

**HYBRIDIZED SUPPORT VECTOR
MACHINE_ARTIFICIAL NEURAL NETWORK
MALARIA PREDICTION MODEL**

By

ABISOYE, Opeyemi Aderiike

(00/55EC017)

B Sc., M.Sc.(UNILORIN)2010

**A THESIS SUBMITTED TO THE DEPARTMENT OF
COMPUTER SCIENCE, FACULTY OF COMMUNICATION
AND INFORMATION SCIENCES, UNIVERSITY OF ILORIN,
ILORIN, NIGERIA.**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE AWARD OF THE DEGREE OF DOCTOR OF
PHILOSOPHY (PH.D) IN COMPUTER SCIENCE**

MAY, 2018.

CERTIFICATION

This is to certify that this Ph.D. thesis titled HYBRIDIZED SUPPORT VECTOR MACHINE_ARTIFICIAL NEURAL NETWORK MALARIA PREDICTION MODEL was written by Abisoye, Opeyemi Aderiike (00/55EC017) has been read and approved as meeting the requirements of the Department of Computer Science, Faculty of Communication and Information Science, University of Ilorin, Ilorin, Nigeria for the award of Ph.D. Degree in Computer Science.

.....
Dr. R.G. Jimoh
(Supervisor)

.....
Date

.....
Dr. D.R. Aremu
(Head of the Department)

.....
Date

.....
(Internal/External Examiner)

.....
Date

DEDICATION

This research work is dedicated to Almighty God for His protection, provision and the knowledge bestown on the successful completion of the programme.

DECLARATION

I, Abisoye, Opeyemi Aderiike, hereby declare that this thesis entitled Hybridized Support Vector Machine_Artificial Neural Network Malaria Prediction Model is a record of my research work. It has neither been presented nor accepted in any previous application for a higher degree. All sources of information have been specifically acknowledged.

In addition, the research work has been ethically approved by the University Ethical Review Committee.

O.A Abisoye

Date

ACKNOWLEDGEMENTS

I hereby express my profound gratitude to Almighty God for His mercy, unquantifiable protection and care for the successful completion of this Ph.D. programme.

My gratitude also goes to my able Supervisor, Dr. Rasheed Gbenga Jimoh, for his mentorship and usual supports given to me. May the Lord continue to lift you high and your forthcoming generations.

My Sincere and profound gratitude goes to my lovely, caring husband, Engr. Abisoye Blessing who always release me for the programme and my lovely children, Queen- Esther and Queen-Elizabeth. Also, my gratitude to my eldest sister, Dr. Mrs. Fadipe Joseph, Olubunmi for her supports and the entire, Amusan and Abisoye Family.

Finally I am personally grateful to my colleagues both at University of Ilorin and Federal University of technology, Minna, Niger State. More especially, to my honourable HOD, Computer Science, Dr. Bashir Abdullahi and Prof. Musa Aibinu, HOD Mechatronics Department Federal University of technology, Minna, Niger State for their supports. Also my gratitude goes to my Reverend, Rev. Dr. Ojekunle, mummy Ojekunle and my able Deacon, Jonathan Samuel. And to those whose name could not be mentioned due to space, I humbly say, thank you all.

ABSTRACT

The increase in the rate of malaria infections called for more rigorous research efforts. Evidences have shown that complexities involved in malaria forecasting required computational approach. Most existing conventional malaria forecasting models examine the dynamics of binary cases of asymptomatic of malaria parasite counts in the thick blood smear. Such existing models are plagued with the problem of over-fitting and prone to local minimum error due to large number of parameters to fix. Therefore, this study aimed at developing hybridized malaria prediction model for a multiclass nature of malaria parasite counts with effect of climatic conditions, given a non-linear malaria incidence cases. The objectives were to: (i) identify the factors that make an

individual susceptible to malaria infection and the threats imposed; (ii) determine likelihood qualitative model of Artificial Neural Network (ANN) and Support Vector Machine (SVM); (iii) develop a thresholded hybridized SVM_ANN malaria model for better performance; and (iv) simulate the model to evaluate its performance compared with existing models using classification accuracy (A_{cc}), sensitivity (S_s), specificity (S_p) and mean square error (mse).

Monthly survey of malarial incidence was collected from five randomly sampled health centers in Minna Metropolis, Niger State, Nigeria. Climatic data was also collected from the Nigerian Environmental and Climate Observation Programme (NECOP) Weather Station, Nigeria. These served as the model input variables. ANN with sigmoid transfer function for classification and SVM with radial basis function for feature selection were employed to predict the severity of multiclass malaria parasite counts.

The findings of the study were that:

- i. the internal factors of frequency of human blood index, duration of sporogony, vector density, vector susceptibility, demography and external factors of temperature, rainfall, relative humidity vegetation, altitude, human behavioural factors and other environmental changes contribute to the possibilities of asymptomatic, symptomatic and climatic based threats;
- ii. SVM feature selection method produced optimal features for ANN classification;
- iii. hybridized SVM_ANN malaria predicting model was developed at an optimum threshold 0.60 and functionally expressed as

$$SVM_ANN_{model} = \phi(b^{mat})$$

where $\phi(x)$ is the transfer function, $b^{mat} = \frac{1}{2}(w^{mat}, x_s + x_c)$,

$w^{mat} = \sum_{i=1}^n \alpha_i y_i = 0$, x_s, x_c are support vectors with constraints of langrange multipliers $\alpha_s, \alpha_c > 0$ and $\alpha_s, \alpha_c \in \alpha_i$ and targets $y_i \in [0, 1]$; and

- iv. SVM_ANN hybridized malaria model achieved a better 98.91% A_{cc} , 100% S_s , 98.68% S_p and 0.14 mse compared with existing ANN and SVM model with 48.33%, 85.60% A_{cc} ; 60.61%, 84.06% S_s ; 45.58%, 86.09% S_p and 0.56, 0.58 mse respectively.

The study concluded that the prediction of multiclass symptomatic malaria infection with effects of climatic conditions using SVM_ANN hybridized model was established to be more efficient than existing models. Thus, this study recommended that the SVM_ANN hybridized model should be adopted by medical personnel and stakeholders to predict malaria incidence occurrences and its severity.

TABLE OF CONTENTS

Page
TITLE PAGE
i

CERTIFICATION

ii

DEDICATION

iii

DECLARATION

iv

ACKNOWLEDGEMENT

v

ABSTRACT

vi

TABLE OF CONTENTS

vii

LIST OF TABLES

xiii

LIST OF FIGURES

xv

LIST OF EQUATIONS

xviii

CHAPTER ONE: INTRODUCTION

1.1. Background to the Study

1

1.2. Statement of the Problem

6

1.3. Justification of the Study

8

1.4. Aim and Objectives of the Study

9

1.5. Scope of the Study

9

1.6. Significance of the Study

10

1.7. Operational Definition of Terms

11

1.8.	Chapter Summary
	13

CHAPTER TWO: LITERATURE REVIEW

2.1	Introduction	14
2.1.1	Malaria	14
2.1.2	Causes of Malaria	16
2.1.3	Lifecycle of Malaria	17
2.1.4	Classes of Malaria	18
2.1.5	Species of Malaria	19
2.1.6	Clinical Symptoms of Malaria	21
2.1.7	Factors Affecting Malaria Transmission	22
	2.1.7.1 Internal Factor	22
	2.1.7.2 External factors	23
2.1.8	Malaria Threats	26
	2.1.8.1 Vector Threat	27
	2.1.8.2 Transmission Threat	27
	2.1.8.3 Asymptomatic/Morphological Based Threat	28

2.1.8.4 Symptomatic and Climatic Based Threat

30

2.2 Predicting Model

31

2.2.1 Mathematical Modeling of Infectious Disease Dynamics

32

2.2.1.1 Statistical-Based Method for Epidemic Surveillance

32

2.2.1.2 Mathematical State Space Model

34

2.2.1.3 Machine Learning Based Model

37

2.3 Malaria Based Models

38

2.3.1 General Based Model

38

2.3.2 SVM Based Model

39

2.3.3 ANN Based Model

41

2.3.4 ANN and SVM Based Models

43

2.4 Review of Malaria Threat Model

45

2.5 Choice of Machine Learning Algorithm

49

2.5.1 Parametric System Learning Process

49

2.5.2 Types of Learning

50

2.5.3 Main Components of Learning

51

2.6 Strength and Weaknesses of Various Computational

53

Intelligence Tools

2.7 Artificial Neural Network (ANN) Model

56

2.7.1 Neuron

56

2.7.2 Network Structures

58

2.7.3 Back Propagation Neural Network(BPN)

59

2.7.4 Feed Forward Neural Network(FFN)

60

2.8 Artificial Neural Network Activation Functions

61

2.8.1 Purelin Function

62

2.8.2 Sigmoid Function

62

2.8.3 Logistic Sigmoid

63

2.8.4 Tangential Sigmoid

63

2.9 Support Vector Machine Model

64

2.9.1 Theory Support Vector Machine

65

2.9.2 Support Vector Machine Classification

67

2.9.3 Support Vector Machine Multiclass Classification

68

2.9.4 Support Vector Machine Feature Selection

68

2.10 Support Vector Machine Kernel Function

71

2.10.1	Linear Kernel Function	71
2.10.2	Gaussian (Radial Basis) Function	71
2.10.3	Polynomial Function	72
2.10.4	Sigmoid Function	73
2.11	Review of Related Work	74
2.12	Chapter Summary	76
 CHAPTER THREE: RESEARCH METHODOLOGY		
3.1	Introduction	77
3.2	Research Design	77
3.2.1	System Analysis	78
3.2.2	Likelihood Quantitative Model	78
3.2.3	ANN and SVM Computational model	79
3.2.3.1	The Flow Diagram of Architecture ANN to Predict Malaria Parasite Count	80
3.2.3.2	The Flow Diagram of Architecture SVM to Predict Malaria Parasite Count	80
3.2.3.3	Research Design of the SVM_ANN Hybridized Malaria to Predict Severity of Malaria Cases	83

3.3 Population Sample and Sampling Procedure	84
3.4 SVM Malaria Model	85
3.5 Instrument for Data Collection and Preparation	90
3.6 Data Analysis Method	91
3.6.1 Preprocessing	91
3.6.1.1 Outlier Removal	91
3.6.1.2 Feature Normalization	91
3.6.1.3 Missing data	92
3.6.2 Feature Selection	92
3.7 Computation of SVM_ANN Qualitative Measures	93
3.8 Multiclass Encoding and Normalization of Dataset Features	93
3.9 Hybridized SVM_ANN Malaria Model Descriptions	94
3.9.1 Error Function	96
3.9.2 Weights from Different Inputs	97
3.10 Hybridized SVM_ANN Malaria Model Quantitative Measure	98
3.11 Adaptive Static Thresholding and Feature Selection Algorithm	102
3.12 Malaria Threat and Severity Algorithm	104

3.13	Analysis of Performance	105
3.13.1	ANN Analysis of Performance	105
3.13.2	SVM Analysis of Performance	107
3.14	Chapter Summary	109

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1	Introduction	110
4.2	Data Analysis	110
4.2.1	Raw Malaria Data	110
4.2.2	Preprocessed Malaria Data	111
4.2.3	Multiclass Encoded and Normalized Dataset Features	112
4.3	ANN Processing Results and Models	112
4.3.1	ANN Target and Simulated Test Result	117
4.3.2	ANN Performance Evaluation	118
4.4	SVM Processing Results and Models	121
4.4.1	SVM Feature Selection Results	122

4.4.2	SVM Target and Simulated Test Result	128
4.4.3	SVM Performance Evaluation	128
4.5	SVM_ANN Adaptive Static Thresholding	131
4.6	Weaknesses and Strength of SVM and ANN	132
4.7	Hybridized SVM_ANN Malaria Model Result	133
4.8.1	Overall Best Results for ANN, SVM and SVM_ANN Hybridized Malaria Model	135
4.8	SVM_ANN Hybridized Malaria Model Formulation	136
	Model Quantitative Measure	137
4.9.1	SVM_ANN Hybridized Malaria Model Result Discussion	138
4.9	Performance Evaluation	138
4.10	Chapter Summary	140

CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATION

5.1	Summary	141
5.2	Conclusion	142
5.3	Recommendation	143

5.4	Future Work	143
5.5	Contribution to Knowledge	144
5.6	List of Published Work	144
5.7	Acknowledgements	145
	References	146

Appendix A

Table 4.3	Threshold Values and Performance for the ANN	164
	Malaria Training Model	
Table 4.5	Threshold Values and Performance for the ANN	166
	Malaria Testing Model	
Table 4.7	Threshold Values and Performance for the ANN	168
	ANN Validation Model	

Appendix B

Table 4.10	ANN Target and Simulated Test Result	170
------------	--------------------------------------	-----

Appendix C

Table 4.14	SVM Target and Simulated Test Result	173
Table 4.15	SVM Malaria Model Optimal Features	176

Appendix D

Table 4.17	Hybridized SVM_ANN Malaria Model Target	183
	and Simulated Test Results	

LIST OF TABLES

	Page	P
Table 2.1	Malaria Threat Model	
45		
Table 2.2	Strength and Weaknesses of Various Computational Intelligence	
53		

	Model Tools
Table 2.3	Related Work
74	
Table 3.1	Malaria model Features Description
91	
Table 3.2	Multiclass Encoding Threat Classes
94	
Table 3.3	Adaptive Thresholding Algorithm
103	
Table 3.4	SVM Feature Selection Algorithm
104	
Table 3.5	Malaria Threat and Severity Algorithm
104	
Table 3.6	Confusion Matrix
105	
Table 4.1	Sampled Raw Malaria Data
111	
Table 4.2	Sampled Preprocessed Malaria Data
111	
Table 4.4	Best Threshold Values and Performance for the
113	
	ANN Malaria Model Training
Table 4.6	Best Threshold Values and Performance for the
115	
	ANN Malaria Model Testing
Table 4.8	Optimum Threshold Values and Performance for the
117	
	ANN Malaria Model
Table 4.9	Best Threshold Values and Performance for the
116	
	ANN Validation Malaria Model
Table 4.11	ANN Classifier Result
118	

Table 4.12	SVM_0 Feature Selection Result	122
Table 4.13	SVM_1 Feature Selection Result	124
Table 4.16	SVM_2 Feature Selection Result	126
Table 4.18	Weaknesses and Strength of SVM and ANN	132
Table 4.19	SVM_ANN Hybridized Malaria Model	133
Table 4.20	Overall Best Results for ANN, SVM and SVM_ANN Hybridized Malaria Model	135
Table 4.21	Root of Equation	136
Table 4.22	Performance Comparison between SVM_ANN Hybridized Malaria Model and Baseline Works	139
Table 4.23	Performance of SVM_ANN Hybridized Malaria Model	140

LIST OF FIGURES

	Page
Figure 2.1 A mosquito	
16	
Figure 2.2 Lifecycle of Malaria	
17	
Figure 2.3 P. Normal RBC, Platelets and White Blood Cells	
18	
Figure 2.4 P. Falciparum	
19	
Figure 2.5 P. Malariae	
19	
Figure 2.6 P. Vivax	
19	
Figure 2.7 P. Ovale	
19	
Figure 2.8 Variations of Temperature in the Study Time	
24	
Figure 2.9 Variations of Rainfall Figure in the Study Time	
25	
Figure 2.10 Variations of Relative Humidity in the Study Time	
25	
Figure 2.11 An Overview of Mathematical Modeling	
33	

of Infectious Disease

Figure 2.12	Parametric System Learning Process
49	
Figure 2.13	Basic Set-up of Learning Problem
52	
Figure 2.14	Modeling Errors.
53	
Figure 2.15	Human Body Network of Neurons
56	
Figure 2.16	A Perceptron
56	
Figure 2.17	Multilayered Artificial Neural Network
57	
Figure 2.18	Malaria Classification Schema
61	
Figure 2.19	Linear Transfer Function
62	
Figure 2.20	Log Sigmoid Transfer Function
63	
Figure 2.21	Tan Sigmoid Transfer Function
64	
Figure 2.22	Linearly separable Two Classes of Data
67	
Figure 2.23	Optimal Separating Hyperplane
67	
Figure 2.24	SVM Algorithm
70	
Figure 3.1	Methods and Steps in Data Mining Process
79	
Figure 3.2	The Flow Diagram of the Architecture of ANN Classifier
81	
Figure 3.3	The Flow Diagram of the Architecture of SVM Selector
82	

Figure 3.4	Research Design SVM_ANN Hybridized Malaria Model	83
Figure 3.5	SVM& ANN Qualitative Measure	93
Figure 3.6	A Fully Connected Schematic Description of Hybridized SVM_ANN Malaria Model	95
Figure 3.7	Weights from Different Inputs	98
Figure 4.1	Thresholded Values and Performance on ANN Malaria Training Model (1)	113
Figure 4.2	Thresholded Values and Performance on ANN Malaria Training Model (2)	113
Figure 4.3	Best Thresholded Values and Performance on ANN Malaria Training Model	113
Figure 4.4	Thresholded Values and Performance on ANN Malaria Testing Model (1)	114
Figure 4.5	Thresholded Values and Performance on ANN Malaria Testing Model (2)	114
Figure 4.6	Best Thresholded Values and Performance on ANN Malaria Testing Model	115
Figure 4.7	Thresholded Values and Performance on ANN Malaria Validation Model (1)	116
Figure 4.8	Thresholded Values and Performance on ANN Malaria Validation Model (2)	116
Figure 4.9	Best Thresholded Values and Performance on ANN	117

Malaria Validation Model

Figure 4.10	ANN_0 Results
118	
Figure 4.11	ANN_1 Results
118	
Figure 4.12	ANN_2 Results
119	
Figure 4.13	Graph of Accuracy Vs Support Vectors
123	
	for ‘SVM_0’ Malaria cases
Figure 4.14	Graph of Accuracy Vs Mean Square Error
123	
	for ‘SVM_0’ Malaria cases
Figure 4.15	<i>SVM_0 (lin)</i>
123	
Figure 4.16	<i>SVM_0 (rbf)</i>
123	
Figure 4.17	<i>SVM_0 (pol)</i>
124	
Figure 4.18	Graph of Accuracy Vs Support Vectors
125	
	for ‘SVM_1’ Malaria Cases
Figure 4.19	Graph of Accuracy Vs Mean Square Error
125	
	for ‘SVM_1’ Malaria Cases
Figure 4.20	<i>SVM_1 (lin)</i>
125	
Figure 4.21	<i>SVM_1 (rbf)</i>
125	

Figure 4.22	<i>SVM_1 (pol)</i>
126	
Figure 4.23	Graph of Accuracy Vs Support Vectors
127	
	for ‘SVM_2’ Malaria Cases
Figure 4.24	Graph of Accuracy Vs Mean Square Error
127	
	for ‘SVM_2’ Malaria Cases
Figure 4.25	<i>SVM_2 (lin)</i>
127	
Figure 4.26	<i>SVM_2 (rbf)</i>
127	
Figure 4.27	<i>SVM_2(pol)</i>
128	
Figure 4.28	Best Results for ANN_0 and SVM_2 Malaria Model
131	
Figure 4.29	SVM_ANN Hybridized Malaria Model Confusion Matrix
134	
Figure 4.30	Performance of ANN, SVM and SVM_ANN Hybridized
136	
	Malaria Model

LIST OF EQUATIONS

Equation (2.1)

27

Equation (2.2)

27

Equation (2.3)

53

Equations (2.4 - 2.6)

57

Equation (2.7)

61

Equation (2.8)

63

Equations (2.9 - 2.10)

65

Equations (2.11 - 2.12)

66

Equation (2.13)

71

Equations (2.14 – 2.16)

72

Equation (2.17)

73

Equation (3.0)

84

Equations (3.10 – 3.14)

85

Equation (3.15 – 3.19)

86

Equations (3.20 – 3.25)

87

Equations (3.26 – 3.30)

88

Equations (3.31 – 3.35)

89

Equations (3.36 – 3.37)

94

Equations (3.38 – 3.39)

96

Equation (3.40)

97

Equations (3.41 – 3.46)

99

Equations (3.47 – 3.51)

100

Equations (3.52 – 3.54)

101

Equations (3.55 – 3.56)

102

Equations (3.57 – 3.62)

106

Equations (3.63– 3.64)

108

CHAPTER ONE

INTRODUCTION

This chapter introduces the concepts of the study, the critical background to the study, statement of the problem, research aim and objectives, scope of the study, significance of the study, limitation of the research, definition of operational terms and chapter summary.

1.1. Background to the Study

Health is the most important aspect of human life. Poor health condition affects one's life span and achievements. Several diseases affect the proper functioning of human's health; one of such common diseases is malaria. Malaria is a parasitic infectious disease caused by *Plasmodium Species (Spp)*. It was revealed that 200 to 300 million of people are infected by malaria every year and about 3 million people face fatal death yearly due to malaria (WHO, 2014). Malaria is a major public health issue since virtually all humans are being diagnosed of malaria infections at one time or the other and its effects are monumental; thus causing great harm to human lives. Across 95 countries, about 3.2 billion people are at risk and 1.2 billion are at high risk (WHO, 2015). There are an estimated 250 million cases of malaria yearly leading to approximately (1,000,000) one million deaths (WHO, 2014).

Malaria is a parasitic infection being caused by a parasite known as *Plasmodium spp* and transmitted by a vector known as *Anopheles mosquito* (Mueller, Galinski, Baird, Carlton, Kochar, Alonso & Portillo, 2009). The parasites invade the blood and causes adverse effect on the blood cells. Within 48 to 72 hours, the parasites

multiply inside the red blood cells and break open, infecting more red blood cells (Thomas, 2014). The first symptoms usually occur between 10-14 days to 4 weeks after infection (Bannister & Mitchell, 2003). Malaria parasites can also be transmitted from a mother to her unborn baby (congenitally) by blood transfusions and by sharing needles used during injection (Singh, 2006). In some parts of the world, malaria parasites have developed resistance to insecticides and antibiotics (White, 2004).

Recently, the WHO has given priority to malaria reduction and prevention through several aid agencies and donors (Brown, Cueto & Fee, 2006; WHO, 2015). Hundreds of millions of dollars has been donated by Bill and Melinda Gates Foundation who called malaria “the worst thing on the planet”. Also, the Bush Administration pledged 1.2 billion dollars (Finkel, 2007). Funds devoted to malaria prevention have doubled since 2003 (WHO, 2015). The idea is to combat the disease headlong by combining virtually every known malaria-fighting techniques, from the ancient Chinese herbal medicines to the old bed nets and to the ultramodern multidrug cocktails such as artemisinin-based combination therapies and indoor residual spraying with insecticide to control the vector mosquitoes. Likewise, malaria researchers are pursuing a vaccine and methods that would curb the disease for good (Cherif, Movahedzadeh, Michel, Hill & Jedlicka, 2011).

Malaria transmission is site specific due to different climatic variations of a region (Yusuf, 2013). Temperature, rainfall, relative humidity fluctuations affects the life cycle of malaria parasite (*Depinay, Mbogo, Kileen, Knols, Beier, Calson & McKenzie, 2004*). Several non-climatic factors, such as human/behavioural

factors can also affect the pattern of malaria transmission and severity (Randolph, 2008).

Temperature affects the life cycle of the malaria parasite (CDC, 2015). The lifespan of the parasite in the gut of the mosquito is about 10 days (Moreira, Iturbe-Ormaetxe, Jeffery, Lu, Pyke, Hedges, & Hugo, 2009). Development of the mosquito larva also depends on temperature – The ranges of minimum and maximum temperature greatly affect the development of the malaria parasite and its mosquito vector, which determines malaria transmission (Martens, Niessen, Rotmans, Jetten, & McMichael, 1995). Malaria parasite develops more quickly at higher temperatures (McMichael, Woodruff, & Hales, 2006). Higher temperature increases the number of blood meals taken and the number of eggs laid by the mosquitoes, which increases the number of mosquitoes in a given area (Yang, Macoris, Galvani, Andrighetti, & Wanderley, 2009). The minimum temperature for mosquito development is between 8–10°C; the optimum temperature is 25–27°C, and the maximum temperature is 40°C (Harvell, 2002).

Rainfall affects the breeding habitat of mosquitoes. Too much rainfall flushes away breeding habitats temporarily, but mosquitoes start breeding as soon as the rain stops (Koenraadt, & Harrington, 2008). Flushing occurs frequently in highlands and hilly areas than in the lowland plains, so a bigger impact on vector breeding habitats. Not all water collections are suitable for the mosquito life cycle. Vegetation cover increases after rainfall, which in turn increases the relative humidity of the environment (Yusuf, 2013).

Relative humidity is the amount of moisture in the air, expressed in percentage; (0% humidity would mean the air is completely free of moisture and 100%

humidity would mean the air is completely saturated with moisture). Mosquitoes need to live at least 8–10 days to be able to transmit malaria thus mosquitoes survive better under conditions of high humidity (Thomas & Read, 2007).

The other non-climatic factor that contributes to malaria is *human behaviour* (Craig, Kleinschmidt, Le Sueur, & Sharp, 2004). Travelers from non-endemic areas may choose not to use insect repellent or medicines to prevent malaria. Reasons include cost, inconvenience, or a lack of knowledge. Poor rural populations in malaria-endemic areas often cannot afford the housing and bed nets that would protect them from exposure to mosquitoes. Human activities can create breeding sites for larvae (standing water in irrigation ditches, burrow pits). Agricultural work such as harvesting (also influenced by climate) may force increased night time exposure to mosquito bites. Raising domestic animals near the household may provide alternate sources of blood meals for *Anopheles* mosquitoes and thus decrease human exposure (Cohen & Gürtler, 2001). The governments in malaria-endemic countries often lack financial resources (Mutabingwa, 2005). As a consequence, health workers in the public sector are often underpaid and overworked. They lack equipment, drugs, training, and supervision. So prediction of the possibility of malarial occurrences in the nearest future is essential so as to take preventive measure by individuals and governments.

Dynamics of malaria parasite transmission is complex and have been widely studied. Research is ongoing on how the risk of asymptomatic and symptomatic malaria infection changes (Alonso, Brown, Arevalo, Binka, Chitnis, Collins & Mendis, 2011; Bousema, Okell, Felger & Drakeley, 2014). Malaria diagnosis can

be asymptotically or symptomatically low, mild and high. Many symptoms of different patients may even overlap. A malaria patient may even have characteristics of other diseases. Therefore, medical problems cannot be generalized and analyzed by imagination. An analytical programme should be desperately conducted to integrate this complex network of problems and devise individualized solutions (Onuwa, 2014).

Diagnosing asymptomatic malaria transmission is not straightforward due to the obvious lack of clinical manifestations and often sub-patient levels of parasites are undetectable under microscopes (Bottius, 1996). The extent of the prevalence of asymptomatic parasitemia in a given population is inversely related to the population's susceptibility to clinical disease (Bereczky Liljander, Rooth, Faraja, Granath, Montgomery & Fanert, 2007; Mwangi, Fegan, Williams, Kinyanjui & Snow, 2008). Asymptomatic malaria infection cases are prevalent in malaria endemic regions and have become a serious cause for concern as efforts are increasing towards eliminating the parasite (Trape, Zoulani & Quinet, 1987). The presence of both symptomatic and asymptomatic diagnostic measure is very vital in detecting the transmission dynamics of malaria infection (Wu, Hoogen, Slater, Walker, Ghani, Drakeley & Okell, 2015). To avoid the occurrence of new malaria outbreaks in both endemic and non-endemic areas, improved methods are required to decrease the parasite sources of infection by active prediction and treatment of symptomatic and asymptomatic parasite carriers (Patel, Shortliffe, Stefanelli, Szolovits, Bellazzi, & Abu-Hanna, 2009).

There is a huge amount of data which is hard to understand and to interpret by humans difficulty arises; a typical example is malarial incidences (Keeling &

Rohani, 2008). Therefore, the need for a machine learning method arise to processes the data and automatically find structures in the data (Witten, Frank, Hall & Pal, 2016). The knowledge about the extracted structure can be used to solve the problem at hand. Problems being solved by machine learning methods range from classifying observations, predicting values, structuring data (e.g. clustering), compressing data, visualizing data, filtering data, selecting relevant components from data, extracting dependencies between data components, modeling the data generating systems, constructing noise models for the observed data, integrating data from different sensors, using classification and drawing inferences. Thus, machine learning focuses on prediction based on known properties learned from the trained data sets (Nilsson, 2005).

1.2 Statement of the Problem

The rate at which malaria prevails in the concerned areas seems exponential despite the fact that several strategies and therapies have been proposed to combat the spread and even eradicate the disease (Owoseni & Ogundahunsi, 2016). Virtually all humans are being diagnosed of malaria infections at one time or the other especially in tropical area in Africa and its negative effects increase drastically and cause great harm to human life (Mendis, Sina, Marchesini & Carter (2001).

Globally, recent researches are focusing on the new applications in the medical field and particularly in prediction and diagnosis (Oguntimilehin & Adetumbi, 2015). Medical experts are now demanding for an intelligent, knowledge-based

system in finding solutions to medical problems (Djam, Wajiga, Kimbi & Blamah, 2011). Most of the existing conventional malaria prediction models examine the dynamics of asymptomatic and morphological characteristics of malaria parasite (Khoshmanesh, Dixon, Kenny, Tilley, McNaughton & Wood, 2014). The binary cases of malaria incidences were examined in the literature whereas the multiclass nature of malaria cases is still yet unaddressed. One of the great determinants of malaria severity is the prediction of multiclass symptomatic characteristics of malaria parasite combined with effects of climatic factor which is the focus of this study.

Several models have been proposed to forecast the incidence of malaria in different locations (Smith, Killen, Maire, Ross, Mollineaux, Tediosi & Tanner, 2006). Various old time models like Auto-Regressive Integrated Moving Average (ARIMA), Auto-Regressive Moving Average (ARMA) and Auto-Regressive (AR) were in existence to forecast malaria (Box & Jenkins, 1976; Ch, Sohani, Kumar, Malik, Chahar, Nema & Dhiman, 2014). These models fail to address the non-linear nature of malaria which is of geometric progression (Anderson, May & Gupta, 1989). Developing proper algorithms and approaches to achieve further reliability, accuracy and convenience in computational process is still highly demanding (Mohammadi, Shamshirband, Tong, Arif, Petkovic & Ch, 2015).

Existing models are prone to local minimum error and they suffer from over-fitting problem due to a large number of parameters to fix and the little prior user knowledge about the relevance of the inputs in the analyzed problem (Hegazy, Soliman & Salam, 2013). Also, to produce an efficient computing systems, an accurate, reliable malarial incidences predicting model is required (Patel,

Shortliffe, Stefanelli, Szolovits, Bellazzi, & Abu-Hanna, 2009; Abeku, Hay, Ochola, Langi, Beard, de Vlas, & Cox, 2004). The quest for higher prediction accuracy performance of malaria incidences is of great importance (Ch *et al.*, 2014; Munakata, 2008). This study therefore focuses on the prediction of multiclass, symptomatic and climatic based models that are not prone to local minimum error and free from over-fitting.

1.3 Justification of the Study

Recent researches globally focus on the new applications in the medical field and particularly in prediction and diagnosis (Oguntimilehin & Adetumbi, 2015). Medical experts are now demanding for an intelligent, knowledge based system to be knowledge intensive in finding solutions to medical problems (Djam, Wajiga, Kimbi & Blamah, 2011). The prediction of the multiclass symptomatic characteristics of malaria parasite combined with effects of climatic factor a great determinant of malaria severity is of great importance. Also, the new system incorporates feature selection and global thresholding to avoid over-fitting and local minimum error.

1.4 Aim and Objectives of the Study

The aim of this research is to develop hybridized malaria prediction model based on Support Vector Machine and Artificial Neural Network to predict the severity of malaria parasite counts given a non-linear malaria incidence cases. Specific objectives are to:

1. identify the factors that make an individual susceptible to malaria infection and the threats of the infection;
2. construct appropriate relationships in (1) above to determine likelihood qualitative model to predict malaria incidence;
3. develop a hybridized SVM_ANN malaria model for predicting the severity of malaria parasite count; and
4. simulate the model to evaluate its performance compared with existing models using classification accuracy (A_{cc}), sensitivity (S_s), specificity (S_p) and mean square error (mse).

1.5 Scope of the Study

This study covers the prediction of multiclass malaria based on symptoms and climatic factors. The study covers the monthly surveys of malarial incidences from sampled health centers in Minna metropolis, Niger State, Nigeria. Also, monthly averages of rainfall, temperature and relative humidity of those areas with monthly malaria incidences were considered as influential climatic factors. The asymptomatic and morphological characteristics of malaria cases were not considered.

1.6 Significance of the Study

The cost of providing local and advanced medical equipment and health care services may not be affordable by some citizens and especially people in the underdeveloped and rural areas. Sometimes, prescription of drugs by some medical experts is based on intuitions. Medication should be administered based on age and severity of the disease. This has led to poor healthcare services especially in our rural areas. Huge amount of malaria incidences are captured every year and there are difficulties in predicting its occurrence and analysis of its possible threats. The need for an alternative cheap, fast healthcare hybrid solution that will employ the unique features of the two models and complement the weakness and strength of each other is of great importance. This research will benefit patient, medical personnel and any individual to perform malaria severity prediction in remote or underdeveloped areas. Taking effective decision like appropriate medication to prescribe will be guarded by the model. Also this research will benefit the patients by reducing the waiting response time being spent in the hospital.

1.7 Operational Definition of Terms

Hybrid: Combination of two or more techniques to form a single technique and each technique performs different functions.

Threat: A possible danger that might exploit a vulnerability to breach security and thus cause possible harm.

Factor: A constituent or element that brings about certain effects or results, or indicates a specific multiple, number or quantity.

Infection: The invasion of an organism by a pathogen such as bacteria, viruses, or parasites. Some, but not all, infections lead to disease.

Malaria: Disease caused by malaria parasite (*plasmodium spp*).

Anopheles: specie of mosquito, which transmits malaria parasites into blood cells.

Parasite: Any organism that lives in or on another organism without benefiting the host organism; commonly refers to pathogens, most commonly in reference to protozoans and helminthes.

Plasmodium Spp.: The genus of the parasite that causes malaria. The genus includes many species. The four species that naturally infect humans are *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale* and *Plasmodium malariae*.

Predict To state, tell about, or make known in advance, especially on the basis of special knowledge.

Model: is a representation of a proposed structure or thing.

Symptomatic: Prediction that is based on clinical symptoms emanated from malaria cases such as headache, fever, dizziness, body pain etc.

Asymptomatic: Prediction that is based on behavioural patterns emanated from malaria cases such as size of infected red blood cells, shape of parasite, number of

chromatin dots, texture of red blood cells, location of chromatin and number of parasite

Climatic: Patterns influenced by rainfall, temperature, relative humidity and other climate factors

Classification: to predict the category of some input values. For instance, to predict the class of patient that will have malaria in the nearest future.

Regression: to predict the value of some input values; for instance, to predict whether a patient will have malaria in the nearest future.

Quinine: A drug used against malaria, obtained from the bark of the cinchona tree. Quinine is used for treatment but not prevention of malaria.

Vaccine: A preparation that stimulates an immune response that can prevent an infection or create resistance to an infection.

Vector: An organism (e.g., *Anopheles* mosquitoes) that transmits an infectious agent (e.g. malaria parasites) from one host to the other (e.g. humans).

Virus: A microorganism composed of a piece of genetic material - RNA or DNA - surrounded by a protein coat. To replicate, a virus must infect a cell and direct its cellular machinery to produce new viruses.

LLINs: long-lasting insecticidal nets

IRS: indoor residual spraying

1.8 Chapter Summary

Strategies for fighting infectious disease rely on vector control, transmission blocking and medical assistance. The vast outbreak of malaria threat on human has not received adequate therapy or solutions. This chapter identifies the malaria threats, determinant factors and proposed two machine learning methods that serve as an alternative Healthcare Solution.

The introduction covers the background to the study, motivation of the study, problem statement, research aim, research objectives, scope of the study, significance of the study, limitation of the study, definition of operational terms and chapter summary. The next chapter gives the literature review of the study. Chapter three presents the methodology adopted for the research while chapter four discusses the results obtained for the research, analysis and evaluation of the model with the existing models. Finally chapter five gives the concluding remark, makes some recommendation and future prospects of the research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter contains review of concepts of malaria, factors affecting malaria transmission and malaria threats. Related works on malaria threat, strength and weaknesses of various computational intelligence tools and baseline models were reviewed. then the rationale for the choice of machine learning algorithms is made.

2.1.1 Malaria

Infections are diseases caused by virus, bacterial and fungal organisms. Such diseases could be sexually transmitted, air borne or water borne, examples are glaucoma, tuberculosis, measles, chickenpox, influenza, malaria and so on. In the recent decades; emerging and re-emerging infections such as AIDS, measles, malaria, and tuberculosis cause death to millions of people each year. Problem arises mainly from two reasons: the continuous and perennial mutations of the viruses, and the complexity in the disease transmission mechanism (Siettos & Russo, 2013).

In India, there are approximately 20,000,000 confirmed cases of malaria yearly. Also, malaria is a major public health challenge in our country, Nigeria (MDHIL, 2014). In Nigeria, on the average, each Nigerian suffers at least two or more attacks of malaria every year. Although millions recover, hundreds of thousands are not so lucky. Malaria accounts for about 60 percent of outpatient visits and 30

percent of hospitalizations; 25 percent of deaths in children under one year old; and 11 percent of maternal deaths —a heavy burden on Nigerian families, communities, health system, and workforce (Sola & Chioma, 2010).

A scoping review of malaria prediction in past work and future directions was presented in 2013. The motivation behind the review was the growing body of literature on malaria prediction methods and the objective of our review is to identify and assess methods, including predictors, used to forecast malaria 29 different studies that met the inclusion criteria. Applying different prediction methods to the same data, exploring the predictive ability of non-environmental variables, including transmission reducing interventions and using common forecast accuracy measures will allow malaria researchers to compare and improve models and methods, which should improve the quality of malaria prediction. The limitation to the study was that the forecast accuracy of different models from the different studies was not reported as the evaluation measures differed across the studies (Zinszer Verma, *Charland*, Brewer, Brownstein, Sun & Buckeridge, 2013).

The Global Technical Strategy for Malaria 2016-2030 was adopted by the World Health Assembly in May 2015. It provides a comprehensive framework to guide countries in their efforts to accelerate progress towards malaria elimination. The strategy sets the target of reducing global malaria incidence and mortality rates by at least 90% by 2030. It emphasizes the need for universal coverage of core malaria interventions for all populations at risk and highlights the importance of using high-quality surveillance data for decision-making. It also identifies areas

where innovative solutions will be essential for attaining the goals, and summarizes the estimated global costs of implementation (WHO, 2015).

2.1.2 Causes of Malaria

a. Mosquito: Malaria is caused by a parasite known as *Plasmodium* specie being transmitted by *Anopheles mosquito*. The parasite invades the blood as shown in Figure 2.1 and causes adverse effect on the blood cells. Within 48 to 72 hours, the parasites multiply inside the red blood cells and break open, infecting more red blood cells. The first symptoms usually occur between 10-14 days to 4 weeks after infection (Thomas, 2014).



Figure 2.1: A mosquito (Thomas, 2014)

b. Cogenital: Malaria parasites can also be transmitted from a mother to her unborn baby (Marchesini & Crawley, 2004).

c. Transfusion: Malaria parasite can be transmitted by blood transfusions and by sharing needles used to inject drugs (Dodd, 1998; Kitchen & Chiodini, 2006).

2.1.3 Lifecycle of Malaria:

When a mosquito takes a blood meal, it injects sporozoites into the blood cells which is later transmitted to the liver cells. The sporozoites infect the liver cells and mature to schizonts. Schizonts thus release merozoites. Merozoites can undergo a dormant stage (hypnozoites) and persist in the liver. Then they undergo sexual reproduction and form gametocytes.

Mosquito is a P. Malaria vector. **Gametocytes** are ingested by an Anopheles mosquito during a blood meal. This parasite's multiplication in the mosquito is known as sporogonic cycle. Mating of micro and macrogametes produces zygotes which become elongated and form **ookinetes** which invade the midgut wall and matured to **Oocytes**. The **Oocytes** grow and release sporozoites that are transmitted to the salivary gland. The sporozoites' inoculation to a new human host completes the life cycle as shown in Figure 2.2.

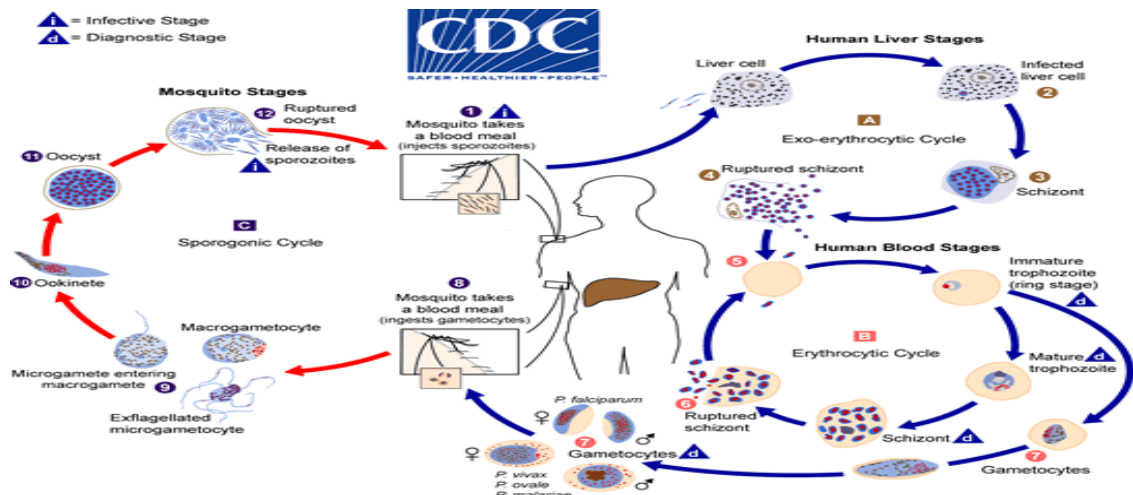


Figure 2.2: Lifecycle of Malaria (CDC, 2015)

2.1.4 Classes of malaria

There are two classes of Malaria:

- a. ***Benign malaria:*** Benign malaria is milder and relatively easy to treat. A very high temperature called "*fever*" is the most common symptom of Benign malaria. Benign malaria are also of two types : Benign Tertian, the fever occurs every second day with the "tertian" parasite - *P. malariae* (Mali, Steele, Slutsker & Arguin, 2009). Benign Quartan, the fever occurs every third day with the "quartan" parasite (*P. malariae*) (Mali et al; 2009). Other symptoms of benign malaria may include headaches, body-pain, cough, dizziness and feeling generally unwell. Diarrhoea and vomiting can occur especially in children (Thomas, 2014).

- b. ***Malignant malaria:*** Malignant malaria can be very severe and can sometimes be fatal. Malignant malaria usually begins with benign malaria and later progressed to complicated breathing problems, liver failure and shock. Malignant malaria can also affect the brain and central nervous system, and thus lead to death (Thomas, 2014).

2.1.5 .Species of Malaria

Human blood has three kind of cells; Red Blood Cells (RBCs), White Blood Cells(WBCs) and Blood Platelets. Malaria affects the RBCs, infects the liver and multiplies. Figure 2.4 -2.7 depicts the Normal WBC and species of affected red blood cells respectively

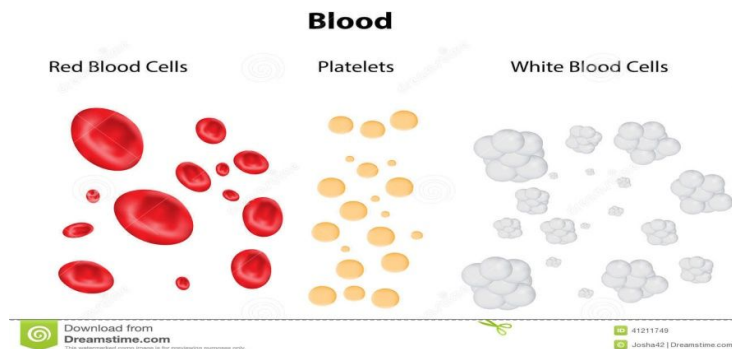


Figure 2.3 P. Normal RBCs, Platelets and White Blood Cells.



