

**ANALISIS SENTIMEN TERHADAP BRAND XIAOMI PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

SKRIPSI



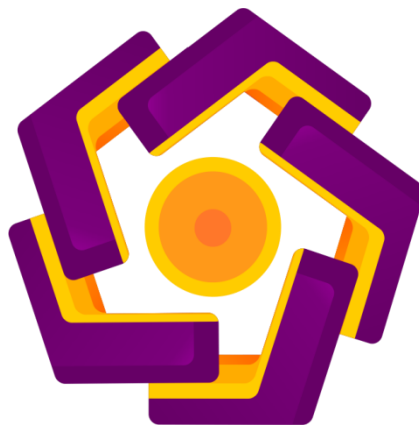
**disusun oleh
Wahyu Permadi
19.21.1364**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMEN TERHADAP BRAND XIAOMI PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Wahyu Permadi

19.21.1364

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP BRAND XIAOMI PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

yang dipersiapkan dan disusun oleh

Wahyu Permadi

19.21.1364

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 Juli 2021

Dosen Pembimbing,

Anggit Dwi Hartanto, M.Kom

NIK. 190302163

PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP BRAND XIAOMI PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE**

yang dipersiapkan dan disusun oleh

Wahyu Permadi

19.21.1364

telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Juli 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

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

Erni Seniwati, S.Kom., M.Cs.
NIK. 190302231

Anggit Dwi Hartanto, M.Kom
NIK. 190302163

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 2 Agustus 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom
NIK. 190302096

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, 2 Agustus 2021



Wahyu Permadi

NIM. 19.21.1364

MOTTO

“Sesungguhnya sesudah kesulitan itu ada kemudahan, sesungguhnya sesudah kesulitan itu ada kemudahan”.

(Q.S Asy Syarh ayat 5-6)

“Jika saya mencoba yang terbaik dan saya gagal, yah, setidaknya saya sudah mencoba yang terbaik”

(Steve Jobs)

“Ketika dunia ternyata jahat padamu, maka kau harus menghadapinya. Karena tidak seorangpun yang akan menyelamatkanmu jika kau tak berusaha.”

(Roronoa Zoro)

“Keajaiban tidak datang dari orang lain, tapi keajaiban ada karena kita sendiri yang membuatnya”

(Emporio Ivankov)

“Kebahagiaan adalah ketika apa yang Anda pikirkan, apa yang Anda katakan, dan apa yang Anda lakukan selaras.”

(Mahatma Gandhi)

PERSEMBAHAN

Segala puji dan puji penulis panjatkan kepada Allah SWT yang telah memberi nafas hingga akhir penyusunan skripsi ini, solawat serta salam saya panjatkan atas limpahan rahmat dan hidayahnya hingga sampai detik ini masih diberi umur panjang.

Skripsi ini saya persembahkan untuk :

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangnya sampai sejauh ini.
2. Kedua Orang tua, Ayah dan ibu yang selalu memberikan Doa, dorongan dan motifasi, pengorbanan serta perjuangan hingga sampai detik ini.
3. Dosen pembimbing Bapak Anggit Dwi Hartanto, M.Kom yang telah membimbing dan mengarahkan dalam pembuatan tugas akhir ini.
4. Bapak dan Ibu Dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama kuliah.
5. Teman – teman di perkuliahan : Ikhsan Subagyo, Lukman Dwi Yulianto, Arian Wahyu Dewantara, Ruri, Mirchan Prasetyo, Dimar, Nopus, Syifa Fauziah, Fitri yang sudah menghibur, memberi semangat dan saling membantu dalam perkuliahan dan skripsi.
6. Teman – teman dalam referensi Skripsi : Andia Enggar Mayasari dan Mas Sidiq, yang telah memberikan ilmunya, semangat dan suportnya untuk tercapainya skripsi ini.
7. Ikhsan Subagyo terimakasih telah memberikan semangat, teman disaat terpuruk dan sudah menjadi teman yang baik dalam perkuliahan dan skripsi ini.
8. Orang – orang baik yang bersedia untuk memberikan bantuan kepada saya dalam mengerjakan skripsi.

KATA PENGANTAR

Segala puji dan syukur kepada Allah SWT. Karena berkat rahmat, karuniaNya, hidayah dan kekuatan sehingga penulis dapat menyelesaikan Skripsi ini dengan judul “Analisis Sentimen Terhadap Brand Xiaomi Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine”.

