

**ANALISIS SENTIMEN TERHADAP DATA TWITTER PERUSAHAAN
GRAB INDONESIA MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

SKRIPSI



disusun Oleh :

**Adhe Zhaputri
15.11.8669**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2018**

**ANALISIS SENTIMEN TERHADAP DATA TWITTER PERUSAHAAN
GRAB INDONESIA MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

SKRIPSI

Untuk memenuhi sebagian persyaratan
Mencapai gelar sarjana
Pada Program Studi Informatika



disusun Oleh :
Adhe Zhaputri
15.11.8669

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2018**

PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP DATA TWITTER PERUSAHAAN
GRAB INDONESIA MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

yang dipersiapkan dan disusun oleh

**Adhe Zhaputri
15.11.8669**

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 2 Agustus 2018

Dosen Pembimbing,



Hartatik, S.T., M.CS.
NIK. 190302232

PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP DATA TWITTER PERUSAHAAN
GRAB INDONESIA MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

yang dipersiapkan dan disusun oleh

Adhe Zhaputri

15.11.8669

telah dipertahankan didepan Dewan Penguji
pada tanggal 23 Agustus 2018

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Bayu Setiaji, M.Kom
190302216



Dina Maulina, M.Kom
190302250



Windha Mega Pradnya D, M.Kom
190302185



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
tanggal 24 Agustus 2018

DEKAN FAKULTAS ILMU KOMPUTER



Krisnawati, S.Si, M.T.
NIK/190302038

III

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 24 Agustus 2018



Adhe Zhaputri

NIM. 15.11.8669

MOTTO

You Can cause You Do

-Z

Segala sesuatu yang baik, selalu datang saat terbaiknya.

Persis waktunya.

Tidak datang lebih cepat, pun tidak datang lebih lambat.

Itulah kenapa rasa sabar itu harus disertai keyakinan.

-Tere Liye

Jangan menyerah saat doa – doamu belum terjawab. Jika kamu mampu bersabar, Allah akan memberikan lebih dari apa yang kamu minta

- Anon

PERSEMBAHAN

Alhamdulillah segala puji syukur atas berkat rahmat dan karunia Allah SWT yang telah memberikan kemudahan dan kelancaran bagi penulis dalam menyelesaikan skripsi ini dengan sebaik - baiknya. tidak lepas dari bantuan dan dukungan berbagai pihak. Skripsi ini saya persembahkan untuk Allah SWT yang telah selalu membantu penulis dimanapun dan kapanpun.

Terimakasih untuk Bapak, Ibu, Adik – Adikku yang selalu mendoakan, mendukung dan menyemangati tiada hentinya sehingga skripsi ini bisa terselesaikan dan juga terimakasih kepada semua pihak yang membantu penyusunan skripsi ini yang tidak dapat disebutkan satu persatu.

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul *“Analisis Sentimen terhadap Data Twitter Perusahaan Grab Indonesia Menggunakan Algoritma Support Vector Machine”* dengan sebaik-baiknya. Tidak lupa sholawat serta salam penulis haturkan kepada junjungan umat Nabi Muhammad SAW.

Dengan selesainya skripsi ini, maka penulis mengucapkan terima kasih kepada :

1. Allah SWT yang selalu setia memberikan petunjuk dan membantu disaat – saat getir dan kesulitan dalam menyelesaikan skripsi ini.
2. Bapak M. Suyanto, Prof., Dr., M.M selaku Rektor Universitas Amikom Yogyakarta
3. Ibu Hartatik, ST, M.CS selaku dosen pembimbing yang telah memberikan pengarahan bagi penulis serta telah sabar membimbing dalam pembuatan skripsi ini .
4. Ibu Windha Mega Pradnya D, M.Kom selaku selaku dosen yang membantu dalam pembuatan skripsi ini.
5. Bapak M. Fairul Filza, M.Kom selaku dosen yang membantu dalam penelitian skripsi ini.
6. Bapak I Made Artha Agastya, S.T., M.Kom selaku dosen yang membantu dalam penelitian skripsi ini.