Figure 2.4: P. Falciparum

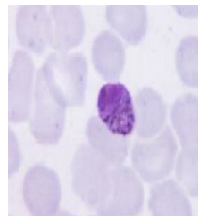


Figure 2.5: P. Malariae

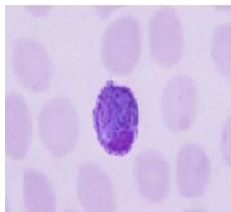


Figure 2.6: P. Vivax

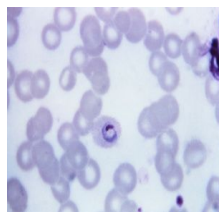


Figure 2.7: P. Ovale

Source: Gunder & Martin, 2011

There are five major types of malaria (Nordqvist, 2013)

a. *Plasmodium falciparum* (*P. falciparum*) – It is the most serious form of the disease. It is most common in Africa, especially sub-Saharan Africa. *P. Faliparum*

is shown in Figure 2.4. Current data indicates that cases are now being reported in areas of the world where this type was thought to have been eradicated.

b. *Plasmodium malariae (P. malariae)* - Not all that fatal, milder form of the disease. However, the infected human still needs treatment because lack of treatment can also lead to a host of health problems. *Plasmodium malariae* parasite has been known to stay in the blood of some people for several decades. It is shown in Figure 2.5.

c. *Plasmodium vivax (P. vivax)* – Globally, this type has the widest geographic distribution. It is a milder form of the disease, generally not fatal. However, infected people still need treatment because if untreated it can also cause a host of other health problems. About 60% of infections in India are due to *P. vivax*. This parasite has a liver stage and can remain in the body for years without causing sickness. It is shown in Figure 2.6. If the patient is not treated, the liver stage may re-activate and cause malaria attacks after months, or even years without symptoms. It is the most dangerous form of malaria, with the highest rates of complications and mortality.

d. *Plasmodium ovale (P. ovale)* – This is a milder form of the disease which is generally not fatal. However, the infected human still needs to be treated because it may progress and cause a host of health problems. This parasite has a liver stage and can remain in the body for years without causing sickness. It is shown in Figure 2.7. Without treatment, there is a risk that the liver stage can re-activate and cause relapses after very long periods without symptoms. Recently, *P. ovale* has been shown genetically to consist of two subspecies, *P. ovale curtisi* and *P. ovale wallikeri*. It causes tertian malaria in humans.

2.1.6 Clinical Symptoms of Malaria

All the clinical symptoms associated with malaria are caused by the asexual erythrocytes or blood stage parasites. When the parasite develops in the erythrocyte, numerous known and unknown waste substances such as hemozoin pigment and other toxic factors accumulate in the infected red blood cell. These are dumped into the bloodstream when the infected cells lyse and release invasive merozoites. The hemozoin and other toxic factors such as glucose phosphate isomerase (GPI) stimulate macrophages and other cells to produce cytokines and other soluble factors which act to produce fever and rigours and probably influence other severe pathophysiology associated with malaria. The clinical symptoms of malaria are:

- a. **Headache:** Headache is a common symptom of malaria, either cerebral type or not. The cytokine is believed to be an important factor leading to headache in acute malaria. In addition, headache is one of the symptoms of post malaria neurologic syndrome. Headache usually lasts for the whole duration of illness (Suyaphun, Wiwanitkit, Suwansaksiri, Nithiuthai, Srita, Suksirisampant & Fongsungnern, 2002).
- b. **Fever:** Fever is a high temperature which comes and goes. The patterns of fever vary according to the species of malaria. However, malaria does not have a specific pattern. Initially, malaria presents itself like the flu with high fever, fatigue, and body aches, with high body temperature and cold stages. Signs and symptoms in children may be non-specific, leading to delays in diagnosis.

- c. ***Dizziness***: Some patients may present with dizziness or vertigo, with or without fever. Associated with dizziness may also be vomiting and/or diarrhea (UNICEF, 2000).
- d. ***Body pain***: This is a feeling of general discomfort, uneasiness or pain, of being "out of sorts", often the first indication of an infection or other diseases.
- e. ***Vomiting***: Vomiting is the involuntary, forceful expulsion of the contents of one's stomach through the mouth. Vomiting can be caused by a specific response to ailments like gastritis or poisoning or malaria (Glass & White, 2007).

2.1.7 Factors Affecting Malaria Transmission

2.1.7.1 Internal Factors

Five critical internal factors responsible for Malaria transmission are identified:

- i. ***Frequency of Human Blood Index (HBI)***: This represents the proportion of blood meals derived from humans by mosquito vectors. It is dependent on the host and the temperature. It determines the likelihood of malaria infection and transmission of the infection back to a human host (Molineaux, Wernsdorfer & McGregor, 1988).
- ii. ***Duration of Sporogony***: This is the time taken from infection to sporozoite development as determined by temperature and parasite genetic (Molineaux, Wernsdorfer & McGregor, 1988).

- iii. **Vector Density:** This *is* influential for transmission because human hosts are sporadically infected (Molineaux, Wernsdorfer & McGregor, 1988).
- iv. **Vector Susceptibility:** It varies with capacity of different species and geographical strains of anophelids to transmit different parasites (Molineaux, Wernsdorfer & McGregor, 1988).
- v. **Demography.** Resistance to malaria infection can be acquired with age. Therefore, the age structure of a population may have consequences for the rate and distribution of malarial transmission (Sallares, 2002).

2.1.7.2 External Factors

- a. **Temperature:** The minimum temperature for mosquito development is between 8–10°C; the optimum temperature is 25–27°C, and the maximum temperature is 40°C. Higher temperatures increase the number of blood meals taken and the number of eggs laid by the mosquitoes, which increases the number of mosquitoes in a given area. Figure 2.8 indicates the Variations of Temperature in the Study Time.

Mosquito Survival Probability (MSP): The survival probability of mosquito is the proportion by which each blood meal survive and complete each gonotrophic cycle.

Sporogonic factors: The sporogonic cycle has a complex relationship which is completely dependent on mosquito survival and temperature. The measurement of sporogonic cycle is degree days, which are calculated by adding the heat from each day, which comes from the temperature. A certain amount of energy needed to complete the cycle is added up.

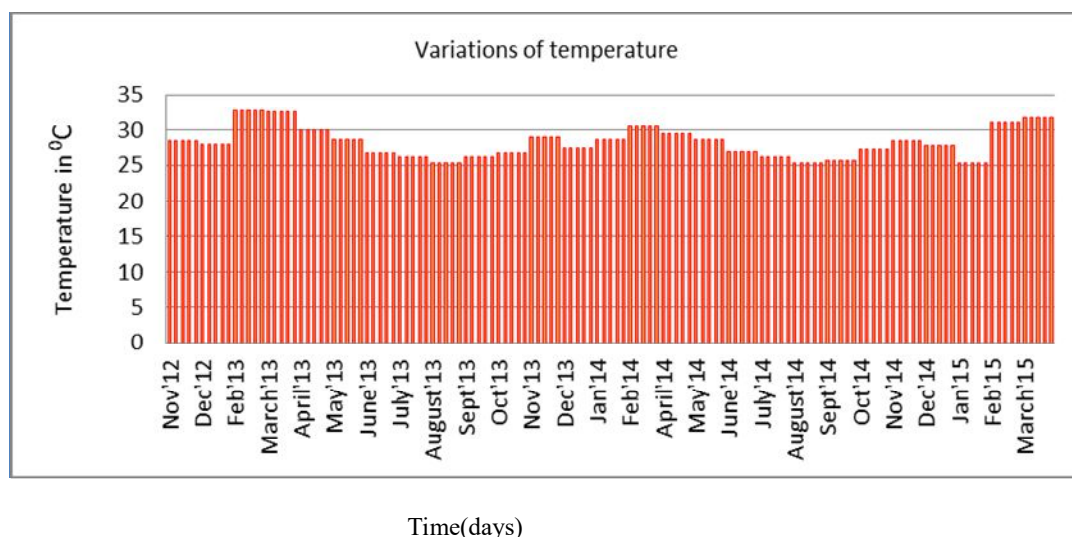


Figure 2.8: Variations of Temperature in the Study Time

up. The required amount of heat must be reached within the mosquito's lifetime, unless the sporogonic cycle will not occur.

Gonotrophic factors: Temperature also controls the gonotrophic cycle. It is measured also in degree days and the gonotrophic threshold is the minimum temperature required for mosquito reproduction.

b. Rainfall: Rainfall is needed to provide stagnant water, in which mosquitoes lay their eggs. Mainly malaria vectors breed in stagnant water collections, rarely in slightly moving waters and never in rapidly flowing streams and rivers (Sinka, Bangs, Mangun, Coetzee, Mbogo, Hemingway & Okara 2010). Figure 2.9 is the Variations of Rainfall with the Study Time. In dry areas, rainfall affects malaria transmission indirectly through its effect on humidity (Harvell, 2002).

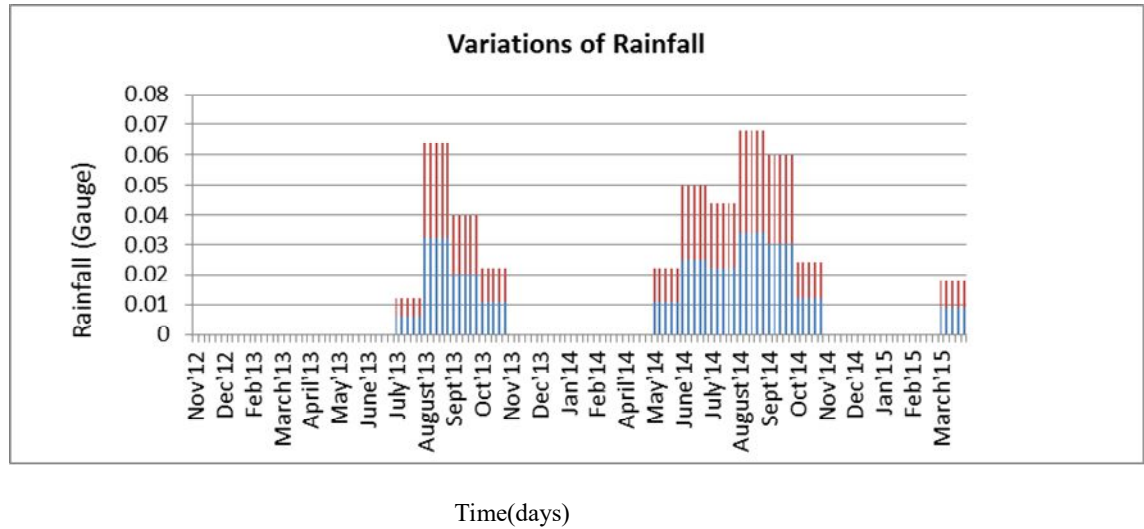


Figure 2.9: Variations of Rainfall in the Study Time

- c. **Relative Humidity:** Relative humidity affects malaria transmission through its effect on the activity and survival of mosquitoes. Malaria parasites become more active when humidity rises (Yusuf, 2013). Figure 2.10 is the Variations of Relative Humidity with the Study Time

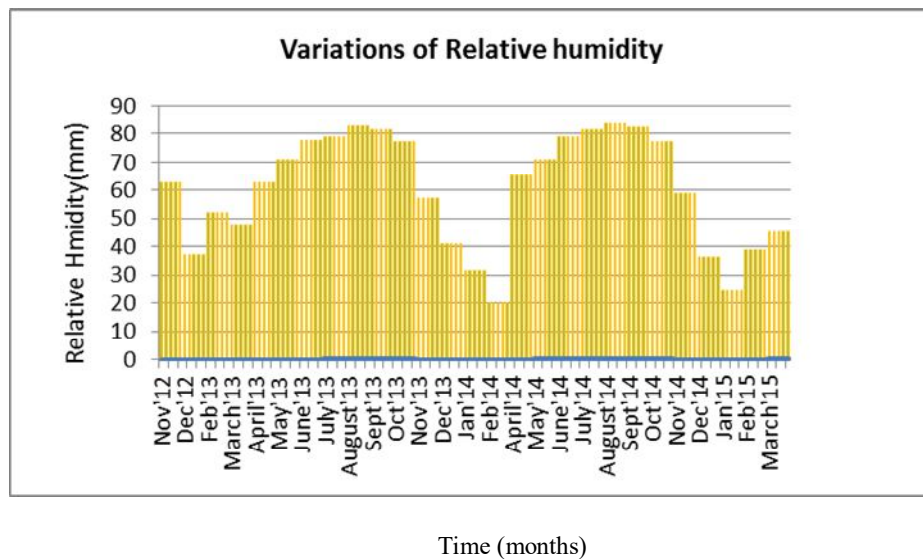


Figure 2.10: Variations of Relative Humidity in the Study Time

- d. ***Vegetation:*** Vegetation cover increases after rainfall, which in turn increases the relative humidity of the environment (Yusuf, 2013).

- e. ***Altitude:*** Altitude influences the distribution and transmission of malaria indirectly, through its effect on temperature. As altitude increases, temperature decreases, so highlands are colder and lowlands are warmer. In the Ethiopian highlands, with altitudes between 2,000 and 2,400 metres, malaria transmission occurs for short periods only when temperatures rise unusually high (Tesi, 2011; Yusuf, 2013).

- f. ***Human/behavioural factors:*** Several non-climatic factors, such as human/behavioural factors can also affect the pattern of malaria transmission and the severity of the problem.

- g. ***Environmental changes:*** Environmental changes such as increased global mobility and increased drug resistance, global warming, civil disturbances, are also considered among the causes of the severity of malaria cases worldwide (Greenwood, 1997).

2.1.8 Malaria Threats

The severity and complicated cases of malaria are commonly found among infants and pregnant women (Warrel, 1997). From research, pregnancy exacerbates malaria through a non-specific hormone-dependent depression of the immune system. At pregnancy, the protective anti-plasmodia function is suppressed at pregnancy which has clinical consequences with important public health

implications on pregnant women (Okwa, 2003). Premature deliveries of infants can be caused by Malaria infection with low birth weight due to intra-uterine growth retardation attributed to placental parasitization (Stekeetee, Wirima, Hightower, Slutsker, Heliman & Breman, 1996).

2.1.8.1 Vector threat: Vector threat in equation (2.1) is the mass tendency to be infected by malaria parasite. This is influenced by Human Blood Index(frequency with which the vector sucks blood meal) in relation with population size and environmental or climatic factors

$$V_{tr} = H_{bd}P_s + E_f \quad (2.1)$$

Where V_{tr} = Vector Threat

H_{bd} = Human Blood Index

P_s = Population Size

E_f = Environmental or Climatic factors

2.1.8.2 Transmission threat: Transmission threat in equation (2.2) is the rate at which vector threat causes mortality and morbidity in a population.

$$T_{tr} = (H_{bd}P_s + E_f)\ell \quad (2.2)$$

Where T_{tr} = Transmission Threat

ℓ = other internal and external factors contributed to the threat

2.1.8.3 Asymptomatic/Morphological Based Threat

The most widely used criteria for diagnosis of asymptomatic malaria infection are presence of parasites in peripheral thick blood smears, at auxiliary temperature $< 37.5^{\circ}\text{C}$, and an absence of malaria-related symptoms (de Mast, Syafruddin, Keijmel, Riekerink, Asih & Van der (2010); Leoratti, Durlacher, Lacerda, Alecrim, Ferreira, Sanchez & Moraes, 2008); Males, Gaye & Garcia, (2008); Duarte, (2007)). Some studies do include other criteria, such as longitudinal follow-up and parasite quantification. Longitudinal follow-up is particularly important for differentiating between infections that appear asymptomatic at the time of detection, but may become symptomatic after the initial detection (Martha, 2007; Males, Gaye & Garcia, (2008); Alves, Durlacher, Menezes, Krieger, Silva & Camargo, 2002; Cucunuba, Guerra, Rahirant, Rivera, Cortes & Nicholls, 2008).

Asymptomatic malaria is not only limited to regions of high transmission where exposure-related immunity is expected to develop; it has also been reported in the low transmission regions. But few reports are available on the study of asymptomatic malaria caused by species other than *P. falciparum*. The diagnosis of asymptomatic carriers has become a new and increasingly important challenge for malaria control programme. Major obstacle in the study of asymptomatic malaria infection is the lack of standard diagnostic criteria. For example, infected individuals may be in a pre-symptomatic period with parasitaemia, and present with clinical manifestations at a subsequent date (Coura, Suárez & Ladeia-Andrade, 2006).

In thin blood images, morphology of cells can be observed clearly. Parasitemia focuses on segmenting blood cells and classify them as infected or normal. So a Giemsa stained thin blood images is used. The staining process slightly colorizes the red blood cells (RBCs) and identifies spp parasites, white blood cells (WBC), and platelets or artefacts. Giemsa stains nuclei, chromatin in blue tone and erythrocytes in pink color. The techniques used in counting of infected and normal erythrocytes for purpose of estimating parasitemia (number of infected blood cells over total red blood cell count by correctly segmenting and separating erythrocytes was proposed (Savkare & Narote, 2011).

A morphological method for detection of parasites in Giemsa stained blood slides was conducted (Ruberto, *et.al.*, 2000). Dimension and colour of different objects in blood are identified an automatic thresholding based on morphological approach. Using Granulometrices was used to detect the parasites and evaluate size of RBCs and nuclei of the parasite. A segmentation method was used to identify morphological operators combined with the watershed algorithm. Some of the morphological characteristics are Size of Infected Red Blood Cells (SRBC), Shape of Parasite (SP), Number of Chromatin Dots (NCD), Texture of RBC (TRBC), Location of Chromatin (LC), Number of Parasite/RBC (NRBC). Halim *et al.*, proposed a template matching technique for detection of Parasitemia in RBCs. Variance-based techniques were used to detect parasites from grayscale images and color co-occurrence matrix which is based on the individual colour index of pixel and colour indices of its eight neighboring pixels (Halim, Bretschneider, Li, Preiser & Kuss, 2006).

Díaz G *et al.*, proposed a method for quantification and classification of erythrocytes in stained thin blood films infected with *Plasmodium Falciparum*. Three (3) main phases of preprocessing step, which corrects luminance differences are used. A segmentation step that uses the normalized RGB colour space for classifying pixels either as erythrocyte or background, followed by an Inclusion-Tree representation that structures the pixel information into objects, from which erythrocytes are found and classified as infected or normal erythrocytes and differentiates the infection stage, using a trained bank of classifiers (Díaz, González & Romero, 2009).

2.1.8.4 Symptomatic and Climatic Malaria Based Threat

In a study on risk factor for symptomatic malaria in a cohort of Ugandan Children, each time a child presented to the clinic with a new history of fever within previous 48 hours or temperature $\geq 38^{\circ}\text{C}$, a thick blood smear was performed. Patients were diagnosed with malaria infection if they fulfilled any of the following criteria: severe malaria (Warrel, 1997) or danger signs (WHO, 2015) temperature $\geq 38^{\circ}\text{C}$, a ≥ 500 asexual parasites/ μl blood and any parasitaemia features. Patients with symptomatic malaria cases were treated with their assigned treatment regimen and followed for 14 days using a standardized WHO protocol. All patients classified as clinical failures were treated with standard doses of quinine. Any case of malaria diagnosed more than 14 days after a previous episode was considered a new case (Njama-Meya, Kamya, & Dorsey, 2004).

The impact of climate in malaria resurgence is difficult due to high spatial and temporal climate variability and the lack of large datasets on malaria cases from different sites. Climate variability, a short-term fluctuation around the mean climate state, may be epidemiologically more relevant than mean temperature change, but its effects on malaria epidemics have not been rigorously examined. Nonlinear and synergistic effects of temperature and rainfall on the number of malaria outpatient's cases were found in all the seven sites. The net variance of the number of monthly malaria outpatients caused by auto-regression and seasonality varied among sites and ranged from 18 to 63% (mean = 38.6%), whereas 12–63% (mean = 36.1%) of variance is attributed to climate variability. The study results indicate that there was a high spatial variation in the sensitivity of malaria outpatient number to climate fluctuations in the highlands, and that climate variability played an important role in initiating malaria epidemics in the East African highlands (Zhou, Minakawa, Githeko, & Yan, 2004).

2.2 Predicting Models

Applying different predicting methods to the same data, exploring the predictive ability of non-environmental variables, including transmission reducing interventions and using common forecast accuracy measures will allow malaria researchers to compare and improve models and methods, which should improve the quality of malaria prediction. The prediction approaches included statistical modeling, mathematical modeling and machine learning methods (Zinszer, Kigozi, Charland, Dorsey, Kamya & Buckeridge, 2013). Mathematical, statistical

and computational engineering models are playing a most valuable role in shedding light on the problem and for helping to make decisions.

2.2.1 Mathematical Modeling of Infectious diseases Dynamics

In the time past, series of approaches have been proposed looking at the problem from different perspectives. These encompasses three general categories in Figure 2.11: (1) surveillance outbreaks statistical methods and identification of real epidemics spatial patterns; (2) State-Space models mathematical models within the context of dynamic systems used to predict the evolution of a “hypothetical” or on-going epidemic spread; and (3) machine learning/ expert methods for the prediction of the evolution of an ongoing epidemic. For all three of these categories, there are again different approaches from different literatures.

2.2.1.1 Statistical-Based Methods for Epidemic Surveillance

One of the most important aspects in epidemics revolves around the surveillance, early detection of possible outbreaks and patterns that may help in controlling a spread. One of the very first success stories in the area is the modeling of cholera epidemic that swept through London in 1854. At that time, John Snow, a physician, collected spatiotemporal data and by visualizing them in a map found that there was a particular pattern around the Broad Street water which actually was the zero point of transmission (McLeod, 2000). His analysis helped eradicate the disease. In the dawn of 20th century Greenwood, an epidemiologist and statistician was the first Professor of Epidemiology and Statistics at the London

School of Hygiene and Tropical Diseases establishing a rigorous mathematical connection between fields (Greenwood, 1916).

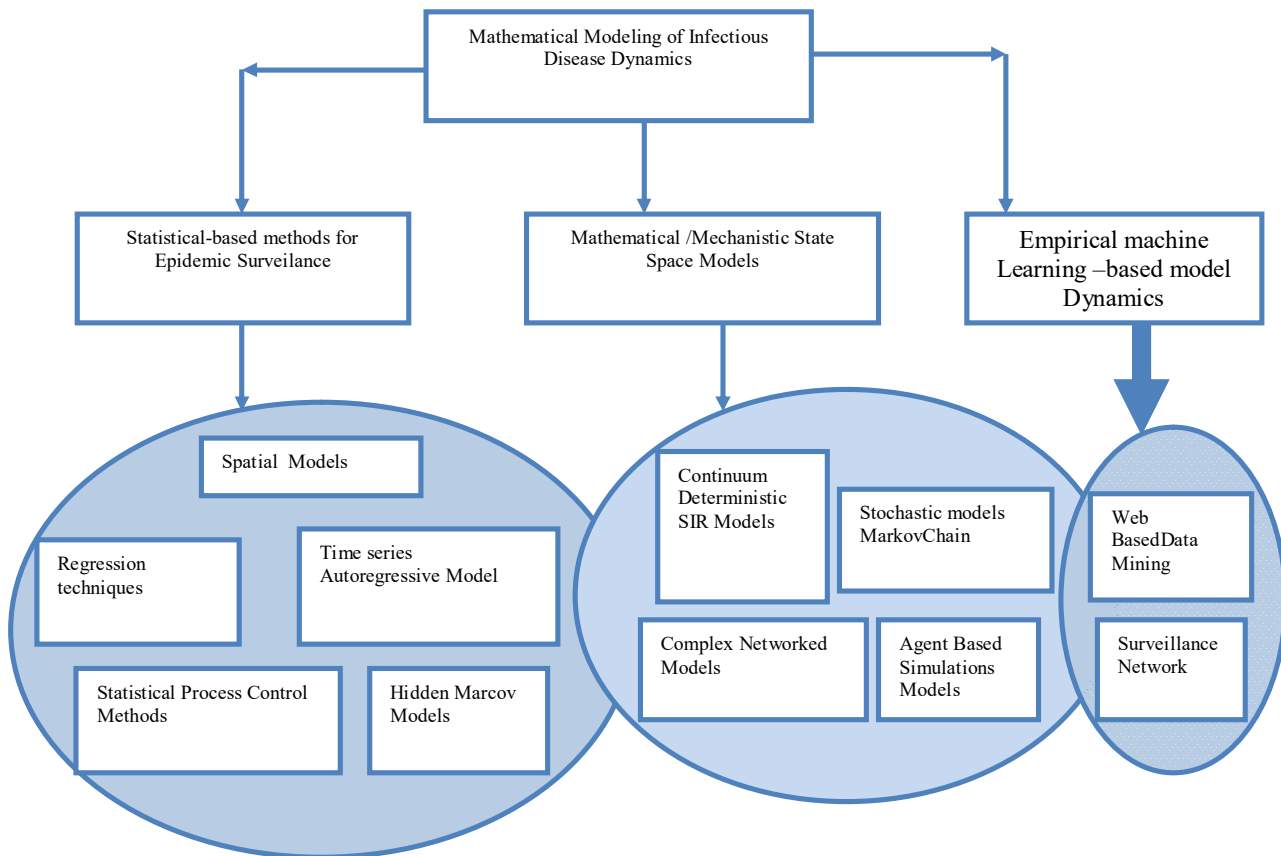


Figure 2.11: An Overview of Mathematical Modeling of Infectious Diseases Dynamics (Siettos & Russo, 2013).

Today, global initiatives to combat epidemics require effective domestic action mechanisms and preparedness. An intensive worldwide effort led by World Health Organization and Centers for Disease Control is speeding up the developments for the establishment of a global surveillance network. New emerged pandemics such as the Acquired Immune Deficiency Syndrome (AIDS), the Severe Acute Respiratory Syndrome (SARS) of 2002–2003 and the H1N1

swine flu of 2009 pandemics remind us about the importance of surveillance and prompt outbreak detection. Toward this aim, statistical methods have enhanced our potential in fighting epidemics allowing for rapid assessment of emerging situations. Obviously, the correctness of the data and the selection of the appropriate methodology are crucial for the construction of statistical models that can capture, in an efficient robust way, the communicable disease characteristics.

To date, several statistical methods have been proposed (Unkel, Farrington, Paul, Garthwaite, Robertson & Andrew, 2012) for a review of statistical methods for the detection of disease outbreaks.

2.2.1.2 Mathematical State Space Model

According to the level of the approximation of the reality and increasing complexity mathematical models may be categorized into the following categories:

i. Continuum models in the form of partial differential equations

Continuum models describe the coarse-grained dynamics of the epidemics in the population (Brauer, Castillo & Castillo, 2001; Feng, Dieckmann & Levin, 2006), for example, a study on a model for the evolution of the disease as a function of the age and the time since vaccination or investigation on the influence of quarantine or isolation of the infected part of the population (Hethcote, Zhien & Shengbing, 2002). Such models can be explored using powerful analysis techniques for ordinary or partial differential equations. However, due to the complexity and the stochasticity of the phenomena, most available continuum

models are often only qualitative caricatures that cannot capture all of the details, therefore compromising epidemiological realism.

Within this context, the population is divided in compartments in accordance with the state of their health, such as susceptible (S), infected (I), and recovered (R). Other states of the population linked with control policies such as vaccinated (V) and quarantined (Q) are also used.

ii. **Stochastic models including discrete and continuous-time individual based Markov-chain model** (Prosperi *et al*; Vardavas & Blower, 2007)

These are usually individual-level models that relax the hypothesis of the mean field approximations of infinite population and perfect mixing introducing the uniqueness of the individual behaviour including multiple heterogeneous characteristics. The main representative in the category is the Discrete Markov chains (DMC). In DMC both time and states are defined on a discrete set of values. The states of the individuals change at every discrete time step in a probabilistic manner according to simple rules involving their own states and the states of their links satisfying the Markov property, i.e., that the future values of the states at time $t + \Delta t$ depend only on the values of the states at the previous time step t .

For example, for a stochastic model these transition rules may read:

Rule 1: An infected individual (I) infects a susceptible (S) link with a probability $p_{S \rightarrow I} = \lambda$ if an active physical communication exists between them.

Rule 2: An infected individual (I) recovers with a probability $p_{I \rightarrow R} = \delta$.

Rule 3: A recovered individual (R) becomes susceptible (S) with a probability $p_{R \rightarrow S} = \gamma$. This condition expresses the case of temporal immunity.

iii. **Complex network models that are relaxing the hypothesis of the stochastic models that the interactions between individuals are instantaneous and homogeneous** (Keeling, Danon, Vernon & House, 2010; Keeling & Eames, 2005; Read, Eames & Edmunds, 2008)

One of the most critical problems in epidemics concerns the dynamic effects of the contact network heterogeneity. Contacts between individuals evolve under numerous complicated and strongly heterogeneous modes that are influenced by a broad spectrum of factors, ranging from the pathogen inherent variability and host–pathogen interaction stochasticity characterizing the transmission mechanisms of a particular disease, to the population-level ones complicated by environmental, seasonal, economic, and demographic conditions.

iv. **Agent-Based Simulations**

In contemporary mathematical epidemiology, agent-based modeling represents the state-of-the-art for reasoning about and simulating complex epidemic systems. These take into account details such as the transportation, infrastructure of the simulated area, the mobility of the population, demographics, and epidemiological aspects such as the evolution of the disease within a host and transmission between hosts.

2.2.1.3 Machine Learning Based Model

A computer programme is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance of tasks in T , as measured by P , improves with experience E . Machine learning (ML) is the science and art of making computers to work without being explicitly programmed. Recently, there is very rapid development of the use of machine learning because it provides powerful computing environment for the analysis of big data and have efficient learning algorithms. For instance, ML is used in medical science to check health condition (Paokanta, Ceccarelli & Srichairatanakool, 2010; Martinez *et al.*, 2014) and diagnose a disease such as cancer (Kourou, Exarchos, Exarchos, Karanonzix, & Fotiadis, 2014; Asadi, Dowling, Yan & Mitchell, 2014). In pharmacology ML find the right formula and reliable drugs to incapacitate a disease virus (Danger, Segura-Bedmar, Martinez, & Rosso, 2010; Urquiza, Rojas, Pomares, Herrera, Florido, Valenzuela, & Cepero, 2012). ML is also used to choose the effective therapeutic treatment (Caravaca, Soria, Bataller, Serrano, Miquel, France & Gurrero, 2013). Also ML can also be used in agriculture to increase agricultural production as with predicting pest plants (Worner & Gevrey, 2006). In the business world, ML is used to predict the stock market and stock price index movement (Patel, Shah, Thakkar & Kotecha, 2014).

2.3 *Malaria Based Models*

Malaria based framework and technique serves as the general practical implementation and analysis of evidence of malaria disease. Some are generally identified frameworks, SVM based framework and ANN based framework.

2.3.1 *General Based Model*

An early-warning system for malaria epidemics in Punjab based on rainfall, fever-related deaths and wheat prices was developed (Christophers, 1911). Since that initial system, researchers and practitioners have continued to search for determinants of temporal and spatial variability of malaria to improve systems for predicting disease burden. Malaria prediction is now being conducted in many countries and typically uses data on environmental risk factors, such as climatic conditions, to forecast malaria incidence for a specific geographic area over a certain period of time (Zinszer *et al.*,2013).

Predictive model helps healthcare providers design and implement preventive life saving measures given susceptibility towards a particular diseases. The availability of "big data" and cost-efficient processing power is expanding the applicability of predictive data-driven techniques in different industries. In doing that, clever mathematics is helping more and more companies realize the true potential hidden in their data. The study was a review of common predictive modeling techniques (Alex, 2012).

A Malaria Incidence Prediction from Incidence Record and Weather Pattern Using Polynomial Neural Network (PNN) was proposed. Inductive Group Method of Data Handling (GMDH) algorithms give possibility to find automatically interrelations in data, to select optimal structure of model or network and to

increase the accuracy of existing algorithms by developing a Polynomial Neural Network model. Based on the experiments, it was proven that the modified GMDH PNN was able to reduce the learning time by 72% and improve the accuracy up to 88.02% compared to the original GMDH PNN. The observation was that any other normalization and data complexity reduction such as PCA, quality measurement function and technique may help the learning process to make a better prediction model (Arifianto, Barmawi & Wibowo, 2014).

A Malaria Disease Identification and Analysis Using Image Processing research was conducted in 2014, since analysis of malaria using manual microscopy diagnostic method is prone to human error and time consuming. The analysis shows the identification of malaria disease using image processing to detect various features in the blood image. The statistical analysis of Skewness, Kurtosis, Energy and standard deviation was deployed on the blood image. The analysis generated an accuracy of 98.25% for SVM classifier and 78.53% for ANN (Chavan & Sutkar, 2014).

2.3.2 SVM Based Model

A vital informative aspect of clinical practice is the relative counting and assessment of the blood cells in the bone marrow of patients. This is very important for Leukemia patients for the observation of the development stage of the illness and the preparation of the treatment of patients. Recognition of the cells at different stages of their development is needed to achieve proper diagnosis of the disease and calculate their relative quantity in the aspirated bone marrow. Genetic Algorithm (GA) was applied with Support Vector Machine (SVM) for the

recognition of blood cells in the image of bone marrow (Osowski, Siroic, Markiewicz, & Siwek, 2009).

An automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage was proposed in 2012. Precise diagnosis and detection of parasitemia are necessary for accurate medication and manual counting and classification of RBCs is laborious and time consuming. Otsu's threshold for segmentation of cells from Giemsa stained blood images was used. Watershed Transform was also used for separation of overlapping cells. The geometrical and statistical features were extracted and given to SVM binary classifier and this gives 96.26% sensitivity and 99.09% specificity. Consequently, SVM multiclass classifier (i.e. RBF Kernel) gives 96.42% accuracy but other tools for classification were not considered (Savkare & Narote, 2012).

A predictive symptom-based system for malaria and typhoid coinfection using support vector machines to improve the classification accuracy and results of the coinfection was developed. The system aimed at providing a computer based alternative quality healthcare solution that is effective, relatively accurate and affordable especially for developing nations like Nigeria that are challenged with shortages of medical specialists and facilities. Similarly, the available few medical resources are not available for rural dwellers. Twenty (20) patient malaria cases were collected to serve as the datasets. Out of the twenty cases, 16 malaria cases were classified correctly and four malaria cases were misclassified thereby making the system 80% accuracy for classifying malaria. Beside, Twelve (12) typhoid cases were classified correctly and 8 were misclassified thereby the

system has 60% accuracy for classifying typhoid. The overall model accuracy of 90% was attained for the typhoid and malaria and captured relatively low dataset. Other performance metrics were not considered. No global thresholding and consideration of effects of climatic conditions (Aminu, Onyebuchi & Ibrahim, 2016).

2.3.3 ANN Based Model.

There is a need for an alternative approach for medical practitioners to detect malaria and accurately identify the level of severity. So a malaria severity classification through Jordan-Elman Neural Network based on features extracted from thick blood smear was conducted. Out of total of 57 cases of Mildly Severe (MS) malaria, 54 were correctly classified and three were wrongly classified as Highly Severe (HS) malaria. In the case of HS malaria, of a total of 33 cases, 31 were correctly classified and two were incorrectly classified as MS malaria. This result indicated that the proposed model did not perform perfectly but has very high classification accuracy. The need for more accurate detection and classification of malaria severity is urgent to assist medical practitioners in making decisions (Chiroma, 2014)

In 2010, an Artificial Neural Networks and Bayesian Networks were employed as Supporting Tools for diagnosis of asymptomatic malaria. Since asymptomatic carriers generally present with very low parasitemia, the microscopy test does not generate a reliable diagnosis for this form of malaria. Thus, innovative tools are urgently needed to support a feasible approach to overcome the menace of this disease. The result of the study shows that ANN has 67.5% of sensitivity and

92.5% of specificity, and has correctly predicted 80.0% of the 80 individuals separated for testing the method. The BN presented 37.5% of sensitivity and 97.5% of specificity, and has correctly predicted 67.0% of the 80 individuals separated for testing the method (Júnior, Duarte, Netto & Andrade., 2010).

A study on the comparison of the performance of different networks for malaria classification was conducted in 2014. The motivation was that traditionally, light microscopy is used for malaria detection which is less accurate and time consuming process. Malaria detection in early stage is not possible with microscopy test. A trained operator is required for testing, thus it increases the cost of test. The result of the study shows that probabilistic model of Naive Bayes classifier is more accurate than Bayesian regularization. Hidden neurons, Number of epochs can be selected in Back Propagation (BPNN). So, Naive Bayes classifier is proved to be better technique. Sensitivity of ANN and BN is 45% and 15% greater than microscopy test (Dipti & Dilip, 2014).

Blood image processing for detecting malarial parasites in images of Giemsa Stained Blood slides to evaluate the parasitemia of the blood was introduced (Raviraja, Geethanjali, Chethana & Kanthesh 2015).. Artificial Neural Network was used to detect infected red blood cells by parasitemia. The reason for the research was that most existing malaria diagnosis methods were based on clinical notion and explicitly on clinical request. No technique has been used for routine clinical automated screening. The model diagnosed 95 patient's clinical data of plasmodium parasites and obtained an accuracy of 60-96% based on selected dependent variables. The model was validated with ROC characteristics.

2.3.4 ANN and SVM Based Models

Mistakes are inevitable in manual counting diagnosis and it is time consuming; thus we need to develop an image processing algorithm to automate diagnosis of malaria on thin blood smears. Morphological and novel threshold selection technique can be used to identify the parasites on microscopic slides. Behavioural image features such as colour, texture and the geometry of the cells and parasite were generated. Image processing was used to identify malaria parasite with the use of Phase of Image, Mean of Greenplane, Skewness, Kurtosis, Standard Deviations and Energy. ANN classifier gives an accuracy of 80% for affected and 77% for not affected and SVM gives an accuracy of 90% for affected and 100% for not affected. But the researchers were unaware that the performance of a classifier depends on the domain under discussion (Shruti & Shirgan, 2015).

Most of the conventional microscopy used in diagnosis of diseases are occasionally proving inefficient and results are difficult to reproduce. Three (3) classifiers SVM, Naïve Bayes and Neural network classifier and two feature extraction techniques - Discrete Wavelength Transform (DWT) and Gray-Level Co-Occurrence Matrix (GLCM) were used. The system obtained 100% accuracy of disease detection with the use of SVM classifier and 92.85% accuracy with Naïve Bayes. An accuracy of 92% and 85.41% were obtained when DWT and GLCM Feature Extraction methods were used respectively with Neural Network. The methods adopted made use of the morphological, colour and texture features of plasmodium parasites and erythrocytes (Chaudhari & Agrawal, 2015).

Early prediction of a Malaria outbreak is a key for control of malaria morbidity. This will help various health organizations to better target medical resources to areas of greatest need. Two popular data mining classification algorithms Support Vector Machine (SVM) and Artificial Neural Network (ANN) were used for Malaria prediction. The features of the model are average monthly rainfall, temperature, humidity, total number of positive cases, total number of Plasmodium Falciparum cases and the outbreak occurrence in binary values of Yes or No. The accuracy of the prediction needs to be improved upon by using more training data. Also, in the model, the individual positive cases needed to be considered not total number of positive cases as one of the training, testing and validation features (Sharma, Kumar, Lakshmi & Karajkhede, 2015).

2.4 Review of *Malaria Threat Model*

Table 2.1: Malaria Threat Model

	Author	Title	Motivation	Results	Limitation
1.	Raviraja S, Geethanjali S.D, Chethana and Kanthesh B.M. (2015)	The classification and recognition of Plasmodium parasite in Prediction of Malaria Infected Blood Smears using Artificial Intelligence Technique.	Most existing malaria diagnosis methods are based on clinical notion and explicitly on clinical request. No technique has been used for routine clinical automated screening.	Blood image processing for detecting malarial parasites in images of Giemsa Stained Blood slides to evaluate the parasitemia of the blood was introduced. Artificial Neural Network are also used to detect infected red blood cells by parasitemia. The model diagnosed 95 patient's clinical data plasmodium parasites and obtained an accuracy of 60-96% based on selected dependent variables. The model was validated with ROC characteristics.	The exact accuracy of the model was not stated. The one given was in ranges.
2.	Vijeta Sharma, Ajai Kumar, Lakshmi Panat, Dr. Ganesh Karajkhede, Anuradha Lele (2015)	Malaria Outbreak Prediction Model Using Machine Learning	Early prediction of a Malaria outbreak is key for control of malaria morbidity. This will help various health organizations to better target medical resources to areas of greatest need.	Two popular data mining classification algorithms Support Vector Machine (SVM) and Artificial Neural Network (ANN) are used for Malaria Prediction. Parameters used are average monthly rainfall, temperature, Humidity, total number of positive cases, total number of Plasmodium Falciparum cases and outbreak occurrence in binary values of Yes or No. The SVM model can predict the outbreak 15 -20 days in advance.	The accuracy of the prediction needs to be improved on by using more training data.
3.	Tejashri Chaudhari & Prof D.G. Agrawal (2015)	Automatic Detection of malaria parasites for estimating parasitemia	Most of the conventional microscopy used in diagnosis of diseases are occasionally proving inefficient and results are difficult to reproduce.	Three (3) classifiers SVM, Naïve Bayes and Neural Network classifier and two feature extraction techniques DWT and GLCM (Gray Level Cooccurrence Matrix) were used. The system obtained 100% accuracy of disease detection with the use of SVM classifier and 92.85% accuracy with Naïve Bayes. An accuracy of 92% and 85.41% were obtained when DWT and GLCM Feature Extraction method were used respectively with Neural Network.	The method makes use of the morphological, colour and texture features of Plasmodium parasites and erythrocytes.

4.	WHO (2015)	<i>The Global Technical Strategy for Malaria 2016-2030</i>	It provides a comprehensive framework to guide countries in their efforts to accelerate progress towards malaria elimination. The strategy sets the target of reducing global malaria incidence and mortality rates by at least 90% by 2030.		
5.	Shruti Annaldas & Prof S.S Shirgan (2015)	Automatic Diagnosis of Malaria Parasites using Neural Network and Support Vector machine	Mistakes are inevitable in manual counting diagnosis and time consuming. So we need to develop an image processing algorithm to automate diagnosis of malaria on thin blood smears. Morphological and novel threshold selection technique can be used to identify the parasites on microscopic slides.	Behavioural image features such as colour, texture and the geometry of the cells and parasite were generated. Image processing was used to identify malaria parasite with the use of Phase of Image, Mean of greenplane, Skewness, Kurtosis, Standard Deviations and energy. ANN classifier gives an accuracy of 80% for affected and 77% for not affected and SVM gives an accuracy of 90% for affected and 100% for not affected.	The researchers was unaware that the performance of a classifier depends on the domain under discussion.
6.	Anditya Arifianto, Ari Moesriami Barmawi, and Agung Toto Wibowo (2014)	Malaria Incidence Predicting from Incidence Record and Weather Pattern Using Polynomial Neural Network	Inductive Group Method of Data Handling (GMDH) algorithms give possibility to find automatically interrelations in data, to select optimal structure of model or network and to increase the accuracy of existing algorithms	Based on the experiments, it was proven that the modified GMDH PNN was able to reduce the learning time by 72% and improve the accuracy to 88.02% compared to the original GMDH PNN.	Any other normalization and data complexity reduction such as PCA, quality measurement function and technique may help the learning process to make a better prediction of the model
7.	H. Chiroma_y, S. et al., (2014)	Malaria Severity Classification Through Jordan-Elman Neural Network Based On Features Extracted From Thick Blood Smear	There is a need for an alternative approach for medical practitioners to detect malaria and accurately identify the level of severity.	Out of total of 57 cases of Mildly Severe(MS) malaria, 54 were correctly classified and three were wrongly classified as Highly Severe (HS) malaria. In the case of HS malaria, of a total of 33 cases, 31 were correctly classified and two were incorrectly classified as MS malaria. This result indicated that the proposed model did not perform perfectly but has very high classification accuracy.	The need for more speedy detection and classification of malaria severity is urgent to assist medical practitioners in making decisions.

