

**Hybridized Model of an Improved Genetic Algorithm with Local Linear Embedding
Algorithm for Feature Selection in Intrusion Detection Systems**

**Tolulope Olushola OLUFEMI
LCU/PG/000138**

**Being a PhD Thesis Submitted to the Department of Computer Science, Faculty of Natural
and Applied Sciences, Lead City University, Ibadan, Oyo State, Nigeria**

**In Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy Degree
(PhD) in Computer Science**

2023

Certification

This is to certify that Tolulope Olushola OLUFEMI with matriculation number LCU/PG/000138 carried out this research work titled “Intrusion Detection System Classification Using a Hybridised Model of an Improved Genetic Algorithm with Local Linear Embedding” in the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State, for the award of Doctor of Philosophy Degree (Ph.D.) in Computer Science and that this has not been previously submitted.

Dr. Wilson Sakpere
(Supervisor)

Date

Dr Wilson Sakpere
(Head of Department)

Date

Dedication

This research work is dedicated to God Almighty, my super mum, Madam Florence Folorunsho Folayegbe Ogunleye, and my very inspiring brother, Dr Ezekiel Olayiwola Ogunleye for their sacrifices over my life.

Do Not Copy, Lead City University, Nigeria

Acknowledgement

I want to thank this prestigious institution, Lead City University, for the opportunity to learn and complete my PhD program and the University Library for providing the adequate materials needed. I am grateful to the Department of Computer Science for giving me the privilege and opportunity to study and learn.

I appreciate my supervisor Dr W. Sakpere for his thorough guidance throughout this work. Thanks to the Departmental postgraduate Coordinator, Dr. A. A. Waheed, for his steadfast encouragement and support. I appreciate your efforts, Sir. God bless you immensely.

I also acknowledge the invaluable contributions of my lecturers, Prof P.O. Achimugu, Prof (Mrs) T.O. Achimugu, Prof S. O Akinola, Dr. Agbaje, Dr W. Ajayi, Dr A. M. Ayoade, Dr (Mrs) Oyekunle, Mrs. K. Okesola, Mrs Afe and other departmental staffs for their support throughout the programme.

My hearty thanks to my Mr Abiodun Olufemi and my loving children Joy, Joshua, John, and Joan who have always supported me through this programme. I greatly appreciate my friends Chinonyelum Nwufoh, Dr Kolapo Ridwan, Bukola Gaji, Olubode Ojajuni, Victor Owonifaari and Olabisi Arinola. Lastly, my in-depth appreciation goes to my big family, Mr Dele Durojaye, Mr & Mrs Dupe-Olayebi, Sefunmi Ola, and Wale Ajayi, this feat wouldn't have been possible if not for your encouraging words always.

“Even though the above-mentioned institutions and persons have assisted in the process of this research work, I alone stand responsible for the errors, if any, found in the work.

Table of Contents

Content	Page
Certification	2
Dedication	3
Acknowledgement	4
Table of Contents	5
List of Tables	18
List of Figures	19
List of Algorithms	xv
List of Acronyms	21i
List of Appendix	24
Chapter One: Introduction	
1.1	Background to the Study
Error! Bookmark not defined.	
1.2	Statement of the Problem
Error! Bookmark not defined.	
1.3	Aim and Objectives of the Study
Error! Bookmark not defined.	
1.4	Research Questions
Error! Bookmark not defined.	

1.5 Justification of the Study

Error! Bookmark not defined.

1.6 Scope of the Study

Error! Bookmark not defined.

1.7 Limitations of the Study

Error! Bookmark not defined.

1.8 Operational Definition of Terms

Error! Bookmark not defined.

Endnotes

Error! Bookmark not defined.

Chapter Two: Literature Review

2.1 Intrusion Detection

Error! Bookmark not defined.

2.2 Steganography

Error! Bookmark not defined.

2.2.1 Investigation of Networks

Error! Bookmark not defined.

2.2.2 Network Techniques for Botnet Investigation

Error! Bookmark not defined.

2.2.3 Machine Learning Role in Botnet Investigations

Error! Bookmark not defined.

