

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN BEBAS PLAGIASI	iv
PERNYATAAN PEMBUATAN PROGRAM	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
INTISARI	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan	4
1.5 Hasil yang Diharapkan	4
1.6 Manfaat	4
1.7 Kontribusi	4
1.8 Metodologi Penelitian	5
1.9 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
BAB III LANDASAN TEORI	12
3.1 Informasi Palsu	12
3.2 <i>Social Context Features</i>	13
3.3 Twitter	14

3.4	<i>REST API</i>	15
3.5	<i>Data Mining</i>	16
3.6	<i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i>	18
3.7	<i>Text Mining</i>	19
3.8	Normalisasi	20
3.9	Diskritisasi Menggunakan <i>Data Binning</i>	21
3.10	<i>Feature Extraction</i>	21
3.11	<i>Feature Selection</i>	22
3.12	<i>Information Gain</i>	22
3.13	Klasifikasi	23
3.14	<i>Naïve Bayes</i>	24
3.15	<i>Term Frequency</i> untuk Teks	28
3.16	Akurasi	28
3.17	Presisi/ <i>Precision</i>	29
3.18	<i>Recall</i>	30
3.19	<i>F-Measure</i>	30
3.20	<i>K-Fold Cross Validation</i>	30
BAB IV ANALISIS DAN PERANCANGAN		33
4.1	Gambaran Umum Penelitian	33
4.2	Dataset	33
4.3	Rancangan Penelitian	33
4.4	Rancangan Proses	35
BAB V IMPLEMENTASI		52
5.1	Deskripsi Implementasi	52
5.2	Implementasi Pengumpulan Data	52
5.3	Dataset	60
5.4	Implementasi Praproses	62

5.5	Implementasi Ekstraksi dan Seleksi Fitur	66
5.6	Pemodelan dan Pengujian	67
5.7	Evaluasi	69
BAB VI HASIL DAN PEMBAHASAN		71
6.1	Hasil untuk Skenario Pertama	71
6.2	Hasil untuk Skenario Kedua	73
6.3	Hasil untuk Skenario Ketiga	75
6.4	Pembahasan	75
BAB VII KESIMPULAN DAN SARAN		81
7.1	Kesimpulan	81
7.2	Saran	81
DAFTAR PUSTAKA		83
LAMPIRAN DATASET		86

DAFTAR GAMBAR

Gambar 3.1	Relasi Dimensi dalam Ekosistem Media Sosial	13
Gambar 3.2	Ilustrasi <i>K-Fold Cross Validation</i> dengan $k=4$	31
Gambar 4.1	Rancangan Alur Proses Penelitian	34
Gambar 4.2	Diagram Alir Pengumpulan Data	35
Gambar 4.3	Diagram Alir Praproses	40
Gambar 4.4	Diagram Alir Skema Evaluasi	51
Gambar 5.1	<i>Form Input URL Tweet</i>	53
Gambar 5.2	Potongan Kode Proses <i>Scraping</i>	53
Gambar 5.3	Potongan Kode Proses Penentuan Sentimen	56
Gambar 5.4	Potongan Kode Penentuan Atribut Status Penangguhan pada 21 <i>tweets</i> Terakhir	58
Gambar 5.5	Potongan Kode Penentuan Atribut Waktu Rata-Rata Penulisan 21 <i>tweets</i> Terakhir	58
Gambar 5.6	Potongan Kode Penentuan Atribut Jumlah Rata-Rata <i>Retweet</i> 21 <i>tweets</i> Terakhir	59
Gambar 5.7	Potongan Kode Penentuan Atribut Jumlah Rata-Rata <i>Like</i> 21 <i>tweets</i> Terakhir	60
Gambar 5.8	Visualisasi Jumlah Data Berdasarkan Kelas/Label	61
Gambar 5.9	Visualisasi Jumlah Data Berdasarkan Kategori	61
Gambar 5.10	Potongan Kode Proses <i>Casefolding</i> dan <i>Cleaning</i>	62
Gambar 5.11	Potongan Kode Proses Tokenisasi	63
Gambar 5.12	Potongan Kode Proses Penghapusan <i>Stop Word</i>	63
Gambar 5.13	Potongan Kode Proses <i>Stemming</i>	64
Gambar 5.14	Potongan Kode Proses Normalisasi Data	65
Gambar 5.15	Potongan Kode Proses Diskritisasi Data	65
Gambar 5.16	Potongan Kode Proses Ekstraksi Fitur	66
Gambar 5.17	Proses Kalkulasi Information Gain	67
Gambar 5.18	Implementasi <i>Naïve Bayes</i> dan <i>K-Fold Cross Val.</i>	68
Gambar 5.19	Contoh Skor Hasil Pemodelan dan Pengujian	70
Gambar 6.1	Perbandingan Nilai Akurasi Skenario Pertama	73
Gambar 6.2	Perbandingan Nilai Akurasi Skenario Kedua	74

DAFTAR TABEL

Tabel 2.1	Tinjauan Pustaka	10
Tabel 3.1	<i>Twitter API Endpoint</i>	15
Tabel 3.2	<i>Confusion Matrix</i> pada Klasifikasi Dua Kelas	29
Tabel 3.3	Skema <i>10-Fold Cross Validation</i>	32
Tabel 4.1	Pembagian Kategori Dataset	36
Tabel 4.2	Contoh <i>Tweet</i> Hasil <i>Scraping</i>	36
Tabel 4.3	Contoh <i>Casefolding</i> dan <i>Cleaning</i>	41
Tabel 4.4	Contoh Tokenisasi	41
Tabel 4.5	Contoh Penghapusan <i>Stop Word</i>	42
Tabel 4.6	Contoh <i>Stemming</i>	42
Tabel 4.7	Contoh <i>Min-Max Normalization</i>	37
Tabel 4.8	Contoh Diskritisasi Menggunakan <i>Binning</i>	43
Tabel 4.9	Praproses Atribut	44
Tabel 4.10	Fitur Klasifikasi	44
Tabel 4.11	Penggunaan Fitur Klasifikasi	45
Tabel 4.12	Contoh Data	46
Tabel 4.13	<i>Count Table</i> Fitur A1	47
Tabel 5.1	Statistik <i>Dataset</i>	60
Tabel 5.2	Nilai <i>Information Gain</i>	68
Tabel 6.1	Hasil untuk Skenario Pertama	72
Tabel 6.2	Hasil untuk Skenario Kedua	73
Tabel 6.3	Perbandingan Akurasi Terhadap Penggunaan Dimensi ...	75
Tabel 6.4	Parameter/Koefisien Model untuk Skenario Pertama	76
Tabel 6.5	Parameter/Koefisien Model untuk Skenario Kedua	77
Tabel 6.6	Perbandingan Peningkatan Akurasi Classifier	79