ANDROID BASED AUTOMATIC E-MAIL DETECTION SYSTEM USING NAIVE BAYES ALGORITHM/CLASSIFIER

**BY**

**ANIMASAUN DAMILARE ADERINWALE**

***1605022010***

**SUBMITTED TO**

**DEPARTMENT OF COMPUTER SCIENCE**

**LAGOS STATE POLYTECHNIC, IKORODU, LAGOS.**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF HIGHER NATIONAL DIPLOMA (HND) IN COMPUTER SCIENCE.**

**AUGUST, 2019.**

**CERTIFICATION**

This is to certify that this project was written and compiled by **ANIMASAUN DAMILARE ADERINALE** with matric no. ***1605022010*** and was approved for meeting the requirement of the award of Higher National Diploma in Computer Science.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**ANIMASAUN DAMILARE ADERINWALE**  **DATE**

STUDENT

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**MR. AKANJI WASIU A. DATE**

SUPERVISOR

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**MR. ADELANWA S.O.A. DATE**

COORDINATOR

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**MR. AKANJI W.A. DATE**

HEAD OF DEPARTMENT

**ACKNOWLEDGEMENT**

My appreciation goes to my parents for their support throughout my years of studies, my supervisor **Mr, Akanji W.A**, my family and friends for all their support and co-operation. I appreciate you all.

**DEDICATION**

I dedicate this project to God almighty for giving me the grace to complete this project. I also dedicate this project to the gallant Nigerian Army officers fighting insurgency in the North-Eastern Nigeria.

**Abstract**

Electronic Mail (E-mail) has established a significant place in information user’s life. Mails are used as a major and important mode of information sharing because emails are faster and effective way of communication. Spam comes in email through unsolicited messages sent over the Internet.

Email Message Filtering (EMF) is the process which is used in order to classify the emails into various categories on the basis of their content.

**TABLE OF CONTENTS**

Title Page i

Certification ii

Acknowledgement iii

Dedication iv

Table of Contents v-vi

Abstract vii

**CHAPTER ONE: INTRODUCTION**

1.0 Introduction 1

1.1 Statement of the Problem 4

1.2 Aims of the Study 5

1.3 Objectives of the study 5

1.4 Methodology 5

1.5 Scopes of the Study 5

* 1. Definition of Terms 5

**CHAPTER TWO: LITERATURE REVIEW**

2.0 Literature Review 7

2.1 Summary of Review Related Works 11

2.2 Survey of Android System Security 13

2.3 Android Malware Detection 13

2.4 Spam 14

2.4.1 Where Spam Appears 14

2.4.2 What Spam does 15

2.5 Malware 16

2.6 Types of Malware 16

2.6.1 Virus 17

2.6.2 Worm 17

2.6.3 Spyware 17

2.6.4 Adware 17

2.6.5 Trojan 17

2.6.7 Botnet 17

2.7 Malware Obfuscation 17

2.8 Window API Calls 18

2.9 Malware Propagation 18

2.9.1 Through Operating System 18

2.9.2 Through Wireless Networks 19

2.9.3 Through File Sharing 19

2.9.4 Through Social Networking 19

2.9.5 Inbound and Outbound Filtering of Email 19

2.10 Spam Filtering Using NB-SVM 19

2.11 Outline of Spam Filtering 20

* 1. Rundown of Naïve Bayes Algorithm 21

**CHAPTER THREE: RESEARCH METHODOLOGY**

3.0 Introduction 22

3.1 Pre-processing 23

3.2 Data Set 23

3.3 System Algorithm 23

* 1. System Flow Chart 24

**CHAPTER FOUR: SYSTEM IMPLEMENTATION**

4.1 Choice of Programming 25

4.2 System Implementation 26

4.3 System Testing 26

4.4 System Modules Evaluation 27

4.5 System Documentation 30

4.6 User Documentation 31

4.7 System Requirement 31

**CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS**

5.1 Summary 33

5.2 Conclusion 33

5.3 Recommendation 33

References

Appendix

**LIST OF FIGURES**

Steps of Knowledge Discovery (2.1.)

Comparison graph for accuracy and execution time analysis (2.2)

Types of Malware (2.3)

Filtration Process (2.4)

Proposed System Architecture (3.1)

General Design (3.2)

**LIST OF TABLES**

Summary of Review Related Works (2.1)

Software and Hardware Specification (4.1)