

## DAFTAR ISI

HALAMAN JUDUL .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN .....	iv
MOTTO .....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
INTISARI .....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA .....	6
BAB III LANDASAN TEORI .....	9
3.1 Media Sosial .....	9
3.2 <i>Twitter</i> .....	10
3.3 <i>Cyberbullying</i> .....	11
3.4 Teknik <i>Text Mining</i> .....	12
3.5 <i>Natural Language Processing</i> .....	13
3.6 Analisis Teks Bahasa Indonesia .....	13
3.7 <i>Part of Speech (POS) Tagging</i> .....	14
3.8 Analisis LSF .....	15
3.8.1 Analisis <i>Lexical Feature</i> .....	15
3.8.2 Analisis <i>Syntactic Feature</i> .....	16
3.8.3 Nilai <i>Cyberbullying Tweet</i> .....	17
3.9 <i>TF-IDF</i> .....	17
3.10 <i>Naïve Bayes Classification</i> .....	18
BAB IV ANALISIS DAN PERANCANGAN SISTEM.....	21
4.1 Analisis Sistem.....	21
4.2 Arsitektur Sistem.....	21
4.3 Analisis Data .....	23
4.4 Preprocessing.....	26
4.4.1 Tokenisasi.....	26
4.4.2 Penghapusan <i>Stopword</i> .....	27

4.4.3	<i>Spell-Normalize</i> .....	27
4.4.4	Konversi <i>Multiword</i> .....	28
4.4.5	Konversi <i>Slangword</i> .....	28
4.5	Analisis LSF .....	29
4.5.1	<i>POS Tagging</i> .....	29
4.5.2	<i>Rule Matching LSF</i> .....	30
4.5.3	Pembobotan <i>Tweet</i> .....	31
4.6	Proses Klasifikasi dengan <i>Machine Learning</i> .....	32
4.6.1	TF-IDF.....	32
4.6.2	Proses Pelatihan dengan Metode <i>Naïve Bayes Classifier</i> .....	36
4.6.3	Proses Klasifikasi dengan Metode <i>Naïve Bayes Classifier</i> .....	40
4.7	Pengujian dan Evaluasi .....	43
BAB V	IMPLEMENTASI.....	44
5.1	Deskripsi Implementasi .....	44
5.2	Implementasi Pengumpulan Data <i>Tweet</i> .....	44
5.3	Implementasi <i>Preprocessing</i> .....	46
5.3.1	Tokenisasi.....	47
5.3.2	Penghapusan <i>Stopword</i> .....	48
5.3.3	<i>Spell-Normalize</i> .....	49
5.3.4	Konversi <i>Multiword</i> .....	49
5.3.5	Konversi <i>Slangword</i> .....	50
5.4	Deteksi <i>Cyberbullying Tweet</i> dengan analisis LSF .....	50
5.4.1	Analisis <i>Lexical Feature</i> .....	52
5.4.2	Analisis <i>Syntactic Feature</i> .....	52
5.4.3	Prediksi Label <i>Tweet</i> berdasarkan metode LSF .....	53
5.5	Klasifikasi <i>Tweet</i> dengan <i>Naïve Bayes Classifier</i> .....	54
5.5.1	Proses Pelatihan <i>Naïve Bayes Classifier</i> .....	54
5.5.2	Proses Klasifikasi <i>Naïve Bayes Classifier</i> .....	55
5.6	Klasifikasi <i>Tweet</i> dengan Kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	56
5.6.1	Proses Pelatihan Kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	56
5.6.2	Proses Klasifikasi Kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	58
5.7	Pengujian Metode Klasifikasi .....	60
5.7.1	Pengujian Klasifikasi <i>Naïve Bayes Classifier</i> .....	60
5.7.2	Pengujian Klasifikasi Kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	62
BAB VI	HASIL DAN PEMBAHASAN .....	65
6.1	Hasil Pengujian Deteksi <i>Cyberbullying Tweet</i> dengan LSF .....	65
6.2	Hasil Pengujian Metode Klasifikasi .....	66
6.2.1	Hasil Pengujian Metode <i>Naïve Bayes Classifier</i> .....	67
6.2.2	Hasil Pengujian Metode Kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	69
6.2.3	Perbandingan Pengujian Metode Klasifikasi.....	71
BAB VII	KESIMPULAN DAN SARAN.....	73
7.1	Kesimpulan.....	73
7.2	Saran .....	73
DAFTAR PUSTAKA	.....	75
LAMPIRAN	.....	78