8.	<i>Dipti D. Patankar, Dilip G. Khairnar (2014)</i>	Comparing the Performance of Different Networks for Malaria Classification	Traditionally, light microscopy is used for malaria detection which is less accurate and time consuming process. Malaria detection in early stage is not possible with microscopy test. A trained operator is required for testing, thus it increases the cost of test. To make a system automated and improve the accuracy of system Neural Networks are used.	A probabilistic model Naive Bayes classifier is more accurate than Bayesian regularization and Hidden neurons. Number of epochs can be selected in Back Propagation (BPNN). So, Naive Bayes classifier is proved to be better technique. Sensitivity of ANN and BN is 45% and 15% greater than microscopy test.	Malaria images are classified only from their morphological characteristics .
9.	<i>Sneha Chavan, Ashok Sutkar (2014)</i>	Malaria Disease Identification and Analysis Using Image Processing	Analysis of malaria using manual microscopy diagnostic method is prone to human error and it is time consuming	The analysis of the identification of malaria disease using image processing to detect various features in the blood image. The statistical analysis of Skewness, Kurtosis, Energy and Standard Deviation was conducted on the blood image. The analysis generated an accuracy of 98.25% for SVM classifier and 78.53% for ANN Classifier.	The method was based on data image processing not on real life datasets
10.	<i>Sudheer Ch et al., (2013)</i>	A Support Vector Machine-Firefly Algorithm based predicting model to determine malaria transmission	The performance of SVM mainly depends on its appropriate parameters selection which is very complex in nature and quite hard to solve by conventional optimization techniques	The results indicate that the proposed SVM-FFA model provides more accurate prediction compared to the other traditional techniques	The utility of SVM-FFA in detecting the malarial cases for other geographical regions was not considered. Also, other lead times (such as bi-monthly, quarterly or yearly prediction) and comparison of the SVM-FFA with the other new methods such as extreme learning machines were not considered.
11.	<i>Kate Zinszer et al., (2013)</i>	A scoping review of malaria prediction: past works and future directions	There is a growing body of literature on malaria predicting methods and the objective of our	29 different studies that met the inclusion criteria were reviewed. Applying different Predicting methods to the same data, exploring the predictive ability of non-environmental variables, including transmission reducing	Comparison of the prediction accuracy of models from the different studies as the

			review is to identify and assess methods, including predictors, used to predict malaria	interventions and using common forecast accuracy measures will allow malaria researchers to compare and improve models and methods, which should improve the quality of malaria predicting.	evaluation measures differed across the studies was not conducted
12.	Savkare and Narote (2012)	Automatic System for Classification of Erythrocytes Infected with Malaria and Identification of Parasite's Life Stage.	Manual Counting and Classifications of infected erythrocytes is laborious and time consuming. There is need for an automatic system to detect malaria parasite infected erythrocytes and detection of life stage of parasites.	Otsu's threshold for segmentation of cells from Giemsa stained blood images was used. Watershed Transform was also used for separation of overlapping cells. The Geometrical and Statistical features are extracted and given to SVM binary classifier and gives 96.26% sensitivity and 99.09% specificity. SVM multiclass classifier (i.e. RBF Kernel) gives 96.42% accuracy.	Other tools for classification were not considered.
13.	Savkare and Narote 2011.	Automatic Clasiffication of Normal and Infected Blood cells for Parasitemia Detection.	Precise diagnosis and detection of parasitemia are necessary for accurate medication. Manual counting and classification of RBCs is laborious and time consuming.	It presents an automatic method for quantification and classification of erythrocytes in Giemsa stained infected thin blood films with Plasmodium Falsiparum. Staistical parameters werer used to extract the features and SVM classifier was used for classification infected or normal blood cells. Watershed algorithm was used to separate overlapped cells.	The method recognizes only two classes of parasitemia infection.
14.	Austeclino et al., (2010)	Artificial Neural Networks and Bayesian Networks as supporting tools for diagnosis of Asymptomatic Malaria	Since asymptomatic carriers generally present with very low Parasitemia, the microscopy test does not generate a reliable diagnosis for this form of malaria. Thus, innovative tools are urgently needed to support a feasible approach to overcome the menace of this disease.	The ANN presented 67.5% of sensitivity and 92.5% of specificity, and has correctly diagnosed 80.0% of the 80 individuals separated for testing the method. The BN presented 37.5% of sensitivity and 97.5% of specificity, and has correctly diagnosed 67.0% of the 80 individuals separated for testing the method.	The need to find other relevant features that can be used to construct a better classifier

2.5 Choice of Machine Learning Algorithm

Data Mining is the act of exploring hidden knowledge in data. Due to immense medical data in health care system, data mining evolves to analyze the complex nature of the biological systems (Kulkarni, Joshi & Brown, 2016). Conceptualized data mining and knowledge discovery on the unstructured data is needed to be carried out, but the choice of ML algorithm to solve a problem always depends on the size, quality, and nature of the data (Cherkassky & Mulier, 2007; Goswami & Chakrabarti, 2014).

2.5.1 Parametric System Learning Process

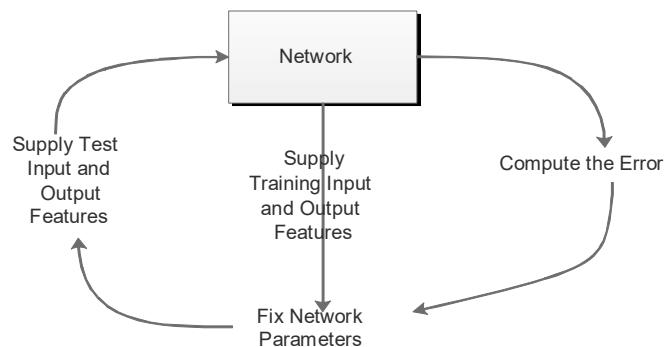


Figure 2.12 : Parametric System Learning Process

The Figure 2.12 is the learning process in a parametric system of modeling and the types of learning are discussed. A raw medical system data is collected and divided into testing and training feature then undergoes most suitable preprocessing technique.

2.5.2 Types of Learning:

1. **Supervised Learning:** In supervised learning given a set of input data that corresponds to a set of output data, a rule is learned which mapped input dataset to the output data set. Supervised learning has four(4) algorithms:
 - a. **Classification:** the data is classified into groups or set of classes for categorical response values. The classification learning scheme is subjected to a set class data from which a rule is learned to classify new data. Classification algorithms include; Neural Networks, Support Vector Machines (SVM), Nearest Neighbors (k NN), Naïve Bayes classifier, Decision trees and Discriminant analysis.
 - b. **Regression:** the data is classified into groups for continuous-response values. This is a variant of classification learning of which the outcome is a numeric value, not categorical values. Regression algorithms include: Linear Regression, Non-Linear Regression, Support Vector Regression (SVR), Generalized linear models, Decision trees and Neural networks.
 - c. **Clustering:** Features that are commonly related are grouped together in clustering learning scheme. Rules are also learned to cluster new instances.
 - d. **Association:** in association learning scheme, ore or more features can be predicted, not just the class and association among these features are sought.

2. **Reinforcement Learning:** in reinforcement learning, the training example does not explicitly contain the correct target output for each input but some output with how good the output is. It is commonly used for learning how to play a game.
3. **Unsupervised Learning:** In unsupervised learning giving a set of input data the corresponding set of output data in prior is not known but the rule that mapped input dataset to the output data set is known. So the output target has to be found.

2.5.3 Main Components of Learning

The Main Components of Learning are:

- i. X : Instance space or input space where an observation x belongs to this space
- ii. Y : Output Space or label space which is the set of all output where an output y belongs to this space. Examples: Binary Classification $y = \{\pm 1\}$, Multiclass Classification $y = \{1, \dots, K\}$, Regression $y \in \mathbb{R}$
- iii. $l: y^* y = R$: Loss Function (measures prediction error)
- iv. $f: X \rightarrow Y$: Target Function/Model/Hypothesis Class (Set of functions from input space to output space). The unknown target function is the correct formula that correctly mapped x to the corresponding output y .
- v. *Training Data Set $D(\mathbf{D})$* : This is the set of historical records x with their corresponding output y (input-output examples): $\{(x_1, y_1), \dots, (x_n, y_n)\}$ where $y_n = f(x_n)$, $n = 1, 2, \dots, l$.

- vi. *Learning Algorithm g*: A learning algorithm uses the data set D to pick a formula $g: X \rightarrow Y$ that approximates f . the algorithm chooses g from a set of hypothesis set H (candidates formula under consideration).
- vii. $g(x)$:the hypothesis that the learning algorithm produced is represented mathematically as $g(x) = \text{sign}((w_i, x) + b)$ where w is the weight and b is the bias.

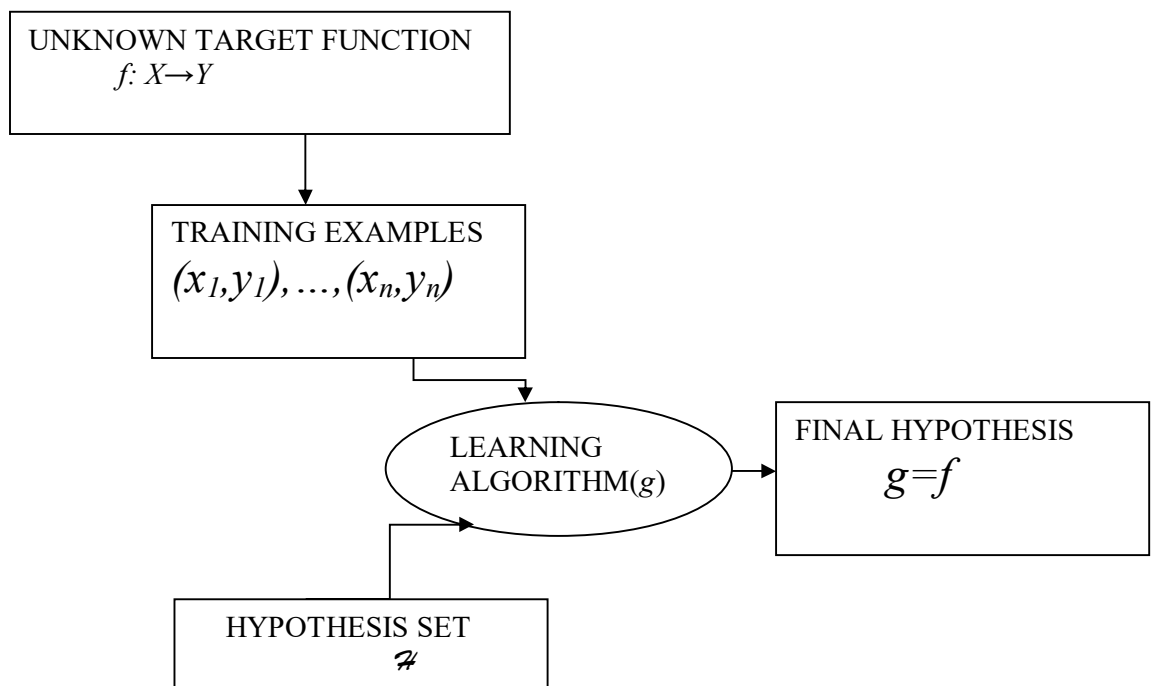


Figure 2.13: Basic Set up of Learning Problem (Abu-Mostafa et al., 2012)

Figure 2.13 portrays the components of learning problem where the target function f is the object of learning which is unknown.

Modeling goal is to choose a model from the hypothesis space, which is closest to the underlying function in the target space (Gunn, 1998). The outcome of learning process is being affected by the choice of an error measure. So, the choice an error measure has consequences on what we learn.

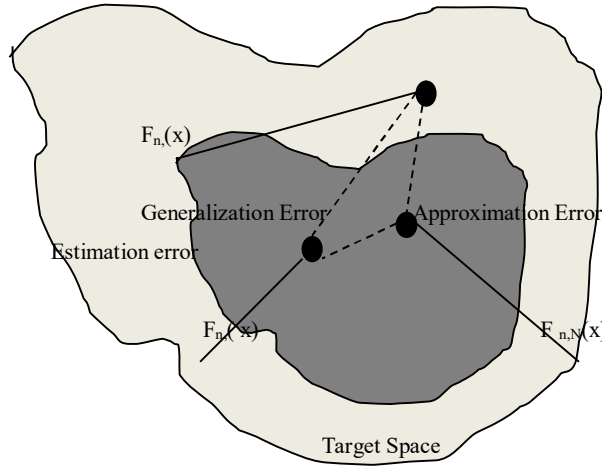


Figure 2.14. Modeling Errors.(Vapnik, 1998).

$$Error = E(h,f) \quad (2.3)$$

Formally, an error measure in equation (2.3) quantifies how well each hypothesis h in the model approximates the target function f as shown in Figure 2.14.

2.6 Strength and Weaknesses of Various Computational Intelligence Tools

Table 2.2: Strength and Weaknesses of Various Computational Intelligence Model Tools

Sources: [Sharma, V., Rai, S., & Dev, A.(2012); Castillo et al., 2001a,b). Karnik and Mendel, 1999; Mendel, 2003, Anifowose, F., & Abdulraheem, A. (2011); Jian & Wenfen, 2006]

Characteristics Model	FL	ANN	SVM	GA
1. Function Approximate Reasoning	Mild	High	High	High
2. Learning Abilities	Low	High	High	High
3. Handling of Uncertainties	High	High	High	High
4. Fault Tolerance	Low	High	High	High
5. Massive Parallelism	Low	High	High	High
6. Capability to solve Multiclass Problem	Low	High	Low	Mild
7. Threshold Metrics (Accuracy)	Mild	High	High	Mild
8. Probability Metrics (MSE)	Low	Low	Low	Low
9. Rank Metrics (ROC)	Low	High	Mild	Mild
10. Training Time	High	Low	Low	High
11. Number of Parameters	Low	High	High	High
12. Number of Features	Low	High	Mild	High
13. Application Capability	Mild	High	High	High

1. ***Approximate Reasoning:*** This minimizes the sum of squared errors between the input and the target output by suggesting an approximation to each of the functions and selecting the best among them. Fuzzy Logic is mild while others are high (Castillo *et al.*, 2001a,b).
2. ***Learning Abilities:*** Ability to extract rules directly from input data. FL ability is low in the process of extracting rules, directly from the input data. ANN learning process consists of obtaining the neural functions from a set of training data and does relatively well (Anifowose & Abdulraheem, 2011). While GA and SVM also have high learning abilities and good generalization ability (Savkare & Narote, 2012)
3. ***Handling of Uncertainties:*** The strengths of FL, lie in its ability to determine an exact membership function for a fuzzy set; hence, it is useful for handling uncertainties. ANN, SVM and GA also do relatively well (Karnik & Mendel, 1999; Mendel, 2003).
4. ***Fault Tolerance:*** SVM and GA explicitly control the trade-off between complexity and error. ANN minimizes the sum of squared errors between the input and the target output well because if any part of the network fails the network fails automatically While Fuzzy has low fault tolerance. (Sharma, Rai, & Dev, 2012).
5. ***Massive parallelism:*** As portrayed the parallel processing in neurons of human body, ANN and SVM and GA do the same but fuzzy does not.
6. ***Capability to solve Multiclass Problem:*** ANN is highly capable of solving multiclass problems while SVM is capable of solving binary problems. GA and FL are averagely capable of solving multiclass problems.

7. ***Threshold Metrics (accuracy):*** SVM needs a “good” kernel function to have high accuracy and ANN need a good activation function. GA needs a good objective function but Fuzzy logic does not need any function.
8. ***Probability Metrics (MSE):*** The acceptable probability metrics of mean square error must be low as it depends on the nature of the data.
9. ***Rank Metrics (ROC):*** The Rank metrics especially for biological system must be high. This is the ratio of specificity to the sensitivity.
10. ***Training Time:*** Is the complexity of implementation high or low? The strengths of SVM lie mainly in its relative ease of training, and absence of local optima unlike ANN. FL has complexity of implementation and hence too much time spent in tuning the parameters used for inferencing during the process of training, compared to other techniques (Anifowose & Abdulraheem, 2011). GA takes a lot of time to tune the parameters.
11. ***Number of Parameters:*** Does the algorithm obtain good performance when the number of training data parameters is small? FL performs better when the number of training prototypes is large. SVM does well when the number of parameters is small while ANN does not. GA does well whether the parameters are small or not (Mendel, 2003).
12. ***Number of Training Data:*** Does the algorithm obtain good performance when the number of training data is small? SVM scales relatively well to high dimensional data and obtain good performance when the number of training data is small (Anifowose & Abdulraheem, 2011). FL does not obtain good performance when the number of training data is small. GA does well whether the features are small or not.

13. ***Application Capability:*** This capability of GA and SVM is good for variable selection and dimensionality reduction of input variables. ANN and SVM application involves pattern recognition, regression estimation, and time series prediction (Jian & Wenfen, 2006). FL is strong in the process of extracting rules, which is used for inferencing.

2.7 Artificial Neural Network (ANN) Model

2.7.1 A Neuron

A neuron is a special biological cell that processes information from one neuron to another by electrical and chemical charge. An Artificial Neuron is basically an engineering approach of biological neurons in Figure 2.15

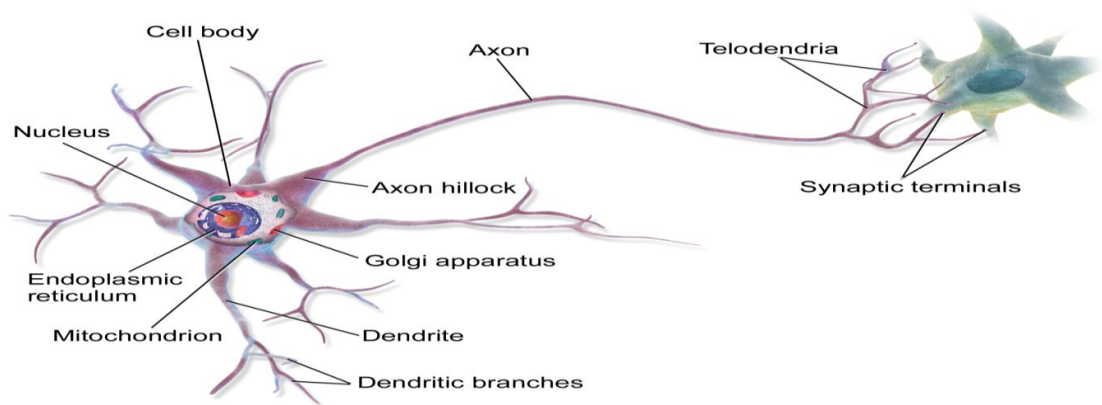


Figure 2.15: Human Body Network of Neurons

The history of neural networks can be traced back to the modeling of neurons. The first model of a neuron was by physiologists - McCulloch and Pitts (1943). The perceptron is a very simple mathematical representation of neurons as shown in Figure 2.16.

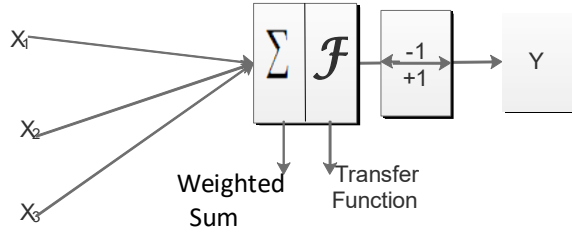


Figure 2.16: A Perceptron

In order to solve an n-separable problem perceptrons show that n-1 nodes are needed. A perceptron could then only solve a 2-separable problem, or a linearly separable problem.

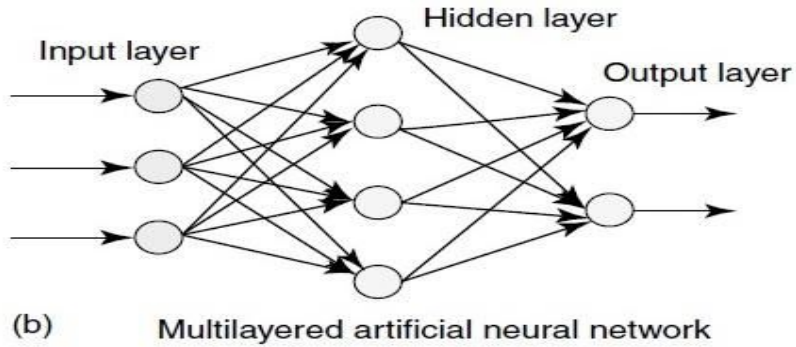


Figure 2.17

For non-linear inseparable problems multi-layered artificial neural network in figure (b) is used. Mathematically we may describe a neuron r with this pair of equation (2.4) and (2.5)

$$N_r = \sum_{j=0}^n w_{rj} x_j \quad (2.4)$$

$$y_k = \varphi(N_r) \quad (2.5)$$

or in this matrix form in equation (2.6):

$$N_r = [w_{r0} \quad w_{r1} \quad \dots \quad w_{rn}] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = w_r^t x \quad (2.6)$$

Where φ is the activation function that mapped input x to output y.

An *Artificial* Neural network is a web of interconnected millions of neurons where parallel processing takes place (Sharma *et al.*, 2012). ANN tool solves classification problems, pattern recognition problems, non-linear mapping, classification learning and real time optimization (Lizhe *et al.*, 2010, QI *et al.*, 2010). Prediction models are built with ANN because of their ability to learn from experience without prior knowledge of the system (Gundin, 2002).

Artificial Neural Networks (ANNs) and statistical methods in literature face similar classification and prediction problems. ANNs have an underlying statistical methodology sometimes, appropriately designated statistical neural networks. With the appearance of ANNs, some statistical procedures such as regression, discriminant analysis, cluster analysis and time series are being defined (Jain, Mao & Mohiuddin, 1996).

2.7.2 Network Structures

ANN network structure are of two types Feed-Forward Neural Network (FFN) also called associative or Recurrent Structure and Back-propagation Neural Network (BPN) also called Auto-Associative or Non-Recurrent Structure.

2.7.3 Back-propagation Neural Network (BPN)

N-separable problems using back propagation algorithm was developed by Werbos (1974). Widrow-Hoff LMS algorithm is the basis for the generalization of back-propagation and allowed perceptrons to be trained in a multilayered configuration, thus a n-1 node neural network could be constructed and trained. Adjustments are made to the weights based on the error between the output and some known desired output (Werbos, 1988).

The back-propagation algorithm uses supervised learning, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea of the back-propagation algorithm is to reduce this error, until the ANN *learns* the training data. The training begins with random weights, and the goal is to adjust them so that the error will be minimal.

Back propagation algorithm involves adjusting the weights backwardly through the neural network, starting with the output layer and working through each hidden layer until the input layer is reached. Parker (1985), Rumelhart & McClelland (1986) then independently rediscovered Back propagation algorithm simultaneously (Jacobs, Jordan & Barto, 1991). The back-propagation algorithm is used in layered feed-forward ANNs. This means that the artificial neurons are organized in layers and they send their signals “forward”, and then the errors are propagated backwards. Then network receives inputs by neurons in the *input layer*, and the output of the network is given by the neurons on an *output layer*. There may be one or more intermediate *hidden layers*.

2.7.4 Feed-Forward Neural Network (FFN)

A network of perceptrons with a differentiable squashing function, usually the sigmoid (logsig) function is called feed-forward neural network. The back propagation algorithm (BPn) adjusts the weights based on the idea of minimizing the squared error. The differentiable squashing function of feed-forward neural network enables the back propagation algorithm to adjust the weights across multiple hidden layers.

By having multiple nodes on each layer, the Exclusive-OR, or the XOR problem and n-separable problems can be solved by FFN and BPN could not be solved with only the perceptron. Figure 2.17 shows a fully connected feed-forward neural network; from input to output, each node is connected to every node on the adjacent layers. In this research, a malaria severity analysis artificial neural network (ANN), based on feedforward-error back-propagation algorithm is proposed.

By applying the same procedure the test patterns belonging to normal, non-severe and severe malaria (N, NS, S) are identified as Class 0, Class 1 and Class 2 as shown in the diagram in Figure 2.18:

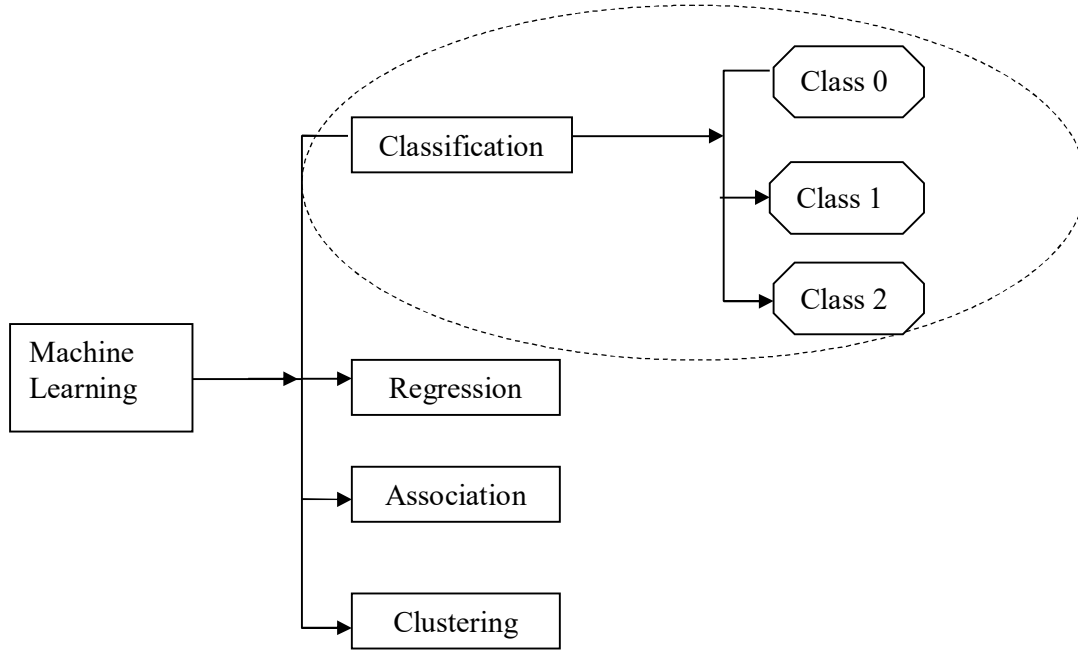


Figure 2.18: Malaria Classification Schema

2.8 Artificial Neural Network Activation Functions

Activation function translates the input signals to output signals. The activation function of the artificial neurons in ANNs implementing the j^{th} back-propagation algorithm is a weighted sum of the inputs x multiplied by their respective weights $w_{r,j}$ in equation (2.7) :

$$\phi_r(\bar{x}, \bar{w}) = \sum_{j=0}^n w_{r,j} x_j \quad (2.7)$$

Three types of transfer functions are commonly used *purelin*, *sigmoid*, and *tansig* functions for multilayer networks, but other differentiable transfer functions can be created and used if desired.

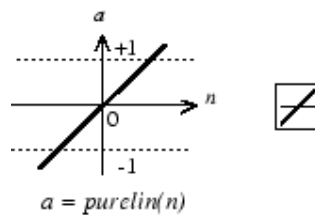
2.8.1 Purelin Function:

We can see that the activation depends only on the inputs and the weights. If the output function would be the identity (output=activation), then the neuron would be called linear and it is called Purelin.

A linear activation function transforms the weighted sum inputs of the neuron to an output using a linear function. In the linear function:

$f(x) = x$ the inputs and outputs range from minus infinity to plus infinity, which

X is being used in the output layer of the network. The linear transfer function *purelin* is shown in Figure 2.19.



Linear Transfer Function **Figure 2.19**

2.8.2 Sigmoid Function

The sigmoid function consists of 2 functions, logistic and tangential. The sigmoid function is very close to one for large positive numbers, 0.5 at zero, and very close to zero for large negative numbers. This allows a smooth transition between the low and high output of the neuron (close to zero or close to one). We can see that the output depends only in the activation, which in turn depends on the values of the inputs and their respective weights. Now, the goal of the training process is to obtain a desired output when certain inputs are given. Since the error is the

difference between the actual and the desired output, the error depends on the weights, and we need to adjust the weights in order to minimize the error.

2.8.3 Logistic Sigmoid Function(Log-Sigmoid)

The values of logistic function range from 0 and 1. Multilayered networks often use the log-sigmoid transfer function in *Figure 2.20* with equation (2.8)

$$S_r(\bar{x}, \bar{w}) = \frac{1}{1 + e^{-\phi_r(\bar{x}, \bar{w})}} \quad (2.8)$$

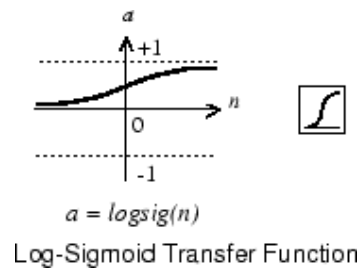


Figure 2.20

2.8.4 Tangential Sigmoid Function (Tansig Function)

The values of tangential function range from -1 to +1 as shown in Figure 2.21. Tan-Sigmoid output neurons are often used for pattern recognition problems while linear output neurons are used for function fitting problems.

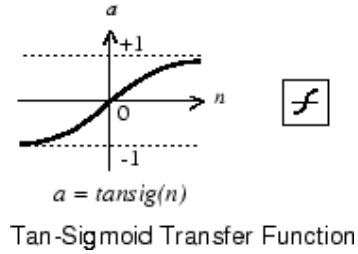


Figure 2.21

2.9 Support Vector Machine Model

Support Vector Machine a well-known machine learning approach which is recently applied in the variety of fields such as computing, hydrology and environmental researches (Asefa, Kemblowski, McKee, Khalil, 2006; Ji & Sun, 2013; Joachims, 1998; Mukkamala, Janoski & Sung, 2002). SVM is mainly utilized in pattern recognition, prediction, classification and regression analysis with its root in structural risk minimization. It has been proved that its applications show superior performance compared to prior developed methodologies such as neural network and other conventional statistical models (Sung & Mukkamala 2003; Vapnik, Golowich, Smola, 1996; Huang, Davis, Townshend, 2002; Collobert & Bengio, 2000; Mohammadi *et al.*, 2015; Rajasekaran, Gayathri, & Lee, 2008).

SVM are supervised machine learning in which the computer is presented with inputs and their desired outputs and the goal is to learn a general rule that maps inputs to outputs (Nilsson, 1998). SVM is a machine learning model that focuses on prediction based on known properties (historical data) learned from the training data. SVMs are a set of related supervised learning (labeled training) data used for classification, recognizing patterns and regression analysis from a given data set.

SVM is based on the principle of statistical machine learning and structural risk minimization, which minimizes the upper bound generalization error compared to local training error, which is common technique in previously used machine learning methodologies (Vapnik & Vapnik, 1998). The mentioned technique proved advantages over other machine learning algorithms because (1) applying high dimensional spaced set of kernel equations, which discreetly include non-linear transformation; thus, there is no assumption in functional transformation which makes data linearly separable indispensable and (2) unique solution due to the convex nature of the optimal problem (Mohammadi *et al.*, 2015).

2.9.1 Theory of Support Vector Machine

The detail of theory and evolution of SVM developed by Vapnik's can be found in Vapnik, 2000 and Vapnik & Vapnik, 1998.

Let $\{x_i, y_i\}_{i=1}^n$ be the assumed training data sets where $x_i \in R_n$ is the input space vector of the sample data $y_i \in R_n$ is the respective target value for $i=1 \dots n$ and n is the data size.

SVM approximates the function as depicted in the equation 2.9 and 2.10

$$f(x) = w\phi(x) + b \quad (2.9)$$

$$R_{svms}(C) = \frac{1}{2}\|w\|^2 + C \frac{1}{n} \sum_{i=1}^n L(x_i - b_i) + C \frac{1}{n} \sum_{i=1}^n L(x_i - b_i) \quad (2.10)$$

Where $\phi(x)$ represents high dimensional space features that mapped the input space vector x , w is a normal vector, b_i is a scalar quantity and

$C \frac{1}{n} \sum_{i=1}^n L(x_i - b_i)$ depicts the empirical error where the parameter w can be estimated by minimization of regularized risk function in equation (2.11) after introduction of positive slack variables ε_i and ε_i^* which represents upper and lower excess deviation (Vapnik & Vapnik, 1998).

$$R_{svms}(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{n} \sum_{i=1}^n (\varepsilon_i + \varepsilon_i^*) \quad (2.11)$$

Subject to

$$\begin{cases} y_i - w\phi(x_i) + b_i \leq \varepsilon_i + \varepsilon_i^* \\ w\phi(x_i) + b_i - y_i \leq \varepsilon_i + \varepsilon_i^* \\ \varepsilon_i, \varepsilon_i^* \leq 0, i, \dots, l \end{cases}$$

where $\frac{1}{2} \|w\|^2$ represents the regularization term.

C is the error penalty factor or regularization constant.

Equation 2.10 can be solved with the introduction of Lagrange Multipliers ($\alpha, \alpha_i^*, n, n^*$) and optimality constraints, hence obtained a generic function in equation (2.12);

$$f(x_i, \alpha_i, \alpha_i^*) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x, x_i) \quad (2.12)$$

where $k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ and $k(x_i, x_j)$ is the kernel function which is the product of the two inner vector x_i and x_j in the feature space $\phi(x_i)$ and $\phi(x_j)$ respectively.

2.9.2 Support Vector Machine Classification

A Support Vector Machine (SVM) does prediction of classes by finding the hyperplane that maximizes the margin between the two classes. Support Vector Classification deals with binary class problem as shown in Figure 2.23 without loss of generality. Its goal is to separate the two classes by a function induced from an available examples by a classifier called '*optimal separation hyperplane*'. This optimal classifier separates the data by finding the maximum margin among the two classes of the data as shown in Figure 2.23.

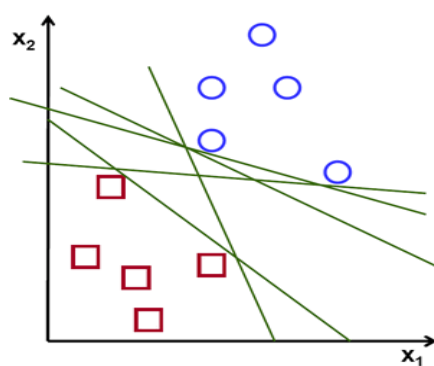


Figure 2.22: Linearly separable

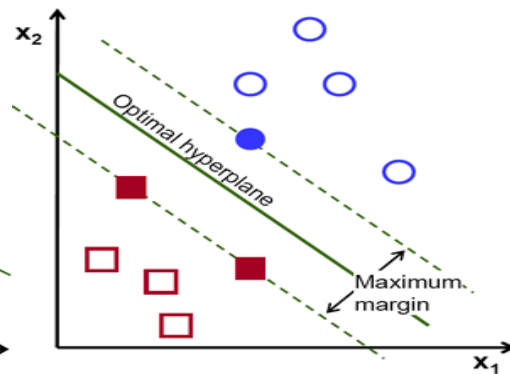


Figure 2.23: Optimal separating two hyperplane class of data

(Cortes & Vapnik 1995)

2.9.3 Support Vector Machine Multiclass Classification

Support vector machine is a binary classifier but can also handle multiclass problems by using either one-against-all or one-against-one algorithm. The winning class is then determined by the highest output function. This makes the multiclass SVM to perform slower than the MLPs. Neural networks computational complexity depends on the dimensionality of the input space. But one of the drawbacks of the SVM is the large number of support vectors used from the training set to perform classification task. However, SVM is still considered to be powerful classifier, soon to be replacing the ANNs.

2.9.4 Support Vector Machine Feature Selection

Feature Selection aim is to select features that lead to a large between class distance and small within class variance in the feature vector space. It finds a subset of the original features, that will generate a classifier with the highest possible accuracy. There are quantitative (continuous), ordinal and categorical (nominal/discrete) types of features. Some classifiers like Naïve Bayes, decision trees, treat categorical and quantitative features differently.

- i. Significance of feature selection**
- ii. To have a better understanding of the data and the classification rule.
- iii. To avoid computational complexity, we need to reduce the number of features to a sufficient minimum.
- iv. To improve classifier performance; because some machine learning algorithms, are known to degrade in performance when faced with many irrelevant/noisy features.

b. Feature Selection Approaches

- i. **Filter-** Features or groups of features are scored by some measure of correlation with the labels using scoring functions like Pearson correlation between a feature and the labels, Fisher scores, or mutual information.
- ii. **Wrapper.** Wrapper uses the classifier to guide the process of selection. A feature subset is typically evaluated using cross validation.
- iii. **Embedded.** In embedded method, the selection method uses properties of the classifier.
- iv. **Support Vector Machine:** The hyper parameters are searched and selected by SVM.

The performance of SVM depends on appropriate parameters selection which is very complex in nature and quite hard to solve by conventional optimization techniques (Bao, Hu, & Xiong, 2013; Friedrichs, & Igel, 2005; Lorena & Carvalho. 2008). In general the hyper parameters are determined using sequential grid search (Hsu, Chang & Lin, 2003). In this method, the hyper parameters are searched in sequence with a fixed step size. The major drawback of grid search is the computational complexity. Hence, grid search is applicable only in those areas where very few parameters are involved. Further, researchers have applied gradient descent techniques for searching out the hyper parameters (Chung, Kao, Sun, Wang, Lin, 2003; Chapelle, Vapnik, Bousquet, & Mukherjee, 2002). But the gradient descent techniques are prone to local minima. In most of the optimization problems, there may be multiple local solution set off parameters. Evolutionary algorithms have the capability to provide global solutions to such problems.

Figure 2.11 is an SVM training algorithm builds a model that predicts whether a new example falls into one category or the other or learned from historical examples. SVM models take a set of input data and predict for each given input the possible output which makes them non probabilistic binary linear class as shown in Figure 2.24 (Ch *et al.*,2013). The operation of SVM algorithm is based on finding the hyperplane that gives the highest minimum distance to the training examples. So SVM is a discriminative selector formally defined by a separating hyperplane.

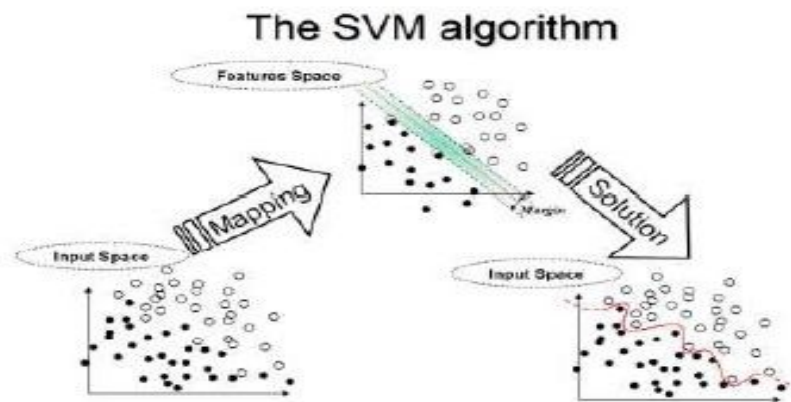


Figure 2.24: An SVM Algorithm (Vapnik & Vapnik, 1998)

The operation of the SVMs algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples. Therefore, the optimal separating hyperplane maximizes the margin of the training data.

2.10 Support Vector Machine Kernel Function:

Training vectors x_i is mapped into a higher dimensional space by a function ϕ . Then SVM finds a linear separating hyper plane with the maximal margin in this higher dimensional space. Furthermore, $k(x_i, y_i) = \phi(x_i) \phi(y_i)$ is called the kernel functions. There are number of kernels that can be used in SVM models. These include linear polynomial, RBF and sigmoid (Rejani & Selvi, 2009); Sigmoid, lineal, polynomial and radial basis functions are the four basic kernel functions which are provided by SVMs.

2.10.1 Linear Kernel Function

The Linear kernel in equation (2.13) is the simplest kernel function. It is given by the inner product x_i and y_j .

$$K(x, y) = (x_i * y_j) + c \quad (2.13)$$

plus an optional constant c . K represents the linear kernel with a function of x and y .

2.10.2 Gaussian (Radial Basis)Function

The Gaussian kernel is an example of radial basis kernel function. Among all, radial basis function (RBF) has been considered as the best kernel features due to its computationally effectiveness, reliability, simplicity, ease of adaption for optimization as well as its adaptability in handling factors which are more complicated (Rajasekaran, Gayathri, & Lee, 2008; Yang, Huang, King & Lyu, 2009). Only the solution of a set of linear functions is required for the training of RBF kernel equation is represented in equation (2.14) rather than the lengthy and

complicated demanding quadratic programming problem (Shamshirband *et al.*,2009).

$$k(x,y) = \exp(-\gamma|\mathbf{x}_i - \mathbf{y}_j|^2) \quad (2.14)$$

Alternatively, it could also be implemented using equation (2.15)

$$k(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (2.15)$$

2.10.3 Polynomial Kernel Function

Polynomial kernel is a non-stationary kernel function. Polynomial kernel function is represented in equation (2.16) and well suited for normalized training data problems.

$$k(x,y) = (\alpha x^T y + c)^d \quad (2.16)$$

Alpha(α) , constant (c) and polynomial degree (d) are the adjustable parameters.

2.10.4 Sigmoid Kernel Function

Sigmoid Kernel Function is also known as the Hyperbolic Tangent Kernel being used for Multilayer Perceptron (MLP). The Sigmoid Kernel function in equation (2.17) comes from the Neural Networks field, where the bipolar sigmoid function is often used as an activation function for artificial neurons.

$$k(x, y) = \tanh(\alpha x^T y + c) \quad (2.17)$$

SVM model using a Sigmoid Kernel function is equivalent to a two-layer, perceptron neural network. This kernel function was quite popular for support vector machines due to its origin from neural network theory. Also, it has been found to perform well in practice.

There are two adjustable parameters in the Sigmoid Kernel function, the intercept constant c and the slope α ($1/N$) where N is the data dimension that is defined by the user (Rejani & Selvi, 2009).

2.11 Review of Related Work

The existing researches related to this work have some limitations that influenced the benchmark for this research. They serve as the baseline of this research work as shown in Table 2.3. Consequently, they form the basis for performance evaluation of the new method.

Table 2.3: Related Work

Author	Methodology	Motivation	Results	Limitation
Aminu, Onyebuchi & Ibrahim(2016)	SVM	High cost of medical equipment and insufficient number of medical specialists.	The model achieved 90% accuracy.	The model captured relatively low dataset. Other Performance metrics were not considered. No global thresholding and consideration of effects of climatic conditions.
Shruti & Shirgan (2015)	ANN and SVM	We need to secure and protect our lives from malaria threat that maims millions of lives.	On 70 images of blood cells ANN gave accuracy of 78.53 and running time of 6.58 seconds while SVM gave accuracy of 98.25	The study focuses only on morphological characteristics and novel threshold technique
Chaudhari & Agrawal, 2015	SVM and Naive bayes classifier and two feature extraction techniques DWT and GLCM feature extraction technique were used	Conventional microscopy used in malaria diagnosis s occasionally proved is time consuming and results are difficult to reproduce. Alternative diagnosis techniques which yield superior results are quite expensive and hence inaccessible to developing countries where the disease is endemic	The study revealed that artificial neural network (ANN) classifiers are suitable for detection and classification of Plasmodium parasites into their respective stages and species.	The method makes use of the morphological, colour and texture features of Plasmodium parasites and erythrocytes.
H. Chiroma et al.,(2014)	Jordan-Elman Neural Network	There is a need for an alternative approach for medical practitioners to detect malaria and accurately identify the level of severity.	Out of total of 57 cases of Mildly Severe(MS) malaria, 54 were correctly classified and three were wrongly classified as Highly Severe (HS) malaria. In the case of HS malaria, of a total of 33 cases, 31 were correctly classified and two were incorrectly classified as MS malaria.	The study considered only the morphological characteristics of malaria parasite.
Chavan, & Sutkar (2014)	Malaria Disease Identification and Analysis Using Image Processing	Analysis of malaria using manual microscopy diagnostic method is prone to human error and time consuming	The analysis of the identification of malaria disease using image processing todetect various features in the blood image. The statistical analysis of Skewness, Kurtosis, Energy and standard	The method was based on data image processing not on real life datasets.

deviation was conducted on the blood image. The analysis generated an accuracy of 98.25% for SVM classifier and 78.53% for ANN Classifier.

