

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP  
LAYANAN INTERNET BIZNET MENGGUNAKAN  
ALGORITME SUPPORT VECTOR MACHINE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**DJUPHA DWI CAHYO**

**16.11.0511**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2023**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP  
LAYANAN INTERNET BIZNET MENGGUNAKAN  
ALGORITME SUPPORT VECTOR MACHINE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**DJUPHA DWI CAHYO**

**16.11.0511**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP  
LAYANAN INTERNET BIZNET MENGGUNAKAN  
ALGORITME SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

**Djupha Dwi Cahyo**

**16.11.0511**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 27-01-2023

**Dosen Pembimbing,**



**Sumarni Adi, S.Kom, M.Cs**

**NIK. 190302256**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP LAYANAN  
INTERNET BIZNET MENGGUNAKAN ALGORITME SUPPORT  
VECTOR MACHINE**

yang disusun dan diajukan oleh

**Djupha Dwi Cahyo**

**16.11.0511**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 27-01-2023

**Susunan Dewan Penguji**

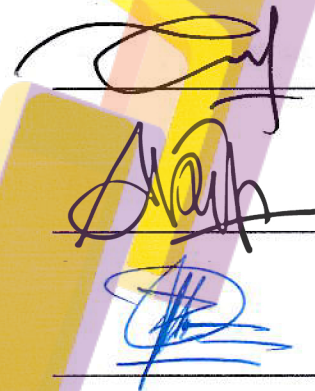
**Nama Penguji**

**Tanda Tangan**

Arif Akbarul Huda, S.Si, M.Eng  
NIK. 190302287

Arif Dwi Laksito, M.Kom  
NIK. 190302150

Arifiyanto Hadinegoro, S.Kom, MT  
NIK. 190302289



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 27-01-2023

**DEKAN FAKULTAS ILMU KOMPUTER**



Hanif Al Fatta, S.Kom., M.Kom.  
NIK. 190302096

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

**Nama mahasiswa : Djupha Dwi Cahyo**  
**NIM : 16.11.0511**

Menyatakan bahwa Skripsi dengan judul berikut:

### **Tuliskan Judul Skripsi**

Dosen Pembimbing : Sumarni Adi, S.Kom., M.Cs.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 27-01-2023

Yang Menyatakan,



Djupha Dwi Cahyo

## KATA PENGANTAR

Puji dan Syukur kita panjatkan kepada Tuhan. Puji Tuhan atas segala pertolongan, rahmat, dan kasih sayang-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Biznet Menggunakan Algoritme Support Vector Machine”. Syukur kepada Tuhan Yang Maha Esa yang senantiasa menjadi sumber inspirasi dan teladan terbaik untuk umat manusia.

Selama menyusun skripsi ini telah banyak hambatan yang penyusun lewati dan tanpa bantuan banyak pihak tentu akan sulit untuk penyusun menyelesaikan skripsi ini, untuk itu penyusun mengucapkan ribuan terima kasih pada :

1. Prof. Dr. M. Suyanto, MM selaku Rektor Universitas AMIKOM Yogyakarta.
2. Hanif Al-Fatta, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Program Studio Informatika Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
4. Sumarni Adi, S.Kom, M.Cs selaku pembimbing Tugas Akhir, terimakasih atas waktu, petunjuk, bimbingan dan ilmu yang telah diberikan kepada penulis untuk dapat menyelesaikan Tugas Akhir ini.
5. Arif Akbarul Huda, S.Si, M.Eng selaku dosen penguji 1, terimakasih juga untuk ilmu-ilmunya, saran-sarannya, perbaikan-perbaikannya, dan masukan yang Bapak berikan untuk penyempurnaan laporan ini.
6. Arif Dwi Laksito, M.Kom selaku dosen penguji 2, terimakasih juga untuk ilmu-ilmunya, saran-sarannya, perbaikan-perbaikannya, dan masukan yang Bapak berikan untuk penyempurnaan laporan ini.
7. Arifiyanto Hadinegoro, S.Kom, MT selaku dosen penguji 3, terimakasih juga untuk ilmu-ilmunya, saran-sarannya, perbaikan-



perbaikannya, dan masukan yang Bapak berikan untuk penyempurnaan laporan ini.

8. Ayahanda Andreas Agus Prasetyo dan Ibunda Fransiska Haryanti, serta saudara Titus Gatot Bayu Setiawan yang telah mendoakan, memberikan dukungan dan semangatnya sehingga penulis dapat menyelesaikan tugas akhir ini.
9. Seluruh keluarga besar TI8 dan teman-teman seperjuangan Teknik Informatika angkatan 2016. Terima kasih atas doa dan semangatnya.

