

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN TUGAS AKHIR</b> .....	ii
<b>HALAMAN PENGESAHAN PENGUJI SIDANG</b> .....	iii
<b>HALAMAN PERNYATAAN KEASLIAN</b> .....	iv
<b>KATA PENGANTAR</b> .....	v
<b>ABSTRAK</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR RUMUS</b> .....	xi
<b>DAFTAR TABEL</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xv
<b>DAFTAR SINTAK</b> .....	xvii

### **BAB 1 PENDAHULUAN**

1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah .....	3
1.3. Batasan Masalah .....	3
1.4. Tujuan Penelitian.....	4
1.5. Metode Pendekatan.....	5
1.6. Sistematika Penulisan.....	6

### **BAB 2 LANDASAN TEORI**

2.1. Data Mining .....	7
2.2. Text Mining .....	8
2.2.1. Case Folding .....	10
2.2.2. Tokenizing .....	10
2.2.3. Filtering / Stopword .....	10
2.2.4. Stemming .....	11
2.2.5. Tagging Dan Feature Selection .....	16

2.2.5.1. Feature Selection Data Mining .....	16
2.2.5.2. Feature Selection Text Mining .....	18
2.2.5.3. Chi Squared ( $x^2$ ) .....	21
2.2.5.3.1. Chi Squared Satu Sampel ( $x^2$ ) .....	21
2.2.5.3.2. Chi Squared Dua Sampel ( $x^2$ ) .....	22
2.2.5.3.3. Chi Squared k Sampel ( $x^2$ ) .....	23
2.2.5.3.4. Chi Squared Feature Selection .....	23
2.3. Klasifikasi .....	36
2.4. Microsoft Visual Basic .....	49
2.5. Microsoft SQL Server. ....	50
2.6. Model Perancangan Dan Pengujian .....	50
2.6.1. Model Perancangan .....	50
2.6.2. Pengujian .....	52

### **BAB 3 ANALISIS DAN PERANCANGAN SISTEM**

3.1. Bahasa Pemrograman yang Digunakan .....	53
3.2. Gambaran Kerja Sistem .....	53
3.2.1. Pembacaan Dokumen Teks .....	53
3.2.2. Pembuatan Classifier .....	54
3.2.2.1. Tokenizing Dan Case Folding .....	55
3.2.2.2. Stopword .....	56
3.2.2.3. Stemming .....	56
3.2.2.4. Feature Selection .....	58
3.2.2.5. Deklarasi Naïve Bayes .....	62
3.2.3. Klasifikasi Dokumen .....	63
3.2.3.1. Persentase Kata .....	64
3.2.2.5. Pengkategorian Dokumen .....	65
3.3. Interface .....	68
3.3.1. Form Utama .....	69

3.3.2. Form Persiapan Dokumen .....	71
3.3.3. Form Machine Learning .....	72
3.3.4. Form View Table .....	73
3.3.5. Form Chart .....	73
3.4 Database .....	74
3.4.1. Table Kamus .....	74
3.4.2. Table Imbuhan .....	74
3.4.3. Table Dokumen .....	75
3.4.4. Table Stopword .....	76
3.4.5. Table Stemming .....	77
3.4.1. Table Feature Selection .....	79
3.4.1. Table Hasil .....	82

#### **BAB 4 IMPLEMENTASI DAN PENGUJIAN**

4.1. Implementasi Sistem .....	83
4.1.1. Implementasi Sistem .....	83
4.1.2. Pembuatan Machine Learning .....	84
4.1.2.1. Tahap Tokenisasi Dan Case Folding .....	84
4.1.2.2. Tahap Stopword .....	85
4.1.2.3. Tahap Stemming .....	86
4.1.2.4. Feature Selection .....	89
4.1.2.5. Naïve Bayes Classifier .....	89
4.1.2.6. Klasifikasi Dokumen .....	90
4.1.2.6.1. Proses Presentase .....	91
4.1.2.6.2. Pengkategorian Dokumen .....	94
4.2. Pengujian Dan Analisis Sistem .....	97
4.2.1. Pengujian Pertama .....	98
4.2.2. Pengujian Kedua .....	99
4.2.3. Pengujian Ketiga .....	100

4.2.4. Pengujian Keempat .....	101
4.2.5. Pengujian Kelima .....	102
4.2.6. Pengujian Keenam .....	103
4.2.7. Pengujian Ketujuh .....	104
4.2.8. Pengujian Kedelapan .....	105
4.2.9. Pengujian Kesembilan .....	106
4.2.10. Pengujian Kesepuluh .....	107
4.2.11. Pengujian Rata-Rata .....	108
4.2.12. Rata-Rata Dan Hasil Pengujian .....	108

#### **BAB 5 KESIMPULAN DAN SARAN**

5.1. Kesimpulan .....	109
5.2. Saran .....	110

#### **DAFTAR PUSTAKA**

#### **LAMPIRAN**

#### **DAFTAR RIWAYAT HIDUP**