7. Ibu, Bapak, Adik – adik penulis yang selalu setia mendoakan, membimbing, mendukung, sehingga skripsi ini terlaksana dengan lancar dan sesuai target.
8. Rizal Afifi Latukomsina yang selalu sabar dalam membantu, memotivasi dan memberikan dukungan dalam pembuatan skripsi ini.
9. Oktavia yang selalu mendukung, membantu dan telah menjadi sahabat seperjuangan dalam menyelesaikan skripsi ini.
10. Devy Kumalasari dan Sri Suryati yang telah membantu dan mendukung dari awal perkuliahan hingga sekarang ini.
11. Wahyu Setiawan dan Fregy Damara selaku alumni AMIKOM Yogyakarta yang membantu dalam penelitian skripsi ini.
12. Seluruh Pengurus dan Alumni - Alumni UKM FOSSIL AMIKOM Yogyakarta.
13. Para Dosen dan Staff Universitas AMIKOM Yogyakarta yang telah membantu memberikan ilmu pengetahuan, pengalaman selama masa.
14. Seluruh Teman Kelas 15 S1 Informatika 03 dan Semua pihak yang telah membantu dan tidak dapat disebutkan satu persatu.

Pembuatan skripsi ini masih banyak sekali kekurangan. Oleh karena itu, kepada semua pihak agar dapat menyampaikan kritik,saran untuk menambah kesempurnaan skripsi ini. Semoga skripsi ini dapat bermanfaat bagi pembaca pada umumnya.

Tanggal, Agustus 2018

Adhe Zhaputri

15.11.8669

DAFTAR ISI

PERNYATAAN.....	ii
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiv
INTISARI.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penulisan.....	4
1.5 Manfaat Penulisan.....	4
1.6 Metode Penelitian.....	4
1.6.1. Metode Pengumpulan Data.....	5
1.6.2. Metode Analisis.....	5
1.6.3. Metode Perancangan.....	7
1.6.4. Metode Implementasi.....	7
1.6.5. Metode Testing.....	7
1.7 Sistematika Penulisan.....	7

BAB II LANDASAN TEORI	9
2.1 Tinjauan pustaka.....	9
2.2 Pengertian Data Mining.....	11
2.3 Pengertian Sentiment Analysis.....	12
2.4 Preprocessing.....	13
2.5 Term Frequency-Inverse Document Frequency(TF-IDF).....	13
2.6 Support Vector Machine	14
2.7 Consine Similarity	18
2.8 K-Fold Cross Validation	19
2.9 Confusion Matrix	19
2.10 Flowchart.....	21
2.10.1 System Flowchart (Diagram Alir Sistem).....	22
2.10.2 Program Flowchart (Diagram Alir Program).....	23
2.11 Bahasa Pemograman Python	24
2.12 Langkah – langkah pengembangan aplikasi.....	25
BAB III ANALISIS DAN PERANCANGAN	31
3.1. Analisis Masalah	31
3.2. Solusi yang Dipilih.....	32
3.3. Analisis Kebutuhan	32
3.3.1. Analisis Kebutuhan Fungsional	33
3.3.2. Analisis Kebutuhan Non-Fungsional	33
3.4. Analisa Sentimen dengan Klasifikasi Algoritma SVM.....	34
3.4.1. Cleaning Data.....	34

3.4.2.	Case Folding	36
3.4.3.	Tokenization.....	37
3.4.4.	Steaming	38
3.4.5.	TF-IDF	39
3.4.6.	Consine Similarity.....	43
3.4.7.	Support Vector Machine	48
3.4.8.	K-Fold Cross Validation	51
3.5.	Perancangan Aplikasi	53
3.5.1	Diagram Alir Preprocessing	53
3.5.2	Proses Klasifikasi Support Vector Machine (SVM)	54
3.6.	Perancangan File Crawling Data Twitter	56
3.7.	Perancangan File Training	56
BAB IV	IMPLEMENTASI DAN PEMBAHASAN	58
4.1	Deskripsi Implementasi	58
4.2	Implementasi Pengumpulan Data Tweet	58
4.3	Implementasi Pelebelan Data Training	60
4.4	Implementasi Perancangan Fungsi dan Sistem	61
4.4.1	Mengakses Dataset Training	61
4.5	Pembuatan Interface	64
4.6	Evaluasi Program	67
BAB V	KESIMPULAN	73
5.1.	Kesimpulan.....	73
5.2.	Saran	73

