

DAFTAR ISI

Halaman Judul	ii
Halaman Pengesahan	iii
Halaman Pernyataan	iv
Halaman Persembahan	v
Halaman Motto	vi
PRAKATA	vii
INTISARI	xv
ABSTRACT	xvi
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
1.6 Sistematika Penulisan	2
II TINJAUAN PUSTAKA	4
III LANDASAN TEORI	8
3.1 Analisis Sentimen	8
3.1.1 Pengertian Analisis Sentimen	8
3.1.2 Tingkat Analisis Sentimen	8
3.1.3 Tokenisasi	9
3.1.4 Stemming dan lemmatisasi	10
3.1.5 Algoritma Nazief dan Adriani	10
3.1.6 Pengurangan Dimensi Data	14
3.2 Random Forest	16

3.2.1	Random forest menggunakan fitur acak	18
3.2.2	Random forest menggunakan seleksi input acak	18
3.2.3	Perhitungan Gini Gain pada CART	19
3.3	Support Vector Machine	27
3.3.1	SVM Hyperplane	28
3.3.2	Perhitungan jarak dari titik ke hyperplane	29
3.4	Majority Vote Classifier	33
3.5	Cross Validation	34
3.5.1	Metode holdout	35
3.5.2	K-fold cross validation	35
3.5.3	Leave-one-out cross validation	35
3.6	Precision, Recall, dan Accuracy	36
IV	ANALISIS DAN PERANCANGAN SISTEM	38
4.1	Gambaran Umum Sistem	38
4.2	Perancangan Sistem	38
4.2.1	Pengumpulan Data	39
4.2.2	Praproses	41
4.2.3	Pelatihan	43
4.2.4	Klasifikasi	45
4.2.5	Rancangan Pengujian	46
V	IMPLEMENTASI	48
5.1	Pengumpulan Data	48
5.2	Praproses	51
5.3	Pelatihan	56
5.4	Klasifikasi	57
5.5	Implementasi Pengujian	58
VI	PENGUJIAN DAN PEMBAHASAN	60
6.1	Pengujian Klasifikasi Sentimen pada Data Latih	60
6.2	Pengujian Precision, Recall, dan Accuracy	61
6.2.1	Pengujian tahap pertama	63
6.2.2	Pengujian tahap kedua	73

VII KESIMPULAN DAN SARAN	76
7.1 Kesimpulan	76
7.2 Saran	76
A Stopwords Bahasa Indonesia	82
B Kode Pengumpulan Data pada Situs Detik.com	84
C Kode Pengumpulan Data pada Situs Kompas.com	86
D Kode Pengumpulan Data pada Situs Okezone.com	88
E Kode Praproses Data Berita	90
F Kode Pelatihan Data Berita	92
G Kode Pengujian Hasil Klasifikasi	93
H Aturan dan Variasi Bentuk Awalan	97