

**ANALISIS SENTIMEN TERHADAP OPINI PUBLIK MELALUI JEJARING
SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES**

Alfarizy M. G. (0927050)

Jurusan Sistem Komputer, Fakultas Teknik,

Universitas Kristen Maranatha

Jalan Prof. Drg. Suria Sumantri MPH No. 65

Bandung 40164

ABSTRAK

Manusia adalah makhluk sosial dan mereka berkomunikasi satu sama lain. Saat ini, orang banyak berinteraksi menggunakan media sosial salah satunya Twitter. Bahkan Presiden Indonesia menggunakan Twitter untuk mengungkapkan perasaan, sentimen, dan ide menurut aspek yang berbeda setiap hari. Banyak tokoh politik menggunakan Twitter di Indonesia seperti Gubernur DKI Jakarta Basuki Tjahaja Purnama (Ahok) dengan *basuki_btp* sebagai *username*-nya.

Ahok adalah tokoh politik kontroversial dengan gayanya memimpin Jakarta. Seringkali ia bersemangat mengungkapkan perasaannya bahkan cenderung marah. Penelitian ini dapat menentukan sentimen masyarakat terhadap Ahok. Penelitian ini menggunakan tiga kategori sentimen: positif, negatif, dan netral. Algoritma Naive Bayes dapat mengklasifikasikan sentimen orang terhadap Ahok. Dengan metode *cleansing*, *case folding*, *parsing*, *stopping*, dan metode *term frequency* untuk menyaring *input* teks.

**SENTIMENT ANALYSIS TO PUBLIC FIGURE BY SOCIAL MEDIA TWITTER
USING NAIVE BAYES METHODE**

Alfarizy M. G. (0927050)

Jurusan Sistem Komputer, Fakultas Teknik,

Universitas Kristen Maranatha

Jalan Prof. Drg. Suria Sumantri MPH No. 65

Bandung 40164

ABSTRACT

Humans are social beings and they loved to communicate with each other. Nowadays people interact using social media such as Twitter. Even Indonesian president used Twitter to express their feelings, sentiments, and idea of different aspects everyday. Many political figures in Indonesia used Twitter such as Governor of DKI Jakarta Basuki Tjahaja Purnama (Ahok) with @basuki_btp as his username.

Ahok is a controversial political figure with his style to lead Jakarta. He often express his feelings passionately even more angry. This research can determined sentiment of people to Ahok. This research used three class of sentiment: positive, negative, and netral. Naive bayes algorithm can classify sentiment of people to Ahok. This cleansing, case folding, parsing, stopping, and term frequency method to filter the text input.

DAFTAR ISI

LEMBAR PENGESAHAN.....	i
PERNYATAAN ORISINALITAS LAPORAN TUGAS AKHIR.....	v
PERNYATAAN PUBLIKASI LAPORAN TUGAS AKHIR	i
ABSTRAK	vi
<i>ABSTRACT</i>	viii
KATA PENGANTAR.....	vi
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiv
DAFTAR ISI	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi masalah	3
1.3 Tujuan	3
1.4 Pembatasan Masalah.....	3
1.5 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Analisis Sentimen	6
2.2 <i>Text Mining</i>	8
2.2.1 <i>Text Preprocessing</i>	8
2.2.2 Pemilihan Fitur	8
2.3 Twitter.....	11
2.4 Pembobotan	14

2.5	Algoritma Naive Bayes.....	17
2.6	Flowchart (Diagram Alir).....	17
2.7	Database MySQL.....	17
2.7.1	SQLyog.....	21
2.8	Pemrograman Java.....	22
2.8.1	Kelebihan Pemrograman Java	23
2.8.2	Kekurangan Pemrograman Java	25
2.8.3	Konsep Pemrograman Berorientasi Objek	26
2.8.3.1	Abstraksi.....	26
2.8.3.2	Pembungkusan.....	27
2.8.3.3	Pewarisan.....	27
2.8.3.4	Polimorfisme	28
2.9	Regex	28
2.9	NetBeans.....	29
BAB III PERANCANGAN.....		31
3.1	Perancangan Database	31
3.1.1	Tabel TWEETMASTER	32
3.1.2	Tabel TWEETNGRAM.....	32
3.1.3	Tabel STOPLIST	33
3.1.4	Tabel SENTIMEN.....	34
3.1.5	Tabel PROBABILITAS	34
3.1.6	Tabel INTISARI	35
3.2	Data Tweet.....	36
3.3	Pengambilan Data dari Twitter.....	37
3.4	Text Preprocessing.....	45