DAFTAR GAMBAR

Gambar 2. 1 Sample Penyebaran Pemasalahan Dua Dimensi	16
Gambar 2. 2 Hyperplane pada support vector machine	16
Gambar 2. 3 System Flowchart (Diagram Alir System)	22
Gambar 2. 4 Program Flowchart (Diagram Alir Program)	23
Gambar 2. 5 Langkah – Langkah Analisis Sentimen.....	26
Gambar 2. 6 Langkah – Langkah Pre-Processing.....	26
Gambar 2. 7 Langkah – Langkah Training Data.....	28
Gambar 2. 8 Langkah – Langkah Testing Data	30
Gambar 3. 1 Diagram Alir Pre-processing.....	53
Gambar 3. 2 Konfigurasi File Crawling Data	56
Gambar 3. 3 Isi pada File Training	57
Gambar 4. 2 Script Crawling Data Twitter	59
Gambar 4. 3 Hasil Crawling Data.....	60
Gambar 4. 4 Dataset Training yang akan digunakan.....	61
Gambar 4. 5 Script Akses Data Training	61
Gambar 4. 6 Script Preprocessing Data Training	62
Gambar 4. 7 Script Pembobotan dan Klasifikasi SVM.....	62
Gambar 4. 8 Script Pembersihan Data Testing	63
Gambar 4. 9 Script Stemming Data Testing	63
Gambar 4. 10 Script Prediksi Tweet	64
Gambar 4. 11 Tampilan awal dari User Interface Prediksi Data Twitter.....	65
Gambar 4. 12 User Interface Hasil Prediksi.....	66

