

TABLE OF CONTENTS

TABLE OF CONTENTS.....	I
ABSTRACT.....	III
INTISARI	IV
CHAPTER I INTRODUCTION.....	1
1.1 BACKGROUND	1
1.2 RESEARCH PROBLEM.....	3
1.3 RESEARCH SCOPE	3
1.4 RESEARCH OBJECTIVE.....	4
1.5 RESEARCH BENEFIT.....	4
1.6 THESIS ORGANIZATION	4
CHAPTER II LITERATURE REVIEW	6
CHAPTER III THEORETICAL BASIS	11
3.1 POLYCYSTIC OVARIAN SYNDROME	11
3.2 GRADIENT BOOSTED DECISION TREE	12
3.3 SMOTE-TOMEK.....	16
3.4 FEATURE SELECTION ALGORITHM.....	18
3.4.1 ANOVA <i>F</i> -test	19
3.4.2 Kendall's <i>Correlation Coefficient</i>	20
3.5. GRID SEARCH	21
3.6 MODEL EVALUATION METRICS	22
3.6.1 <i>Accuracy</i>	23
3.6.2 <i>Precision</i>	24
3.6.3 <i>Recall</i>	24
3.6.4 <i>F1 Score</i>	24
CHAPTER IV RESEARCH METHODOLOGY	25
4.1 RESEARCH OVERVIEW	25
4.2 DATA COLLECTION.....	25
4.3 PROPOSED MODEL.....	27
4.4 DATA PREPROCESSING	28
4.5 MODEL BUILDING.....	30
4.6 MODEL EVALUATION	32
CHAPTER V IMPLEMENTATION	34

5.1 LIBRARY PREPARATION	34
5.2 DATA PREPARATION	35
5.2.1 Data Preprocessing	35
5.2.2 Data Resampling with SMOTE-Tomek.....	37
5.3 MODEL BUILDING.....	38
5.3.1 Rank Feature Importance	38
5.3.2 Data Splitting.....	40
5.3.3 Hyperparameter Optimization.....	41
5.3.4 Gradient Boosted Decision Tree Model	42
5.4 MODEL EVALUATION	43
CHAPTER VI RESULTS AND DISCUSSION	45
6.1 CLASSIFIERS WITH FEATURES RANKED BY ANOVA F-TEST	45
6.2 CLASSIFIERS WITH FEATURES RANKED BY KENDALL'S CORRELATION COEFFICIENT	47
CHAPTER VII CONCLUSION AND SUGGESTION	50
7.1 CONCLUSION	50
7.2 SUGGESTION	51
REFERENCES.....	52