Penulis menyadari dalam penulisan laporan ini masih terdapat kekurangan. Oleh karena itu, saran dan kritik sangat penulis harapkan untuk kemajuan penulis menuju ke arah yang lebih baik lagi. Terimakasih.

Yogyakarta, 27-01-2023

Penulis

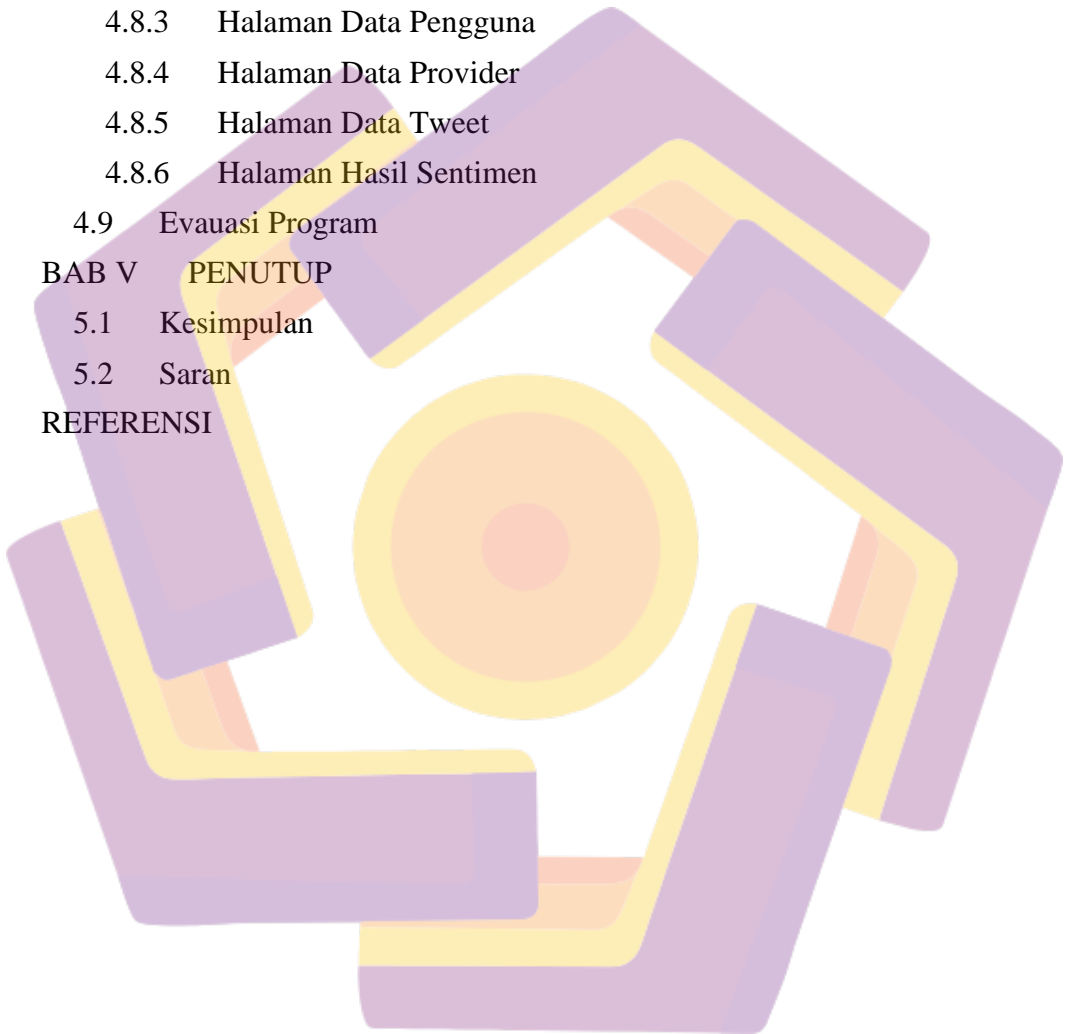
## DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
INTISARI	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penulisan	3
1.5 Manfaat Penulisan	3
1.6 Metode Penelitian	3
1.6.1 Metode Pengumpulan Data	4
1.6.2 Metode Analisis	4
1.6.3 Metode Implementasi	5
1.6.4 Metode Testing	5
1.7 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Pengertian Data Mining	10
2.3 Pengertian Sentiment Analysis	11
2.4 Preprocessing	11
2.5 Term Frequency-Inverse Document Frequency(TF-IDF)	12
2.6 Support Vector Machine	13
2.7 K-Fold Cross Validation	15
2.8 Confusion Matrix	15
2.9 Flowchart	17



