

## SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM

Auliya Rahman Isnain<sup>1)</sup>, Adam Indra Sakti<sup>2)</sup>, Debby Alita<sup>3)</sup>, Nurman Satya Marga<sup>4)</sup>

<sup>1,2,3,4</sup>Program Studi S1 Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia  
<sup>1,2,3,4</sup>Jl.ZA.Pagaralam No.9-11 Kedaton Bandarlampung

Email: <sup>1</sup>auliyarahman@teknokrat.ac.id, <sup>2</sup>adamindrasakti8@gmail.com, <sup>3</sup>debbyalita@teknokrat.ac.id

### Abstract

*Social media makes people experience a shift in behavior, both culture, ethics, and existing norms so that they can express their opinions. The opinion is an opinion from the public's thinking about a problem that is happening, currently, Indonesia is being faced with a problem regarding the Covid-19 virus which has taken so many lives so that people issue their opinions about the virus and the policies implemented by the government to deal with the virus. This study aims to determine how public sentiment is on policies that will be carried out by the government regarding lockdown policies or large-scale social restrictions using the Support Vector Machine method with the extraction of the TF-IDF feature with testing which will see the accuracy, precision, recall, and F1-Score values. . The use of the Support Vector Machine method and feature extraction with TF-IDF which divides the class into positive sentiment 68.75% and negative 31.25% results in an accuracy value of 74%, precision of 75%, recall of 92% and F1-Score of 83 %.*

**Keyword:** Sentiment Analysis, Support Vector Machine, Tf-idf, Twitter, Lockdown, Covid-19

### Abstrak

Media sosial menjadikan masyarakat mengalami pergeseran perilaku baik budaya, etika dan norma yang ada, sehingga mereka dapat mengeluarkan opini-opini yang mereka miliki. Opini merupakan suatu pendapat dari pemikiran masyarakat mengenai suatu permasalahan yang sedang terjadi, saat ini Indonesia sedang dihadapkan oleh masalah mengenai virus Covid-19 yang memakan begitu banyak korban jiwa sehingga masyarakat mengeluarkan opini mereka mengenai virus tersebut dan kebijakan yang dilakukan pemerintah menghadapi virus tersebut.

Penelitian ini bertujuan untuk mengetahui bagaimana sentiment publik terhadap kebijakan yang akan dilakukan pemerintah mengenai kebijakan lockdown ataupun pembatasan sosial berskala besar menggunakan metode *Support Vector Machine* dengan ekstraksi fitur *tf-idf* dengan pengujian yang nantinya akan dilihat bagaimana nilai *accuracy*, *precision*, *Recall* dan *F1-Score*.

Penggunaan metode *Support Vector Machine* dan ekstraksi fitur dengan *tf-idf* yang membagi kelas menjadi sentiment positif 68,75% dan negative 31,25% menghasilkan nilai *accuracy* sebesar 74%, *precision* sebesar 75%, *recall* sebesar 92% dan *F1-Score* sebesar 83%.

**Kata Kunci:** Sentiment Analisis, Support Vector Machine, Tf-idf, Twitter, Lockdown, Covid-19

### 1. Pendahuluan

Lahirnya media sosial menjadikan pola perilaku masyarakat mengalami pergeseran baik budaya, etika dan norma yang ada, sehingga dari berbagai kalangan dan usia semua masyarakat Indonesia memiliki dan menggunakan media sosial sebagai salah satu sarana guna memperoleh dan menyampaikan informasi ke masyarakat [1]. Saat ini media sosial sebagai media

komunikasi telah menjadi bagian penting dalam kehidupan masyarakat luas sehari-hari [2], dengan munculnya media sosial telah merubah cara pemikiran seseorang dalam mengekspresikan pendapat. Selain itu, teknologi internet juga berperan aktif dalam masyarakat mengakses informasi dari manapun[3]. Dari hasil survey yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) tahun 2018, mengungkapkan bahwa jumlah penduduk Indonesia mencapai 264,16 juta

