

**ANALISIS SENTIMENT PADA TWITTER TERHADAP KEBIJAKAN
SOCIAL DISTANCING DI INDONESIA MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE**

SKRIPSI



disusun oleh

Ilham Naufal Arbani

17.11.1210

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMENT PADA TWITTER TERHADAP KEBIJAKAN
SOCIAL DISTANCING DI INDONESIA MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE**

SKRIPSI

**untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika**



disusun oleh :

Ilham Naufal Arbani

17.11.1210

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

ANALISIS SENTIMENT PADA TWITTER TERHADAP KEBIJAKAN SOCIAL DISTANCING DI INDONESIA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

yang dipersiapkan dan disusun oleh

Ilham Naufal Arbani

17.11.1210

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 10 April 2021

Dosen Pembimbing,

Sumarni Adi, S.Kom, M.Cs

NIK. 190302256

PENGESAHAN

SKRIPSI

ANALISIS SENTIMENT PADA TWITTER TERHADAP KEBIJAKAN SOCIAL DISTANCING DI INDONESIA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

yang dipersiapkan dan disusun oleh

Ilham Naufal Arbani

17.11.1210

telah dipertahankan di depan Dewan Penguji

pada tanggal 21 April 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya D, M.Kom

NIK. 190302185

Yoga Pristyanto, S.Kom, M.Eng

NIK. 190302412

Sumarni Adi, S.Kom, M.Cs

NIK. 190302256

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 April 2021

DEKAN UNIVERSITAS AMIKOM YOGYAKARTA

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 tanggung jawab saya pribadi.

Yogyakarta, 11 Mei 2021



Ilham Naufal Arbani

NIM. 17.11.1210

MOTTO

“Sebaik baik manusia adalah manusia yang bermanfaat bagi orang lain”



PERSEMBAHAN

Alhamdulillah penulis panjatkan puji-syukur kepada Allah SWT atas segala rahmat, taufiq, serta hidayah-Nya, sehingga diberi kesempatan untuk dapat menyelesaikan skripsi ini dengan sebaik-baiknya dengan segala kekurangan penulis. Segala syukur penulis ucapkan kepada-Mu karena telah menghadirkan mereka yang memberikan semangat dan doa disaat menjalani proses pembuatan skripsi ini. Dengan segala kerendahan hati saya persembahkan skripsi ini kepada :

1. Kedua Orang Tua, Bapak Parikesi Arif Chandra Dewa Dan Ibu Nur Khasanah yang selalu mendoakan, memberi semangat serta motivasi supaya dapat menyelesaikan skripsi ini dengan lancar serta bermanfaat bagi semua.
2. Saudariku Zahra Khoirun Nisa yang tidak berhenti untuk memberikan semangat.
3. Ibu. Sumarni Adi, S.Kom, M.Cs. selaku dosen pembimbing dalam skripsi ini yang tidak lelah untuk tetap membimbing dan mengingatkan penulis dari awal hingga akhir proses pembuatan skripsi.
4. Dosen-dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu baik ilmu akademik maupun ilmu non-akademik selama kuliah.
5. Keluarga besar kelas 17-S1IF-05 yang telah bersama-sama menemani selama kuliah. Semoga silaturahmi kita tetap terjaga.
6. Serta orang-orang yang selalu membantu peneliti dalam mengerjakan skripsi yang tidak bisa disebut namanya satu-persatu.

Saya ucapkan terima kasih yang sebesar-besarnya untuk kalian semua. Sukses untuk kalian semua, semoga Allah SWT memberikan rahmat dan hidayah-Nya kepada kita semua. Dan semoga skripsi ini dapat bermanfaat dan berguna untuk kemajuan ilmu pengetahuan kedepannya.