Skripsi ini diselesaikan dengan maksud dan tujuan untuk memenuhi syarat menyelesaikan studi jenjang Starta Satu (S1) dan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Saya menyadari bahwa masih ada kekurangan dalam penulisan skripsi ini, oleh sebab itu saya menerima segala kritik dan saran yang membangun dari pembaca. terselesainya Skripsi ini tidak terlepas dari dukungan dari berbagai pihak yang telah memberikan petunjuk, bimbingan dan dorongan baik moril maupun materi. Untuk ini penulis mengucapkan terimakasih kepada :

1. Prof. Dr. M. Suyanto, M. M, selaku Rektor UNIVERSITAS AMIKOM YOGYAKARTA.
2. Bapak Hanif Al Fatta, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega PD, M.Kom selaku Ketua Program Studi Informatika UNIVERSITAS AMIKOM YOGYAKARTA.
4. Anggit Dwi Hartanto, M.Kom, selaku Dosen Pembimbing.
5. Dosen penguji Ibu Sumarni Adi, S.Kom, M.Cs, Ibu Erni Seniwati, S.Kom., M.Cs. dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
6. Kedua Orang Tua, yang selalu mendoakan, serta memberi semangat dan dukungan.
7. Sahabat – sahabat seangkatan dan seperjuangan yang memberi inspirasi.

DAFTAR ISI

JUDUL.....	i
PERSETUJUAN.....	ii
PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I Pendahuluan.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Maksud dan Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Metode Penelitian.....	5
1.6.1 Metode Pengumpulan Data.....	5
1.6.2 Metode Analisis.....	5
1.6.3 Metode Perancangan.....	6
1.6.4 Metode Pengujian.....	6
1.7 Sistematika Penulisan.....	6
BAB II Landasan Teori.....	8

2.1	Kajian Pustaka	8
2.2	Dasar Teori.....	12
2.2.1	Data Mining	12
2.2.2	Text Mining	12
2.2.3	Analisis Sentimen	12
2.2.4	Preprocessing	13
2.2.5	Term Frequency-Inverse Document Frequency (TF-IDF)	15
2.2.6	Support Vector Machine (SVM).....	17
2.2.7	Cosine Similarity	22
2.2.8	K-Fold Cross Faldation.....	22
2.2.9	Confusion Matrix	23
2.2.10	Flowchart	25
2.2.11	Bahasa Pemrograman Python	26
2.2.12	Langkah-langkah Pengembangan Aplikasi.....	27
BAB III METODE PENELITIAN		33
3.1	Alur Penelitian	33
3.1.1	Alur Preprocessing	33
3.1.2	Alur Klasifikasi Support Vector Machine (SVM).....	35
3.2	Analisis Kebutuhan.....	36
3.2.1	Analisis Kebutuhan Fungsional	36
3.2.2	Analisis Kebutuhan Non Fungsional	37
3.3	Analisis Sentimen Menggunakan Algoritma Support Vector Machine....	37
3.3.1	Cleaning Data.....	37
3.3.2	Case Folding	39
3.3.3	Tokennizing	40
3.3.4	Stemming	41
3.3.5	TF-IDF	42
3.3.6	Cosine Similarity	46
3.3.7	Support Vector Machine	53
3.3.8	K-Fold Cross Validation	55
3.4	Perancangan Crawling Data Twitter	56

3.5 Perancangan File Training	57
BAB IV HASIL dan Pembahasan.....	58
4.1 Deskripsi Implementasi	58
4.2 Implementasi Pengumpulan Data Tweet	58
4.3 Implementasi Pelabelan Data Training	60
4.4 Implementasi Perancangan Fungsi Sistem.....	61
4.4.1 Pembuatan Fungsi Pre-Processing dan Stemming Data Training	61
4.4.2 Pembuatan Fungsi Klasifikasi SVM Data Training.....	62
4.4.3 Pembuatan Fungsi Pembersihan Data Testing.....	62
4.4.4 Pembuatan Fungsi Stemming Data Testing	63
4.4.5 Pembuatan Fungsi Prediksi Sentimen.....	63
4.5 Pembuatan Interface.....	64
4.5.1 Halaman Utama	64
4.5.2 Halaman Hasil.....	65
4.6 Evaluasi Program	66
BAB V Penutup	71
5.1 Kesimpulan	71
5.2 Saran	71
Daftar Pustaka.....	72

DAFTAR TABEL

Tabel 2.1 Perbandingan penelitian.....	10
Tabel 2.2 Confusion Matrix	24
Tabel 3.1 Data Tweet Sebelum Dilakukan Cleaning Data	38
Tabel 3.2 Data yang sudah dilakukan cleaning Data	39
Tabel 3.3 Data Tweet Setelah Dilakukan Case Folding	39
Tabel 3.4 Hasil Tokenization Dari Case Folding.....	40
Tabel 3.5 Hasil Stemming Dari Tokenizing	41
Tabel 3.6 Feature List	42
Tabel 3.7 TF IDF Positif	43
Tabel 3.8 TF IDF Negatif	45
Tabel 3.9 Cosine Similarity Positif	48
Tabel 3.10 Cosine Similarity Negatif	49
Tabel 3.11 Hasil Cosine Similarity	52
Tabel 3.12 Persamaan Positif Terhadap Training Data Negatif	54
Tabel 3.13 Persamaan w_1 , w_2 dan b Pada Setiap Label.....	54
Tabel 3.14 Persamaan Negatif Terhadap Training Data Positif	55
Tabel 3.15 Predik Pada Data Training	55
Tabel 3.16 Fold Pertama	55
Tabel 3.17 Fold Kedua.....	56
Tabel 3.18 Fold Ketiga	56
Tabel 4.1 Confusion Matrix SVM	66
Tabel 4.2 Tabel Hasil Cross Validation SVM	68