<i>Dipti D. Patankar, et al., (2014)</i>	Naive Bayes classifier and Bayesian regularization	Traditionally, light microscopy is used for malaria detection which is less accurate and time consuming and costly	A probabilistic model Naive Bayes classifier is more accurate than Bayesian regularization.	Bayesian Regularization requires more memory to build the network than naive Bayes classifier also it requires 10% more time for its execution.
<i>Júnior et al., (2010)</i>	Artificial Neural Networks and Bayesian Networks as Supporting Tools for diagnosis of asymptomatic malaria	Since asymptomatic carriers generally present with very low parasitemia, the microscopy test does not generate a reliable diagnosis for this form of malaria. Thus, innovative tools are urgently needed to support a feasible approach to overcome the menace of this disease.	The ANN presented 67.5% of sensitivity and 92.5% of specificity, and has correctly diagnosed 80.0% of the 80 individuals separated for testing the method. The BN presented 37.5% of sensitivity and 97.5% of specificity, and has correctly diagnosed 67.0% of the 80 individuals separated for testing the method.	Use of a l different computational technique (like fuzzy logic) to create a more robust diagnosis technique for asymptomatic malaria. Further data of other relevant features that can be used to construct a better classifier.
<i>Sudheer Ch et al., (2013)</i>	Support Vector Machine-Firefly Algorithm	The performance of SVM mainly depends on its appropriate parameters selection which is very complex in nature and quite hard to solve by conventional optimization techniques	The results indicate that the proposed SVM-FFA model provides more accurate predictions compared to the other traditional techniques.	The utility of SVM-FFA in detecting the malarial cases for other geographical regions was not considered. Also other lead times (such as bi-monthly, quarterly or yearly prediction) and comparison of the SVM-FFA with the other new methods such as extreme learning machines were not considered.
<i>SavKare & Narote(2012)</i>	Otsu Threshold, SVM Binary Classifier and SVM multi classifier	Manual Counting and classifications	SVM binary classifier gives 96.26% sensitivity and 99.09% specificity on 71 images. SVM multiclass classifier (i.e. RBF kernel) gives 96.42% accuracy for correct identification rate of life stage of parasite	Thin blood Imaging system was used.
<i>Osowski et al.,(2009)</i>	GA and SVM	Genetic Algorithm(GA) was applied with Support vector machine(SVM) for the recognition of blood cells in the image of bone marrow .	NA	NA

2.12 Chapter Summary

This chapter explored the basic concepts of malaria, factors affecting malaria transmission, and malaria threats. Various existing malaria threat models were reviewed alongside their strength and weaknesses including those of the various computational intelligence tools. The choice of machine learning algorithm depends on the complex nature of the biological systems. Review of ANN and SVM malaria based existing models was explored. The study revealed that most of the baseline models considered the morphological characteristics of the thick and thin blood smear of malaria cases but the symptomatic nature of the disease was not considered. Also, the existing studies focused on the binary nature of malaria cases and influences of climatic factors were not considered. So a choice of SVM_ANN hybridized malaria model was made to predict the severity of malaria cases. The next chapter presents the methodology adopted for the development of SVM_ANN hybridized malaria model.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

This chapter consists of these sections: the block diagram of the research design, an optimal separation hyperplane for malaria, dataset collection, feature selection, normalization approach used, malaria model features description, quantitative and qualitative measure, computation of quantitative and qualitative measure and binary encoding for malaria threat severity.

3.2 Research Design

This research adopts descriptive and experimental case study design method. The entire design framework of the ANN and SVM model is divided into sub phases training, testing and validation each phase is directly mapped to each research objective stated in Chapter One (Refaeilzadeh, Tang, & Liu, 2009; Reitermanova, Z, 2010). The Design Framework of the ANN and SVM is presented in Figure 3.1(a) and Figure 3.1(b). The Architecture of the Proposed SVM_ANN Hybridized Malaria Prediction Model for training, testing and validation phase is also presented in Figure 3.1(c).

3.2.1 System Analysis

The first step undergone was the qualitative research of patients' observations (symptoms) and Laboratory test results documented and get report analysis to identify malaria threats and attacks on human health.

3.2.2 Likelihood Quantitative model

According to the mathematical modeling of infectious disease dynamics technique which involved statistical based methods, mathematical/mechanistic state- space model and empirical/machine learning based model literature were reviewed in Figure 2.4 (Siettos & Russo, 2013). Then the likelihood quantitative model was selected. Recently, various computational intelligence techniques are efficient approaches for the problem of global malarial prediction in many locations across the globe. The Choice of machine learning algorithm to solve a problem always depends on the size, quality, and nature of the data (Djam *et al.*, 2011). In this study after consideration of strength and weaknesses of various Computational Intelligence (CI) Tools; Fuzzy Logic (FL), Artificial Neural Network (ANN), Support Vector Machine (SVM) and Genetic Algorithm (GA), as shown in Table Figure 3.1, the choice of SVM and ANN were made.

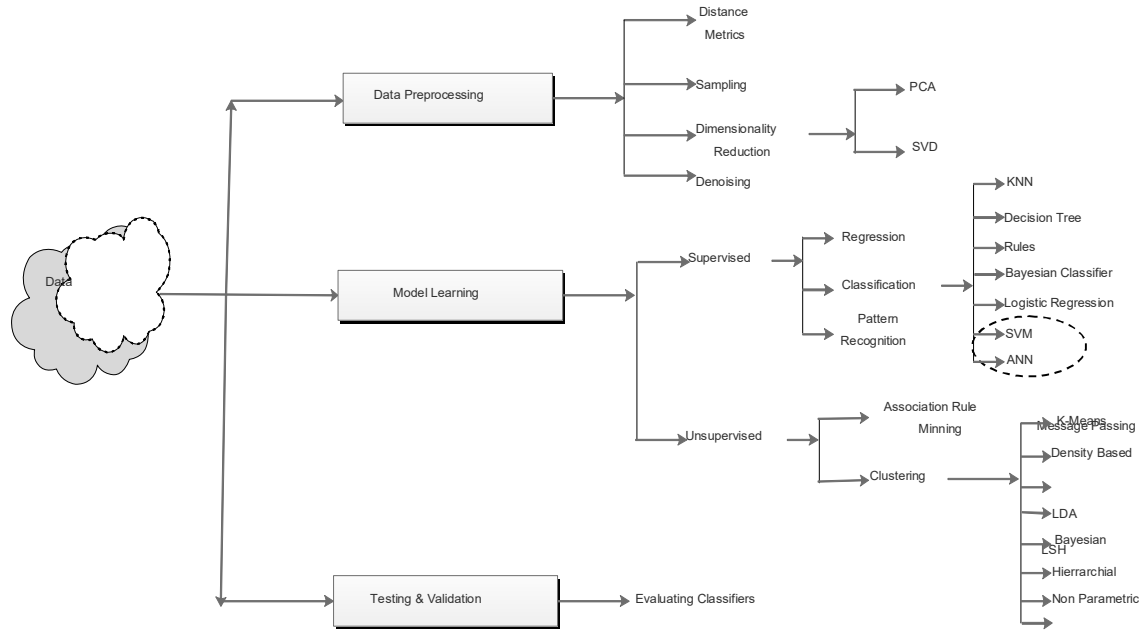


Figure 3.1: Method and Steps in data Mining Process

Then, the data is subjected to machine learning models such as Neural network, Support Vector Machine, Naïve Bayes, Fuzzy Logic and Genetic Algorithm etc.

Therefore, the likelihood computational models selected were Artificial Neural Network and Support Vector Machine.

3.2.3 ANN and SVM Computational Model

ANN needs a good activation function to choose an adaptive threshold value. Also, SVM needs a good kernel function to choose the best parameters. From the literature purelin, log-sigmoid and tansig functions are the three activation functions provided by ANN. Consequently, Sigmoid, linear, polynomial and radial basis functions are the four basic kernel functions which are provided by SVMs (Hsu, Chang & Lin, 2008; Joachims,1998). Among all, radial basis function (RBF) has been considered as the best kernel feature due to its

computationally effectiveness, reliability, simplicity, ease of adaption for optimization as well as its adaptability in handling factors which are more complicated (Rajasekaran, Gayathri & Lee, 2008; Yang, Huang, King & Lyu.,2009; Wu & Wang, 2009).

3.2.3.1 The Flow Diagram of ANN to Predict Malaria Parasite (MP)Count

The malaria cases data are collected consisting of input features and target features. The outliers are removed and reprocessed, then further divided into training, testing and validation according to hold out cross validation method. The data is then subjected to the ANN classifier as shown in Figure 3.2 which depicts the flow diagram of the architecture of ANN to predict Malaria Parasite Count. ANN network is created and trained with the data to get the optimal threshold values. This optimal threshold values are then used to simulate the network and get the appropriate results after several iterations.

3.2.3.2 The Flow Diagram of SVM to Predict Malaria Parasite (MP)Count

SVM classifier requires cross validation to detect relevant features and divide the data into training, testing and validation after pre-processing (Refaeilzadeh, Tang, & Liu, 2009; Reitermanova, 2010). The SVM network is also trained to get the best threshold values to simulate the network. The appropriate results for the model is then derived after several simulations as shown in Figure 3.3 which depict The Flow Diagram of the architecture of SVM to predict Malaria Parasite (MP) Count.

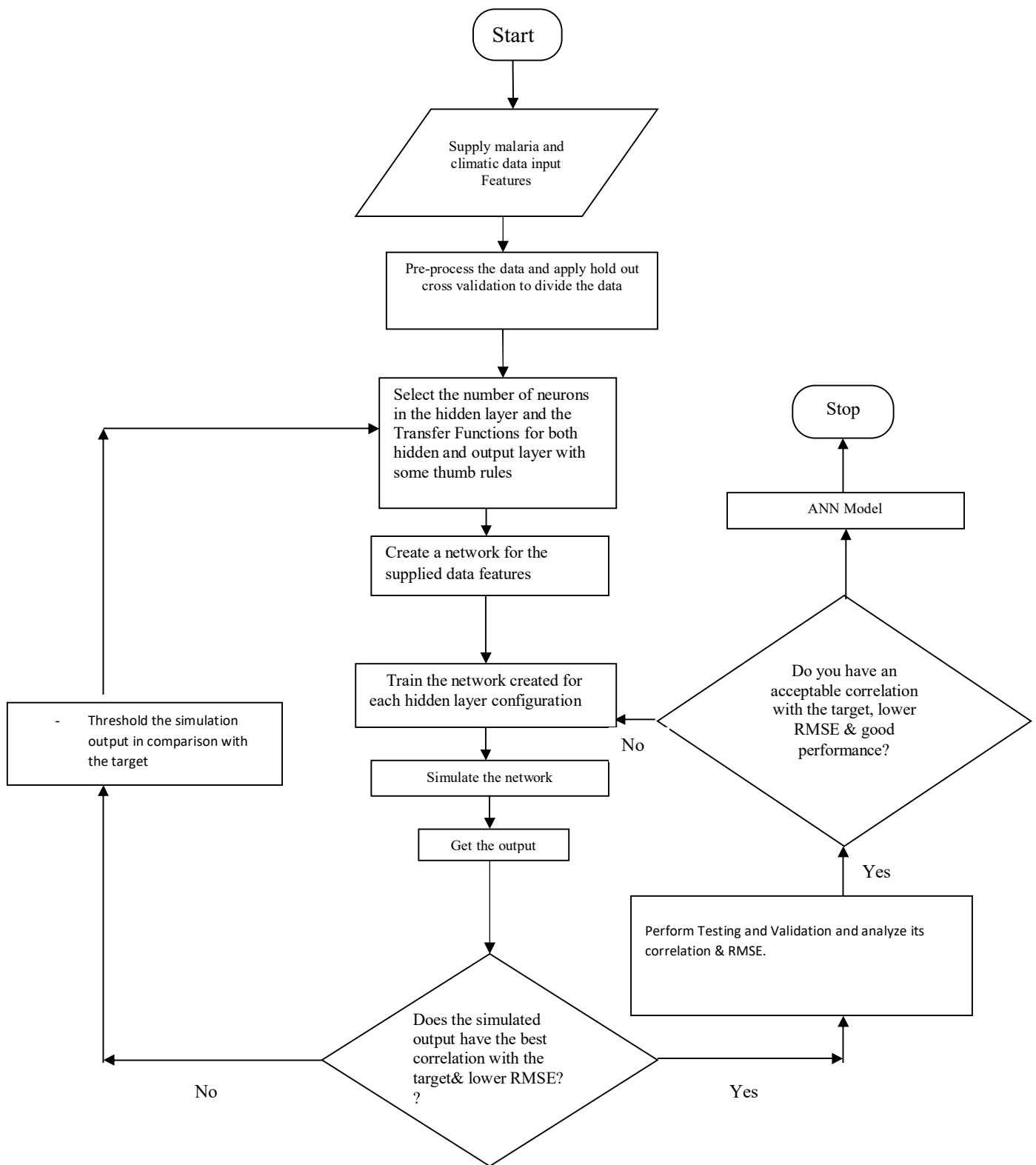


Figure 3.2: Flow Diagram of of ANN Predictor

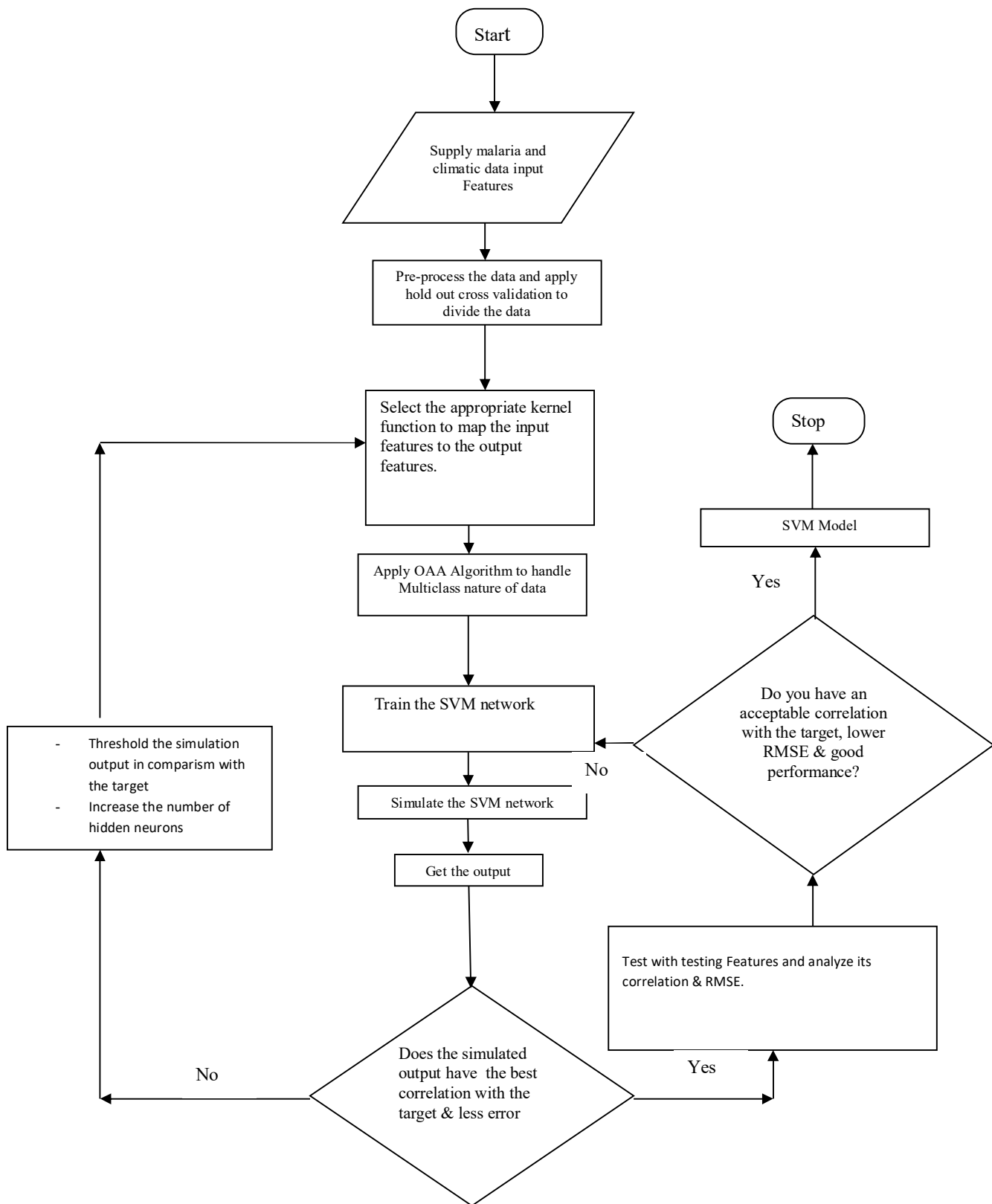


Figure 3.3: Flow Diagram of the architecture of SVM Selector

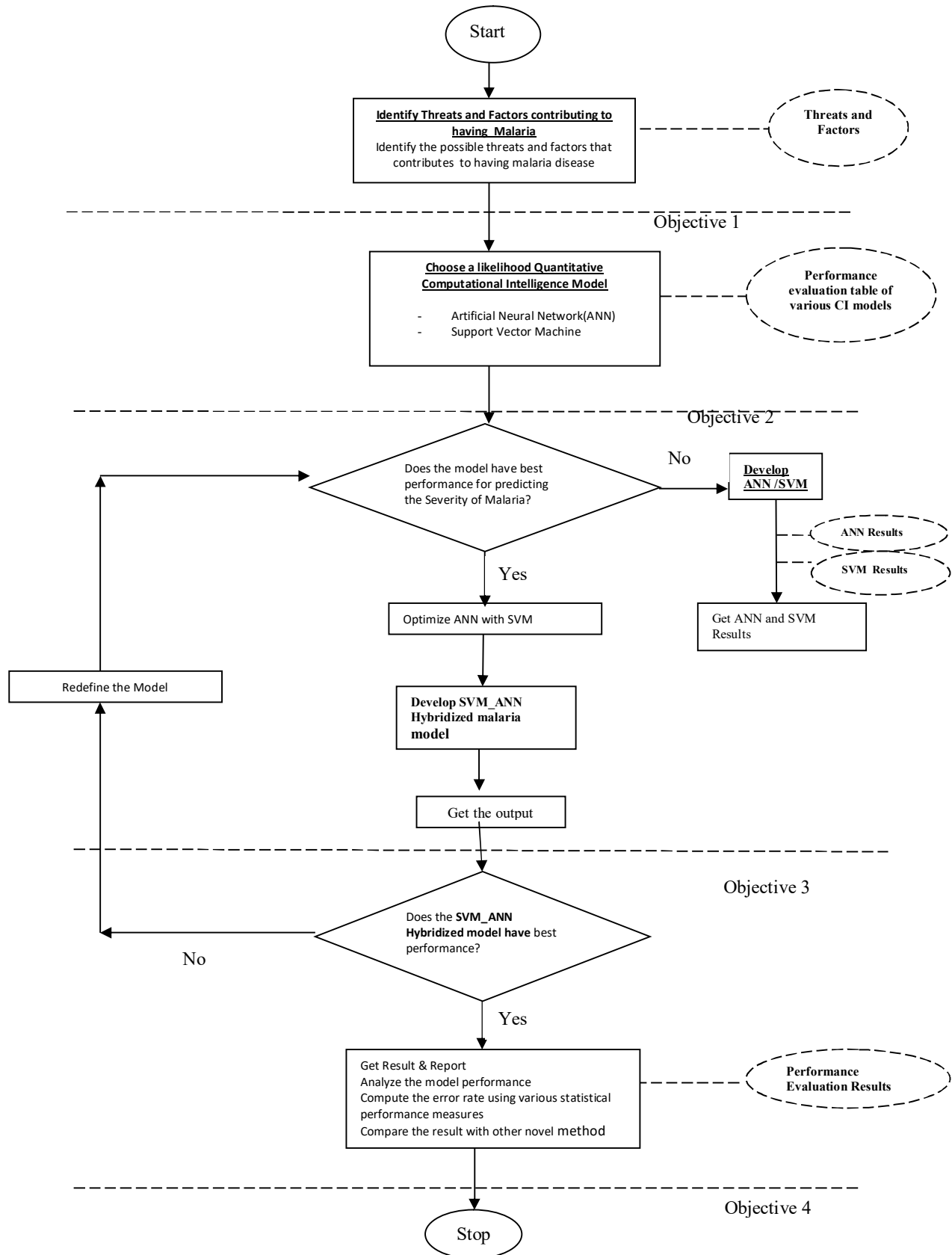


Figure 3.4: Research Design of SVM_ANN Hybridized Malaria Model

3.2.3.3 Research Design of the SVM_ANN Hybridized Model

The possible threats and factors that contribute to having malaria disease were identified; the possible computational intelligence model is selected based on the nature of the data and choice of application. The data is simulated and the optimal results were collected. The developed SVM and ANN models did not give higher accuracy because of the problem of multiclass nature of data, large number of parameters to fix and prone to local minimum error. Therefore, the Hybridized SVM_ANN Malaria Model is proposed. The hyper parameters (support vectors) are searched in SVM and then fed into the ANN network to form Hybridized SVM_ANN Malaria Model. In this model, SVM does feature selection and ANN predicts the class of malaria cases. Figure 3.4 depicts the research design of the hybridized SVM_ANN model.

3.3 Population Sample and Sampling Procedure

Consider the problem of classifying this set of training vectors of two classes of malaria patterns

$$M = \{(x_1, y_1), \dots, (x_n, y_n)\}, x_i \in \mathbb{R}_n, y_i \in \{-1, 1\} \quad (3.0)$$

where $\{x_i, y_i\}_{i=1}^n$ is the assumed training data sets where $x_i \in \mathbb{R}_n$ is the input space vector of real number $\mathbb{R}_n$ and the sample data $y_i \in \{-1, 1\}$ is the target value.

3.4 SVM Malaria Model

In this research, a linear classifier that can perfectly separate the features is deduced as follows:

$$(w,x)+b = 0 \quad (3.10)$$

This linear classifier function is:

$$f(x) = \sum_{i=1}^n (w_i \cdot x_i) + b \quad (3.11)$$

or

$$f(x) = w\phi(x) + b \quad (3.12)$$

Where $\phi(x)$ represents high dimensional space features that mapped the input space vector x , w is a weight vector, b is a scalar quantity or bias.

A linear classifier (hyperplane) that can optimally separates the malaria data is the one that minimizes

$$M(w) = \frac{1}{2} ||w||^2 \quad (3.13)$$

Equation 3.13 is independent of b provided this constraint in Equation 3.14 is satisfied (it is a separating hyperplane), changing b will move it in the normal direction to itself.

$$y_i[(w,x_i)] + b \geq 1, \quad i=1, \dots, n \quad (3.14)$$

The solution to the optimization problem of equation 3.13 subjected to the constraint in equation 3.14 is given by the introduction of Langrange function

$$M(w,b,\alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^n \alpha_i (y_i [(w, x_i) + b] - 1), \quad (3.15)$$

Where α is the langrange multiplier. The Langragian will be minimized with respect to w, b and maximized with respect to $\alpha \geq 0$. Thus equation 3.15 is transformed to its dual problem as thus:

$$\max_{\alpha} \text{Mal}(\alpha) = \max_{\alpha} [(\min M(w,b,\alpha))] \quad (3.16)$$

The minimum with respect to w and b of the Langrangian, M , is given by:

$$\frac{\partial M}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.17)$$

$$\frac{\partial M}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i x_i \quad (3.18)$$

Thus, the dual problem from equation (3.16),(3.17) and (3.18) is given by:

$$\max_{\alpha} \text{Mal}(\alpha) = \max_{\alpha} -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i, x_j) + \sum_{k=1}^n \alpha_k. \quad (3.19)$$

Hence, the solution to the problem is given by

$$\alpha^{mal} = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i, x_j) - \sum_{k=1}^n \alpha_k, \quad (3.20)$$

with constraints

$$\alpha_i \geq 0 \quad i = 1, \dots, n \quad \sum_{j=1}^n \alpha_j y_j = 0 \quad (3.21)$$

Thus Equation 3.20 with constraints in equation 3.21 determines the Lagrange

Multipliers and the optimal separating hyperplane is given by:

$$w^{mal} = \sum_{j=1}^n \alpha_j y_j x_j \quad (3.22)$$

$$b^{mal} = \frac{1}{2} (w^{mal}, x_c + x_s) \quad (3.23)$$

Where $x_c + x_s$ are any support vector from each class satisfying $\alpha_c, \alpha_s > 0$, $y_c = -1$, $y_s = 1$.

To generalize the optimal separating hyperplane method, we introduced positive variables $\xi_i \geq 0$ and a penalty function:

$$F_{\sigma}(\xi) = \sum_i \xi_i^{\sigma} \quad \sigma > 0 \quad (3.24)$$

Where the ξ_i are the measure of misclassification errors.

From equation (3.13), we have :

$$M(w, \xi) = \frac{1}{2} \|w\|^2 + c \sum_i \xi_i^{\sigma} \quad (3.25)$$

Subjected to the constraint in 3.14

To deal with data that are not linearly separable, the previous analysis can be generalized by introducing some non-negative variables $\xi_i \geq 0$ such that (3.14) is modified to:

$$y_i(w, x_i) + b \geq 1 + \xi_i, \quad \xi_i \geq 0, \quad i=1, \dots, n \quad (3.26)$$

Where C is a regularization constant that has to be determined to prevent over-fitting or reflect the knowledge of the noise on the data and minimized the training error.

By applying Lagrangian theory, the Lagrangian will be minimized with respect to w, b, ξ and maximized with respect to α and β as done in equation 3.16 -3.18. Thus the dual problem is:

$$\max_{\alpha} \text{Mal}(\alpha) = \max_{\alpha} [(\min M(w, b, \alpha, \xi, \beta))] \quad (3.27)$$

The minimum with respect to w, b, ξ and of the Lagrangian, M is given by:

$$\frac{\partial M}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.28)$$

$$\frac{\partial M}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i x_i \quad (3.29)$$

$$\frac{\partial M}{\partial \xi} = 0 \Rightarrow \alpha_i \beta_i = C \quad (3.30)$$

Thus, the dual problem from equation (3.28),(3.29) and (3.30) is given by:

$$\max_{\alpha} \text{Mal}(\alpha) = \max_{\alpha} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i, x_j) + \sum_{k=1}^n \alpha_k, \quad (3.31)$$

Hence, the solution to the problem is given by:

$$\alpha^{mal} = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i, x_j) - \sum_{k=1}^n \alpha_k, \quad (3.32)$$

with constraints

$$0 \leq \alpha_i \leq C \quad i = 1, \dots, n \quad \sum_{j=1}^n \alpha_j y_j = 0 \quad (3.33).$$

To make generalization into high dimensional feature space, training vectors x_i is mapped into a higher dimensional space by a kernel function K . Then SVM finds a linear separating hyper plane with the maximal margin in this higher dimensional space. So $k(x_i, x_j) = \phi(x_i) \phi(x_j)$ is called the kernel functions.

Thus far, the optimization problem in equation 3.32 now becomes:

$$\alpha^{mal} = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{k=1}^n \alpha_k, \quad (3.34)$$

with constraints

$$0 \leq \alpha_i \leq C \quad i = 1, \dots, n \quad \sum_{j=1}^n \alpha_j y_j = 0 \quad (3.35)$$

Where $K(x_i, x_j)$ is the kernel function mapping (x_i, x_j) into feature Space X and

where $K(x_i, x_j)$ is an explicit dot product of $\phi : K(x_i, x_j) = [(\phi x_i), (\phi x_j)]$.

3.5 Instrument for Data Collection and Preparation

Monthly surveys of malarial incidences were collected from five (5) randomly sampled health centers in Minna metropolis, Niger State. Climatic data consisting of monthly averages of rainfall, temperature and relative humidity were collected from Nigerian Environmental and Climate Observation Programme (NECOP) Weather Station, Bosso Campus, Federal University of Technology, Minna, Niger state. Each patient has a set of symptoms and MP count known as Patients' malaria data symptoms and lab test results. This climatic data combined with monthly malaria incidences were considered as input variables was trained and simulated using Microsoft Excel and libSVM in MATLAB 2015a.

In the sampled hospitals laboratories, Giemsa staining was used for the laboratory tests. The red blood cells (RBCs), Plasmodium spp, platelets and other artifacts were identified. This Plasmodium spp is measured in count being called Malaria Parasite Count (MPcount). Each patient has a set of symptoms and MP count. This malaria dataset is used to train and test the performance of the developed Hybridized SVM_ANN malaria model. The dataset consists of 1,200 malaria cases between January, 2012 and December 2015. The dataset has these prevalent features: Headache (Hd), Fever(F), Dizziness(D), Body Pain(Bp) and Vomiting. The climatic factor; temperature, relative humidity and rainfall are contributing factors. The research features were presented in Table 3.1. According to hold out

cross validation procedure, random 800, 200, 200 patients data were used to train, test and validate the model respectively.

3.6 Data Analysis Method

Table 3.1: Malaria Model Features Description

Input	Variable	Description(Malaria Demographic)
	Age	Adult (1) Children(2)
	Gender	Male(1) Female(2)
	Headache(H_d)	+ve(1) -ve(0)
	Fever(F_v)	+ve(1) -ve(0)
	Dizziness(D_z)	+ve(1) -ve(0)
	Body Pain(B_p)	+ve(1) -ve(0)
	Vomiting(V_m)	+ve(1) -ve(0)
	Temperature(T_{emp})	$\{0 \leq T_{emp} \leq 32.83\}$
	Relative Humidity(R_h)	$\{0 \leq R_h \leq 83.74\}$
	Rainfall(R_r)	$\{0 \leq R_r \leq 0.034\}$
Output	Variable	Class
	MP Count	- (Class 0) + (Class 1) $\geq ++$ (Class2)

3.6.1 Preprocessing

3.6.1.1 Outlier Removal

Outliers are points lying very far to the mean of the corresponding random variables. The outliers are discarded.

3.6.1.2 Feature Normalization

Features with large values usually have a larger influence in the objective function than features with smaller values. This problem is overcome by normalizing the features so that their values lie within similar ranges. Data normalization or feature scaling is used to standardize the range of independent variables or features of data. Thus, feature normalization neutralizes the effect of different scales across features and makes the objective function accurate. Also feature normalization makes gradient descent converge much faster than without it.

3.6.1.3 Missing Data

There are different approaches to handling missing data. It varies from assigning missing values with zero or getting the unconditional and conditional mean from respective features value or discard feature vectors with missing values. In this research, features with missing data are assigned zero. A typical example is the rainfall data features.

3.6.2 Feature Selection

Given a number of features, wrapper method and Support Vector Machine were used to select subset of features that have the greatest predictive power and still carry their class discriminatory properties. This research features is thus restricted to five (5) predominant malarial symptomatic features: Headache (H_d), Fever (F_v), Dizziness (D_z), Body Pain (B_p), Vomiting (V_m) and two (3) significant climatic factors that contribute to having malaria; Temperature (T_{emp}), Relative humidity (R_h) and Rainfall (R_f).

3.7. Computation of SVM_ANN Qualitative Measures

In this research, the perceptrons (individual nodes) are representative of the neuron. Figure 5 represents the input vector to the network. The eight (8) nodes H_d (Headache), F_v (Fever), D_z (Dizziness), B_p (Body pain), V_m (Vomiting), T_{emp} (Temperature), R_h (Relative humidity) and R_f (Rainfall) serves as input to the SVM_ANN model.

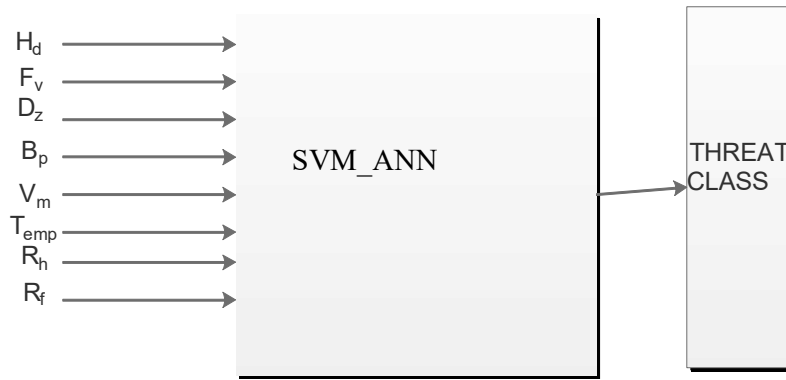


Figure 3.5: SVM_ANN Qualitative Measure

3.8 Multiclass Encoding and Normalization of Dataset Features

In this research features with missing data are assigned zero. A typical example is the rainfall data features. In order to standardize the range of independent variables or data features, feature scaling and unitary method were used and the binary encoding threat classes are represented in Table 3.2 . OPT 1 is the encoded qualitative measure for target Y with class 0 depicting insignificant malaria parasite counts cases, class 1 depicting significant but low malaria parasite counts cases and class 2 depicting significant but high malaria parasite counts cases.

- a. Feature Scaling

$$x' = \frac{x}{||x||} \quad (3.36)$$

- b. Unitary Method/Divide by Maximum

It involves dividing the column or curve by the dataset maximum value.

$$x' = \frac{x}{x_{\max}} \quad (3.37)$$

where x is an original value, and x' is the normalized value.

Table 3.2: Multiclass Encoding Threat Classes

Malaria parasite Count	Malaria parasite Count	Output for Qualitative
Multiclass Output(class)	Binary-class Output	Computation
		(OPT1)
Insignificant (0) { }	Insignificant (0) { }	0
Significant(1) +	Significant(1 and above) $\geq +$	1
Highly Significant(2)++		2

3.9 Hybridized SVM_ANN Malaria Model Descriptions

Climatic factors such as temperature, rainfall and relative humidity affect the lifecycle of malaria parasite *Plasmodium spp.* So malaria incidences depend on these climatic factors and also on antecedent malarial incidences. In this study, a model based on antecedent random malaria cases, antecedent temperature (T_{emp}), antecedent rainfall (R_f) and relative humidity (R_h) is used. The model will be developed to predict malarial incidences:

Here is the model description:

Model 1: $M_1 = f[(Mp), T_{emp}, R_h, R_f]$

From the description above, in the model, the antecedent malaria incidences will be trained with antecedent temperature, antecedent relative humidity and rainfall simultaneously.

This Hybridized SVM_ANN Malaria Model was developed with the use of Support Vector Machine and Multi-Layer Feed-Forward Malaria Neural Network (FFN) to detect and classify symptomatic malaria threats. Eight (8) input features were trained with ten (10) hidden neurons for each iteration and three 3 output features of 1 bits representation as shown in Figure 3.6.

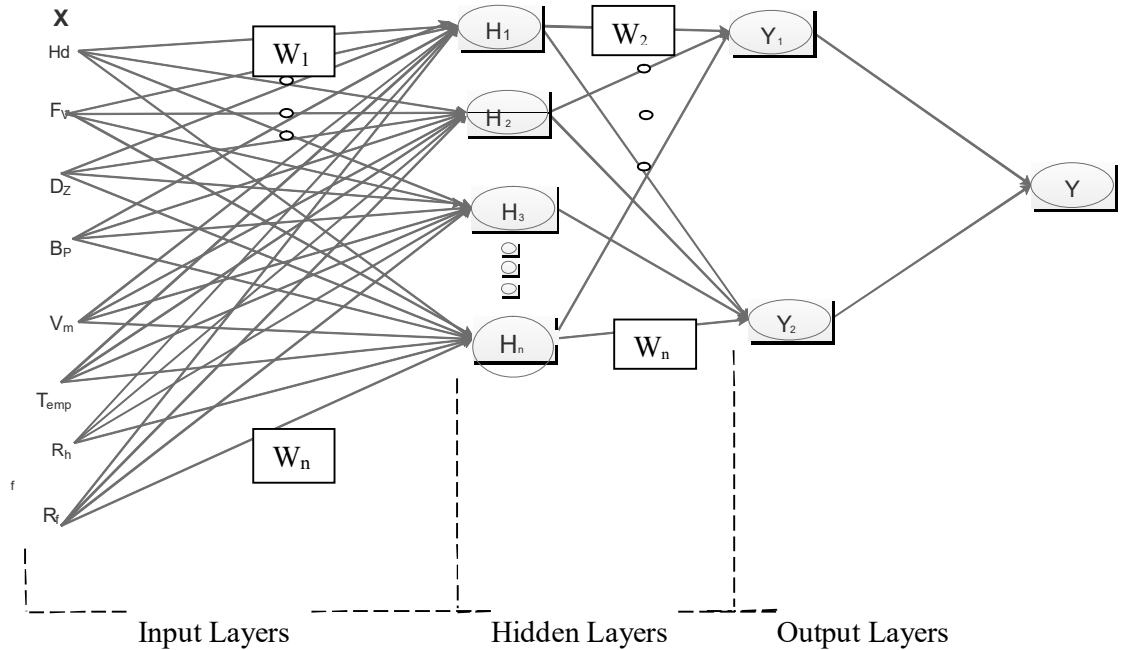


Figure 3.6: A fully connected Schematic Description of Hybridized SVM_ANN Malaria Model

The derivation of the back propagation algorithm (BP_a) which tries to minimize the square of the error (Rumelhart & McClelland, 1986) is shown. The variables for the derivation of back propagation are being defined as follows:

- a. X are the support vectors of the input features $\{H_d(\text{Headache}), F_v(\text{Fever}), D_z(\text{Dizziness}), B_p(\text{Body pain}), V_m(\text{Vomiting}), T_{\text{emp}}(\text{Temperature}), R_h(\text{Relative humidity}) \text{ and } R_f(\text{Rainfall})\}$ from Support Vector machine;
- b. Y is the target of the SMM_ANN network; Y is the output simulated response of Support Vector Machine;
- c. W is the weight of the SMM_ANN network;
- d. E is the error which is the difference between the Target of the network node and the Output
- e. S is the Squashing Function Value;
- f. μ is the Learning Rate.

3.9.1 Error Function

By using the log-sigmoid transfer function and equation (3.13) the Output function of this ANN Multilayer network is

$$S_r(\bar{x}, \bar{w}) = \frac{1}{1 + e^{-\phi_r(\bar{x}, \bar{w})}} \quad (3.38)$$

And the error function for the output of each neuron is thus:

$$E_r(\bar{x}, \bar{w}, T) = (S_r(\bar{x}, \bar{w}) - T_r)^2 \quad (3.39)$$

The square of the difference between the output $S_r(\bar{x}, \bar{w})$ and the desired target

T_r is taken because it will be always positive, and because it will be greater if the difference is much, and lesser if the difference is small.

The error of the network will simply be the sum of the errors of all the neurons in the output layer:

$$E_r(\bar{x}, \bar{w}, \bar{T}) = \sum_r (S_r(\bar{x}, \bar{w}) - T_r)^2 \quad (3.40)$$

3.9.2 Weights from Different Inputs

Before training rands function a symmetric random weight/bias initialization function was used to initialize the weights to random values. Weights were initialized to small random valued to prevent saturation. A network training function 'traingsecg' was also used to update weight and bias values according to scaled conjugate gradient method. The network topology used for this study is 8-10-1 structure. Weights from different inputs H_d (Headache), F_v (Fever), D_z (Dizziness), B_p (Body pain), V_m (Vomiting), T_{emp} (Temperature), R_h (Relative humidity) and R_f (Rainfall) are depicted in Figure 3.5. SOM training shows the visualization of the weights that connect each input to each of the neuron. Fig 3.7 SOM training identified the weight vector associated with each neuron and moves to become the center of cluster of input vector. Darker colours reveal larger weights. Inputs are highly correlated if connections of two inputs were very similar.

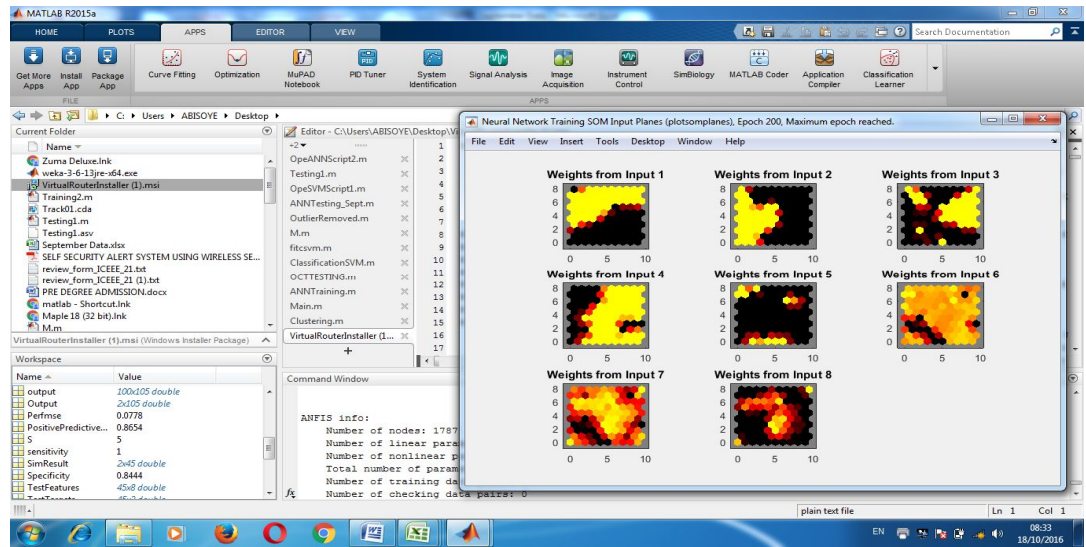


Figure 3.7: Weights from different inputs

It was revealed that the sampled data set at the study time have prevalence of vomiting feature and were great determinant on severity of malaria infection. Rainfall is also the next determinant followed by fever, dizziness and headache. Temperature and relative humidity have lesser effect on the severity of malaria parasite counts.

3.10 Hybridized SVM_ANN Malaria Model Quantitative Measure

Developed Hybridized SVM_ANN Malaria Model is a $8 \times 10 \times 2$ network consisting of eight (8) input nodes, 10 neurons in the layer and two (2) nodes in the output layers. Thus, in each iteration, only 10 hidden neurons were used.

The input layer and the output layer of the Hybridized SVM_ANN Malaria Model as shown in Figure 3.6 consist of Y as

$$Input = \begin{bmatrix} H_d \\ F_v \\ B_p \\ D_z \\ V_m \\ T_{emp} \\ R_h \\ R_f \end{bmatrix} \quad Output = [OP1] \quad (3.42)$$

Thus, if $\{x_i\}$ is a set of support vectors where $\{i = 0, 1, 2, \dots, 8\}$ then $x_1 = H_d$, $x_2 = F_v$, $x_3 = B_p$, $x_4 = D_z$, $x_5 = V_m$, $x_6 = T_{emp}$, $x_7 = R_h$, $x_8 = R_f$. Also, if $\{y_i\}$ is a set of langrange multiplier then $y_1 = OP1$ This implies that the equation (3.41) and (3.42) can be expressed as (3.43) and (3.44)

$$x_i = \begin{bmatrix} H_d \\ F_v \\ B_p \\ D_z \\ V_m \\ T_{emp} \\ R_h \\ R_f \end{bmatrix} \quad y_i = [OP1] \quad (3.43)$$

Let the set of input features x_i and the target features (t_i) of SVM_ANN

hybridized malaria model be k then the total set of input x_k and target (t_k) are

$$x_k = [x_{k1}, x_{k2}, \dots, x_{km}] \quad \text{where } \{k = 1, 2, 3, \dots, m\} \quad (3.45)$$

$$t_k = [t_{k1}, t_{k2}, \dots, t_{kn}] \quad \text{where } \{k = 1, 2, 3, \dots, n\} \quad (3.46)$$

Since, the error function for the output of each neuron in equation (3.39) is :

$$E_r(\bar{x}, \bar{w}, T) = (S_r(\bar{x}, \bar{w}) - T_r)^2$$

The back-propagation algorithm now calculates how the error depends on the output, inputs, and weights by using the *gradient descent method*:

$$\Delta w_{rj} = -\mu \frac{\partial E}{\partial w_{rj}} \quad (3.47)$$

Thus adjustment of each weight denoted by Δw_{rj} is the negative of constant mew (μ) multiplied by derivative of the previous weight dependence on the error of the network. This means that the size of the adjustment depends on mew (μ) and contribution of the weight to the error function. The goal of back-propagation algorithm is to find the derivative of E in respect to w_{rj} .

Firstly there is need to calculate how much error depends on the output which is:

$$\frac{\partial E}{\partial S_r} = 2(S_r - T_r) \quad (3.48)$$

Secondly, how much of the output depends on the activation function which in turn depends on the weight from equation (3.13) and (3.14)

$$\frac{\partial S_r}{\partial w_{rj}} = \frac{\partial S_r}{\partial \varphi_r} \frac{\partial \varphi_r}{\partial w_{rj}} = S_r (1 - S_r) x_j \quad (3.49)$$

Then from (3.48) and (3.49) above

$$\frac{\partial E}{\partial w_{rj}} = \frac{\partial E}{\partial S_r} \frac{\partial S_r}{\partial w_{rj}} = 2(S_r - T_r) S_r (1 - S_r) x_j \quad (3.50)$$

So the adjustment to each weight from gradient descent function in (3.48) is

$$\Delta w_{rj} = -2\mu(S_r - T_r) S_r (1 - S_r) x_j \quad (3.51)$$

Equation (3.52) is for training with two layers.

For training with one more layer, there is need to adjust the y weights say Z_{jk} of previous layer. Firstly, calculation on how the error depends on the input from previous layer rather than on the weight should be conducted.. Change of x_j in equation (3.49), (3.50), (3.51) was done with weight Z_{jk} . Also, how the error of the network depends on the adjustment of Z_{jk} in equations (3.53) and (3.54) was depicted in :

$$\Delta Z_{jk} = -\mu \frac{\partial E}{\partial z_{jk}} = -\mu \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial z_{jk}} \quad (3.52)$$

Where,

$$\frac{\partial E}{\partial w_{rj}} = 2(S_r - T_r)S_r(1 - S_r)w_{rj} \quad (3.53)$$

Consequently, to add another layer to the network, calculation on how the error depends on the weights and inputs of the previous layers was performed..

Practically, implementing back-propagation algorithm in ANN should not have too many layers since the time for training grows exponentially.

