

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
PRAKATA.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
DAFTAR PERSAMAAN.....	xiv
DAFTAR LAMPIRAN.....	xv
INTISARI.....	xvi
ABSTRACT.....	xvii
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 Metodologi Penelitian.....	4
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	7
BAB III LANDASAN TEORI.....	11
3.1 Twitter.....	11
3.1.1 Twit.....	11
3.1.2 Twitter API.....	12
3.2 Identifikasi <i>Writeprint / Author Identification</i>	14
3.3 <i>Preprocessing</i>	14
3.4 Ekstraksi Fitur.....	15
3.5 <i>Support Vector Machine (SVM)</i>	19
3.6 <i>K-Fold Cross Validation</i>	24
3.7 Python.....	27
3.7.1 Keunggulan Python.....	28
3.7.2 Library Python.....	29
BAB IV ANALISIS DAN PERANCANGAN SISTEM.....	31
4.1 Deskripsi Sistem.....	31
4.2 <i>Dataset</i>	32
4.3 Perancangan Sistem.....	34
4.3.1 Perancangan <i>Preprocessing</i>	35
4.3.2 Perancangan Ekstraksi Fitur.....	39
4.3.3 Perancangan Klasifikasi.....	42
4.3.4 Perancangan Evaluasi Sistem Klasifikasi.....	43
BAB V IMPLEMENTASI.....	44
5.1 Implementasi Sistem.....	44



5.2	Implementasi <i>Preprocessing</i>	44
5.3	Implementasi Ekstraksi Fitur	46
5.4	Implementasi Klasifikasi	53
5.5	Implementasi Pengujian	59
BAB VI HASIL PENELITIAN DAN PEMBAHASAN		62
6.1	<i>Dataset</i>	62
6.2	Hasil <i>Preprocessing</i>	64
6.3	Hasil Ekstraksi Fitur	64
6.4	Hasil Klasifikasi	65
6.5	Hasil Evaluasi Sistem Klasifikasi	67
6.6	Penggunaan Model	69
BAB VII KESIMPULAN DAN SARAN		72
DAFTAR PUSTAKA		73
LAMPIRAN		76

DAFTAR GAMBAR

Gambar 3.1 Tampilan twit	11
Gambar 3.2 Contoh tokenisasi	16
Gambar 3.3 Contoh proses POS <i>Tagging</i>	17
Gambar 3.4 Proses <i>named-entity recognition</i> menggunakan Stanford NER.....	19
Gambar 3.5 (a) <i>Hyperplane</i> yang mungkin (b) <i>Hyperplane</i> margin maksimal ...	20
Gambar 3.6 Hasil SVM linear.....	21
Gambar 3.7 (a) Data latih logika AND (b) Batas keputusan logika AND.....	22
Gambar 3.8 <i>Hyperplane</i> pada logika AND.....	23
Gambar 4.1 Garis besar sistem	32
Gambar 4.2 Contoh daftar twit yang belum melewati tahap <i>preprocessing</i>	36
Gambar 4.3 Hasil penghapusan twit	36
Gambar 4.4 Hasil proses pembersihan twit	38
Gambar 4.5 Alur <i>Preprocessing</i>	38
Gambar 4.6 Hasil tokenisasi	39
Gambar 4.7 Hasil twit setelah pemberian kelas kata	40
Gambar 4.8 Diagram alir proses ekstraksi fitur	41
Gambar 4.9 Proses Klasifikasi	43
Gambar 5.1 Kode implementasi <i>preprocessing</i>	44
Gambar 5.2 Kode untuk membersihkan <i>noise</i>	45
Gambar 5.3 Kode untuk memanggil fungsi <code>word_tokenize</code> dari library <code>nlTK</code> ..	46
Gambar 5.4 Kode menyiapkan <i>environment</i> Java.....	46
Gambar 5.5 Kode untuk mengimpor Stanford POS-Tagger.....	46
Gambar 5.6 Kode untuk mengimpor Stanford NER.....	46
Gambar 5.7 Kode implementasi tokenisasi.....	47
Gambar 5.8 Kode ekstraksi fitur kelompok pertama	47
Gambar 5.9 Kode menghitung jumlah karakter spesial	48
Gambar 5.10 Kode menghitung huruf kapital.....	49
Gambar 5.11 Kode menghitung jumlah singkatan pada twit.....	49
Gambar 5.12 Kode menghitung jumlah akronim pada twit.....	50
Gambar 5.13 Kode implementasi menghitung jumlah kata dalam twit.....	50
Gambar 5.14 Kode implementasi Stanford POS-Tagger	51
Gambar 5.15 Kode mencari fitur-fitur kelompok kedua.....	51
Gambar 5.16 Kode implementasi pembuatan model klasifikasi 2 kelas	54
Gambar 5.17 Data model klasifikasi <i>dataset</i> 2 kelas	55
Gambar 5.18 Kode implementasi <i>testing dataset</i> 2 kelas	55
Gambar 5.19 Kode implementasi pembuatan model klasifikasi 12 kelas	57
Gambar 5.20 Data model klasifikasi 12 kelas.....	58
Gambar 5.21 Kode implementasi <i>testing dataset</i> 12 kelas	58
Gambar 5.22 Kode implementasi preparasi pengujian	60
Gambar 5.23 Kode implementasi pengujian (lanjutan)	60



