

TABLE OF CONTENT

APPROVAL SHEET	2
TABLE OF CONTENT	3
TABLE OF FIGURES	5
TABLE OF TABLES	6
ABSTRACT	7
CHAPTER I INTRODUCTION	8
1.1 Research Background	8
1.2 Research Problem	10
1.3 Research Scope	11
1.4 Research Objective	11
1.5 Research Advantage	12
CHAPTER II LITERATURE REVIEW	13
2.1 English Social Media Summarization	13
2.2 English Twitter-specific Language Model	14
2.3 Social Media Research in Indonesian	14
2.4 Summarization methods in Indonesian	16
CHAPTER III THEORETICAL BASIS	19
3.1 Neural Network	19
3.2 Sequence-to-sequence (Seq2Seq)	19
3.3 Transformer-based Language Model	20
3.3.1 IndoBART	20
3.3.2 Pegasus-xsum	21
3.3.3 mT5	21
3.4 Text Normalization	22
3.5 Sentence Vectorization	22
3.6 Semantic Search	22
3.7 ROUGE metric	23
CHAPTER IV RESEARCH METHODOLOGY	24
4.1 Research Description	24
4.2 Research Step	24
4.3 Dataset Construction	25

4.3.1 Data Collection	26
4.3.2 Data Cleaning	27
4.3.3 Fetching Thread	28
4.3.4 Data Preprocessing	28
4.3.5 Dataset Labelling	28
4.3.6 Dataset Insight	29
4.4 Fine-tuning	31
4.4.1 The Language Models	31
4.4.2 Fine-tuning Steps	32
4.5 Performance Evaluation and Model Selection	34
4.5.1 ROUGE Metric	35
4.5.2 ROUGE Scores during Fine-tuning	36
4.5.3 ROUGE Scores on All XLSum-Indonesian Test Set	37
4.5.4 Human Judgement on 10 XLSum-Indonesian Test Samples	38
4.5.5 Human Judgement on 10 TweetSum-Indonesian Samples	38
4.5.6 Model Selection	39
4.6 Designing Twitter Summarization Algorithm	40
4.6.1 Data Retrieval	40
4.6.2 Pre-processing	41
4.6.3 Text Normalization	41
4.6.4 Semantic Search	42
CHAPTER V IMPLEMENTATION	47
5.1 Building Website	47
CHAPTER VI RESULT AND DISCUSSION	51
6.1 ROUGE Score on All TweetSum-Indonesian Data	51
6.2 Human Judgement on Real-World Twitter Thread	52
6.3 Conclusion	54
CHAPTER VII CONCLUSION	58
LIST OF REFERENCE	59
APPENDIX	62

TABLE OF FIGURES

Figure 3.1 Transformer model architecture.....	20
Figure 3.2 Pegasus model architecture	21
Figure 3.3 T5 model concept.....	21
Figure 4.1 Key steps of experiment.....	25
Figure 4.2 Dataset Construction	26
Figure 4.3 Thread collection pseudocode	28
Figure 4. 4 Comparison of colloquial lexicons contained in the TweetSum- Indonesian vs. XLSum dataset	31
Figure 4.5 ROUGE-1 score for mt5-base, pegasus-xsum, and indobart-v2	36
Figure 4.6 ROUGE-2 score for mt5-base, pegasus-xsum, and indobart-v2	36
Figure 4.7 ROUGE-L score for mt5-base, pegasus-xsum, and indobart-v2.....	36
Figure 4.8 ROUGE-LSUM score for mt5-base, pegasus-xsum, and indobart-v2.	36
Figure 4.9 Twitter Summarization Algorithm.....	40
Figure 4.10 Sentence embedding using a language model	43
Figure 4.11 Cosine similarity formula	44
Figure 4.12 Illustration of vector space	44
Figure 4.13 Heatmap of Cosine Similarity Matrix	45
Figure 5.1 Fetching Twitter Thread.....	47
Figure 5.2 Thread preprocessing	48
Figure 5.3 Normalized colloquial lexicons	48
Figure 5.4 Removed unimportant sentences	49
Figure 5.5 Summary result.....	49
Figure 5.6 WordCloud	49
Figure 5.7 Bar graph of most common words in thread	50
Figure 6.1 ROUGE Score comparison with algorithm and without algorithm.....	54
Figure 6.2 Grammar score comparison.....	55
Figure 6.3 Relevance score comparison	55

TABLE OF TABLES

Table 2.1 Literature Review	17
Table 4.1 Example of kept thread and discarded thread.....	27
Table 4.2 A snippet of TweetSum-Indonesian dataset	29
Table 4.3 Noise comparison from 200 data of TweetSum-Indonesian vs. 200 training data of XLSum-Indonesian	30
Table 4.4 XLSum-Indonesian data split	32
Table 4.5 Training Hyperparameter.....	33
Table 4. 6 Maximum ROUGE score values from fine-tuning result.....	37
Table 4.7 ROUGE scores on all XLSum-Indonesian test set	37
Table 4.8 Human evaluation on 10 XLSum-Indonesian test set	38
Table 4.9 Human evaluation on 10 TweetSum-Indonesian test set	39
Table 4.10 Overall result obtained from performance evaluation on all of the cases	39
Table 4.11 Colloquial Lexicon and its equivalences	41
Table 4.12 Calculating similarity scores.....	45
Table 6.1 ROUGE Score on all TweetSum-Indonesian.....	51
Table 6.2 Results of 5 human evaluation on summary with algorithm and without algorithm	52
Table 6.3 Relevance and Grammar scores on real-world Twitter thread	53