Thus, Hybridized SVM_ANN Malaria based mathematical model can be expressed from the equation (3.51) as:

$$\begin{bmatrix} \text{OP1} \end{bmatrix} = \begin{bmatrix} W_{rj} \end{bmatrix} \begin{bmatrix} H_d \\ F_v \\ D_z \\ B_P \\ V_m \\ T_{emp} \\ R_h \\ R_f \end{bmatrix}$$

where $r = 1, 2$

$j = 1, 2, \dots, 8$ (3.54)

W_{rj} are the Hybridized SVM_ANN Malaria model weights matrix. They are constant values between input layer and hidden layer (input weights), between hidden layer and output layer(output weights). Then,

$$[OPI] = \begin{bmatrix} w_{1,1} & w_{1,2} & w_{1,3} & w_{1,4} & w_{1,5} & w_{1,6} & w_{1,7} & w_{1,8} \end{bmatrix} \begin{bmatrix} H_d \\ F_v \\ D_z \\ B_p \\ V_m \\ T_{emp} \\ R_h \\ R_f \end{bmatrix} \quad (3.55)$$

Where:

$$\begin{aligned} \text{Hybridized SVM_ANN Malaria Model matrix} &= [IW_{rj}^T] [LW_{rj}^T] \\ &= \begin{bmatrix} w_{1,1} & w_{1,2} & w_{1,3} & w_{1,4} & w_{1,5} & w_{1,6} & w_{1,7} & w_{1,8} \end{bmatrix} \end{aligned} \quad (3.56)$$

And $[IW_{rj}^T]$ are input weight matrix and $[LW_{rj}^T]$ are output weight matrix

3.11 Adaptive Static Thresholding and Feature Selection Algorithm

The Hybridized SVM_ANN Malaria model was thresholded to get the optimum threshold value of the vector density that will yield best result for the model using the adaptive thresholding algorithm in Table 3.3. At training section of the model, the thresholding ranges start from 0.1 to 1.5 to search for the optimal threshold that gives highest accuracy with low error rate and low standard deviation. At the testing stage, the thresholding algorithm gave the best trade off search of the

threshold optimal parameters of accuracy, sensitivity, specificity, false negative and false positive rates.

Table 3.3 Adaptive Thresholding Algorithm

<i>Input: Training Malaria Datasets</i>	
<i>Output: Thresholds Values(population), Accuracy, Mean and Standard Deviation</i>	
1.	<i>Begin</i>
2.	<i>Supply the Malaria Data Features</i>
3.	<i>Preprocess the data by using the most suited normalization method</i>
4.	<i>Divide the data into Training , Testing and Validation Malaria Datasets in ratio 70:15:15</i>
5.	<i>Perform Data transpose section in preparation for feeding into ANN</i>
6.	While <i>Threshold_Value</i> >= 0.2800 <i>Step</i> 0.005 Do
7.	While <i>Accuracy_Instances</i> <= <i>No_of_Runs</i>
8.	<i>Perform ANN Training</i>
9.	<i>Get Error and Performance Computation</i>
10.	<i>Simulate with Transposed Testing Malaria Datasets</i>
11.	<i>Get Simulation Results thresholding</i>
12.	<i>Get Optimun Threshold_Value</i>
13.	<i>Compute Threat rate and Severity Accuracy</i>
14.	EndWhile
15.	<i>Tabulation of Threshold_Value, Accuracy, Mean and Standard Deviation</i>
16.	EndWhile <i>Threshold_Value</i> = 0.7350
17.	End

After several tuning of the model, a good configuration of step size 0.005 was taken to get the local and global maxima concentrated within the region 0.2800 to 0.7350, output ranges. Several threshold values were searched to get the best threshold values. The optimal threshold value is calculated by selection algorithm in Table 3.4 and the optimal features are used to test the Hybridized SVM_ANN Malaria Model using the Malaria Threat and Severity Algorithm in Table 3.5.

Table 3.4: SVM Feature Selection Algorithm

<i>Input: Training Malaria Datasets</i>	
<i>Output: Optimal Features</i>	
1.	<i>Begin</i>
2.	<i>Supply the Malaria Data Features</i>
3.	<i>Preprocess the data by using the most suited normalization method</i>
4.	<i>Divide the data into Training Malaria Datasets and Testing Malaria Datasets in ratio 70:30</i>
5.	<i>Perform One_Against_All(OAA) algorithm to convert Multiclass to Binary class in preparation for feeding into SVM</i>
6.	WhileThreshold_Value >= 0.2800 Step 0.05 Do
7.	While Accuracy_Instances <= No_of_Runs
8.	<i>Train an SVM</i>
9.	<i>Simulate SVM</i>
10.	<i>Recall Simulated SVM</i>
11.	<i>Simulate with Transposed Testing Malarial Datasets</i>
12.	<i>Get Simulation Results</i>
13.	<i>Compute Optimal Features, Accuracy ,Performance Evaluation</i>
14.	EndWhile
15.	EndWhile
16.	End

3.12 Malaria Threat and Severity Algorithm

The model finally derived good result at threshold 0.600 as shown in Table 3.5.

Table 3.5: Malaria Threat Severity Algorithm

<i>Input: Training Malaria Datasets</i>	
<i>Output: Threat, Severity and Accuracy</i>	
1.	<i>Begin</i>
2.	<i>Supply the Malaria Data Features</i>
3.	<i>Preprocess the data by using the most suited normalization method</i>
4.	<i>Divide the data into Training Malaria Datasets and Testing Malaria Datasets in ratio 70:30</i>
5.	<i>Perform Data transpose section in preparation for feeding into ANN</i>
6.	WhileThreshold_Value = 0.600 Do
	a. While Accuracy_Instances <= No_of_Runs
	i. <i>Recall Simulated network</i>
	ii. <i>Simulate with Transposed Testing Malaria Datasets</i>
	1. <i>Get Simulation Results</i>
	iii. <i>Compute Threat rate, Severity, Accuracy , and Performance Evaluation</i>
7.	EndWhile
8.	EndWhile
9.	End

3.13 Analysis of Performance

3.13.1 ANN Analysis of Performance

There is the need for evaluation of the model is based on its ability to give correct classification (Wu & Banzhaf, 2008). To analyze the performance criterion for the Neural Network classifier for malaria disease severity performance evaluation metrics based on confusion matrix are used. Confusion matrix is a table of classification performance which cross-maps the real nature of expected result against the predicted result from the developed model as shown in Table 3.6.

The matrix consists of four possible outcomes; True Positives, True Negatives, False Positives and False Negatives.

Table 3.6: Confusion matrix

<i>Predicted Class</i>	<i>Actual Class</i>	
	<i>Infected (Positive Class)</i>	<i>Non-Infected Negative Class</i>
<i>Infected (Positive Class)</i>	<i>True Positive(TP)</i>	<i>False Negative(FN)</i>
<i>Non Infected Negative Class</i>	<i>False Positive(FP)</i>	<i>True Negative(TN)</i>

The criteria to measure the performance based on confusion matrix are given in equations (3.57- 3.62) as follows:

- a. ***Sensitivity (True Positive Rate):*** *Sensitivity* is the ability of test to identify the infected cases.

$$\begin{aligned}
\text{Sensitivity(Recall)} &= \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} * 100 \\
&= \frac{TP}{TP + FN} * 100
\end{aligned} \tag{3.57}$$

- b. **Specificity (True Negative Rate):** *Specificity* is the ability to separate the infected cases from the non-infected cases

$$\begin{aligned}
\text{Specificity} &= \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} * 100 \\
&= \frac{TN}{TN + FP} * 100
\end{aligned} \tag{3.58}$$

c. False Positive Rate(FPR): $= \frac{FP}{TN + FP} * 100 = 1 - \text{Specificity}$ (3.59)

d. False Negative Rate(FNR): $= \frac{FN}{TP + FN} * 100 = 1 - \text{Sensitivity}$ (3.60)

- e. **Accuracy:** Accuracy is the percentage of correct predictions that are correct.

$$\text{Accuracy} = \frac{\text{Correct Classified Patterns}}{\text{Total Patterns}} = \frac{TP + TN}{TP + TN + FP + FN} * 100 \tag{3.61}$$

- f. **Mean Square Error:** Mean squared error (MSE) is used for obtaining efficient estimators, a widely used class of estimators.

$$MSE = \frac{1}{2} \sum_{i=1}^n (S_r(\bar{x}, \bar{w}) - T_r)^2 \tag{3.62}$$

3.13.2 SVM Analysis of Performance

To analyze the performance criterion for the Support Vector classifier for malaria disease, the support vector parameters of number of Support Vectors, Langrange Multipliers, Regularization Constant measures, mean square error and accuracy were used. Where

a. **Support Vectors:** Support Vectors are closest data points to the hyperplane. These points exert the maximum force on the hyperplane. A large support vectors indicate good model.

b. **Langrangian Multipliers (Alpha):** Langrange multipliers are a mathematical method used to solve constrained optimization problems of differentiable functions. It is denoted by α_i the vector of nonnegative Lagrange multipliers associated with the optimization problem constraints $y_i[(w, x_i)] + b \geq 1$, $i=1, \dots, n$

c. **Regularization Constant (C):** is the soft margin cost function of classification or penalty factor. A large C indicates low bias and high variance. Low bias because you penalize the cost of misclassification high “**hard margin**”. A small C indicates higher bias and lower variance and makes the cost of misclassification low”soft margin” and the decision surface smooth.

d. **Gamma:** Gamma is the parameter of a Gaussian Kernel to handle non-linear classification. Large gamma indicates a small variance implying the

support vector does not have widespread influence. Small gamma indicates a Gaussian with large variance implying support vectors having widespread influence.

g. **Accuracy:** Accuracy is the percentage of correct predictions that are correct.

$$\begin{aligned} \text{Accuracy} &= \frac{\text{Correct Classified Patterns}}{\text{Total Patterns}} \\ &= \frac{TP + TN}{TP + TN + FP + FN} * 100 \end{aligned} \quad (3.63)$$

h. **Mean Square Error:** Mean squared error (MSE) is used for obtaining efficient estimators, a widely used class of estimators.

$$MSE = \frac{1}{2} \sum_{i=1}^n (S_r(\bar{x}_i, \bar{w}) - T_r)^2 \quad (3.64)$$

3.14 Chapter Summary

This chapter covers the symptomatic analysis, likelihood quantitative model, simulation, optimization of Hybridized SVM_ANN Malaria and performance evaluation metrics. Two dimensional malaria threat models were hybridized to validate the security of the models. The developed Hybridized SVM_ANN Malaria Model in this study predicts multiclass cases of malaria incidence. At training section, the model was thresholded with adaptive thresholding algorithm ranging from 0.1 to 1.5 to get the optimum threshold value that gives the best accuracy, low standard deviation and low error rate. After several testing, the model optimum threshold values lie within 0.2800 and 0.7350 with a step size of 0.005. This optimum threshold values were then used to get the local and global maxima concentrated within the region.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

Vector threat and transmission threat are the major threat of malaria infections. Factors that contributes to having malaria such as climatic factors; temperature, rainfall and relative humidity were all collected. These factors combined with the symptomatic clinical features served as input features to the SVM and thus generated optimal features for Hybridized SVM_ANN model.

4.2 Data Analysis

4.2.1 Raw Malaria Data

Monthly malarial incidences data cases were obtained from five different hospitals in Niger State. Climatic data were also collected from Nigerian Environmental and Climate Observation Programme (NECOP) Weather Station, Bosso Campus, Federal University of Technology, Minna, Niger State. Each patient has a set of symptoms and MP count known as patients' malaria data symptoms and lab test results. Monthly 1,200 malarial incidences cases between January, 2012 and December 2015 were collected. Table 4.1 is the sample of raw malaria datasets collected.

Table 4.1: Sampled Raw Malaria Data

Age	Sex	Headache	Fever	BodyPain	Dizziness	Vomt	Temp	Rel Humid	Rainfall	MP_Count
14	2	1	1	0	0	0	0	0	0	1
9	2	1	0	1	1	0	0	0	0	1
25	1	0	1	0	0	0	0	0	0	0
63	1	0	0	1	0	0	0	0	0	1
40	1	1	1	0	1	0	0	0	0	1
20	2	1	0	0	1	0	28.46	62.8	0	1
28	1	0	0	0	1	0	28.46	62.8	0	0
34	1	0	0	0	1	1	28.46	62.8	0	2
25	2	0	1	0	0	0	28.46	62.8	0	1
19	2	1	0	0	0	1	28.46	62.8	0	2
49	2	0	1	0	1	1	28.02	37.27	0	2
60	1	0	0	0	1	1	28.02	37.27	0	2
32	1	0	1	0	1	0	28.02	37.27	0	1
24	2	0	1	0	1	1	28.02	37.27	0	2

4.2.2 Preprocessed Malaria Data

The data was preprocessed before training to remove the outliers, irrelevant features and incorrectly classified samples. One thousand, two hundred (1,200) malaria datasets were sampled and used to develop the model. This sample was further divided into training, testing and validation datasets. Table 4.2 shows the sample of the preprocessed malaria data.

Table 4.2: Sampled Preprocessed Malaria Data

Headache	Fever	Body-Pain	Dizziness	Vomiting	Temperature	Rel- Humid	Rainfall	OP1	OP2
0	0	0	1	0	0.769114	1	1	0	0
0	1	0	1	0	0.769114	1	1	0	0
0	1	0	0	0	0	0	0	0	0
0	0	0	1	1	0.798051	0.941963	0.176471	1	0
0	0	1	0	1	0.798051	0.941963	0.176471	1	0
0	1	0	1	1	0.798051	0.941963	0.176471	1	0
0	0	1	0	1	0.798051	0.941963	0.176471	1	0
1	1	1	0	1	0.772464	0.988894	0.941176	1	0
1	0	0	0	1	0.772464	0.988894	0.941176	1	0
1	1	0	0	0	0	0	0	0	1
1	0	1	1	0	0	0	0	0	1

0	0	1	0	0	0	0	0	0	1
1	1	0	1	0	0	0	0	0	1
1	0	0	1	0	0.86689	0.74994	0	0	1
0	1	0	0	0	0.86689	0.74994	0	0	1

4.2.3 Multiclass Encoded and Normalized Features Dataset

In order to standardize the range of independent variables, data feature scaling and unitary method were used as stated in section 3.4. The dataset was normalized by using scaling for the malaria symptoms and divide by maximum for the meteorological data. The multiclass encoded features and normalized features are shown in Table 4.2.

4.3 ANN Processing Results and Models

The Training, Testing and Validation Features are divided in ratio 70:15:15 according to stratified sampling method. The Training, Testing and Validation Features are Malaria Symptoms; Headache, Fever, Dizziness, Body-pain, Vomiting and Climatic factors; Temperature, Relative Humidity and Rainfall while the Target Features are MP_Count. Having 1,200 data instances, 840 are for Training, 180 for Testing and 180 for Validation. Maximum Normalization method normalizes each feature of the sample.

An adaptive thresholding algorithm was developed to search for global maximum threshold value with higher accuracy and low standard deviation as shown in *Appendix A: Table 4.3*. A step size of 0.005 was taken to get the local and global maxima concentrated within the region 0.2800 to 0.7350, output ranges. Several threshold values were searched to get the best threshold values. The learning rate of 0.9 and momentum rate 0.1 gave the optimum accuracy and least error.

After several iterations of the model with the malaria data, certain threshold values at training stages tagged “Threshold Values and Performance for the developed ANN model on Training Dataset were obtained as shown in *Table 4.3 in Appendix A* and depicted in *Figure 4.1 and Figure 4.2*. Thus the Optimum Threshold Values for the developed Neural Network Model on Training Dataset is between 0.475 and 0.580 as shown *Table 4.4*.

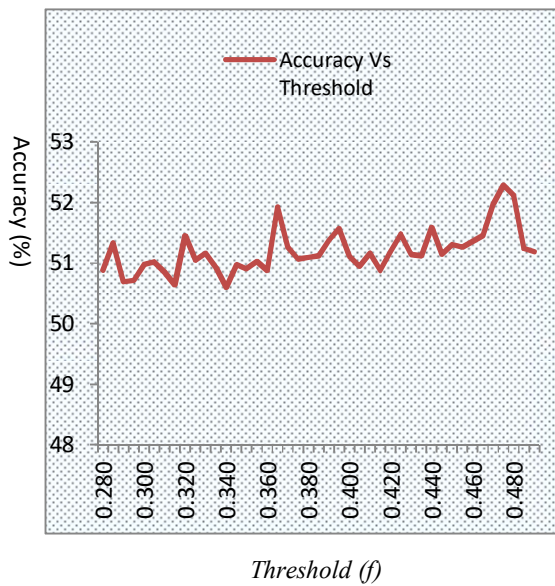


Figure 4.1: Threshold Values and Performance Training (1)

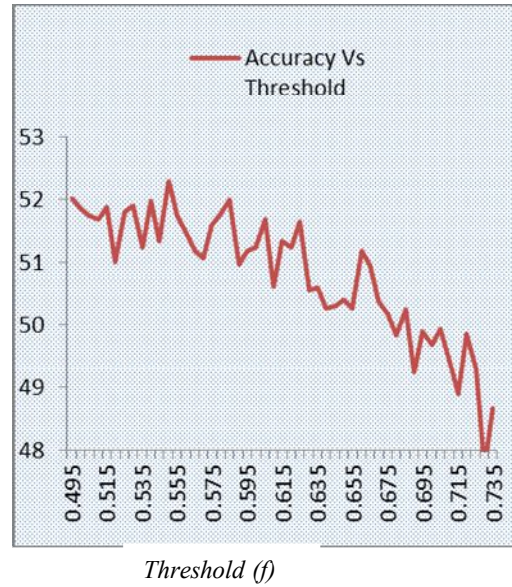


Figure 4.2: Threshold Values and Training (2)

Table 4.4: Best Threshold Values and Performance for the ANN Training Model

Threshold	Accuracy 1	Accuracy 2	Accuracy 3	Accuracy 4	Accuracy 5	Average Accuracy	Standard Deviation
0.4750	50.4762	52.3810	53.2143	52.2619	53.0952	52.2857	1.0956
0.4800	54.4048	53.3333	50.4762	50.8333	51.5476	52.1190	1.6865
0.4950	50.9524	50.9524	51.6667	53.9286	52.6190	52.0238	1.2655
0.5500	52.5000	52.3810	50.0000	53.8095	52.8571	52.3095	1.4076
0.5850	50.7143	51.5476	51.7857	54.0476	51.9048	52.0000	1.2355

In the training section, after five (5) runs of several iterations at different threshold, the optimum accuracy of 52.31 at threshold 0.55 with standard deviation 1.4076 was

obtained as shown in *Table 4.4* and Figure 4.3.

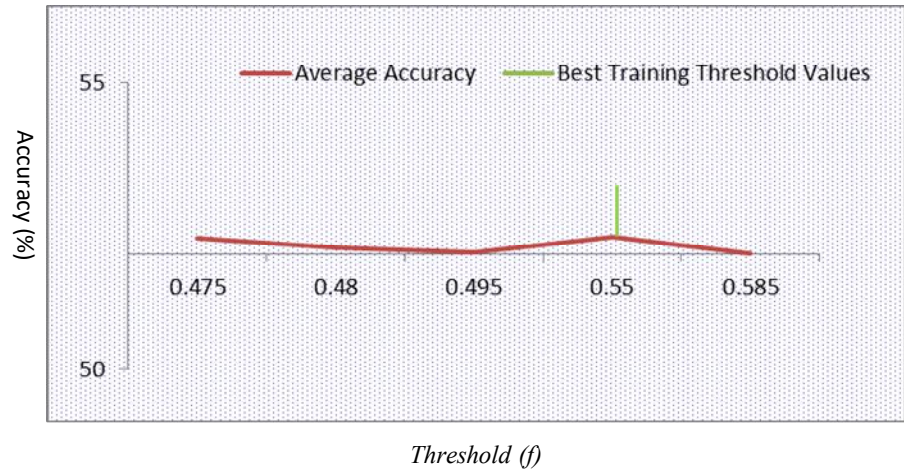


Figure 4.3: Best Threshold Values and Performance for the ANN Malaria Training Model

From *Table 4.4* and Figure 4.3, the best accuracy for ANN Malaria Training model is between 0.475 and 0.5850. Also, Threshold Values and Performance for the ANN Malaria model on Testing Dataset is shown in *Table 4.5 Appendix A*.

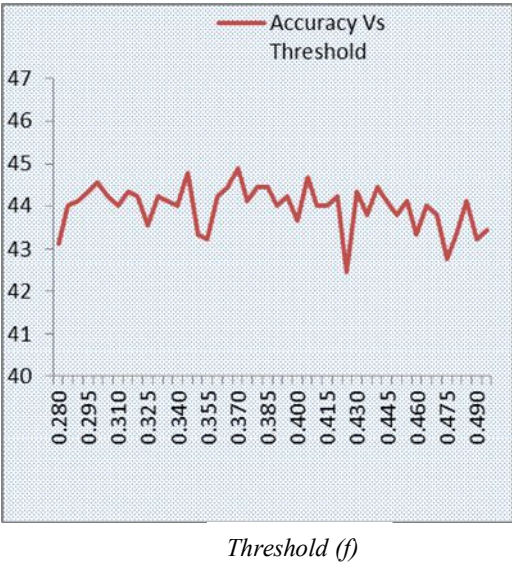


Figure 4.4: Threshold Values and Performance for the ANN Malaria Testing Model (1)

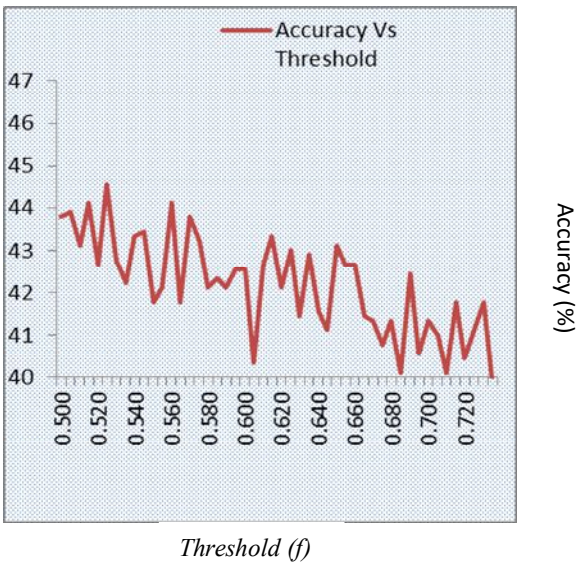


Figure 4.5: Threshold Values and Performance for the ANN Malaria Testing Model (2)

Table 4.6: Best Threshold Values and Performance for the Developed ANN Malaria Testing Model

Threshold	Accuracy 1	Accuracy 2	Accuracy 3	Accuracy 4	Accuracy 5	Average Accuracy	Standard Deviation
0.3000	42.7778	45.0000	45.0000	45.0000	45.0000	44.5556	0.9938
0.3450	43.8889	45.5556	44.4444	45.0000	45.0000	44.7778	0.6334
0.3700	45.0000	45.0000	45.0000	45.0000	44.4444	44.8889	0.2485
0.4050	45.0000	45.0000	45.0000	43.8889	44.4444	44.6667	0.4969
0.5250	45.0000	43.3333	45.5556	44.4444	44.4444	44.5556	0.8240

Table 4.6 depicts *Best Threshold Values and Performance for ANN model on Testing Dataset* which is an excerpt of Table 4.5 in Appendix A. The best accuracy of 44.889 at threshold 0.3700 with standard deviation 0.2485 obtained is shown in Figure 4.6.

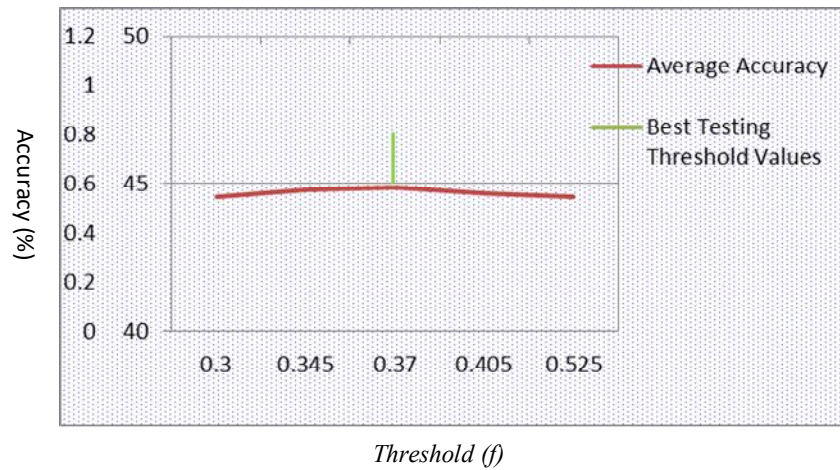


Figure 4.6: Best Threshold Values and Performance for the developed ANN Malaria Testing Model

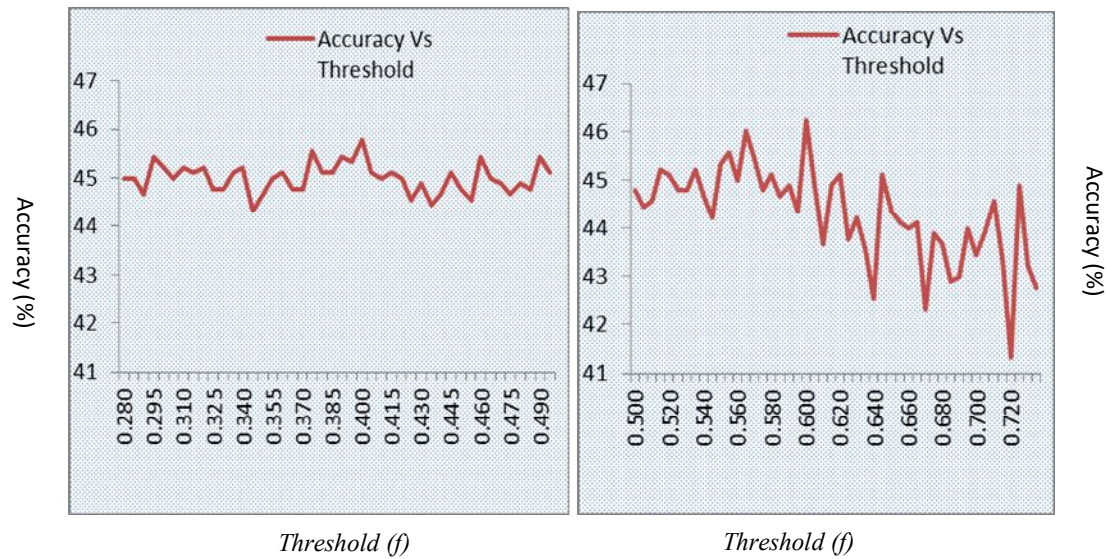


Figure 4.7: Threshold Values and Performance for the ANN Malaria Validation Model (1) Figure 4.8: Threshold Values and Performance for the ANN Malaria Validation Model (2)

Table 4.8, Figure 4.7 and Figure 4.8 depicts Threshold Values and Performance for the developed ANN Malaria Model on Validation Data set which was an excerpt of Table 4.7 in Appendix A. To obtain the best accuracy of the model, a benchmark of Standard Deviation of not greater than 1.50 should not be exceeded as shown for best runs in Table 4.8. This was confirmed in the overall best threshold value and performance depicted in Table 4.9 and Figure 4.9.

Table 4.8: Optimum Threshold Values and Performance for the developed ANN Malaria Validation Model

Threshold	Accuracy 1	Accuracy 2	Accuracy 3	Accuracy 4	Accuracy 5	Average Accuracy	Standard Deviation
0.3750	47.2222	44.4444	45.0000	45.5556	45.5556	45.5556	1.0393
0.4000	45.0000	46.1111	46.6667	46.1111	45.0000	45.7778	0.7454
0.5550	46.6667	47.2222	43.3333	43.8889	46.6667	45.5556	1.8002
0.5650	48.8889	43.8889	44.4444	45.0000	47.7778	46.0000	2.2013
0.6000	48.3333	46.1111	45.0000	45.0000	46.6667	46.2222	1.3833

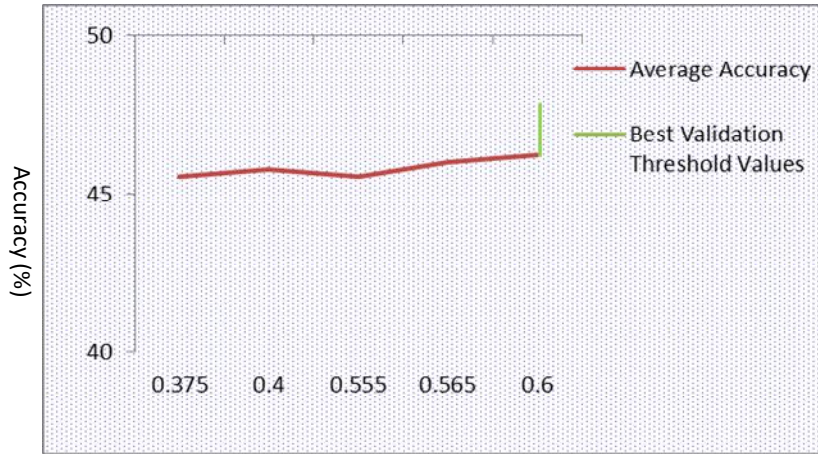


Figure 4.9: The Threshold Values and Performance for ANN Malaria Validation Model

Table 4.9 Best Threshold Values and Performance for ANN Malaria Model on Validation Phase

Threshold	Accuracy 1	Accuracy 2	Accuracy 3	Accuracy 4	Accuracy 5	Avg Acc	Std Dev
0.6000	48.3333	46.1111	45.0000	45.0000	46.6667	46.2222	1.3833

Finally, In Validation stage, the overall best accuracy of 46.222 at threshold 0.6000 with standard deviation 1.3833 was obtained as shown in Figure 4.9.

4.3.1 ANN Target and Simulated Malaria Test Result

Table 4.11 is an excerpt of Table 4.10 in *Appendix B* which shows the Target and Simulated Thresholded Test Results. However, when each Class ANN_0, ANN_1 and ANN_2 was subjected to test and validated, accuracies 48.33%, 47.22% and 46.11% respectively were obtained within the threshold 0.2800 and 0.7350 as shown in Table 4.11.

4.3.2 ANN Performance Evaluation

Confusion Matrixes showing ANN_0, ANN_1 and ANN_2 in Figures 4.10 - 4.12 were derived from Table 4.11.

Table 4.11: ANN Classifier Results

Performance Metrics	Accuracy (%)	True Positive (TP)	True Negative(TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FPR)	False Negative Rate FN _R	Mean Square Error (MSE)	Total Positive	Total Negative	Total
ANN_0	48.33	20	67	80	13	0.6061	0.4558	0.5442	0.3939	0.5550	33	147	180
ANN_1	47.22	68	17	85	10	0.8718	0.1667	0.8333	0.1282	0.5611	78	102	180
ANN_2	46.11	0	83	28	69	0	0.7477	0.2523	1	0.5556	69	111	180

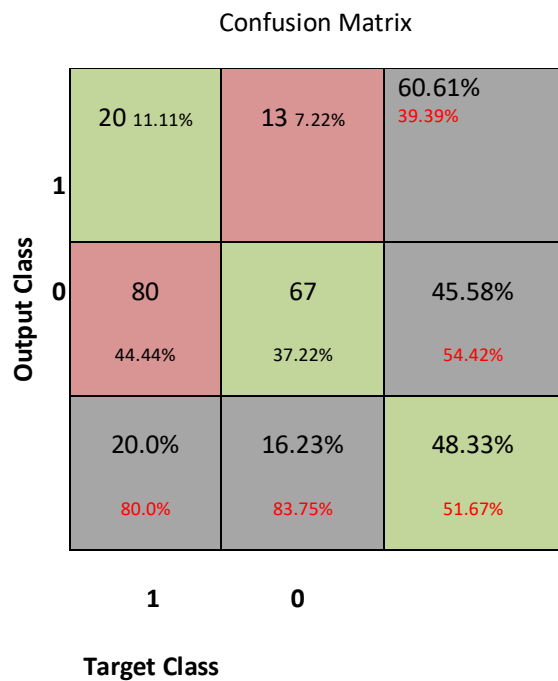


Figure 4.10 ANN_0 Results

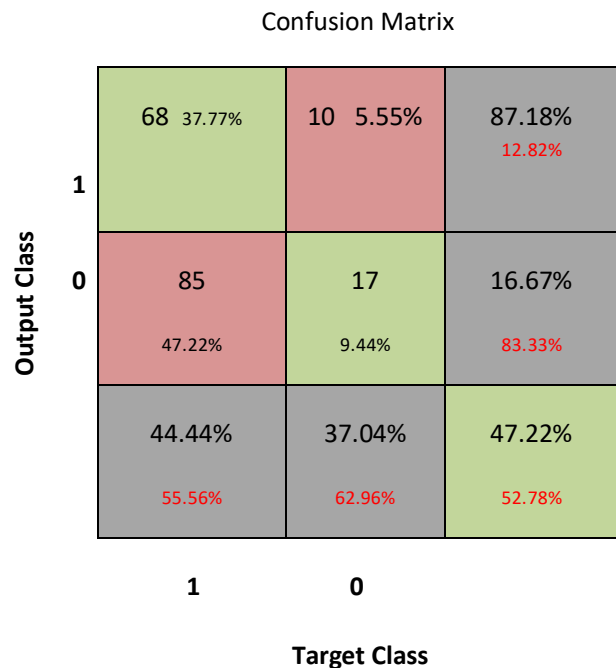


Figure 4.11 ANN_1 Results

Output Class	1	0	
	0		
	1		
	1	0	

In the confusion matrix ANN_0 result, twenty (20) cases of Class_0 infected were correctly classified as positive. This corresponds to 11.11% of all one hundred and eighty (180) malaria cases. Similarly, sixty seven (67) cases of Class_1 and Class_2 non infected were correctly classified as negative. This corresponds to 37.22% of all malaria cases. Also, thirteen (13) cases of Class_1 and Class_2 non infected cases which correspond to 7.22% were incorrectly classified as negative. Similarly, eighty (80) cases of Class_0 infected cases were incorrectly classified as positive. This corresponds to 44.44% of all malaria cases.

144

correspond to 39.39% were wrongly classified. Similarly, out of one hundred and forty seven (147) non infected cases, only sixty-seven (67) cases which corresponds to 45.58% were correctly classified as non-infected cases while eighty (80) which corresponds to 54.42% were incorrectly classified.

In the confusion matrix ANN_1 result, sixty-eight (68) cases of Class_1 infected were correctly classified as positive. This corresponds to 37.77% of all one hundred and eighty (180) malaria cases. Similarly, seventeen (17) cases of Class_0 and Class_2 non infected cases were correctly classified as negative. This corresponds to 9.44% of all malaria cases. Also, ten (10) cases of Class_0 and Class_2 non infected cases which correspond to 5.55% were incorrectly classified as negative. Similarly, eighty-five (85) cases of Class_1 infected cases were incorrectly classified as positive. This corresponds to 47.22% of all malaria cases.

Out of seventy-eight (78) infected cases, sixty-eight (68) were correctly classified. This corresponds to 87.18% correctly classified while ten (10) cases which correspond to 12.82% were wrongly classified. Similarly, out of one hundred and two (102) non infected cases only seventeen (17) cases which correspond to 16.67% were correctly classified as non-infected cases while eighty-five (85) which corresponds to 83.33% were incorrectly classified.

In the confusion matrix ANN_2 result, zero (0) cases of Class_2 infected cases were correctly classified as positive. Eighty-three (83) cases of Class_0 and Class_1 non infected cases were correctly classified as negative. This corresponds to 46.11% of all malaria cases. Also, sixty-nine (69) cases of Class_0 and Class_1

non infected cases which correspond to 39.33% were incorrectly classified as negative. Similarly, twenty-eight (28) cases of Class_2 infected cases were incorrectly classified as positive. This corresponds to 15.56% of all malaria cases.

Out of sixty-nine (69) infected cases of Class_2, zero (0) were correctly classified, all sixty-nine (69) infected cases were wrongly classified. Similarly, out of one hundred and eleven (111) non infected cases Class_0, Class_1 eighty-three (83) cases which correspond to 74.22% were correctly classified as non-infected cases while twenty-eight (28) which corresponds to 25.23% were incorrectly classified.

Conclusively, regardless of different activation function used; logsig, tansig and purelin used in ANN Malaria Model, it results to low performance. The results show that ANN is prone to local error and over-fitting due to large number of parameters to fit. So, a quest for a feature selection method: Support Vector Machine that employs a unique kernel function to map the input cases to target output correctly is needed.

4.4 SVM Processing Results and Models

The Training, Testing and Validation Features are divided in ratio 70:15:15 according to stratified sampling method. The Training, Testing and Validation Features are Malaria Symptoms; Headache, Fever, Dizziness, Body-pain, Vomiting and Climatic factors; Temperature, Relative Humidity and Rainfall while the Target Features are MP_Count.

4.4.1 SVM Feature Selection Results

The Multiclass malaria data was handled by one-against-all algorithm. The result of various classes of SVM Feature Selection with 1200 malaria cases; 840:180:180 were used for Training, Testing and Validation are presented in *Tables 4.12 – 4.14* respectively. Also the Graph of the Support vectors Vs. Accuracy and Accuracy Vs. mse for SVM_0, SVM_1, SVM_2 are depicted in *Figures 4.13 – 4.14*, *Figures 4.18 – 4.19* and *Figures 4.23 – 4.24* respectively. Likewise the confusion matrixes for the classes are depicted in *Figures 4.15 - 4.17*, *Figures 4.20 – 4.22*, *Figures 4.25 – 4.27*.

Table 4.12: SVM_0 Feature Selection Results

Performance Metrics	Accuracy (%)	Support Vectors (double)	Lang-range Multiplier	True Positive (TP)	True Negative(TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FP _R)	False Negative Rate FN _R	Mean Square Error (MSE)	Total Positive	Total Negative	Total
SVM_0 (lin)	81.67	232 x 8	232 x 1	0	147	0	33	0	1	0	1	0.7333	33	147	180
SVM_0 (rbf)	86.11	324x8	324 x 1	11	144	3	22	0.3333	0.9797	0.0204	0.6667	0.5556	33	147	180
SVM_0 (pol)	86.11	208 x 8	208 x 1	15	140	7	18	0.4545	0.9524	0.0400	0.5455	0.5556	33	147	180

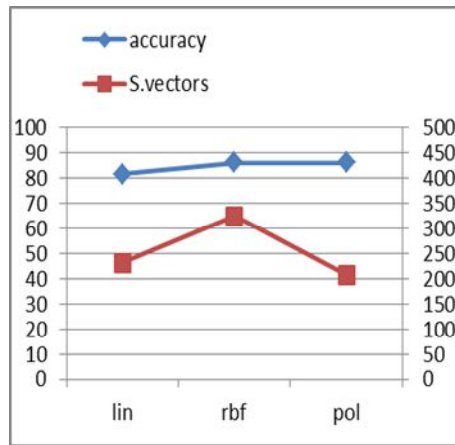


Figure 4.13 Graph of Accuracy Vs Support Vectors for ‘SVM_0’ Malaria Cases

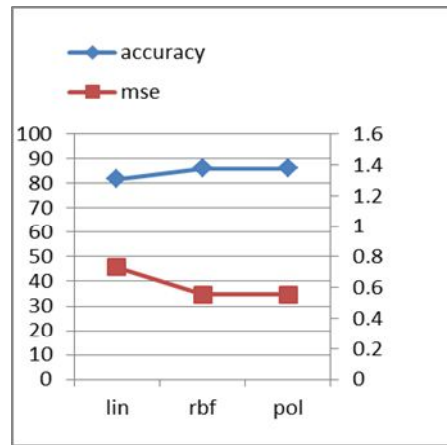


Figure 4.14 Graph of Accuracy Vs Mean Square Error for ‘SVM_0’ Malaria Cases

Confusion Matrix

Output Class	1	0	
	0	1	
	1	0	
	1	0	Target Class

0 (0%)	33 (18.33%)	0%
0 (0%)	147 (81.67%)	100%
0%	81.67%	18.33%

Figure 4.15 : SVM_0 (lin)

Confusion Matrix

Output Class	1	0	
	0	1	
	1	0	
	1	0	Target Class

11 (6.11%)	22 (12.22%)	33.33%
3 (01.67%)	144 (80%)	97.97%
78.57%	86.75%	86.11%

Figure 4.16 : SVM_0 (rbf)

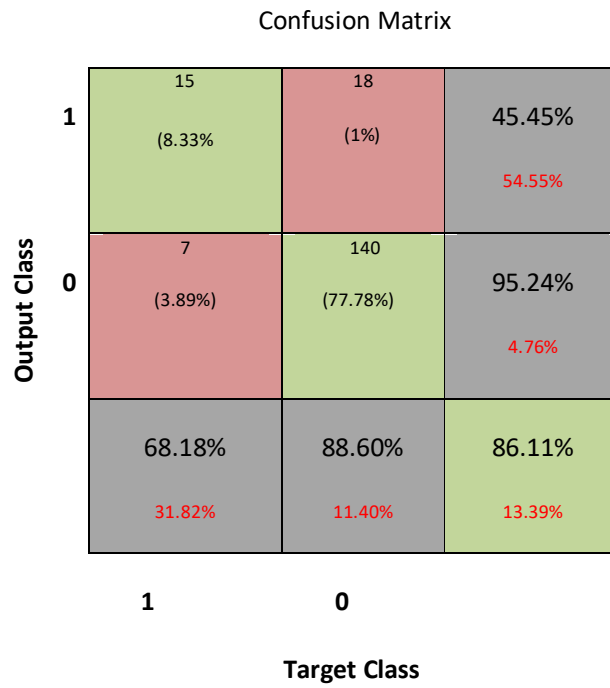


Figure 4.17 : SVM_0 (pol)

Table 4.13: SVM_1 Feature Selection Results

Performance metrics	Accuracy (%)	Support Vectors (double)	Lang-range Multiplier	True Positive (TP)	True Negative(TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FP _R)	False Negative Rate FN _R	Mean Square Error (MSE)	Total Positive	Total Negative	Total
SVM_1 (lin)	66.67	378 x 8	378 x 1	63	57	45	15	0.8077	0.5588	0.4412	0.1923	1.3333	78	102	180
SVM_1 (rbf)	80.55	435 x8	435 x 1	60	85	17	18	0.7692	0.8333	0.1667	0.2308	0.5556	78	102	180
SVM_1 (pol)	83.89	282 x 8	282x 1	63	88	15	14	0.8077	0.8627	0.1311	0.1311	0.6444	78	102	180

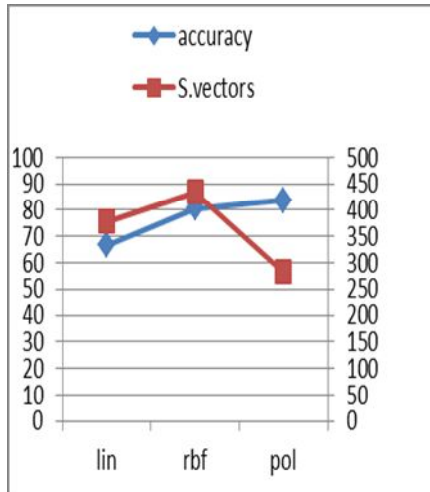


Figure 4.18 Graph of Accuracy Vs Support Square Vectors for 'SVM_1' Malaria Cases

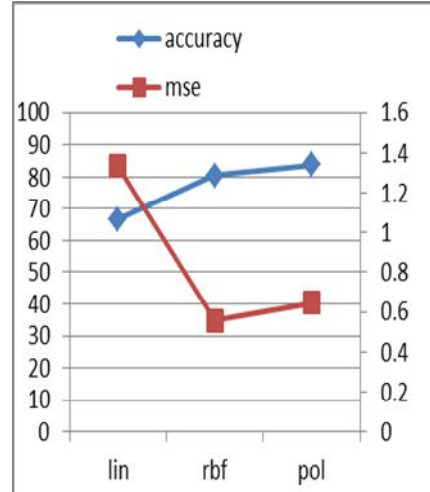


Figure 4.19 Graph of Accuracy Vs Mean Square Error for 'SVM_1' Malaria Cases

Output Class	1	63 (35%)	15 (83.33%)	80.77% 19.23%
	0	45 (25%)	57 (31.67%)	55.88% 44.12%
		58.33% 41.67%	79.17% 20.83%	66.67% 33.33%
	1		0	
		Target Class		

Figure 4.20 : SVM_1 (lin)

Confusion Matrix

Output Class	1	60 (33.33%)	18 (10%)	76.92% 23.08%
	0	17 (9.44%)	85 (47.22%)	83.33% 16.67%
		68.97% 31.03%	82.52% 17.48%	80.55% 19.45%
	1	0		
		Target Class		

Figure 4.21 : SVM_1 (rbf)

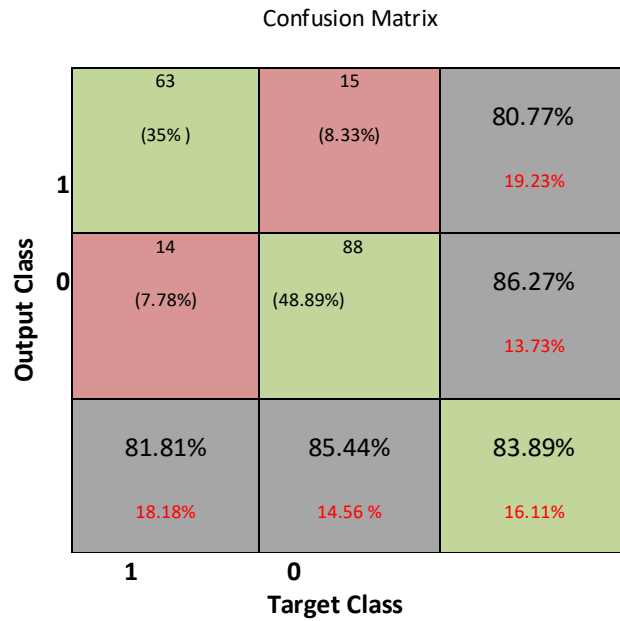


Table 4.14: SVM_2 Feature Selection Results

Performance Metrics	Accuracy (%)	Support Vectors (double)	Lang-range Multiplier	True Positive (TP)	True Negative(TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FP _R)	False Negative Rate FN _R	Mean Square Error (MSE)	Total Positive	Total Negative	Total
SVM_2 (lin)	79.44	195 x 8	195 x 1	52	9	20	17	0.7536	0.8198	0.1802	0.2464	0.8222	69	111	180
SVM_2 (rbf)	85.60	308 x8	308 x 1	58	96	15	11	0.8406	0.8649	0.1351	0.1594	0.5778	69	111	180
SVM_2 (pol)	88.89	147x 8	147 x 1	63	97	14	06	0.9130	0.8736	0.1262	0.0870	0.444	69	111	180

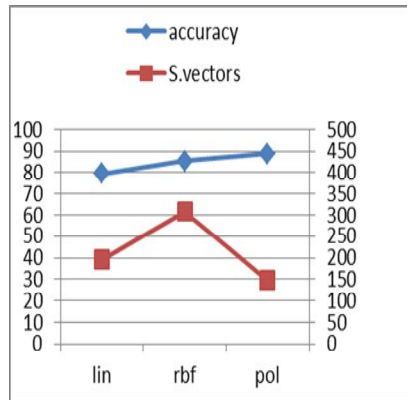


Figure 4.23 Graph of Accuracy Vs Support Square Vectors for 'SVM_2' Malaria Cases

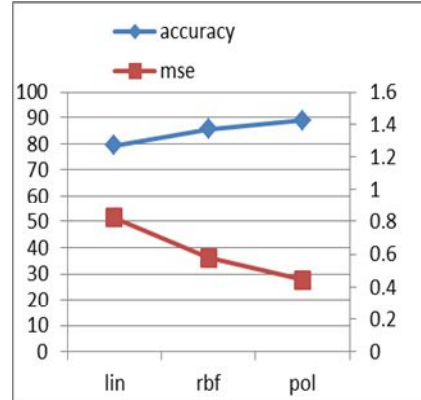


Figure 4.24 Graph of Accuracy Vs Mean Square Error for 'SVM_2' Malaria Cases

Confusion Matrix

Output Class	1	52 (28.89%)	17 (9.44%)	75.36% 24.64%
	0	20 (11.11%)	91 (50.56%)	81.98% 68.97%
		72.22% 27.28%	84.26% 65.38%	79.44% 20.56%
	1	0		
		Target Class		

Figure 4.25: SVM_2 (lin)

Confusion Matrix

Output Class	1	58 (32.22%)	11 (6.11%)	84.06% 15.94%
	0	15 (8.33%)	96 (53.33%)	86.49% 13.51%
		79.45% 20.55%	89.72% 10.28%	85.60% 14.40%
	1	0		
		Target Class		

Figure 4.26 : SVM_2 (rbf)

Confusion Matrix			
Output Class	1	06 (3.33%)	91.30% 8.70%
	0	14 (7.78%)	87.36% 12.64%
		81.81% 18.19%	94.17% 5.83%
		1	0
		Target Class	

Figure 4.27 : SVM_2 (pol)

4.4.2 SVM Target and Simulated Test Result

Tables 4.12 – 4.14 were derived from Appendix Table 4.15 tagged SVM target and Simulated Test Result. However, when each Class SVM_0, SVM_1 and SVM_2 with Linear, Radial basis function and polynomial kernel function was subjected to test and validated, different accuracies within the threshold 0.2800 and 0.7350 were obtained as shown in Tables 4.12 – 4.14.

