

**ANALISIS SENTIMEN PADA TWITTER
MENGUNAKAN METODE *MAXIMUM ENTROPY*
DAN *SUPPORT VECTOR MACHINE***



**OLEH :
MONA CINDO
09042681721006**

**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2019**

**ANALISIS SENTIMEN PADA TWITTER
MENGUNAKAN METODE *MAXIMUM ENTROPY*
DAN *SUPPORT VECTOR MACHINE***

TESIS

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister**



**OLEH :
MONA CINDO
09042681721006**

**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2019**

LEMBAR PENGESAHAN
ANALISIS SENTIMEN PADA TWITTER
MENGGUNAKAN METODE *MAXIMUM ENTROPY* DAN
SUPPORT VECTOR MACHINE

TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

OLEH :
MONA CINDO
09042681721006

Palembang, Juli 2019

Pembimbing 1,



Dian Palupi Rini, M.Kom., Ph.D.
NIP.197802232006042002

Pembimbing 2,



Dr. Ermatita, M.Kom.
NIP.196709132006042001

Mengetahui,
Ketua Program Studi Magister Teknik Informatika



Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

HALAMAN PERSETUJUAN

Pada hari kamis tanggal 18 Juli 2019 telah dilaksanakan ujian sidang tesis oleh Magister Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Mona Cindo

NIM : 09042681721006

Judul : Analisis Sentimen pada Twitter Menggunakan Metode
Maximum Entropy dan Support Vector Machine

1. Pembimbing I

Dian Palupi Rini, M.Kom., Ph.D.
NIP. 197802232006042002



.....

2. Pembimbing II

Dr. Ermatita, M.Kom.
NIP. 196709132006042001



.....

3. Penguji I

Dr. Reza Firsandaya Malik, M.T.
NIP. 197604252010121001



.....

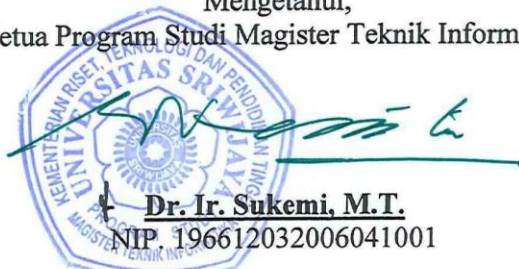
4. Penguji II

Dr. Iwan Pahendra, M.T.
NIP. 19740322200212002



.....

Mengetahui,
Ketua Program Studi Magister Teknik Informatika



Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

LEMBAR PERYATAAN

Yang bertanda tangan di bawah ini :

Nama : Mona Cindo
NIM : 09042681721006
Program Studi : Magister Teknik Informatika
Judul Tesis : Analisis Sentimen pada Twitter Menggunakan Metode
Maximum Entropy dan *Support Vector Machine*

Hasil Pengecekan Software iThenticate/Turnitin : 8%

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis ini, maka saya bersedia menerima sanksi akademik dari universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, Juli 2019



Mona cindo
NIM. 09042681721006

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT atas rahmat-nya sehingga penulis dapat menyelesaikan tesis yang berjudul “**Analisis Sentimen pada Twitter Menggunakan Metode *Maximum Entropy* dan *Support Vector Machine***” di susun untuk memenuhi sebagian persyaratan kelulusan untuk memperoleh gelar Magister Komputer pada Program Studi Magister Teknik Informatika Universitas Sriwijaya.

Pada kesempatan ini penulis menyadari keterbatasan dan kelemahan yang ada dalam menyelesaikan tesis ini sehingga penulis ingin menyampaikan ucapan terima kasih kepada pihak-pihak yang telah memberikan dukungan, bimbingan dan motivasi kepada penulis untuk menyelesaikan tesis ini, kepada :

1. Bapak Jaidan Jauhari, S.Pd, M.T selaku dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Bapak Dr. Ir. Sukemi, M.T., sebagai Ketua Program Studi Magister Teknik Informatika Universitas Sriwijaya.
3. Ibu Dian Palupi Rini, M.Kom., Ph.D dan Dr. Ermatita, M.Kom selaku pembimbing yang telah meluangkan waktu, bantuan serta saran dan kritiknya dalam penyusunan tesis ini.
4. Terkhusus kepada kedua orang tua, keluarga dan fadli atas semua dukungan dan bantuan yang tidak dapat dihitng dan dituliskan satu persatu.
5. Seluruh teman-teman dan staf Magister Teknik Informatika yang telah membantu dan memberikan semangat pada masa-masa perkuliahan.

