

## DAFTAR ISI

ABSTRAK .....	i
ABSTRACK .....	ii
LEMBAR PENGESAHAN .....	iii
LEMBAR PERNYATAAN ORISINILITAS .....	iv
KATA PENGANTAR .....	v
Daftar Isi .....	vi
Daftar Gambar.....	x
Daftar Tabel .....	xi
Daftar Simbol.....	xiii
Daftar Istilah .....	xiv
Bab I   Pendahuluan .....	1
I.1   Latar Belakang .....	1
I.2   Perumusan Masalah.....	5
I.3   Tujuan Penelitian.....	5
I.4   Batasan Penelitian .....	5
I.5   Manfaat Penelitian.....	6
I.6   Sistematika Penulisan.....	6
Bab II   Tinjauan Pustaka .....	8
II.1   Penelitian Terdahulu .....	8
II.2   Dasar Teori .....	11
II.2.1   PDAM .....	12
II.2.2 <i>Web Scrapping</i> .....	12
II.2.3 <i>Natural Language Processing (NLP)</i> .....	13
II.2.4 <i>Text Mining</i> .....	13
II.2.5   Analisis Sentimen .....	14

II.2.6	<i>Supervised Learning</i> .....	14
II.2.7	<i>Support Vector Machine (SVM)</i> .....	15
II.2.8	<i>Text Pre-processing</i> .....	19
II.2.9	<i>SentiStrength</i> .....	19
II.2.10	<i>Imbalanced Class</i> .....	20
II.2.11	<i>Synthetic Minority Over-sampling Technique (SMOTE)</i> .....	21
II.2.12	<i>TF-IDF</i> .....	22
II.2.13	<i>K-Fold Cross Validation</i> .....	23
II.2.14	<i>Confusion Matrix</i> .....	24
II.2.15	<i>Receiver Operating Characteristic (ROC)</i> .....	26
II.2.16	<i>Python</i> .....	27
II.2.17	<i>FLASK</i> .....	28
BAB III	<i>Metodologi Penelitian</i> .....	29
III.1	<i>Kerangka Berpikir</i> .....	29
III.2	<i>Sistematika Penyelesaian Masalah</i> .....	30
III.3	<i>Pengumpulan Data</i> .....	33
III.4	<i>Pengolahan Data</i> .....	34
III.5	<i>Metode Evaluasi</i> .....	34
Bab IV	<i>IDENTIFIKASI DAN ANALISIS KEBUTUHAN</i> .....	36
IV.1	<i>Pemahaman Bisnis (Business Understanding)</i> .....	36
IV.2	<i>Pemahaman Data (Data Understanding)</i> .....	36
IV.3	<i>Persiapan Data (Data Preparation)</i> .....	39
IV.3.1	<i>Labelling</i> .....	39
IV.3.2	<i>Analisis Dataset</i> .....	40
IV.3.3	<i>Data Cleaning</i> .....	41
IV.3.4	<i>Spelling Correction</i> .....	45

IV.3.5 <i>Stemming</i> .....	46
IV.3.6 <i>Tokenizing</i> .....	46
IV.3.7 <i>Stopword Removal</i> .....	47
IV.3.8 <i>TF-IDF (Term Frequency-Inverse Document Frequency)</i> .....	48
IV.4 <i>Pemodelan (Modelling)</i> .....	51
IV.4.1 <i>Data Splitting</i> .....	51
IV.4.2 <i>Class Balancing</i> .....	52
IV.4.3 <i>Implementasi Algoritma Support Vector Machine (SVM)</i> .....	52
IV.5 <i>Evaluasi (Evaluation)</i> .....	55
IV.5.1 <i>Confusion Matrix</i> .....	56
IV.5.2 <i>Classification Report</i> .....	56
IV.5.3 <i>ROC Curve</i> .....	56
IV.5.4 <i>K-Fold Cross Validation</i> .....	57
IV.6 <i>Pengembangan (Deployment)</i> .....	57
IV.6.1 <i>Framework Deployment</i> .....	57
IV.6.2 <i>Arsitektur Model Deployment</i> .....	58
Bab V <b>IMPLEMENTASI DAN PENGUJIAN</b> .....	59
V.1 <i>Implementasi Algoritma Support Vector Machine</i> .....	59
V.2 <i>Evaluasi Model</i> .....	59
V.2.1 <i>Confusion Matrix</i> .....	60
V.2.2 <i>Classification Report</i> .....	61
V.2.3 <i>ROC-AUC</i> .....	63
V.2.4 <i>K-Fold Cross Validation</i> .....	65
V.3 <i>Analisis Performa Model</i> .....	66
V.4 <i>Analisis Data dengan Wordcloud dan N-Gram</i> .....	68
V.4.1 <i>Wordcloud</i> .....	68

V.7 <i>N-Gram</i> .....	69
V.8 <i>Deployment Model Machine Learning</i> .....	70
V.8.1 Analisis Hasil Prediksi .....	71
Bab VI KESIMPULAN DAN SARAN .....	73
VI.1 Kesimpulan .....	73
VI.2 Saran .....	74
Daftar Pustaka .....	75