

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN PENGGUNA
JASA LAYANAN ANTAR BARANG JNE PADA TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI



disusun oleh

Andia Enggar Mayasari

16.11.0161

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN PENGGUNA
JASA LAYANAN ANTAR BARANG JNE PADA TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Andia Enggar Mayasari

16.11.0161

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**

PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN
PENGGUNA JASA LAYANAN ANTAR BARANG JNE PADA TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE
(SVM)**


yang dipersiapkan dan disusun oleh

Andia Enggar Mayasari

16.11.0161

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 12 November 2019

Dosen Pembimbing,



Anggit Dwi Hartanto, M.Kom.
NIK. 190302163

PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN
PENGGUNA JASA LAYANAN ANTAR BARANG JNE PADA TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE
(SVM)**

yang dipersiapkan dan disusun oleh

Andia Enggar Mayasari

16.11.0161

telah dipertahankan di depan Dewan Penguji
pada tanggal 12 November 2019

Susunan Dewan Penguji

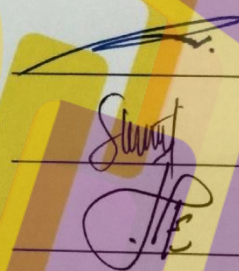
Nama Penguji

Tanda Tangan

Bambang Sudaryatno, Drs., M.M.
NIK. 190302029

Sumarni Adi, S.Kom, M.Cs.
NIK. 190302256

Anggit Dwi Hartanto, M.Kom.
NIK. 190302163



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 November 2019



DEKAN FAKULTAS ILMU KOMPUTER

Prisnawati, S.Si, M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan tidak 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 tanggung jawab saya pribadi.

Yogyakarta, 4 November 2019



Andia Enggar Mayasari
NIM. 16.11.0161

MOTTO

“The best revenge is massive success” –Frank Sinatra

“Satu contoh lebih baik daripada seribu nasehat”



PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangnya sampai sejauh ini.
2. Kedua orang tua saya dan keluarga, yang selalu mendoakan, selalu menyemangati dan memberikan uang jajan kepada saya.
3. Bapak Anggit Dwi Hartanto, M.Kom yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
4. Dosen – dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama kuliah.
5. Teman skripsi saya Aiman Muklishah, Linda Kumala Sari, Francisca Ayu Lestari, Cidalia yang selalu memberikan solusi, tempat bertanya ketika dimasa sulit dan tempat curhat keluh kesah selama dalam pembuatan skripsi.
6. Teman – teman kelas IF 03 2016 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini. Semoga kita selalu bahagia dan menjadi pribadi yang lebih baik lagi.
7. Teman – teman FOSSIL yang selalu mendukung dan memberikan semangat sampai saat ini. Semoga kita selalu bahagia dan menjadi pribadi yang lebih baik lagi.
8. Orang – orang baik yang senantiasa mendoakan, memberi semangat dan bantuan kepada saya.

KATA PENGANTAR

Puji dan syukur panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Analisis Sentimen Terhadap Tingkat Kepuasan Pengguna Jasa Antar Brang JNE Pada Twitter Menggunakan Algoritma Support Vector Machine (SVM).

Skripsi ini saya buat guna menyelesaikan studi jenjang Starta Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata satu dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini, maka pada kesempatan ini saya mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Anggit Dwi Hartanto, M.Kom selaku dosen pembimbing yang selalu bijaksana dalam memberikan bimbingan, nasehat serta waktunya selama pembuatan skripsi ini.
4. Dosen penguji Ibu Sumarni Adi, S.Kom, M.Cs, Bapak Bambang Sudaryatno, Drs., M.M dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
5. Kedua orang tua saya yang telah mendoakan, mendukung dan memberikan semangat kepada saya.
6. Orang – orang baik yang senantiasa mendoakan, memberi semangat dan bantuan kepada saya.
7. Teman – teman FOSSIL yang selalu memberi semangat dan bantuan kepada saya.

