

**AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS
ANOMALY-BASED NETWORK INTRUSION DETECTION
SYSTEM**

**OWOKA, EMMANUEL OLUSOLA
(20PCG02184)**

AUGUST, 2022

**AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS
ANOMALY-BASED NETWORK INTRUSION DETECTION
SYSTEM**

BY

**OWOKA, EMMANUEL OLUSOLA
(20PCG02184)**

B.Sc Computer Science, University of Benin, Benin-City

**A DISSERTATION SUBMITTED TO THE SCHOOL OF
POSTGRADUATE STUDIES IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE
(M.Sc) DEGREE IN MANAGEMENT INFORMATION SYSTEMS IN
THE DEPARTMENT OF COMPUTER AND INFORMATION
SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY,
COVENANT UNIVERSITY.**

AUGUST, 2022

ACCEPTANCE

This is to attest that this dissertation is accepted in partial fulfilment of the requirements for the award of the degree of MASTER of Sciences in Management Information Systems in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Mr. Taiwo B. Erewunmi
(Secretary, School of Postgraduate Studies)

Signature and Date

Prof Akan B. Williams
(Dean, School of Postgraduate Studies)

Signature and Date

DECLARATION

I, **OWOKA, EMMANUEL OLUSOLA (20PCG02184)**, declare that this research was carried out by me under the supervision of Dr. Aderonke A. Oni of the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria. I attest that the dissertation has not been presented either wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

OWOKA, EMMANUEL OLUSOLA

Signature and Date

CERTIFICATION

We certify that this dissertation titled “**AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS ANOMALY-BASED NETWORK INTRUSION DETECTION SYSTEM**” is an original research work carried out by **OWOKA, EMMANUEL OLUSOLA (20PCG02184)** in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria under the supervision of Dr. Aderonke A. Oni. We have examined and found this work acceptable as part of the requirements for the award of Master of Science in Management Information Systems.

Dr. Aderonke A. Oni
(Supervisor)

Signature and Date

Prof. Olufunke O. Oladipupo
(Head of Department)

Signature and Date

Prof. Olufunke R. Vincent
(External Examiner)

Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies)

Signature and Date

DEDICATION

I dedicate this work to the Almighty God, for His infinite wisdom, grace, and love over my life. Also, this work is dedicated to my loving parents who have both worked exceptionally hard to set me up for success.

ACKNOWLEDGEMENTS

I appreciate God almighty for the strength, grace, wisdom and understanding to execute this study. Many thanks to the head of department, Prof. Olufunke Oladipupo, and the entire members of the faculty for this opportunity. I appreciate my supervisor, Dr. Aderonke Oni for her guidance, support, and encouragement. I also appreciate my family, friends and colleagues for their help and support.

TABLE OF CONTENTS

CONTENTS	PAGES
COVER PAGE	i
TITLE PAGE	ii
ACCEPTANCE	iii
DECLARATION	iv
CERTIFICATION	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	x
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xv
ABSTRACT	xviii
CHAPTER ONE: INTRODUCTION	1
1.1 Background to the Study	1
1.2 Statement of the Problem	4
1.3 Aim and Objectives of the Study	4
1.4 Significance of the Study	5
1.5 Scope of the Study	5
1.6 Organization of the Study	5
CHAPTER TWO: LITERATURE REVIEW	6
2.1 Introduction	6
2.2 Conceptual Review	6
2.2.1 Anomaly Detection	6
2.2.2 Machine Learning	9
2.2.3 Imbalanced Dataset	24
2.2.4 Cybersecurity	25
2.2.5 Intrusion Detection System	29
2.3 Methodological Review	31
2.3.1 Machine Learning Approach	31
2.3.2 Deep Learning Approach	32
2.3.3 Ensemble Learning Approach	34
2.3.4 Existing Model	36
2.4 Related Works	37

CHAPTER THREE: METHODOLOGY	43
3.1 Introduction	43
3.2 Proposed Model	45
3.3 Data Collection	46
3.4 Data Pre-processing	48
3.5 Feature Selection	49
3.6 Modelling	50
3.7 Evaluation	53
3.8 Development Environment	55
CHAPTER FOUR: RESULTS AND DISCUSSION	58
4.1 Introduction	58
4.2 Data Collection and Pre-processing	58
4.3 Modelling	61
4.3.1 Models Without Feature Selection	64
4.3.2 Models with Feature Selection	71
4.3.3 Models with Feature Selection and Balancing Technique	80
4.3.4 Other Models implemented	88
4.3.5 Proposed Model	90
4.3.6 Unknown Attack Evaluation	92
4.4 Summary of Results	93
4.5 Discussion of Findings	98
CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS	101
5.1 Summary	101
5.2 Conclusion	101
5.3 Contributions to knowledge	102
5.4 Recommendations	102
5.5 Limitations to Study	102
REFERENCES	103

