

DAFTAR ISI

ABSTRAK	ii
<i>ABSTRACT</i>	iii
LEMBAR PENGESAHAN	iv
LEMBAR PERNYATAAN ORISINALITAS	v
Kata Pengantar	vi
Daftar Isi.....	vii
Daftar Gambar.....	xi
Daftar Tabel	xiv
Daftar Istilah dan Singkatan.....	xvi
Bab I PENDAHULUAN.....	1
I.1 Latar Belakang.....	1
I.2 Perumusan Masalah.....	6
I.3 Tujuan Tugas Akhir.....	7
I.4 Batasan Tugas Akhir	7
I.5 Manfaat Tugas Akhir.....	8
I.6 Sistematika Penulisan.....	8
Bab II TINJAUAN PUSTAKA	11
II.1 Twitter	11
II.2 <i>Natural Language Processing (NLP)</i>	11
II.3 <i>Text Mining</i>	12
II.4 Analisis Sentimen.....	12
II.5 Praproses Data	13
II.6 <i>Term Frequency – Inverse Document Frequency (TF-IDF)</i>	17
II.7 <i>Machine Learning</i>	19

II.8	Support Vector Machine (SVM)	20
II.9	<i>Confusion Matrix</i>	24
II.10	<i>Receiver Operating Characteristic Curve (ROC Curve)</i>	26
II.11	Teknik <i>Resampling Data</i>	27
II.11.1	<i>Synthetic Minority Over-Sampling Technique (SMOTE)</i>	31
II.11.2	<i>Random Undersampling (RUS)</i>	31
II.11.3	<i>Synthetic Minority Over-Sampling Technique-Edited Nearest Neighbors (SMOTE-ENN)</i>	32
II.11.4	<i>Synthetic Minority Over-Sampling Technique-Tomek Links (SMOTE-Tomek Links)</i>	33
II.12	Perbandingan Algoritma Klasifikasi	34
II.13	Penelitian Terdahulu	38
Bab III Metodologi Penelitian		43
III.1	Kerangka Pemecahan Masalah / Pengembangan Model Konseptual	43
III.2	Metode Penelitian	44
III.3	Sistematika Penyelesaian Masalah	46
III.3.1	Tahap Inisiasi	48
III.3.2	Tahap <i>Data Selection</i>	48
III.3.3	Tahap <i>Data Preprocessing</i>	49
III.3.4	Tahap <i>Data Transformation</i>	50
III.3.5	Tahap <i>Data Mining/Modelling</i>	50
III.3.6	Tahap <i>Interpretation/Evaluation</i>	50
Bab IV Analisis dan perancangan		52
IV.1	Analisis Ruang Lingkup Studi Kasus	52
IV.2	Pengumpulan Data	52
IV.3	Persiapan Data	53

IV.3.1	<i>Case Folding</i>	53
IV.3.2	Penghapusan <i>Whitespace</i>	55
IV.3.3	Penghapusan <i>Username</i>	56
IV.3.4	Penghapusan Tautan atau <i>Uniform Resource Locator (URL)</i>	58
IV.3.5	Penghapusan Tanda Baca	59
IV.3.6	Penghapusan <i>Emoji</i> dan <i>Emoticon</i>	60
IV.3.7	Tokenisasi	62
IV.3.8	Normalisasi	64
IV.3.9	Penghapusan <i>Stop Words</i>	65
IV.3.10	<i>Stemming</i>	66
IV.3.11	Pelabelan Data.....	67
IV.3.12	<i>Splitting Data</i>	72
IV.3.13	TF-IDF	75
IV.4	<i>Imbalance Handling</i>	80
IV.5	Klasifikasi Dengan Algoritma SVM	85
IV.6	Evaluasi Performa Model Klasifikasi.....	89
IV.6.1	Evaluasi Performa Model Klasifikasi Dengan <i>Confusion Matrix</i>	90
IV.6.2	Evaluasi Performa Model Klasifikasi Dengan Metrik Evaluasi.....	93
IV.6.3	Evaluasi Performa Model Klasifikasi Dengan Kurva ROC dan Nilai AUC	93
Bab V	Hasil dan Pembahasan	96
V.1	Proses Klasifikasi Menggunakan SVM.....	96
V.2	Evaluasi Performa Model Klasifikasi.....	98
V.2.1	Evaluasi Performa Model Menggunakan <i>Confusion Matrix</i>	98
V.2.1.1	Evaluasi Menggunakan <i>Confusion Matrix</i> Pada Rasio 70:30.....	98
V.2.1.2	Evaluasi Menggunakan <i>Confusion Matrix</i> Pada Rasio 75:25.....	100

V.2.1.3	Evaluasi Menggunakan <i>Confusion Matrix</i> Pada Rasio 80:20.....	101
V.2.2	Evaluasi Performa Model Menggunakan Metrik Evaluasi.....	102
V.2.3	Evaluasi Performa Model Menggunakan Kurva ROC dan Nilai AUC 109	
V.2.3.1	Evaluasi Performa Model Menggunakan Kurva ROC dan Nilai AUC Pada Rasio 70:30.....	110
V.2.3.2	Evaluasi Performa Model Menggunakan Kurva ROC dan Nilai AUC Pada Rasio 75:25.....	111
V.2.3.3	Evaluasi Performa Model Menggunakan Kurva ROC dan Nilai AUC Pada Rasio 80:20.....	112
V.3	Visualisasi <i>Word Cloud</i> Terhadap Komentar Positif dan Negatif.....	115
Bab VI	Kesimpulan dan Saran.....	119
VI.1	Kesimpulan.....	119
VI.2	Saran.....	120
Daftar Pustaka	121