

TABLE OF CONTENT

COVER PAGE

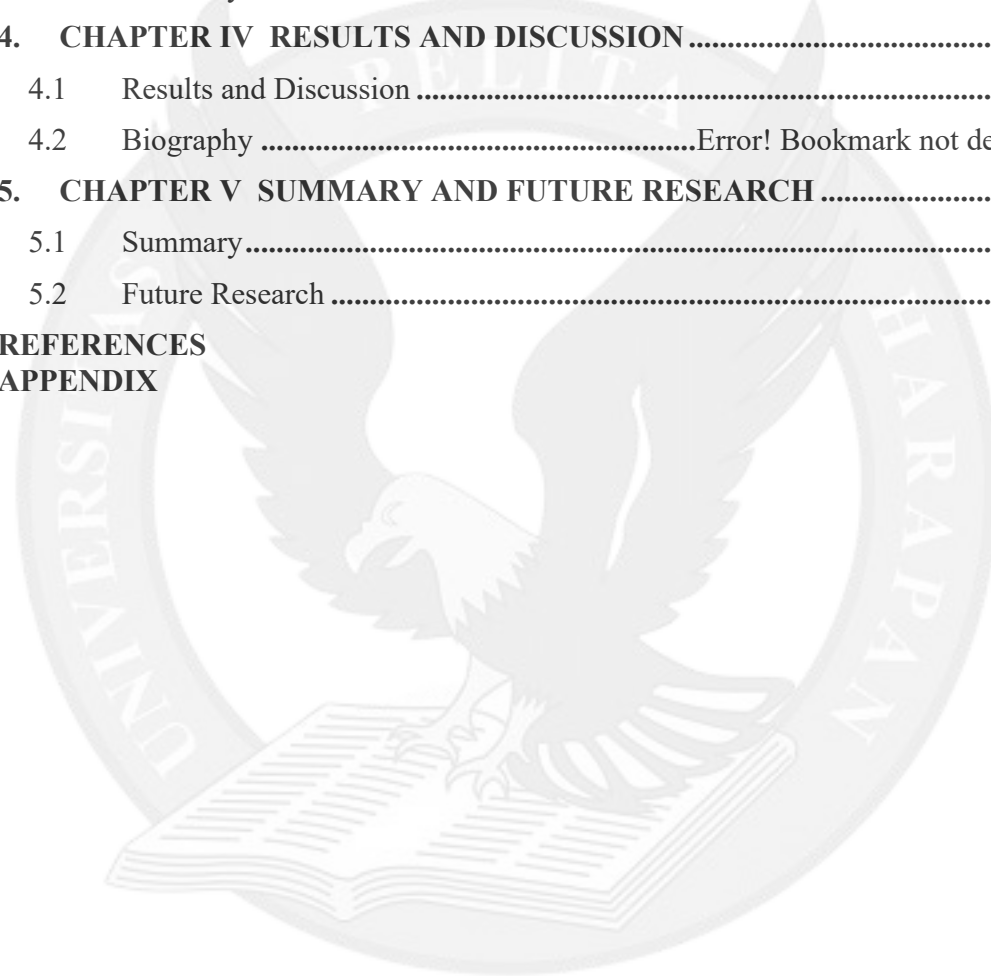
ORIGINALITY STATEMENT OF THE THESIS

THESIS SUPERVISOR'S APPROVAL

THESIS EXAMINATION COMMITTEE

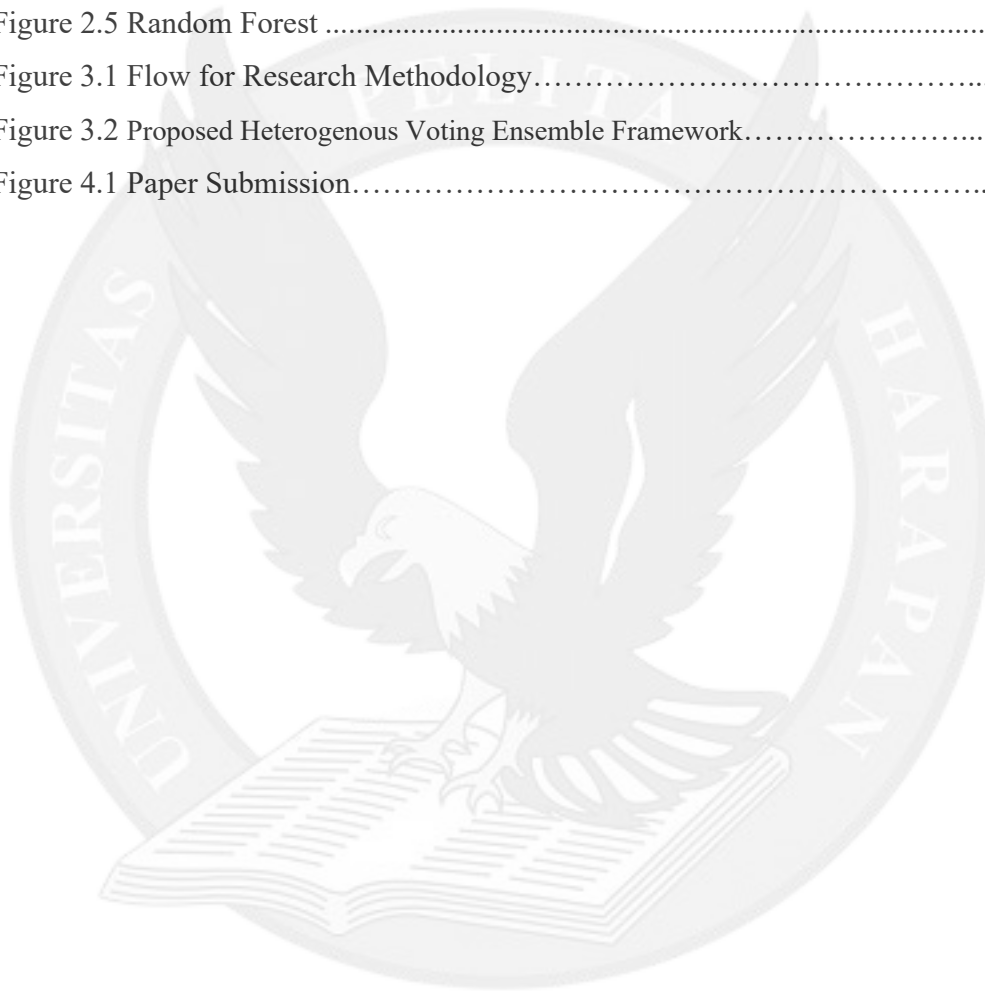
| | |
|--|-------------|
| ABSTRACT | v |
| FOREWORD | ix |
| TABLE OF CONTENT | xi |
| LIST OF FIGURES | xiii |
| LIST OF TABLES | xiv |
| 1. CHAPTER I INTRODUCTION | 1 |
| 1.1 Background..... | 1 |
| 1.2 Problem Identification | 2 |
| 1.3 Problem Limitation..... | 3 |
| 1.4 Problem Definition | 3 |
| 1.5 Research Purpose..... | 3 |
| 1.6 Outline of the Thesis..... | 4 |
| 2. CHAPTER II THEORITICAL REVIEW | 5 |
| 2.1 Software Defect | 5 |
| 2.2 Heterogenous Parallel Ensemble | 6 |
| 2.3 Voting | 8 |
| 2.4 SMOTE and LASSO for Feature Selection..... | 10 |
| 2.4.1 Synthetic Minority Over-sampling Technique (SMOTE)..... | 10 |
| 2.4.2 Least Absolute Shrinkage and Selection Operator (LASSO)..... | 11 |
| 2.5 NASA Software Defect Dataset | 12 |
| 2.6 Classification Methods | 13 |
| 2.6.1 Support Vector Machine..... | 13 |
| 2.6.2 Decision Trees | 14 |
| 2.6.3 Gaussian Process | 16 |
| 2.6.4 k-Nearest Neighbors | 17 |
| 2.6.5 Random Forest..... | 18 |
