Contents

Abstract iii

Zusammenfassung v

List of Figures ix

List of Tables xi

List of Abbreviations xiii

1 Introduction 1
- 1.1 Study Background and Motivation . . . 1
- 1.2 Research Questions and Objectives . . . 5
- 1.3 Conceptual Architecture Design . . . 10
- 1.4 Research Achievements and Contribution . . . 12
- 1.5 Structure of the Thesis . . . 16

2 Fundamentals and Related Work 19
- 2.1 Statistics in Diagnostic Medicine . . . 19
- 2.2 Supervised ML Techniques . . . 23
- 2.3 Classification Model Evaluation and Selection . . . 30
- 2.4 Ensemble Methods for Classifiers . . . 39
- 2.5 Probability Calibration Methods . . . 43
- 2.6 Data Visualization . . . 49

3 Empirically-Based Disease Prediction Model 51
- 3.1 System Diagnostic Procedure . . . 51
- 3.2 Classification Model Selection and Calibration . . . 55
- 3.3 Ensemble Method for Binary Classification Tasks . . . 58
- 3.4 Probability Inspection Table for Local Validation . . . 59
- 3.5 Validation of the Classification Model . . . 61
- 3.6 Feasibility of the Disease Classification Model . . . 63

4 Interactive Visual Inspection of Research Findings 65
- 4.1 Symptom Patterns of Different Rare Disease Groups . . . 66
- 4.2 System Diagnosis Distribution with Heatmaps . . . 75
- 4.3 Advantages of Diagnostic Data Visualization . . . 81

5 Dynamic Adaptive Decision-Making Mechanism 83
- 5.1 Dimension Reduction Analysis . . . 84

5.2 Dynamic Adaptive Questioning Concept 89
5.3 Validating the Dynamic Adaptive Questioning Method 93
5.4 Discussion of Dynamic Diagnosing Method 95

6 Discussion and Conclusion **97**
6.1 Discussion of the Research Findings 98
6.2 Research Limitations 101
6.3 Research Prospects and Future Work 104
6.4 Summary of the Research 107

Appendices **109**

A Rare Disease Data Acquisition **111**

B Data Storage Structure **115**

C The Diagnosis–Question Correlation Ranking **119**

D Publications in the Medical Field **121**

E Questionnaire Content **127**

Bibliography **139**

Curriculum Vitae **159**

List of Scientific Publications **161**