

## DAFTAR ISI

LEMBAR PERSETUJUAN	
LEMBAR PENGESAHAN	
LEMBAR PERNYATAAN KARYA SENDIRI	
LEMBAR PERSEMBAHAN	
ABSTRAK .....	i
<i>ABSTRACT</i> .....	ii
KATA PENGANTAR .....	iii
DAFTAR ISI .....	vi
DAFTAR TABEL .....	ix
DAFTAR GAMBAR .....	x
BAB I .....	1
PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan .....	5
1.4 Batasan Masalah .....	6
1.5 Kerangka Pemikiran .....	7
1.6 Sistematika Penulisan .....	8
BAB II .....	10
KAJIAN LITERATUR .....	10
2.1 Tinjauan Pustaka .....	10
2.2 Landasan Teori .....	14
2.2.1 <i>Text Mining</i> .....	14
2.2.2 Analisis Sentimen .....	14

2.2.3	Python .....	15
2.2.4	Web Scraping .....	15
2.2.5	Twitter .....	16
2.2.6	Bansos (Bantuan Sosial) .....	17
2.2.7	<i>Ensemble Learning</i> .....	18
2.2.8	Algoritma Random Forest .....	18
2.2.9	<i>TF-IDF</i> .....	21
2.2.10	<i>Polarity dan Subjectivity</i> .....	22
2.2.11	<i>Confusion Matrix</i> .....	23
2.2.12	<i>Cross Validation</i> .....	26
2.2.13	CRISP-DM .....	28
BAB III .....		32
METODOLOGI PENELITIAN .....		32
3.1	Penggambaran Umum Metode CRISP-DM .....	32
3.2	<i>Business Understanding</i> .....	33
3.3	<i>Data Understanding</i> .....	34
3.4	<i>Data Preparing</i> .....	34
3.5	<i>Modelling</i> .....	35
3.6	Evaluasi .....	36
BAB IV .....		37
HASIL DAN PEMBAHASAN .....		37
4.1	<i>Business Understanding</i> .....	37
4.2	<i>Data Understanding</i> .....	41
4.3	<i>Data Preparing</i> .....	45
4.3.1.	Pemilihan Data .....	45
4.3.2.	Pembersihan Data .....	46

4.3.3.	<i>Data Construct/Labelling</i> .....	50
4.3.4.	<i>Formatting</i> .....	51
4.4	Modelling Random Forest.....	52
4.5	Evaluasi .....	58
4.5.1.	Persentase Analisis Sentimen.....	58
4.5.2.	Topik Analisis Sentimen .....	60
4.5.3.	Performa Random Forest .....	62
BAB V	.....	66
PENUTUP	.....	66
5.1	Kesimpulan.....	66
5.2	Saran.....	68
DAFTAR PUSTAKA	.....	69
LAMPIRAN	.....	75
DAFTAR RIWAYAT HIDUP	.....	89

## DAFTAR TABEL

Tabel 2. 1 State of The Art.....	13
Tabel 2. 2 Parameter Random Forest [23] .....	20
Tabel 2. 3 Contoh Polarity dan Subjectivity [29] .....	23
Tabel 4. 1 Ketentuan Data Tweet.....	39
Tabel 4. 2 Process, Tools, and Libraries .....	40
Tabel 4. 3 Contoh Hasil Case Folding .....	46
Tabel 4. 4 Contoh Hasil Data Cleaning .....	47
Tabel 4. 5 Contoh Hasil Tokenizing .....	48
Tabel 4. 6 Contoh Hasil Stopwords Removal.....	48
Tabel 4. 7 Contoh Hasil Stemming.....	49
Tabel 4. 8 Contoh Hasil Remove Punctuation.....	49
Tabel 4. 9 Nilai Rata-Rata Subjectivity .....	50
Tabel 4. 10 Perhitungan Pembobotan TF-IDF.....	51
Tabel 4. 11 Contoh Sederhana Data Training Random Forest .....	53
Tabel 4. 12 Contoh perhitungan gini 1.....	54
Tabel 4. 13 Contoh Perhitungan gini 2 .....	55
Tabel 4. 14 Contoh Perhitungan gini 3 .....	56
Tabel 4. 15 Contoh Voting Keputusan Prediksi .....	57
Tabel 4. 16 Hasil Pengujian .....	63

## DAFTAR GAMBAR

Gambar 1. 1 Kerangka Pemikiran.....	7
Gambar 2. 1. Standar Parameter Random Forest [22] .....	21
Gambar 2. 2 Algoritma Random Forest[23] .....	22
Gambar 2. 4 Ilustrasi Cross Validation.....	29
Gambar 2. 5 Metode CRISP-DM[10] .....	30
Gambar 3. 1 Penggambaran Umum CRISP-DM.....	32
Gambar 4. 1 Detail Atribut Data Tweet.....	41
Gambar 4. 2 Tweet/Hari 5-12 Mei.....	42
Gambar 4. 3 Tweet/Hari 13-20 Mei.....	42
Gambar 4. 4 Tren Hashtag .....	43
Gambar 4. 5 Source Terbanyak.....	44
Gambar 4. 6 User Terbanyak .....	44
Gambar 4. 7 Perbandingan Jumlah <i>Tweet</i> dan <i>Retweet</i> .....	45
Gambar 4. 8 Persentase Sentimen Labelling.....	50
Gambar 4. 9 Contoh Cabang Pohon Perhitungan gini 1 .....	54
Gambar 4. 10 Contoh Cabang Pohon Perhitungan gini 2 .....	55
Gambar 4. 11 Contoh Cabang Pohon Perhitungan gini 3 .....	56
Gambar 4. 12 Persentase Sentimen Setelah Klasifikasi.....	57
Gambar 4. 13 Word Cloud Positif.....	58
Gambar 4. 14 Word Cloud Negatif .....	58
Gambar 4. 15 Word Cloud Netral.....	58