

**ANALISIS SENTIMEN PADA KOMENTAR VIDEO  
YOUTUBE MENGGUNAKAN METODE  
SUPPORT VECTOR MACHINE**

**SKRIPSI**



disusun oleh:

**Iqbal Rilo Pambudi**

**17.11.1259**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

**ANALISIS SENTIMEN PADA KOMENTAR VIDEO  
YOUTUBE MENGGUNAKAN METODE  
SUPPORT VECTOR MACHINE**

**SKRIPSI**

Untuk memenuhi sebagian persyaratan  
mencapai gelar sarjana pada program studi Informatika



disusun oleh:

**Iqbal Rilo Pambudi**

**17.11.1259**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

# PERSETUJUAN

## SKRIPSI

**ANALISIS SENTIMEN PADA KOMENTAR VIDEO  
YOUTUBE MENGGUNAKAN METODE  
SUPPORT VECTOR MACHINE**

yang dipersiapkan dan disusun oleh

**Iqbal Rilo Pambudi**

**17.11.1259**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 27 Maret 2021

**Dosen Pembimbing,**

**Ferian Fauzi Abdulloh, M.Kom.**

**NIK. 190302276**

# PENGESAHAN

## SKRIPSI

### ANALISIS SENTIMEN PADA KOMENTAR VIDEO YOUTUBE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

yang dipersiapkan dan disusun oleh

**Iqbal Rilo Pambudi**

**17.11.1259**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 17 Maret 2021

#### Susunan Dewan Penguji

Nama Penguji

Ferian Fauzi Abdullah, M.Kom

NIK. 190302276

Agung Nugroho, M.Kom

NIK. 190302242

Ali Mustopa, M.Kom

NIK. 190302192

Tanda Tangan

---

---

---

Skrripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 17 Maret 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom., 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, 17 Maret 2021