KATA PENGANTAR

Alhamdulillah penulis panjatkan puji syukur kepada Allah SWT atas segala rahmat, taufiq, serta hidayah-Nya kepada penulis sehingga dapat menyelesaikan skripsi yang berjudul “analisis sentiment pada twitter terhadap kebijakansocial distancing di indonesia menggunakan metode support vector machine”.

Selama proses pengerjaan skripsi ini penulis menyadari bahwa dalam proses penulisan skripsi ini banyak mengalami kendala, namun berkat bantuan, bimbingan, kerjasama dari berbagai pihak dan berkah dari Allah SWT sehingga kendala-kendala tersebut bisa diatasi. Selanjutnya ucapa terima kasih penulis sampaikan kepada :

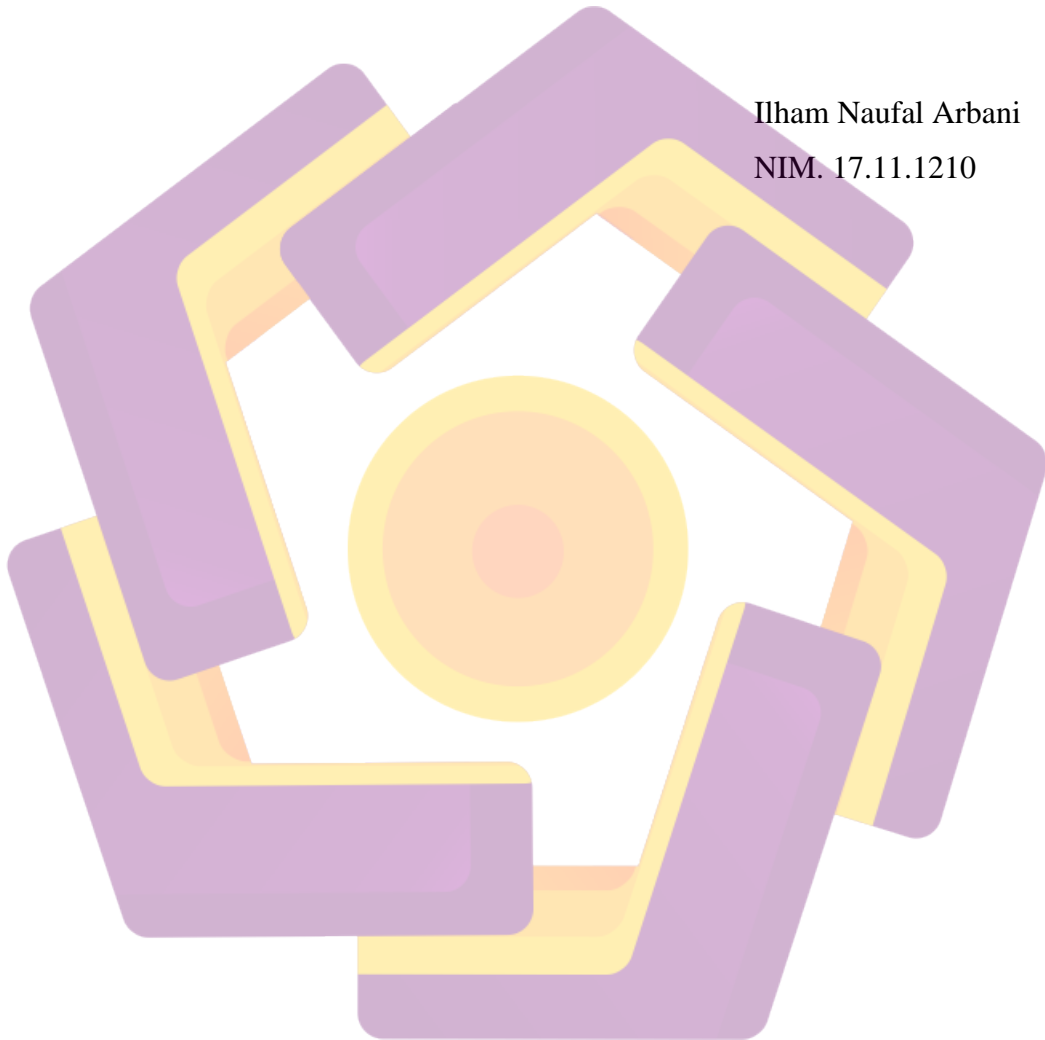
1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu. Sumarni Adi, S.Kom, M.Cs. selaku dosen pembimbing yang telah memberikan banyak masukan yang membantu membimbing dalam menyelesaikan skripsi ini.
3. Ibu Krisnawati, S.Si, M.T selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Ibu Windha Mega D, M.Kom selaku Kepala Prodi Informatika Universitas Amikom Yogyakarta.
5. Dosen Penguji (Bpk/Ibu Dosen Penguji, Bpk/Ibu Dosen Penguji, Bpk/Ibu Dosen Penguji, Bpk/Ibu Dosen Penguji) yang telah memberikan masukan terhadap penelitian ini.
6. Kedua orang tua dan keluarga yang selalu memberikan doa, dukungan dan semangat.
7. Keluarga Besar 17-S1IF-05.
8. Serta semua pihak yang tidak bisa penulis sebutkan satu-persatu yang telah membantu dalam penyusunan skripsi ini.