Akhir kata, penulis menyadari bahwa tesis ini masih banyak kekurangan baik dari isi maupun susunan. Semoga tesis ini dapat bermanfaat untuk kita semua.

Palembang, Juli 2019

Penulis

LEMBAR PERSEMBAHAN

*“Allah menghendaki kemudahan bagimu, dan tidak
meghendaki kesukaran bagimu”*

(QS. Al-Baqarah: 185)

*“Allah tidak membebani seseorang melainkan sesuai dengan
kesanggupannya”*

(QS. Al-Baqarah: 286)

Kupersembahkan untuk

Ibu dan ayah

SENTIMENT ANALYSIS ON TWITTER USING MAXIMUM ENTROPY AND SUPPORT VECTOR MACHINE METHOD

ABSTRACT

Social media such as Twitter are often used by users to give opinions about the quality of products they use, express their emotions every day and some companies such as airlines use to gain competitive advantage by continuing to improve service and experience of service users by using twitter to become an important source for tracking sentiment analysis. Sentiment analysis is done to see opinions or tendencies of opinion on a problem. This study uses two types of datasets, namely public opinion data and airline opinion data to test the best features and methods. Feature extraction is expected to improve the accuracy of sentiment analysis. To get the best results this research uses two methods, namely maximum entropy and support vector machine as a comparison. The best results are found in the maximum entropy method which has an average accuracy of 85.8 % for the general opinion dataset while in the airline's opinion dataset the accuracy reaches 92.6%.

Keywords: *Twitter, Support Vector Machine, Maximum Entropy, Feature Extraction*

Palembang, Juli 2019

Pembimbing 1,



Dian Palupi Rini, M.Kom., Ph.D.
NIP.197802232006042002

Pembimbing 2,



Dr. Ermatita, M.Kom.
NIP.196709132006042001

Mengetahui,
Ketua Program Studi Magister Teknik Informatika



† Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

ANALISIS SENTIMEN PADA TWITTER MENGUNAKAN METODE *MAXIMUM ENTROPY* DAN *SUPPORT VECTOR MACHINE*

ABSTRAK

Media sosial seperti twitter sering dimanfaatkan penggunanya untuk memberikan pendapat tentang kualitas produk yang gunakan, mengekspresikan emosi mereka sehari-hari dan beberapa perusahaan seperti maskapai penerbangan memanfaatkan untuk mendapatkan keunggulan kompetitif dengan terus meningkatkan layanan dan pengalaman pengguna layanan dengan menggunakan twitter untuk menjadi sumber penting untuk melacak analisis sentimen. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah. Penelitian ini menggunakan dua tipe dataset yaitu data opini umum dan data opini maskapai untuk menguji fitur serta metode terbaik. Pada ekstraksi fitur diharapkan dapat meningkatkan akurasi dari analisis sentimen. Agar mendapatkan hasil yang terbaik penelitian ini menggunakan dua metode yaitu *maximum entropy* dan *support vector machine* sebagai pembandingan. Hasil terbaik terdapat pada metode *maximum entropy* yang memiliki akurasi rata-rata 85,8% untuk dataset opini umum sedangkan pada dataset opini maskapai dengan akurasi mencapai 92,6%.

Kata Kunci: *Twitter, Support Vector Machine, Maximum Entropy, Ekstraksi Fitur*

Palembang, Juli 2019

Pembimbing 1,



Dian Palupi Rini, M.Kom., Ph.D.
NIP.197802232006042002

Pembimbing 2,



Dr. Ermatita, M.Kom.
NIP.196709132006042001

Mengetahui,
Ketua Program Studi Magister Teknik Informatika



Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
LEMBAR PERNYATAAN	iv
KATA PENGANTAR	v
LEMBAR PERSEMBAHAN	vi
ABSTRACT	vii
ABSTRAK	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR SINGKATAN	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terkait	6
2.2 Landasan Teori	8
2.2.1 Twitter	8
2.2.2 Analisis Sentimen	9
2.2.2.1 Analisis Sentimen pada Twitter	10
2.2.2.2 Metode Klasifikasi Analisis Sentimen	11
2.2.3 <i>Maximum Entropy</i>	11
2.2.4 <i>Support Vector Machine</i>	13
2.2.5 <i>Cross Validation</i>	15
2.2.6 <i>Confusion Matrix</i>	16
2.3 Kerangka Pemikiran	18
BAB III METODE PENELITIAN	20
3.1 Pengumpulan Data	21
3.2 <i>Pre-procesing</i> Data	21
3.3 Ekstraksi Fitur	24
3.4 Klasifikasi Kalimat	27
3.6 Evaluasi	27

