

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PENGESAHAN PEMBIMBING	iii
HALAMAN PENGESAHAN PENGUJI.....	iv
SURAT PERNYATAAN KARYA TULIS TUGAS AKHIR	v
PERNYATAAN BEBAS PLAGIASI.....	vi
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR MODUL PROGRAM.....	xiv
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 Tahapan Penelitian	3
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN LITERATUR	5
2.1 Klasifikasi.....	5
2.2 Jurnal	5
2.3 Text Mining.....	5
2.4 Text Preprocessing	5
2.4.1. Casefolding.....	5
2.4.2. Tokenization	6
2.4.3. Stopwords Removal.....	6
2.4.4. Stemming.....	6
2.5 K-Nearest Neighbor	6
2.6 Cosine Similarity	6
2.7 Euclidean Distance	7
2.8 TF-IDF.....	7
2.9 Cross Validation	8
2.10 Confusion Matrix	8
2.11 Studi Literatur	9
BAB III METODOLOGI PENELITIAN	15
3.1 Metodologi Penelitian	16
3.1.1 Studi Literatur	16
3.1.2 Persiapan Dataset.....	16
3.1.3 Pre-Processing Data	16
3.1.4 Pembobotan Kata	22
3.1.5 K-Nearest Neighbor	25
3.1.6 Cosine Similarity	27
3.1.7 Euclidean Distance.....	30

3.1.8 Evaluasi Model	33
3.2 Metodologi Pengembangan Sistem	33
3.2.1 Analisis Kebutuhan Sistem	34
3.2.2 Perancangan Sistem	32
3.3 Perancangan Antarmuka.....	36
3.4 Pengujian Sistem	39
BAB IV HASIL DAN PEMBAHASAN.....	40
4.1 Implementasi Klasifikasi Jurnal	40
4.2 Implementasi Sistem	43
4.3 Pengujian	46
4.4 Pembahasan	50
BAB V KESIMPULAN SARAN.....	51
5.1 Kesimpulan.....	51
5.2 Saran	51
DAFTAR PUSTAKA.....	52

DAFTAR TABEL

Tabel 2.1 State of The Art	12
Tabel 2.2 State of The Art (lanjutan).....	13
Tabel 3.1 Contoh Penggunaan Cleansing.....	17
Tabel 3.2 Contoh Penggunaan Case Folding.....	17
Tabel 3.3 Contoh Penggunaan Tokenization.....	18
Tabel 3.4 Contoh Penggunaan Stopword Removal.....	19
Tabel 3.5 Contoh Penggunaan Stemming	21
Tabel 3.6 Contoh Penggunaan Stemming (lanjutan).....	22
Tabel 3.7 Perbandingan Sebelum dan Setelah Tahap Pre-Processing.....	23
Tabel 3.8 Tabel TF	23
Tabel 3.9 Tabel TF (lanjutan).....	24
Tabel 3.10 Tabel IDF.....	24
Tabel 3.11 Tabel IDF (lanjutan).....	25
Tabel 3.12 Tabel TF-IDF.....	25
Tabel 3.13 Perhitungan Skalar	28
Tabel 3.14 Perhitungan Skalar (lanjutan).....	29
Tabel 3.15 Perhitungan Akar Kuadrat TF-IDF	29
Tabel 3.16 Perhitungan Akar Kuadrat TF-IDF (lanjutan).....	30
Tabel 3.17 Hasil Cosine Similarity	30
Tabel 3.18 Hasil Euclidean Distance.....	31
Tabel 3.19 Hasil Euclidean Distance (lanjutan).....	32
Tabel 3.20 Kebutuhan Perangkat Keras (Hardware).....	34
Tabel 3.21 Kebutuhan Perangkat Lunak (Software)	34
Tabel 3.22 Confusion Matrix.....	39
Tabel 4.1 Hasil Performa.....	47
Tabel 4.2 Pengujian Confusion Matrix Model K-NN dengan Cosine Similarity	47
Tabel 4.3 Pengujian Confusion Matrix Model K-NN dengan Euclidean Distance.....	48

DAFTAR GAMBAR

Gambar 3.1 Metodologi Penelitian.....	15
Gambar 3.2 Flowchart Cleansing	16
Gambar 3.3 Flowchart Case Folding	17
Gambar 3.4 Flowchart Tokenization	18
Gambar 3.5 Flowchart Stopwords Removal	19
Gambar 3.6 Flowchart Stemming.....	21
Gambar 3.7 Flowchart TF-IDF.....	22
Gambar 3.8 Flowchart K-NN dengan Cosine Similarity	26
Gambar 3.9 Flowchart K-NN dengan Eucidean Distance.....	27
Gambar 3.10 Flowchart Cosine Similarity	28
Gambar 3.11 Flowchart Euclidean Distance	31
Gambar 3.12 Rancangan Arsitektur	35
Gambar 3.13 DFD Level 0	35
Gambar 3.14 DFD Level 1	36
Gambar 3.15 Rancangan Halaman Dashboard.....	37
Gambar 3.16 Rancangan Halaman Pre-Processed	37
Gambar 3.17 Rancangan Halaman Performa	38
Gambar 3.18 Rancangan Halaman Klasifikasi.....	38
Gambar 4.1 Halaman Dashboard	44
Gambar 4.2 Halaman Dashboard (lanjutan)	44
Gambar 4.3 Halaman Pre-Processed	44
Gambar 4.4 Halaman Confusion Matrix	45
Gambar 4.5 Halaman Performa	45
Gambar 4.6 Halaman Klasifikasi.....	46
Gambar 4.7 Halaman Klasifikasi setelah Input Data Uji	46

DAFTAR MODUL PROGRAM

Modul Program 4.1 Implementasi Cleansing	40
Modul Program 4.2 Implementasi Casefolding.....	40
Modul Program 4.3 Implementasi Stopwords Removal	41
Modul Program 4.4 Implementasi Tokenizing.....	41
Modul Program 4.5 Implementasi Stemming	41
Modul Program 4.6 Implementasi Pembobotan Kata TF-IDF	41
Modul Program 4.7 Implementasi K-NN	42
Modul Program 4.8 Implementasi Perhitungan Jarak Metriks.....	42
Modul Program 4.9 Implementasi Perhitungan Jarak Metriks (lanjutan)	43
Modul Program 4.10 Implementasi Klasifikasi Jurnal.....	43