LIST OF FIGURES

FIGURES	TITLE OF FIGURES	PAGES
2.1	The relationship between AI, ML, NN, and DL	10
2.2	A standard machine learning pipeline	11
2.3	Structure of a confusion matrix	19
2.4	The five functions of the NIST cybersecurity framework version 1.1	28
2.5	Intrusion detection system overview	30
2.6	A diagram of a host-based IDS in a network.	30
2.7	A diagram of a network-based IDS in a network	30
2.8	Existing model architecture	36
3.1	The process flow	44
3.2	Proposed model architecture	45
3.3	Architecture of a feed-forward neural network	50
4.1	Class distribution of the pre-processed CICIDS2018 dataset	60
4.2	Correlation heatmap of the cicids2018 dataset	61
4.3	Loss curve of the 128-64-32-16 architecture on the CICIDS2018 dataset	62
4.4	Classification report of the 128-64-32-16 architecture on the CICIDS2018 dataset	62
4.5	Loss curve of the 256-128-64-32-16 architecture on the CICIDS2018 dataset	63
4.6	Classification report of the 256-128-64-32-16 architecture on the CICIDS2018 dataset	63
4.7	Loss curve of a single multi-class model	65
4.8	Confusion matrix of a single multi-class model	65
4.9	Classification report of single multi-class model	66

4.10	Loss curve of the individual models used for the one-vs-all modelling	67
4.11	Confusion matrix of method_A	68
4.12	Classification report of method_A	68
4.13	Confusion matrix of method_B	69
4.14	Classification report of method_B	69
4.15	Confusion matrix of method_C	70
4.16	Classification report of method_C	71
4.17	Feature importance score of the selected 15 features of each class and single model	72
4.18	Feature importance score of the selected 15 features of each class and single model (Contd.)	73
4.19	Loss curve of the single multi-class model with feature selection	74
4.20	Confusion matrix of the single multi-class model with feature selection	74
4.21	Classification report of the single multi-class model with feature selection	75
4.22	Loss curve of the individual models used for the one-vs-all modelling with feature selection	76
4.23	Confusion matrix of method_A with feature selection	77
4.24	Classification report of method_A with feature selection	77
4.25	Confusion matrix of method_B with feature selection	78
4.26	Classification report of method_B with feature selection	78
4.27	Confusion matrix of method_C with feature selection	79
4.28	Confusion matrix of method_C with feature selection	80
4.29	Confusion matrix of the single multi-class model with feature selection and balancing techniques implemented	82
4.30	Classification report of the single multi-class model with feature selection and balancing techniques implemented	82

4.31	Confusion matrix of method_A with feature selection and balancing techniques implemented	84
4.32	Classification report of method_A with feature selection and balancing technique implemented	85
4.33	Confusion matrix of method_B with feature selection and balancing techniques implemented	85
4.34	Classification report of method_B with feature selection and balancing technique implemented	86
4.35	Confusion matrix of method_C with feature selection and balancing technique	87
4.36	Classification report of method_C with feature selection and balancing technique	87
4.37	Hyperparameter of the xgboost model	88
4.38	Confusion matrix of the xgboost model	88
4.39	Classification report of the xgboost model	89
4.40	Hyperparameters of the random forest model	89
4.41	Confusion matrix of the random forest model	89
4.42	Classification report of the random forest model	90
4.43	Confusion matrix of the proposed model	91
4.44	Confusion matrix of the proposed model	91
4.45	Unknown attack (Type-B) evaluation of the proposed model	92
4.46	Unknown attack (Type-A) evaluation of the proposed model	93

LIST OF TABLES

TABLES	TITLE OF TABLES	PAGES
1.1	Cyber-attacks and their definitions	3
1.2	Summary of Objectives	5
3.1	Snapshot of the CICIDS2018 features	47
3.2	Confusion matrix	54
3.3	Evaluation metrics	55
4.1	The csv files of the CICIDS2018 dataset	58
4.2	Removed features from the CICIDS2018 dataset	59
4.3	The class distribution of the CICIDS2018 dataset before and after the removal of null and infinite values	59
4.4	Completely pre-processed CICIDS2018 dataset's class distribution	60
4.5	Selected hyperparameters	64
4.6	False alarm rate and specificity of the single multi-class model	66
4.7	False alarm rate and specificity of method_A	68
4.8	False alarm rate and specificity of method_B	70
4.9	False alarm rate and specificity of method	71
4.10	False alarm rate and specificity of the implemented model	75
4.11	False alarm rate and specificity of the method_A	77
4.12	False alarm rate and specificity of method_B	79
4.13	False alarm rate and specificity of method_C with feature selection	80
4.14	Performance of the single multi-class model on various balancing techniques	81
4.15	False alarm rate and specificity of the single multi-class model with feature selection and balancing techniques implemented	82