DAFTAR GAMBAR

Gambar 2.1 Sample Penyebaran Pemasalahan Dua Dimensi	19
Gambar 2.2 Hyperlane pada Support Vector Machine	20
Gambar 2.3 Langkah – langkah Analisis Sistem	28
Gambar 2.4 Langkah – langkah Training Data.....	30
Gambar 2.5 Langkah – langkah Testing Data.....	32
Gambar 3.1 Diagram Alur Pre-processing.....	33
Gambar 3.2 Diagram Alur Klasifikasi SVM	35
Gambar 3.3 Konfigurasi Crawling Data	56
Gambar 3.4 File Training.....	57
Gambar 4.1 Script Crawling Data Twitter	59
Gambar 4.2 Hasil Crawling Data Twitter	59
Gambar 4.3 Dataset yang akan digunakan.....	60
Gambar 4.4 Script akses masuk data training	61
Gambar 4.5 Script Preprocessing Data Training	61
Gambar 4.6 Script Pembobotan dan Klasifikasi SVM	62
Gambar 4.7 Script Pembersihan Data Testing	62
Gambar 4.8 Script Stemming Data Testing	63
Gambar 4.9 Script Prediksi Tweet	64
Gambar 4.10 Halaman Utama Sistem.....	65
Gambar 4.11 Halaman Hasil.....	66
Gambar 4.12 Hasil Confusion Matrix pada Sistem	67

INTISARI

Xiaomi merupakan perusahaan elektronik smartphone berasal dari Tiongkok yang didirikan pada tahun 2010 dan berkantor pusat di Beijing. Xiaomi mampu bersaing dalam perkembangan bisnis dan sudah terkenal dengan pengembangan mulai dari kualitas, fitur dan harga. Opini dari masyarakat sangat berpengaruh, diperlukan metode untuk mengatasi masalah tersebut. Analisis Sentimen adalah penelitian yang mengolah tentang opini dari masyarakat dalam bentuk data teks.

Data yang digunakan untuk diproses didapat dari platform media sosial Twitter. Twitter merupakan media sosial yang biasa digunakan untuk menyampaikan opini dari penggunanya. Data yang ada pada Twitter berupa data teks, gambar dan video. Pada penelitian ini data yang digunakan adalah data teks. Data tweet akan diproses menggunakan algoritma Support Vector Machine. Support Vector Machine merupakan salah satu metode penggalian teks yang dapat mengkategorikan opini. Penelitian yang dilakukan ini bertujuan untuk mengklasifikasi sehingga memperoleh opini positif dan negatif menggunakan algoritma Support Vector Machine.

Dataset yang digunakan terdiri dari 2 data yang merupakan training dan testing. Pada data training diproses labeling secara manual. Dalam melakukan pengujian performa algoritma menggunakan metode K-Fold Cross Validation. Untuk proses data testing melakukan crawling data twitter secara realtime dan data yang dihasilkan kemudian melakukan proses pelabelan dan proses Preprocessing menggunakan system. Maka dari itu output akhir yang ditampilkan berbentuk sentiment positif dan negative dari hasil crawling data twitter yang sebelumnya. Hasil evaluasi proses pengujian kemudian mendapatkan nilai akurasi rata – rata sebesar 79%.

Kata Kunci: Python, Text Mining, Analisis Sentimen, Support Vector Machine

ABSTRACT

Xiaomi is a smartphone electronics company from China that was founded in 2010 and is headquartered in Beijing. Xiaomi is able to compete in business development and is well known for its development from quality, features and price. Public opinion is very influential, a method is needed to overcome this problem. Sentiment analysis is research that processes public opinion in the form of text data.

The data used for processing is obtained from the social media platform Twitter. Twitter is a social media that is commonly used to convey opinions from its users. The data on Twitter is in the form of text, image and video data. In this study the data used is text data. Tweet data will be processed using the Support Vector Machine algorithm. Support Vector Machine is one method of extracting text that can categorize opinions. This research aims to classify so as to obtain positive and negative opinions using the Support Vector Machine algorithm.

The dataset used consists of 2 data which are training and testing. In manual labeling management data training. In testing the performance of the algorithm using the K-Fold Cross Validation method. For the data testing process, crawl Twitter data in real time and the resulting data then carry out the labeling process and the Preprocessing process using the system. Therefore, the final output displayed is in the form of positive and negative sentiments from the results of the previous Twitter data crawl. The results of the evaluation of the testing process then get an average accuracy value of 79%.

Keyword: Python, Text Mining, Sentiment Analysis, Support Vector Machine