BAB IV HASIL DAN ANALISA	29
4.1 Hasil Pengolahan Dataset	29
4.2 Hasil Analisis Sentimen	33
4.2.1 Hasil Analisis Sentimen Opini Umum	34
4.2.2 Hasil Analisis Sentimen Opini Maskapai	37
4.3 Hasil Ekstraksi Fitur	39
4.3.1 Hasil Ekstraksi Fitur <i>Pragmatic</i>	40
4.3.1.1 Fitur <i>Pragmatic</i> Data Opini Umum	40
4.3.1.2 Fitur <i>Pragmatic</i> Data Opini Maskapai	42
4.3.2 Hasil Ekstraksi Fitur <i>Lexical-grams</i>	43
4.3.2.1 Fitur <i>Lexical-grams</i> Data Opini Umum	44
4.3.2.2 Fitur <i>Lexical-grams</i> Data Opini Maskapai	46
4.3.3 Hasil Ekstraksi Fitur <i>POS-Grams</i>	48
4.3.3.1 Fitur <i>POS-Grams</i> Data Opini Umum	48
4.3.3.2 Fitur <i>POS-Grams</i> Data Opini Maskapai	50
4.3.4 Hasil Ekstraksi Fitur Sentimen <i>Lexical</i>	52
4.3.4.1 Fitur Sentimen <i>Lexical</i> Data Opini Umum	52
4.3.4.2 Fitur Sentimen <i>Lexical</i> Data Opini Maskapai	54
4.4 Analisis Ekstraksi Fitur	55
4.5 Analisis Fitur Kombinasi	58
4.5.1 Analisis Dua Fitur Kombinasi	58
4.5.1.1 Hasil Matrik Dua Fitur Kombinasi Opini Umum	59
4.5.1.2 Hasil Akurasi Dua Kombinasi Fitur Opini Umum	63
4.5.1.3 Hasil Matrik Dua Fitur Kombinasi Opini Maskapai	64
4.5.1.4 Hasil Akurasi Dua Kombinasi Fitur Opini Maskapai	67
4.5.2 Analisis Tiga Fitur Kombinasi	69
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	76
5.2 Saran	76
DAFTAR PUSTAKA	78

DAFTAR GAMBAR

	Halaman
Gambar 2.1 <i>Linier Karnel Support Vector Machine</i>	14
Gambar 2.2 <i>Non-Linier Karnel Support Vector Machine</i>	15
Gambar 2.3 Pembagian Data pada <i>Cross Validation</i>	16
Gambar 2.4 Kerangka Pemikiran pada Analisis Sentimen pada Twitter	19
Gambar 3.1 Alur Proses Analisis Sentimen	20
Gambar 4.1 Developer Dalam Mengambil Data dari Twitter	30
Gambar 4.2 <i>Consumer API keys</i> dan Akses Token Data dari Twitter	30
Gambar 4.3 Tampilan dari <i>Library Tweepy</i>	31
Gambar 4.4 Dataset Twitter Mentah Opini Umum	32
Gambar 4.5 Dataset Twitter Mentah Opini Maskapai	32
Gambar 4.6 Alur Proses <i>10 Fold Cross Validation</i>	34
Gambar 4.7 Hasil Akurasi Analisis Sentimen Opini Umum	35
Gambar 4.8 Hasil Akurasi <i>Fold 4</i> Dastaset Opini Umum	36
Gambar 4.9 <i>Confusion Matrix</i> Opini Umum	36
Gambar 4.10 Hasil Akurasi Analisis Sentimen Opini Maskapai	38
Gambar 4.11 <i>Confusion Matrix</i> Opini Maskapai	38
Gambar 4.12 <i>Confusion Matrix</i> Fitur <i>Pragmatic</i> Opini Umum	40
Gambar 4.13 Hasil Akurasi Fitur <i>Pragmatic</i> Opini Umum	41
Gambar 4.14 <i>Confusion Matrix</i> Fitur <i>Pragmatic</i> Opini Maskapai	42
Gambar 4.15 Hasil Akurasi Fitur <i>Pragmatic</i> Opini Maskapai	43
Gambar 4.16 <i>Confusion Matrix</i> Fitur <i>Lexical-Grams</i> Opini Umum	44
Gambar 4.17 Hasil Akurasi Fitur <i>Lexical-Grams</i> Opini Umum	45
Gambar 4.18 <i>Confusion Matrix</i> Fitur <i>Lexical-Grams</i> Opini Maskapai	46
Gambar 4.19 Hasil Akurasi Fitur <i>Lexical-Grams</i> Opini Maskapai	47
Gambar 4.20 <i>Confusion Matrix</i> Fitur <i>POS-Grams</i> Opini Umum	48
Gambar 4.21 Hasil Akurasi Fitur <i>POS-Grams</i> Data Opini Umum	49
Gambar 4.22 <i>Confusion Matrix</i> Fitur <i>POS-Grams</i> Opini Maskapai	50
Gambar 4.23 Hasil Akurasi Fitur <i>POS-Grams</i> Opini Maskapai	51
Gambar 4.24 <i>Confusion Matrix</i> Fitur Sentimen <i>Lexical</i> Opini Umum	52
Gambar 4.25 Hasil Akurasi Fitur Sentimen <i>Lexical</i> Opini Umum	53
Gambar 4.26 <i>Confusion Matrix</i> Fitur Sentimen <i>Lexical</i> Opini Makapai	54
Gambar 4.27 Hasil Akurasi Fitur Sentimen <i>Lexical</i> Opini Maskapai	55
Gambar 4.28 Hasil Akurasi Ekstraksi Fitur Opini Umum	56
Gambar 4.29 Hasil Akurasi Ekstraksi Fitur Opini Maskapai	57
Gambar 4.30 Hasil Akurasi Kombinasi 2 Fitur Dataset Opini Umum	63
Gambar 4.31 Hasil Akurasi Kombinasi 2 Fitur Opini Maskapai	68
Gambar 4.32 Hasil Akurasi Kombinasi 3 Fitur Opini Umum	70
Gambar 4.33 Hasil Akurasi Kombinasi 3 Fitur Opini Maskapai	73

