CONTENTS

SELF DECLARATION AGAINST PLAGIARISM ..... **iii**
ABSTRACT ..... **iv**
ABSTRAK ..... **v**
DEDICATION ..... **vi**
ACKNOWLEDGMENTS ..... **vii**
CONTENTS ..... **viii**
LIST OF TABLES ..... **x**
LIST OF FIGURES ..... **xi**
LIST OF TERMS ..... **xii**
LIST OF NOTATIONS ..... **xiii**
CHAPTER 1 INTRODUCTION **1**
1.1 Introduction 1
1.2 Theoretical Framework 2
1.3 Conceptual Framework/Paradigm 3
1.4 Statement of the Problem 4
1.5 Objective and Hypotheses 4
1.5.1 Objective *4*
1.5.2 Hypotheses *4*
1.6 Assumption 4
1.7 Scope and Delimitation 4
1.8 Significance of the Study 5
CHAPTER 2 REVIEW OF LITERATURE AND STUDIES **6**
2.1 Data Balancing for Drug Target Interaction 6
2.2 SMILES for Drug Representation 7
2.3 Random Forest on Drug Target Interaction Prediction 8
CHAPTER 3 RESEARCH METHODOLOGY **9**
3.1 Dataset 9
3.2 Feature Extraction 10
3.3 Data Pre-processing 10
3.4 Firefly Algorithm Under-sampling (FAUS) 11
3.5 Random Forest Algorithm 13
3.6 Analysis of Evaluation Method 15
CHAPTER 4 PRESENTATION, ANALYSIS, AND INTERPRETATION OF DATA **16**
4.1 Presentation of Data 16
4.1.1 Tabular Presentation *16*

4.1.2 Textual Presentation *16*
4.1.3 Graphical Presentation *17*
4.2 Result and Discussion 17
4.2.1 Performance with No Balancing Method (NBM) *17*
4.2.2 Performance with Random Under-sampler (RUS) *18*
4.2.3 Performance with All KNN-Under-sampling (KUS) *18*
4.2.4 Performance with Firefly Algorithm Under-sampling *19*
4.3 Performance Comparisons 27
CHAPTER 5 CONCLUSION AND RECOMMENDATIONS 29
5.1 Conclusions 29
5.2 Recommendations 29
BIBLIOGRAPHY 30
APPENDIX A 32
A.1 The Drug Features 32
A.2 The Target Features 34
APPENDIX B 41