Ministry of Higher Education and
Scientific Research
University of Diyala
College of Science
Department of Computer Science

Thalassemia Disease Classification Based On Machine Learning Techniques

A Research

Submitted to the Department of Computer Science\ College of Sciences\ University of Diyala in a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science

By

Muna Qais Mohammed

Supervised By

Assist. Prof. Dr. Jamal Mustafa Al-Tuwaijari

2021A.D ***1443 A.H***

بسم الله الرحمن الرحيم

﴿نَرْفَعُ دَرَجَاتٍ مَنْ نَشَاءُ وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ﴾

صدق الله العظيم

سورة يوسف آية (76)

Examination Committee Certification

We certify that we have read the thesis entitled ***"Thalassemia Disease Classification And Detection Based On Artificial Intelligence Technique"*** *and an examination committee, examined the student* ***"Muna Qais Mohammed"*** *in the thesis content and our opinion it is adequate as fulfill the requirement for Degree of Master in Computer Science at the Computer Science Department, University of Diyala*.

Signature:

Name: **Prof.Dr. Ziad Tariq Mustafa** (**Chairman**)

Date: **/ /2021**

Signature:

Name: **Assist.Prof.Dr. Khalid Shaker Jasim** (**Member**)

Date: **/ /2021**

Signature:

Name: **Assist.Prof.Dr. Bashar Talib Hameed** (**Member**)

Date: **/ / 2021**

Signature:

Name: **Assist. Prof. Dr. Jamal Mustafa Al-Tuwaijari** (**Member/Supervisor**)

Date: **/ / 2021**

Approved by the **Dean** of College of Science, University of Diyala.

(**The Dean**)

Signature:

Name: **Prof. Dr. Tahseen H. Mubarak**

Date: **/ / 2021**

(Supervisor Certification)

I certify that this research entitled ***"Thalassemia Disease Classification And Detection Based On Artificial Intelligence Technique"*** *was prepared by* ***"Muna Qais Mohammed"*** *Under my supervision at the Department of Computer Science\ College of Sciences\ University of Diyala in a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science.*

(Supervisor)
Name: **Assist. Prof. Dr. Jamal Mustafa Al-Tuwaijari**

Signature:
Date: / /2021
Approved by Department of Computer Science\ College of Sciences\ University of Diyala

Signature:

Name: **Assist. Prof. Dr.Bashar Talib Hameed**
Date: / /2021
Head of Department of Computer Science\ College of Sciences\ University of Diyala

(Linguistic Certification)

I certify that the thesis entitled ***"Thalassemia Disease Classification And Detection Based On Artificial Intelligence Technique"*** *was prepared by* ***"Muna Qais Mohammed"*** *and was reviewed linguistically. Its language was amended to meet the style of the english language.*

Signature:

Name:

Date: / /2021

(Scientific Amendment)

I certify that the thesis entitled ***"Thalassemia Disease Classification And Detection Based On Artificial Intelligence Technique"*** *was prepared by* ***"Muna Qais Mohammed"*** *has been evaluated scientifically ; therefore, it is suitable for debate by examining committee*.

Signature:

Name:

Date: / / 2021

ACKNOWLEDGMENT

Initially and foremost I deeply thank God, Lord of the worlds, for all the grace and helping to bring this thesis to its end.

I would like to grateful my supervisor Dr Jamal Mustafa AL-Tuwaijari for his support and encouragement. I am extremely thankful for his effort, completing and revising the thesis. My sincere thanks and gratitude is for the Council of the College of Science at Diyala University and the staff of Computer Science Department for all the supports to perform this study and for giving me the opportunity to have the M.Sc.

I would like to express my gratitude to my father and mother, my daughter , my brothers and their families and my sister who had unlimited support and patience.

There are not enough words to thank my dear husband *Bassam* for his logistic support and belief and encouragement throughout my studies. I pass many Thanks to my colleagues *Hayder Al-Bawe and Hafssa* .

Finally, I apologize for all those who not mentioned here and I thank them a lot.

muna q. mohammed

DEDICATION

To the flower of life and its light and the most precious person in my life, my tender mother.