DAFTAR TABEL

	Halaman
Tabel 1 Macam-Macam Fungsi <i>Kernel Support Vector Machine</i>	13
Tabel 2 Model <i>Confusion Matrix</i>	17
Tabel 3 Contoh Kalimat pada <i>Pre-posesing</i> Tokenisasi	22
Tabel 4 Contoh Kalimat pada <i>Part-of-Speech Tagging</i>	22
Tabel 5 Contoh Simbol <i>Emoticon</i> pada <i>Part-of-Speech Tagging</i>	23
Tabel 6 Contoh <i>Output</i> dari Fitur <i>Pragmatic</i>	25
Tabel 7 Contoh Hasil <i>Output</i> Kalimat Fitur <i>Lexical Grams</i>	26
Tabel 8 Contoh Hasil <i>Output</i> Kalimat Fitur <i>POS-grams</i>	26
Tabel 9 Contoh Hasil <i>Output</i> Kalimat Fitur Sentimen <i>Lexical</i>	27
Tabel 10 Pembagian Dataset Twitter	29
Tabel 11 Data Sebelum dan Sesudah di <i>Preprocessing</i>	33
Tabel 12 Hasil Evaluasi Matrik pada Opini Umum	37
Tabel 13 Hasil Evaluasi Matrik pada Opini Maskapai	39
Tabel 14 Hasil Evaluasi Matrik Fitur <i>Pragmatic</i> Opini Umum	40
Tabel 15 Hasil Evaluasi Matrik Fitur <i>Pragmatic</i> Opini Maskapai	42
Tabel 16 Hasil Evaluasi Matrik Fitur <i>Lexical-Grams</i> Opini Umum	45
Tabel 17 Hasil Evaluasi Matrik Fitur <i>Lexical-Grams</i> Opini Maskapai	46
Tabel 18 Hasil Evaluasi Matrik Fitur <i>POS-Grams</i> Opini Umum	49
Tabel 19 Hasil Evaluasi Matrik Fitur <i>POS-Grams</i> Opini Maskapai	50
Tabel 20 Hasil Evaluasi Matrik Fitur Sentimen <i>Lexical</i> Opini Umum	53
Tabel 21 Hasil Evaluasi Matrik Fitur Sentimen <i>Lexical</i> Opini Maskapai	54
Tabel 22 Susunan dan Total Atribut Ekstraksi Fitu	56
Tabel 23 Susunan dan Total dari Atribut Kombinasi 2 Fitur	59
Tabel 24 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>Lexical-Grams</i> Opini Umum	60
Tabel 25 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>Sentimen lexical</i> Opini Umum	60
Tabel 26 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>POS-gram</i> Opini Umum	61
Tabel 27 Evaluasi Matrik Fitur <i>Lexical-grams</i> dan <i>Sentimen lexical</i> Opini Umum	61
Tabel 28 Evaluasi Matrik Fitur <i>POS-grams</i> dan <i>Sentimen lexical</i> Opini Umum	62
Tabel 29 Evaluasi Matrik Fitur <i>Lexical-grams</i> dan <i>POS-grams</i> Opini Umum	62
Tabel 30 Hasil Evaluasi Matrik Kombinasi 2 Fitur Dataset Opini Umum	64
Tabel 31 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>Lexical-Grams</i> Opini Maskapai	64
Tabel 32 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>Sentimen lexical</i> Opini Maskapai	65
Tabel 33 Evaluasi Matrik Fitur <i>Pragmatic</i> dan <i>POS-gram</i> Opini Maskapai	65
Tabel 34 Evaluasi Matrik Fitur <i>Lexical-grams</i> dan <i>Sentimen lexical</i> Opini Maskapai	66
Tabel 35 Evaluasi Matrik Fitur <i>POS-grams</i> dan <i>Sentimen lexical</i> Opini Maskapai	66
Tabel 36 Evaluasi Matrik Fitur <i>Lexical -grams</i> dan <i>POS-grams</i> Opini Maskapai	67
Tabel 37 Hasil Evaluasi Matrik Kombinasi 2 Fitur Opini Maskapai	68
Tabel 38 Susunan dan Total dari Atribut Kombinasi Ekstraksi Fitur	69
Tabel 39 Hasil Evaluasi Matrik dari Kombinasi 3 Fitur Opini Umum	72
Tabel 40 Hasil Evaluasi Matrik dari Kombinasi 3 Fitur Opini Maskapai	74
Tabel 41 Contoh Twitter yang sulit diprediksi dan mudah di prediksi	75

