

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION.....	1
1.1 Research Background.....	1
1.2 Research Problem Formulation.....	2
1.3 Research Purpose.....	3
1.4 Research Benefits.....	3
1.5 Scope of Problem.....	3
1.6 Writing Schematics.....	4
CHAPTER 2 LITERATURE REVIEW.....	5
2.1 Previous Studies.....	5
2.1.1 Styawati, Nirwana Hendrastuty, Auliya Rahman Isnain, Ari Yanti Rahmadhani (2021).....	5
2.1.2 Majid Rahardi, Afrig Aminuddin, Ferian Fauzi Abdullah, Rizky Adhi Nugroho (2022).....	5
2.1.3 Imam Rasyidin Muqsith Rizqi Prasetyo, Aziz Musthafa, Taufiqurrahman (2023).....	6
2.1.4 Eva Darwisah Harahap, Rakhmat Kurniawan (2024).....	6
2.1.5 Ade Dwi Dayani, Yuhandri, Gunadi Widi Nurcah (2024).....	7
2.2 Theoretical Review.....	9
2.2.1 Public Housing Savings (Tapera).....	9
2.2.2 Twitter Social Media.....	9
2.2.3 Sentiment Analysis.....	10
2.2.4 Support Vector Machine.....	10
2.2.5 Confusion Matrix.....	10
CHAPTER 3 RESEARCH METHODOLOGY.....	12
3.1 Research Time and Place.....	12
3.1.1 Time.....	12
3.1.2 Place.....	12
3.2 Needs Analysis.....	12
3.2.1 Hardware.....	12
3.2.2 Software.....	12
3.3 Research Stages.....	13

3.3.1 Business Understanding	13
3.3.2 Data Understanding	13
3.3.3 Data Preparation	14
3.3.4 Modeling	15
3.3.5 Evaluation	16
3.3.6 Deployment	16
CHAPTER 4 RESULTS AND DISCUSSION	17
4.1 Data Collection	17
4.1.1 Data Labeling	17
4.2 Preprocessing	18
4.2.1 Oversampling	21
4.2.2 Pembobotan TF-IDF	22
4.3 Modeling	22
4.4 Evaluation	22
4.4.1 Confusion Matrix	22
4.4.2 Cross Validation	23
4.4.3 Wordcloud	24
4.5 Deployment	25
CHAPTER 5 CLOSING	26
5.1 Conclusion	26
5.2 Suggestions	26
APPENDIX	30



LISTS OF TABLES

Table 2.1 Previous Studies	8
Tabel 3.1 Confusion Matrix	10
Tabel 3.1 Execution Time	12
Table 4.1 Data Collection Results Samples	17
Table 4.2 Data Labeling Examples	17
Table 4.4 Case-Folding Results Samples	18
Table 4.5 Tokenizing Samples	19
Table 4.6 Normalizing Samples	19
Tabel 4.7 Stopwords Samples	20
Table 4.8 Stemming Samples	20
Table 4.9 Data Amount after Oversampling	21
Table 4.10 TF-IDF Results Samples	22
Table 4.11 SVM Accuracy Metric	22
Tabel 4.12 K-fold Cross-Validation Results	23



TABLE OF FIGURES

Figure 3.1 CRISP-DM Diagram	13
Figure 3.2 Dirty Data Samples	14
Figure 3.3 Model SVM	15
Figure 4.1 Data Oversampling Pie-chart	21
Figure 4.2 Confusion Matrix	23
Figure 4.3 10 Frequent Words	24
Figure 4.4 Wordcloud	25