jiwa, dari jumlah tersebut sebanyak 171,17 juta adalah pengguna internet atau sebanyak 64,8 % dari populasi penduduk Indonesia. Meningkatnya aktivitas pengguna di media sosial ini tidak menutup kemungkinan akan adanya saran, pendapat ataupun tanggapan dari masyarakat luas mengenai sesuatu yang sedang terjadi [4]. Twitter merupakan media sosial yang dapat dipakai oleh berbagai pihak untuk mengekspresikan pendapat mereka secara bebas [5]. Saat ini Twitter merupakan sebuah media sosial yang baik untuk memberikan pengaruh dalam penelitian [6] Namun sejauh ini masih belum banyak aplikasi dan metode *sentiment analisis* yang dikembangkan untuk bahasa Indonesia, hal tersebut yang mendorong perlunya dilakukan *sentiment analisis* untuk dokumen bahasa Indonesia. Penelitian *sentiment analisis* ini dilakukan untuk mengetahui *sentiment* mengenai sesuatu dengan menggunakan pendekatan dalam *machine learning* yang dikenal dengan nama *support vector machine* yang dikhususkan dengan dokumen teks berbahasa Indonesia. Pemilihan metode *support vector machine* dalam penelitian ini karena memiliki kemampuan generalisasi dalam mengklasifikasikan suatu *pattern*, dan tidak termasuk data yang dipakai dalam fase pembelajaran metode tersebut [7].

## 2. Landasan Teori

### 2.1 Text Mining

*Text mining* merupakan bagian dari data mining, yaitu proses untuk memperoleh suatu pengetahuan menggunakan seperangkat alat analisis dimana pengguna berinteraksi dengan sekumpulan dokumen dari waktu ke waktu. Seperti halnya *data mining*, *text mining* berusaha untuk mengekstraksi informasi dari suatu sumber data (sekumpulan dokumen) melalui identifikasi dan eksplorasi pola yang ada. Teks mining merupakan teknik yang digunakan untuk menangani *classification*, *information extraction* dan *information retrieval*. Konsep teks mining yang digunakan dalam klasifikasi yaitu dokumen tekstual dengan tujuan untuk mengklasifikasi dokumen yang sesuai dengan topic pembahasan. Perbedaan antara *data mining* dan *text mining* terletak pada preprocessing, pada data mining preprocessing berfokus pada penomoran (*indexing*) dan normalisasi data, sedangkan *text mining* berfokus pada identifikasi dan ekstraksi fitur [8]

### 2.2 Text Preprocessing

Dalam melakukan *text mining*, *text* dokumen harus dilakukan persiapan terlebih dahulu sehingga dapat digunakan untuk proses utama. Proses mempersiapkan *text* atau data mentah ini disebut *text preprocessing*. *Text preprocessing* berfungsi mengubah data yang tidak terstruktur menjadi terstruktur (Han dan Kamber, 2006). Proses yang dilakukan dalam preprocessing adalah sebagai berikut :

- Case folding**, merupakan proses merubah kalimat data teks menjadi seragam
- Cleansing**, yaitu proses membersihkan dokumen dan menyeleksi kata yang tidak diperlukan seperti html, emoticon, hashtag, mention dan url.
- Stopwords**, yaitu menghilangkan kata yang kurang efektif
- Steaming**, yaitu proses untuk menyaring kata yang terdapat kata sambung, kata ganti, kata depan, menjadi kata dasar dengan menghilangkan awalan atau akhiran.
- Tokenization**, merupakan proses seleksi pemotongan kata dalam kalimat. Diberikan pemisah seperti tanda koma (,), titik (.), dan tanda pemisah lainnya.

### 2.3 Sentiment Analysis

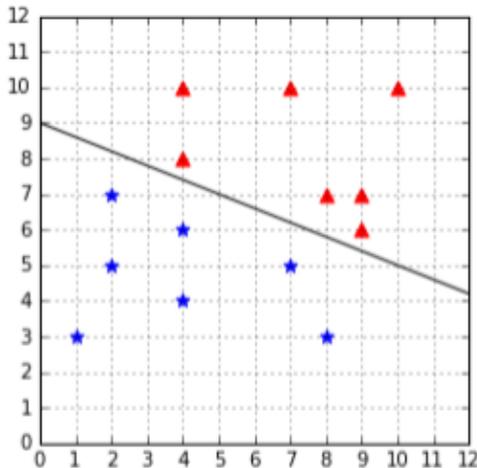
*Sentiment Analisis* adalah *sentiment* dari teks subjektif tersebut menganalisis, memproses, meringkas dan proses inferensial. *Sentiment analisis* saat ini dibagi menjadi penggunaan klasifikasi pembelajaran mesin dan metode klasifikasi berdasarkan aturan, metode pembelajaran mesin menggunakan kata kata emosi sebagai klasifikasi fitur, dan kamus emosi dapat digunakan untuk mewujudkan pemilihan karakteristik *sentiment* dengan cepat dan efisien. Dalam kombinasi dengan pengklasifikasian dalam menyelesaikan tugas metode yang umumnya digunakan adalah *Naïve Bayes (NB)*, *Support Vector Machine (SVM)* dan *Entropi Maksimum (EM)* [9].