Gambar 6.1 Akun asli milik Ariana Grande	62
Gambar 6.2 Akun yang meniru akun Ariana Grande	62
Gambar 6.3 Cuplikan data twit milik Ariana Grande yang belum diproses	63
Gambar 6.4 Cuplikan data twit yang telah melewati <i>preprocessing</i>	64
Gambar 6.5 Cuplikan hasil klasifikasi data <i>testing</i> milik Ariana Grande.....	66
Gambar 6.6 Cuplikan hasil klasifikasi <i>dataset</i> 12 kelas	67
Gambar 6.7 Hasil klasifikasi percobaan twit pertama	69
Gambar 6.8 Hasil klasifikasi percobaan twit kedua.....	70
Gambar 6.9 Hasil penggunaan model <i>dataset</i> 12 kelas pada twit pertama.....	71
Gambar 6.10 Hasil penggunaan model <i>dataset</i> 12 kelas pada twit kedua.....	71

DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Penelitian Terdahulu	9
Tabel 3.1 Contoh <i>Part-Of-Speech</i>	16
Tabel 3.2 Kelas pada model Stanford NER	18
Tabel 3.3 Operasi logika AND dengan notasi SVM	22
Tabel 3.4 Fungsi kernel	24
Tabel 3.5 Tabel <i>confusion matrix</i>	26
Tabel 4.1 Daftar nama figur publik dan jumlah twit yang digunakan	33
Tabel 4.2 Label kelas <i>dataset</i> 12 kelas	34
Tabel 4.3 Daftar Fitur	35
Tabel 4.4 Perubahan kata-kata yang disingkat	37
Tabel 4.5 Pemberian Label	37
Tabel 5.1 Tabel karakter spesial	48
Tabel 6.1 Penjelasan fitur pada twit	65
Tabel 6.2 Hasil pengujian semua akun	68
Tabel 6.3 Hasil pengujian <i>dataset</i> 12 kelas	68

DAFTAR PERSAMAAN

Persamaan (3.1).....	20
Persamaan (3.2).....	21
Persamaan (3.3).....	21
Persamaan (3.4).....	22
Persamaan (3.5).....	23
Persamaan (3.6).....	23
Persamaan (3.7).....	24
Persamaan (3.8).....	26
Persamaan (3.9).....	27
Persamaan (3.10).....	27
Persamaan (3.11).....	27
Persamaan (3.12).....	27
Persamaan (3.13).....	27

DAFTAR LAMPIRAN

Lampiran 1. Daftar Nama Akun	76
------------------------------------	----