

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN PEMBIMBING	ii
HALAMAN PENGESAHAN PENGUJI	iii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR MODUL	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian dan Pengembangan Sistem	3
1.6.1 Metodologi Penelitian	3
1.6.2 Metodologi Pengembangan Sistem	4
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	6
2.1 Twitter	6
2.2 Spam	7
2.3 Klasifikasi Teks dan Text Mining	8
2.4 Crawling	9
2.5 Text Preprocessing	9
2.5.1 Cleansing	9
2.5.2 Slangword	9
2.5.3 Stopword	9
2.5.4 Stemming	10
2.5.5 Tokenizing	12
2.6 Pembobotan TF-IDF	12
2.7 Artificial Neural Network	12
2.8 Multilayer Perceptron	14
2.9 Evaluasi atau Pengujian	16
2.9.1 K-Fold Cross Validation	16
2.9.1 Confusion Matrix	16
2.9 Studi Pustaka (Staet of the Art)	17
BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM	19
3.1 Metodologi Penelitian	20
3.1.1 Pengumpulan Data	20
3.1.2 Text Preprocessing	21
3.1.2.1 Cleansing	22
3.1.2.2 Slangword	22
3.1.2.3 Stopword	22
3.1.2.4 Stemming	23
3.1.2.5 Tokenizing	23
3.1.3 Pembobotan TF-IDF	24
3.1.3.1 Menghitung TF-IDF	25
3.1.4 Algoritma Multilayer Perceptron	26
3.1.5 Pengujian Model	29

3.2 Metodologi Pengembangan Sistem.....	31
3.2.1 Pengumpulan Data	32
3.2.2 Analisa Kebutuhan Sistem.....	32
3.2.3 Perancangan Sistem	34
3.2.4 Perancangan Proses.....	36
3.2.5 Perancangan Antarmuka	45
BAB IV HASIL DAN PEMBAHASAN	46
4.1 Hasil Penelitian	46
4.1.1 Pengumpulan Data	46
4.1.2 Implementasi Labeling	47
4.1.3 Implementasi Preprocessing	47
4.1.4 Implementasi TF-IDF	49
4.1.5 Implementasi Algoritma Multilayer Perceptron	50
4.1.6 Implementasi Sistem.....	51
4.2 Hasil Pengujian	51
4.2.1 Pengujian Cross Validation	52
4.2.2 Pengujian Confusion Matrix.....	52
4.3 Pembahasan.....	54
BAB V PENUTUP	56
5.1 Kesimpulan	56
5.2 Saran	56
DAFTAR PUSTAKA	xiv

DAFTAR GAMBAR

Gambar 2.1	Desain ANN Secara Umum	13
Gambar 2.2	Arsitektur Multilayer Perceptron	14
Gambar 3.1	Metodologi Penelitian	20
Gambar 3.2	Hasil Observasi Tweet Spam dan Bukan Spam	21
Gambar 3.3	Flowchart Pembobotan TF-IDF	25
Gambar 3.4	Arsitektur Multilayer Perceptron	27
Gambar 3.5	Arsitektur Multilayer Perceptron Dari Sampel Perhitungan	28
Gambar 3.6	Metodologi Pengembangan Sistem	29
Gambar 3.7	Arsitektur Sistem	34
Gambar 3.8	Use Case Diagram	35
Gambar 3.9	Activity Diagram	35
Gambar 3.10	Sequence Diagram	36
Gambar 3.11	Flowchart Preprocessing	37
Gambar 3.12	Flowchart Cleansing	38
Gambar 3.13	Flowchart Slangword	39
Gambar 3.14	Flowchart Stopword	40
Gambar 3.15	Flowchart Stemming	41
Gambar 3.16	Flowchart Tokenizing	42
Gambar 3.17	Flowchart TF-IDF	43
Gambar 3.18	Flowchart Multilayer Perceptron	44
Gambar 3.19	Perancangan Antarmuka	45
Gambar 4.1	Hasil Crawling Data	47
Gambar 4.2	Dataset Hasil Labelisasi	47
Gambar 4.3	Halaman Deteksi Tweet Spam	51
Gambar 4.4	Pengujian Cross Validation	52