Gambar 4. 13 Userinterface Hasil Prediksi dalam Bentuk Grafik 66

Gambar 4. 14 User interface Hasil Prediksi dalam Bentuk Pie Chart 67

Gambar 4. 15 Grafik Akurasi dari hasil Proses 10-Fold Cross Validation 72



DAFTAR TABEL

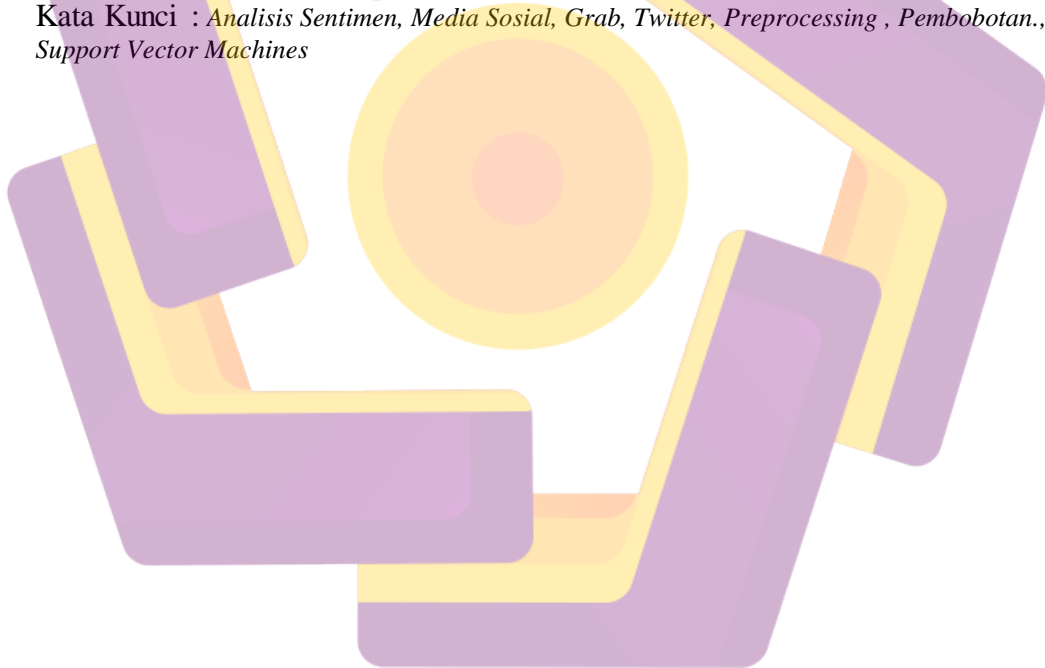
Tabel 2. 1 Perbandingan penelitian terkait dengan sistem yang dibangun	10
Tabel 2. 2 <i>Confusion Matrix</i>	20
Tabel 3. 1 Data Tweet Sebelum Dilakukan Cleaning Data	34
Tabel 3. 2 Data Setelah Dilakukan Cleaning Data.....	35
Tabel 3. 3 Data Tweet Setelah Case Folding	36
Tabel 3. 4 Hasil Tokenization Dari Case Folding.....	37
Tabel 3. 5 Hasil Steaming Dari Tokenizing.....	38
Tabel 3. 6 Feature List	39
Tabel 3. 7 TF IDF Positif	41
Tabel 3. 8 TF IDF Negatif.....	42
Tabel 3. 9 Hasil Consine Similarity Positif.....	45
Tabel 3. 10 Hasil Consine Similarity Negatif	46
Tabel 3. 11 Hasil Consine Similarity Negatif	47
Tabel 3. 12 Perhitungan Persamaan Positif Terhadap Training Data Positif.....	49
Tabel 3. 13 Persamaan w_1 , w_2 dan b Pada Setiap Label.....	50
Tabel 3. 14 Perhitungan Persamaan Negatif Terhadap Training Data Positif	50
Tabel 3. 15 Predik Pada Data Training	51
Tabel 3. 16 Fold Pertama	51
Tabel 3. 17 Fold Kedua.....	52
Tabel 3. 18 Fold Ketiga.....	52
Tabel 4. 1 Confusion Matrix SVM	68
Tabel 4. 2 Table Hasil Cross Validation SVM	69

INTISARI

Informasi saat ini sangat mudah ditemukan salah satunya di Media Sosial. Media sosial merupakan tempat bertukar informasi / opini dari berbagai sumber. Salah satu media sosial yang diminati adalah Twitter. Menurut kominfo, berdasarkan data PT Bakrie Telecom, Twitter memiliki 19,5 juta pengguna di Indonesia dari total 500 juta pengguna global. Dengan demikian dapat disimpulkan bahwa twitter memiliki dampak baik untuk konsumen maupun produsen dalam mengambil keputusan, saran, opini atau ulasan – ulasan.

Oleh karena itu opini – opini tersebut dapat diimplementasikan untuk analisis sentimen terhadap pelayanan perusahaan Grab yang bertujuan guna meningkatkan strategi dalam pemasaran, peningkatan produk/ pelayanan untuk menjadi lebih baik. Dalam analisis sentimen data akan diolah dengan pre-processing, pembobotan dan diklasifikasi dengan Support Vector Machine.

Kata Kunci : Analisis Sentimen, Media Sosial, Grab, Twitter, Preprocessing , Pembobotan., Support Vector Machines



ABSTRACT

Information is very easy to find one of them on Social Media. Social media is a place to exchange information / opinions from various sources. One of the most popular social media is Twitter. According to Kominfo, based on data from PT Bakrie Telecom, Twitter has 19.5 million users in Indonesia out of a total of 500 million global users. Thus it can be concluded that twitter has a good impact for consumers and producers in making decisions, suggestions, opinions or reviews - reviews.

Therefore, these opinions can be implemented for sentiment analysis on the services of Grab's company which aims to improve strategies in marketing, product / service improvement to be better. In the data sentiment analysis will be obtained by pre-processing, weighting and classified with the Support Vector Machine.

Keywords: *Sentiment Analysis, Social Media, Grab, Twitter, Preprocessing, Weighting, Support Vector Machines*