To whom I proudly carry your name, teach me how to make success and instill confidence in myself my dear father, may God extend your life.

To those who supported me in adversity and the source of my happiness, my companion to my path and my love, my dear husband.

To whom their love and blood flow in my veins, and I lived with them the most beautiful moments, my brothers and sisters.

To the hony of my love and my soul, my dear my daughter.

To the candles that illuminate the path of knowledge in my path, my distinguished teachers.

To everyone who loved me and supported me in my scientific and practical life.

muna q. mohammed

ABSTRACT

Thalassemia is considered one of the most common genetic blood disorders that has received excessive attention in the medical research fields worldwide. It cannot be cured , but an early detection and classification using screening process is the best way to prevent the disease. If early classification is done, patients can get the right treatment. It helps them increase their life expectancy and reduce the risk of thalassemia to the next generation.

In this thesis, efficient thalassemia classification system have been design to increase the accuracy and decrease the error rate in the diagnosis process. This system based on two proposed approaches for classifying thalassemia disease. The first proposed approach is based on four machine learning technique which include artificial neural network (ANN), decision tree (DT), k-nearest neighbor (KNN) and logistic regression (LR). This approach consist of two main stages: pre-processing and classification of thalassemia disease .The second proposed approach based on two deep learning techniques , these are consist of convolutional neural network (CNN) and deep neural network (DNN).

The proposed system has been tested by using two thalassemia dataset .The first type of dataset contain 391 sample with nine features and split into 30% for testing and 70% for training and the second type of dataset contain 7108 image with nine type of erythrocytes and split into 20% for testing and 80% for training. The comparison results show that the proposed system has efficient diagnosis performance and can be used as a promising tool for thalassemia disease diagnosis compared to another existing method where the accuracy rate of first proposed approach for first dataset using ANN is 99%, DT is 99% ,LR is 99% and KNN is 98%. The result of second proposed approach for the second dataset using deep learning technique has an average accuracy of 99% for CNN and 83% for DNN.

List of Contents

List of Abbreviations

Abbreviations	*Description*
AI	*Artificial Intelligent*
ANN	*Artificial Neural Networks*
BN	*Bayesian Networks*
CAD	*Computer Aided Diagnosis*
CBC	*Complete Blood Count*
CNN	*Convolutional Neural Network*
CT	*Computed Tomography*
DL	*Deep Learning*
DNN	*Deep Neural Network*
DT	*Decision Tree*
FC	*Fully Connected*
FCM	*Fuzzy C-Means*
FKRCM	*Fuzzy Kernel C-Means*
FRCM	*Fuzzy Robust C-Means*
HB	*Hemoglobin Electrophoresis*
HCDP	*hierarchical clustering based on density peaks*
HCT	*Hematocrit Test*
HCTM	*Hospital Canselori Tuanku Muhriz*
K-NN	*K-Nearest Neighbor*
LR	*Logistic Regression*
MCV	*Mean Corpuscular Volume*
ML	*Machine Learning*
MLNN	*Multilayer Neural Network*
MLP	*Multi-Layer Perceptron*

MRI	*Magnetic Resonance Imaging*
NB	*Naive Bayes*
PCA	*Principal Components Analysis*
PET	*Positron Emission Tomography*
PTPP	*Punjab Thalassaemia Prevention Programme*
RBC	*Red Blood Cell*
RBFN	*Radial Basis Function Network*
ReLU	*Rectified Linear Unit*
ResNet	*Residual Neural Network*
RGB	*Red, Green, Blue*
SEM	*Scanning Electron Microscope*
SGR-VC	*SVM, GBM , RF-voting classifier*
SMOTE	*Synthetic Minority Oversampling Technique*
SVM	*Support Vector Machine*

List of Tables

List of Figures

List of Algorithms

Chapter One

General Introduction

Chapter One

General Introduction

1.1 Introduction

[illegible] (1). [illegible] (2).

[illegible] 300 [illegible] and functionality of a red blood cell. A globin molecule contains two parts: α [illegible] globin and β globin. The α [illegible] 141 [illegible] are regulated [illegible] by genes on chromosome 16. On the other hand, the β [illegible] 146 [illegible] 11 (3).