### 2.4 Support Vector Machine

*Support Vector Machine* merupakan sistem pembelajaran yang menggunakan hipotesis berupa fungsi-fungsi linear dalam sebuah fitur yang berdimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran yang didasarkan dengan teori optimasi [10] *Support Vector Machine* diperkenalkan pertama kali oleh Vapnik pada tahun 1992 sebagai rangkaian harmonis konsep konsep unggulan dalam bidang *pattern recognition* [11] Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan pada *svm* sangat bergantung pada fungsi kernel dan parameter yang digunakan. Berdasarkan dengan karakteristiknya metode *SVM* dibagi menjadi dua yaitu linear dan non linear, *SVM* linear merupakan data yang dipisahkan secara linear yaitu memisahkan dua kelas pada *hyperplane* dengan *soft margin*. Sedangkan non linear yaitu merupakan fungsi dari kernel trick terhadap ruang yang berdimensi tinggi [12]. *SVM* sangat cepat dan efektif untuk masalah klasifikasi teks, dalam istilah geometris sebuah klasifikasi biner dapat dilihat sebagai *hyperplane* dalam ruang fitur yang memisahkan titik-titik yang mewakili contoh positif dari kategori yang mewakili keadaan negative. Klasifikasi ini dipilih selama pelatihan sebagai *hyperplane* unik yang memisahkan *instance* positif yang diketahui dari *instance* negative, dalam klasifikasi *SVM* memiliki keunggulan penting dalam pendekatan teorinya yang dibenarkan atas masalah

overfitting yang memungkinkannya bekerja dengan baik [8].

Support Vector Machine menggunakan 2 titik (vector) yang selanjutnya dua titik tersebut akan membentuk garis pembatas (sisi pembatas jika 3 dimensi atau lebih) garis pembatas yang dibentuk dari dua buah vector ini disebut hyperplane.



Gambar 1 Hyperplane memisahkan dua kelas.

Dua titik yang menjadi patokan hyperplane disebut dengan support vector. Dapat dilihat bahwa memiliki dua kelompok data yang disebut klasifikasi, kemudian tugas SVM adalah membagi dua kelompok ini sebaik mungkin atau menentukan hyperplane terbaik, pembagian dimana garis batasnya dapat memisahkan dua kelompok dengan jarak terjauh antara titik terluar di masing masing kelompok dengan garis pembatas itu sendiri[13]. Permasalahan non linear dapat diatasi dengan memodifikasi trick kernel ke dalam SVM yang akan menjadi pemisah kelas atau hyperplane menjadi dua kelas didalam ruang vector dalam penelitian ini kernel yang akan digunakan adalah kernel linear. Seperti yang dapat dilihat persamaannya pada Tabel 1 di bawah ini.

