

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
PERNYATAAN BEBAS PLAGIASI.....	ii
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	xi
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Keaslian Penelitian.....	3
BAB II TINJAUAN PUSTAKA.....	4
BAB III LANDASAN TEORI.....	13
3.1 Text Mining	13
3.2 Analisis Sentimen	14
3.3 <i>Twitter</i>	14
3.4 <i>Feature Extraction</i>	15



3.5	<i>Naïve Bayes</i>	16
3.6	Deteksi Sarkasme.....	18
3.7	<i>Random forest</i>	24
3.8	Evaluasi Kinerja <i>Classifier</i>	27
BAB IV Analisis dan Perancangan.....		28
4.1	Gambaran Umum Sistem	28
4.2	Data Penelitian.....	29
4.3	Perancangan Deteksi Sarkasme.....	30
4.3.1	Pelabelan Manual oleh Ahli Bahasa.....	31
4.3.2	<i>Preprocessing</i>	31
4.3.3	Ekstraksi Fitur Sarkasme.....	32
4.3.3.1	Unigram.....	32
4.3.3.2	Fitur-Fitur untuk Deteksi Sarkasme.....	32
4.3.4	Klasifikasi Sarkasme.....	37
4.3.5	Pengujian Hasil Algoritma dan Ahli Bahasa.....	41
4.4	Penggabungan Analisis Sentimen dan Deteksi Sarkasme.....	41
4.4.1	Pelabelan Manual.....	41
4.4.2	<i>Preprocessing</i>	42
4.4.3	Ekstraksi Fitur.....	43
4.4.4	Klasifikasi.....	45
4.4.5	Pembalikan.....	46
4.4.6	Pengujian Hasil Analisis Sentimen.....	47

BAB V IMPLEMENTASI.....	48
5.1 Deskripsi Implementasi.....	48
5.2 Implementasi Preprocessing.....	48
5.2.1 Daftar Library dan Pengambilan Data Set.....	49
5.2.2 Implementasi Mengambil Daftar Kata Slang.....	49
5.2.3 Implementasi Mengambil Stopwords.....	50
5.2.4 <i>Translate</i>	51
5.2.5 <i>Convert</i> Angka ke Karakter.....	52
5.2.6 <i>Stemming</i>	53
5.2.7 Proses Inti Preprocessing Sarkasme.....	53
5.2.8 Proses Inti Preprocessing Sentimen.....	54
5.2.9 Menyimpan Hasil Preprocessing.....	55
5.3 Implementasi Ekstraksi Fitur Sarkasme.....	56
5.3.1 Unigram.....	56
5.3.2 Empat Set Fitur Boazizi.....	58
5.3.2.1 Fitur punctuation-relate.....	58
5.3.2.2 Fitur Lexical and syntactic.....	60
5.3.2.3 Fitur Sentiment-relate.....	60
5.3.2.4 Fitur Top Word.....	62
5.4 Implementasi Pelatihan dan Pengujian Sarkasme.....	63
5.5 Implementasi Pelatihan dan Pengujian Kombinasi Sentimen dan Sarkasme.....	64

5.6 Implementasi Program untuk Data Masukan dari User.....	67
BAB VI HASIL DAN PEMBAHASAN.....	68
6.1 Pendeteksian Sentimen.....	68
6.1.1 <i>Preprocessing</i> Sentimen.....	68
6.1.2 Ekstraksi Fitur.....	69
6.1.3 Klasifikasi.....	69
6.2 Pendeteksian Sarkasme.....	70
6.2.1 <i>Preprocessing</i> Sarkasme.....	70
6.2.2 Ekstraksi Fitur Sarkasme.....	71
6.2.3 Klasifikasi Sarkasme.....	71
6.2.4 Pembalikan.....	76
6.3 Pengujian Deteksi Sarkasme Menggunakan Cross Validation.....	77
6.3.1 Penentuan Jumlah 10 Tree.....	77
6.3.2 Pengujian 5 Fold.....	77
6.3.3 Pengujian 10 Fold.....	78
6.3.4 Pengujian 15 Fold.....	79
6.3.5 Analisis Pengujian Sarkasme.....	79
6.4 Perbandingan Uji Sentimen dengan Kombinasi (Sentimen- Sarkasme).....	80
6.4.1 Pengujian 5 Fold.....	80
6.4.2 Pengujian 10 Fold.....	81
6.4.3 Pengujian 15 Fold.....	82

6.4.4 Analisis Pengujian Sentimen dengan Kombinasi (Sentimen-Sarkasme).....	84
6.5 Pengujian 4 fitur Boazizi.....	84
6.5.1 Pengujian Punctuation Related Feature.....	84
6.5.2 Pengujian Sentiment Related Features.....	85
6.5.3 Pengujian Lexical and Syntactic Features.....	85
6.5.4 Pengujian Top Words dan Punctuation Related Features.....	86
6.5.5 Pengujian Top Words dan Sentiment Related Features.....	86
6.5.6 Pengujian Top Words dan Lexical and Syntactic Features....	87
6.5.7 Analisis Fitur Boazizi.....	87
6.6 Pencarian Top words Terbaik.....	88
6.7 Pengujian terkait Misklasifikasi.....	88
6.7.1 Contoh Pengujian Beberapa Data Tweet.....	88
6.7.2 Hasil Misklasifikasi Semua Data Training.....	91
6.7.3 Hasil Misklasifikasi 5 Fold.....	91
6.8 Pengujian Salang dan Translate.....	92
6.9 Pembahasan.....	92
6.10 Kelebihan dan Kekurangan Kombinasi Sentimen-Sarkasme.....	94
BAB VII KESIMPULAN DAN SARAN.....	95
7.1 Kesimpulan	95
7.2 Saran.....	95
DAFTAR PUSTAKA.....	96



LAMPIRAN A.....	101
LAMPIRAN B.....	103
LAMPIRAN C.....	106
LAMPIRAN D.....	111
LAMPIRAN E.....	112
LAMPIRAN F.....	112
LAMPIRAN G.....	113
LAMPIRAN H.....	115
LAMPIRAN I.....	118