[illegible]

□g□ri□□ □□e□ □□□er □□e □□□□□□ i□e re□uire□ □r □e□e □□i□ie□□□ e□□□ □i□□er re□e□□□re□□e□□□r□egie□□1□.

□□□i□ □□e □e□i□□□r□e□i□□ i□□uge□□□□□□□ □□□ □i□□e□□e□i□i□ □□□ u□i□□e□ □□i□□□ □i□u□ □□e□i□. □□e i□□r□ □i□□ □□□□e□e□ □□ i□e□e□i□ □□□□e□i□□re□ure r□□e □□□i□e□□□□e□ge. □r□i□i□□□□e□ige□e □□r□□□e□ □□□e u□i□□e□ □□□□□er□□e□e □□□e□□□□i□□□u□□e □□□□i□r□er □□i□g□□e □□e □□ie□□4□. □r□i□i□□i□e□ige□e (□□) i□□e r□e□□□□i□u□□i□g □u□□□ i□e□ige□e i□□□□i□e□□□□□re r□i□e□□□i□□□□□□i□e □u□□□. □e□ri□r□ □ur□□e □□□□□□□i□□i□□i□□ □□e □e□□□ie□□i□□□i□e□ig□e □□e re□i□□□i□□e□ee□ □re□e□i□e □r re□□e□□e□□i□ue□ □□□ □□ie□□ u□□□e□□□i□□ i□ □□□□□ □i□e□□r□ug□□e u□e □□□g□ri□□ □□□□ r□gr□ □□□□ri□u□r□□□□i□e e□r□i□g □g□ri□□□□5□.

□□□i□e □e□r□i□g (□□) □ r□□□□ □□ □□i□ □i□□ □□□□u□er □r□gr□□□ (□g□ri□□□□) □ui□□ re□i□i□e □□□er □□□□e□i□□□r□□ □□□□i□□□□□e□. □□□i□e □e□r□i□g i□ □□e □□□ u□er□□□i□e□ □□□i□□i□□ □□ □ie□i□i□ □□e□□ □□□ □□□□. □□□i□e □e□r□i□g e□□□□□□ □ □i□er r□□ge □□ □□□i□i□□□ □□□r□□□e□ □□□□ □re □□□ □□□□□e□ □□□□e□i□ □e□□□e□i□e. F□r e□□□ □e □ee□ □e□r□i□g(□□) i□□ui□□□□ □□e□□□□□□□□e □e□er □□□u□i□□□ □□□u□□e u□er□i□g □□□□ □□□□ e□□□ □□□□e □□re □□□ □e□ □□□□. □□□i□e □e□r□i□g □g□ri□□□□□ □e u□i□□e□ i□ □□e □e□i□□□r□e□i□□□r □re□i□i□□ □□□□□□□□□□□□□□i□□□□□□. □□e□□□□□□□□□□□□i□□□□ i□g□□e □e□□□□□□gi□□□i□r□er□ □□□ □□□□er □□□□□□□ □e□i□er □□re □□□ur□e i□□i□g□□□ □□□ie□□□6□.

□e□u□e □□□e□□□□e□□□i□ □ee□ □e□r□i□g □e□□□□□g□□□e u□e □□r□i□i□□ i□e□ige□e (□□) i□□e□i□i□e □□□re□e□□□g□e□□□□□□□□e□i□□. □ee□ □e□r□i□g i□ □ □□□e □□ □□□□i□e □e□r□i□g i□ r□i□i□i□□i□e□ige□e (□□) □□□□u□e□ □eur□□ □e□□r□□ □□ e□r□u□u□er□i□e□ r□□ u□□ru□□ure□ r u□□□□e□e□ □□□□. □□□□□□□□□ □ere i□ □□□□□□i□ere□i□ u□i□□i□g □□□□ □i□g□□e □ri□u□i□□□ □□□e□i□□

i ge ri ri r ug u eur e r ( ) e ee
e r i g e g u er i i . gi g u er ie e
e e e gre e e e i er re i e i i ge
i e i e r i g i . e e i e ig u er i e i g i
( ) e e i u e i ge re r e i g e e i i i regi ( ) e
i ere e ure e r i e e i i i i i 7 .

e g ri u g ri i g i u e
i re e i er re i u i e r i r e r i g u er
er r er i u eri g i e u e u er i 8 . i
e i r e e e e e e i i i i i i r r e
r e ie e e i e e e i r er er i r i .
r i i i eur e r ( ) gi i egre i ( ) e re eig r
( ) e i i ree ( ) re u e e i ier e i ie e
e i i e e ie r e ee e r i g e i ue u e
u eur e r ( ) ee eur e r ( ) re
u e .