2.2.4 Frameworks for IoT and Smart Systems

Error! Bookmark not defined.

2.2.5 The NIST Digital Investigative Procedure

Error! Bookmark not defined.

2.2.6 Denial of Service (DoS) and Distributed (DDoS) (DoS and DDoS)

Error! Bookmark not defined.

2.2.7 Discussion of the Different Types of Attacks

Error! Bookmark not defined.

2.3 Intrusion Detection Systems (IDS)

Error! Bookmark not defined.

2.3.1 Assaults against Cyber Security

Error! Bookmark not defined.

2.3.2 Identifying an Identity Clone Attack

Error! Bookmark not defined.

2.3.3 Detection Systems That Are Based on Intrusion Detection Systems

Error! Bookmark not defined.

2.3.4 A Strategy for the Detection of Anomalies

Error! Bookmark not defined.

2.4 Artificial Intelligence (AI)

Error! Bookmark not defined.

2.4.1 Artificial intelligence (AI)

Error! Bookmark not defined.

2.4.2 Machine Learning (ML)

Error! Bookmark not defined.

2.4.2.1 Intensive Education

Error! Bookmark not defined.

2.4.2.2 Inductive Learning

Error! Bookmark not defined.

2.4.2.3 Learning through Deduction

Error! Bookmark not defined.

2.4.3 Machine Learning's Importance in Intrusion Investigations

Error! Bookmark not defined.

2.4.4 Approaches to Supervised Machine Learning

Error! Bookmark not defined.

2.4.4.1 k-Nearest Neighbour

Error! Bookmark not defined.

2.4.4.2 Support Vector Machine (SVM)

Error! Bookmark not defined.

2.4.4.2.1 Machine with Linear Separable Support Vectors **Error! Bookmark not defined.**

2.4.4.2.2 Nonlinear Support Vector Machines (NSVMs) **Error! Bookmark not defined.**

2.4.5 Naive Bayes Algorithm

Error! Bookmark not defined.

2.4.6 Decision Tree

Error! Bookmark not defined.

2.4.7 Random Forests

Error! Bookmark not defined.

2.4.8 AdaBoost Classifier

Error! Bookmark not defined.

2.4.9 Logistic Regression

Error! Bookmark not defined.

2.4.10 Convolutional Neural Network 59

2.4.11 Recurrent Neural Networks (RNN)

Error! Bookmark not defined.

2.5 Preparation Stages for Data

Error! Bookmark not defined.

2.5.1 Particle Swarm Optimization (PSO)

Error! Bookmark not defined.

2.5.2 Chi-Square (Chi²)

Error! Bookmark not defined.

2.5.3 Mutual Information (MI)

Error! Bookmark not defined.

2.5.4 Genetic Algorithm

Error! Bookmark not defined.

2.5.4.1 Levy Flight Algorithm

Error! Bookmark not defined.

2.5.4.2 Other Enhancements for Evolutionary Algorithms

Error! Bookmark not defined.

2.5.4.3 Particle Swarm Optimization (PSO)

Error! Bookmark not defined.

2.5.5 Principal Components Analysis

Error! Bookmark not defined.

2.5.6 Deep Learning

Error! Bookmark not defined.

2.5.7 Chi Square

Error! Bookmark not defined.

2.5.8 ANOVA (Analysis of Variance)

Error! Bookmark not defined.

2.5.9 Principal Component Analysis (PCA)

Error! Bookmark not defined.

2.6 A Classification Model Based on Deep Neural Networks

Error! Bookmark not defined.

2.7 Deep Neural Network

Error! Bookmark not defined.

2.8 Datasets

Error! Bookmark not defined.

2.8.1 CICIDS2017 Dataset

Error! Bookmark not defined.

2.8.2 NSL/KDD

Error! Bookmark not defined.

2.8.3 DARPA (Lincoln Laboratory 1998,1999)

Error! Bookmark not defined.

2.8.4 DEFCON-8

Error! Bookmark not defined.

2.8.5 Center for Applied Internet Data Analysis (2002/2016)

Error! Bookmark not defined.