DAFTAR SINGKATAN

SVM	=	Support Vector Machine
Max Ent	=	Maximum Entropy
TP	=	True Positif
TN	=	True Negatif
FP	=	False Positif
FN	=	False Negatif
POS	=	Part Of Speech

DAFTAR LAMPIRAN

- Lampiran 1** Source Code Python
- Lampiran 2** Kartu Konsultasi Tesis II
- Lampiran 3** Form Perbaikan Ujian Tesis II
- Lampiran 4** Jurnal

BAB I. PENDAHULUAN

Pada bab ini berisi latar belakang penelitian yang berjudul : “Analisis Sentimen pada Twitter Menggunakan Metode *Maximum Entropy* dan *Support Vector Machine*”. Kemudian dari latar belakang pada bab ini dirumuskan permasalahan, agar permasalahan pada penelitian ini tidak meluas maka dibuat batasan masalah. Kemudian diberikan tujuan serta manfaat yang akan dibuat dan metodologi yang digunakan dalam penelitian ini.

1.1 Latar Belakang

Media sosial merupakan konten *online* yang dibuat menggunakan teknologi penerbitan yang sangat mudah diakses dan terukur. terpenting dari teknologi adalah terjadinya pergeseran cara mengenali orang lain, berbagi dan membaca berita yang ada, serta mencari informasi. Ada banyak ratusan media sosial yang beroperasi di seluruh dunia saat ini, salah satu dari media sosial tersebut yaitu twitter. Twitter merupakan *mikroblog* cukup populer di Indonesia dengan memungkinkan setiap pengguna untuk mengirim dan membaca pesan yang disebut kicauan (*tweets*) berupa teks 140 karakter untuk ditampilkan pada setiap halaman penggunaanya (Badri, 2011).

Twitter merupakan salah satu *web* terbesar bagi orang-orang mengekspresikan perasaan mereka, berbagi pemikiran, pendapat, memberikan komentar dan melaporkan peristiwa yang dialami secara *real-time* (Soler dkk., 2011). Meningkatnya popularitas twitter dan media sosial lainnya banyak dimanfaatkan oleh perusahaan-perusahaan untuk meningkatkan layanan produknya, dan memberikan layanan *after sales* yang memungkinkan pelanggannya dapat *me-review* kualitas produknya sehingga banyak perusahaan yang mengandalkan *tools HootSuite* untuk menganalisis data dan menyediakan layanan *costumer servic*. Akan tetapi *tools* ini memiliki kekurangan dalam mendeteksi bentuk kalimat emosi positif dan negatif dalam setiap penggunaanya. Beberapa perusahaan penerbangan mencoba untuk mendapatkan keunggulan

kompetitif dengan terus meningkatkan layanan mereka. Ada banyak perusahaan maskapai penerbangan fokus pada kualitas layanan dan pengalaman pengguna layanan mereka serta mengetahui kepuasan dari pelanggan yang menggunakan produknya. Dengan menggunakan media sosial twitter untuk menjadi sumber penting untuk melacak analisis sentimen (Hakh dkk., 2017).