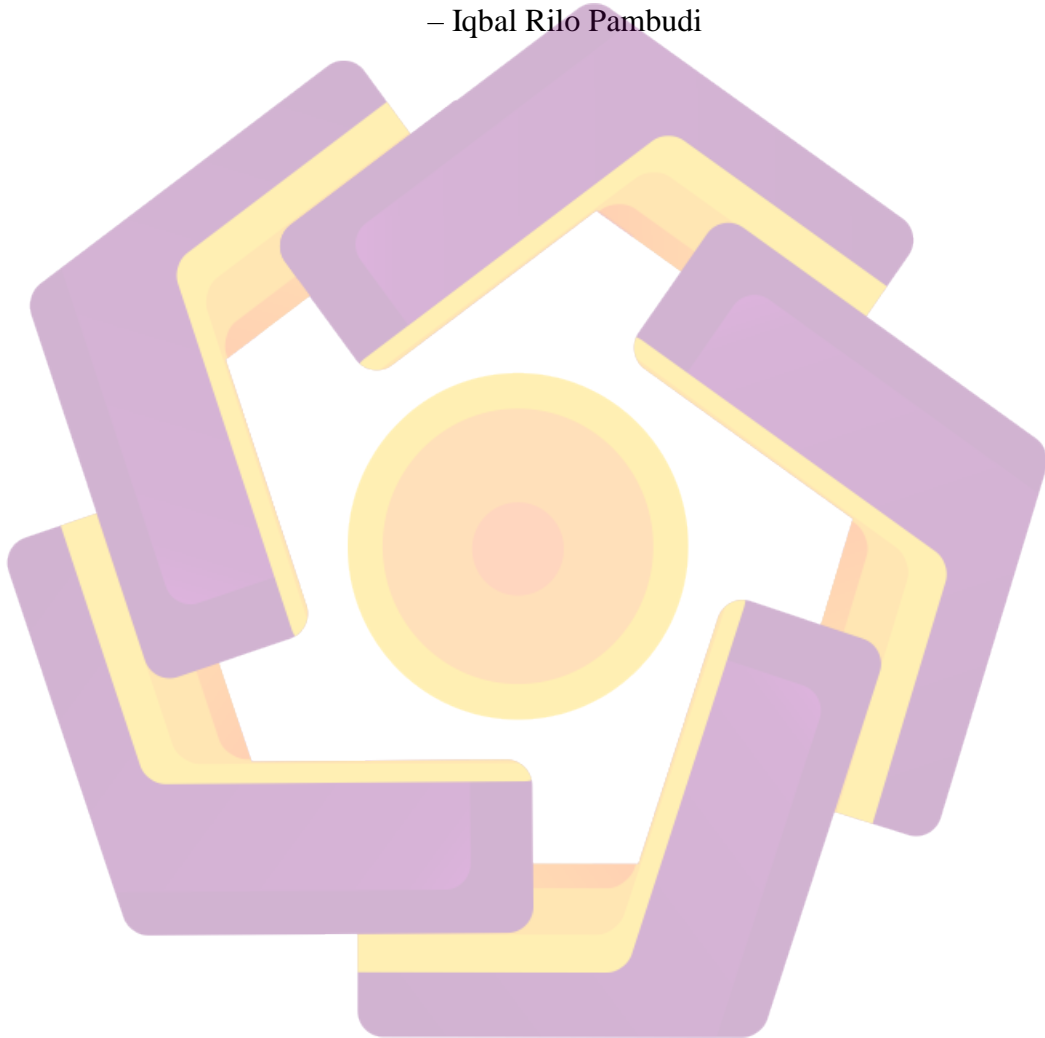
Iqbal Rilq Pambudi

NIM. 17.11.1259

## MOTTO

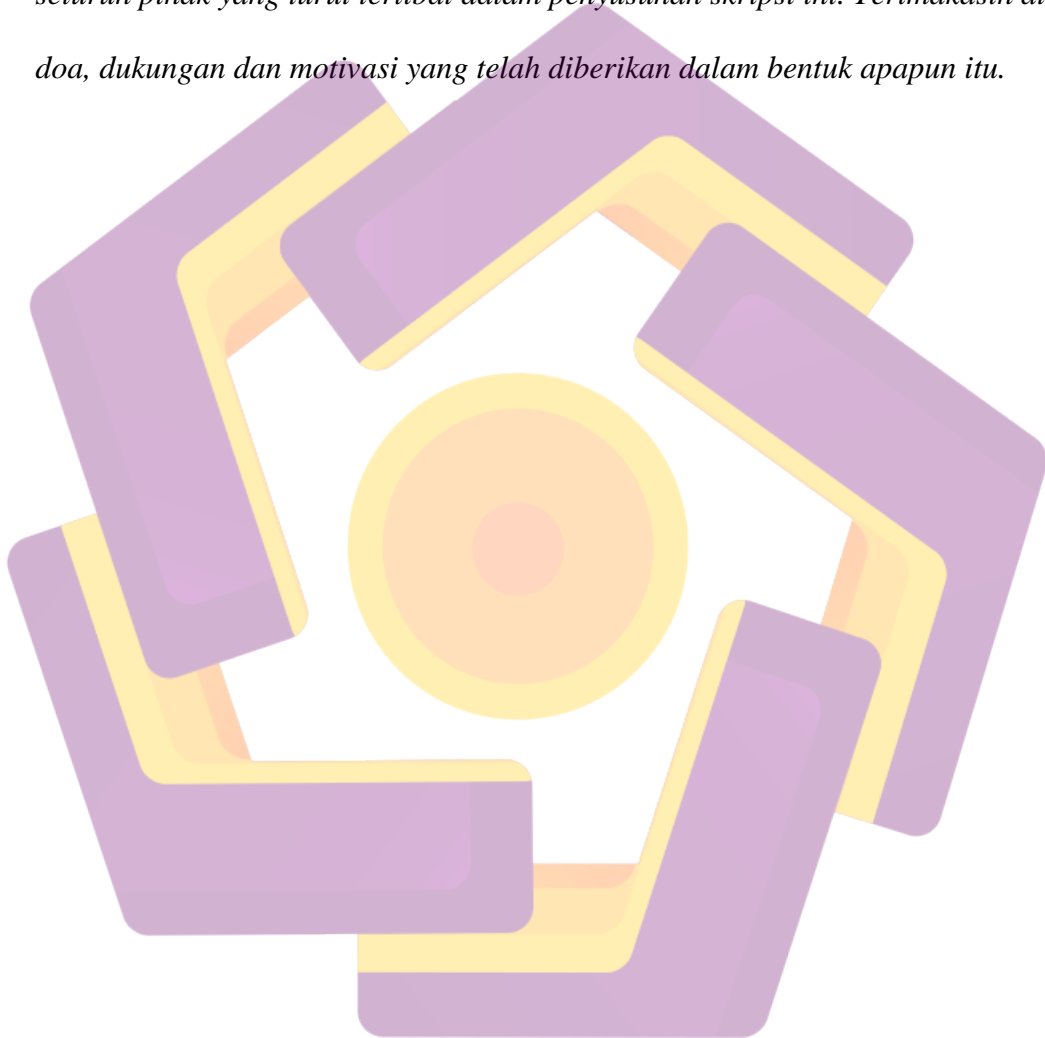
”Tetaplah berbuat baik dimanapun kamu berada”

