

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN DOSEN PEMBIMBING.....	ii
HALAMAN PENGESAHAN DOSEN PENGUJI.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iii
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 Penelitian.....	2
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian.....	3
1.7 Sistematis Penulisan	5
BAB II LANDASAN TEORI.....	6
2.1 Hate Speech	6
2.2 Twitter.....	6
2.3 Sentimen Analysis	7
2.4 Text Mining	7
2.5 Naive Bayes Classifier.....	9
2.6 Cross Validation	9
2.7 Performance Evaluation Measure	10
2.8 Penelitian Serupa	13
BAB III METODOLOGI PENELITIAN	15
3.1 Alur Pengerjaan Tugas Akhir	15
3.2 Uraian Metodologi.....	16
3.2.1 Pengambilan Data.....	16
3.2.2 Tahapan Proses Preprocessing.....	19
3.2.3 Ekstraksi Fitur.....	22
3.2.4 Klasifikasi.....	23
3.2.5 Uji Model.....	23
3.2.6 Evaluasi Model	24
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Pengambilan Data (Crawling data)	25
4.2 Preprocessing.....	28
4.3 Ekstraksi Fitur	32
4.4 Implementasi Klasifikasi Naive Bayes.....	35
4.5 Uji Model.....	36
4.6 Evaluasi Model.....	37
BAB V KESIMPULAN.....	43
5.1 Kesimpulan.....	43

5.2	Saran.....	43
	DAFTAR PUSTAKA.....	44
	LAMPIRAN.....	46

DAFTAR TABEL

Tabel 2.1 Confusion Matrix.....	11
Tabel 2.2 Contoh Hasil Confusion Matrix.....	12
Tabel 3.1 Contoh Data Hasil Labelling	18
Tabel 3.2 Contoh Hasil Cleaning.....	19
Tabel 3.3 Contoh Data Hasil Remove Stopword.....	20
Tabel 3.4 Contoh Data Hasil Tokenization.....	21
Tabel 3.5 Contoh Data Hasil Stemming	22
Tabel 4.1 Pembuatan Word Vector.....	32
Tabel 4.2 Proses Perhitungan TF (Term Frequency).....	33
Tabel 4.3 Proses Perhitungan DF (Document Frequency)	33
Tabel 4.4 Proses IDF (Inverse Document Frequency).....	34
Tabel 4.5maka Contoh Proses Perhitungan TF-IDF.....	34
Tabel 4.6 Contoh Word Vector yang sudah dibobotkan.....	34
Tabel 4.7 Model Confusion Matrix	36
Tabel 4.8 Hasil Confusion Matrix	37
Tabel 4.9 Hasil dari Nilai Precision, Recall, dan F-1 score.....	39
Tabel 4.10 Hasil Precision, Recall, dan F-1 score	40
Tabel 4.11 Perbandingan Metode penelitian	41

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Gambar Precision dan Accuracy	11
Gambar 2.2 Perbandingan Precision, nilai Recall dan nilai Accuracy	12
Gambar 3.1 Alur Pengerjaan Tugas Akhir	15
Gambar 3.2 Halaman Awal Twitter	16
Gambar 3.3 Halaman Developer Twitter	17
Gambar 3.4 Halaman Request API Key Twitter	17
Gambar 3.5 Hasil Crawling Data	18
Gambar 4.1 API Key Twitter	25
Gambar 4.2 Tampilan Anaconda	25
Gambar 4.3 Tampilan Jupyter Notebook	26
Gambar 4.4 Source Code Pemanggilan Python Library Proses Crawling	26
Gambar 4.5 Source Code Proses Crawling	27
Gambar 4.6 File Excel Hasil Crawling	27
Gambar 4.7 Hasil data crawling dan labelling	28
Gambar 4.8 kode program proses cleaning	29
Gambar 4.9 proses install library nltk	30
Gambar 4.10 Pendeklarasian library nltk	30
Gambar 4.11 Kode Program Proses Remove Stopword	30
Gambar 4.12 Kode Program Proses Tokenization	31
Gambar 4.13 Proses Instalasi library sastrawi	31
Gambar 4.14 Pendeklarasian library sastrawi	31
Gambar 4.15 Kode Program Proses Stemming	31
Gambar 4.16 Hasil Preprocessing	32
Gambar 4.17 Proses pendeklarasian library yang digunakan	35
Gambar 4.18 Proses memanggil data set	35
Gambar 4.19 Proses Pengimplementasian Class Pipeline	36
Gambar 4.20 Hasil Akurasi	36
Gambar 4.21 Nilai Akurasi dan Confusion Matrix 2x2	37
Gambar 4.22 Proses Menghitung dari Nilai Presisi, Recall dan F-1 score	38
Gambar 4.23 Hasil dari Proses Pengevaluasian Model	39
Gambar 4.24 Hasil Pengujian 5 K-Fold Cross Validation	40
Gambar 4.25 Hasil Fold Validation	41