4.4.3 SVM Performance Evaluation

- a. **Support Vectors:** Support Vectors are closest data points to the hyperplane.

It was observed from Tables 4.12 – 4.14 that the developed SVM model of Class 0, Class 1 and Class 2 with radial basis function has highest support vectors compared to other kernel function. Also SVM model Class 1 with radial basis function has the highest number of support vectors [435x8] in the model.

- b. **Langrangian Multipliers (Alpha):** Alpha are the nonnegative Lagrange multipliers associated with the malaria optimization problem constraints $y_i[(w, x_i)] + b \geq 1, \quad i=1, \dots, n..$ It was observed from Tables 4.12 – 4.14 that the developed SVM model of Class 0, Class 1 and Class 2 with radial basis function has high Langrangian Multipliers compared to other kernel function. Also SVM model Class 1 with radial basis function has the highest number of Langrangian Multipliers [435x8] in the model.
- c. **Regularization Constant (C):** This is the soft margin cost function of classification or penalty factor. A large C indicates low bias and high variance; low bias because the penalty for the cost of misclassification is high; “**hard margin**”. A small C indicates higher bias and lower variance and this makes the cost of misclassification low; “**soft margin**” because the decision surface is smooth. The cost function of the SVM model for Class 0, Class 1 and Class 2 irrespective of the kernel function used is 2×2 double vector.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- d. **MSE:** It was observed in Tables 4.12 – 4.14 that radial and polynomial kernel function gave the lowest mean square error compared with linear kernel function. In the model, polynomial function has highest bias of -3.7651 compared with radial basis function with the lowest bias -0.0817.

- e. **Accuracy:** It was observed in Table 4.12 that polynomial kernel function has the highest accuracy of 88.89% but lowest support vectors, high bias and lowest alpha. Thus, the accuracy of the SVM model as shown in Figures 4.15 – 4.17, Figures 4.20 – 4.22 and Figures 4.25 – 4.27 was derived from Class_2 with radial basis function of 85.60% accuracy and 308x8 double support vectors. This result indicates that performance analysis should not depend on accuracy alone but on other criteria as well. Thus, Class 2 with radial basis function is the best chosen for the model.

f. False Positive Rate and False Negative Rate

Tables 4.12 – 4.14 shows the predicted result obtained and the target values for linear, radial basis and polynomial function. It was observed that:

- i. With linear kernel function, out of 69 positive cases, 52 were identified positive while 17 were false negative. Thus, class 2 linear function gave TP_R of 75.36%, TN_R of 81.89%, FP_R of 0.1802 and FN_R of 0.2464.
- ii. With radial basis function, out of 69 positive cases, 58 were identified positive while 11 were false negative. Thus, class 2 radial basis function gave TP_R of 84.06% TN_R of 0.8649 FP_R of 0.1351 and FN_R of 0.1594.
- iii. With polynomial kernel function out of 69 positive cases, 63 were identified positive while 6 were false negative. Thus, class 2 polynomial kernel function gave TP_R of 91.30% TNR of 87.36%, FP_R of 0.1262 and FN_R of 0.0872.

4.5 SVM_ANN Adaptive Static Thresholding

Initially, ANN and SVM models were trained single handedly but they generated low performance as shown in *Figure 4.28*. Artificial Neural Network_Class 0 (ANN_0) with feed forward and back-propagation algorithm was employed and the model produced optimal 48.33% accuracy, 60.61% Sensitivity and 45.58% Specificity. Also Support Vector Machine Class_2 (SVM_2) generates optimal 85.60 % accuracy, 84.06% Sensitivity and 86.49% Specificity.

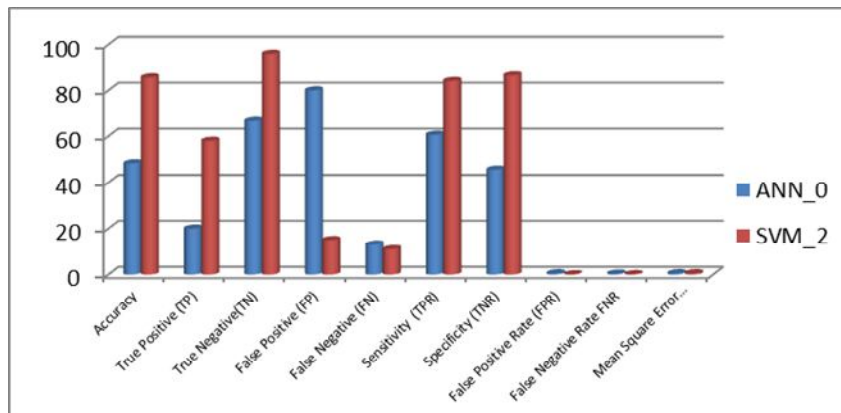


Figure 4.28: Best Results for ANN_0 and SVM_2 Malaria Model

4.6 Weaknesses and Strength of SVM and ANN

The following *Table 4.18* depicts the weaknesses and strength of the adopted models when trained single handedly. But in this research, the weaknesses of each model can be strengthened with the strength of the other one.

Table 4.18: Strength and Weaknesses of SVM and ANN

Methodology	Strength	Weaknesses
SVM	✓ Handles Bivariate prediction, pattern recognition, feature selection and classification ✓ Handles small and large dataset well ✓ Uses predefined activation function ✓ Solves the problems of over-fitting by optimizing the model parameters to feature selection	✓ Handles only binary prediction, pattern recognition, classification, and regression analysis ✓ Needs a 'good' kernel function. ✓ Chooses appropriately hyper parameters that will allow for sufficient generalization performance
ANN	✓ Creates network to have hidden neurons ✓ Handle Multivariate prediction, pattern recognition, classification, regression analysis. ✓ Uses predefined activation function ✓ Requires less formal statistical training ✓	✓ No general framework to design most suited network for particular problems ✓ Threshold frequency, number of hidden layers and hidden neurons are searched in the network by trial and error ✓ Greater computational burden because large parameters are needed to fit a good network structure ✓ Prone to local minima ✓ Over fitting often occurs because of large data to fix.

A need for Hybridized SVM_ANN Malaria model that gives higher prediction values resulting to greater reliability and robustness then arises. The Hybridized SVM_ANN Malaria model was thresholded to get the optimum threshold value that will yield best result for the model using the adaptive thresholding algorithm in Table 3.3.

4.7 Hybridized SVM_ANN Malaria Model Results

In this research, SVM Support Vectors serves as the input features to the Neural Network algorithm. The features are divided training (70%), testing (15%) and validation (15%) according to hold out cross validation sampling. At training section of the model, the thresholding ranges from 0.2800 to 1.000 to search for the optimal threshold that gives the highest accuracy with low error rate and low standard deviation. At the testing stage, the thresholding algorithm gave the best trade off search of the threshold optimal parameters of accuracy, sensitivity, specificity, false negative and false positive rate. Table 4.19 shows the results of SVM_2 with 308x8 double support vectors with radial basis function subjected to SVM_ANN Hybridized Malaria model.

Table 4.19 SVM_ANN Hybridized Malaria Model

SVM_ANN	Performance Metrics Accuracy (%)	Support Vectors (double)	Lang-range Multiplier	True Positive (TP)	True Negative(TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FP _R)	False Negative Rate (FN _R)	Mean Square Error (MSE)	Total Positive	Total Negative	Total
98.91		308 x 8	308 x 1	16	75	0	1	100	98.68	0.0131	0	0.1364	16	76	92

Table 4.19 shows the SVM_ANN Hybridized Malaria Model generates 100% TP_R , 94.74% TN_R , 5.26% FP_R and 0 (zero) FN_R . Fig 4.29 depicts the confusion matrix of the output class versus target class for SVM_ANN Hybridized Malaria Model.

Confusion Matrix

Output Class	1	16 17.39%	1 1.11%	98.68% 1.32%
	0	0 0%	75 81.52%	100% 0%
		100% 0%	98.68% 1.32%	98.91% 1.09%
		1	0	
		Target Class		

Figure 4.29: SVM_ANN Hybridized Malaria Model Confusion Matrix

In Figure 5.2 SVM_ANN Hybridized Malaria Confusion Matrix, sixteen (16) cases of Class_2 infected cases were correctly classified as positive. This corresponds to 17.39% of all ninety-two (92) malaria cases. Similarly, seventy-five (75) cases of Class_1 and Class_0 non infected cases were correctly classified

as negative. This corresponds to 81.52% of all malaria cases. Also, zero (0) case of Class_2 infected cases are incorrectly classified as non-infected cases. Similarly, only one (1) case of Class_1 and Class_0 non infected cases were incorrectly classified as infected. This corresponds to 1.11% of all malaria cases. Out of seventy-six (76) non infected cases, seventy-five (75) were rightly classified. This corresponds to 98.68% correctly classified while only one (1) which corresponds to 1.32% was wrongly classified. Similarly, all sixteen (16) infected cases were rightly classified as infected which corresponds to 100% correctly classified.

4.7.1 Overall Best Results for ANN, SVM and SVM_ANN Hybridized Malaria Model

Table 4.20: Overall Best Results for ANN, SVM and SVM_ANN Hybridized Malaria Model

Performance Metrics	Accuracy (%)	Support Vectors (double)	Lang-range Multiplier	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)	Sensitivity (TP _R)	Specificity (TN _R)	False Positive Rate (FP _R)	False Negative Rate FN _R	Mean Square Error (MSE)	Total Positive	Total Negative	Total
ANN_0	48.33	-	-	20	67	80	13	60.61	45.58	0.5442	0.3939	0.5550	33	47	180
SVM_2	85.60	308 x 8	308 x 1	58	96	15	11	84.06	86.49	0.1351	0.1594	0.5778	69	111	180
SVM_ANN	98.91	308 x 8	308 x 1	16	75	0	1	100	98.68	0.0131	0	0.1364	16	76	92

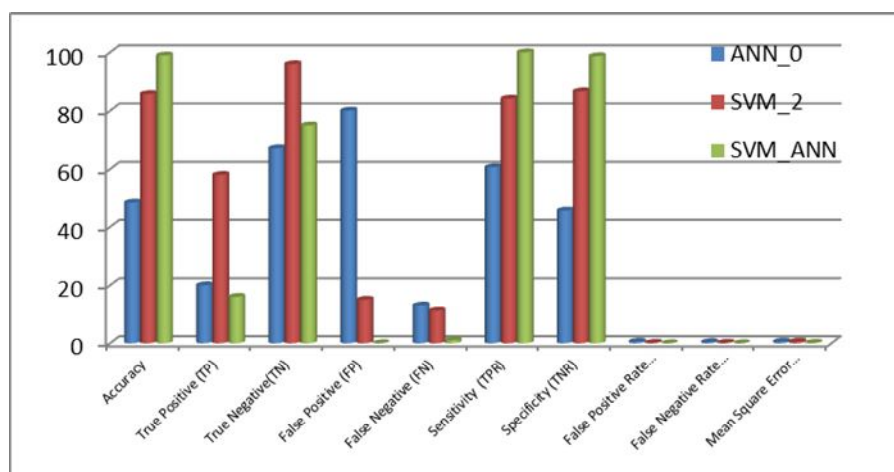


Figure 4.30: Performance of ANN, SVM and SVM_ANN Hybridized Malaria Model

From *Table 4.20* and *Figure 4.30*, it was observed that The ANN_0 Malaria Model generates the best results of 60.61% TP_R , 45.58% TN_R , 54.42% FP_R and 39.45% FN_R . Also, SVM_2 Malaria Model generates 84.06% TP_R , 86.49% TN_R , 13.51% FP_R and 15.94% FN_R . The optimal features selected by SVM_2 predicted well the occurrence of infected and non-infected cases in SVM_ANN Hybridized Malaria Model and generates the best results of 100% TP_R , 98.68% TN_R , 1.31% FP_R and 0 (zero) FN_R .

4.8 SVM_ANN Hybridized Malaria Model Formulation

Table 4.21 : Root of Equation

Input-Weight	Output-Weight
Input-Weight = net. IW	Output-Weight = net. LW
Input-Weight = $[5 \times 8 \text{ double}]$	Output-Weight = $[] []$
$[]$	$[1 \times 5 \text{ double}] []$
Root of the Equation = $[IW^T] [LW^T]$	
= 8 x 1 matrix	

From the *Table 4.21* the root of the malaria model is $[IW^T] [LW^T]$ as shown in equation (5.1) 8 x 1 matrix that follows:

$$Root = \begin{bmatrix} 0.038340456 \\ 0.072184123 \\ -0.176426924 \\ -0.023345842 \\ -0.086651486 \\ -0.456497928 \\ -0.209310882 \\ 0.10363675 \end{bmatrix} \quad (5.1)$$

4.8.1 *SVM_ANN Hybridized Malaria Quantitative Measure*

$$\begin{bmatrix} OP1 \end{bmatrix} = \begin{bmatrix} 0.0383 & 0.0722 & -0.1764 & -0.0233 & -0.0867 & -0.4565 & -0.2093 & 0.1036 \end{bmatrix} \begin{bmatrix} H_d \\ F_v \\ D_z \\ B_p \\ V_m \\ T_{emp} \\ R_h \\ R_f \end{bmatrix} \quad (5.2)$$

In SVM_ANN Hybridized Malaria model a logsig transfer function mapped the input features into the hidden layer while a Purelin activation function mapped the input from hidden layer to the output layer. By using the transfer functions for each channel, then the output layer OP1 of the model is a linear function:

$$\begin{aligned} \begin{bmatrix} OP1 \end{bmatrix} = & 0.0383 H_d + 0.0722 F_v - 0.1764 D_z - 0.0233 B_p - 0.0867 V_m - 0.4565 T_{emp} \\ & - 0.2093 R_h + 0.1036 R_f \end{aligned}$$

4.8.2 SVM_ANN Hybridized Malaria Model Result Discussion

From *Table 4.20* SVM_ANN Hybridized Malaria Model gave the best accuracy of 98.91% at threshold 0.600 with the least standard deviation 1.38 and MSE 0.14 thus:

- i. From SVM_ANN hybridized malaria model, 0.600 threshold values of the vector density means that only 60% of the female Anopheles mosquitoes that go for a blood meal survive to lay their eggs at the end of that cycle as they are subjected to monthly average daily temperature, rainfall and relative humidity.
- ii. The effects of rainfall, temperature and relative Humidity are the contributing climatic factors in the model.
- iii. The model was able to handle symptomatic and climatic based multiclass malaria infection with feed forward accurate measurement of 98.91%, 1.38 variations rate among the data and 0.14 back propagation error rate of the vector population.

4.9 Performance Evaluation

The performance evaluation of the developed SVM_ANN Hybridized Malaria model was evaluated based on the threshold metrics; accuracy, sensitivity, specificity, probability metrics; cited in *Table 2.1* under literature review. The model was comparatively evaluated with the existing baseline work as shown in *Table 4.22*.

Table 4.22 : Performance Comparison between SVM_ANN Model and Baseline Works

Author /Year	Methodology	Accuracy (%)	Sensitivity (%)	Specificity (%)	False Negative Rate (%)	False Positive Rate (%)
Júnior et al., (2010)	ANN & BN	80	67.5(ANN) 37.5(BN)	92.5(ANN) 97.5(BN)	32.5(ANN) 62.5(BN)	17.5(ANN) 2.5(BN)
Sakvare & Narote(2012)	Otsu Threshold, SVM Binary Classifier and SVM Multi Classifier	96.42% (SVM Multiclass Classifier	SVM Binary Classifier (96.26%) Sensitivity on 71 Images.	99.09% Specificity on 71 Images	3.74	0.91
Chavan, & Sutkar (2014)	ANN & SVM	78.53 (ANN) 98.25% (SVM)	100	88	0	12
H. Chiroma et al. (2014)	Jordan-Elman Neural Network	90	96.4 (MS) 91.2 (HS)	NA	3.6(HS) 8.8(MS)	NA
Dipti & Dilip, 2014	Naive Bayes & Bayesian Regularization	97.5 (NV) 92.5 (BR)	NA	NA	NA	NA
Shruti & Shirgan (2015)	ANN and SVM	78.53 (ANN) 98.25 (SVM)	NA	NA	NA	NA
Raviraja, et al., (2015)	ANN	96	NA	NA	NA	NA
Chaudhari& Agrawal, 2015	SVM & Naïve Bayes Classifier	100 (SVM) 92.85 (Naive Bayes)	NA	NA	NA	NA
Aminu, et al.,(2016)	SVM	90	NA	NA	NA	NA
Abisoye & Jimoh (2017)	SVM_ANN	98.91	100	98.68	0	1.32

From Table 4.23 at threshold 0.600, this research obtained an accuracy of 98.91%, with 100% ability to detect malaria infection and 98.68% ability to separate infected malaria cases from non-infected cases. The model obtained error rate of 0.14% and correct rate of 99.86%. This multiclass, symptomatic and climatic based research has shown better performance compared with other existing models.

Performance Parameter	Performance Value (%)
<i>Correct Rate</i>	<i>99.86</i>
<i>Error rate</i>	<i>0.14</i>
<i>Last Correct rate</i>	<i>99.86</i>
<i>Last Error Rate</i>	<i>0.14</i>
<i>Classified Rate</i>	<i>100</i>
<i>Sensitivity</i>	<i>100</i>
<i>Specificity</i>	<i>98.68</i>
<i>Positive Predictive Value</i>	<i>98.68</i>
<i>Negative Predictive Value</i>	<i>100</i>

Table 4.23 Performance of SVM_ANN Hybridized Malaria Model

4.10 Chapter Summary

This study has developed hybridized SVM_ANN malaria model to predict the classes of malaria infection according to its severity using Artificial Neural Network (ANN) and Support Vector Machine (SVM) in MATLAB Environment. The malaria symptomatic threats identified with climatic factors serves as the input features in the proposed model. Multilevel thresholding was conducted with Artificial Neural Network to get the optimum threshold with the best accuracy at lowest standard deviation and mean square error. In ANN, the best accuracy of 48.33% was achieved at the threshold level 0.600 with ability of test to identify the infected cases of 60.61% and ability to separate the infected cases from non-infected cases is 45.58%. In SVM, after several testing of the model with different kernel functions, the radial basis function gave the best SVM result. SVM_2 with radial basis function gave the best result for the developed model with an accuracy of 85.60% and 308 X 8 support vectors. These support vectors were used as input variables to the hybridized SVM_ANN malaria model. Hybridized SVM_ANN malaria model the produces an accuracy of 98.91%, 100% TP_R, 98.68% TN_R, 1.31% FP_R and zero (0) FN_R.

CHAPTER FIVE

5.0 SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Summary

The vast outbreaks of malaria cases and threat being imposed on humans have not received adequate prediction and therapy. Specifically, the existing models studied the asymptomatic and morphological characteristics of malaria infection. They could not predict symptomatic malaria incidence cases under effects of different climatic conditions.

This study has identified the internal and external factors that make an individual susceptible to the malaria infection. The internal factors include: frequency of human blood index (HBI), duration of sporogony, vector density, vector susceptibility and demography. Likewise, the external factors involve effects of temperature, rainfall, relative humidity vegetation, altitude, human behavioural factors and other environmental changes. The threats of the infection are vector threat and transmission threat which may be asymptomatic, symptomatic and climate based.

Appropriate relationship between factors and threat was constructed by selecting likelihood qualitative model for symptomatic and climate based malaria prediction. ANN was selection for feature selection but it has the weakness of over-fitting due to large number of parameters to fix and prone to local minimum error. Therefore, a feature selection algorithm SVM was

adopted which incorporated one-against-all algorithm to handle multiclass nature of malaria cases.

This study has developed a hybridized SVM_ANN malaria model to predict the severity of malaria parasite count. Developed hybridized SVM_ANN malaria model is capable of predicting multiclass nature of malaria infection. This research has explored the climatic effects on multiclass symptomatic malaria cases. The hybridized malaria model worked best at an optimum threshold value of 0.6000. An accuracy of 98.91%, 100% TP_R , 98.68% TN_R , 1.31% FP_R and zero (0) FN_R were derived from the model at this optimum threshold.

5.2 Conclusion

In this research work, Hybridized SVM_ANN malaria model was used to predict the multiclass nature of malaria infection. The validated test features greatly enhanced the parametric training and fine tuning of the developed model. Support Vector Machine (SVM) incorporated one-against-all approach to generate optimal features from the multiclass malaria features. These optimal features then served as input features for the multi-layered Neural Network (ANN).

The output of the dual model was thresholded. An Adaptive Thresholding Algorithm was developed to get optimum threshold level at best model performance accuracy. The algorithm was applied on a known multiclass

malaria features. The algorithm worked on the principle of search for all possible local and global maxima.

The model will always perform best with this adaptive thresholding algorithm for symptomatic malaria incidence cases compared to other existing models that do not consider global maximum threshold. Regardless of the dataset used, the developed Hybridized SVM_ANN malaria model will give optimum prediction of multiclass malaria incidence cases. The model will be able to handle low, mild and high nature of malaria infections cases.

5.3 Recommendation

In case of any dynamic multiclass system that aims to predict infection severity, study the effects of climatic factors, Hybridized SVM_ANN model is recommended. Re-training of the developed model is recommended to further ensure global maximum threshold and improve detection rate. The model will be helpful in taking appropriate prediction before diagnosis and medication. It will also help in taking appropriate actions by individuals, World Health Organizations and Government Agencies and its devastating impact will be reduced.

5.4 Future Work

This research work has focused on predicting multiclass symptomatic nature of malaria incidence cases with effects of climatic condition using Hybridized SVM_ANN malaria model. Due to inaccurate specificity rate, an enhanced model may be needed to improve the ability to separate infectious cases from

non-infectious cases. Further researches can focus on using other feature selection approaches to get optimal features to improve a new hybridized method.

5.5 Contribution to Knowledge

This study has contributed to knowledge by predicting the classes of severity of malaria cases using hybridized SVM_ANN approach that:

- i. Handles Symptomatic Nature of Malaria Incidence Cases;
- ii. Considers Effect of Climatic Factors on Malaria Incidence Cases;
- iii. Handles the Multiclass Nature of Malaria Incidence Cases;
- iv. Considers Global Maximum Threshold on Malaria Incidence Cases.

5.6 List of Published Work

1. R.G Jimoh, O.A. Abisoye & M.BM.B. Uthman, (2017). “Towards Classification of Symptomatic and Climatic Malaria Parasite-Count”, A paper presented at 11th International Multi-Conference on ICT Application, AICTTRA, 2017 Obafemi Awolowo University, Ile Ife, Nigeria.
2. Opeyemi Abisoye & Rasheed Jimoh (2017). “Symptomatic and Climatic Based Malaria Threat Detection using Multilevel Thresholding Feed-Forward Neural Network”. International Journal of Information Technology and Computer Science (IJITCS), 9,(8), 34-45, DOI: 10.5815/ijitcs.2017.08.05

3. Abisoye O.A & Jimoh R.G., (2015) “A Hybrid Intelligent Predicting Model To Determine Malaria Transmission”. A paper presented at International conference on Applied Information Technology(AIT), 2015, NCS, Ogun State Chapter, Nigeria .

5.7 Acknowledgement

1. Abdul Salam International Center for Theoretical Physics (ICTP), Trieste, Italy for organizing a Complex System Research Training between 30th July and 11th August, 2017.
2. University of Ilorin, Ilorin, Nigeria, TETFund Institutional Based Research Intervention (IBRI) Grant offered.
3. Federal University of Technology, Minna, Niger State, Local TetFund Award.

References

- Abraham, A., & Nath, B. (2000). Hybrid intelligent systems design: A review of a decade of research. *IEEE Transactions on Systems, Man and Cybernetics (Part C)*, 3(1), 1-37.
- Abu-Mostafa, Y. S., Magdon-Ismael, M., & Lin, H. T. (2012). *Learning from data* (Vol. 4). New York, NY, USA:: AMLBook.
- .Acampora, G., & Loia, V. (2005). Using FML and fuzzy technology in adaptive ambient intelligence environments. *International Journal of Computational Intelligence Research*, 1(1), 171-182.
- Adehor, A. B., & Burrell, P. R. (2008). The integrated management of health care strategies and differential diagnosis by expert system technology: A single-dimensional approach. *World Academy of Science, Engineering and Technology*, 44, 533-538.
- Adekoya, A. F., Akinwale, A. T., & Oke, O. E. (2008). A medical expert system for managing tropical diseases. *College of Natural Sciences Proceedings*, 74-86.
- Adetunmbi, A. O., Oguntimilehin, A., & Falaki, S. O. (2012). Web-based medical assistant system for malaria diagnosis and therapy. *GESJ: Computer Science and Telecommunications*, 1(33), 42-53.
- Agboizebeta, I. A., & Chukwuyeni, O. J. (2012). Cognitive analysis of multiple sclerosis utilizing fuzzy cluster means. *International Journal of Artificial Intelligence and Applications*, 3(1), 33.
- Aguilar, R. M., Muñoz, V., & Callero, Y. (2012). Control application using fuzzy logic: Design of a fuzzy temperature controller. In *Fuzzy Inference System-Theory and Applications*. InTech.
- Aladag, C. H., Egrioglu, E., & Kadilar, C. (2009). Predicting nonlinear time series with a hybrid methodology. *Applied Mathematics Letters*, 22(9), 1467-1470.
- Alonso, P. L., Brown, G., Arevalo-Herrera, M., Binka, F., Chitnis, C., Collins, F., ... & Mendis, K. (2011). A research agenda to underpin malaria eradication. *PLoS Medicine*, 8(1), e1000406.
- Alves, F. P., Durlacher, R. R., Menezes, M. J., Krieger, H., Silva, L. H. P., & Camargo, E. P. (2002). High prevalence of asymptomatic Plasmodium vivax and Plasmodium falciparum infections in native Amazonian populations. *The American Journal of Tropical Medicine and Hygiene*, 66(6), 641-648.
- Aminu, E. F., Ogbonnia, E. O., & Shehu, I. S. (2016). A Predictive Symptoms-based System using Support Vector Machines to enhanced Classification Accuracy of Malaria and Typhoid Coinfection.

- Anderson, R. M., May, R. M., & Gupta, S. (1989). Non-linear phenomena in host—parasite interactions. *Parasitology*, 99(S1), S59-S79.
- Anguita, D., Ghio, A., & Ridella, S. (2011). Maximal discrepancy for support vector machines. *Neurocomputing*, 74(9), 1436-1443.
- Anifowose, F. A., & Abdulraheem, A. (2010, January). Prediction of porosity and permeability of oil and gas reservoirs using hybrid computational intelligence models. In *North Africa Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Anifowose, F., & Abdulraheem, A. (2011). Fuzzy logic-driven and SVM-driven hybrid computational intelligence models applied to oil and gas reservoir characterization. *Journal of Natural Gas Science and Engineering*, 3(3), 505-517.
- Anifowose, F., Ewenla, A., & Eludiora, S. (2012, February). Prediction of Oil and Gas Reservoir Properties using Support Vector Machines. In *IPTC 2012: International Petroleum Technology Conference*.
- Anifowose, F., Labadin, J., & Abdulraheem, A. (2013). A least-square-driven functional networks type-2 fuzzy logic hybrid model for efficient petroleum reservoir properties prediction. *Neural Computing and Applications*, 23(1), 179-190.
- Arifianto, A., Barmawi, A. M., & Wibowo, A. T. (2014). Malaria incidence predicting from incidence record and weather pattern using Polynomial Neural Network. *International Journal of Future Computer and Communication*, 3(1), 60.
- Asadi, H., Dowling, R., Yan, B., & Mitchell, P. (2014). Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PloS one*, 9(2), e88225.
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: the support vector machines approach. *Journal of Hydrology*, 318(1), 7-16.
- Bai, Y., & Jin, Z. (2005). Prediction of SA+RS epidemic by BP neural networks with online prediction strategy. *Chaos, Solitons & Fractals*, 26(2), 559-569.
- Baig, F., Khan, M. S., Noor, Y., & Imran, M. (2011). Design model of fuzzy logic medical diagnosis control system. *International Journal on Computer Science and Engineering*, 3(5), 2093-2108.
- Bannister, L., & Mitchell, G. (2003). The ins, outs and roundabouts of malaria. *Trends in Parasitology*, 19(5), 209-213.
- Bao, Y., Hu, Z., & Xiong, T. (2013). A PSO and pattern search based memetic algorithm for SVMs parameters optimization. *Neurocomputing*, 117, 98-106.

- Bereczky, S., Liljander, A., Rooth, I., Faraja, L., Granath, F., Montgomery, S. M., & Färnert, A. (2007). Multiclonal asymptomatic *Plasmodium falciparum* infections predict a reduced risk of malaria disease in a Tanzanian population. *Microbes and Infection*, 9(1), 103-110.
- Bottius, E., Guanzirolli, A., Trape, J. F., Rogier, C., Konate, L., & Druilhe, P. (1996). Malaria: even more chronic in nature than previously thought; evidence for subpatent parasitaemia detectable by the polymerase chain reaction. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 90(1), 15-19.
- Bousema, T., Okell, L., Felger, I., & Drakeley, C. (2014). Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nature Reviews. Microbiology*, 12(12), 833.
- Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: predicting and control, revised ed.* Holden-Day.
- Brauer, F., Castillo-Chavez, C., & Castillo-Chavez, C. (2001). *Mathematical models in population biology and epidemiology* (Vol. 1). New York: Springer.
- Brown, T. M., Cueto, M., & Fee, E. (2006). The World Health Organization and the transition from “international” to “global” public health. *American Journal of Public Health*, 96(1), 62-72.
- Caravaca Moreno, J., Soria Olivas, E., Bataller Mompeán, M., Serrano López, A. J., Such Miquel, L., Vila Francés, J., & Guerrero Martínez, J. F. (2014). Application of machine learning techniques to analyse the effects of physical exercise in ventricular fibrillation. *Computers in Biology and Medicine*, 2014, vol. 45, num. 1, p. 1-7.
- Castillo, O., & Melin, P. (2008). Intelligent systems with interval type-2 fuzzy logic. *International Journal of Innovative Computing, Information and Control*, 4(4), 771-783.
- Centers for Disease Control and Prevention. (2013). *CDC Health Information for International Travel 2014: The Yellow Book*. Oxford University Press.
- Ch, S., & Mathur, S. (2010). Modeling uncertainty analysis in flow and solute transport model using Adaptive Neuro Fuzzy Inference System and particle swarm optimization. *KSCE Journal of Civil Engineering*, 14(6), 941-951.
- Ch, S., & Mathur, S. (2012). Particle swarm optimization trained neural network for aquifer parameter estimation. *KSCE Journal of Civil Engineering*, 16(3), 298-307.
- Ch, S., Sohani, S. K., Kumar, D., Malik, A., Chahar, B. R., Nema, A. K., ...& Dhiman, R. C. (2014). A support vector machine-firefly algorithm based predicting model to determine malaria transmission. *Neurocomputing*, 129, 279-288.

- Chapelle, O., Vapnik, V., Bousquet, O., & Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Machine Learning*, 46(1-3), 131-159.
- Chaudhari, t., & Agrawal, (2015). The Automatic Detection of Malaria Parasites for Estimating Parasitemia.
- Chavan, S. N., & Sutkar, A. M. (2014). Malaria disease identification and analysis using image processing. *International. Journal of Latest Trends Engineering. Technology*, 3, 218-223.
- Cherif, A. H., Movahedzadeh, F., Michel, L., Hill, A., & Jedlicka, D. M.(2011) Environmental Release of Genetically Engineered Mosquitoes
- Cherkassky, V., & Mulier, F. M. (2007). *Learning from data: concepts, theory, and methods*. John Wiley & Sons.
- Chiroma, H., Abdul-kareem, S., Ibrahim, U., Ahmad, I. G., Garba, A., Abubakar, A. & Herawan, T. (2015). Malaria severity classification through Jordan-Elman neural network based on features extracted from thick blood smear. *Neural Network World*, 25(5), 565.
- Christophers, S. R. (1911). Epidemic malaria of the Punjab: with a note of a method of predicting epidemic years. *Trans Committee Stud Malaria India*, 2, 17-26.
- Chung, K. M., Kao, W. C., Sun, C. L., Wang, L. L., & Lin, C. J. (2003). Radius margin bounds for support vector machines with the RBF kernel. *Neural computation*, 15(11), 2643-2681.
- Cohen, J. E., & Gürtler, R. E. (2001). Modeling household transmission of American trypanosomiasis. *Science*, 293(5530), 694-698.
- Collins, F. H., & Paskewitz, S. M. (1995). Malaria: current and future prospects for control. *Annual Review of Entomology*, 40(1), 195-219.
- Collins, F. H., Mendez, M. A., Rasmussen, M. O., Mehaffey, P. C., Besansky, N. J., & Finnerty, V. (1987). A ribosomal RNA gene probe differentiates member species of the *Anopheles gambiae* complex. *The American Journal of Tropical Medicine And Hygiene*, 37(1), 37-41.
- Collobert, R., & Bengio, S. (2000). *Support vector machines for large-scale regression problems* (No. EPFL-REPORT-82604).IDIAP.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Coura, J. R., Suárez-Mutis, M., & Ladeia-Andrade, S. (2006). A new challenge for malaria control in Brazil: asymptomatic *Plasmodium* infection-a review. *Memórias do Instituto Oswaldo Cruz*, 101(3), 229-237.

- Craig, M. H., Kleinschmidt, I., Le Sueur, D., & Sharp, B. L. (2004). Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: part II. The impact of non-climatic factors. *Tropical Medicine & International Health*, 9(12), 1258-1266.
- Cucunubá, Z. M., Guerra, Á. P., Rahirant, S. J., Rivera, J. A., Cortés, L. J., & Nicholls, R. S. (2008). Asymptomatic plasmodium spp. infection in Tierralta, Colombia. *Memórias do Instituto Oswaldo Cruz*, 103(7), 668-673.
- Danger, R., Segura-Bedmar, I., Martínez, P., & Rosso, P. (2010). A comparison of machine learning techniques for detection of drug target articles. *Journal of biomedical informatics*, 43(6), 902-913.
- de Mast, Q., Syafruddin, D., Keijmel, S., Riekerink, T. O., Deky, O., Asih, P. B., ... & van der Ven, A. J. (2010). Increased serum hepcidin and alterations in blood iron parameters associated with asymptomatic *P. falciparum* and *P. vivax* malaria. *Haematologica*, 95(7), 1068-1074.
- Depinay, J. M. O., Mbogo, C. M., Killeen, G., Knols, B., Beier, J., Carlson, J., ... & McKenzie, F. E. (2004). A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission. *Malaria Journal*, 3(1), 29.
- Di Ruberto, C., Dempster, A., Khan, S., & Jarra, B. (2000). Automatic thresholding of infected blood images using granulometry and regional extrema. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (Vol. 3, pp. 441-444). IEEE.
- Díaz, G., González, F. A., & Romero, E. (2009). A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. *Journal of Biomedical Informatics*, 42(2), 296-307.
- Diaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., & Moncada-Herrera, J. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile. *Atmospheric Environment*, 42(35), 8331-8340.
- Dipti D. Patankar & Dilip G Khairnar (2014). Comparison of the Performance of Different Networks for Malaria Classification, *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181, 3(4).
- Djam, X. Y., Wajiga, G. M., Kimbi, Y. H., & Blamah, N. V. (2011). A fuzzy expert system for the management of Malaria. *International Journal of Pure and Applied Sciences and Technology*, 84-108, ISSN 2229-610
- Dodd, R. Y. (1998). Transmission of parasites by blood transfusion. *Vox Sanguinis*, 74(S2), 161-163.

- Donfack Kana A.F, Abdullahi M, Ezugou A.E and Alkali S.A (2010), Online System for diagnosis and treatment of malaria . International Journal for the Application of Wireless and Mobile Computing. 2.
- Duarte, J., Deshpande, P., Guiyedi, V., Mécheri, S., Fesel, C., Cazenave, P. A., ... & Pied, S. (2007). Total and functional parasite specific IgE responses in Plasmodium falciparum-infected patients exhibiting different clinical status. *Malaria journal*, 6(1), 1.
- Duenyas, S., & Margaliot, M. (2011, April). Knowledge extraction from a class of support vector machines using the fuzzy all-permutations rule-base. In *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on* (pp. 1-7). IEEE.
- Duodu, Q., Panford, J. K., & Hayfron-acquah, J. B. (2014). Designing Algorithm for Malaria Diagnosis using Fuzzy Logic for Treatment (AMDFLT) in Ghana..*International Journal of Computer Applications*, 91(17).
- Faruk, D. Ö. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, 23(4), 586-594.
- Feng, Z., Dieckmann, U., & Levin, S. A. (Eds.). (2006). *Disease evolution: models, concepts, and data analyses* (Vol. 71). American Mathematical Soc.
- Finkel, M. (2007). Stopping a global killer. *National Geographic*, July.
- Friedrichs, F., & Igel, C. (2005). Evolutionary tuning of multiple SVM parameters. *Neurocomputing*, 64, 107-117.
- Gillies, Michael Thomas, and Botha De Meillon. "The anophelinae of Africa south of the Sahara (Ethiopian zoogeographical region)." *The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region)*. (1968).
- Glass, P. S., & White, P. F. (2007). Practice guidelines for the management of postoperative nausea and vomiting: past, present, and future. *Anesthesia & Analgesia*, 105(6), 1528-1529.
- Goswami, S., & Chakrabarti, A. (2014). Feature selection: A practitioner view. *International Journal of Information Technology and Computer Science (IJITCS)*, 6(11), 66.
- Greenwood, B. M. (1997). The epidemiology of malaria. *Annals of tropical medicine and parasitology*, 91(7), 763-770.
- Greenwood, M. (1916). The application of mathematics to epidemiology. *Nature*, 97, 243-244.
- Guazzelli, A. (2012). "Predicting the future, part 2: Predictive modeling techniques.
- Gunder, L., & Martin, S. A. (2011). *Essentials of medical genetics for health professionals*. Jones & Bartlett Learning.

- Gundin, D. A., Garcia, C., Gomez, E., Dimitriadis, Y. A., & Vega, G. (2002). Short-term load predicting for industrial customers using FasArt and FasBack neuro-fuzzy systems.
- Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report*, 14.
- Halim, S., Bretschneider, T. R., Li, Y., Preiser, P. R., & Kuss, C. (2006). Estimating malaria parasitaemia from blood smear images. In *Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on* (pp. 1-6). IEEE.
- Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S., & Samuel, M. D. (2002). Climate warming and disease risks for terrestrial and marine biota. *Science*, 296(5576), 2158-2162.
- Hay, S. I., Gething, P. W., & Snow, R. W. (2010). India's invisible malaria burden. *Lancet*, 376(9754), 1716.
- Hay, S. I., Rogers, D. J., Randolph, S. E., Stern, D. I., Cox, J., Shanks, G. D., & Snow, R. W. (2002). Hot topic or hot air? Climate change and malaria resurgence in East African highlands. *Trends in parasitology*, 18(12), 530-534.
- Hegazy, O., Soliman, O. S., & Salam, M. A. (2013). A machine learning model for stock market prediction. *International Journal of Computer Science and Telecommunications*, 4(12), 16-23.
- Hethcote, H., Zhien, M., & Shengbing, L. (2002). Effects of quarantine in six endemic models for infectious diseases. *Mathematical Biosciences*, 180(1), 141-160.
- Hochreiter, S. (2013). Basic Methods of Data Analysis.
- Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to support vector classification.
- Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, 23(4), 725-749.
- Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science*, 15(2), 219-250.
- Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *IEEE computer*, 29(3), 31-44.
- Jain, R., (2003). The Third International Conference on Hybrid Intelligent Systems. HIS'03 <http://his03.hybridsystem.com> 14-17, December 2003 , Melbourne, Australia.

