

## DAFTAR ISI

HALAMAN JUDUL .....	ii
HALAMAN PENGESAHAN PEMBIMBING .....	iii
HALAMAN PENGESAHAN PENGUJI.....	iv
SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR.....	v
SURAT PERNYATAAN BEBAS PLAGIAT .....	vi
HALAMAN PERSEMBAHAN .....	vii
ABSTRAK .....	viii
KATA PENGANTAR.....	ix
DAFTAR ISI .....	x
DAFTAR TABEL .....	xii
DAFTAR GAMBAR.....	xv
DAFTAR MODUL PROGRAM.....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Tugas Akhir.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metodologi Penelitian dan Pengembangan Sistem .....	3
1.6.1 Metodologi Penelitian .....	3
1.6.2 Metode Pengembangan Sistem .....	4
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN LITERATUR .....	5
2.1 <i>Natural Language Processing</i> (NLP).....	5
2.2 <i>K-Nearest Neighbor</i> (KNN) .....	5
2.2.1 Kalkulasi <i>K-Nearest Neighbor</i> (KNN).....	6
2.3 <i>Latent Semantic Analysis</i> (LSA) .....	8
2.3.1 <i>Singular Value Decomposition</i> (SVD) .....	9
2.3.2 Kalkulasi <i>Latent Semantic Analysis</i> (LSA).....	10
2.4 Penelitian Terdahulu.....	14
BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM .....	14
3.1 Metodologi Penelitian .....	14
3.1.1 Pengumpulan Data .....	14
3.1.2 <i>Text Pre-Processing</i> .....	16
3.1.2.1 Sub-Proses <i>Case Folding</i> .....	17
3.1.2.2 Sub-Proses <i>Remove Punctuation</i> .....	18
3.1.2.3 Sub-Proses <i>Remove Number</i> .....	19
3.1.2.4 Sub-Proses <i>Tokenizing</i> .....	20
3.1.2.5 Sub-Proses <i>Stopwords Removal</i> .....	21
3.1.2.6 Sub-Proses <i>Stemming</i> .....	22
3.1.3 Ekstraksi Fitur .....	23

3.1.3.1	Sub-Proses Perhitungan TF, DF dan IDF.....	24
3.1.3.2	Sub-Proses Perhitungan TF-IDF .....	26
3.1.4	Pemilihan Fitur .....	27
3.1.4.1	Sub-Proses Perhitungan <i>Singular Value Decomposition</i> (SVD).....	27
3.1.4.2	Sub-proses Reduksi Fitur .....	33
3.1.5	Klasifikasi Pertanyaan <i>K-Nearest Neighbor</i> (KNN) .....	35
3.1.5.1	Model <i>K-Nearest Neighbor</i> dengan TF-IDF .....	36
3.1.5.2	Model <i>K-Nearest Neighbor</i> dengan <i>Latent Semantic Analysis</i> .....	38
3.1.6	Pengujian Penelitian .....	40
3.1.6.1	<i>K-Fold Cross Validation</i> .....	40
3.1.6.2	<i>Confusion Matrix</i> .....	41
3.1.6.3	Kurva ROC.....	43
3.2	Metode Pengembangan Sistem .....	44
3.2.1	Analisis Kebutuhan .....	45
3.2.1.1	Kebutuhan Pengguna.....	45
3.2.1.2	Domain .....	45
3.2.1.3	Kebutuhan Sistem.....	45
3.2.2	Perancangan Sistem.....	47
3.2.2.1	Arsitektur Sistem.....	47
3.2.2.2	Perancangan Proses .....	48
3.2.2.3	Perancangan Database .....	49
3.2.2.4	Perancangan Antarmuka.....	53
3.2.3	Pengujian Aplikasi .....	64
BAB IV	HASIL DAN PEMBAHASAN.....	65
4.1	Hasil Penelitian.....	65
4.1.1	Implementasi Klasifikasi Pertanyaan .....	65
4.1.1.1	Modul <i>Text Pre-Processing</i> .....	65
4.1.1.2	Modul Ekstraksi Fitur <i>Term Frequency - Inverse Document Frequency</i> .....	70
4.1.1.3	Modul Pemilihan Fitur <i>Latent Semantic Analysis</i> .....	73
4.1.1.4	Modul Klasifikasi Pertanyaan <i>K-Nearest Neighbor</i> (KNN) .....	78
4.1.1.5	Modul Fungsi <i>Retrieval</i> dan Ekstraksi Jawaban .....	84
4.1.2	Implementasi Aplikasi.....	85
4.1.2.1	Halaman User .....	85
4.1.2.2	Halaman Admin .....	87
4.1.3	Pengujian .....	96
4.1.3.1	Pengujian Penelitian .....	96
4.1.3.2	Pengujian Aplikasi .....	103
4.2	Pembahasan .....	104
BAB V	PENUTUP .....	105
5.1	Kesimpulan.....	105
5.2	Saran.....	105
DAFTAR PUSTAKA	.....	107
LAMPIRAN	.....	xvii