2.9.1	System Flowchart	17
2.9.2	Program Flowchart	19
2.10	Bahasa Pemrograman Python	19
2.11	Langkah – langkah pengembangan aplikasi	20
2.11.1	Pre- Processing	21
2.11.2	Transformation	23
2.11.3	Classification	24
2.11.4	Interpretation/Evaluation	24
2.11.5	Testing Data	24
<b>BAB III</b>	<b>ANALISIS DAN PERANCANGAN</b>	<b>26</b>
3.1	Analisis Masalah	26
3.2	Solusi yang Dipilih	26
3.3	Analisis Kebutuhan	27
3.3.1	Analisis Kebutuhan Fungsional	27
3.3.2	Analisis Kebutuhan Non-Fungsional	27
3.4	Analisa Sentimen dengan Klasifikasi Algoritma SVM	28
3.4.1	Cleaning Data	28
3.4.2	Case Foling	30
3.4.3	Tokenization	30
3.4.4	Steaming	31
3.4.5	TF-IDF	32
3.4.6	Consine Similarity	36
3.4.7	Support Vector Machine	40
3.4.1	K-Fold Cross Validation	42
3.5	Perancangan Aplikasi	44
3.5.1	Diagram Alir Preprocessing	44
3.5.2	Proses Klasifikasi Support Vector Machine ( SVM )	45
3.6	Perancangan File Crawling Data Twitter	46
3.7	Perancangan Interface	48
<b>BAB IV</b>	<b>IMPLEMENTASI DAN PEMBAHASAN</b>	<b>51</b>
4.1	Implementasi Pengumpulan Data Tweet	51
4.2	Cleaning Data	73
4.3	Pelabelan Data Training	74
4.4	Proses Olah Data	75

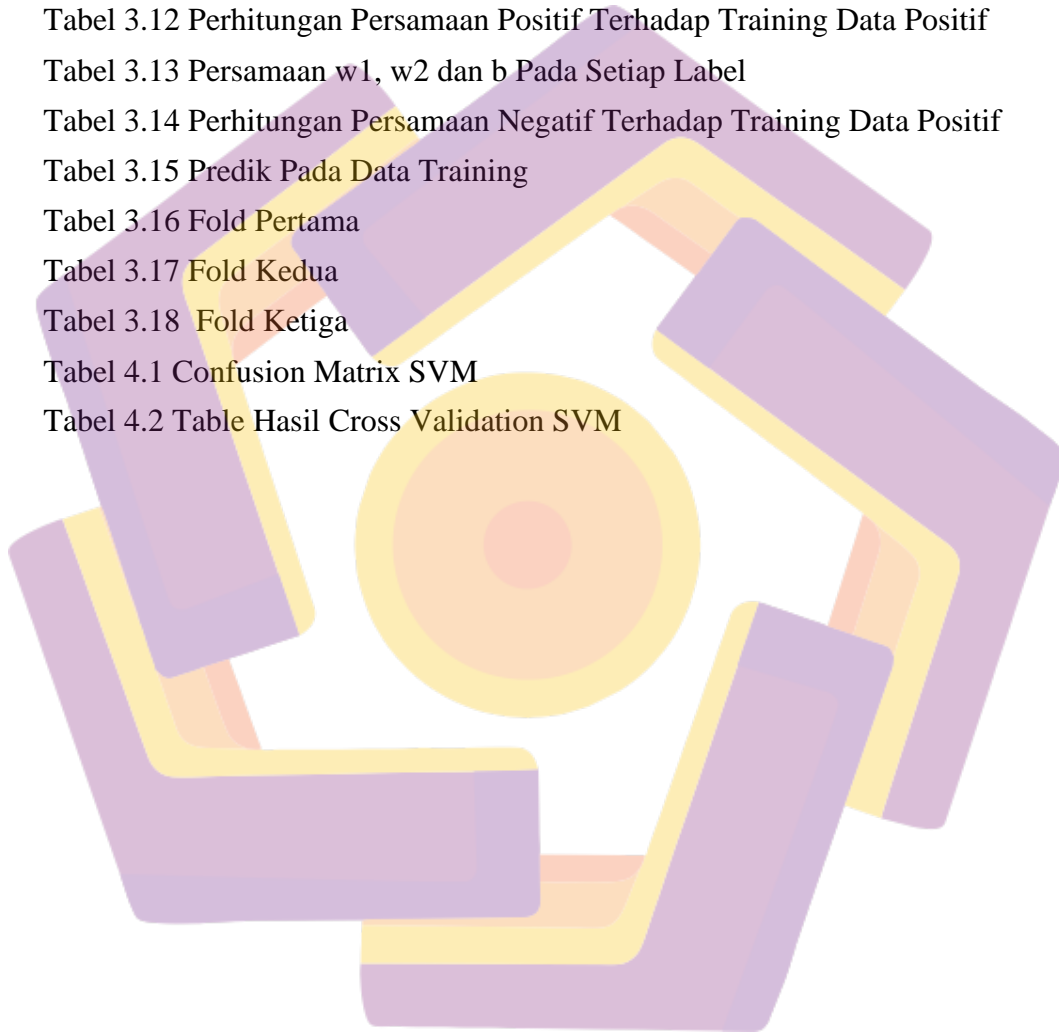
4.5	Proses Import Data	76
4.6	Proses Analisis Sentimen	77
4.7	Proses Pengujian	78
4.8	Pembuatan interface	80
4.8.1	Halaman Login	80
4.8.2	Halaman Utama	81
4.8.3	Halaman Data Pengguna	81
4.8.4	Halaman Data Provider	82
4.8.5	Halaman Data Tweet	82
4.8.6	Halaman Hasil Sentimen	82
4.9	Evauasi Program	83
<b>BAB V</b>	<b>PENUTUP</b>	92
5.1	Kesimpulan	92
5.2	Saran	92
<b>REFERENSI</b>		93



## DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian terkait dengan sistem yang dibangun	9
Tabel 2.2 Confusion Matrix	16
Tabel 2.3 System Flowchart	18
Tabel 2.4 Program Flowchart	19
Tabel 3.1 Data Tweet Sebelum dilakukan Cleaning Data	29
Tabel 3.2 Data Tweet Setelah dilakukan Cleaning data	29
Tabel 3.3 Data Tweet Setelah Case Folding	30
Tabel 3.4 Hasil Tokenization Data Tweet	31
Tabel 3.5 Hasil Steaming Data Tweet	32
Tabel 3.6 Feature List	32
Tabel 3.7 TF IDF Positif	34
Tabel 3.8 TF IDF Negatif	35
Tabel 3.9 Hasil Consine Similarity Positif	37
Tabel 3.10 Hasil Consine Similarity Negatif	38
Tabel 3.11 Hasil Akhir Consine Similarity	39
Tabel 3.12 Perhitungan Persamaan Positif Terhadap Training Data Positif	41
Tabel 3.13 Persamaan $w_1$ , $w_2$ dan $b$ Pada Setiap Label	41
Tabel 3.14 Perhitungan Persamaan Negatif Terhadap Training Data Positif	41
Tabel 3.15 Predik Pada Data Training	42
Tabel 3.16 Fold Pertama	42
Tabel 3.17 Fold Kedua	43
Tabel 3.18 Fold Ketiga	43
Tabel 4.1 Confusion Matrix SVM	84
Tabel 4.2 Table Hasil Cross Validation SVM	86
Tabel 2.2 Confusion Matrix	16
Tabel 2.3 System Flowchart	18
Tabel 2.4 Program Flowchart	19
Tabel 3.1 Data Tweet Sebelum dilakukan Cleaning Data	29
Tabel 3.2 Data Tweet Setelah dilakukan Cleaning data	29
Tabel 3.3 Data Tweet Setelah Case Folding	30
Tabel 3.4 Hasil Tokenization Data Tweet	31
Tabel 3.5 Hasil Steaming Data Tweet	32

Tabel 3.6 Feature List	32
Tabel 3.7 TF IDF Positif	34
Tabel 3.8 TF IDF Negatif	35
Tabel 3.9 Hasil Consine Similarity Positif	37
Tabel 3.10 Hasil Consine Similarity Negatif	38
Tabel 3.11 Hasil Akhir Consine Similarity	39
Tabel 3.12 Perhitungan Persamaan Positif Terhadap Training Data Positif	41
Tabel 3.13 Persamaan $w_1$ , $w_2$ dan $b$ Pada Setiap Label	41
Tabel 3.14 Perhitungan Persamaan Negatif Terhadap Training Data Positif	41
Tabel 3.15 Predik Pada Data Training	42
Tabel 3.16 Fold Pertama	42
Tabel 3.17 Fold Kedua	43
Tabel 3.18 Fold Ketiga	43
Tabel 4.1 Confusion Matrix SVM	84
Tabel 4.2 Table Hasil Cross Validation SVM	86