Kehidupan emosi memang merupakan wilayah yang dapat ditangani dengan keterampilan-keterampilan yang lebih tinggi atau lebih rendah, dan membutuhkan keahlian tersendiri (Goleman, 2015). Emosi atau perasaan merupakan suasana psikis atau suasana batin yang dihayati seseorang pada suatu saat. Deteksi kalimat emosi dalam konten media *online* twitter dapat dilakukan dengan menganalisis sentimen para pengguna yang disampaikan melalui *tweets* dalam situs jejaring sosial twitter. Analisis sentimen merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi. sehingga analisis sentimen ini dapat digunakan untuk mendapatkan informasi emosional seseorang yang terdapat dalam pesan para pengguna jejaring sosial twitter terhadap topik yang di bahas penggunanya (Bo Pang dkk., 2008). Banyak pembahasan tentang analisis sentimen pada emosi seperti melakukan penambahan fitur dalam set uji coba analisis sentimen dengan menggunakan pengklasifikasian *Naive Bayes* (Saif dkk., 2012). Mengevaluasi kegunaan sumber daya leksikal yang ada serta fitur yang menangkap informasi tentang bahasa informal dan kreatif digunakan dalam *microblogging* dengan memanfaatkan *hashtag* yang ada di twitter data untuk membangun data pelatihan (Kouloumpi dkk., 2011).

Penambahan fitur dalam analisis sentimen seperti fitur *pragmatic* terkait dengan ekspresi kebencian dan menggunakannya sebagai fitur tambahan untuk deteksi kebencian dengan fitur-fitur lain dalam mendeteksi ucapan kebencian pada pesan teks pendek di Twitter dapat mempermudah dalam mendeteksi ekspresi kebencian (Watanabe, dkk., 2018). Berdasarkan dari penelitian tersebut kami menemukan yaitu fitur *pragmatic* yang dapat menggabungkan dengan menghitung, tanda baca, huruf kapital, hastag, simbol @ maupun emotikon untuk

menegaskan emosi seseorang. Emotikon digunakan oleh pengguna sebagai cara mudah untuk mengekspresikan emosi secara ringkas. Dengan demikian, emotikon adalah cara mudah untuk membedakan *polarity* (pelawanan kata) dari pesan *non-polarity* dan positif dari pesan negatif dalam analisis sentimen. Sedangkan Tanda baca juga dapat memberikan wawasan pada *polarity* pesan. Misalnya, tanda seru digunakan untuk mengekspresikan penekanan kuat yang biasanya merupakan pesan *polarity* (Gokulakrishnan dkk., 2012).

Penelitian membahas Analisis sentimen pada facebook dengan sentimen yang biasa dan untuk mendeteksi perubahan emosional yang signifikan dalam mengekstrak informasi tentang polaritas sentimen pengguna (positif atau negatif) hasil yang diperoleh melalui pendekatan dengan menggunakan metode *Support Vector Machine* mendapatkan nilai akurasi terbaik (Ortigosa dkk., 2014). Dengan analisis sentimen menggunakan twitter pada opini politik menggunakan metode *Support Vector Machine* dan *Maximum Entropy* menyatakan metode *Support Vector Machine* memiliki hasil lebih baik (Ramadhan dkk., 2017). Pada percobaan analisis sentimen, di temukan bahwa metode *Maximum Entropy* memiliki tingkat efektifitas yang lebih baik dari pada metode *Support Vector Machine* dan metode lainnya (Gupta dkk., 2017). Analisis sentimen opini mengenai produk dan layanan pada *e-commerce* menggunakan metode *Support Vector Machine* dan *Maximum Entropy* sebagai metode klasifikasi menunjukkan hasil terbaik pada metode *Maximum Entropy* (Praptiwi, 2018). Maka dalam penelitian ini akan dilakukan suatu percobaan dalam analisis sentimen pada twitter dengan empat ekstraksi fitur dan metode tersebut untuk mendeteksi emosional pengguna sehari-hari (opini umum) dan opini maskapai penerbangan sebagai dataset simulasi perbandingan, yang nantinya akan di uji menggunakan metode algoritma *Maximum Entropy* dan *Support Vector Machine*.