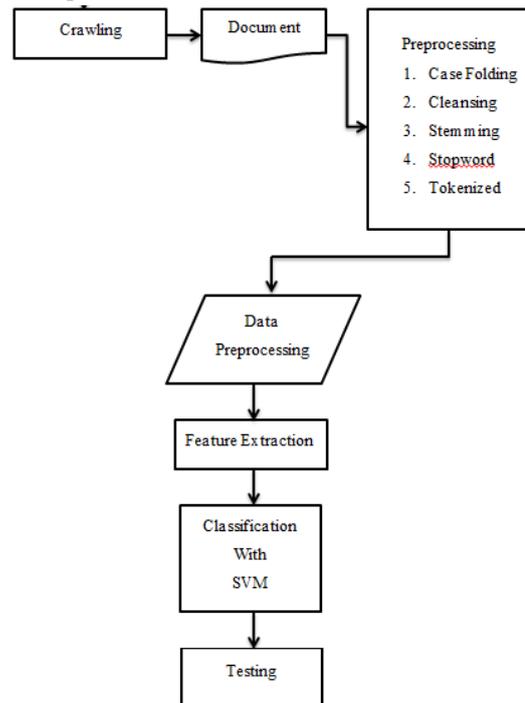
Tabel 1 Rumus Kernel

Jenis Kernel	Model
Linear	$K(x, x') = x \cdot x'$
Polynomial	$K(x, x') = (x \cdot x' + c)'$
RBF Gaussian	$K(x, x') = \exp(-\gamma   x - x'  ^2)$
Sigmoid	$K(x, x') = \tanh(\alpha x \cdot x' + \beta)$

### 3. Metode Penelitian

Pada Penelitian ini terdapat beberapa tahapan yang

dilakukan untuk menyelesaikan penelitian ini dapat dilihat pada Gambar 2.



Gambar 2 Tahapan Penelitian

### 3.1 Crawling

Crawling merupakan proses pengunduhan data yang berfungsi sebagai pengumpulan dataset dan menjadi tahap awal dalam penelitian ini. Data yang dikumpulkan berupa data tweet berbahasa Indonesia mengenai kebijakan lockdown atau pembatasan wilayah yang dilakukan oleh pemerintah Jakarta demi menangani Covid-19. Dari hasil pengambilan data tersebut yang berjumlah 2000 data yang selanjutnya perlu dilakukan pengolahan agar menjadi data yang lebih mudah digunakan dalam proses sentiment analisis.

Tabel 2 Hasil Crawling

Hasil Crawling
@diaraherlina : Jangan nunggu korban lebih banyak gerak sekarang, tegas dan melihat visi yang lebih jauh #LockdownJakarta pic.twitter.com
Kesal sama para pelanggar psbb !! <a href="https://t.co/NfhmfMjtXw">https://t.co/NfhmfMjtXw</a>
Haruskah Indonesia lockdown #lockdown
Sebenarnya psbb ini efektif gak sih #lockdown #psbb
Indonesia terserah
Dukung lockdown Jakarta !!

### 3.2 Preprocessing

Dalam Preprocessing data, kita perlu melakukan pembersihan data dengan tujuan supaya data digunakan dapat digunakan pada tahap selanjutnya. Adapaun

tahapan yang dilakukan pada penelitian ini adalah Menghilangkan teks html, *case folding*, *cleansing*, *stemming*, *stopword*, dan tokenisasi.

### 3.3 Seleksi Fitur

Pada tahap ini untuk meningkatkan kinerja *classifier* yang berguna untuk meningkatkan akurasi dan mengurangi waktu komputasi. Dalam hal ini fitur yang digunakan yaitu TF-IDF. Dapat dilihat pada Tabel 3 data training.

Tabel 3 Data Training

Dokumen	Tweet
D1	jangan nunggu korban lebih banyak gerak sekarang tegas lihat visi lebih jauh
D2	kesal sama para pelanggar psbb
D3	haruskah Indonesia lockdown
D4	sebenarnya psbb ini efektif gak sih
D5	indonesia terserah
D6	dukung lockdown Jakarta

Perhitungan TF-IDF

Tabel 4 Perhitungan TF-IDF

Term	TF						IDF
	D1	D2	D3	D4	D5	D6	
Korban	1	0	0	0	0	0	0.778
Banyak	1	0	0	0	0	0	0.778
Gerak	1	0	0	0	0	0	0.778
Tegas	1	0	0	0	0	0	0.778
Kesal	1	0	0	0	0	0	0.778
Para	1	0	0	0	0	0	0.778
Pelanggar	1	0	0	0	0	0	0.778
Psbb	0	1	0	1	0	0	0.477
Indonesia	0	1	0	0	1	0	0.477
Lockdown	0	0	1	0	0	1	0.477
Jakarta	0	0	0	0	0	1	0.778

Keterangan :

TF(D) = jumlah kata di setiap dokumen-D

IDF =  $\log(D/df)$

### 3.4 Training and Testing

Training merupakan tahapan pada algoritma SVM untuk menyesuaikan data yang telah melewati tahap *cleansing* dan *preprocessing* sehingga dapat mencapai set data yang telah ditentukan. *Training* akan digunakan untuk pembuatan model machine learning sementara *testing* berguna untuk menguji performa pada metode *Support Vector Machine*.