– Iqbal Rilo Pambudi



## **PERSEMBAHAN**

*Skripsi ini saya persembahkan teruntuk kedua orang tua saya, adik serta seluruh pihak yang turut terlibat dalam penyusunan skripsi ini. Terimakasih atas doa, dukungan dan motivasi yang telah diberikan dalam bentuk apapun itu.*



## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, shalawat serta salam tak lupa kepada junjungan kita Nabi Muhammad SAW yang telah memberikan tauladan dan menuntun ummatnya sehingga penulis dapat menyelesaikan skripsi ini.

Skripsi yang berjudul **“Analisis Sentimen pada Komentar Video Youtube dengan Metode Support Vector Machine”** ini disusun sebagai salah satu syarat dalam mendapat kelulusan untuk mendapatkan gelar sarjana program Strata Satu (S1) di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Penulis mengucapkan banyak terimakasih kepada seluruh pihak yang telah membantu dalam penyusunan skripsi ini. Penulis sadar keberhasilan skripsi ini tak luput dari perhatian, dukungan dan bimbingan dari masing-masing pihak. Maka dari itu penulis ingin menyampaikan ucapan terimakasih kepada :

1. Allah SWT yang telah memberikan rahmat dan karunia dan ridho-Nya sampai sekarang kepada penulis.
2. Kedua orang tua tercinta, adik, serta keluarga besar. Terimakasih atas dukungan dan doa yang diberikan kepada penulis.
3. Pak Ferian Fauzi Abdullah, M. Kom. yang telah membimbing penulis dari awal sampai skripsi ini selesai.
4. Dosen-dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu selama kuliah.



5. Seluruh teman-teman Informatika 5 angkatan 2017 yang telah membantu belajar dan memberikan pengalaman yang berharga selama berkuliah dan tidak bisa penulis sebutkan satu persatu.
6. Seluruh teman-teman FOSSIL yang selalu mendukung dan memberikan semangat kepada penulis.
7. Teman-teman kos ali imron yang selalu menemani dari semester I sampai sekarang. Terimakasih atas doa, dukungan serta semangat yang diberikan kepada penulis (Vicky, Indra, Yhogi, Emen, Ari).
8. Sahabat-sahabat anggun yang telah kebersamai sejak 2014 sampai sekarang. Penulis mengucapkan terimakasih banyak atas doa, dukungan, motivasi dan semangat yang telah kalian berikan (Yuvina, Riska, Lutfi, Danang, Rezha, Febri, Galih, Dimas dan Burhan).
9. Seluruh pihak yang telah telah memberikan doa dan dukungan yang tidak bisa penulis sebutkan satu persatu. Semoga Allah SWT membalas semua kebaikan kalian semua. Aamiin.

Penulis sadar masih banyak terdapat kekurangan karena keterbatasan pengetahuan penulis dalam penyusunan skripsi ini. Maka dari itu penulis mengharapkan kritik dan saran yang membangun dari semua pihak. Semoga penelitian ini bermanfaat dan dapat digunakan untuk menambah wawasan dan informasi bagi pihak yang membutuhkan.

