

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN DOSEN PEMBIMBING.....	ii
HALAMAN PENGESAHAN DOSEN PENGUJI.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTO	vi
KATA PENGANTAR.....	vii
SARI	ix
GLOSARIUM.....	x
DAFTAR ISI	xi
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan	2
1.4 Manfaat.....	2
1.5 Metodologi	3
1.6 Sistematika Penulisan	4
BAB II LANDASAN TEORI.....	6
2.1 Machine Learning.....	6
2.2 Pengolahan Bahasa Alami (PBA).....	6
2.3 <i>Text Classification</i>	7
2.4 <i>Languange Detection</i>	9
2.5 <i>Classification</i>	9
2.6 <i>Naïve Bayes Classiffication</i>	10
2.7 <i>Performance Evaluation Measure</i>	11
2.8 <i>Application Programming Instruction (API), Twitter, dan Twitter API</i>	15
2.9 Penelitian Serupa	19
BAB III METODE PENELITIAN	21
3.1 Analisis Kebutuhan.....	21
3.2 Data dan Sumber Data	22

3.3	Gambaran Sistem.....	24
3.3.1	Pengambilan Data.....	25
3.3.2	<i>Pre-processing</i>	26
3.3.3	Uji Model.....	27
BAB IV HASIL DAN PEMBAHASAN		28
4.1	Hasil Pengolahan <i>Dataset</i>	28
4.2	Hasil Pembentukan Model	33
4.3	Performa Algoritma	35
BAB V KESIMPULAN.....		40
5.1	Kesimpulan.....	40
5.2	Saran.....	40
DAFTAR PUSTAKA		41
LAMPIRAN		42

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	13
Tabel 2.2 Contoh soal <i>accuracy</i> , <i>precisi</i> dan <i>recall</i>	14
Tabel 3.1 Tabel contoh <i>word vector</i>	27
Tabel 4.1 <i>Confusion Matrix</i> Bahasa Jawa	36
Tabel 4.2 <i>Confusion Matrix</i> Bahasa Sunda	36
Tabel 4.3 <i>Confusion Matrix</i> Bahasa Indonesia.....	36
Tabel 4.4 <i>Confusion Matrix</i> keseluruhan	37

DAFTAR GAMBAR

Gambar 2.1 Contoh Penggunaan N-Gram	8
Gambar 2.2 Proses kerja klasifikasi.....	10
Gambar 2.3 Gambaran akurasi dan presisi data	13
Gambar 2.4 gambaran data perbandingan <i>accuracy</i> , <i>precision</i> dan <i>recall</i>	15
Gambar 2.5 Logo Twitter.....	16
Gambar 2.6 Jumlah Pengguna Aktif Twitter di Indonesia (dalam juta)	17
Gambar 2.7 Gambaran kerja API	18
Gambar 3.1 Halaman awal <i>Apps</i> Twitter	22
Gambar 3.2 Contoh akun Twitter Bahasa Jawa.....	23
Gambar 3.3 Contoh akun Twitter Bahasa Sunda.....	24
Gambar 3.4 Gambaran Sistem.....	25
Gambar 4.1 <i>API Key</i> Twitter	28
Gambar 4.2 Contoh code pengambilan data.....	30
Gambar 4.3 Data JSON.....	30
Gambar 4.4 Code menghilangkan tanda baca dan akun	31
Gambar 4.5 Teks Bahasa Jawa	32
Gambar 4.6 Teks Bahasa Sunda	32
Gambar 4.7 Teks Bahasa Indonesia.....	32
Gambar 4.8 Data Salah <i>Tag</i> dan Teks Kosong.....	32
Gambar 4.9 Data Excel	33
Gambar 4.10 Kata Setelah Melalui Proses Tokenisasi	33
Gambar 4.11 Data <i>Word Vector</i>	34
Gambar 4.12 Data yang Sudah Terbobot dengan N-Gram dan TF-IDF.....	34
Gambar 4.13 <i>Word vector</i> untuk <i>data test</i>	35
Gambar 4.14 <i>Confusion Matrix</i>	36
Gambar 4.15 Code penghitungan <i>accuracy</i>	37
Gambar 4.16 Code penghitungan <i>precision</i> dan <i>recall</i>	37
Gambar 4.17 Contoh Penggunaan Bahasa Daerah yang Bercampur dan Mirip	38
Gambar 4.18 Input nama akun Twitter	38
Gambar 4.19 Input nama akun Twitter Bahasa Jawa.....	39