1.2 Perumusan Masalah

Dari latar belakang terdapat beberapa permasalahan yang akan dibahas dalam penelitian ini :

1. Bagaimana cara menganalisis sentimen pada dataset opini umum dan maskapai dengan metode *Maximum Entropy* dan *Support Vector Machine*.
2. Bagaimana cara membandingkan hasil analisis sentimen pada dataset opini umum dan maskapai dengan metode *Maximum Entropy* dan *Support Vector Machine*.
3. Bagaimana hasil akurasi analisis sentimen dengan fitur *Pragmatic*, *Lexical-Grams*, *POS-Grams* dan sentimen *lexical* menggunakan metode *Maximum Entropy* dan *Support Vector Machine*.

1.3 Batasan Masalah

1. Penelitian hanya digunakan terhadap pengguna media sosial twitter dengan dataset opini umum dan dataset opini maskapai penerbangan.
2. *Tweets* yang diteliti merupakan *tweets* berbahasa Inggris.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menganalisis pesan para pengguna jejaring sosial twitter terhadap opini umum dan opini maskapai untuk mengetahui emosi positif dan negatif dengan mengimplementasikan ekstraksi fitur dan membandingkan akurasi, presisi, *recall*, dan *f1-score* pada metode *Maximum Entropy* dan *Support Vector Machine* dalam dataset opini umum dan opini maskapai.

1.5 Manfaat Penelitian

Sedangkan manfaat yang ingin dicapai dari penelitian ini adalah:

1. Mengklasifikasikan analisis sentimen pada twitter dalam jumlah yang besar secara otomatis.
2. Dapat menjadi alat bantu dalam pengambilan keputusan strategi pengembangan layanan informasi yang dimiliki oleh lembaga penelitian.

3. Dari segi praktisi, mendapatkan salah satu metode untuk mengetahui opini dari twitter.

1.6 Sistematika Penulisan

Dalam mempermudah dalam penyusunan penelitian dan memuat uraian secara garis besar isi dari setiap bab, maka dibuatlah sistematika penulisan :

BAB I - PENDAHULUAN

Bab 1 ini berisi tentang latar belakang, perumusan masalah penelitian, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II - TINJAUAN PUSTAKA

Bab II berisi tentang penelitian terkait, landasan teori dan kerangka pemikiran yang berhubungan dengan permasalahan pada tesis ini.

BAB III – METODOLOGI PENELITIAN

Bab ini berisi penjelasan secara bertahap dan terperinci tentang langkah-langkah metodologi penelitian seperti Pengumpulan data, *Pre-procesing* Data, Ekstraksi fitur, Klasifikasi Kalimat serta evaluasi dalam menyelesaikan tesis.

BAB IV – HASIL DAN ANALISA

Bab IV berisi tentang analisis dan pembahasan dari tiap-tiap diagram perencanaan rangkaian data-data hasil pengukuran dalam penelitian.

BAB V – KESIMPULAN

Bab V ini berisi kesimpulan tentang hasil penelitian yang telah diperoleh dalam penelitian, serta memberi jawaban dari tujuan yang ingin dicapai pada BAB I (Pendahuluan) dan disampaikan saran yang perlu dilakukan pada penelitian selanjutnya.