1.2 Related Work

ere re r e e i i i i e u i g i ere i e
e r i g g ri i e i e er u r i e i e ie i g
ur e u i e e i i i i i e er e i i i g i e
e i e er.

E. Elshami et al.(2015) 9 i u ree r e e i ier (e i i
ree e e e eur e r) ere ie . i re e r i e ig e
e urre e e i u i g e er i ier e e e . e
e i ier u e i i gui e ee ie i e i r i —

[illegible] 4600 [illegible] 500 [illegible] 0.3 [illegible] 0.2 [illegible] 93.64 [illegible] 93.7 [illegible] 95.71 [illegible].

D .Tyas et al .(2017) [illegible] 256 [illegible] 43 [illegible] 6 [illegible] 0.1 [illegible] 0.05 [illegible] 93.24 [illegible]. [illegible] 92.55 [illegible].

I.Ahmed et al .(2018) [illegible]

[illegible] These patients' peripheral blood smear were [illegible] 725 [illegible] 99 [illegible] 31 [illegible] 83.5 [illegible] 83.5 [illegible] 83.3 [illegible]

A. AlAgha et al. (2018) [illegible] β [illegible] β [illegible] 4.3 [illegible] β [illegible] 320 β [illegible] 10 [illegible] 400 [illegible] β [illegible] 99.47 [illegible] 98.81 [illegible]

D .Tyas et al .(2019) [illegible] 12 [illegible] Fu [illegible]

e[illegible] 7108 [illegible] i[illegible]g[illegible]e er[illegible]r[illegible]e [illegible]i[illegible]u[illegible]e[illegible] i[illegible]e [illegible]i[illegible]ere[illegible]e[illegible] [illegible]er[illegible]r[illegible]e[illegible]. [illegible]e [illegible]er[illegible]r[illegible]e [illegible]e [illegible]ur [illegible]i[illegible]er[illegible] i[illegible] [illegible]i[illegible]i[illegible]g [illegible]i[illegible]e [illegible]i[illegible]ere[illegible]e[illegible] [illegible]er[illegible]r[illegible]e[illegible] [illegible]re[illegible]. F[illegible]r [illegible]e [illegible]i[illegible]i[illegible] [illegible]r[illegible]i[illegible]i[illegible]g [illegible]e[illegible]i[illegible]g [illegible]e [illegible]r[illegible]i[illegible]i[illegible]r[illegible]u[illegible]i[illegible]i[illegible]e[illegible]. [illegible]e i[illegible]u[illegible]i[illegible]er [illegible]i[illegible] 38 [illegible]ri[illegible]u[illegible]e [illegible]ue[illegible]e[illegible]r[illegible]e[illegible] [illegible]i[illegible] [illegible]e [illegible]e[illegible]ure e[illegible]r[illegible]i[illegible]. [illegible] [illegible]e [illegible]e[illegible]er[illegible]r[illegible]e i[illegible] [illegible]e e[illegible]eri[illegible]e[illegible]i[illegible] 89.6 [illegible]er[illegible]e[illegible]ur[illegible] 89.3 [illegible]er[illegible]e[illegible]re[illegible]i[illegible]i[illegible] 89.6 [illegible]er[illegible]e[illegible]re[illegible]. [illegible]r[illegible]e[illegible]e[illegible]e[illegible] [illegible]F [illegible]e[illegible]r[illegible]e [illegible]e [illegible]e[illegible]i[illegible]er.