- Data training yang digunakan adalah data yang telah melakukan proses *labeling* secara manual oleh peneliti yang akan ditemukan nilai dari perkata, dan nilai ini yang akan digunakan sebagai data masukan untuk proses *training* dan *testing*.
- Data training yang digunakan adalah data yang telah melakukan proses *labeling* secara manual oleh peneliti yang akan ditemukan nilai dari perkata, dan

nilai ini yang akan digunakan sebagai data masukan untuk proses *training* dan *testing*.

### 3.5 Klasifikasi

Klasifikasi di penelitian ini menggunakan metode *Support Vector Machine* dengan seleksi fitur. Seleksi fitur dalam pengklasifikasian diharapkan dapat lebih efisien dengan mengurangi jumlah data yang di analisis dengan mengidentifikasi fitur yang selanjutnya akan diproses berdasarkan model *classifier* yang telah dihasilkan dari proses *training* pada penelitian inidata dibagi menjadi dua bagian yaitu data training dan data testing langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

- Menginisiasi nilai awal untuk nilai  $\alpha$ , C, epsilon gamma dan lamda  
 $\alpha = 0$ , C = 1, epsilon = 0,001, gamma = 0,5, lamda = 0,5
- Memasukkan data training (latih) berdasarkan kemunculan keyword dalam satu kalimat.
- Menentukan *dot product* setiap data dengan memasukkan fungsi kernel [K]. Rumus fungsi kernel yang umum sama pada Tabel 1. fungsi kernel yang digunakan adalah fungsi *linear*. Sebelumnya data di transpose karena menggunakan perkalian matriks  $A \times A^T$  seperti Tabel 5:

Tabel 5 Transpose Data

D1	D2	D3	D4	D5	D6
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	1	0	0
0	1	0	0	1	0
0	1	1	0	0	0
0	0	0	0	0	1

Pada metode kernel data tidak direpresentasikan secara individu, melainkan melalui perbandingan antara sepasang data. setiap data akan dibandingkan pada dirinya dan kata lain.

- Menghitung matriks dengan rumus :

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2)$$

$D_{ij}$  = elemen matriks ke i j

$y_i$  = kelas data ke-i

$y_j$  = kelas data ke-j

$\lambda$  = batas teoritis

Contoh perhitungan D1 dengan D1 :

$$D_{ij} = [1] [1] [8] + 0,5^2 = 5,25$$

- Mencari nilai error dengan rumus :

$$E_i = \sum_j^i = 1 \alpha_j D_{ij}$$

Dimana  $E_i$  = nilai error data ke-i

$\alpha_j = 0,5$

$$E_1 = (0,5 * 8,25) + (0,5 * 0,25) + (0,5 * 0,25) + (0,5 * 1,25) + (0,5 * 0,25) + (0,5 * 0,25) = 5,25$$

6. Menghitung nilai delta alpha  
 $\delta\alpha_i = \min\{\max[\gamma(1-E_i) - \alpha_i] C - \alpha_i$   
 data pertama :  
 $\text{Min}\{\max[0,5(1-5,75) - 0,5] 1 - 0,5\}$   
 $= -3,375$
7. Menghitung nilai alpha baru dengan rumus :

$$\alpha_i = \alpha_i + \delta\alpha_i$$

Contoh perhitungan pada D1 :  
 $\alpha_i = 0,5 + (-1,312) = -0.812$

8. Mencari nilai bias

$$b = -\frac{1}{2} (< w \cdot x-1 > + < w \cdot x+1 >)$$

terlebih dahulu mencari nilai w  
 $w_{x+} = (-0,812 \times (1) \times 8,25) + (-0,312 \times (1) \times 0,25) + (0,063 \times (-1) \times 0,25) + (0,188 \times (-1) \times 1,25) + (0,063 \times (-1) \times 0,25) + (0,188 \times (1) \times 0,25)$   
 $= -6.99$   
 $w_{x-} = (-0,812 \times (1) \times 0,25) + (-0,312 \times (1) \times 2,25) + (0,063 \times (-1) \times 1,25) + (0,188 \times (-1) \times 0,25) + (0,063 \times (-1) \times 0,25) + (0,188 \times (1) \times 0,25)$   
 $= -0.99$