Penulis Menyadari bahwa masih banyak terdapat kekurangan-kekurangan dalam mengerjakan skripsi ini, sehingga penulis mengharapkan adanya saran dan kritik yang membangun demi kesempurnaan skripsi ini.

Yogyakarta, 11 Mei 2021

Ilham Naufal Arbani

NIM. 17.11.1210



DAFTAR ISI

PERSETUJUAN	iii
PENGESAHAN	iv
PERNYATAAN.....	v
MOTTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
INTISARI.....	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.5.1 Bagi Penulis	4
1.5.2 Bagi Universitas	4
1.6 Metode penelitian.....	4
1.6.1 Metode Pengumpulan Data.....	4
1.6.2 Metode Perancangan	4
1.6.3 Metode Analisis	4
1.6.4 Metode Perancangan	5
1.6.5 Metode Implementasi.....	5
1.6.6 Metode Pengujian.....	5
1.7 Sistematika Penelitian	6
BAB II LANDASAN TEORI	7

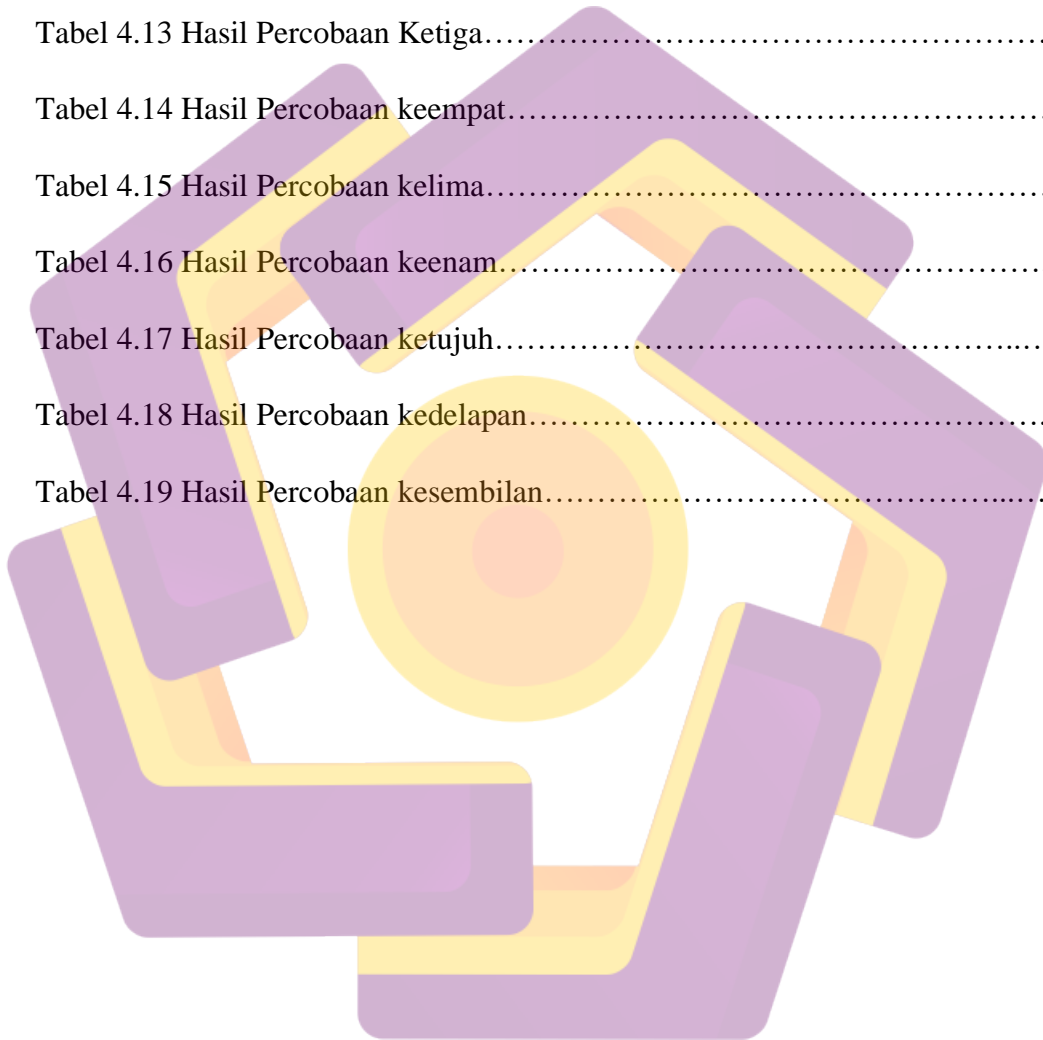
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	10
2.2.1 Analisis Sentimen	10
