APPLICATION OF PATTERN RECOGNITION TO EARLY DETECTION OF SYMPTOMS OF SICKLE CELL PATIENT CRISIS

BY

ADEGOKE TOLUWANI B.

(151789)

BEING A PROJECT REPORT SUBMITTED

TO

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

FACULTY OF ENGINEERING AND TECHNOLOGY

LADOKE AKINTOLA UNIVERSITY OF TECHNOLOGY

OGBOMOSO, NIGERIA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF BACHELOR OF TECHNOLOGY (B.TECH)

DEGREE IN COMPUTER SCIENCE

FEBRUARY 2022

CERTIFICATION

This project report with the title "**Application of Pattern Recognition to Early Detection of Symptoms of Sickle Cell Crisis"** submitted by ADEGOKE Toluwani B. was carried out in the Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso.

_____ _____

Prof. A.O. Afolabi Date
(Project Supervisor)

_____ _____

Prof. J.B. Oladosu Date
(Head of Department)

DEDICATION

This Project is dedicated to Almighty God, the creator of all and the knower of all things. If not for His help, this work would not have been successful. All glory, honor and adoration be unto His name.

ACKNOWLEDGEMENTS

I would like to acknowledge God, my Creator for bringing me this far in this journey because without Him, I am nothing.

My supervisor, Professor Afolabi has been really amazing and supportive, so much thanks to him for his love, support and fatherly care. To my H.O.D, Prof. J.B. Oladosu, I am grateful for the privilege and opportunity you've granted me, thank you sir. I am also grateful to every lecturer in the department, without their knowledge and wealth of experience, I wouldn't have come this far. Professor Justice has been a father, he stretched me and made me realize how much I could achieve no matter the time limit, thank you so much sir.

Abdullah, Taiwo and Gabriel have been really awesome, they made working on the project nice, it was such an honor to lead a group of talented individuals in the course of the project.

I would also like to acknowledge Mr. and Mrs. Adegoke, together with my siblings (Oluwatosin, Miracle, Samuel, Esther) who believed in me more than I did in myself. I am also grateful to the Zion family. Clement, for how they were able to contribute one way or the other to the success of my journey in LAUTECH.

TABLE OF CONTENTS

CERTIFICATION **2**

DEDICATION **3**

ACKNOWLEDGEMENTS **4**

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT **7**

CHAPTER ONE **8**

INTRODUCTION **8**

1.1 Statement of the problem 8
1.2 Background of the Study 9
1.3 Significance of the Study 10
1.4 Aims and Objectives 10
1.5 Methodology 11
1.6 Limitations of Study 12
1.7 Definition of Terms 12

CHAPTER TWO **7**

LITERATURE REVIEW **7**

2.1 Pattern Recognition 7
2.2 Pattern Recognition Techniques used in Medical Diagnosis 8
2.3 Application of Early Detection of Sickle Cell Crisis System 14
2.4 Overview of Sickle Cell Crisis 22
2.4.1 Types of Sickle Cell Crisis 22
2.4.2 Symptoms of Crisis in Sickle Cell Patients 23
2.5 Early Detection 25
2.6 Related Works 25

CHAPTER THREE **31**

METHODOLOGY **31**

3.1 Process and Methods 31

3.2 Dataset 41
3.3 Dataset Visualization 42
3.4 Feature Selection 46
3.5 Classification 48
3.6 Model Evaluation 50
3.7 Design 51
3.7.1 State Diagram 51
3.7.2 Communication Diagram 52
3.7.3 UML Sequence Diagram 53
3.7.4 UML Activity Diagram 54
3.7.5 UML Use Case Diagram 55
3.7.6 UML Data Flow Diagram 56
3.7.7 UML Class Diagram 56
3.7.8 HIPO Diagram 57
3.7.9 Flowchart Diagram 62
3.8 Backend and Frontend Development 63
3.8.1 Backend Development 63
3.8.2 Frontend Development 64
3.9 Specification and Configuration 65