Nilai b =  $-\frac{1}{2} (-6.9) + (-0.99) = 3.99$

9. Setelah mendapatkan nilai  $\alpha$ , w, dan b maka akan dilakukan pengujian  
 Langkah pertama untuk menguji yaitu menghitung dot product anatar data uji dengan semua data latih dengan fungsi kernel  
 $K(x, xi) = x \cdot xi$   
 Dimana x adalah data uji dan xi adalah data latih  
 Data ke -1 =  $(1x1) + (1x1) + (0x0) + (0x0) + (0x0) = 8$   
 Data ke -2 =  $(1x0) + (1x0) + (1x1) + (1x1) + (1x0) = 2$

### 3.6 Pengujian

Tahap pengujian digunakan untuk menghitung nilai akurasi, recall, precision. Didapatkan dari proses pelabellan secara manual dan hasil klasifikasi menggunakan metode *Support Vector Machine* pada proses *testing data*.  
 Pengujian dilakukan dengan menggunakan metode Confusion Matrix sebagai berikut :

Tabel 6 Confusion Matrix

No	Jenis	Model
1	Accuracy	$\frac{TP + TN}{TP + TN + FP + FN} * 100$
2	Precision	$\frac{TP}{FP + TP} * 100$

3	Recall	$\frac{TP}{TP + FN} * 100$
4	F1-Score	$2 * (\text{recall} * \text{precision}) / (\text{recall} + \text{precision})$

## 4. Implementasi

### 4.1 Crawling Data

Penelitian ini memakai data twitter yang dikumpulkan dalam satu file berformat CSV menggunakan library twint yang dijalankan melalui commant prompt yang selanjutnya diberi label sesuai dengan kelasnya dan data yang digunakan adalah tweet dengan kata kunci #Lockdownjakarta #Karantinawilayah #psbb.

```
twint -s "#lockdownjakarta" -o
lockdown.csv --limit 2000 --since
2020-04-01 --until 2020-05-31 --
csv --lang id
```

```
from Sastrawi.Stemmer.StemmerFactory
import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()
stemming = stemmer.stem(tweet)
```

### 4.2 Implementasi Preprocessing

```
import re
def process_tweet(tweet):
    tweet = tweet.lower()
    tweet = re.sub('@[\s]+', '', tweet)
    tweet = re.sub('((www\.[^\s]+)|(https?://[\s]+))',
'', tweet)
    tweet = re.sub(r'\d+', ' ', str(tweet))
    tweet = re.sub('&quot;', ' ', tweet)
    tweet = re.sub(r'\b[a-zA-Z]\b', '', str(tweet))
    for word in tweet.split():
        if word.lower() in contractions:
            tweet = tweet.replace(word,
contractions[word.lower()])
    tweet = re.sub(r'[\w\s]', ' ', str(tweet))
    tweet = re.sub(r'(\.|\!|\,)', ' ', str(tweet))
    tweet = re.sub(r'\s+', ' ', str(tweet))
return tweet
```

1. tweet = tweet.lower() berfungsi untuk transformasi huruf kapital menjadi huruf kecil
2. tweet = re.sub('@[\^s]+', '', tweet), Berfungsi untuk menghapus username
3. tweet = re.sub('((www\.[\^s]+)|(https?://[\^s]+))', '', tweet), Berfungsi untuk menghapus Url
4. tweet = re.sub(r"[d+", " ", str(tweet)), Berfungsi untuk menghapus digit
5. tweet = re.sub('&quot;', " ", tweet) , Berfungsi untuk menghapus (&quot;)
6. tweet = re.sub(r"[a-zA-Z]", "", str(tweet)), berfungsi menghapus karakter tunggal seperti a,b,c, dll
7. tweet = re.sub(r"[^\w\s]", " ", str(tweet)), menghapus tanda baca
8. tweet = re.sub(r'(\.|\!|\,)', ' ', tweet), Berfungsi untuk menghapus data ganda
9. tweet = re.sub(r"\s+", " ", str(tweet)), Berfungsi untuk mengubah 2 spasi atau lebih menjadi 1 spasi.