## DAFTAR GAMBAR

Gambar 2.1 Hyperlane pada Support Vector Machine	14
Gambar 2.2 Langkah – Langkah Analisis Sentimen	21
Gambar 2.3 Langkah – Langkah Pre-Processing	21
Gambar 2.4 Langkah – Langkah Training Data	23
Gambar 2.5 Langkah – Langkah Testing Data	25
Gambar 3.1 Diagram Alir Pre-processing	44
Gambar 3.2 Diagram Alir Klasifikasi SVM	45
Gambar 3.3 Konfigurasi File Crawling Data	47
Gambar 3.4 Diagram alir proses login dan menu utama	48
Gambar 3.5 Diagram alir menu data pengguna	49
Gambar 3.6 Diagram alir menu data provider	49
Gambar 3.7 Diagram alir menu data tweet	50
Gambar 3.8 Diagram alir menu hasil sentimen	50
Gambar 4.1 Script Crawling Data Twitter	51
Gambar 4.2 Hasil Crawling Data	73
Gambar 4.3 Data Bersih	74
Gambar 4.4 Pelabelan Data	74
Gambar 4.5 Halaman Login	81
Gambar 4.6 Halaman Utama	81
Gambar 4.7 Halaman Data Pengguna	81
Gambar 4.8 Halaman Data Provider	82
Gambar 4.9 Halaman Data Tweet	82
Gambar 4.10 Halaman Hasil Sentimen	83
Gambar 4.11 Data Hasil Sentimen dan Grafik Persebaran	83
Gambar 4.12 Grafik Akurasi dari hasil Proses 10-Fold Cross Validation	91

## INTISARI

Internet adalah hal yang saat ini menjadi salah satu kebutuhan pokok bagi masyarakat, selain untuk mencari informasi, hiburan, juga digunakan untuk menghasilkan uang. Biznet merupakan salah satu provider penyedia layanan internet di Indonesia yang memiliki jumlah pengguna terbanyak. Meskipun diminati, tidak semua pengguna berkomentar positif bahkan negatif, kini pengguna dapat menyampaikan opini di berbagai media salah satunya Twitter. Dari berbagai ragam komentar Twitter diperlukan teknik untuk membagi ke dalam kelas opini positif, negatif, maupun netral.

Maka dari itu di penelitian ini akan melakukan analisis sentimen menggunakan metode *Support Vector Machine* untuk klasifikasi. Data yang digunakan berupa opini tentang ulasan Biznet dari media sosial Twitter. Nantinya data dari twitter itu akan terbagi dalam tiga kelas yaitu positif, negatif, dan netral.

Adapun hasil analisis sentimen dari data tweet yang yang dikumpulkan sebanyak 503 ada 373 tweet positif, 118 tweet positif, dan 12 tweet netral. Dan akurasi analisis sentimen yang didapatkan adalah sebesar 75.53%.

**Kata kunci :** Analisis Sentimen, *Support Vector Machine*, Twitter

## **ABSTRACT**

*The internet is currently one of the basic needs for the community, in addition to finding information, and entertainment, it is also used to make money. Biznet is one of the internet service providers in Indonesia with the biggest number of users. Even though it is in demand, not all users comment positively or negatively. Now users can express opinions in various media, one of which is Twitter. From the various kinds of Twitter comments, techniques are needed to divide opinions into positive, negative, and neutral classes.*

*Therefore, in this study, sentiment analysis will be carried out using the Support Vector Machine method for classification. The data used is in the form of opinions about Biznet's reviews from social media Twitter. Later the data from Twitter will be divided into three classes positive, negative, and neutral. The results of the analysis of the tweets data collected were 503, there were 373 positive tweets, 118 positive tweets, and 12 neutral tweets. And the accuracy of the sentiment analysis obtained is 75.53%.*

*The sentiment analysis results from the collected tweet data were 503, there were 373 positive tweets, 118 positive tweets, and 12 neutral tweets. And the accuracy of the sentiment analysis obtained is 75.53%.*

**Keyword :** *Sentiment Analysis, Support Vector Machine, twitter*