M. Hajipour et al .(2019) [13] [illegible]i[illegible]e[illegible]r[illegible] [illegible]gi[illegible]i[illegible] [illegible]egre[illegible]i[illegible] [illegible]e[illegible] [illegible]e[illegible]e[illegible]i[illegible]g[illegible]ri[illegible] [illegible]u[illegible]i[illegible]r[illegible]e[illegible]. [illegible]e [illegible]ur[illegible]e [illegible]i[illegible]u[illegible]u[illegible]e [illegible]e [illegible]re[illegible]e[illegible]e[illegible] [illegible]g[illegible]ri[illegible] [illegible]r [illegible]i[illegible]i[illegible]g [illegible]i[illegible]u[illegible]i[illegible] [illegible]ru[illegible] [illegible]re[illegible]i[illegible]i[illegible]e [illegible]e[illegible]r [illegible]ie[illegible]i[illegible]e[illegible]i[illegible]r[illegible]i[illegible]. [illegible]ie[illegible]i[illegible]e[illegible]i[illegible] [illegible]ere [illegible]e[illegible]e[illegible]r[illegible] [illegible]ie[illegible]re[illegible]r[illegible] [illegible]e[illegible]i[illegible]er [illegible]10[illegible] [illegible]e[illegible]r [illegible]eri[illegible]r [illegible]i[illegible]u[illegible]. [illegible]ie[illegible]ere [illegible]e[illegible]r[illegible]e[illegible]i[illegible]gr[illegible]u[illegible]e [illegible]ie[illegible]e [illegible]ur[illegible]i[illegible]e[illegible]i[illegible]e 26 [illegible]ri[illegible]e[illegible]re[illegible]rie[illegible]e[illegible]r[illegible] e[illegible]gr[illegible]u[illegible]. [illegible]e [illegible]er[illegible]ge [illegible]ge [illegible]e [illegible]e[illegible]i[illegible]ie[illegible] 25.7[illegible]9.04 [illegible]e[illegible]r[illegible] 185 [illegible]e[illegible]e [illegible]ie[illegible]e e[illegible]e [illegible]u[illegible]. [illegible]ere [illegible]ere [illegible]26 i[illegible]e[illegible]e[illegible]e[illegible]ri[illegible]e[illegible]i[illegible]er. Fur[illegible]er[illegible]re[illegible]e [illegible]e[illegible]e [illegible]i[illegible]i[illegible]g [illegible]ri[illegible]e[illegible]r e[illegible]ie[illegible] 1.8[illegible]0.81. [illegible]i[illegible] 94.35 [illegible]er[illegible]e[illegible]ur[illegible]e [illegible]i[illegible]e[illegible]re[illegible]i[illegible]i[illegible]e [illegible]e[illegible]e [illegible]re[illegible] [illegible]e [illegible]ie[illegible]ur[illegible]i[illegible]r[illegible]e. [illegible]i[illegible]u[illegible]i[illegible]i[illegible]ere[illegible] 26 i[illegible]e[illegible]e[illegible]e[illegible] [illegible]ri[illegible]e[illegible] [illegible]ere [illegible]ie[illegible]r[illegible]i[illegible] [illegible]re[illegible]i[illegible]r[illegible]i[illegible] [illegible]ere [illegible]eri[illegible]e[illegible] [illegible]r[illegible] 12 [illegible]ri[illegible]e[illegible]. [illegible]i[illegible]i[illegible]g [illegible]i[illegible]u[illegible]i[illegible]i[illegible]i[illegible]r[illegible]r[illegible]eg[illegible]r[illegible]i[illegible]r[illegible]i[illegible]g [illegible]e e[illegible]i[illegible]ie[illegible]i[illegible]i[illegible]g [illegible]g[illegible]ri[illegible].