### 4.3 Implementasi TF-IDF

```
from sklearn.feature_extraction.text import TfidfVectorizer
final_vectorized_data=tf_idf_vectorizer.fit_transform(total_data['processed_tweet'])
final_vectorized_data
```

### 4.4 Implementasi Support Vector Machine

```
from sklearn.svm import SVC
model_svm = SVC(kernel = 'linear', random_state = 0)
model_svm.fit(X_train, y_train)
predicted_svm = model_svm.predict(X_test)
```

## 5. Hasil dan pembahasan

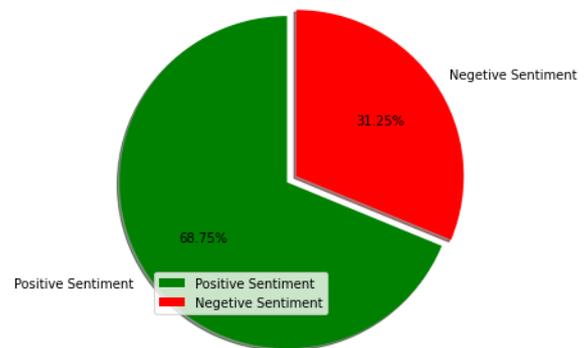
Data yang digunakan dalam penelitian ini diambil dari twitter.com menggunakan Twitterscrapper dan Twint sebanyak 1377 data tweet berbahasa Indonesia mengenai kebijakan lockdown Jakarta. Data tersebut dibagi menjadi dua kelas yaitu positif dan negatif yang berarti setuju atau tidak setuju dengan kebijakan lockdown yang dilakukan pemerintah Jakarta.

Tahap pengujian ini dilakukan dengan pengujian confusion matrix yang mengacu pada TP, FP, TN, FN untuk mengetahui seberapa baik akurasi metode svm dengan ekstraksi fitur *tf-idf* dalam kebijakan tersebut.

### 5.1 Perhitungan Kelas Sentimen

Dalam pengujian data sentimen tweet tentang kebijakan pemerintah mengenai kenijakan Lockdown didapatkan hasil sentimen negatif dan sentimen positif secara manual. Dari hasil pelabelan data tersebut

kemudian dilakukan *training data* dan *testing data* yaitu 80% *data Training* dan 20% *data testing* yang dilakukan secara acak. Hasil dari perhitungan data sentimen kelas Negatif berjumlah 275 data, dan sentimen kelas positif sebesar 1100 data. perbandingan dalam bentuk persentase dapat dilihat pada Gambar 3:



Gambar 3 Kelas Sentimen

bisa kita lihat perbandingan antara tweet masyarakat yang berkomentar positif mengenai kebijakan lockdown dengan tweet masyarakat yang berkomentar negatif yaitu 68,75 % tweet positif berbanding 31,25% tweet negatif dengan jumlah kelas Negatif sebanyak 275 data, dan sentimen kelas positif sebesar 1100 data dari hasil perhitungan sentimen didapat kata - kata yang memiliki kemunculan terbesar yang terdapat pada Gambar 4 :



Gambar 4 Kemunculan Kata Terbesar

### 5.2 Pengujian TF-IDF Terhadap Support Vector Machine

Pengujian yang digunakan yaitu dengan fitur *tf-idf* yang bertujuan mengetahui seberapa besar nilai *Accuracy*, *Precision*, *recall* dan *F1-Score* berdasarkan pengujian memiliki hasil yang ditunjukkan pada tabel 5.1.

Tabel 5.1 Hasil Pengujian Dengan TF-IDF

TF-IDF	Hasil (%)
Accuracy	74
Precision	75
Recall	92
F1-Score	83

Berdasarkan tabel diatas dapat dilihat bahwa hasil metode *Support Vector Machine* dengan fitur *tf-idf*

didapat dengan tinggak nilai *accuracy* sebesar 74%, *precision* sebesar 75%, *recall* sebesar 92%, dan *F1-Score* sebesar 83% dengan akurasi metode *Support Vector Machine* sebesar 74% yang artinya cukup baik dalam penelitian mengenai kebijakan lockdown pemerintah Jakarta.