4.16	Results of balancing techniques performed on the infiltration class model	83
4.17	Results of balancing techniques performed on the web class	84
4.18	False alarm rate and specificity of method_A with feature selection and balancing technique implemented	85
4.19	False alarm rate and specificity of method_B with feature selection and balancing technique implemented	86
4.20	False alarm rate and specificity of method_C with feature selection and balancing technique implemented	87
4.21	False alarm rate and specificity of the XGBoost model with feature selection and balancing technique implemented	89
4.22	False alarm rate and specificity of the Random Forest model with feature selection and balancing technique implemented	90
4.23	False alarm rate and specificity of the proposed model with feature selection and balancing technique implemented	91
4.24	Comparative summary of models without feature selection	94
4.25	Comparative summary of models with feature selection	95
4.26	Comparative summary of models with feature selection and balancing techniques implemented	96
4.27	Comparative summary of the proposed model, state-of-the-art machine learning algorithms and existing models in literature	97

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
A-NIDS	Anomaly Network Intrusion Detection Systems
AWS CLI	Amazon Web Service Command Line Interface
BiDLSTM	Bidirectional Long Short-Term Memory
CES-CIC-IDS	Communications Security Establishment and the Canadian Institute for Cybersecurity Intrusion Detection System
CNN	Convolution Neural Network
CRISP-DM	CRoss Industry Standard Process for Data Mining
CSV	Comma Separated Value
DBN	Deep Belief Network
DDoS	Distributed Denial of Service
DL	Deep Learning
DNN	Deep Neural Network
DoS	Denial of Service
DVWA	Damn Vulnerable Web App
EFC	Energy-Based Flow Classification
ELM	Extreme Learning Machine
EOT	Edge-of-Things
FAR	False Alarm Rate
FN	False Negative
FP	False Positive
FTP	File Transfer Protocol
HOIC	High Orbit Ion Cannon
IDS	Intrusion Detection System
Interpol	International Criminal Police Organization
IoT	Internet of Things

IR	Imbalance Ratio
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbours
LASSO	Least Absolute Shrinkage and Selection Operator
LightGBM	Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NB	Naïve Bayes
NIDS	Network Intrusion Detection
NIST	National Institute of Standards and Technology
NMUS	NearMiss Under-sampling
NN	Neural Network
PCA	Principal Component Analysis
PKI	Public Key Infrastructure
R^2	Coefficient of Determination
R2L	Root to Local
RBM	Restricted Boltzmann Machines
RF	Random Forest
RNN	Recurrent Neural Network
SC	Silhouette Coefficient
SCADA	Supervisory Control and Data Acquisition
SFSDT	Sequence Forward Selection algorithm with Decision Tree
SMO	Sequential Minimal Optimization
SMOTE	Synthetic Minority Over-sampling Technique
SSH	Secure Shell

SVM	Support Vector Machine
TN	True Negative
TP	True Positive
U2R	User to Root
WELM	Weighted Extreme Learning Machine
XGBoost	eXtreme Gradient Boosting
XML	eXtensible Markup Language
XSS	Cross-Site Scripting

ABSTRACT

The internet is advancing at a fast pace, and it is very essential to individuals and organizations. Also, there are a lot of malicious actors on the internet and a successful attack on a victim can be very devastating. Hence, the growing need for cybersecurity. Network security helps protect computer networks from attackers and this can be achieved with the help of intrusion detection systems (IDS). Over the years researchers have proposed improvements to IDSs, however, the problem of low detection rate especially towards minority classes within the available datasets plagues the research area. This study builds and evaluates an ensemble anomaly-based network intrusion detection system for multi-class classification using an argmax one-vs-all approach. The Communications Security Establishment and the Canadian Institute for Cybersecurity Intrusion Detection System 2018 dataset (CSE-CIC-IDS2018), referred to as CICIDS2018, was used in this study. The eXtreme Gradient Boosting (XGBoost) was used for feature selection and the Minority Oversampling Technique (SMOTE) alongside cost-sensitive learning were utilized to address the imbalanced nature of the CICIDS2018 dataset. The Multilayer Perceptron (MLP), Random Forest (RF), and XGBoost were used to build the ensemble model. A one-vs-all approach was adopted to design an ensemble of the classifiers tailored to detecting a specific class within the dataset. This means that the feature selection process was done for each class, producing multiple datasets based on the number of classes within the dataset. The results of the classifiers are then combined and aggregated using the argmax function. Finally, the proposed model was evaluated against other models, existing works in literature and unknown attacks to see how well the model performs. The results showed that the proposed approach performs better than other approaches achieving a better macro average F1-score of 83.50% and an improved classification of the minority classes, attaining an F1-score of 29.95% and 75.98% in the infiltration and web classes respectively. The infiltration class was seen to be hard to decipher from the benign class and so approaches to properly separate and oversample the infiltration class should be taken to improve the detection of the class.

Keywords: Intrusion Detection System, CICIDS2018, Cyber Security, Machine Learning, Deep Learning