

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN PEMBIMBING .....	ii
HALAMAN PENGESAHAN PENGUJI .....	iii
SURAT PERNYATAAN .....	iv
KARYA ASLI TUGAS AKHIR .....	iv
PERNYATAAN BEBAS PLAGIASI .....	v
HALAMAN PERSEMBAHAN .....	vi
ABSTRAK.....	vii
ABSTRACT.....	viii
KATA PENGANTAR .....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL.....	xiv
DAFTAR MODUL.....	xv
BAB I.....	1
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.....	4
BAB II.....	6
TINJAUAN LITERATUR.....	6
2.1. Landasan Teori .....	6
2.1.1. Lesbian, Gay, Biseksual, Transgender (LGBT) di Indonesia .....	6
2.1.2. Text Mining, Klasifikasi Opini dan Analisis Sentimen .....	6
2.1.3. Algoritma Nazief Adriani .....	7
2.1.4. Naïve Bayes Classifier .....	7
2.1.5. Algoritma k-Nearest Neighbor (k-NN).....	9

2.1.6. Matriks Evaluasi .....	10
2.2. Studi Literatur .....	10
2.3. State of Art .....	12
<b>BAB III .....</b>	<b>14</b>
<b>METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM .....</b>	<b>14</b>
3.1. Metodologi Penelitian .....	14
3.1.1. Business Understanding .....	14
3.1.2. Data Understanding .....	15
3.1.3. Data Preprocessing .....	16
3.1.4. Model Building .....	27
3.1.5. Testing & Evaluation .....	38
3.1.6. Deployment .....	39
4.2. Metodologi Pengembangan Sistem (Prototype) .....	39
4.2.1. Communication .....	39
4.2.2. Quick Plan .....	40
4.2.3. Modeling Quick Design .....	41
4.2.4. Construction Prototype .....	42
4.2.5. Deployment and Delivery Feedback .....	42
<b>BAB IV .....</b>	<b>43</b>
<b>HASIL, PENGUJIAN, DAN PEMBAHASAN .....</b>	<b>43</b>
4.1. Hasil Penelitian .....	43
4.1.1. Implementasi Penelitian .....	43
a. Data Understanding .....	43
b. Data Preparation .....	44
c. Pemodelan .....	47
d. Evaluasi .....	50
4.1.2. Implementasi Pengembangan Perangkat Lunak .....	50
a. Halaman Eksplorasi dan Preprocessing .....	51
b. Halaman Komparasi Model .....	51
c. Halaman Pengujian Kalimat .....	52
4.2. Pengujian .....	53
a. Pengujian Algoritma .....	53
b. Pengujian Fungsionalitas Sistem .....	56
4.3. Pembahasan .....	56

BAB V .....	60
KESIMPULAN DAN SARAN.....	60
5.1. Kesimpulan .....	60
5.2. Saran .....	60
DAFTAR PUSTAKA .....	61

## DAFTAR GAMBAR

Gambar 3. 1 Metodologi Penelitian (CRISP-DM) .....	14
Gambar 3. 2 Diagram alir proses <i>Scrapping</i> .....	15
Gambar 3. 3Diagram alir proses <i>clean data</i> .....	17
Gambar 3. 4 Diagram alir <i>stopword removal</i> .....	23
Gambar 3. 5 <i>Diagram Alir Algoritma Nazief Adriani</i> .....	24
Gambar 3. 6 Diagram Alur TF-IDF .....	28
Gambar 3. 7 Diagram alir K-Nearest Neighbor Classifier .....	31
Gambar 3. 8 Diagram alir algoritma <i>naïve bayes</i> .....	34
Gambar 3. 9 Metodologi pengembangan sistem <i>Prototype</i> .....	39
Gambar 3. 10 Rancangan antarmuka halaman Eksplorasi.....	41
Gambar 3. 11 Rancangan antarmuka halaman Model .....	42
Gambar 4. 1 Halaman Eksplorasi dan Preprocessing .....	51
Gambar 4. 2 Halaman Komparasi Model (bagian K-NN).....	52
Gambar 4. 3 Halaman Komparasi Model (bagian Naive Bayes).....	52
Gambar 4. 4 Halaman pengujian kalimat ( <i>userable system</i> ) .....	53
Gambar 4. 5 Ilustrasi presentase labeling dataset .....	57