8. Teman – teman IF 03 2016 yang telah menemani selama proses perkuliahan.
9. Teman – teman kontrakan yang selalu memberikan semangat dan masukan dalam pembuatan skripsi.
10. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Semoga Allah SWT memberikan balasan yang lebih kepada semua yang telah ikut membantu saya hingga menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik membangun diterima dengan senang hati dan terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 4 November 2019

Andia Enggar Mayasari

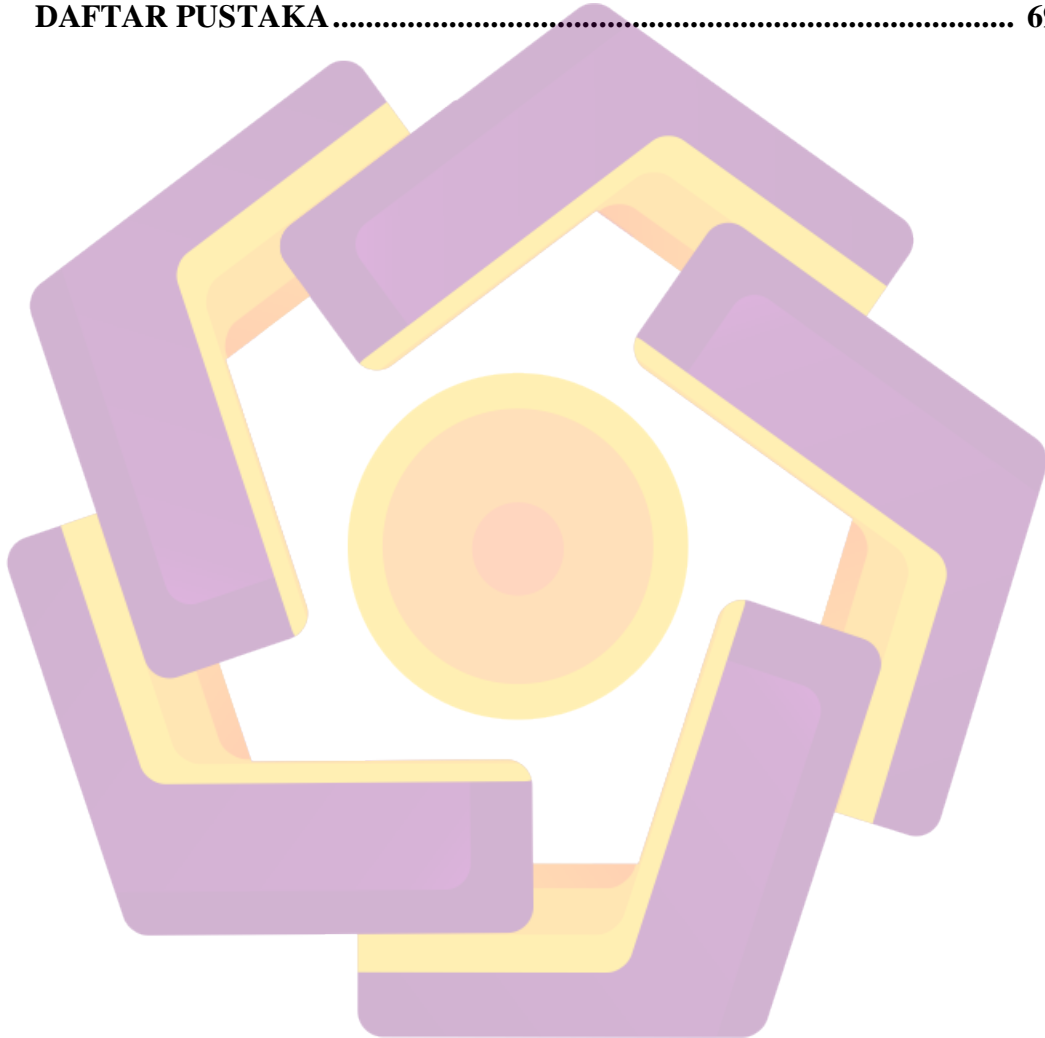
DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiv
INTISARI	xv
ABTRACT	xvi
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH	2
1.3 BATASAN MASALAH	2
1.4 MAKSUD DAN TUJUAN PENELITIAN	3
1.5 MANFAAT PENULISAN	3
1.6 METODE PENELITIAN	3
1.7 SISTEMATIKA PENULISAN	6
BAB II LANDASAN TEORI	8
2.1 KAJIAN PUSTAKA	8