N .Chidozie Egejuru et al .(2019) [14] [illegible]i[illegible]r[illegible]u[illegible]er[illegible]i[illegible]e[illegible]i[illegible]e [illegible]e[illegible]r[illegible]i[illegible]g [illegible]e[illegible] ([illegible]e [illegible]e[illegible]u[illegible]i[illegible]er [illegible]er[illegible]e[illegible]r[illegible]) [illegible]ugge[illegible]e[illegible]re[illegible]e [illegible]e [illegible]re[illegible]i[illegible]i[illegible]e [illegible]e[illegible]r ri[illegible]e[illegible]i[illegible]. [illegible]e [illegible]i[illegible]ir[illegible]e[illegible]r [illegible]e[illegible]ge [illegible]i[illegible] [illegible]u[illegible]e[illegible]i[illegible]u[illegible]e [illegible]e [illegible]re[illegible]i[illegible]i[illegible]e [illegible]e[illegible]r [illegible]e[illegible]i[illegible]ri[illegible] ([illegible]). [illegible]e i[illegible]ri[illegible]uire[illegible]r[illegible] [illegible]e [illegible]i[illegible]e[illegible]i[illegible]i[illegible]g [illegible]e [illegible]r[illegible]e[illegible]er[illegible]e ri[illegible]

□ □e□i□ □u□e□ □ □i□e □e □i□u□e□ □e□. F□i□g □e □e□i□ □ □r□ 51 □ie□ □e re□e□r□er□ □i□ere□ □e□gr□i□ □ri□e□u□ □ge□er□ge□ □ri□u□e□i□i□ □i□ □e□ □i□i□ri□e□u□ □i□ □i□r□ee□e□rge□e□i□e□e□uri□e □r □ge□ □re□rrier□i□ □ri□ □i□ri□u□i□ □43 □er□e□ □e□10 □er□e□e□10 □er□e□e. □e re□e□r□er□u□ □i□g □u□i□ □er □er□e□r□ □r □e□i□re□i□i□ □u□i□ □r□e □e □e□i□i□ □i□g □r□e□i□i□e □e □e□re □e□.

D .Tyas et al.(2020) □15□i□ □i□ □u□ □u□i□er □er□e□r□ (□ □) □i□ □r□g□i□ □e□r□i□g □e□ □ere □r□e□ □ □er□r□ □e □i□i□i□ □ge i□ □i□ □u□. □e g□ □i□ □u□i□ □ □i□ □r□eri□i□ □e u□i□i□e□ □ □i□i□gui□ □i□e □i□ere□ □e□ □er□r□e□. □ □e□ □ere u□e□i□ □i□i□e□ig□i□. □e □1 □e□i□u□e□7108 □i□g□e er□r□e □u□ i□ □ge□eri□e□r□ □e□eg□e□i□re□u□. □e □2 □e□i□u□e□e □1 □e□i□2247 □i□g□e er□r□e □u□i□ge. □e□ □i□i□ □e□ure□r□ □r□gi□e□ure□e □u□ig□e□ur□ □98.11 □er□e□i□u□ □e□i□i□i□ □99.75 □er□e□e□i□i□i□ □86.67 □er□e□ □e □re□0.99.

F. Aszhari et al .(2020) □16□e r□ □re□g□ri□ □r□e□ □u□i□ □r □i□i□g □e□i□i□e□. □ere □re 150 □e□e□i□e □e□ □i□ 82 □e□ □i□g □e□i□ □ 68 □i□g □e□i□. □e □e□i□ □e□i□ □e□ □ri□e□ □rge□ □ri□e□ □r □e□ (□e□i□ □e□i□). □i□g □u□i□e □i□e i□ □e r□ge □70 □85 □er□e□ □ r□i□i□g □e r□ □re□g□ri□ □ □ur□re□ □re□i□i□ □re □100 □er□e□. □e□ □i□e□e □e□re□ □ue □e□r□ □e e□ire □er□e□ge □r□i□i□g □i□ i□ 100 □er□e□. □e r□ □re□ □g□ri□ □er□ge □ur□re□ □re□i□i□u□i□i□g □u□i□e □i□e i□ □e

r□ge □□50 □□ 85 □er□e□□□□r□i□i□g □□□□ re 98.99 □er□e□□ 100 □er□e□□□□□ 98.20 □er□e□□re□e□i□e□□.