2.8.6 Lawrence Berkeley National Laboratory (LBNL)

Error! Bookmark not defined.

2.8.7 CDX

Error! Bookmark not defined.

2.8.8 Kyoto 2009

Error! Bookmark not defined.

2.8.9 Twente

Error! Bookmark not defined.

2.8.10 UMASS (University of Massachusetts - 2011)

Error! Bookmark not defined.

2.8.11 ISCX2012

Error! Bookmark not defined.

2.8.12 ADFA

Error! Bookmark not defined.

2.8.13 AWS: CSE-CIC-IDS2018

Error! Bookmark not defined.

2.8.14 2019 Dataset of the CIC-DDoS2019

Error! Bookmark not defined.

2.8.15 MQTT-IOT-IDS2020

Error! Bookmark not defined.

2.9 Related Works

Error! Bookmark not defined.

2.10 Summary of the Review **Error!**

Bookmark not defined.

Endnotes **Error!**

Bookmark not defined.

Chapter Three: Methodology

3.1 Introduction 120

3.1.1 Dataset **Error!**

Bookmark not defined.

3.1.2 Python as an Implementation Tool **Error!**

Bookmark not defined.

3.2 Optimization Error!

Bookmark not defined.

3.2.1 Improved Genetic Algorithm Error!

Bookmark not defined.

3.2.1.1 Genetic Algorithm (GA) Error!

Bookmark not defined.

3.2.1.2 Lévy Flight Error!

Bookmark not defined.

3.2.2 Locally Linear Embedding (LLE) Error!

Bookmark not defined.

3.3 Classification Error!

Bookmark not defined.

3.3.1 Support Vector Machine Error!

Bookmark not defined.

3.3.2 K Nearest Neighbor Error!

Bookmark not defined.

3.4 Performance Evaluation Error!

Bookmark not defined.

3.5 Improved Model with classifiers (I-GA-LLE-SVM) and (I-GA-LLE-KNN) Error!

Bookmark not defined.

Endnotes Error!

Bookmark not defined.

Chapter Four: Results and Discussion

4.1 Introduction **Error!**

Bookmark not defined.

4.1.1 Data Processing **Error!**

Bookmark not defined.

4.2 Optimization Results **Error!**

Bookmark not defined.

4.2.1 I-GA Optimization Model **Error!**

Bookmark not defined.

4.2.2 LLE Optimization Model **Error!**

Bookmark not defined.

4.2.3 I-GA-LLE Optimization Model **Error!**

Bookmark not defined.

4.3 Classification Result **Error!**

Bookmark not defined.

4.3.1 I-GA-SVM and I-GA-KNN **Error!**

Bookmark not defined.

4.3.2 LLE-SVM and LLE-KNN **Error!**

Bookmark not defined.

4.3.3 I-GA-LLE-SVM and I-GA-LLE-KNN **Error!**

Bookmark not defined.

4.4 Evaluation Results **Error!**

Bookmark not defined.

4.4.1 I-GA with the Classifiers (SVM, KNN) **Error!**

Bookmark not defined.

4.4.2 LLE with the Classifiers **Error!**

Bookmark not defined.

4.4.3 I-GA-LLE with the Classifiers **Error!**

Bookmark not defined.

4.5 Discussion (Optimization) **Error!**

Bookmark not defined.

4.5.1 I-GA Subset (Reduced Dataset) **Error!**

Bookmark not defined.

4.5.2 LLE Subset (Reduced Dataset) **Error!**

Bookmark not defined.

4.5.3 I-GA-LLE (Reduced Dataset) **Error!**

Bookmark not defined.

4.6 Discussion (Classification) **Error!**

Bookmark not defined.

4.6.1 I-GA-SVM and I-GA-KNN Classification **Error!**

Bookmark not defined.

4.6.2 LLE-SVM and LLE-KNN Classification **Error!**

Bookmark not defined.

4.6.3 I-GA-LLE-SVM and I-GA-LLE-KNN Classification **Error!**

Bookmark not defined.

4.7 Evaluation Discussion **Error!**

Bookmark not defined.