2.2 DASAR TEORI	10
2.2.1 DATA MINING	10
2.2.2 SENTIMEN ANALISIS	10
2.2.3 PREPROCESSING.....	10
2.2.4 TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY	12
2.2.5 SUPPORT VECTOR MACHINE	13
2.2.6 COSINE SIMILARITY.....	16
2.2.7 K-FOLD CROSS VALIDATION	16
2.2.8 CONFUSION MATRIX	17
2.2.9 FLOWCHART	18
2.2.10 BAHASA PEMROGRAMAN PYTHON	19
2.2.11 LANGKAH-LANGKAH PENGEMBANGAN APLIKASI	22
BAB III ANALISIS DAN PERANCANGAN	25
3.1 ANALISIS MASALAH	25
3.2 SOLUSI YANG DIPILIH	25
3.3 ANALISIS KEBUTUHAN	26
3.3.1 ANALISIS KEBUTUHAN FUNGSIONAL.....	26
3.3.2 ANALISIS KEBUTUHAN NON-FUNGSIONAL	26
3.4 ANALISIS SENTIMEN DENGAN ALGORITMA SVM	27
3.4.1 CLEANING DATA.....	27
3.4.2 CASE FOLDING.....	29
3.4.3 TOKENIZATION.....	29

3.4.4 STEMMING	30
3.4.5 TF-IDF	32
3.4.6 CONSINE SIMILARITY.....	38
3.4.7 SUPPORT VECTOR MACHINE	45
3.5 PERANCANGAN APLIKASI	49
3.5.1 DIANGRAM ALIR PREPOCESSING	49
3.5.2 PROSES KLASIFIKASI SUPPORT VECTOR MACHINE	50
3.6 PERANCANGAN FILE CRAWLING DATA TWITTER	51
3.7 PERANCANGAN FILE TRAINING	52
BAB IV IMPLEMENTASI DAN PEMBAHASAN.....	53
4.1 DESKRIPSI IMPLEMENTASI.....	53
4.2 IMPLEMENTASI PENGUMPULAN DATA TWEET	53
4.3 IMPLEMENTASI PELABELAN DATA TRAINING	55
4.4 IMPLEMENTASI PERANCANGAN FUNGSI DAN SISTEM	56
4.4.1 PEMBUATAN FUNGSI PRE-PROCESSING DAN STEMMING DATA TRAINING	56
4.4.2 PEMBUATAN FUNGSI KLASIFIKASI SVM DATA TRAINING	57
4.4.3 PEMBUATAN FUNGSI PEMBERSIHAN DATA TRAINING	57
4.4.4 PEMBUATAN FUNGSI STEMMING DATA TESTING	58
4.4.5 PEMBUATAN FUNGSI PREDIKSI SENTIMEN.....	58
4.5 PEMBUATAN INTERFACE	59

4.6 EVALUASI PROGRAM	61
BAB V KESIMPULAN.....	68
5.1 KESIMPULAN	68
5.2 SARAN.....	68
DAFTAR PUSTAKA.....	69



DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian	9
Tabel 2.2 Confusion Matrix	17
Tabel 2.3 Data Tweet Sebelum Dilakukan Cleaning Data	27
Tabel 2.4 Data Tweet yang Sudah Dilakukan Cleaning	28
Tabel 2.5 Data Tweet Setelah Case Folding	29
Tabel 2.6 Hasil Tokenization Dari Case Folding	30
Tabel 2.7 Hasil Stemming Dari Tokenizing	31
Tabel 2.8 Feature List	31
Tabel 2.9 TF IDF Positif	33
Tabel 3.0 TF IDF Negatif	35
Tabel 3.1 Consine Similarity Positif	40
Tabel 3.2 Consine Similarity Negatif	41
Tabel 3.3 Hasil Consine Similarity	45
Tabel 3.4 Perhitungan Persamaan Positif Terhadap Training Data Negatif	46
Tabel 3.5 Persamaan w_1 , w_2 dan b Pada Setiap Label	46
Tabel 3.6 Perhitungan Persamaan Negatif Terhadap Training Data Positif	47
Tabel 3.7 Predik Pada Data Training	47
Tabel 3.8 Fold Pertama	48
Tabel 3.9 Fold Kedua	48
Tabel 4.0 Fold Ketiga	48
Tabel 4.1 Confusion Matrix SVM	62
Tabel 4.2 Tabel Hasil Cross Validation SVM	63