3.5	Pemilihan Fitur	48
3.6	Pembobotan	52
3.6	Contoh Penerapan Metode Naive Bayes Classifier	54
BAB IV PENGUJIAN.....		59
4.1	Spesifikasi <i>Hardware</i> dan <i>Software</i> yang Digunakan	59
4.2	Database.....	60
4.2.1	Tabel TWEETMASTER	60
4.2.2	Tabel TWEETNGRAM.....	61
4.2.3	Tabel PROBABILITAS	62
4.2.4	Tabel INTISARI.....	63
4.2.5	Tabel STOPLIST	64
4.2.6	Tabel SENTIMEN.....	66
4.3	Implementasi Sistem.....	66
4.3.1	Tampilan Tweet Ngram	67
4.3.2	Tampilan Hasil Pemilihan Fitur & Perhitungan Probabilitas.....	68
4.3.3	Tampilan Perhitungan Probabilitas Tertinggi.....	69
4.3.4	Tampilan Data <i>Training</i>	70
4.3.5	Tampilan Kumpulan <i>Stopword</i>	71
4.4	Hasil Akhir Klasifikasi	72
BAB V KESIMPULAN & SARAN.....		74
5.1	Kesimpulan	74
5.2	Saran	74
Daftar Pustaka		75
LAMPIRAN A		A1
LAMPIRAN B.....		B1

LAMPIRAN C.....	C1
LAMPIRAN D	D1
LAMPIRAN E.....	E1



DAFTAR GAMBAR

Gambar 2.1 Logo Twitter	11
Gambar 2.2 Logo SQLyog	21
Gambar 2.3 Tampilan SQLyog	22
Gambar 2.4 Logo Java.....	23
Gambar 2.5 Logo NetBeans	29
Gambar 3.1 Relasi antar tabel	31
Gambar 3.2 Skema pengambilan tweet	37
Gambar 3.3 Halaman <i>Application Management</i>	38
Gambar 3.4 <i>Keys and access tokens</i>	39
Gambar 3.5 <i>Consumer keys and secret</i>	40
Gambar 3.6 <i>Access tokens and secret</i>	40
Gambar 3.7 Implementasi pengambilan data dari Twitter	41
Gambar 3.8 flowchart penelitian	44
Gambar 3.9 Flowchart Text Preprocessing.....	46
Gambar 3.10 Contoh kalimat yang akan di input.....	47
Gambar 3.11 Kalimat setelah melalui tahap <i>cleansing</i>	47
Gambar 3.12 Kalimat setelah melalui tahap <i>case folding</i>	47
Gambar 3.13 Implementasi <i>text preprocessing</i> menggunakan REGEX	48
Gambar 3.14 flowchart pemilihan fitur <i>stopping</i>	50
Gambar 3.15 Hitung $P(x_i V_j)$	54
Gambar 4.1 Tabel TWEETMASTER	61
Gambar 4.2 Tabel TWEETNGRAM.....	62
Gambar 4.3 Tabel PROBABILITAS	63
Gambar 4.4 Tabel INTISARI	64
Gambar 4.5 Tabel STOPLIST	65
Gambar 4.6 Tabel SENTIMEN	66
Gambar 4.7 Menjalankan program.....	67
Gambar 4.8 Tampilan tweet ngram	68
Gambar 4.9 Tampilan probabilitas	69

Gambar 4.10 Tampilah V_{MAP}	70
Gambar 4.11 Tampilan data training.....	71
Gambar 4.12 Tampilan kumpulan stopword	72
Gambar 4.13 Tampilan hasil klasifikasi	73



DAFTAR TABEL

Tabel 2.1 Contoh pemotongan n-gram berbasis kata	10
Tabel 2.2 Simbol-simbol <i>flowchart</i>	17
Tabel 2.3 Simbol-simbol <i>flowchart</i> (lanjutan)	18
Tabel 3.1 Field pada tabel TWEETMASTER.....	32
Tabel 3.2 Field pada tabel TWEETNGRAM	33
Tabel 3.3 Field pada tabel STOPLIST	34
Tabel 3.4 Field pada tabel SENTIMEN	34
Tabel 3.5 Field pada tabel PROBABILITAS	35
Tabel 3.6 Field pada tabel INTISARI	36
Tabel 3.7 Hasil dari <i>text preprocessing</i>	48
Tabel 3.8 Hasil dari pemilihan fitur <i>stopping</i>	51
Tabel 3.9 Kumpulan <i>stopword</i>	51
Tabel 3.10 Hasil pemilihan fitur	52
Tabel 3.11 Tabel SENTIMEN	53
Tabel 3.12 Contoh perhitungan nilai probabilitas	55
Tabel 3.13 Nilai probabilitas kata	56
Tabel 3.14 Isi tabel PROBABILITAS	56
Tabel 3.15 Isi tabel INTISARI	57
Tabel 3.16 Isi tabel TWEETMASTER	58