2.2.2 Text Mining	10
2.2.3 Preprocessing	11
2.2.4 Term Frequency – Inverse Document Frequency (TF-IDF)	13
2.2.5 One vs One.....	14
2.2.5 Support Vector Machine (SVM).....	15
2.2.5.1 Klasifikasi <i>Linear</i>	16
2.2.5.2 Klasifikasi <i>Non Linier</i> SVM	19
2.2.6 Confusion Matrix.....	22
2.2.7 Twitter.....	23
BAB III ANALISIS DAN PERNCANGAN	25
3.1 Tinjauan Umum.....	25
3.2 Pengumpulan Data.....	27
3.2 Preprocesssing	28
3.2.1 Diagramn Alir Penghilang <i>Stopword</i>	30
3.2.2 Steaming	31
3.3 Pembobotan	32
3.3.1 Term Weighting	32
3.3.2 Term frequency (TF)	33
3.3.3 Inverse Document Frequency (IDF).....	33
3.3.4 TF-IDF	34
3.4 Klasifikasi (SVM)	35
3.5 Confusion Matrix.....	40
BAB IV HASIL DAN PEMBAHASAN	42
4.1 Proses Pengumpulan Data	42
4.2 Peprocessing	44
4.2.1 Proses Case Flooding	44
4.2.2 Proses Pembersiahan Dokuman (Cleansing)	45
4.2.3 Proses Tokenizing.....	46
4.2.4 Normalization	47