Yogyakarta, 17 Maret 2021

Penulis,  
Iqbal Rilo Pambudi  
NIM. 17.11.1259

## DAFTAR ISI

JUDUL.....	i
PERSETUJUAN.....	ii
PENGESAHAN .....	iii
PERNYATAAN.....	iv
MOTTO .....	v
PERSEMBAHAN .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiv
BAB I.....	1
1.1. LATAR BELAKANG.....	1
1.2. RUMUSAN MASALAH .....	2
1.3. BATASAN MASALAH .....	3
1.4. TUJUAN PENELITIAN .....	3
1.5 MANFAAT PENELITIAN .....	4
1.6 METODOLOGI PENELITIAN .....	4
1.7 SISTEMATIKA PENULISAN .....	6
BAB II.....	9
2.1 KAJIAN PUSTAKA .....	9
2.2 DASAR TEORI.....	11
2.2.1 Youtube.....	11
2.2.2 Analisis Sentimen .....	12
2.2.3 Text Mining .....	12
2.2.2 Text Preprocessing.....	13
2.2.3 Pembobotan Kata ( <i>Term Weighting</i> ) .....	16
2.2.4 Support Vector Machine (SVM).....	18
2.2.5 Confusion Matriks .....	22
BAB III .....	25

3.1 Analisis Masalah .....	26
3.2 Perancangan Sistem.....	26
3.2.1 Sumber Data .....	28
3.2.2 Pengumpulan Data.....	28
3.2.3 Pelabelan Data .....	32
3.2.2 Text Preprocessing.....	33
3.2.3 Pembobotan Term.....	38
3.2.4 Support Vector Machine.....	40
3.3 Implementasi Sistem .....	42
3.3.1 Pengumpulan Data .....	42
3.3.5 Text Preprocessing.....	49
3.3.6 Pembobotan Term.....	54
3.3.7 Support Vector Machine.....	55
BAB IV .....	57
4.1 Perancangan Perangkat Keras .....	57
4.2 Perancangan Perangkat Lunak .....	57
4.3 Pengujian Model.....	58
4.4 Pengujian Data .....	63
BAB V.....	70
5.1 Kesimpulan.....	70
5.2 Saran.....	71

