

TABLE OF CONTENTS

COVER

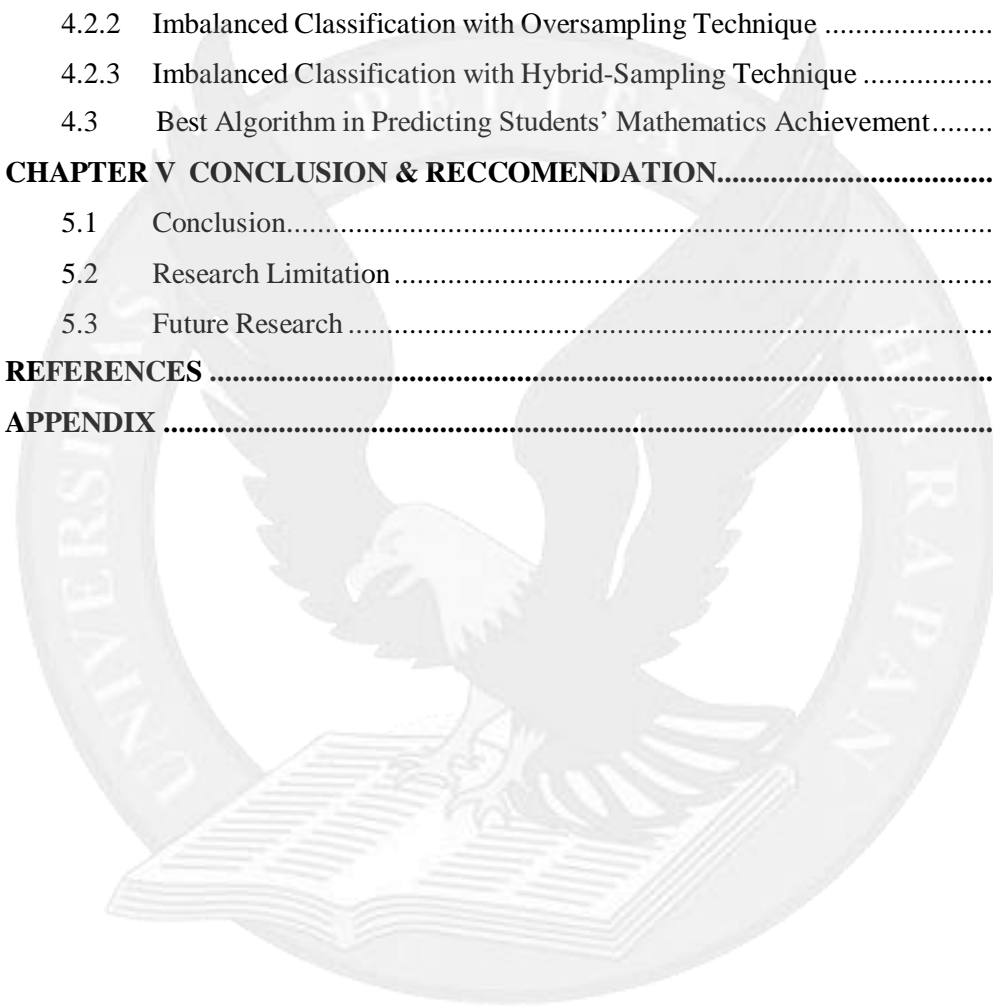
ORIGINALITY STATEMENT OF THE THESIS

THESIS SUPERVISOR'S APPROVAL

THESIS EXAMINATION COMMITTEE

ABSTRACT	v
ABSTRAK.....	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	ix
LIST OF FIGURES	xi
LIST OF TABLES	xii
CHAPTER I INTRODUCTION.....	1
1.1 Background	1
1.2 Problem Identification.....	9
1.3 Problem Limitation	10
1.4 Research Problem	11
1.5 Research Objectives.....	11
1.6 Research Outline.....	12
CHAPTER II LITERATURE REVIEW.....	14
2.1 Students' Academic Achievement	14
2.2 Mathematics Achievement in Elementary School	15
2.3 Factors that Influence Academic Achievement in Mathematics	17
2.4 Data Mining.....	18
2.5 The Role of Data Mining in Education	20
2.6 Predicting Students' Academic Achievement in Mathematics using Data Mining Techniques	21
2.7 Research Gap and Conceptual Framework.....	24
CHAPTER III RESEARCH METHODOLOGY.....	26
3.1 Research Design	27
3.2 Environment	27
3.3 Data Collection.....	28
3.4 Data Mining.....	28

3.4.1	Data Selection	28
3.4.2	Data Pre-Processing	29
3.4.3	Machine Learning Algorithm.....	33
3.4.4	Evaluation and Interpretation.....	36
CHAPTER IV RESULT & DISCUSSION		39
4.1	Classification Algorithms without Resampling Methods.....	39
4.2	Imbalanced Classification with Resampling Methods	44
4.2.1	Imbalanced Classification with Undersampling Technique.....	44
4.2.2	Imbalanced Classification with Oversampling Technique	46
4.2.3	Imbalanced Classification with Hybrid-Sampling Technique	50
4.3	Best Algorithm in Predicting Students' Mathematics Achievement.....	52
CHAPTER V CONCLUSION & RECCOMENDATION.....		57
5.1	Conclusion.....	57
5.2	Research Limitation	57
5.3	Future Research	58
REFERENCES		60
APPENDIX		A-1



LIST OF FIGURES

Figure 1. 1 Indonesian 2003-2018 PISA score trends	3
Figure 1.2 Students' grade distribution on all subjects.....	4
Figure 3. 1 Flow diagram of the Research Methodology.....	26
Figure 3. 2 Class label distribution	32
Figure 3. 3 Hold-out validation	33
Figure 3. 4 5-k fold cross validation	33
Figure 3. 5 Confusion Matrix 3x3	37
Figure 4. 1 AUC-ROC Curve of ESD.....	43
Figure 4. 2 AUC-ROC Curve of L-SVM.....	43
Figure 4. 3 AUC-ROC Curve of EBT.....	44
Figure 4. 4 AUC-ROC Curve of ERT.....	45
Figure 4. 5 AUC-ROC Curve of S-ESD	48
Figure 4. 6 AUC-ROC Curve of S-EBT	49
Figure 4. 7 AUC-ROC Curve of ERT.....	51
Figure 4. 8 Observation Result using S-ERT	55

LIST OF TABLES

Table 1. 1. Mathematics ranking of Indonesian students by PISA and TIMSS	2
Table 3. 1 Features and description.....	29
Table 3. 2 Classification Algorithm	34
Table 4. 1 10 Classification algorithms without resampling	40
Table 4. 2 Hyperparameter of ESD, L-SVM, EBT.....	41
Table 4. 3 Comparison of ESD, L-SVM, and EBT	41
Table 4. 4 Imbalanced Classification with undersampling technique (ERT).....	45
Table 4. 5 Imbalanced Classification on different SMOTE-Balanced training set	47
Table 4. 6 Hyperparameter of S-ESD and S-EBT	48
Table 4. 7 Comparison between ERT and ESD on different SMOTE balanced- training set	48
Table 4. 8 Imbalanced Classification with hybrid-sampling (S-ERT).....	50
Table 4. 9 Comparison of classification algorithms without resampling and with resampling method.....	53
Table 4. 10 Comparison between ERT and S-ERT	54