4.2.5 Proses Stopword Removal	48
4.2.5 Proses <i>Stemming</i>	49
4.3 Proses Pembobotan.....	49
4.4 Klasifikasi Support Vector Machine (SVM).....	50
4.5.1 Percobaan Dengan Data Pelatihan 90% dan Data Pengujian 10%	53
4.5.2 Percobaan Dengan Data Pelatihan 80% dan Data Pengujian 20%	54
4.5.3 Percobaan Dengan Data Pelatihan 70% dan Data Pengujian 30%	56
4.5.4 Percobaan Dengan Data Pelatihan 60% dan Data Pengujian 40%	57
4.5.5 Percobaan Dengan Data Pelatihan 50% dan Data Pengujian 50%	59
4.5.6 Percobaan Dengan Data Pelatihan 40% dan Data Pengujian 60%	60
4.5.7 Percobaan Dengan Data Pelatihan 30% dan Data Pengujian 70%	62
4.5.8 Percobaan Dengan Data Pelatihan 20% dan Data Pengujian 80%	63
4.5.9 Percobaan Dengan Data Pelatihan 10% dan Data Pengujian 90%	65
BAB V KESIMPULAN DAN SARAN.....	70
5.1 Kesimpulan.....	70
5.3 Saran	70
DAFTAR PUSTAKA	71

DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Penelitian.....	8
Tabel 2.2 Tabel <i>Confusion Matrix</i>	22
Tabel 3.1 Tabel Data Tweet.....	27
Tabel 3.2 Tabel Data Set.....	27
Tabel 3.3 Tabel Data Preprocessing.....	31
Tabel 3.4 Proses Pembobotan.....	33
Tabel 3.5 Label Data.....	38
Tabel 3.6 Perhitungan $x_i - x$	38
Tabel 3.7 Perhitungan Panjang Vektor.....	39
Tabel 3.8 Hasil Perhitungan Nilai K Dengan Kernel.....	40
Tabel 3.9 Hasil Perhitungan $y_i y_j$	40
Table 3.10 Table <i>Confusion Matrix</i>	41
Tabel 4.1 Data Tweet.....	43
Tabel 4.2 Case Flooding.....	45
Tabel 4.3 Proses <i>Cleansing</i>	46
Tabel 4.4 Proses <i>Tokenizing</i>	47
Tabel 4.5 Proses <i>Normalization</i>	48
Tabel 4.6 Proses Stopword Removal.....	49
Tabel 4.7 Proses <i>Stemming</i>	50
Tabel 4.8 Proses Pembobotan.....	50

Tabel 4.9 Split Data Pelatihan dan Pengujian.....	52
Tabel 4.10 Confusion Matrix.....	52
Tabel 4.11 Hasil Percobaan Pertama.....	54
Tabel 4.12 Hasil Percobaan Kedua.....	55
Tabel 4.13 Hasil Percobaan Ketiga.....	57
Tabel 4.14 Hasil Percobaan keempat.....	59
Tabel 4.15 Hasil Percobaan kelima.....	61
Tabel 4.16 Hasil Percobaan keenam.....	62
Tabel 4.17 Hasil Percobaan ketujuh.....	64
Tabel 4.18 Hasil Percobaan kedelapan.....	66
Tabel 4.19 Hasil Percobaan kesembilan.....	67



DAFTAR GAMBAR

Gambar 2.1 Metode One vs One	14
Gambar 2.2 Klasifikasi <i>Linear SVM</i>	15
Gambar 2.3 Klasifikasi <i>Linear SVM</i>	15
Gambar 2.4 Klasifikasi <i>Non Linear SVM</i>	19
Gambar 2.5 Transformasi dari Vektor Input ke <i>feature space</i>	20
Gambar 3.1 Alur penelitian <i>sentiment analysis</i>	26
Gambar 3.2 Alur Diagram Preprocessing.....	29
Gambar 3.3 Alur Diagram Penghilang Stopword.....	29
Gambar 3.4 Alur Diagram Stemming.....	30
Gambar 3.5 Diagram Alur Klasifikasi Dengan SVM.....	35
Gambar 3.7 Visualisasi Data.....	37
Gambar 4.1 Perbandingan Data Set dan Akurasi.....	69
Gambar 4.2 Perbandingan Data Set dan Precision.....	69
Gambar 4.3 Perbandingan Data Set dan Recall.....	70
Gambar 4.4 Perbandingan Data Set dan F-Measure.....	70

