EFFECTIVE ARCHITECTURE FOR CLINICAL DECISION SUPPORT SYSTEM

Kalutharage Chamlini Vidyarthi Dayathilake

199315B

Master of Science in Computer Science

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

July 2022

EFFECTIVE ARCHITECTURE FOR CLINICAL DECISION SUPPORT SYSTEM

Kalutharage Chamlini Vidyarthi Dayathilake

199315B

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science.

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

July 2022

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

……………………………………… ……………………………….

K.C.V. Dayathilake Date

The above candidate has carried out research for the Masters thesis/ dissertation under my supervision.

……………………………………… ……………………………….

Prof. Indika Perera Date

ABSTRACT

The Healthcare domain is a very sensitive domain where the direct stakeholders are patients in the public, which makes a healthcare system directly dealing with patients' lives. There could be heard of several cases in a year where it leads to critical damage or even loss of life, because of misdiagnosis or delays in the diagnosis process or the delay or ignorance of new treatment methods. A clinical decision support system facilitating support for diagnosis and therapeutic decisions could greatly help healthcare professionals when identifying diseases by going through patients' biometrics, life cycle, and symptoms, and when deciding on necessary clinical tests to arrive at confirmation of the diagnosis and decide on treatment methods. The main focus of this study is mapping the real-world diagnosis process to a digitalized system. Senior clinicians use their own experience to derive medical diagnoses accurately. This study proposes an architecture for an evidence-based clinical decision support system, where the system infers knowledge from past knowledge, using machine learning algorithms, and use for future predictions, which could infer and use the medical incidents of the past for future diagnosis, just like experienced doctors. In a practical scenario, diagnosis of disease happens step by step, going through several stages starting from an initial level and digging deeper. To incorporate this behavior, a layered knowledge modeling system is proposed with an ensemble classifier of Random Forest classifier, Support Vector Machine, and Naïve Bayes classifier, and organized into a tree structure based on disease classification hierarchy. Additionally, the proposed system provides feedback and suggestions for clinical tests using feature selection, and a rationale for the diagnosis derived by incorporating explainable machine learning concepts.

Keywords: Clinical decision support system architecture, CDSS architecture, Machine learning, Layered architecture for CDSS, Digital medicine, Disease classification

ACKNOWLEDGEMENTS

My sincere gratitude is extended to Prof. Indika Perera, my supervisor, Head of the Department of Computer Science and Engineering, Faculty of Engineering, University of Moratuwa, for the valuable guidance and support throughout this journey.

I would like to thank my family members and friends for the motivation and continuous support provided throughout this period of my life.

My appreciation goes to all my batch mates in the MSc batch and colleagues at my workplace for the assistance extended in managing work.

TABLE OF CONTENTS

DECLARATION i

ABSTRACT ii

ACKNOWLEDGMENTS iii

LIST OF FIGURES vi

LIST OF TABLES ix

LIST OF APPENDICES ix

INTRODUCTION 1

1.1 Background 1

1.2 Problem 3

1.3 Objectives 4

1.4 Scope 5

1.5 Proposed solution 5

1.6 Outline 8

LITERATURE REVIEW 9

2.1 Research carried out under architecture for CDSS 9

2.2 Research carried out under knowledge engineering 11

2.3 Machine learning approaches 16

METHODOLOGY 20

3.1 Dataset 21

3.2 Knowledge engine 21

3.2.1 Layered Knowledge Models 22

3.2.2 Train knowledge model by machine learning algorithms 26

3.2.3 Generate result in an ensemble classifier 27

3.2.4 Deriving rationale for the decision 31

3.3 CDSS interfaces 33

IMPLEMENTATION 35

4.1 Modeling layered knowledge engine 35

4.2 Implementation of Training Process of CDSS 37

4.3 Implementation of diagnosis support/ Single instance classification 46

4.4 Implemented Prototype – User Interfaces and Sample Scenarios 56

EVALUATION 62

5.1 Clinical-diagnosis-result-based evaluation 62

5.2 Performance-based evaluation 66

CONCLUSION 74

6.1 Research contribution 74

6.2 Limitations and future work 76

REFERENCES 79

Appendix A : CD/DVD

LIST OF FIGURES

Figure 2.3.1 Equation of posterior probability in Naïve Bayes 16

Figure 2.3.2 Support Vector Machine 17

Figure 2.3.3 Decision Tree 18

Figure 3.2.1 Service architecture of the CDSS 22

Figure 3.2.1.1 Disease hierarchy of circulatory system according to ICD-11 (not the full hierarchy) 24

Figure 3.2.1.2 Graphical representation of the disease hierarchy and traversing through the tree structure 25

Figure 3.2.2.1 High-level architecture of the knowledge model training 26

Figure 3.2.3.1 Generate result in an ensemble classifier 30

Figure 4.1.1 Code snippet of Node object 35

Figure 4.1.2 Snippet of node hierarchy JSON 36

Figure 4.2.1 Rest API for training classifier for a single node 38

Figure 4.2.2 Rest API for training classifier for multiple nodes 38

Figure 4.2.3 Base folder for a node 38

Figure 4.2.4 TrainingInputDTO 39

Figure 4.2.5 TrainingResultDTO 39

Figure 4.2.6 Implementation of training classifier for a single node 44-45

Figure 4.2.7 Implementation of comparing the evaluation summaries of old and new models before saving new model 46

Figure 4.3.1 REST API for classifying a single instance 47

Figure 4.3.2 ClassifierInputDTO 47

Figure 4.3.3 CollectiveClassifierResultDTO 48

Figure 4.3.4 Implementation of single instance classification 50-51

Figure 4.3.5 Implementation of the calculation of ensemble classifier result 54-55

Figure 4.4.1 CDSS (Prototype implementation) - Home Page 56

Figure 4.4.2 CDSS (Prototype implementation) – Diagnosis Support Window 57

Figure 4.4.3 CDSS (Prototype implementation) – Training Window 57

Figure 4.4.4 CDSS (Prototype implementation) – System Evaluation Window 58

Figure 4.4.5 CDSS (Prototype implementation) – Single Node Evaluation Window 58

Figure 4.4.6 Sample scenario 1 – Heart Attack Negative 59

Figure 4.4.7 Sample scenario 2 – Heart Attack Negative 59

Figure 4.4.8 Sample scenario 3 – Heart Attack Positive 60

Figure 4.4.9 Sample scenario 4 – Heart Attack Positive 60

Figure 4.4.10 Sample scenario 5 – Missing mandatory information 61

Figure 5.1.1 Graphical visualization of result comparison of classifiers 64

Figure 5.2.1 Graphical representation of response time values for the training of a given node 69

Figure 5.2.2 Server resource usage by training process by each classifier – Root Node 69

Figure 5.2.3 Server resource usage by training process by each classifier – Node Heart Diseases 70

Figure 5.2.4 Graphical representation of response time values for classifying a given instance 71

Figure 5.2.5 Server resource usage by the process for classifying instances given by each classifier 72

LIST OF TABLES

Table 5.1.1 Result comparison of classifiers for heart disease classification 64

Table 5.2.1 Response time values for the training of a given node 68

Table 5.2.2 Response time values for classifying a given instance 71

LIST OF APPENDICES

Appendix A CD/DVD