silva  
Reg. No: MS24034036

A THESIS  
SUBMITTED TO  
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY  
(CYBER SECURITY)

September 2025

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

---

Dr. Harinda Fernando

Approved for MSc. Research Project:

---

MSc in IT Programme Co-ordinator, SLIIT

Approved for MSc:

---

Head of Graduate Studies, FoC, SLIIT

# DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

Sign: .....

GAAIS DE SILVA

Date: .....04/11/2025.....

# ABSTRACT

G.A.A.I.S De Silva

MSc. in Information Technology Cyber Security

**Supervisor:** Dr. Harinda Fernando

December 2025

Ransomware attacks pose a significant and evolving threat to data security and operational integrity, necessitating advanced detection mechanisms. This project aims to develop an effective ransomware detection system leveraging machine learning techniques, specifically Recurrent Neural Networks (RNN) and auto encoders, to analyze network traffic for anomalies indicative of ransomware activity. Utilizing the UNSW-NB15 datasets, we undertook extensive data preprocessing, including handling missing values and normalizing features, to prepare the datasets for training. The model employs Long Short-Term Memory (LSTM) layers to capture temporal dependencies and patterns within the network traffic data. The training and validation processes focused on normal traffic data to establish a baseline for detecting deviations caused by ransomware. Our results demonstrate high accuracy in distinguishing between normal and ransomware-infected traffic, with a clear ability to identify potential threats in real-time. This innovative approach showcases the potential of RNN-based auto encoders in enhancing cyber security measures. The conclusion emphasizes the system's effectiveness in providing early warnings of ransomware attacks, thereby significantly aiding in the protection of valuable data assets and maintaining operational continuity.

## ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my research supervisor, **Dr. Harinda Fernando**, for her invaluable guidance, patience, and support throughout my research journey. Her expertise and insightful feedback have been instrumental in shaping this work.

I am also grateful to the **Sri Lankan Institute of Information Technology** for providing me with the necessary resources to carry out my research.

# TABLE OF CONTENTS

DECLARATION.....	4
ABSTRACT .....	5
ACKNOWLEDGEMENT.....	6
TABLE OF CONTENTS .....	7
List of Figures.....	9
List of Tables.....	9
1. Introduction .....	10
1.1 What is a ransomware detection system?.....	10
1.2. Background.....	13
1.3 The problem.....	14
1.4. The proposed solution .....	14
1.5 Deliverable and milestones.....	15
2. Related work.....	17
3. Literature Review .....	18
3.1 Review of Existing Literature.....	18
3.2 Challenges in Ransomware Detection.....	18
3.3 Static Analysis Techniques.....	19
3.4 Dynamic Analysis Techniques .....	20
3.5 Feature Extraction and Selection.....	21
3.6 Machine Learning Models for Ransomware Detection.....	21
3.7 Real-Time Ransomware Detection.....	22
3.8 Adversarial Attacks and Robust Machine Learning.....	23
3.9 Continuous and Adaptive Learning Frameworks.....	23
3.10 Comparative Studies and Evaluations.....	24
3.11 Future Directions and Emerging Trends.....	25
4. Methodology.....	26
4.1 System Architecture .....	29
4.2 Dataset Description (UNSW-NB15) .....	31
4.3 Data Preprocessing .....	31
4.4 Model Development .....	33
4.4.1 AutoEncoder based Anomaly detection model .....	33
4.4.2 RNN-based Classification Model.....	34
4.5 Training and Optimization.....	34
4.6 Machine learning process .....	35

4.7 Datasets training process .....	36
4.8 Comparative Analysis: Supervised Classification Models .....	44
4.8.1 Random Forest Classifier .....	44
4.8.2 XGBoost Classifier.....	45
4.8.3 Isolation Forest .....	45
4.8.4 Neural Network (MLP). .....	45
4.8.5 SVM 1 Class Uniform.....	45
4.8.5 Deep Learning, Unsupervised- Recurrent Autoencoder.....	45
4.9 Evaluation Metrics.....	46
4.10 Risk Assessment Framework.....	46
5. Experimental Setup and Implementation .....	47
5.1 Research instruments.....	47
5.2 Data manipulation and Testing.....	47
4.3 Pitfalls and workarounds .....	48
6. Results and Analysis.....	49
5.1 Supervised Model Comparison .....	51
5.2 Analysis of Ransomware Class Detection.....	52
5.3 Feature Importance Analysis .....	54
6. Conclusions and Future Works.....	55
7. References .....	i

## List of Figures

Figure 1.1 Adding Table of Contents .....	<b>Error! Bookmark not defined.</b>
Figure 1.2 Styles .....	<b>Error! Bookmark not defined.</b>
Figure 2.1 Page Layout.....	<b>Error! Bookmark not defined.</b>
Figure 2.2 Set up the page margin .....	<b>Error! Bookmark not defined.</b>
Figure 3.1 Page Break .....	<b>Error! Bookmark not defined.</b>
Figure 3.2 Adding New Numbering System .....	<b>Error! Bookmark not defined.</b>
Figure 3.3 Format Numbers .....	<b>Error! Bookmark not defined.</b>
Figure 3.4 Change Options.....	<b>Error! Bookmark not defined.</b>
Figure 4.1 Add a Reference 1 .....	<b>Error! Bookmark not defined.</b>
Figure 4.2 Create Source .....	<b>Error! Bookmark not defined.</b>
Figure 4.3 Source Manager .....	<b>Error! Bookmark not defined.</b>

## List of Tables

Table 1.1 Style Table.....	<b>Error! Bookmark not defined.</b>
----------------------------	-------------------------------------