DAFTAR TABEL

Tabel 2.1	Contoh Kata-kata Spam.....	8
Tabel 2.2	Kombinasi Awalan dan Akhiran Tidak Diiijinkan	10
Tabel 2.3	Aturan Peluruhan Kata Dasar	11
Tabel 2.4	Confusion Matrix	16
Tabel 2.5	State of the Art.....	18
Tabel 3.1	Rincian Data Spam	21
Tabel 3.2	Contoh Data Spam Beserta Label.....	22
Tabel 3.3	Perhitungan TF-IDF.....	26
Tabel 3.4	Penerapan Algoritma Multilayer Perceptron	27
Tabel 3.5	Pengujian Confusion Matrix Model TF-IDF dan Multilayer Perceptron.....	30
Tabel 3.6	Pengujian Cross Validation Model TF-IDF dan Multilayer Perceptron.....	30
Tabel 3.7	Spesifikasi Kebutuhan Perangkat Keras dan Perangkat Lunak	33
Tabel 3.8	Spesifikasi Kebutuhan Perangkat Keras dan Perangkat Lunak	33
Tabel 4.1	Confusion Matrix TF dan Multilayer Perceptron	53
Tabel 4.2	Confusion Matrix TF-IDF dan Multilayer Perceptron	53
Tabel 4.3	Confusion Matrix TF Orthography dan Multilayer Perceptron.....	53
Tabel 4.4	Confusion Matrix TF-IDF Orthography dan Multilayer Perceptron	54

DAFTAR MODUL

Modul Program 4.1	Pengumpulan Dataset	46
Modul Program 4.2	Inisiasi Dataset.....	47
Modul Program 4.3	Cleansing	48
Modul Program 4.4	Slangword dan Stopword	48
Modul Program 4.5	Stemming.....	49
Modul Program 4.6	Preprocessing.....	49
Modul Program 4.7	Proses TF	49
Modul Program 4.8	Proses TF-IDF	50
Modul Program 4.9	Proses Pelatihan, Pengujian dan Akturasi Multilayer Perceptron.....	50

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN PEMBIMBING	ii
HALAMAN PENGESAHAN PENGUJI	iii
KATA PENGANTAR	iv
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	viii
DAFTAR MODUL	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian dan Pengembangan Sistem	3
1.6.1 Metodologi Penelitian	3
1.6.2 Metodologi Pengembangan Sistem	4
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	6
2.1 Twitter	6
2.2 Spam	7
2.3 Klasifikasi Teks dan Text Mining	8
2.4 Crawling	9
2.5 Text Preprocessing	9
2.5.1 Cleansing	9
2.5.2 Slangword	9
2.5.3 Stopword	9
2.5.4 Stemming	10
2.5.5 Tokenizing	12
2.6 Pembobotan TF-IDF	12
2.7 Artificial Neural Network	12
2.8 Multilayer Perceptron	14
2.9 Evaluasi atau Pengujian	16
2.9.1 K-Fold Cross Validation	16
2.9.1 Confusion Matrix	16
2.9 Studi Pustaka (Staet of the Art)	17
BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM	19
3.1 Metodologi Penelitian	20
3.1.1 Pengumpulan Data	20
3.1.2 Text Preprocessing	21
3.1.2.1 Cleansing	22
3.1.2.2 Slangword	22
3.1.2.3 Stopword	22
3.1.2.4 Stemming	23
3.1.2.5 Tokenizing	23
3.1.3 Pembobotan TF-IDF	24
3.1.3.1 Menghitung TF-IDF	25
3.1.4 Algoritma Multilayer Perceptron	26
3.1.5 Pengujian Model	29