

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
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 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	10
3.1 Data Mining.....	10
3.1.1 Text mining.....	10
3.2 Twitter.....	11
3.2.1 Web Scraping.....	11
3.3 Banjir.....	12
3.4 Data Pre-Processing.....	12
3.4.1 Cleansing.....	13
3.4.2 Case Folding.....	13
3.4.3 Stemming.....	13
3.4.4 Tokenisasi.....	17
3.4.5 Pembuangan <i>Stopword</i>	17
3.4.6 N-gram.....	17
3.5 Fitur dan Pembobotan.....	18
3.5.1 Term Presence (TP).....	18

3.5.2	Term Frequency (TF).....	19
3.5.3	Term Frequency Inverse Document Frequency (TF-IDF).....	19
3.6	Klasifikasi.....	20
3.6.1	Multinomial Naive Bayes	20
3.6.2	Support Vector Machine	22
3.7	Evaluasi Performa Model.....	24
3.7.1	K-Fold Cross Validation	25
BAB IV ARSITEKTUR DAN PERANCANGAN SISTEM		27
4.1	Rancangan Sistem	27
4.2	Deskripsi Sistem.....	28
4.3	Data dan Kelas Data	30
4.3.1	Data Stopword.....	30
4.3.2	Data <i>Tweet</i>	31
4.3.3	Data Kata Dasar	33
4.4	<i>Preprocessing</i> Data (Prapemrosesan Data).....	33
4.4.1	Perancangan <i>Cleansing</i>	34
4.4.2	Perancangan <i>Case Folding</i>	35
4.4.3	Perancangan <i>Stemming</i>	36
4.4.4	Perancangan Tokenisasi	38
4.4.5	Perancangan Pembuangan <i>Stopword</i>	38
4.4.6	Perancangan Running Time Preprocessing	40
4.5	Perancangan Pembuatan <i>Vocabulary</i>	40
4.6	Perancangan TF-IDF	40
4.7	Klasifikasi.....	40
4.7.1	Pelabelan Data.....	41
4.7.2	Perancangan <i>Training</i>	41
4.7.3	Perancangan <i>Testing</i>	43
4.7.4	Perancangan Evaluasi dan Validasi Klasifikasi	44
BAB V IMPLEMENTASI SISTEM.....		45
5.1	Spesifikasi Sistem Pengembangan	45
5.2	Implementasi Sistem	45
5.3	Implementasi Pelabelan Data	46
5.3.1	Implementasi <i>Cleansing</i> dan <i>Case Folding</i>	46
5.3.2	Implementasi <i>Stemming</i>	47

5.3.3	Implementasi Tokenisasi.....	48
5.3.4	Implementasi Pembuangan <i>Stopword</i>	48
5.4	Implementasi Running Time Preprocessing.....	49
5.5	Implementasi Pembuatan <i>Vocabulary</i>	49
5.6	Implementasi Pembobotan TF-IDF.....	50
5.7	Implementasi <i>Training</i>	51
5.8	Implementasi <i>Testing</i>	52
5.9	Implementasi Klasifikasi dan Evaluasi Model.....	52
5.9.1	Model Evaluasi Multinomial Naïve Bayes.....	53
5.9.2	Model Evaluasi Support Vector Machine.....	55
BAB VI HASIL PENELITIAN DAN PEMBAHASAN.....		57
6.1	Pengujian Sistem dan Hasil.....	57
6.2	Hasil Preprocessing.....	57
6.3	Hasil Vocabulary.....	58
6.4	Hasil Pembobotan TF-IDF.....	60
6.5	Hasil Identifikasi.....	61
6.6	Hasil Pengujian.....	64
BAB VII KESIMPULAN DAN SARAN.....		65
6.7	Kesimpulan.....	65
6.8	Saran.....	65
DAFTAR PUSTAKA.....		67
LAMPIRAN.....		70
A.	Rata-rata Akurasi Hasil Pengujian Multinomial Naïve Bayes.....	70
B.	Rata-rata Akurasi Hasil Pengujian Support Vector Machine.....	70
C.	Data Stopwords.....	71