4.7.1 I-GA with the Classifiers (Evaluation Discussion) **Error!**

Bookmark not defined.

4.7.2 LLE with the Classifiers (Discussion) **Error!**

Bookmark not defined.

4.7.3 I-GA-LLE with the Classifiers (Evaluation Discussion) **Error!**

Bookmark not defined.

4.7.4 I-GA with KNN and LLE-KNN 172

4.8 Comparison with Previous Works **Error!**

Bookmark not defined.

4.9 Table 4.6 Summary of the Results **Error!**

Bookmark not defined.

Chapter Five: Conclusion

5.1 Summary of Results **Error!**

Bookmark not defined.

5.1.1 Conclusion **Error!**

Bookmark not defined.

5.2 Contribution to Knowledge **Error!**

Bookmark not defined.

5.3 Suggestions for Further Studies

Error!

Bookmark not defined.

Bibliography

Error!

Bookmark not defined.

Appendix

Error!

Bookmark not defined.

Biodata

Error!

Bookmark not defined.

The University Compliance Certification

248

Do Not Copy, Lead City University, Nigeria

List of Tables

Table	Title	Page
3.1	CICIDS2017 Dataset	123
4.1	Reduced Data	151
4.2	Results obtained from IGA with Classifiers	160
4.3	Results obtained from LLE with Classifiers	161
4.4	Performance measures for IGA-LLE with classifiers	174
4.5	Comparison of Results	174
4.6	Summary of Results	175

List of Figures

Figure	Title	Page
3.1	Architectural Workflow	120
3.2	Model Design	121
3.3	Improved GA Model	125
3.4	Flowchart for I-GA	133
3.5	I-GA-LLE Model	138
4.1	I-GA Model Snippet	152
4.2	I-GA Optimization Result	152
4.3	LLE Model Snippet	153
4.4	LLE Optimization Result	153
4.5	I-GA-LLE Model Snippet	154
4.6	I-GA-LLE Optimization Result	154
4.7	IGA-SVM Confusion Matrix	155
4.8	IGA-KNN Confusion Matrix	156
4.9	LLE-SVM Confusion Matrix	156
4.10	LLE-KNN Confusion Matrix	157
4.11	IGA-LLE-SVM Confusion Matrix	157
4.12	IGA-LLE-KNN Confusion Matrix	158
4.13	IGA ROC	159
4.14	Bar Chart IGA	159
4.15	LLE ROC	160
4.16	Bar Chart LLE	160

4.17	IGA-LLE ROC	161
4.18	Bar Chart IGA-LLE	161

Do Not Copy, Lead City University, Nigeria

List of Algorithms

Algorithm	Title	Page
2.1:	KNN Algorithm	50
2.2:	SVM Algorithm	53
2.3:	Naive Bayes Algorithm	54
2.4:	Decision Tree Algorithm	56
2.5	Random Forest Algorithm	57
2.6	AdaBoost Algorithm	57
2.7	Logistics Regression	58
2.8	CNN Algorithm	60
2.9	UMAP Algorithm	65
2.10	PSO Algorithm	65
2.11	Mutual Information Algorithm	67
2.12	Genetic Algorithm	68
3.1	Improved Genetic Algorithm	130
3.2	Locally Linear Embedding Algorithm	137
3.3	SVM Algorithm	139
3.4	KNN Algorithm	142
3.5	Improved Model	147

List of Pseudocodes

Pseudocode	Title	Page
2.1	Pseudocode for Levy Flight Algorithm	70
2.2	Pseudocode for Differential Evolution (DE) Algorithm	70
2.3	Pseudocode for Particle Swarm Optimisation (PSO)	71
2.4	Pseudocode for PCA	72

List of Acronyms

Abbreviation

Meaning

GA

Genetic Algorithm

LLE

Local Linear Embedding

PSO

Particle Swarm Optimization

SVM

Support Vector Machine

KNN

K Nearest Neighbour

Do Not Copy, Lead City University, Nigeria

List of Appendix

Appendix	Title	Page
1	Source Code	197

Do Not Copy, Lead City University, Nigeria