- Jayavardhana Rama, G. L., Shilton, A., Parker, M. M., & Palaniswami, M. (2005, December). Disulphide bridge prediction using fuzzy support vector machines. In *Intelligent Sensing and Information Processing, 2005.ICISIP 2005. Third International Conference on* (pp. 48-54). IEEE.
- Ji, Y., & Sun, S. (2013). Multitask multiclass support vector machines: model and experiments. *Pattern Recognition*, 46(3), 914-924.
- Jian, H., & Wenfen, H. (2006). Novel approach to predict potentiality of enhanced oil recovery. In *Proc. Soc. Petrol. Engineers Intelligent Energy Conf. Exhibition*.
- Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Springer, Berlin, Heidelberg.
- Joshi, S., Ramakrishnan, G., & Chandra, S. (2012). Using sequential unconstrained minimization techniques to simplify SVM solvers. *Neurocomputing*, 77(1), 253-260.
- Júnior, A. M. B., Duarte, A. A., Netto, M. B., & Andrade, B. B. (2010, July). Artificial Neural Networks and Bayesian Networks as supporting tools for diagnosis of asymptomatic malaria. In *e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on* (pp. 106-111). IEEE.
- Kahraman, C., Ertay, T., & Büyüközkan, G. (2006). A fuzzy optimization model for QFD planning process using analytic network approach. *European Journal of Operational Research*, 171(2), 390-411.
- Kamukama, I. (2009). Clinical protocol-based decision support system for malaria treatment.
- Karnik, N. N., Mendel, J. M., & Liang, Q. (1999). Type-2 fuzzy logic systems. *IEEE transactions on Fuzzy Systems*, 7(6), 643-658.
- Kecman, V. (2001). *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*. MIT press, Cambridge, Massachusetts.
- Keeling, M. J., & Eames, K. T. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4), 295-307.
- Keeling, M. J., & Rohani, P. (2008). *Modeling infectious diseases in humans and animals*. Princeton University Press, New Jersey 08540 USA
- Keeling, M. J., Danon, L., Vernon, M. C., & House, T. A. (2010). Individual identity and movement networks for disease metapopulations. *Proceedings of the National Academy of Sciences*, 107(19), 8866-8870.
- Kermack, W. O., & McKendrick, A. G. (1927, August). A contribution to the mathematical theory of epidemics. In *proceedings of the royal society of*

London A: Mathematical, physical and engineering sciences (Vol. 115, No. 772, pp. 700-721). The Royal Society.

- Keyvanfar, F., Shoorehdeli, M. A., & Teshnehlal, M. (2011, June). Feature selection and classification of breast MRI lesions based on Multi classifier. In *Artificial Intelligence and Signal Processing (AISP), 2011 International Symposium on* (pp. 54-58). IEEE.
- Khasnis, A. A., & Nettleman, M. D. (2005). Global warming and infectious disease. *Archives of Medical Research*, 36(6), 689-696.
- Kim, Y. H., Yoo, S. J., Gu, Y. H., Lim, J. H., Han, D., & Baik, S. W. (2014). Crop Pests Prediction Method Using Regression and Machine Learning Technology: Survey. *IERI Procedia*, 6, 52-56.
- Kitchen, A. D., & Chiodini, P. L. (2006). Malaria and blood transfusion. *Vox sanguinis*, 90(2), 77-84.
- Koenraadt, C. J. M., & Harrington, L. C. (2008). Flushing effect of rain on container-inhabiting mosquitoes *Aedes aegypti* and *Culex pipiens* (Diptera: Culicidae). *Journal of medical entomology*, 45(1), 28-35.
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology journal*, 13, 8-17.
- Kulkarni, P., Joshi, S., & Brown, M. S. (2016). *Big Data Analytics*. PHI Learning Pvt. Ltd..
- Lancaster, S. S. (2004, June). A fuzzy logic controller for the application of skin pressure. In *Fuzzy Information, 2004. Processing NAFIPS'04. IEEE Annual Meeting of the* (Vol. 2, pp. 686-689). IEEE.
- Leoratti, F. M., Durlacher, R. R., Lacerda, M. V., Alecrim, M. G., Ferreira, A. W., Sanchez, M. C., & Moraes, S. L. (2008). Pattern of humoral immune response to *Plasmodium falciparum* blood stages in individuals presenting different clinical expressions of malaria. *Malaria Journal*, 7(1), 1.
- Lizhe, Y., Bo, X., & Xianjie, W. (2010, May). BP network model optimized by adaptive genetic algorithms and the application on quality evaluation for class teaching. In *Future Computer and Communication (ICFCC), 2010 2nd International Conference on* (Vol. 3, pp. V3-273). IEEE.
- Lorena, A. C., & De Carvalho, A. C. (2008). Evolutionary tuning of SVM parameter values in multiclass problems. *Neurocomputing*, 71(16), 3326-3334.
- Mahfouf, M., Abbod, M. F., & Linkens, D. A. (2001). A survey of fuzzy logic monitoring and control utilisation in medicine. *Artificial Intelligence in Medicine*, 21(1-3), 27-42.

- Males, S., Gaye, O., & Garcia, A. (2008). Long-term asymptomatic carriage of *Plasmodium falciparum* protects from malaria attacks: A prospective study among Senegalese children. *Clinical Infectious Diseases*, 46(4), 516-522.
- Mali, S., Steele, S., Slutsker, L., & Arguin, P. M. (2009). Malaria surveillance—United States, 2007. *MMWR Surveill Summ*, 58(SS-2), 1-16.
- Marchesini, P., & Crawley, J. (2004). Reducing the burden of malaria in pregnancy. *MERA IV, supporting agency—Roll Back Malaria, WHO*.
- Martens, W. J., Niessen, L. W., Rotmans, J., Jetten, T. H., & McMichael, A. J. (1995). Potential impact of global climate change on malaria risk. *Environmental Health Perspectives*, 103(5), 458.
- Martha, R. C. D., Tada, M. S., Ferreira, R. G. D. M., Silva, L. H. P. D., & Wunderlich, G. (2007). Microsatellite characterization of *Plasmodium falciparum* from symptomatic and non-symptomatic infections from the Western Amazon reveals the existence of non-symptomatic infection-associated genotypes. *Memórias do Instituto Oswaldo Cruz*, 102(3), 293-298.
- Martínez-Martínez, J. M., Escandell-Montero, P., Barbieri, C., Soria-Olivas, E., Mari, F., Martínez-Sober, M., ...& Gatti, E. (2014). Prediction of the hemoglobin level in hemodialysis patients using machine learning techniques. *Computer Methods And Programs in Biomedicine*, 117(2), 208-217.
- Mathur, S. (2012). Groundwater level Predicting using SVM-PSO. *International Journal of Hydrology Science and Technology*, 2(2), 202-218.
- May, R. M., & Anderson, R. M. (1979). Population biology of infectious diseases: Part II. *Nature*, 280(5722), 455-461.
- McLeod, K. S. (2000). Our sense of Snow: the myth of John Snow in medical geography. *Social Science & Medicine*, 50(7), 923-935.
- McMichael, A. J., Woodruff, R. E., & Hales, S. (2006). Climate change and human health: present and future risks. *The Lancet*, 367(9513), 859-869.
- Mendel, J. M. (2003). Type-2 fuzzy sets: some questions and answers. *IEEE Connections, Newsletter of the IEEE Neural Networks Society*, 1, 10-13.
- Mendis, K., Sina, B. J., Marchesini, P., & Carter, R. (2001). The neglected burden of *Plasmodium vivax* malaria. *The American Journal Of Tropical Medicine And Hygiene*, 64(1 suppl), 97-106.
- Mohammadi, K., Shamshirband, S., Tong, C. W., Arif, M., Petković, D., & Ch, S. (2015). A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation. *Energy Conversion and Management*, 92, 162-171.

- Molineaux, L., Wernsdorfer, W. H., & McGregor, I. (1988). The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. *Malaria: Principles and Practice Of Malariology. Volume 2.*, 913-998.
- Moreira, L. A., Iturbe-Ormaetxe, I., Jeffery, J. A., Lu, G., Pyke, A. T., Hedges, L. M., ... & Hugo, L. E. (2009). A Wolbachia symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and Plasmodium. *Cell*, 139(7), 1268-1278.
- Mueller, I., Galinski, M. R., Baird, J. K., Carlton, J. M., Kochar, D. K., Alonso, P. L., & del Portillo, H. A. (2009). Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. *The Lancet Infectious Diseases*, 9(9), 555-566.
- Mukkamala, S., Janoski, G., & Sung, A. (2002). Intrusion detection using neural networks and support vector machines. In *Neural Networks, 2002.IJCNN'02.Proceedings of the 2002 International Joint Conference on* (Vol. 2, pp. 1702-1707). IEEE.
- Munakata, T. (2008). *Fundamentals of the new artificial intelligence: Neural, evolutionary, fuzzy and more*. Springer Science & Business Media, New York.
- Mutabingwa, T. K. (2005). Artemisinin-based combination therapies (ACTs): best hope for malaria treatment but inaccessible to the needy!. *Acta tropica*, 95(3), 305-315.
- Mwangi, T. W., Fegan, G., Williams, T. N., Kinyanjui, S. M., & Snow, R. W. (2008). Evidence for Over-Dispersion in the Distribution of Clinical Malaria Episodes.
- Niksaz, P., & Mohammad Latif, A. (2014). Rainfall Events Evaluation Using Adaptive Neural-Fuzzy Inference System. *International Journal of Information Technology and Computer Science (IJITCS)*, 6(9), 46.
- Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann, Burlington, Massachusetts, San Francisco, California.
- Nizamuddin, M. (2010). *Hybrid application of AVHRR based satellite remote sensing and ENSO signals for early warning and monitoring of Malaria in Asia and South America*. City University of New York.
- Njama-Meya, D., Kamya, M. R., & Dorsey, G. (2004). Asymptomatic parasitaemia as a risk factor for symptomatic malaria in a cohort of Ugandan children. *Tropical Medicine & International Health*, 9(8), 862-868.

- Nordqvist, C. (2013). Medical News Today. *Sexual Health/STDs*, "What are Crabs (pubic lice).
- Obot, O. U., & Uzoka, F. M. E. (2008). Fuzzy rule-based framework for the management of tropical diseases. *International Journal of Medical Engineering and Informatics*, 1(1), 7-17.
- Ogundipe, S., & Obinna, C. (2010). Malaria: Killer at large. *Vanguard Newspaper*, September, 26, 2010.
- Oguntimilehin, A., Adetunmbi, A. O., & Abiola, O. B. (2015). A Review of Predictive Models on Diagnosis and Treatment of Malaria Fever. *International Journal of Computer Science and Mobile Computing*, 4(5), 1087-1093.
- Okwa, O. O. (2003). The status of malaria among pregnant women: a study in Lagos, Nigeria. *African Journal of Reproductive Health*, 77-83.
- Olabiyisi, S. O., Omidiora, E. O., Olaniyan, M. O., & Dorikoma, O. (2011). A Decision Support System Model for Diagnosing Tropical Diseases Using Fuzzy Logic. *African Journal of Computing & ICT September*.
- Oluwagbemi, O. O., Ofoezie, U., & Nwinyi, O. (2010). A Knowledge-based data mining system for diagnosing malaria related cases in Healthcare Management. *Egyptian Computer Science Journal*, 34(4).
- Omidiora, E. O., Olaniyan, M. O., & Derikoma, O. (2011). A Decision Support System Model A Decision Support System Model for Diagnosing Tropical Diseases Using Fuzzy Logic Diagnosing Tropical Diseases Using Fuzzy Logic.
- Onuwa, O. B. (2014). Fuzzy Expert System for Malaria Diagnosis.
- Osowski, S., Siroic, R., Markiewicz, T., & Siwek, K. (2009). Application of support vector machine and genetic algorithm for improved blood cell recognition. *IEEE Transactions on Instrumentation and Measurement*, 58(7), 2159-2168.
- Owoseni, A. T., & Ogundahunsi, I. O. (2016). Mobile-Based Fuzzy Expert System for Diagnosing Malaria (MFES). *International Journal of Information Engineering and Electronic Business*, 8(2), 14.
- Paokanta, P., Ceccarelli, M., & Srichairatanakoo, S. (2010, November). The efficiency of data types for classification performance of Machine Learning Techniques for screening β -Thalassemia. In *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium of Applied Sciences in Biomedical and Communication Techn.* (pp. 1-4).
- Parkhi, V., Pawar, P., & Surve, A. (2013). Computer automation for malaria parasite detection using linear programming. *Int. J. Adv. Res. Electr. Elect. Instr. Eng*, 2, 1984-1988.

- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259-268.
- Peddabachigari, S., Abraham, A., Grosan, C., & Thomas, J. (2007). Modeling intrusion detection system using hybrid intelligent systems. *Journal of Network And Computer Applications*, 30(1), 114-132.
- Phillips-Wren, G., Ichalkaranje, N., & Jain, L. C. (2008). *Intelligent decision making: an AI-based approach* (Vol. 97). Springer Science & Business Media, New York.
- Prosperi, M. C., D'Autilia, R., Incardona, F., De Luca, A., Zazzi, M., & Ulivi, G. (2009). Stochastic modelling of genotypic drug-resistance for human immunodeficiency virus towards long-term combination therapy optimization. *Bioinformatics*, 25(8), 1040-1047.
- Purwanto, C. E., & Logeswaran, R. (2011). An Optimally Configured Hybrid Model for Healthcare Time Series Prediction. *Asian Journal of Information Technology*, 10(6), 209-217.
- Qi, D., Liu, M., Qiu, M., & Zhang, S. (2010). Exponential synchronization of general discrete-time chaotic neural networks with or without time delays. *IEEE Transactions on Neural Networks*, 21(8), 1358-1365.
- Randolph, S. E. (2008). Tick-borne disease systems. *Rev sci tech Off int Epiz*, 27(2), 1-15.
- Rahimikhoob, A. (2010). Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. *Renewable Energy*, 35(9), 2131-2135.
- Rajasekaran, S., Gayathri, S., & Lee, T. L. (2008). Support vector regression methodology for storm surge predictions. *Ocean Engineering*, 35(16), 1578-1587.
- Raviraja, S., Geethanjali, S.D., Chethana, C., & Kanthesh, B. M. The Classification and Recognition of Plasmodium Parasite in Prediction of Malaria Infected Blood Smears Using Artificial Intelligence Technique. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)* , 5(7), 863-886.
- Read, J. M., Eames, K. T., & Edmunds, W. J. (2008). Dynamic social networks and the implications for the spread of infectious disease. *Journal of The Royal Society Interface*, 5(26), 1001-1007.
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. In *Encyclopedia of Database Systems* (pp. 532-538). Springer USA.
- Rejani, Y., & Selvi, S. T. (2009). Early detection of breast cancer using SVM classifier technique. *arXiv preprint arXiv:0912.2314*.

- Reitermanova, Z.(2010). Data Splitting in WDS(Vol 10,pp.31-36).
- Sallares, R. (2002). *Malaria and Rome: A history of malaria in ancient Italy*. Oxford University Press on Demand, Oxford, England, UK.
- Savkare, S. S., & Narote, S. P. (2012). Automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage. *Procedia Technology*, 6, 405-410.
- Shamshirband, S., Petković, D., Saboohi, H., Anuar, N. B., Inayat, I., Akib, S., ...& Gani, A. (2014). Wind turbine power coefficient estimation by soft computing methodologies: comparative study. *Energy Conversion and Management*, 81, 520-526.
- Sharma, P., Singh, D. B. V., Bandil, M. K., & Mishra, N. (2013). Decision Support System for Malaria and Dengue Disease Diagnosis (DSSMD). *International Journal of Information and Computation Technology ISSN0974-2239*, 3(7), 633-640.
- Sharma, V., Kumar, A., Panat, L., Karajkhede, G., & Lele, A. (2015). Malaria outbreak prediction model using machine learning.. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* 4, 4415-4419.
- .Sharma, V., Rai, S., & Dev, A. (2012). A comprehensive study of artificial neural networks. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(10), India
- Shawe-Taylor, J., & Sun, S. (2011). A review of optimization methodologies in support vector machines. *Neurocomputing*, 74(17), 3609-3618.
- Shruti A. & Shirgan S.S (2015). Automatic Diagnosis of Malaria Parasites Using Neural Network and Support Vector Machine., *International Journal of Advanced Foundation in Computer(IJAFRC)*, 2, 62 -65.
- Siettos, C. I., & Russo, L. (2013).Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4), 295-306.
- Singh, S. (2006). New developments in diagnosis of leishmaniasis. *Indian Journal of Medical Research*, 123(3), 311.
- Sinka, M. E., Bangs, M. J., Manguin, S., Coetzee, M., Mbogo, C. M., Hemingway, J., ... & Okara, R. M. (2010). The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis. *Parasites & vectors*, 3(1), 117.
- Smith, T., Killeen, G. F., Maire, N., Ross, A., Molineaux, L., Tediosi, F., ... & Tanner, M. (2006). Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of Plasmodium falciparum malaria: Overview. *The American Journal of Tropical Medicine and Hygiene*, 75(2 suppl), 1-10.

- Soltani, M., Moghaddam, T. B., Karim, M. R., Shamshirband, S., & Sudheer, C. (2015). Stiffness performance of polyethylene terephthalate modified asphalt mixtures estimation using support vector machine-firefly algorithm. *Measurement*, 63, 232-
- Soomro, A. A., Memon, N. A., & Memon, M. S. (2011). Knowledge based expert system for symptomatic automated healthcare. *Sindh University Research Journal-SURJ (Science Series)*, 43(1 (a)).
- Sproule, B. A., Naranjo, C. A., & Türksen, I. B. (2002). Fuzzy pharmacology: theory and applications. *Trends in Pharmacological Sciences*, 23(9), 412-417.
- Stekeetee, R. W., Wirima, J. J., Hightower, A. W., Slutsker, L., Heliman, D. L., & Breman, J. G. (1996). The effects of malaria. Malaria prevention in pregnancy on Offspring birth weight, Prematurity and Intra-uterine Retardation in rural Malawi. *American Journal of Tropical Medicine and Hygiene*, 2, 824-832.
- Sun, S. (2013). A survey of multi-view machine learning. *Neural Computing and Applications*, 23(7-8), 2031-2038.
- Sun, S., Huang, R., & Gao, Y. (2012). Network-scale traffic modeling and Predicting with graphical lasso and neural networks. *Journal of Transportation Engineering*.
- Sun, Z. L., Au, K. F., & Choi, T. M. (2007). A neuro-fuzzy inference system through integration of fuzzy logic and extreme learning machines. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5), 1321-1331.
- Sung, A. H., & Mulkamala, S. (2003, January). Identifying important features for intrusion detection using support vector machines and neural networks. In *Applications and the Internet, 2003.Proceedings. 2003 Symposium on* (pp. 209-216). IEEE.
- Suyaphun, A., Wiwanitkit, V., Suwansaksri, J., Nithiuthai, S., Sritar, S., Suksirisampant, W., & Fongsungnern, A. (2002). Malaria among hilltribe communities in northern Thailand: a review of clinical manifestations. *Headache*, 27, 28-7.
- Symeonidis, A. L., & Mitkas, P. A. (2006). *Agent intelligence through data mining* (Vol. 14). Springer Science & Business Media, New York.
- Tao, X., Renmu, H., Peng, W., & Dongjie, X. (2004, April). Input dimension reduction for load Predicting based on support vector machines. In *Electric Utility Deregulation, Restructuring and Power Technologies, 2004.(DRPT 2004). Proceedings of the 2004 IEEE International Conference on* (Vol. 2, pp. 510-514).IEEE.

- Tesi, M. K. (2011). Discussion Paper Series *Global Warming and Health: The Issue of Malaria in Eastern Africa's Highlands*,(2). Centre for International Governance Innovation.
- Thomas, L. O. (2014). *A Comprehensive Review of Malaria with an Emphasis on Plasmodium Resistance* (Doctoral dissertation, University of Mississippi).
- Thomas, M. B., & Read, A. F. (2007). Can fungal biopesticides control malaria?. *Nature Reviews. Microbiology*, 5(5), 377.
- Trape, J. F., Zoulani, A., & Quinet, M. C. (1987). Assessment of the incidence and prevalence of clinical malaria in semi-immune children exposed to intense and perennial transmission. *American Journal of Epidemiology*, 126(2), 193-201.
- Triantaphyllou, E., & Felici, G. (2006). *Data mining and knowledge discovery approaches based on rule induction techniques* (Vol. 6). Springer Science & Business Media, New York.
- UNICEF. (2000). Promoting rational use of drugs and correct case management in basic health services. *The Prescriber*, (18).
- Unkel, S., Farrington, C., Garthwaite, P. H., Robertson, C., & Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1), 49-82.
- Urquiza, J. M., Rojas, I., Pomares, H., Herrera, J., Florido, J. P., Valenzuela, O., & Cepero, M. (2012). Using machine learning techniques and genomic/proteomic information from known databases for defining relevant features for PPI classification. *Computers in biology and medicine*, 42(6), 639-650.
- Uzoka, F. M., & Barker, K. D. (2005, April). Medical Decision Support Systems Using Analytic Hierarchy Process: A Case Study of Malaria Diagnosis. Presented at Med-e-tel 2005 Conference in Luxembourg.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media, New York.
- Vapnik, V. N., & Vapnik, V. (1998). *Statistical learning theory* (Vol. 1). New York: Wiley.
- Vapnik, V., Golowich, S. E., & Smola, A. (1996). Support vector method for function approximation, regression estimation, and signal processing. In *Advances in Neural Information Processing Systems 9*.
- Vardavas, R., & Blower, S. (2007). The emergence of HIV transmitted resistance in Botswana: “when will the WHO detection threshold be exceeded?”. *PLoS One*, 2(1), e152.

- Wang, L., & Fu, X. (2006). *Data mining with computational intelligence*. Springer Science & Business Media, New York.
- Wang, Y. (2008). Fuzzy clustering analysis by using genetic algorithm. *ICIC Express Letters*, 2(4), 331-337.
- Warrell, D. A. (1997). Cerebral malaria: clinical pathophysiology and treatment features. *Annals of Tropical Medicine and Parasitology*, 91(7), 875-884.
- Werbos, P. J. (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, 1(4), 339-356.
- White, N. J. (2004). Antimalarial drug resistance. *Journal of Clinical Investigation*, 113(8), 1084.
- World Health Organization (WHO), World malaria report 2013. World Health Organization
- World Health Organization (WHO), World malaria report 2014. World Health Organization
- World Health Organization (WHO), World malaria report 2015. World Health Organization
- Worner, S. P., & Gevrey, M. (2006). Modelling global insect pest species assemblages to determine risk of invasion. *Journal of Applied Ecology*, 43(5), 858-867.
- Wu, K. P., & Wang, S. D. (2009). Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space. *Pattern Recognition*, 42(5), 710-717.
- Wu, L., van den Hoogen, L. L., Slater, H., Walker, P. G., Ghani, A. C., Drakeley, C. J., & Okell, L. C. (2015). Comparison of diagnostics for the detection of asymptomatic *Plasmodium falciparum* infections to inform control and elimination strategies. *Nature*, 528(7580), S86.
- Xie, X., & Sun, S. (2012, January). Multitask twin support vector machines. In *Neural Information Processing* (pp. 341-348). Berlin Heidelberg, Springer.
- Xing, Y., Wu, X., & Xu, Z. (2008, June). Multiclass least squares auto-correlation wavelet support vector machines. In *Innovative Computing Information and Control, 2008.ICICIC'08. 3rd International Conference on* (pp. 85-85). IEEE.
- Yang, H., Huang, K., King, I., & Lyu, M. R. (2009). Localized support vector regression for time series prediction. *Neurocomputing*, 72(10), 2659-2669.
- Yang, H. M., M. D. L. D. G. Macoris, K. C. Galvani, M. T. M. Andrighetti, and D. M. V. Wanderley, (2009). "Assessing the effects of temperature on the

population of *Aedes aegypti*, the vector of dengue." *Epidemiology & Infection* 137, no. 8 (2009): 1188-1202.

Ye, N. (Ed.). (2003). *The handbook of data mining* (Vol. 24). Mahwah, NJ/London: Lawrence Erlbaum Associates, Publishers.

Yusuf, S. K. (2013). *Spatial Temporal Impacts of Climate Variability on Malaria Distribution in Gulu and Mpigi Districts in Uganda Using GIS* (Doctoral dissertation, University of Maiduguri, Nigeria).

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.

Zhou, G., Minakawa, N., Githeko, A. K., & Yan, G. (2004). Association between climate variability and malaria epidemics in the East African highlands. *Proceedings of the National Academy of Sciences of the United States of America*, 101(8), 2375-2380.

Zinszer, K., Kigozi, R., Charland, K., Dorsey, G., Kamya, M., & Buckeridge, D. (2013). Predicting Malaria in a Highly Endemic Country using Environmental and Clinical Data Sources. *Online Journal of Public Health Informatics*, 6(1).

Zinszer, K., Verma, A. D., Charland, K., Brewer, T. F., Brownstein, J. S., Sun, Z., & Buckeridge, D. L. (2013). A scoping review of malaria Predicting: past work and future directions. *BMJ Open*, 2(6), e001992.

APPENDIX

Appendix A:

Table 4.3: Threshold Values and Performance for the ANN Malaria Model Training

Threshold	Accuracy					Average Accuracy	Standard Deviation
0.2800	51.7857	50.5952	50.4762	51.0714	50.4762	50.8810	0.5622
0.2850	52.6190	50.4762	50.4762	50.4762	52.6190	51.3333	1.1737
0.2900	50.7143	51.3095	50.4762	50.4762	50.4762	50.6905	0.3611
0.2950	50.4762	51.1905	50.7143	50.5952	50.5952	50.7143	0.2792
0.3000	50.4762	50.4762	50.8333	52.6190	50.4762	50.9762	0.9313
0.3050	52.5000	50.7143	50.4762	50.7143	50.7143	51.0238	0.8316
0.3100	50.4762	50.4762	51.7857	51.0714	50.4762	50.8571	0.5796
0.3150	50.4762	50.4762	51.3095	50.4762	50.4762	50.6429	0.3727
0.3200	51.6667	52.7381	51.0714	50.4762	51.3095	51.4524	0.8393
0.3250	50.4762	50.4762	51.9048	50.9524	51.4286	51.0476	0.6209
0.3300	51.7857	51.4286	50.4762	50.4762	51.6667	51.1667	0.6433
0.3350	51.6667	50.4762	50.5952	50.4762	51.4286	50.9286	0.5734
0.3400	50.4762	50.4762	50.5952	50.4762	50.9524	50.5952	0.2062
0.3450	52.0238	51.3095	50.5952	50.4762	50.4762	50.9762	0.6808
0.3500	50.4762	51.6667	50.4762	51.4286	50.4762	50.9048	0.5929
0.3550	50.4762	51.3095	52.1429	50.4762	50.7143	51.0238	0.7123
0.3600	50.8333	50.4762	51.5476	50.4762	51.0714	50.8810	0.4502
0.3650	50.4762	52.2619	52.3810	52.7381	51.7857	51.9286	0.8805
0.3700	50.4762	51.7857	51.6667	50.9524	51.4286	51.2619	0.5429
0.3750	51.4286	50.4762	51.5476	51.4286	50.4762	51.0714	0.5455
0.3800	50.4762	52.3810	50.4762	51.5476	50.5952	51.0952	0.8477
0.3850	51.0714	50.4762	52.5000	51.0714	50.4762	51.1190	0.8274
0.3900	51.7857	52.5000	50.4762	51.6667	50.4762	51.3810	0.8853
0.3950	53.0952	50.9524	50.4762	51.5476	51.7857	51.5714	0.9939
0.4000	52.9762	50.3571	51.0714	50.4762	50.7143	51.1190	1.0734
0.4050	50.5952	50.4762	51.5476	51.4286	50.7143	50.9524	0.4980
0.4100	50.4762	51.0714	52.0238	50.5952	51.6667	51.1667	0.6703
0.4150	51.6667	51.3095	50.7143	50.1190	50.5952	50.8810	0.6105
0.4200	51.7857	52.7381	50.4762	50.4762	50.4762	51.1905	1.0344
0.4250	49.7619	52.5000	50.7143	51.9048	52.5000	51.4762	1.2041
0.4300	50.5952	50.9524	51.1905	50.9524	52.0238	51.1429	0.5364
0.4350	51.1905	52.3810	50.4762	50.5952	50.9524	51.1190	0.7604
0.4400	50.4762	51.6667	51.6667	53.5714	50.5952	51.5952	1.2418
0.4450	50.5952	50.4762	50.4762	50.4762	53.6905	51.1429	1.4251
0.4500	51.9048	50.9524	51.1905	51.6667	50.8333	51.3095	0.4611
0.4550	50.8333	53.0952	51.0714	50.4762	50.8333	51.2619	1.0467
0.4600	53.3333	50.4762	51.1905	50.4762	51.3095	51.3571	1.1713
0.4650	49.7619	52.2619	52.6190	50.4762	52.1429	51.4524	1.2554
0.4700	51.3095	52.2619	51.9048	52.9762	51.4286	51.9762	0.6766
0.4750	50.4762	52.3810	53.2143	52.2619	53.0952	52.2857	1.0956
0.4800	54.4048	53.3333	50.4762	50.8333	51.5476	52.1190	1.6865
0.4850	51.4286	50.7143	50.7143	52.5000	50.8333	51.2381	0.7651
0.4900	50.7143	50.8333	51.6667	50.9524	51.7857	51.1905	0.4980
0.4950	50.9524	50.9524	51.6667	53.9286	52.6190	52.0238	1.2655
0.5000	50.9524	52.3810	52.3810	51.7857	51.7857	51.8571	0.5868
0.5050	52.6190	50.8333	50.4762	54.0476	50.7143	51.7381	1.5467
0.5100	51.3095	53.9286	50.5952	50.9524	51.6667	51.6905	1.3133
0.5150	51.7857	50.4762	52.2619	53.4524	51.4286	51.8810	1.0956
0.5200	50.4762	51.9048	50.4762	50.1190	52.1429	51.0238	0.9283
0.5250	52.5000	51.7857	52.6190	50.0000	52.1429	51.8095	1.0628
0.5300	52.2619	50.9524	52.1429	51.7857	52.3810	51.9048	0.5771
0.5350	52.7381	50.9524	51.5476	49.5238	51.4286	51.2381	1.1622
0.5400	52.0238	52.1429	52.3810	52.7381	50.5952	51.9762	0.8187
0.5450	50.8333	49.8810	52.6190	51.4286	51.9048	51.3333	1.0426
0.5500	52.5000	52.3810	50.0000	53.8095	52.8571	52.3095	1.4076
0.5550	50.8333	54.0476	50.9524	52.0238	50.8333	51.7381	1.3847
0.5600	51.5476	51.4286	52.6190	49.8810	52.0238	51.5000	1.0192
0.5650	50.4762	51.6667	51.7857	50.5952	51.4286	51.1905	0.6128
0.5700	50.5952	51.4286	50.9524	50.8333	51.5476	51.0714	0.4037
0.5750	51.5476	52.9762	50.8333	50.7143	51.9048	51.5952	0.9167
0.5800	52.3810	52.0238	49.1667	50.4762	54.8810	51.7857	2.1544
0.5850	50.7143	51.5476	51.7857	54.0476	51.9048	52.0000	1.2355
0.5900	51.0714	49.6429	51.6667	51.7857	50.7143	50.9762	0.8642
0.5950	50.8333	51.1905	50.7143	52.7381	50.4762	51.1905	0.9027

0.6000	52.7381	51.1905	50.2381	50.9524	51.0714	51.2381	0.9167
0.6050	52.5000	53.0952	50.8333	49.2857	52.7381	51.6905	1.6003
0.6100	50.3571	52.2619	50.0000	50.3571	50.1190	50.6190	0.9313
0.6150	51.5476	52.2619	50.4762	53.2143	49.1667	51.3333	1.5712
0.6200	51.7857	51.5476	48.2143	52.6190	52.0238	51.2381	1.7366
0.6250	49.7619	53.4524	51.9048	52.2619	50.8333	51.6429	1.4071
0.6300	48.0952	52.0238	52.1429	49.5238	50.9524	50.5476	1.7284
0.6350	50.4762	50.3571	49.7619	49.5238	52.8571	50.5952	1.3257
0.6400	50.0000	49.7619	49.5238	49.5238	52.5000	50.2619	1.2666
0.6450	50.7143	49.6429	49.4048	52.8571	48.9286	50.3095	1.5672
0.6500	50.5952	50.9524	50.4762	49.7619	50.2381	50.4048	0.4422
0.6550	50.5952	50.8333	50.9524	49.5238	49.4048	50.2619	0.7406
0.6600	51.3095	51.1905	52.3810	50.4762	50.5952	51.1905	0.7576
0.6650	50.2381	51.9048	53.5714	49.0476	50.0000	50.9524	1.7897
0.6700	50.2381	50.2381	51.0714	50.2381	50.1190	50.3810	0.3894
0.6750	51.5476	49.2857	49.8810	48.4524	51.6667	50.1667	1.4101
0.6800	50.8333	47.5000	49.6429	50.5952	50.5952	49.8333	1.3822
0.6850	51.5476	50.5952	50.3571	49.1667	49.5238	50.2381	0.9374
0.6900	51.0714	48.5714	49.5238	46.9048	50.1190	49.2381	1.5896
0.6950	50.2381	51.7857	50.4762	49.0476	47.9762	49.9048	1.4517
0.7000	49.4048	50.7143	48.4524	50.0000	49.8810	49.6905	0.8359
0.7050	49.7619	50.9524	47.9762	50.5952	50.3571	49.9286	1.1743
0.7100	51.6667	50.0000	49.6429	47.7381	48.3333	49.4762	1.5352
0.7150	49.0476	49.1667	49.7619	46.7857	49.7619	48.9048	1.2297
0.7200	50.0000	50.7143	48.0952	50.8333	49.6429	49.8571	1.1021
0.7250	44.4048	49.2857	54.2857	49.7619	48.8095	49.3095	3.5098
0.7300	47.3810	47.5000	49.8810	46.0714	47.7381	47.7143	1.3740
0.7350	50.7143	46.9048	48.2143	48.4524	49.0476	48.6667	1.3868

Table 4.5: Threshold Values and Performance for the ANN Malaria model Testing

Threshold	Accuracy						Average Accuracy	Standard Deviation
0.2800	42.2222	42.2222	42.7778	44.4444	43.8889	43.1111	1.0092	
0.2850	41.6667	45.0000	45.0000	43.3333	45.0000	44.0000	1.4907	
0.2900	44.4444	42.7778	44.4444	44.4444	44.4444	44.1111	0.7454	
0.2950	45.0000	44.4444	43.3333	44.4444	44.4444	44.3333	0.6086	
0.3000	42.7778	45.0000	45.0000	45.0000	45.0000	44.5556	0.9938	
0.3050	45.0000	43.8889	45.0000	43.3333	43.8889	44.2222	0.7454	
0.3100	42.7778	45.0000	45.0000	42.2222	45.0000	44.0000	1.3833	
0.3150	45.0000	43.8889	45.0000	43.8889	43.8889	44.3333	0.6086	
0.3200	43.8889	44.4444	45.0000	44.4444	43.3333	44.2222	0.6334	
0.3250	44.4444	42.2222	45.0000	45.0000	41.1111	43.5556	1.7830	
0.3300	42.2222	44.4444	44.4444	45.0000	45.0000	44.2222	1.1520	
0.3350	45.5556	45.0000	45.0000	42.2222	42.7778	44.1111	1.5010	
0.3400	41.6667	45.0000	45.0000	45.0000	43.3333	44.0000	1.4907	
0.3450	43.8889	45.5556	44.4444	45.0000	45.0000	44.7778	0.6334	
0.3500	41.6667	43.8889	43.8889	42.2222	45.0000	43.3333	1.3608	
0.3550	42.7778	43.3333	43.8889	45.0000	41.1111	43.2222	1.4380	
0.3600	44.4444	43.8889	45.0000	43.8889	43.8889	44.2222	0.4969	
0.3650	42.7778	45.0000	45.0000	45.0000	44.4444	44.4444	0.9623	
0.3700	45.0000	45.0000	45.0000	45.0000	44.4444	44.8889	0.2485	
0.3750	43.8889	42.2222	44.4444	45.0000	45.0000	44.1111	1.1520	
0.3800	45.0000	44.4444	43.3333	44.4444	45.0000	44.4444	0.6804	
0.3850	44.4444	43.3333	45.0000	45.5556	43.8889	44.4444	0.8784	
0.3900	43.3333	45.0000	42.2222	44.4444	45.0000	44.0000	1.2044	
0.3950	44.4444	45.0000	44.4444	42.2222	45.0000	44.2222	1.1520	
0.4000	43.8889	40.5556	45.0000	45.0000	43.8889	43.6667	1.8257	
0.4050	45.0000	45.0000	45.0000	43.8889	44.4444	44.6667	0.4969	
0.4100	45.0000	42.2222	42.7778	45.5556	44.4444	44.0000	1.4380	
0.4150	44.4444	41.1111	44.4444	45.0000	45.0000	44.0000	1.6387	
0.4200	44.4444	42.2222	44.4444	45.0000	45.0000	44.2222	1.1520	
0.4250	40.5556	42.7778	44.4444	40.5556	43.8889	42.4444	1.8257	
0.4300	42.7778	45.0000	45.0000	43.8889	45.0000	44.3333	0.9938	
0.4350	42.2222	43.8889	44.4444	43.8889	44.4444	43.7778	0.9129	
0.4400	45.0000	44.4444	44.4444	43.3333	45.0000	44.4444	0.6804	
0.4450	42.2222	44.4444	45.0000	43.8889	45.0000	44.1111	1.1520	
0.4500	44.4444	43.8889	45.0000	42.7778	42.7778	43.7778	0.9938	
0.4550	44.4444	43.8889	43.3333	43.8889	45.0000	44.1111	0.6334	
0.4600	44.4444	42.2222	42.7778	42.2222	45.0000	43.3333	1.3029	
0.4650	45.0000	42.2222	44.4444	43.3333	45.0000	44.0000	1.2044	
0.4700	45.0000	45.0000	40.5556	43.3333	45.0000	43.7778	1.9405	
0.4750	41.6667	42.2222	43.8889	43.8889	42.2222	42.7778	1.0393	
0.4800	42.7778	41.6667	45.5556	42.2222	44.4444	43.3333	1.6197	
0.4850	45.0000	43.8889	45.0000	44.4444	42.2222	44.1111	1.1520	
0.4900	42.2222	41.6667	43.8889	45.0000	43.3333	43.2222	1.3264	
0.4950	45.0000	44.4444	42.7778	41.6667	43.3333	43.4444	1.3264	
0.5000	43.3333	45.0000	42.2222	45.0000	43.3333	43.7778	1.2044	
0.5050	43.8889	45.0000	45.0000	41.6667	43.8889	43.8889	1.3608	
0.5100	43.3333	43.8889	41.6667	42.7778	43.8889	43.1111	0.9296	
0.5150	43.8889	45.0000	43.3333	45.0000	43.3333	44.1111	0.8425	
0.5200	43.8889	40.5556	43.8889	42.2222	42.7778	42.6667	1.3833	
0.5250	45.0000	43.3333	45.5556	44.4444	44.4444	44.5556	0.8240	
0.5300	45.0000	43.8889	41.1111	43.3333	40.5556	42.7778	1.8840	
0.5350	42.7778	42.7778	41.6667	42.2222	41.6667	42.2222	0.5556	
0.5400	44.4444	43.8889	42.2222	46.1111	40.0000	43.3333	2.3241	
0.5450	45.0000	42.7778	42.2222	45.0000	42.2222	43.4444	1.4380	
0.5500	42.2222	40.5556	40.0000	44.4444	41.6667	41.7778	1.7303	
0.5550	43.3333	42.7778	41.6667	40.0000	42.7778	42.1111	1.3264	
0.5600	42.2222	45.0000	43.8889	45.0000	44.4444	44.1111	1.1520	
0.5650	40.0000	40.5556	41.1111	43.8889	43.3333	41.7778	1.7303	
0.5700	45.0000	41.6667	42.7778	44.4444	45.0000	43.7778	1.4907	
0.5750	41.6667	42.2222	42.2222	45.0000	45.0000	43.2222	1.6387	
0.5800	44.4444	41.1111	41.6667	41.6667	41.6667	42.1111	1.3264	
0.5850	40.5556	43.8889	43.3333	41.6667	42.2222	42.3333	1.3264	
0.5900	42.2222	43.3333	40.5556	42.7778	41.6667	42.1111	1.0686	
0.5950	43.8889	44.4444	43.3333	42.2222	38.8889	42.5556	2.2083	
0.6000	43.3333	44.4444	40.5556	43.8889	40.5556	42.5556	1.8675	
0.6050	41.6667	38.3333	40.0000	40.0000	41.6667	40.3333	1.3944	
0.6100	39.4444	45.0000	42.7778	43.3333	42.7778	42.6667	2.0184	
0.6150	40.0000	44.4444	45.0000	44.4444	42.7778	43.3333	2.0412	

0.6200	43.8889	41.1111	41.1111	41.1111	43.3333	42.1111	1.3833
0.6250	38.8889	45.0000	43.8889	43.3333	43.8889	43.0000	2.3766
0.6300	43.3333	41.6667	41.1111	41.1111	40.0000	41.4444	1.2172
0.6350	44.4444	44.4444	39.4444	43.8889	42.2222	42.8889	2.1300
0.6400	41.6667	40.5556	41.1111	40.0000	44.4444	41.5556	1.7303
0.6450	42.2222	42.2222	41.1111	39.4444	40.5556	41.1111	1.1785
0.6500	45.0000	43.3333	42.2222	43.8889	41.1111	43.1111	1.5010
0.6550	45.0000	39.4444	43.8889	41.6667	43.3333	42.6667	2.1660
0.6600	42.7778	42.2222	43.3333	40.0000	45.0000	42.6667	1.8173
0.6650	45.5556	37.7778	42.2222	42.7778	38.8889	41.4444	3.1329
0.6700	41.6667	42.2222	40.5556	43.3333	38.8889	41.3333	1.6942
0.6750	40.5556	43.8889	42.2222	38.3333	38.8889	40.7778	2.3107
0.6800	40.0000	40.0000	42.2222	39.4444	45.0000	41.3333	2.3107
0.6850	39.4444	38.8889	41.1111	41.1111	40.0000	40.1111	0.9938
0.6900	42.7778	43.8889	41.6667	41.1111	42.7778	42.4444	1.0830
0.6950	40.0000	44.4444	40.5556	36.1111	41.6667	40.5556	3.0174
0.7000	40.5556	38.3333	42.2222	43.3333	42.2222	41.3333	1.9484
0.7050	45.0000	38.8889	40.5556	38.3333	42.2222	41.0000	2.7046
0.7100	39.4444	42.7778	36.1111	39.4444	42.7778	40.1111	2.7889
0.7150	42.2222	40.5556	39.4444	43.8889	42.7778	41.7778	1.7743
0.7200	42.2222	39.4444	41.1111	39.4444	40.0000	40.4444	1.2044
0.7250	43.3333	40.0000	43.3333	39.4444	39.4444	41.1111	2.0412
0.7300	44.4444	38.3333	40.0000	43.8889	42.2222	41.7778	2.5880
0.7350	35.0000	43.3333	40.0000	39.4444	41.6667	39.8889	3.1279

Table 4.7: Threshold Values and Performance for the developed ANN Malaria Model Validation

