

TABLE OF CONTENTS

APPROVAL PAGE	1
PLAGIARISM STATEMENT	2
TABLE OF CONTENT	3
ABSTRACT	5
CHAPTER I INTRODUCTION	6
1.1 Background	6
1.2 Research Problem	8
1.3 Research Objectives	9
1.4 Research Scopes	9
1.5 Research Benefit	10
CHAPTER II LITERATURE REVIEW	11
CHAPTER III THEORETICAL BASIS	14
3.1 Text Mining	14
3.2 Pre-processing	15
3.2.1 Case Folding	16
3.2.2 Stop Word Removal	16
3.2.3 Special Word Removal	16
3.2.4 Tokenization	16
3.2.5 Stemming	16
3.2.6 Normalization	17
3.3 Feature Extraction	17
3.3.1 Word Embedding	17
3.3.2 Word2Vec	17
3.3.3 Contextual Word Embedding	17
3.3.4 BERT	18
3.3.4.1 IndoBERT	18
3.4 Convolutional Neural Network	19
3.4.1 Convolutional Layer	20
3.4.2 Pooling Layer	20

3.4.3 Fully Connected	20
3.5 Model Evaluation	20
CHAPTER IV RESEARCH METHODOLOGY	23
4.1 Research Description	23
4.2 Tools and Materials	24
4.3 Research Steps	25
4.3.1 Literature Review	25
4.3.2 Data Collection	26
4.3.3 Pre-Processing	26
4.3.4 Text Representation	29
4.3.5 Model Training	30
4.3.6 Model Evaluation	31
CHAPTER V IMPLEMENTATION	34
5.1 Introduction	34
5.2 Implementation Flow	34
5.2.1 Work Environment Preparation	34
5.2.2 Dataset	34
5.2.3 Dataset Processing	35
5.2.4 BERT Tokenization	35
5.2.5 Feature Extraction	36
5.2.6 CNN Model Architecture	36
5.2.7 Training Model	37
5.2.8 Evaluation and Visualization	37
CHAPTER VI RESULT AND DISCUSSION	39
6.1 Result of the Study	39
6.1.1 Training Model	39
6.1.2 Result Tuning Hyperparameter	40
6.1.3 Model Evaluation	40
6.2 Discussion	42
6.3 Proposed Improvements	42
CHAPTER VII CONCLUSION AND FUTURE WORKS	44
7.1 Conclusion	44
7.2 Future Wrok	44
REFERENCES	46