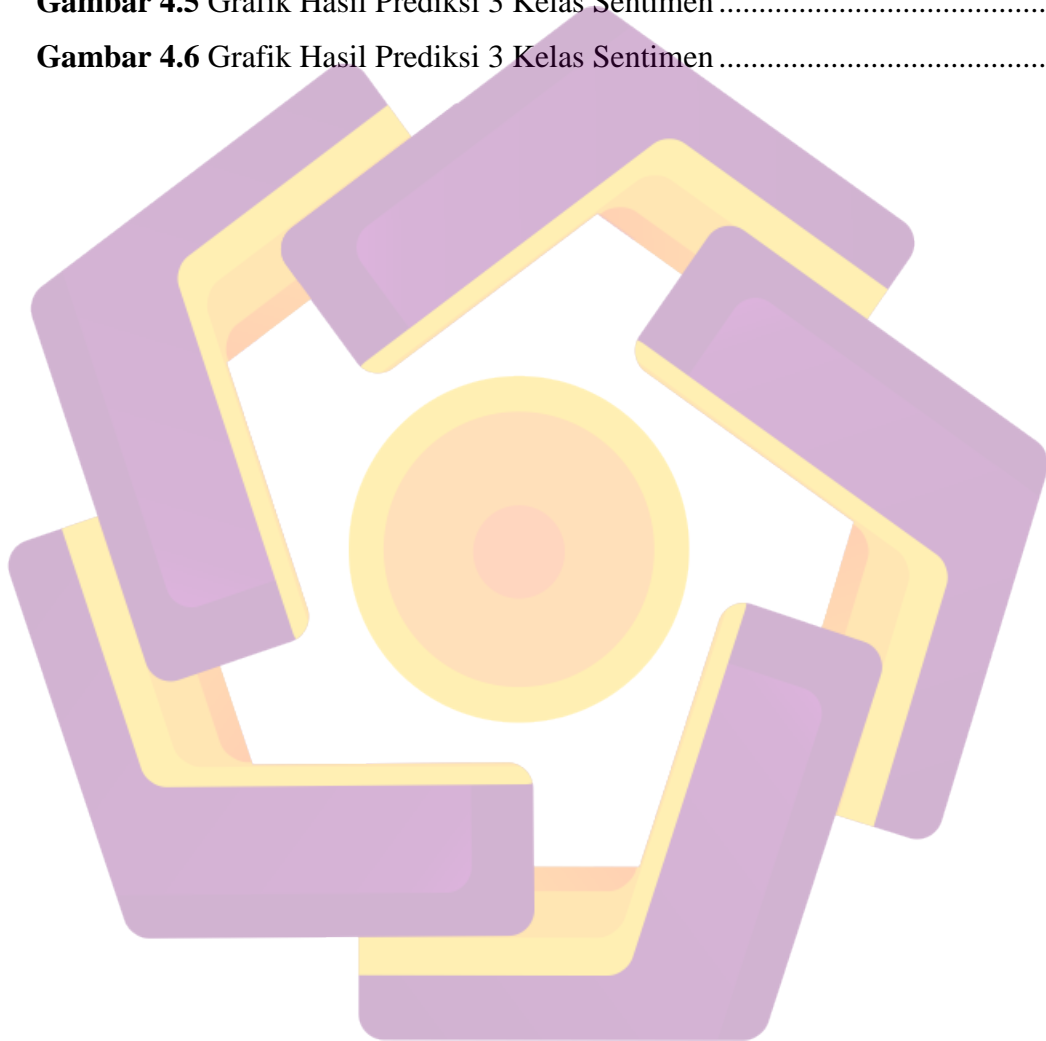
## DAFTAR TABEL

<b>Tabel 2.1</b> Hasil Perbandingan NBC dan SVM.....	10
<b>Tabel 2.2</b> Confusion Matrix .....	23
<b>Tabel 3.1</b> Biaya <i>Credit</i> Youtube Data API v3.....	29
<b>Tabel 3.2</b> Dataset dengan label.....	32
<b>Tabel 3.3</b> Sampel dataset normalisasi .....	33
<b>Tabel 3.4</b> Simulasi Case Folding.....	35
<b>Tabel 3.5</b> Simulasi Text Cleaning.....	35
<b>Tabel 3.6</b> Simulasi Tokenizing.....	36
<b>Tabel 3.7</b> Simulasi Normalisasi.....	36
<b>Tabel 3.8</b> Simulasi Stopword Removal.....	37
<b>Tabel 3.9</b> Simulasi Stemming .....	38
<b>Tabel 3.10</b> Simulasi TF-IDF .....	39
<b>Tabel 3.11</b> Dataset Simulasi Perhitungan SVM.....	40
<b>Tabel 3.12</b> Nilai Akhir Perhitungan SVM.....	41
<b>Tabel 3.13</b> Hasil Klasifikasi Sentimen .....	42
<b>Tabel 3.14</b> Sampel dataset.....	47
<b>Tabel 3.15</b> Sampel Dataset dengan label.....	48
<b>Tabel 4.1</b> Data latih SVM.....	58
<b>Tabel 4.2</b> Hasil pengujian dengan 3 kelas sentimen.....	62
<b>Tabel 4.3</b> Hasil akurasi dengan 2 kelas sentimen.....	62
<b>Tabel 4.4</b> Objek pengujian .....	63
<b>Tabel 4.5</b> Hasil pengujian dengan objek video pada youtube dengan 3 kelas .....	64
<b>Tabel 4.6</b> Confusion Matrix .....	66
<b>Tabel 4.7</b> Hasil pengujian dengan objek video pada youtube dengan 2 kelas .....	66
<b>Tabel 4.8</b> Hasil Prediksi Sentimen dengan 3 kelas.....	67

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Logo Youtube .....	11
<b>Gambar 2.2</b> Proses <i>Case Folding</i> .....	13
<b>Gambar 2.3</b> Proses <i>Text Cleaning</i> .....	14
<b>Gambar 2.4</b> Penerapan <i>Tokenization</i> .....	14
<b>Gambar 2.5</b> Penerapan Normalisasi Teks .....	15
<b>Gambar 2.6</b> Ilustrasi <i>Support Vector Machine</i> .....	19
<b>Gambar 3.1</b> Flowchart Sistem .....	25
<b>Gambar 3.2</b> Flowchart Sistem .....	28
<b>Gambar 3.3</b> Flowchart text preprocessing .....	34
<b>Gambar 3.4</b> Import module dan mendefinisikan service .....	44
<b>Gambar 3.5</b> Mendefinisikan fungsi <code>search_videos_by_keyword</code> .....	45
<b>Gambar 3.6</b> Mendefinisikan fungsi <code>get_videos</code> .....	46
<b>Gambar 3.7</b> Mendefinisikan fungsi <code>get_video_comments</code> .....	46
<b>Gambar 3.8</b> Mendefinisikan fungsi <code>write_to_csv</code> .....	47
<b>Gambar 3.9</b> Membaca dataset .....	50
<b>Gambar 3.10</b> Menghitung jumlah sentimen per kelas .....	50
<b>Gambar 3.11</b> Pembagian sampel dataset .....	50
<b>Gambar 3.12</b> Text Cleaning .....	52
<b>Gambar 3.13</b> Tokenisasi .....	52
<b>Gambar 3.14</b> Normalisasi kata .....	53
<b>Gambar 3.15</b> Stopword Removal .....	53
<b>Gambar 3.16</b> Stemming .....	54
<b>Gambar 3.17</b> Pembagian data uji dan data latih .....	55
<b>Gambar 3.18</b> Pembobotan Term dengan TF-IDF .....	55
<b>Gambar 3.19</b> Support Vector Machine .....	56
<b>Gambar 3.20</b> Prediksi hasil .....	56
<b>Gambar 3.21</b> Nilai akurasi .....	56