Threshold	Accuracy					Average Accuracy	Standard Deviation
0.2800	45.5556	45.0000	44.4444	45.0000	45.0000	45.0000	0.3928
0.2850	45.0000	45.0000	45.0000	45.0000	45.0000	45.0000	0.0000
0.2900	45.0000	45.0000	44.4444	44.4444	44.4444	44.6667	0.3043
0.2950	45.5556	45.0000	45.5556	46.1111	45.0000	45.4444	0.4648
0.3000	45.0000	45.5556	45.5556	45.0000	45.0000	45.2222	0.3043
0.3050	45.0000	45.0000	45.0000	45.0000	45.0000	45.0000	0.0000
0.3100	45.5556	45.0000	45.0000	45.5556	45.0000	45.2222	0.3043
0.3150	44.4444	45.0000	45.0000	46.1111	45.0000	45.1111	0.6086
0.3200	45.0000	46.6667	45.0000	45.0000	44.4444	45.2222	0.8425
0.3250	44.4444	45.0000	45.0000	44.4444	45.0000	44.7778	0.3043
0.3300	45.5556	44.4444	44.4444	44.4444	45.0000	44.7778	0.4969
0.3350	44.4444	44.4444	45.0000	47.2222	44.4444	45.1111	1.2044
0.3400	46.1111	45.0000	45.0000	45.0000	45.0000	45.2222	0.4969
0.3450	43.8889	44.4444	45.0000	45.0000	43.3333	44.3333	0.7244
0.3500	44.4444	44.4444	45.0000	44.4444	45.0000	44.6667	0.3043
0.3550	45.5556	45.0000	45.0000	44.4444	45.0000	45.0000	0.3928
0.3600	45.0000	45.0000	44.4444	46.1111	45.0000	45.1111	0.6086
0.3650	43.8889	45.5556	45.0000	45.0000	44.4444	44.7778	0.6334
0.3700	45.0000	44.4444	45.5556	45.0000	43.8889	44.7778	0.6334
0.3750	47.2222	44.4444	45.0000	45.5556	45.5556	45.5556	1.0393
0.3800	45.0000	45.0000	44.4444	45.0000	46.1111	45.1111	0.6086
0.3850	45.5556	45.0000	45.0000	45.0000	45.0000	45.1111	0.2485
0.3900	46.1111	46.6667	45.0000	45.0000	44.4444	45.4444	0.9129
0.3950	46.1111	43.8889	45.5556	45.5556	45.5556	45.3333	0.8425
0.4000	45.0000	46.1111	46.6667	46.1111	45.0000	45.7778	0.7454
0.4050	45.5556	45.0000	45.0000	44.4444	45.5556	45.1111	0.4648
0.4100	45.0000	45.0000	45.0000	44.4444	45.5556	45.0000	0.3928
0.4150	45.5556	45.5556	45.0000	45.0000	44.4444	45.1111	0.4648
0.4200	45.0000	45.0000	43.8889	45.5556	45.5556	45.0000	0.6804
0.4250	43.3333	45.5556	43.8889	44.4444	45.5556	44.5556	0.9938
0.4300	44.4444	43.8889	46.6667	45.0000	44.4444	44.8889	1.0686
0.4350	43.8889	42.2222	45.0000	45.5556	45.5556	44.4444	1.4164
0.4400	43.8889	45.0000	43.3333	46.1111	45.0000	44.6667	1.0830
0.4450	45.0000	45.5556	45.0000	45.0000	45.0000	45.1111	0.2485
0.4500	45.5556	45.0000	45.0000	43.3333	45.0000	44.7778	0.8425
0.4550	43.8889	45.0000	44.4444	44.4444	45.0000	44.5556	0.4648
0.4600	45.0000	44.4444	46.1111	46.1111	45.5556	45.4444	0.7244
0.4650	46.1111	45.0000	45.0000	44.4444	44.4444	45.0000	0.6804
0.4700	44.4444	43.8889	45.0000	46.1111	45.0000	44.8889	0.8240
0.4750	45.0000	46.1111	44.4444	44.4444	43.3333	44.6667	1.0092
0.4800	45.0000	45.0000	45.0000	45.0000	44.4444	44.8889	0.2485
0.4850	44.4444	44.4444	45.0000	45.0000	45.0000	44.7778	0.3043
0.4900	45.0000	44.4444	47.7778	45.0000	45.0000	45.4444	1.3264
0.4950	45.0000	45.0000	45.0000	45.0000	45.5556	45.1111	0.2485
0.5000	43.8889	45.0000	44.4444	45.0000	45.5556	44.7778	0.6334
0.5050	44.4444	43.8889	43.8889	45.0000	45.0000	44.4444	0.5556
0.5100	43.8889	43.3333	45.5556	45.5556	44.4444	44.5556	0.9938
0.5150	45.0000	44.4444	44.4444	47.7778	44.4444	45.2222	1.4487
0.5200	43.8889	45.0000	45.0000	47.2222	44.4444	45.1111	1.2669
0.5250	45.0000	45.0000	44.4444	45.0000	44.4444	44.7778	0.3043
0.5300	43.3333	44.4444	45.0000	45.5556	45.5556	44.7778	0.9296
0.5350	45.0000	46.6667	43.8889	45.5556	45.0000	45.2222	1.0092
0.5400	43.3333	46.1111	43.8889	45.0000	45.0000	44.6667	1.0830
0.5450	43.8889	43.8889	43.8889	44.4444	45.0000	44.2222	0.4969
0.5500	45.0000	45.5556	45.5556	45.0000	45.5556	45.3333	0.3043
0.5550	46.6667	47.2222	43.3333	43.8889	46.6667	45.5556	1.8002
0.5600	45.0000	45.0000	46.1111	44.4444	44.4444	45.0000	0.6804
0.5650	48.8889	43.8889	44.4444	45.0000	47.7778	46.0000	2.2013
0.5700	46.1111	46.1111	45.0000	45.0000	45.0000	45.4444	0.6086
0.5750	45.0000	45.0000	45.5556	43.3333	45.0000	44.7778	0.8425
0.5800	44.4444	45.0000	43.8889	45.5556	46.6667	45.1111	1.0686
0.5850	43.8889	44.4444	46.1111	45.0000	43.8889	44.6667	0.9296
0.5900	46.1111	43.8889	45.0000	44.4444	45.0000	44.8889	0.8240
0.5950	45.5556	45.0000	41.1111	45.5556	44.4444	44.3333	1.8592
0.6000	48.3333	46.1111	45.0000	45.0000	46.6667	46.2222	1.3833
0.6050	45.0000	45.0000	43.8889	45.5556	45.0000	44.8889	0.6086
0.6100	45.0000	41.1111	41.6667	44.4444	46.1111	43.6667	2.1731
0.6150	46.1111	42.2222	46.6667	46.6667	42.7778	44.8889	2.2013

0.6200	45.0000	45.0000	45.0000	45.5556	45.0000	45.1111	0.2485
0.6250	44.4444	45.0000	43.3333	43.3333	42.7778	43.7778	0.9129
0.6300	46.1111	45.0000	44.4444	43.3333	42.2222	44.2222	1.5010
0.6350	44.4444	42.7778	44.4444	42.2222	43.8889	43.5556	1.0092
0.6400	40.5556	41.1111	43.3333	45.5556	42.2222	42.5556	1.9876
0.6450	43.8889	45.0000	44.4444	45.5556	46.6667	45.1111	1.0686
0.6500	45.0000	43.8889	41.1111	46.6667	45.0000	44.3333	2.0563
0.6550	46.1111	45.0000	43.3333	43.3333	42.7778	44.1111	1.3944
0.6600	43.8889	43.8889	44.4444	43.3333	44.4444	44.0000	0.4648
0.6650	42.2222	45.5556	43.8889	43.3333	45.5556	44.1111	1.4487
0.6700	43.3333	43.3333	42.7778	41.6667	40.5556	42.3333	1.2044
0.6750	44.4444	44.4444	45.0000	41.6667	43.8889	43.8889	1.3029
0.6800	43.3333	41.6667	46.6667	44.4444	42.2222	43.6667	1.9876
0.6850	43.3333	43.3333	43.3333	42.7778	41.6667	42.8889	0.7244
0.6900	41.6667	44.4444	43.3333	43.3333	42.2222	43.0000	1.0830
0.6950	43.8889	45.5556	41.1111	43.8889	45.5556	44.0000	1.8173
0.7000	42.7778	42.2222	43.8889	46.1111	42.2222	43.4444	1.6387
0.7050	41.6667	44.4444	44.4444	47.2222	41.6667	43.8889	2.3241
0.7100	45.5556	43.8889	44.4444	45.5556	43.3333	44.5556	0.9938
0.7150	42.2222	43.3333	42.7778	44.4444	43.8889	43.3333	0.8784
0.7200	40.0000	40.5556	42.2222	40.0000	43.8889	41.3333	1.6942
0.7250	43.3333	47.2222	45.5556	43.8889	44.4444	44.8889	1.5416
0.7300	44.4444	44.4444	41.1111	42.7778	43.3333	43.2222	1.3833
0.7350	41.6667	43.8889	41.1111	41.6667	45.5556	42.7778	1.8840

Appendix B:

Table 4.10: ANN Target and Simulated Malaria Test Results

S/No	ANN_0' Malaria cases Result		ANN_1' Malaria cases Result		ANN_2' Malaria cases Result	
	Target Value	Sim_ Result	Target Value	Sim_ Result	Target Value	Sim_ Result
1.	0	False	0	False	0	True
2.	0	True	0	True	0	True
3.	0	False	0	True	0	True
4.	0	False	0	False	0	False
5.	0	True	0	True	0	True
6.	0	False	0	True	0	False
7.	2	True	2	True	2	True
8.	2	True	2	True	2	True
9.	2	True	2	True	2	True
10.	2	True	2	True	2	True
11.	2	True	2	True	2	True
12.	2	True	2	True	2	True
13.	2	True	2	True	2	True
14.	2	True	2	True	2	True
15.	2	True	2	True	2	True
16.	2	True	2	True	2	True
17.	2	True	2	True	2	True
18.	2	True	2	True	2	True
19.	1	True	1	True	1	True
20.	1	False	1	True	1	False
21.	1	True	1	True	1	True
22.	1	True	1	True	1	True
23.	1	True	1	False	1	False
24.	1	True	1	True	1	True
25.	1	True	1	True	1	True
26.	1	True	1	True	1	True
27.	1	True	1	True	1	True
28.	1	True	1	True	1	True
29.	1	True	1	True	1	True
30.	1	True	1	True	1	True
31.	1	True	1	True	1	True
32.	1	True	1	True	1	True
33.	2	True	2	True	2	True
34.	2	True	2	True	2	True
35.	2	False	2	True	2	True
36.	2	True	2	True	2	True
37.	2	True	2	True	2	True
38.	2	True	2	True	2	True
39.	2	True	2	True	2	True
40.	2	True	2	True	2	True
41.	2	True	2	True	2	True
42.	2	True	2	True	2	True
43.	2	True	2	True	2	True
44.	2	True	2	True	2	True
45.	1	False	1	True	1	True
46.	1	True	1	True	1	True
47.	1	True	1	True	1	True
48.	1	True	1	True	1	True
49.	1	True	1	True	1	True
50.	1	True	1	True	1	True
51.	1	True	1	True	1	True
52.	1	True	1	True	1	True
53.	1	True	1	True	1	True
54.	1	True	1	True	1	False
55.	1	True	1	True	1	False
56.	2	True	2	True	2	True
57.	0	False	0	False	0	False
58.	0	True	0	True	0	True
59.	0	False	0	False	0	False
60.	0	True	0	True	0	True
61.	1	True	1	True	1	True

62.	1	True	1	True	1	False
63.	1	True	1	True	1	True
64.	1	True	1	True	1	True
65.	1	True	1	False	1	True
66.	1	True	1	True	1	True
67.	1	True	1	True	1	True
68.	2	True	2	True	2	True
69.	2	True	2	True	2	True
70.	2	True	2	True	2	True
71.	2	True	2	True	2	True
72.	2	True	2	True	2	True
73.	2	True	2	True	2	True
74.	2	True	2	True	2	True
75.	2	True	2	True	2	True
76.	0	False	0	False	0	True
77.	0	False	0	False	0	True
78.	0	True	0	True	0	True
79.	0	False	0	True	0	True
80.	2	True	2	True	2	True
81.	2	True	2	True	2	True
82.	2	True	2	True	2	True
83.	2	True	2	True	2	True
84.	2	True	2	True	2	True
85.	2	True	2	True	2	True
86.	1	True	1	True	1	True
87.	1	True	1	True	1	True
88.	1	True	1	True	1	True
89.	1	True	1	True	1	True
90.	1	True	1	True	1	True
91.	1	True	1	True	1	False
92.	1	True	1	True	1	True
93.	2	True	2	True	2	True
94.	0	False	0	False	0	False
95.	0	True	0	True	0	True
96.	0	False	0	False	0	False
97.	0	True	0	True	0	True
98.	1	False	1	True	1	True
99.	1	True	1	True	1	True
100.	1	True	1	True	1	True
101.	1	True	1	True	1	True
102.	0	False	0	False	0	False
103.	0	True	0	True	0	True
104.	0	False	0	False	0	False
105.	0	True	0	True	0	True
106.	0	False	0	False	0	False
107.	1	True	1	True	1	True
108.	1	True	1	True	1	True
109.	1	True	1	True	1	True
110.	1	True	1	True	1	True
111.	1	True	1	True	1	True
112.	1	True	1	True	1	True
113.	1	True	1	True	1	True
114.	1	True	1	True	1	True
115.	2	True	2	True	2	True
116.	2	True	2	True	2	True
117.	2	True	2	True	2	True
118.	2	True	2	True	2	True
119.	2	True	2	True	2	True
120.	2	True	2	True	2	True
121.	0	False	0	False	0	False
122.	0	True	0	True	0	True
123.	0	True	0	True	0	True
124.	2	True	2	True	2	True
125.	2	True	2	True	2	True
126.	2	True	2	False	2	True
127.	2	True	2	True	2	True
128.	1	True	1	True	1	True
129.	1	False	1	true	1	False
130.	1	True	1	true	1	True
131.	1	True	1	true	1	True
132.	1	True	1	false	1	False
133.	1	True	1	true	1	True
134.	2	True	2	true	2	True
135.	2	True	2	true	2	True
136.	2	True	2	true	2	True

137.	2	True	2	true	2	True
138.	2	True	2	true	2	True
139.	2	True	2	true	2	True
140.	2	True	2	true	2	True
141.	2	True	2	true	2	True
142.	2	True	2	true	2	True
143.	0	True	0	true	0	True
144.	0	False	0	false	0	True
145.	0	True	0	false	0	False
146.	0	False	0	false	0	False
147.	0	False	0	true	0	False
148.	0	False	0	false	0	False
149.	0	False	0	false	0	False
150.	1	False	1	false	1	True
151.	1	True	1	true	1	True
152.	1	False	1	false	1	True
153.	1	True	1	true	1	True
154.	1	False	1	true	1	False
155.	2	True	2	true	2	True
156.	2	False	2	false	2	False
157.	2	True	2	true	2	True
158.	2	True	2	true	2	True
159.	2	True	2	true	2	True
160.	2	True	2	true	2	True
161.	1	True	1	false	1	True
162.	1	True	1	true	1	True
163.	1	True	1	true	1	True
164.	1	False	1	false	1	True
165.	1	True	1	true	1	True
166.	1	True	1	true	1	True
167.	1	True	1	false	1	True
168.	1	True	1	true	1	True
169.	1	False	1	false	1	False
170.	2	True	2	true	2	True
171.	2	True	2	true	2	True
172.	2	True	2	true	2	True
173.	2	True	2	true	2	True
174.	1	True	1	true	1	True
175.	1	True	1	true	1	True
176.	1	False	1	false	1	True
177.	1	True	1	true	1	True
178.	1	False	1	true	1	True
179.	1	True	1	true	1	True
180.	1	True	1	true	1	True

Appendix C:

Table 4.15: SVM Target and Simulated Test Result

S/No	SVM_2' Malaria cases Result (lin)		SVM_2' Malaria cases Result (rbf)		SVM_2' Malaria cases Result(pol)	
	Target Value	Sim_ Result	Target Value	Sim_ Result	Target Value	Sim_ Result
1.	-1	-1	-1	-1	-1	-1
4.	-1	-1	-1	-1	-1	-1
5.	-1	1	-1	-1	-1	-1
6.	-1	-1	-1	-1	-1	-1
7.	-1	-1	-1	-1	-1	-1
8.	-1	-1	-1	-1	-1	-1
9.	1	1	1	1	1	1
10.	1	1	1	1	1	1
11.	1	-1	1	-1	1	1
12.	1	1	1	1	1	1
13.	1	1	1	1	1	1
14.	1	1	1	1	1	1
15.	1	1	1	1	1	1
16.	1	-1	1	1	1	1
17.	1	1	1	1	1	1
18.	1	1	1	1	1	1
19.	1	-1	1	1	1	1
20.	1	1	1	1	1	1
21.	-1	-1	-1	-1	-1	-1
22.	-1	-1	-1	-1	-1	-1
23.	-1	-1	-1	-1	-1	-1
24.	-1	-1	-1	-1	-1	-1
25.	-1	-1	-1	-1	-1	-1
26.	-1	-1	-1	-1	-1	-1
27.	-1	1	-1	-1	-1	-1
28.	-1	-1	-1	-1	-1	-1
29.	-1	-1	-1	-1	-1	-1
30.	-1	1	-1	1	-1	1
31.	-1	-1	-1	-1	-1	-1
32.	-1	-1	-1	-1	-1	-1
33.	-1	1	-1	1	-1	-1
34.	-1	-1	-1	1	-1	-1
35.	1	1	1	1	1	1
36.	1	1	1	1	1	1
37.	1	1	1	-1	1	-1
38.	1	1	1	1	1	1
39.	1	1	1	1	1	1
40.	1	1	1	1	1	1
41.	1	-1	1	1	1	1
42.	1	1	1	1	1	1
43.	1	1	1	1	1	1
44.	1	1	1	1	1	1
45.	1	-1	1	1	1	1
46.	1	1	1	1	1	1
47.	-1	-1	-1	-1	-1	-1
48.	-1	-1	-1	-1	-1	-1
49.	-1	-1	-1	-1	-1	-1
50.	-1	-1	-1	-1	-1	-1
51.	-1	-1	-1	-1	-1	-1
52.	-1	-1	-1	-1	-1	-1
53.	-1	-1	-1	-1	-1	-1
54.	-1	1	-1	1	-1	-1
55.	-1	-1	-1	-1	-1	-1
56.	-1	-1	-1	-1	-1	-1
57.	-1	-1	-1	-1	-1	-1
58.	1	1	1	1	1	1
59.	-1	-1	-1	-1	-1	-1
60.	-1	1	-1	-1	-1	-1
61.	-1	-1	-1	-1	-1	-1
62.	-1	-1	-1	-1	-1	1
63.	-1	1	-1	1	-1	-1
64.	-1	-1	-1	-1	-1	-1

65.	-1	-1	-1	-1	-1	-1
66.	-1	1	-1	1	-1	1
67.	-1	-1	-1	-1	-1	-1
68.	-1	-1	-1	-1	-1	-1
69.	-1	-1	-1	-1	-1	-1
70.	1	-1	1	-1	1	-1
71.	1	1	1	1	1	1
72.	1	1	1	1	1	1
73.	1	-1	1	1	1	1
74.	1	1	1	-1	1	1
75.	1	1	1	1	1	1
76.	1	-1	1	1	1	1
77.	1	1	1	1	1	1
78.	-1	-1	-1	-1	-1	-1
79.	-1	-1	-1	-1	-1	-1
80.	-1	-1	-1	-1	-1	-1
81.	-1	1	-1	-1	-1	-1
82.	1	-1	1	-1	1	-1
83.	1	1	1	1	1	1
84.	1	1	1	1	1	1
85.	1	1	1	1	1	1
86.	1	-1	1	1	1	1
87.	1	1	1	1	1	1
88.	-1	-1	-1	-1	-1	-1
89.	-1	-1	-1	-1	-1	-1
90.	-1	-1	-1	-1	-1	-1
91.	-1	1	-1	1	-1	1
92.	-1	-1	-1	-1	-1	-1
93.	-1	-1	-1	-1	-1	-1
94.	-1	1	-1	1	-1	1
95.	1	1	1	1	1	1
96.	-1	-1	-1	-1	-1	-1
97.	-1	1	-1	-1	-1	-1
98.	-1	-1	-1	-1	-1	-1
99.	-1	-1	-1	-1	-1	1
100.	-1	-1	-1	-1	-1	-1
101.	-1	-1	-1	-1	-1	-1
102.	-1	-1	-1	1	-1	1
103.	-1	-1	-1	-1	-1	-1
104.	-1	-1	-1	-1	-1	-1
105.	-1	1	-1	1	-1	1
106.	-1	-1	-1	-1	-1	-1
107.	-1	-1	-1	-1	-1	1
108.	-1	-1	-1	-1	-1	-1
109.	-1	1	-1	-1	-1	-1
110.	-1	-1	-1	-1	-1	-1
111.	-1	-1	-1	-1	-1	-1
112.	-1	1	-1	1	-1	1
113.	-1	-1	-1	-1	-1	-1
114.	-1	-1	-1	-1	-1	-1
115.	-1	1	-1	1	-1	-1
116.	-1	-1	-1	1	-1	-1
117.	1	1	1	1	1	1
118.	1	1	1	1	1	1
119.	1	1	1	-1	1	1
120.	1	1	1	1	1	1
121.	1	1	1	1	1	1
122.	1	1	1	1	1	1
123.	-1	-1	-1	-1	-1	-1
124.	-1	-1	-1	-1	-1	-1
125.	-1	-1	-1	-1	-1	1
126.	1	1	1	1	1	1
127.	1	1	1	1	1	1
128.	1	-1	1	-1	1	-1
129.	1	1	1	1	1	1
130.	-1	-1	-1	-1	-1	-1
131.	-1	-1	-1	-1	-1	-1
132.	-1	-1	-1	-1	-1	-1
133.	-1	-1	-1	-1	-1	-1
134.	-1	-1	-1	-1	-1	-1
135.	-1	-1	-1	-1	-1	-1
136.	1	-1	1	1	1	1
137.	1	1	1	1	1	1
138.	1	-1	1	1	1	1
139.	1	1	1	1	1	1

140.	1	1	1	1	1	1
141.	1	1	1	1	1	1
142.	1	-1	1	1	1	1
143.	1	1	1	1	1	1
144.	1	-1	1	-1	1	1
145.	-1	-1	-1	-1	-1	-1
146.	-1	-1	-1	-1	-1	-1
147.	-1	-1	-1	-1	-1	-1
148.	-1	-1	-1	-1	-1	-1
149.	-1	-1	-1	-1	-1	-1
150.	-1	-1	-1	-1	-1	-1
151.	-1	-1	-1	-1	-1	-1
152.	-1	-1	-1	-1	-1	-1
153.	-1	-1	-1	-1	-1	-1
154.	-1	-1	-1	-1	-1	-1
155.	-1	-1	-1	-1	-1	1
156.	-1	-1	-1	-1	-1	-1
157.	1	1	1	1	1	1
158.	1	-1	1	-1	1	-1
159.	1	1	1	1	1	1
160.	1	1	1	-1	1	-1
161.	1	-1	1	-1	1	1
162.	1	1	1	1	1	1
163.	-1	-1	-1	-1	-1	-1
164.	-1	1	-1	-1	-1	-1
165.	-1	-1	-1	-1	-1	-1
166.	-1	-1	-1	-1	-1	-1
167.	-1	1	-1	1	-1	1
168.	-1	-1	-1	-1	-1	-1
169.	-1	-1	-1	-1	-1	-1
170.	-1	1	-1	1	-1	1
171.	-1	-1	-1	-1	-1	-1
172.	1	1	1	1	1	1
173.	1	1	1	1	1	1
174.	1	1	1	1	1	1
175.	1	1	1	1	1	1
176.	-1	-1	-1	-1	-1	-1
177.	-1	-1	-1	-1	-1	-1
178.	-1	-1	-1	-1	-1	-1
179.	-1	1	-1	-1	-1	-1
180.	-1	-1	-1	-1	-1	-1
181.	-1	-1	-1	-1	-1	-1
182.	-1	-1	-1	-1	-1	-1

Table 4.16: SVM Malaria Model Optimal Features

	H _d	F _v	D _z	B _p	V _m	T _{emp}	R _h	R _f	MP _{Count}
1.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.162709	0.9119	-0.65221	1
2.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.085043	0.957267	-0.10316	0.478143
3.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.072292	1.053042	1.177942	0.517028
4.	-1.00657	-0.81804	1.283697	-1.23011	-0.80593	0.406142	0.107695	-0.65221	0.64213
5.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.479172	0.429144	-0.65221	0.236197
6.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	0.479172	0.429144	-0.65221	0.545358
7.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.374844	0.645512	0.354374	0.182296
8.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.374844	0.645512	0.354374	1
9.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.740291
10.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.086202	1.061961	1.360958	0.475562
11.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	-0.02508	1.145716	2.459049	0.729197
12.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	-0.02508	1.145716	2.459049	0.892586
13.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	-2.95207	-2.10135	-0.65221	0.875244
14.	0.992291	-0.81804	1.283697	0.811969	-0.80593	0.280948	-0.69884	-0.65221	0.647264
15.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.879623
16.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.639142	-0.59996	-0.65221	0.094459
17.	0.992291	1.220985	-0.77807	-1.23011	1.239328	-2.95207	-2.10135	-0.65221	0.92193
18.	0.992291	-0.81804	1.283697	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.890881
19.	-1.00657	-0.81804	1.283697	-1.23011	-0.80593	-2.95207	-2.10135	-0.65221	0.904632
20.	0.992291	1.220985	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.890746
21.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.347023	0.333757	-0.65221	0.994243
22.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
23.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	-2.95207	-2.10135	-0.65221	0.877034
24.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	-2.95207	-2.10135	-0.65221	0.890065
25.	0.992291	1.220985	-0.77807	-1.23011	1.239328	-2.95207	-2.10135	-0.65221	0.985197
26.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.853595	-0.08192	-0.65221	0.303031
27.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	1
28.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	1
29.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.534814	0.328716	-0.65221	1
30.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.373684	0.649389	-0.65221	0.26328
31.	0.992291	-0.81804	1.283697	0.811969	-0.80593	0.373684	0.649389	-0.65221	0.395744
32.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.162709	0.9119	-0.65221	1
33.	0.992291	1.220985	-0.77807	0.811969	1.239328	0.162709	0.9119	-0.65221	1
34.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	-0.01233	1.109655	2.276034	0.731247
35.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	-0.01233	1.109655	2.276034	1
36.	0.992291	-0.81804	-0.77807	0.811969	1.239328	-0.01233	1.109655	2.276034	1
37.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.072292	1.053042	1.177942	0.263119
38.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.148799	0.890961	0.354374	0.98352
39.	0.992291	-0.81804	1.283697	0.811969	-0.80593	0.148799	0.890961	0.354374	0.637789
40.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.406142	0.107695	-0.65221	1
41.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
42.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	0.370207	-0.87875	-0.65221	0.662269
43.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.370207	-0.87875	-0.65221	0.989674

44.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	1
45.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	1
46.	-1.00657	1.220985	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.905198
47.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.904179
48.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.479172	0.429144	-0.65221	0.515196
49.	0.992291	1.220985	1.283697	0.811969	-0.80593	0.374844	0.645512	0.354374	0.393407
50.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.374844	0.645512	0.354374	0.122829
51.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.91181
52.	0.992291	-0.81804	1.283697	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.874908
53.	0.992291	-0.81804	1.283697	0.811969	-0.80593	-0.02508	1.145716	2.459049	0.605496
54.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.025923	1.086002	2.093018	0.823278
55.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.025923	1.086002	2.093018	0.998082
56.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.211396	0.900655	0.445882	0.352528
57.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	0.211396	0.900655	0.445882	0.518287
58.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.280948	-0.69884	-0.65221	0.036448
59.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.990215
60.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.185622
61.	0.992291	-0.81804	1.283697	-1.23011	-0.80593	-0.02856	-1.16143	-0.65221	0.763838
62.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.639142	-0.59996	-0.65221	1
63.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.639142	-0.59996	-0.65221	0.474274
64.	0.992291	1.220985	-0.77807	-1.23011	-0.80593	0.728401	-0.35412	0.171359	0.454568
65.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.728401	-0.35412	0.171359	0.502212
66.	-1.00657	-0.81804	1.283697	-1.23011	-0.80593	0.728401	-0.35412	0.171359	0.65216
67.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.347023	0.333757	-0.65221	1
68.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.347023	0.333757	-0.65221	1
69.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
70.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
71.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	0.26137
72.	-1.00657	-0.81804	1.283697	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.996014
73.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	1
74.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.853595	-0.08192	-0.65221	1
75.	0.992291	1.220985	1.283697	-1.23011	1.239328	0.853595	-0.08192	-0.65221	0.785135
76.	0.992291	1.220985	-0.77807	-1.23011	1.239328	0.828092	-0.24284	-0.65221	0.283183
77.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	0.042417
78.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.534814	0.328716	-0.65221	0.393505
79.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.534814	0.328716	-0.65221	1
80.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.373684	0.649389	-0.65221	0.297483
81.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.373684	0.649389	-0.65221	0.315931
82.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.162709	0.9119	-0.65221	1
83.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.162709	0.9119	-0.65221	0.998615
84.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.085043	0.957267	-0.10316	0.416018
85.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.085043	0.957267	-0.10316	0.781805
86.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.085043	0.957267	-0.10316	0.795254
87.	0.992291	1.220985	1.283697	-1.23011	-0.80593	-0.01233	1.109655	2.276034	1
88.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	-0.01233	1.109655	2.276034	1
89.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.072292	1.053042	1.177942	0.997066

90.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.072292	1.053042	1.177942	1
91.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.072292	1.053042	1.177942	1
92.	0.992291	1.220985	1.283697	-1.23011	1.239328	0.148799	0.890961	0.354374	0.804091
93.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
94.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	0.08109
95.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	0.461419
96.	0.992291	1.220985	-0.77807	-1.23011	1.239328	0.370207	-0.87875	-0.65221	0.658623
97.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.370207	-0.87875	-0.65221	1
98.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	0.750135
99.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.590455	-1.32351	-0.65221	0.997572
100.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	0.912067
101.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.074794
102.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	-2.95207	-2.10135	-0.65221	1
103.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.479172	0.429144	-0.65221	0.946937
104.	0.992291	1.220985	1.283697	-1.23011	1.239328	0.374844	0.645512	0.354374	0.155161
105.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.169665	0.957267	1.63548	0.650325
106.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.086202	1.061961	1.360958	1
107.	-1.00657	1.220985	-0.77807	0.811969	1.239328	-0.02508	1.145716	2.459049	0.670289
108.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.025923	1.086002	2.093018	1
109.	0.992291	1.220985	-0.77807	0.811969	1.239328	0.025923	1.086002	2.093018	0.726239
110.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.025923	1.086002	2.093018	0.996476
111.	0.992291	-0.81804	1.283697	-1.23011	1.239328	0.211396	0.900655	0.445882	0.643088
112.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.356296	0.188736	-0.65221	1
113.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.280948	-0.69884	-0.65221	0.86644
114.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.280948	-0.69884	-0.65221	0.996873
115.	0.992291	-0.81804	1.283697	-1.23011	1.239328	-0.02856	-1.16143	-0.65221	0.918954
116.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.639142	-0.59996	-0.65221	0.803595
117.	0.992291	1.220985	-0.77807	-1.23011	1.239328	0.728401	-0.35412	0.171359	0.548329
118.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.148799	0.890961	0.354374	0.00096
119.	-1.00657	1.220985	1.283697	0.811969	1.239328	0.406142	0.107695	-0.65221	1
120.	0.992291	-0.81804	1.283697	0.811969	-0.80593	0.853595	-0.08192	-0.65221	0.317317
121.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.162709	0.9119	-0.65221	0.29175
122.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.085043	0.957267	-0.10316	1
123.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.479172	0.429144	-0.65221	1
124.	0.992291	-0.81804	1.283697	0.811969	-0.80593	0.374844	0.645512	0.354374	0.065023
125.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.374844	0.645512	0.354374	0.250179
126.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	-0.02508	1.145716	2.459049	0.721768
127.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.280948	-0.69884	-0.65221	1
128.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.431024
129.	0.992291	1.220985	-0.77807	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.860151
130.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.347023	0.333757	-0.65221	1
131.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
132.	0.992291	1.220985	1.283697	0.811969	-0.80593	0.373684	0.649389	-0.65221	0.744271
133.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.162709	0.9119	-0.65221	0.347885
134.	0.992291	-0.81804	1.283697	0.811969	-0.80593	-0.01233	1.109655	2.276034	0.250818
135.	0.992291	1.220985	1.283697	0.811969	-0.80593	0.072292	1.053042	1.177942	0.563549

136.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.148799	0.890961	0.354374	0.308123
137.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.406142	0.107695	-0.65221	1
138.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.370207	-0.87875	-0.65221	0.039604
139.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.590455	-1.32351	-0.65221	0.401522
140.	-1.00657	1.220985	1.283697	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.891301
141.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.025923	1.086002	2.093018	0.449146
142.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.347023	0.333757	-0.65221	0.118362
143.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.347023	0.333757	-0.65221	1
144.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
145.	-1.00657	-0.81804	1.283697	0.811969	1.239328	-2.95207	-2.10135	-0.65221	1
146.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.853595	-0.08192	-0.65221	0.226791
147.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.534814	0.328716	-0.65221	1
148.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.373684	0.649389	-0.65221	0.822258
149.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.162709	0.9119	-0.65221	0.623176
150.	0.992291	1.220985	1.283697	-1.23011	1.239328	-0.01233	1.109655	2.276034	1
151.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	-0.01233	1.109655	2.276034	0.068918
152.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.072292	1.053042	1.177942	0.028967
153.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.072292	1.053042	1.177942	0.441393
154.	0.992291	1.220985	-0.77807	0.811969	1.239328	0.072292	1.053042	1.177942	0.777298
155.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
156.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
157.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.370207	-0.87875	-0.65221	1
158.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	1
159.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.590455	-1.32351	-0.65221	0.087654
160.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	-2.95207	-2.10135	-0.65221	1
161.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.479172	0.429144	-0.65221	1
162.	-1.00657	1.220985	1.283697	0.811969	1.239328	0.086202	1.061961	1.360958	1
163.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.025923	1.086002	2.093018	0.050752
164.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.025923	1.086002	2.093018	0.093835
165.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.356296	0.188736	-0.65221	1
166.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.280948	-0.69884	-0.65221	1
167.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.280948	-0.69884	-0.65221	0.150215
168.	-1.00657	1.220985	1.283697	0.811969	1.239328	0.406142	0.107695	-0.65221	0.004115
169.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	-2.95207	-2.10135	-0.65221	0.064331
170.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.479172	0.429144	-0.65221	0.445977
171.	-1.00657	1.220985	1.283697	0.811969	1.239328	0.086202	1.061961	1.360958	0.040905
172.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.356296	0.188736	-0.65221	1
173.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.280948	-0.69884	-0.65221	1
174.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.639142	-0.59996	-0.65221	1
175.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.728401	-0.35412	0.171359	0.30461
176.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.728401	-0.35412	0.171359	0.434631
177.	0.992291	-0.81804	1.283697	-1.23011	1.239328	0.853595	-0.08192	-0.65221	0.799844
178.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.853595	-0.08192	-0.65221	0.519265
179.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	1
180.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.534814	0.328716	-0.65221	1
181.	0.992291	1.220985	-0.77807	-1.23011	1.239328	-0.01233	1.109655	2.276034	0.998109

182.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.072292	1.053042	1.177942	0.50418
183.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.072292	1.053042	1.177942	1
184.	0.992291	1.220985	-0.77807	-1.23011	1.239328	0.148799	0.890961	0.354374	0.659226
185.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	0.521674
186.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.370207	-0.87875	-0.65221	1
187.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	0.952026
188.	-1.00657	-0.81804	-0.77807	-1.23011	1.239328	0.590455	-1.32351	-0.65221	1
189.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.907046
190.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.347023	0.333757	-0.65221	0.603454
191.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
192.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.072292	1.053042	1.177942	1
193.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.406142	0.107695	-0.65221	1
194.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.030993
195.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.280948	-0.69884	-0.65221	1
196.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.347023	0.333757	-0.65221	1
197.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.347023	0.333757	-0.65221	1
198.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
199.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
200.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
201.	-1.00657	-0.81804	1.283697	0.811969	1.239328	-2.95207	-2.10135	-0.65221	1
202.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.085043	0.957267	-0.10316	0.169294
203.	0.992291	-0.81804	1.283697	-1.23011	-0.80593	0.356296	0.188736	-0.65221	0.60066
204.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.492988
205.	-1.00657	1.220985	1.283697	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	1
206.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.639142	-0.59996	-0.65221	0.902564
207.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.639142	-0.59996	-0.65221	1
208.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.085043	0.957267	-0.10316	0.924687
209.	0.992291	1.220985	-0.77807	-1.23011	1.239328	-0.01233	1.109655	2.276034	0.019287
210.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.072292	1.053042	1.177942	1
211.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.072292	1.053042	1.177942	1
212.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.148799	0.890961	0.354374	0.603332
213.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.226466	-0.51465	-0.65221	0.114329
214.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.370207	-0.87875	-0.65221	1
215.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	0.956886
216.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.590455	-1.32351	-0.65221	1
217.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.374844	0.645512	0.354374	1
218.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.347023	0.333757	-0.65221	0.86659
219.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.347023	0.333757	-0.65221	0.017748
220.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
221.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	1
222.	0.992291	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	0.970095
223.	-1.00657	-0.81804	1.283697	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.088401
224.	-1.00657	-0.81804	1.283697	0.811969	-0.80593	0.853595	-0.08192	-0.65221	0.239206
225.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.853595	-0.08192	-0.65221	1
226.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.347023	0.333757	-0.65221	0.566714
227.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.162709	0.9119	-0.65221	0.362088

228.	0.992291	1.220985	1.283697	0.811969	1.239328	0.406142	0.107695	-0.65221	1
229.	0.992291	-0.81804	1.283697	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.940144
230.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
231.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.373684	0.649389	-0.65221	1
232.	0.992291	-0.81804	1.283697	0.811969	1.239328	-0.01233	1.109655	2.276034	1
233.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.406142	0.107695	-0.65221	1
234.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
235.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.356296	0.188736	-0.65221	0.856609
236.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.280948	-0.69884	-0.65221	1
237.	0.992291	-0.81804	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.99602
238.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.853595	-0.08192	-0.65221	1
239.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	1
240.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.373684	0.649389	-0.65221	1
241.	0.992291	1.220985	-0.77807	-1.23011	1.239328	0.162709	0.9119	-0.65221	0.709403
242.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.162709	0.9119	-0.65221	1
243.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.356296	0.188736	-0.65221	1
244.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.280948	-0.69884	-0.65221	1
245.	0.992291	1.220985	1.283697	-1.23011	1.239328	-0.02856	-1.16143	-0.65221	0.908644
246.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.728401	-0.35412	0.171359	1
247.	0.992291	1.220985	1.283697	-1.23011	-0.80593	0.296018	-0.65618	-0.65221	0.991557
248.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.406142	0.107695	-0.65221	1
249.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.347023	0.333757	-0.65221	1
250.	-1.00657	-0.81804	-0.77807	0.811969	-0.80593	0.072292	1.053042	1.177942	0.73615
251.	-1.00657	1.220985	1.283697	0.811969	-0.80593	-0.02856	-1.16143	-0.65221	0.163784
252.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.639142	-0.59996	-0.65221	1
253.	0.992291	-0.81804	1.283697	-1.23011	-0.80593	0.728401	-0.35412	0.171359	0.571762
254.	0.992291	-0.81804	1.283697	-1.23011	1.239328	0.085043	0.957267	-0.10316	0.44255
255.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.085043	0.957267	-0.10316	0.352347
256.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.072292	1.053042	1.177942	0.635124
257.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.072292	1.053042	1.177942	0.164733
258.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.148799	0.890961	0.354374	1
259.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.226466	-0.51465	-0.65221	1
260.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.853595	-0.08192	-0.65221	1
261.	-1.00657	1.220985	-0.77807	-1.23011	1.239328	0.534814	0.328716	-0.65221	1
262.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.590455	-1.32351	-0.65221	0.316507
263.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.374844	0.645512	0.354374	1
264.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.347023	0.333757	-0.65221	0.873727
265.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
266.	0.992291	1.220985	-0.77807	0.811969	1.239328	0.296018	-0.65618	-0.65221	0.145928
267.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	-2.95207	-2.10135	-0.65221	0.847475
268.	-1.00657	-0.81804	-0.77807	-1.23011	-0.80593	0.479172	0.429144	-0.65221	0.358871
269.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.374844	0.645512	0.354374	0.237692
270.	0.992291	1.220985	1.283697	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.978333
271.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.086202	1.061961	1.360958	0.624775
272.	0.992291	-0.81804	-0.77807	0.811969	1.239328	-2.95207	-2.10135	-0.65221	0.091376
273.	-1.00657	-0.81804	1.283697	-1.23011	1.239328	0.853595	-0.08192	-0.65221	0.192909

274.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.853595	-0.08192	-0.65221	1
275.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.828092	-0.24284	-0.65221	0.886788
276.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.373684	0.649389	-0.65221	1
277.	-1.00657	1.220985	-0.77807	-1.23011	-0.80593	0.162709	0.9119	-0.65221	1
278.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.085043	0.957267	-0.10316	0.917106
279.	0.992291	1.220985	1.283697	0.811969	-0.80593	-0.02508	1.145716	2.459049	0.776602
280.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.347023	0.333757	-0.65221	1
281.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
282.	0.992291	-0.81804	-0.77807	0.811969	1.239328	0.373684	0.649389	-0.65221	1
283.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.162709	0.9119	-0.65221	1
284.	0.992291	1.220985	-0.77807	-1.23011	1.239328	-2.95207	-2.10135	-0.65221	1
285.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.025923	1.086002	2.093018	0.47349
286.	-1.00657	1.220985	1.283697	-1.23011	1.239328	0.406142	0.107695	-0.65221	1
287.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.226466	-0.51465	-0.65221	1
288.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.894348
289.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.280948	-0.69884	-0.65221	1
290.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.639142	-0.59996	-0.65221	1
291.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.370207	-0.87875	-0.65221	0.87778
292.	-1.00657	1.220985	-0.77807	0.811969	1.239328	0.590455	-1.32351	-0.65221	0.177714
293.	0.992291	-0.81804	1.283697	0.811969	1.239328	-0.01233	1.109655	2.276034	0.343466
294.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.356296	0.188736	-0.65221	1
295.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.280948	-0.69884	-0.65221	1
296.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.728401	-0.35412	0.171359	0.270432
297.	0.992291	1.220985	1.283697	-1.23011	-0.80593	0.296018	-0.65618	-0.65221	0.104198
298.	-1.00657	1.220985	1.283697	0.811969	-0.80593	0.406142	0.107695	-0.65221	1
299.	0.992291	-0.81804	-0.77807	-1.23011	1.239328	0.639142	-0.59996	-0.65221	1
300.	-1.00657	-0.81804	1.283697	0.811969	1.239328	0.280948	-0.69884	-0.65221	1
301.	-1.00657	1.220985	1.283697	-1.23011	-0.80593	0.639142	-0.59996	-0.65221	1
302.	0.992291	1.220985	-0.77807	0.811969	-0.80593	0.347023	0.333757	-0.65221	1
303.	0.992291	-0.81804	-0.77807	-1.23011	-0.80593	0.148799	0.890961	0.354374	1
304.	0.992291	-0.81804	-0.77807	0.811969	-0.80593	0.226466	-0.51465	-0.65221	1
305.	-1.00657	-0.81804	-0.77807	0.811969	1.239328	0.853595	-0.08192	-0.65221	1
306.	0.992291	-0.81804	1.283697	0.811969	1.239328	0.347023	0.333757	-0.65221	0.788668
307.	-1.00657	1.220985	-0.77807	0.811969	-0.80593	0.296018	-0.65618	-0.65221	1
308.	0.992291	1.220985	1.283697	-1.23011	-0.80593	0.169665	0.957267	1.63548	0.522879

Appendix D:

Table 4.17: Hybridized SVM_ANN Malaria Model Target and Simulated Test Result

S/No	SVM_ANN Malaria Model Result	
	Target Value	Sim_ Result
1.	False	False
2.	False	False
3.	True	True
4.	False	False
5.	False	True
6.	False	False
7.	True	True
8.	False	False
9.	False	False
10.	False	False
11.	False	False
12.	False	False
13.	True	True
14.	False	False
15.	False	False
16.	False	False
17.	False	True
18.	False	False
19.	False	False
20.	False	False
21.	True	True
22.	False	False
23.	False	False
24.	False	False
25.	True	True
26.	False	False
27.	False	False
28.	False	False
29.	True	True
30.	False	True
31.	False	False
32.	False	False
33.	False	False
34.	False	False
35.	False	False
36.	False	True
37.	False	False
38.	True	True
39.	False	False
40.	False	False
41.	False	False
42.	False	False
43.	False	False
44.	False	False
45.	False	True
46.	True	True
47.	False	False
48.	False	False
49.	False	False
50.	False	False

51.	True	True
52.	False	False
53.	False	False
54.	False	False
55.	False	False
56.	True	True
57.	True	True
58.	False	False
59.	False	False
60.	False	False
61.	False	False
62.	False	False
63.	False	False
64.	False	False
65.	False	False
66.	False	False
67.	False	False
68.	False	True
69.	False	False
70.	False	True
71.	False	False
72.	False	False
73.	False	False
74.	False	False
75.	False	False
76.	True	True
77.	False	False
78.	False	False
79.	False	False
80.	True	True
81.	False	False
82.	False	False
83.	False	True
84.	False	False
85.	False	False
86.	False	False
87.	False	False
88.	False	False
89.	False	False
90.	False	False
91.	False	False
92.	False	False