



## DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN BEBAS PLAGIASI .....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR .....	xii
DAFTAR TABEL.....	xiv
DAFTAR PERSAMAAN .....	xv
DAFTAR LAMPIRAN.....	xvi
ABSTRAK.....	xvi
ABSTRACT.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Metodologi Penelitian .....	4
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	7
BAB III LANDASAN TEORI.....	10
3.1 <i>Text Mining</i> .....	10
3.2 Tahapan <i>Text Mining</i> .....	11
3.2.1 <i>Preprocessing</i> teks .....	11
3.2.2 Seleksi fitur .....	12
3.2.3 Penemuan pola .....	18
3.3 N-gram.....	19
3.4 Pembobotan TF-IDF.....	19
3.5 Metode Klasifikasi .....	21
3.5.1 <i>Naïve Bayes Classifier</i> (NBC) .....	21
3.5.2 <i>Support Vector Machine</i> (SVM).....	23



3.6	Evaluasi .....	27
3.6.1	Akurasi .....	28
3.6.2	<i>Precision</i> .....	28
3.6.3	<i>Recall</i> .....	29
3.6.4	<i>F-Measure</i> .....	29
3.6.5	<i>K-Fold Cross Validation</i> .....	29
3.7	Twitter .....	29
3.8	Python.....	30
3.8.1	Pandas .....	32
3.8.2	Sastrawi.....	32
3.8.3	Scikit-learn (sklearn).....	33
3.9	Penghinaan .....	33
<b>BAB IV ANALISIS DAN RANCANGAN SISTEM.....</b>		<b>35</b>
4.1	Deskripsi Sistem.....	35
4.2	Data .....	36
4.2.1	Twit .....	36
4.2.2	Data stopwords .....	37
4.3	Pelabelan Data .....	38
4.4	Perancangan <i>Preprocessing</i> .....	38
4.4.1	Perancangan <i>cleansing</i> .....	39
4.4.2	Perancangan <i>case folding</i> .....	40
4.4.3	Perancangan tokenisasi .....	40
4.4.4	Perancangan perubahan kata tidak baku menjadi kata baku .....	41
4.4.5	Perancangan <i>stemming</i> .....	42
4.4.6	Perancangan <i>stopword remover</i> .....	43
4.4.7	Perancangan <i>filtering</i> .....	43
4.5	Perancangan Pembuatan <i>Vocabulary</i> .....	44
4.6	Perancangan Pembobotan TF-IDF .....	44
4.7	Perancangan Klasifikasi .....	44
4.8	Perencanaan Evaluasi Sistem Klasifikasi .....	46
<b>BAB V IMPLEMENTASI.....</b>		<b>47</b>
5.1	Spesifikasi Sistem.....	47
5.2	Implementasi Sistem .....	47
5.3	Implementasi Pelabelan Data .....	47



5.4	Implementasi <i>Preprocessing</i> .....	48
5.4.1	Implementasi <i>cleansing</i> .....	49
5.4.2	Implementasi <i>case folding</i> .....	49
5.4.3	Implementasi tokenisasi .....	49
5.4.4	Implementasi perubahan kata tidak baku menjadi kata baku .....	50
5.4.5	Implementasi <i>stemming</i> .....	50
5.4.6	Implementasi <i>stopword remover</i> .....	51
5.4.7	Implementasi <i>filtering</i> .....	51
5.4.8	Implementasi <i>remove multiple space</i> .....	51
5.4.9	Implementasi pelatihan .....	52
5.5	Implementasi Pembuatan <i>Vocabulary</i> dan Pembobotan TF-IDF.....	53
5.5.1	Implementasi unigram.....	53
5.5.2	Implementasi bigram.....	55
5.6	Implementasi Pemodelan Klasifikasi .....	56
5.6.1	Pemodelan <i>Multinomial Naïve Bayes</i> .....	57
5.6.2	Pemodelan <i>Support Vector Machine</i> .....	58
5.7	Implementasi Klasifikasi .....	59
5.7.1	Implementasi klasifikasi <i>Multinomial Naïve Bayes</i> .....	59
5.7.2	Implementasi klasifikasi <i>Support Vector Machine</i> .....	60
5.8	Implementasi Evaluasi .....	60
5.8.1	Implementasi evaluasi <i>Multinomial Naïve Bayes</i> .....	60
5.8.2	Implementasi evaluasi <i>Support Vector Machine</i> .....	61
BAB VI HASIL DAN PEMBAHASAN .....		63
6.1	Dataset .....	63
6.2	Hasil <i>Preprocessing</i> .....	65
6.3	Hasil Pembuatan <i>Vocabulary</i> .....	68
6.4	Hasil Pembobotan TF-IDF .....	68
6.5	Hasil Klasifikasi .....	70
6.6	Hasil Evaluasi.....	76
6.6.1	Hasil evaluasi model klasifikasi.....	76
6.6.2	Hasil evaluasi model klasifikasi data uji @tubirfess .....	78
6.6.3	Hasil evaluasi model klasifikasi data uji Ibrohim dan Budi (2018) .....	79
6.7	Penggunaan Model .....	80
BAB VII KESIMPULAN DAN SARAN .....		83



7.1	Kesimpulan.....	83
7.2	Saran.....	83
	DAFTAR PUSTAKA .....	84
	LAMPIRAN A : STOPWORDS.....	88
	LAMPIRAN B : KATA KASAR/KOTOR .....	92
	LAMPIRAN C : KAMUS KATA BAKU .....	93