DAFTAR PUSTAKA

- Badri dan Muhammad. (2011): Komunikasi Pemasaran UMKM Di Era Media Sosial, *Corporate and Marketing Communication*, 8, 1 - 13.
- B. Liu. (2010): Sentiment Analysis and Subjectivity, *Handbook of Natural Language Processing*, 8, 1 - 38.
- B. Liu.(2012): Sentiment Analysis and Opinion Mining, *Morgan & Claypool Publisher*, 8, 5 - 168.
- Chris Brogan. (2010): Social Media 101 : Tactics and Tips to Develop Your Business Online. John Wiley & Sons, *Business & Management Special Topics*, 5, 1 - 352.
- Goleman, Daniel. (2015): Emotional Intelligence : Kecerdasan emosional mengapa EI lebih penting daripada IQ, *PT. Gramedia Pustaka Utama*, 2, 1 - 483.
- Gokulakrishnan, Priyanthan, Ragavan (2012): Opinion Mining and Sentiment Analysis on a Twitter Data Stream (ICTer), *IEEE, Sri Lanka*, 6, 182 - 188.
- Gorunescu, F (2011): Data Mining Concept Model and Techniques, *BerlinSpringer*, 14, 1 - 353.
- Gupta, Negi, Vishwakarma, Rawat, Badhani (2017): Study of Twitter Sentiment Analysis using Machine Learning Algorithms on Python, *International Journal of Computer Applications*, 3, 29 - 34.
- Huang, Chia-Hui., Keng-CHieh, Yang., dan Kao, Han-Ying. (2014): Analyzing Big Data With The Hybrid Interval Regression Methods, *The Scientific World Journal*, 13, 1 - 8.
- Jotheeswaran, J Loganathan, R dan MadhuSudhanan, B. (2012): Feature Reduction using Principal Component Analysis for Opinion Mining, *International Journal of Computer Science and Telecommunications* 120, 118-121.
- J. M. Soler, F. Cuartero, dan M. Roblizo. (2012): Twitter as a tool for predicting elections results, *in Proc. IEEE/ACM ASONAM*, 1, 1226 - 1232.
- Kouloumpis, Wilson and Moore. (2011): Twitter Sentiment Analysis: The Good the Bad and the OMG!, *Association for the Advancement of Artificial Intelligence (AAAI)*, 2, 538 - 541.

- Kharde, Sonawane. (2016): Sentiment Analysis of Twitter Data: A Survey of Techniques, *International Journal of Computer Applications*, 7, 5 - 15.
- Kurniawan, D. (2010): Evaluasi sistem temu kembali informasi model ruang vector dengan pendekatan user judgement, *J. Sains MIPA*, 15, 155 - 162.
- MacKay, D. dan Miller, A.L., (2003): Nutritional Support for Wound Healing, *Altern Med Rev*, 11, 359 - 377.
- Ortigosa, Martín, Carro. (2014): Sentiment analysis in Facebook and its application to e-learning, *Elsevier*, 2, 1 - 15.
- Pak, A. dan Paroubek, P. (2010): Twitter as a Corpus for Sentiment Analysis and Opinion Mining, *Valetta, s.n., pp*, 10, 1320 - 1326.
- Parmana, K.A.B, Sudarma dan Ariastina. (2019): Analisis Rating Sentimen pada Video di Media Sosial Youtube Menggunakan STRUCT-SVM, *Majalah Ilmiah Teknologi Elektro*, 113-118.
- Pendit, P. L. (2008): Perpustakaan Digital Dari A Sampai Z. Jakarta, *Cita karyaarsa*, 15, 1 - 308.
- Pang, Bo dan Lee, Lillian. (2008): Opinion Mining and Sentiment Analysis Computer Science Department, *Cornell University*, 2, 1 - 135.
- Praptiwi, D.Y. (2018): Sentiment Analysis Online Review E-Commerce User Using Support Vector Machine And Maximum Entropy, Universitas Islam Indonesia, 1-107.
- Pravina, A.M, Cholissodin dan Andikara. (2019): Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM), *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, 2789-2797.
- Saif, H., He, Y. dan Alani, H. (2011): Semantic Smoothing for Twitter Sentiment Analysis, *Proceeding of the 10th International Semantic Web Conference (ISWC)*, 2, 508 - 524.
- Saraswati, N.W.S. (2011): Text Mining Dengan Metode Naïve Bayes Classifier Dan Support Vector Machines Untuk Sentiment Analysis, *Seminar Nasional Sistem Informasi Indonesia*, 9, 586 - 591.
- Solakidis, Vavliakis, Mitkas. (2014): Multilingual Sentiment Analysis using Emoticons and Keywords, *IEEE*, 7, 102 - 109.

- Sukrisnadi, Dedy. (2010): Pemakaian Ukuran F-Score dalam Kasus-Kasus Salah Saji Laporan Keuangan di Pasar Modal Indonesia, *Universitas Indonesia*, 15, 1- 63.
- S. Wardoyo. (2012): Opinion Summarization Fitur Produk Elektronik pada Amazon.com dengan Metode Maximum Entropy, *Universitas Telkom*, 12, 1 - 50.
- T. Ahmad. (2011): Identifikasi Pelanggaran Pengguna Listrik Rumah Tangga pada PT PLN Dsitribusi Jawa Barat dan Banten UPJ Bandung Utara Menggunakan Metode Naive Bayes dan Metode Maximum Entropy, *Universitas Telkom*, 12, 1 - 21.
- Zarella Dan. (2010): The Social Media Marketing Book. Oreilly Media. USA, 8, 1 – 227.