## DAFTAR TABEL

Tabel 3. 1 Jumlah data komentar yang digunakan.....	16
Tabel 3. 2 Contoh Dataset.....	16
Tabel 3. 3 Contoh dataset setelah <i>casefolding</i> .....	18
Tabel 3. 4 Contoh dataset setelah menghapus tautan .....	18
Tabel 3. 5 Contoh dataset setelah menghapus <i>mention</i> .....	19
Tabel 3. 6 Contoh dataset setelah menghapus simbol .....	20
Tabel 3. 7 Contoh data ditokenisasi .....	21
Tabel 3. 8 Contoh data terjemahan .....	22
Tabel 3. 9 Contoh data menghapus stopword.....	23
Tabel 3. 10 Contoh menghapus suffix .....	25
Tabel 3. 11 Contoh menghapus derivative suffix .....	25
Tabel 3. 12 Contoh menghapus derivative preffix.....	26
Tabel 3. 13 Contoh pemeriksaan disambiguitas kata.....	27
Tabel 3. 14 Contoh data setelah proses stemming .....	28
Tabel 3. 15 Contoh Hasil Perhitungan DF dan TF .....	29
Tabel 3. 16 Contoh Perhitungan TF-IDF Contoh Perhitungan TF-IDF .....	30
Tabel 3. 17 Contoh perhitungan jarak antar kalimat berdasarkan TF-IDF.....	32
Tabel 3. 18 Euclidean Distance Data Uji terhadap Data Latih .....	33
Tabel 3. 19 Contoh hasil klasifikasi berdasarkan nilai K .....	33
Tabel 3. 20 Contoh perhitungan <i>Prior Probability</i> .....	34
Tabel 3. 21 Data jumlah TF-IDF kata masing-masing kalimat .....	35
Tabel 3. 22 Bobot TF-IDF setiap kata uji untuk masing-masing kelas ( $W_{ct}$ ) .....	36
Tabel 3. 23 Contoh perhitungan logaritma prior probability dan <i>conditional probability</i> ..	37
Tabel 3. 24 Rancangan Tabel Hasil Pengujian .....	38
Tabel 3. 25 Kebutuhan Non Fungsional .....	40
Tabel 3. 26 Tabel Hasil Pengujian dan Evauasi .....	40
Tabel 3. 27 Rancangan matriks evaluasi fungsionalitas sistem.....	42
Tabel 4. 1 Hasil pengujian skenario menggunakan 60% data latih .....	53
Tabel 4. 2 Hasil pengujian skenario menggunakan 70% data latih .....	54
Tabel 4. 3 Hasil pengujian skenario menggunakan 80% data latih .....	54
Tabel 4. 4 Hasil pengujian skenario menggunakan 80% data latih, dan K-NN .....	56
Tabel 4. 5 Hasil pengujian skenario menggunakan 90% data latih .....	56
Tabel 4. 6 Hasil pengujian fungsionalitas sistem.....	56
Tabel 4. 7 Cuplikan hasil pengujian skenario menggunakan Naïve Bayes .....	57
Tabel 4. 8 Cuplikan hasil pengujian skenario terbaik.....	58

## DAFTAR MODUL

Modul 4. 1 Scraping komentar YouTube .....	43
Modul 4. 2 Inialisasi dataset, kamus <i>stopword</i> , dan <i>slangword</i> .....	44
Modul 4. 3 <i>Data cleaning</i> , tokenisasi, dan olah <i>slangword</i> .....	45
Modul 4. 4 Stemming Nazief Adriani.....	45
Modul 4. 5 <i>Remove Stopword</i> .....	46
Modul 4. 6 Preprocessing dengan Stemming Nazief Adriani.....	46
Modul 4. 7 Preprocessing tanpa stemming Nazief Adriani .....	46
Modul 4. 8 Pembobotan TF-IDF .....	47
Modul 4. 9 Spesifikasi Model .....	47
Modul 4. 10 Model K-Nearest Neighbor tanpa stemming Nazief Adriani.....	48
Modul 4. 11 Model K-Nearest Neighbor dengan stemming Nazief Adriani.....	48
Modul 4. 12 Model Naïve Bayes tanpa stemming Nazief Adriani.....	49
Modul 4. 13 Model Naive Bayes dengan stemming Nazief Adriani .....	49
Modul 4. 14 Evaluasi K-Nearest Neighbor tanpa Stemming .....	50