### 6. Kesimpulan

Berdasarkan hasil penelitian dan pengujian metode *Support Vector Machine* untuk mengklasifikasi kebijakan lockdown pemerintah Jakarta dengan tweet berbahasa Indonesia menggunakan ekstraksi fitur *tf-idf* didapat kesimpulan bahwa penggunaan metode *Support Vector Machine* tingkat nilai *Accuracy* sebesar 74%, *Precision* sebesar 75%, *Recall* sebesar 92% dan *F1-Score* sebesar 83%. Yang artinya cukup baik dalam penelitian mengenai kebijakan lockdown pemerintah Jakarta.

### 7. Saran

Penelitian ini dapat dikembangkan dengan menggabungkan metode SVM dengan metode *Firefly* sebagai metode optimasi. Metode SVM dioptimasi dengan metode *Firefly* menghasilkan akurasi lebih tinggi dibandingkan dengan SVM tanpa optimasi[14].

### Daftar Pustaka

- [1] A. S. Cahyono, "Anang Sugeng Cahyono, Pengaruh Media Sosial Terhadap Perubahan Sosial Masyarakat di Indonesia," pp. 140–157, 2016.
- [2] I. Alfina, R. Mulia, M. I. Fanany, and Y. Ekanata, "Hate Speech Detection in the Indonesian Language : A Dataset and Preliminary Study," in *9th Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS 2017)*, 2017, no. October, doi: 10.1109/ICACSIS.2017.8355039.
- [3] S. Styawati and F. Ariany, "Sistem Monitoring Tumbuh Kembang Balita/Batita di Tengah Covid-19 Berbasis Mobile," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 490, 2021, doi: 10.32493/informatika.v5i4.7067.
- [4] G. A. Buntoro and A. E. Purnamasari, "Sentiment Analysis Candidates of Indonesian Presiden 2014 with Five Class Sentiment Analysis Candidates of Indonesian Presiden 2014 with Five Class Attribute," no. May, 2016, doi: 10.5120/ijca2016908288.
- [5] A. Fathan Hidayatullah and A. Sn, "ISSN: 1979-2328 UPN "Veteran," *Semin. Nas. Inform.*, vol. 2014, no. semnasIF, pp. 115–122, 2014.
- [6] J. Weng, E. Lim, J. Jiang, H. I. Search, and R. Information, "TwitterRank: Finding Topic-sensitive Influential Twitterers," pp. 261–270, 2010.
- [7] Monarizqa, Nugroho, and Hantono, "IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP," vol. 15, no. 2, pp. 171–176, 2019, doi: 10.33480/pilar.v15i2.699.
- [8] R. Feldman and J. Sanger, "Book Reviews," 2007.
- [9] J. Li and L. Qiu, "A Sentiment Analysis Method of Short Texts in Microblog," 2017, doi: 10.1109/CSE-EUC.2017.153.
- [10] D. E. Ratnawati and A. W. Widodo, "Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Machine," no. January, 2018.
- [11] E. Susilowati, M. K. Sabariah, A. A. Gozali, J. T. Informatika, U. Telkom, and S. V. Machine, "IMPLEMENTASI METODE SUPPORT VECTOR MACHINE UNTUK MELAKUKAN KLASIFIKASI KEMACETAN LALU LINTAS PADA TWITTER IMPLEMENTATION SUPPORT VECTOR MACHINE METHOD FOR TRAFFIC JAM CLASSIFICATION ON TWITTER," vol. 2, no. 1, pp. 1478–1484, 2015.
- [12] W. Purnami, A. M. Regresi, and L. Ordinal, "Perbandingan Klasifikasi Tingkat Keganasan Breast Cancer Dengan Menggunakan Regresi Logistik Ordinal Dan Support Vector Machine ( SVM )," vol. 1, no. 1, 2012.
- [13] A. Kowalczyk, *Support Vector Machines Succinctly*. SynCFusion Inc., 2017.
- [14] S. Styawati and K. Mustofa, "A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 13, no. 3, p. 219, 2019, doi: 10.22146/ijccs.41302.