S. SADIQ et al.(2021) □17□□□□□□□□e□i□□rrier□□re i□e□i□ie□u□i□g □e□er□□ □u□□i□□□i□e □□□□□ □e□ure□ □□u□□ e□e□i□e□□ u□i□g □ □ig□□□er□r□□□e i□ui□ □r□□ □□gr□□□□(□□□□) e□□□□ i□□i□e□□e□i□e□i□e □□□□u□i□g□□□□re□uire□ □e□i□i□e□e□ui□□e□. □□i□r□□e□ □□□□□e□er□□e□e□i□□e□□□e □e□e□□□□e □□□□□e□e□i□e □□□□ui□□□ree□i□g □□r□□□□e. □□e□g□□□□□□i□□u□□i□□□i□e□i□□ β□□□□□□□□e□i□□rrier□u□i□g re□□□□□□ e□i□i□e□r□□ □u□□□□□□□□□□u□. □□e □u□□□□ □□□□□□□e□i□ □re□e□i□□ □r□e□□□□ □e□r□□ □□□□□e□□□□□u□e□i□ □□i□ i□e□ig□i□□. In this database out of 5066 records 3051 patients are β□ □□□□□□e□i□□□□Carriers and 2015 records are β□□□□□□□e□i□□□rrier□. □u□□r□ □e□□r □ □□□i□e□□r□ie□□□□□□i□g □ □□□i□e□□□□ □□□□□□ F□re□□□re □□e □ree □□□i□e e□r□i□g □g□ri□□ □□□□□□□□e u□□e r□□□□e□□□□□□. □□e □ugge□□e□ e□e□□e □□e□□□□□i□g □□i□□i□e□□re□□□□□□e□□i□□uie e□e□i□e i□β□ □□□□□□□e□i□ □□rrier □ree□i□g □i□□ 93 □er□e□ □□□ur□□□□ □□□□r□i□g □□ □□□□r□i□e □□□□□□i□.

1.3 Problem Statement

□r□□ □□□ □□□ur□e □i□g□□i□□□ □□□□□i□ □□□□□□□e□i□ □i□e□e□□□ □e□□ □□i□ig□i□□□□□□□□i□e□□re□e□i□g □□□ □i□□i□□. □□e □i□g□□i□□□□□□□□e□i□ □□□ee□ □□e□ □□ □e r□□er□e□□□u□□ □□□□□. □□e re□u□□□□□e □□□□e□e □□□□□ □u□□ (□□□) e□□ □□□ e u□i□g □□ □eri□e □e□e □□r□□eri□i□. □□i□ r□e□ure i□ □□□□□□□□i□e□□□□□u□i□g□□□□re□u□□i□ □□e□□i□re□ei□i□g r□□er □er□□□. □□ □□□i□□□□□□ere re □□□e i□□□□□□□□□□□e□i□□□□□re □□□□e□e□□□e □□ □□□□ e□. □□□□re□u□□i□ i□ □r□□□□□ir□u□□□□□e□□□□e □□u□□i□g□□□i□ □□r□□□□ □□□ □□□□e □□e □e□□□□i□□. □□□□re□u□□□□e re□uire□e□□□r □i□□er □□□□□□er □□er□□i□e □e□e□i□□□e□□□□□□ri□e□. □□i□re□e□r□□i□□□□□□□ru□□

[illegible]

1.4 Aim of the Thesis

[illegible]

1.5 Thesis challenges

1[illegible]

2[illegible]

1.6 Thesis Outlines

□e □e i□□□□i□□□ur □□i i□□□□□□□□er□i□□□i i□□ □□e ir□□□□□er □□ □e□ri□e□□e□□□□

Chapter Two□re□e□□□e□rei□□□□□□gr□u□□r □e u□i□i□e□e□□i□ue□□□ □e□e□□□□□□e□i□□□□e□□□□e □□□□□□ge□□□□i□□□□□□ge□□u□i□g e□□□ □□e□□e□e□e□□i□ue□.

Chapter Three: □□u□r□e □e □r□□□e□□□□e□□□□e□e□□i□□.

Chapter Four: □e□ri□e□□e e□eri□e□□□□□□re □□□u□e□□e□□u□e □e □r□□□e□□□□e□□□□e□□□□i□□e □e □□□□□e□i□□□□i□□□r□i□□□i i□□□□e re□u□□□□□e□e□r□□ □e□e e□eri□e□□□.

Chapter Five□□□□□u□i□□□□□□□i□□□□e □ugge□i□□□□r Fu□ure □ □r□.