<b>Gambar 4.1</b> Grafik Dataset.....	59
<b>Gambar 4.2</b> WordCloud Sentimen Negatif .....	60
<b>Gambar 4.3</b> WordCloud Sentimen Netral .....	61
<b>Gambar 4.4</b> WordCloud Sentimen Positif.....	61
<b>Gambar 4.5</b> Grafik Hasil Prediksi 3 Kelas Sentimen .....	67
<b>Gambar 4.6</b> Grafik Hasil Prediksi 3 Kelas Sentimen .....	67



## INTISARI

Vaksinasi yang dilakukan pemerintah Indonesia per tanggal 13 Januari 2021 menjadi momen yang penting. Vaksinasi menjadi pilihan pemerintah Indonesia untuk menekan grafik penularan virus Covid-19. Indonesia telah menyiapkan dosis vaksin sebesar 371 juta vaksin corona. Namun, isu tentang vaksin juga beredar luas di masyarakat. Rumor tentang vaksin yang belum aman dan kurangnya sosialisasi menimbulkan pro dan kontra terhadap vaksinasi ini. Maka dari itu mencari tahu opini masyarakat tentang vaksinasi ini menjadi opsi untuk menentukan sentimen masyarakat terhadap vaksin.

Media sosial menjadi salah satu sumber opini masyarakat dalam menuangkan pendapatnya. Pengolahan opini tersebut dapat menjadi sebuah alternatif untuk menentukan respon publik terhadap suatu peristiwa tertentu. Menurut HootSuite Indonesia tahun 2021 Youtube menjadi social media dengan pengguna terbanyak. Dengan pertimbangan tersebut, Youtube menjadi sebuah pilihan untuk menjadi sumber dataset.

Selanjutnya dataset tersebut akan diolah dan diproses. Menggunakan metode Support Vector Machine data yang diolah tersebut akan dijadikan model klasifikasi untuk melakukan pengklasifikasian teks dari komentar Youtube dengan topik bahasan “vaksin covid-19”. Hasil dari pengklasifikasian tersebut diharapkan dapat menjadi informasi maupun masukan kepada pihak tertentu untuk dijadikan pertimbangan.

**Kata Kunci:** Support Vector Machine, Youtube, Vaksin, Klasifikasi

## ABSTRACT

*The vaccination carried out by the Indonesian government on January 13, 2021 is an important moment. Vaccination is the choice of the Indonesian government to suppress the graph of the transmission of the Covid-19 virus. Indonesia has prepared a vaccine dose of 371 million corona vaccines. However, the issue of vaccines is also widely circulated in the community. Rumors about unsafe vaccines and lack of socialization lead to pros and cons of this vaccination. Therefore, finding out public opinion about vaccination is an option to determine public sentiment towards vaccines.*

*Social media is a source of public opinion in expressing their opinions. Opinion processing can be an alternative to determine the public's response to a particular event. According to HootSuit Indonesia in 2021 Youtube will become the social media with the most users. With these considerations, Youtube is an option to become a dataset source.*

*Furthermore, the dataset will be processed and processed. Using the Support Vector Machine method, the processed data will be used as a classification model to classify the text from Youtube comments with the topic "covid-19 vaccine". The results of the classification are expected to provide information and input to certain parties for consideration.*

**Keywords:** Support Vector Machine, Youtube, Vaccines, Classification