| 2.6.6 Gaussian Naïve Bayes | 19 |

| | |
|---|------------------------------|
| 3. CHAPTER III RESEARCH METHODOLOGY | 21 |
| 3.1 Overview..... | 21 |
| 3.2 Data Collection | 22 |
| 3.3 Pre-processing and Feature Selection..... | 23 |
| 3.4 Data Splitting..... | 24 |
| 3.4 Modelling and Ensembling..... | 24 |
| 3.5 Evaluation | 26 |
| 3.6 Summary..... | 26 |
| 4. CHAPTER IV RESULTS AND DISCUSSION | 27 |
| 4.1 Results and Discussion | 27 |
| 4.2 Biography | Error! Bookmark not defined. |
| 5. CHAPTER V SUMMARY AND FUTURE RESEARCH | 40 |
| 5.1 Summary..... | 40 |
| 5.2 Future Research | 41 |
| REFERENCES | |
| APPENDIX | |



LIST OF FIGURES

| | |
|---|----|
| Figure 2.1 Heterogenous Parallel Ensemble Flow..... | 9 |
| Figure 2.2 SVM Hyperplane..... | 14 |
| Figure 2.3 Decision Trees..... | 15 |
| Figure 2.4 k-Neighbors..... | 18 |
| Figure 2.5 Random Forest | 19 |
| Figure 3.1 Flow for Research Methodology..... | 21 |
| Figure 3.2 Proposed Heterogenous Voting Ensemble Framework..... | 25 |
| Figure 4.1 Paper Submission..... | 27 |



LIST OF TABLES

| | |
|---|----|
| Table 3.1 MDP NASA Software Defect Datasets | 23 |
| Table 3.2 Dataset Description..... | 23 |