DAFTAR GAMBAR

Gambar 2.1 Sample Penyebaran Permasalahan Dua Dimensi	14
Gambar 2.2 Hyperlane Pada Support Vector Machine.....	14
Gambar 2.4 Langkah – langkah Analisis Sistem.....	20
Gambar 2.5 Langkah – langkah Training Data	22
Gambar 2.6 Langkah – langkah Testing Data	24
Gambar 2.7 Diagram Alir Pre-Processing	49
Gambar 2.8 Diagram Alur Klasifikasi SVM.....	50
Gambar 2.9 Konfigurasi Crawling Data	51
Gambar 3.0 File Training	52
Gambar 3.1 Script Crawling Data Twitter	53
Gambar 3.2 Script Hasil Crawling Data Twitter	55
Gambar 3.3 Dataset yang Akan Digunakan	56
Gambar 3.4 Script Akses Masuk Data Training.....	56
Gambar 3.5 Script Preprocessing Data Training	57
Gambar 3.6 Script Pembobotan dan Klasifikasi SVM	57
Gambar 3.7 Script Pembersihan Data Testing	58
Gambar 3.8 Script Stemming Data Testing	58
Gambar 3.9 Scipt Prediksi Tweet	59
Gambar 4.0 Halaman Utama Sistem Prediksi	60
Gambar 4.1 User Interface Hasil Prediksi.....	60
Gambar 4.2 User Interface Hasil Prediksi Dalam Betuk Pie Chart	61
Gambar 4.3 Hasil Confusion Matrix Pada Sistem.....	63

INTISARI

Jumlah data mengalami pertumbuhan yang sangat cepat dalam era sekarang ini. Data dapat berupa *text*, gambar, suara maupun video. Media social menjadi salah satu faktor pertumbuhan data, setiap orang berekspektasi beropini dan mengeluh di media social dari opinion tersebut dapat dilakukan analisis data. Dalam penelitian ini analisis sentiment menggunakan algoritma support vector machine. Langkah pertama adalah dilakukan *crawling* data menggunakan twitter API dengan keyword. Setelah mengumpulkan data, dilakukan proses preprocessing, setelah proses preprocessing dilakukan pengambilan fitur pada setiap tweet, fitur yang didapatkan kemudian dikumpulkan menjadi sebuah list fitur.

List fitur kemudian ditransformasikan menjadi feature vector dengan bentuk binary kemudian ditransformasikan menggunakan metode Tf-idf. Dataset terdiri dari 2 data yaitu training dan testing. Di dalam data training diberikan label secara manual. Untuk pengujian performa algoritma digunakan metode K-Fold Cross Validation. Hasil pengujian adalah akurasi yang diperoleh mencapai rata-rata 80% dengan komposisi data training dan data testing. Dari hasil tersebut metode Support Vector Machine dapat digunakan untuk klasifikasi sentimen terhadap data twitter JNE.

Di dalam data testing dilakukan crawling data secara realtime dan data yang dihasil dilakukan pelabelan dan proses *Preprocessing* melalui sistem sehingga output akhir yang ditampilkan berupa sentiment positif dan negative dari hasil crawling data sebelumnya.

Kata kunci: Sentimen analisis, support vector machine, invers matrix

ABSTRACT

The amount of data is experiencing rapid growth in this era. Data can be in the form of text, images, sound or video. Social media has become one of the factors of data growth, everyone in the opinion opinion and complaining in social media from the opinion can be done data analysis. In this study sentiment analysis uses the support vector machine algorithm. The first step is crawling data using the Twitter API with keywords. After collecting data, the preprocessing process is carried out, after the preprocessing process the feature is retrieved for each tweet, the features obtained are then collected into a feature list.

The feature list is then transformed into a feature vector in binary form then transformed using the Tf-idf method. The dataset consists of 2 data, namely training and testing. In the training data labeled manually. To test the performance of the algorithm used the K-Fold Cross Validation method. The test results are obtained an average accuracy of 80% with the composition of training data and testing data. From these results the Support Vector Machine method can be used for sentiment classification of JNE twitter data.

In data testing, real-time data crawling is done and the resulting data is labeled and the Preprocessing process is done through the system so that the final output displayed is positive and negative sentiment from the results of previous data crawling.

Key words: *Sentiment analysis, support vector machine, invers matrix*