INTISARI

Pemerintah sebagai pelayan masyarakat memiliki peran yang sangat besar dalam meningkatkan kesejahteraan masyarakat. Maka perlu diadakan evaluasi secara bertahap guna meningkatkan pelayanan masyarakat sebagai tugas pemerintah, untuk itu perlu adanya sikap terbuka dari pemerintah untuk menerima kritik, keluhan, dan saran masyarakat mengenai kebijakan yang dilakukan pemerintah yang berhubungan langsung dengan kesejahteraan masyarakat. Media sosial pada saat ini merupakan media komunikasi yang sangat populer dikalangan masyarakat Indonesia. Salah satu media sosial sekaligus media mikro blogging yang sedang digandrungi masyarakat Indonesia saat ini adalah twitter. Dengan media sosial twitter, informasi masyarakat sangat beragam, dari informasi tersebut terdapat data yang dapat diolah menjadi analisis sentimen. Penelitian ini bertujuan untuk membuat sistem analisis sentiment masyarakat yang berupa sentimen positif atau negatif.

Metode yang digunakan untuk klasifikasi sentimen ini adalah Support Vector Machine dengan tahap evaluasi performansi menggunakan metode Confusion Matrix. Masukan sistem ini berupa data tweet dari masyarakat Indonesia dengan pencarian social distancing dan #dirumahaja, sedangkan hasil keluar sistem ini berupa visualisasi data sentiment positif dan negatif. Hasil visualisasi ini dapat digunakan oleh pemerintah Indonesia untuk mengevaluasi kebijakan yang telah dilakukan dan menentukan kebijakan yang diambil pada masa mendatang.

Berdasarkan hasil pengujian, algoritma SVM (Support Vector Machine) dapat mengklasifikasi tweet positif, negative, dan netral. Hasil dari pengujian dengan menggunakan model SVM dengan jumlah dataset sebanyak 1.500 yang terdiri dari tiga kelas yaitu Positif dengan 500 tweet, kelas Negatif dengan 500 tweet, dan kelas Netral dengan 500 tweet diperoleh akurasi tertinggi sebesar 97% dengan data training 80% dan data testing sebanyak 20%.

Kata Kunci : *Support Vector Machine, Analisis Sentimen, Klasifikasi*

ABSTRACT

The government as a public servant has a very big role in improving the welfare of the community. So it is necessary to hold a gradual evaluation in order to improve public services as a government task, for that it is necessary to have an open attitude from the government to accept criticism, complaints, and public suggestions regarding policies carried out by the government which are directly related to the welfare of the community. Social media is currently a very popular medium of communication among Indonesians. One of the social media as well as micro blogging media that is being loved by Indonesians today is Twitter. With the social media twitter, public information is very diverse, from this information there is data that can be processed into sentiment analysis. This study aims to create a public sentiment analysis system in the form of positive or negative sentiments.

The method used for this sentiment classification is the Support Vector Machine with the performance evaluation stage using the Confusion Matrix method. The input of this system is in the form of tweet data from the Indonesian people by searching for social distancing and #dirumahaja, while the output of this system is in the form of visualizing positive and negative sentiment data. The results of this visualization can be used by the Indonesian government to evaluate policies that have been implemented and determine policies to be taken in the future.

Based on the test results, the SVM (Support Vector Machine) algorithm can classify positive, negative, and neutral tweets. The results of the test using SVM model with a total of 1,500 datasets consisting of three classes, namely Positive with 500 tweets, Negative class with 500 tweets, and Neutral class with 500 tweets with the highest accuracy of 97% with 80% training data and 20% testing data.

Key Word : Support Vector Machine, Sentiment Analysis, Clasification