## DAFTAR TABEL

Tabel 2.1	Studi terkait <i>cyberbullying</i> yang menggunakan teknik <i>text mining</i> .....	7
Tabel 3.1	Struktur analisis wacana teks van Dijk .....	14
Tabel 3.2	Contoh <i>tagset</i> Bahasa Indonesia .....	15
Tabel 3.3	Contoh perkataan kotor yang digunakan untuk melakukan <i>bullying</i> di Indonesia .....	16
Tabel 3.4	Contoh sintaksis kalimat <i>cyberbullying</i> dalam Bahasa Indonesia .....	16
Tabel 4.1	Komposisi kelas pelabelan data .....	24
Tabel 4.2	Komposisi data <i>train</i> dan data validasi pada model pelatihan .....	24
Tabel 4.3	Profil responden validator pelabelan data .....	25
Tabel 4.4	Contoh proses pelabelan data validasi .....	25
Tabel 4.5	Contoh tokenisasi dalam <i>preprocessing</i> .....	26
Tabel 4.6	Contoh penghapusan <i>stopword</i> dalam <i>preprocessing</i> .....	27
Tabel 4.7	Contoh <i>spell-normalize</i> dalam <i>preprocessing</i> .....	28
Tabel 4.8	Contoh konversi <i>multiword</i> dalam <i>preprocessing</i> .....	28
Tabel 4.9	Contoh konversi <i>slangword</i> dalam <i>preprocessing</i> .....	29
Tabel 4.10	Contoh Proses <i>POS Tagging</i> pada Analisis LSF .....	30
Tabel 4.11	Contoh <i>lexical feature matching</i> pada metode LSF .....	30
Tabel 4.12	<i>Syntactic feature set</i> pada analisis LSF .....	31
Tabel 4.13	Contoh pembobotan <i>tweet</i> pada analisis LSF .....	32
Tabel 4.14	Contoh data pelatihan untuk <i>Naïve Bayes Classifier</i> .....	33
Tabel 4.15	Contoh perhitungan nilai TF .....	33
Tabel 4.16	Contoh perhitungan nilai IDF .....	34
Tabel 4.17	Contoh perhitungan TF-IDF .....	35
Tabel 4.18	Pembobotan fitur TF-IDF untuk klasifikasi .....	37
Tabel 4.19	Pembobotan kombinasi fitur TF-IDF dan LSF untuk klasifikasi .....	38
Tabel 4.20	<i>Confussion matrix</i> untuk evaluasi hasil pengujian .....	43
Tabel 6.1	<i>Confusion matrix</i> hasil pengujian LSF .....	66
Tabel 6.2	Evaluasi pengujian metode <i>Naïve Bayes Classifier</i> .....	67
Tabel 6.3	Evaluasi pengujian metode kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	70

## DAFTAR GAMBAR

Gambar 4.1	Arsitektur Sistem .....	22
Gambar 5.1	Laman <i>Advanced Search</i> pada <i>Twitter</i> .....	45
Gambar 5.2	Implementasi <i>scraping</i> dengan <i>BeautifulSoup</i> .....	46
Gambar 5.3	Implementasi <i>preprocessing</i> .....	47
Gambar 5.4	<i>Regular Expression</i> untuk <i>tweet</i> .....	47
Gambar 5.5	Fungsi tokenisasi untuk <i>preprocessing</i> .....	48
Gambar 5.6	Pemanggilan <i>dictionary stopwords</i> .....	48
Gambar 5.7	Fungsi <i>spell-normalize</i> untuk <i>preprocessing</i> .....	49
Gambar 5.8	Fungsi konversi <i>multiword</i> untuk <i>preprocessing</i> .....	50
Gambar 5.9	Fungsi konversi <i>slangword</i> untuk <i>preprocessing</i> .....	50
Gambar 5.10	Implementasi deteksi <i>cyberbullying tweet</i> dengan LSF.....	51
Gambar 5.11	Implementasi prediksi label <i>tweet</i> berdasarkan metode LSF .....	53
Gambar 5.12	Implementasi proses pelatihan metode <i>Naïve Bayes Classifier</i> .....	55
Gambar 5.13	Klasifikasi dengan metode <i>Naïve Bayes Classifier</i> .....	56
Gambar 5.14	Implementasi proses pelatihan kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	57
Gambar 5.15	Implementasi proses klasifikasi metode kombinasi <i>Naïve Bayes</i> <i>Classifier</i> dan LSF .....	59
Gambar 5.16	Implementasi pengujian metode klasifikasi <i>Naïve Bayes Classifier</i> ...	62
Gambar 5.17	Implementasi pengujian metode kombinasi <i>Naïve Bayes Classifier</i> dan LSF .....	64
Gambar 6.1	Hasil evaluasi pengujian metode LSF untuk deteksi <i>cyberbullying</i> pada <i>tweet</i> Bahasa Indonesia.....	66
Gambar 6.2	Hasil evaluasi pengujian <i>Naïve Bayes Classifier</i> terhadap data pengujian .....	68
Gambar 6.3	Hasil evaluasi pengujian kombinasi <i>Naïve Bayes Classifier</i> dan LSF terhadap data pengujian .....	71
Gambar 6.4	Perbandingan evaluasi pengujian metode klasifikasi .....	72