# TABLE OF CONTENTS

DEDICATION v

ACKNOWLEDGEMENTS vi

ABSTRACT vii

LIST OF TABLES xii

LIST OF FIGURES xv

LIST OF ABBREVIATIONS xvi

CHAPTER 1: INTRODUCTION 1

1.1 Background of the Study 1

1.2 Problem Statement 3

1.3 Aim and Objectives 5

1.4 Research Questions (IF ANY) 5

1.5 Scope of the Study 6

1.6 Significance of the Study 6

1.7 Structure of the Study 8

CHAPTER 2: LITERATURE REVIEW 11

(Sample Sub-sections)

2.1 Introduction 11

2.2 Data Analytics in Healthcare 12

2.3 Data Mining in Diagnosis of Chronic Diseases 14

2.4 Data Mining in Prediction of Common Types of Cancer 16

2.5 Predictive Modeling in Breast Cancer Diagnosis on Different Datasets 17

2.5.1 WBC Datasets 18

2.5.2 SEER Breast Cancer Datasets 20

2.5.3 Other Breast Cancer Datasets 21

2.6 Visual Analytics in Healthcare 23

2.7 Related Research Publications 24

2.8 Discussion 26

2.9 Summary 29

CHAPTER 3: RESEARCH METHODOLOGY 30

(Sample Sub-sections)

3.1 Introduction 30

3.2 Research Methodology 30

3.2.1 Data Selection 32

3.2.2 Data Pre-processing 33

3.2.3 Data Transformation 34

3.2.4 Interactive Visual Analytics 34

3.2.5 Class Balancing 36

3.2.6 Data Mining 37

3.2.7 Interpretation/Evaluation 38

3.3 Proposed Method (Classification) 41

……

……

……

3.4 Summary 49

CHAPTER 4: ANALYSIS 50

(Sample Sub-sections)

4.1 Introduction 50

4.2 Dataset Description 51

4.3 Data Preparation 54

4.3.1 Elimination of Variables 54

4.3.2 Transformation into Categorical Variables 54

4.3.3 Identification of missing values 54

4.3.4 Univariate analysis 56

4.3.5 Treatment of missing values 64

4.3.6 Splitting of original dataset 65

4.4 Exploratory Data Analysis (Bivariate analysis) 65

4.4.1 Chi-square test 66

……

……

4.5 Data Visualization 87

……

……

……

4.9 Summary 112

CHAPTER 5: RESULTS AND DISCUSSIONS 114

(Sample Sub-sections)

5.1 Introduction 114

5.2 Interpretation of Visualizations 115

5.3 Evaluation of Sampling Methods and Results 119

……

……

……

5.4 Testing on Validation Dataset 142

……

……

……

5.6 Summary 149

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS 151

6.1 Introduction 151

6.2 Discussion and Conclusion 151

6.3 Contribution to knowledge 153

6.4 Future Recommendations 153

REFERENCES 154

APPENDIX A: RESEARCH PROPOSAL 167

(Sample Sub-sections)

APPENDIX B: ETHICS FORMS 179