

Naif Arab University for Security Sciences  
College of Computer and Information Security  
Information Security Department



# PERFORMANCE EVALUATION OF MACHINE LEARNING BASED INTRUSION DETECTION SYSTEMS FOR CLOUD COMPUTING

By **Khawla BinAjlan**

Under the Supervision of

**Dr. Mohammad Shadi Alhakeem**

Submitted in partial fulfillment of the requirements  
for the Degree of Master in Information Security at  
the Department of Information Security at the Collage  
of Computer and Information Security  
Naif Arab University for Security Sciences

2022

## Table of Contents

Thesis Abstract .....	A
مستخلص الرسالة .....	B
DEDICATION .....	C
ACKNOWLEDGMENTS .....	D
Table of Contents .....	E
List of Figures.....	G
List of Tables.....	H
CHAPTER 1: INTRODUCTION.....	1
1.1 Overview .....	1
1.2 Motivation .....	3
1.3 Research Question.....	3
1.4 Problem Statement .....	4
1.5 Aims and Objectives .....	5
1.6 Thesis Structure.....	6
CHAPTER 2: BACKGROUND AND LITERATURE REVIEW.....	7
2.1 Background .....	7
2.1.1 Cloud computing .....	7
2.1.2 Intrusion Detections System (IDS).....	10
2.1.3 Machine Learning.....	12
2.2 Literature Review.....	15

CHAPTER 3: METHODOLOGY .....	24
3.1 Proposed Work.....	24
3.1.1 Data Pre-processing:.....	25
3.1.2 Features selection: .....	28
3.1.3 Classification: .....	31
3.2 Classification Models.....	31
3.2.1 Naive Baye Classifier.....	32
3.2.2 Decision Tree Classifier .....	33
3.2.3 Gradient Boosting Classifier .....	34
3.2.4 Logistic Regression.....	35
3.2.5 Support Vector Machines.....	36
3.3 Data specification.....	37
3.4 Performance measurements.....	38
3.5 Implementation Tools: .....	42
CHAPTER 4: EXPERIMENTS RESULTS AND ANALYSIS .....	43
4.1 Experimental Results and Observations.....	43
4.2 Comparison and Analysis.....	44
4.2.1 Comparison Between Different Classifier on Brute Force Attack.....	44
CHAPTER 5: CONCLUSION .....	48
5.1 Conclusion.....	48
5.2 Future Work .....	49
References .....	50

## List of Figures

FIGURE 2-1 TYPES OF SERVICES PROVIDED BY CLOUD COMPUTING (EL-SAYED, 2014).....	8
FIGURE 3-1 ILLUSTRATES THE ARCHITECTURE OF OUR APPROACH .....	24
FIGURE 3-2 FEATURE SELECTION FIRST FEATURE SUBSET. ....	29
FIGURE 3-3 FEATURE IMPORTANCE FOR THE THREE TYPES OF ATTACK .....	30
FIGURE 3-4 FEATURE SELECTION SECOND FEATURE SUBSET. ....	30
FIGURE 3-5 SAMPLE CODE FOR NAIVE BAYES CLASSIFICATION .....	33
FIGURE 3-6 SAMPLE CODE FOR DECISION TREE CLASSIFICATION .....	34
FIGURE 3-7 SAMPLE CODE FOR GRADIENT BOOSTING CLASSIFIER.....	35
FIGURE 3-8 SAMPLE CODE FOR LOGISTIC REGRESSION.....	36
FIGURE 3-9 SAMPLE CODE FOR SUPPORT VECTOR MACHINES.....	37
FIGURE 3-10 CONFUSION MATRIX FOR BINARY CLASSIFICATION.....	38
FIGURE 3-11 CONFUSION MATRIX FOR MULTI-CLASS CLASSIFICATION.....	38
FIGURE 3-12 TN FOR CLASS B .....	39

## List of Tables

TABLE 2-1 SUMMARY TABLE OF LITERATURE REVIEW .....	20
TABLE 2-2 COMMONLY USED DATASETS IN IDS ARE SUMMARIZED AND COMPARED IN TABLE .....	23
TABLE 3-1 FILES USED FROM THE CSE-CIC-IDS2018 DATASET .....	26
TABLE 3-2 DATA SAMPLING FILES.....	27
TABLE 3-3 DATA SAMPLING .....	28
TABLE 3-4 TRAINING DATA THE TIME AND ACCURACY FOR EACH CLASSIFIER.....	31
TABLE 4-1 TESTING DATA RESULTS THE ACCURACY AND TIME FOR EACH CLASSIFIER.....	44
TABLE 4-2 THE FILE THAT WAS USED IN THIS EXPERIMENT .....	45
TABLE 4-3 COMPARE RESULTS BETWEEN DIFFERENT CLASSIFIER WITH THE FIRST METHOD OF FEATURE SELECTION .....	45
TABLE 4-4 COMPARE RESULTS BETWEEN DIFFERENT CLASSIFIER WITH THE SECOND METHOD OF FEATURE SELECTION .....	46