CHAPTER FOUR **66**

RESULTS AND DISCUSSION **66**
4.1 Performance Metrics for Classification Tasks 66
4.2 Performance of Each Model 68
4.3 Web Interface 76

CHAPTER FIVE **81**

CONCLUSION AND RECOMMENDATION **81**
5.1 Conclusion 81
5.2 Recommendation 81
REFERENCES 83

LIST OF FIGURES

Fig 2.1: A BN for Pneumonia 10

Fig 2.2: Support Vector Machine 12

Fig 2.3: Decision Tree Classifier 13

Fig 2.4: (a) Normal Blood Cell Image (b) Anemic Blood Cell Image 30

Fig.2.5: Objects in cluster after K-Means Clustering based segmentation 31

Fig.2.6: Objects in cluster after Fuzzy C-Means Clustering based segmentation 31

Fig 2.7: Sickle cell admit patient diagnosis pathway and EMR classification process 33

Fig 2.8: Receiver Operating Characteristic (ROC) curve for the ACS class. 34

Fig 3.1: Multinomial Logistic Regression Classification 37

Fig 3.2: Support Vector Machine 38

Fig 3.3: Decision Tree Classifier 39

Fig 3.4: Naive Bayes Classifier 40

Fig 3.5: Random Forest Classifier 42

Fig 3.6: Multilayer Perceptron Classifier 43

Fig 3.7: K Nearest Neighbour Classifier 44

Fig 3.8: A View of the Dataset 46

Fig 3.9: Profile of Sickle Cell Crisis Dataset 47

Fig 3.10: Profile of Sickle Cell Crisis Dataset 2 48

Fig 3.11: Bar Chart Showing the number of records missing from each column 49

Fig 3.12: Heat Map showing the correlation between features in the data. 50

Fig 3.13: Feature Selection Process 52

Fig 3.14:10-Fold cross validation of the blended and tuned model for Model 1 53

Fig 3.15:10-Fold cross validation of the blended and tuned model for Model 2 54

Fig 3.16: State Diagram 55

Fig 3.17: Communication Diagram 56

Fig 3.18: Sequence Diagram 57

Fig 3.19: Activity Diagram 58

Fig 3.20: Use Case Diagram 59

Fig 3.21: Data Flow Diagram 60

Fig 3.22: Class Diagram 61

Fig 3.23: Hierarchical Chart 63

Fig 3.24: Authentication IPO 64

Fig 3.25: Symptom/Diagnosis IPO 65

Fig 3.26: Predict Crisis IPO 66

Fig 3.27: Flow Chart Diagram for the system 67

Fig 3.28: the code script behind the whole experiment process 68

Fig 3.29: Testing Interface for the first model 69

Fig 3.30: Testing Interface for the second model 70

Fig 4.1: Performance of first Logistic Regression Model 73

Fig 4.2: Performance of first MLP Model 74

Fig 4.3: Performance of the blended First Model 75

Fig 4.4: Performance of The first tuned model 76

Fig 4.5: Performance of The second Catboost model 77

Fig 4.6: Performance of The second GBC model 78

Fig 4.7: Performance of The second MLP model 79

Fig 4.8: Performance of The second blended model 79

Fig 4.9: Performance of The second tuned model 80

Fig 4.10: Prediction Interface for Medical Practitioners 81

Fig 4.11: About Crisis Page 82

Fig 4.12: Prediction Interface for Patients and General Public 83

Fig 4.13: About Crisis Page for Patients and General Public 84

ABSTRACT

Pattern recognition is not a new field of research, actually, theories and techniques. It has been developing for a long time. While with the fast advancement of computer architecture, machine learning, and computer vision, computational complexity is possible to be dealt with and more and more new ways of thinking are brought into the research of pattern recognition. In this report, we'd like to introduce the concept and methods of pattern recognition as well as apply it to the early detection of symptoms of sickle cell crisis.