

## **ANALYSIS OF SENTIMENT OF INDONESIAN COMMUNITY ON METAVERSE USING SUPPORT VECTOR MACHINE ALGORITHM**

Siti Sumayah<sup>\*1</sup>, Falentino Sembiring<sup>2</sup>, Wisuda Jatmiko<sup>3</sup>

<sup>1,2,3</sup> Jurusan Sistem Informasi, Universitas Nusa Putra Sukabumi, Indonesia  
Email: <sup>1</sup>[siti.sumayah\\_si18@nusaputra.ac.id](mailto:siti.sumayah_si18@nusaputra.ac.id), <sup>2</sup>[falentino.sembiring@nusaputra.ac.id](mailto:falentino.sembiring@nusaputra.ac.id),  
<sup>3</sup>[wisuda.jatmiko@nusaputra.ac.id](mailto:wisuda.jatmiko@nusaputra.ac.id)

(Naskah masuk: 25 Juni 2022, Revisi: 6 Juli 2022, Diterbitkan: 10 Februari 2023)

### **Abstract**

Metaverse was widely discussed right on October 28, 2021, when Mark Zuckerberg announced the company name change from Facebook to Meta. Regarding the sophisticated Metaverse concept, it is possible to make the pros and cons between the Indonesian people. This information is widely distributed in various media crews, one of which is Twitter social media. Tweets from social media can be used to obtain data and information as material for research on sentiment analysis. This study aims to determine the response of the Indonesian people to the Metaverse. Material data was obtained from Twitter tweets with keywords or Metaverse queries. Research conducted by the Support Vector Machine Algorithm in research with TF-IDF word weighting, the tests carried out with the results of the class division of positive sentiment 70.69%, neutral 15.85%, negative 13.46% will produce values of accuracy, precision, recall, and f1-score. They use various data, namely, 90% training data and 10% test data. The result is high accuracy in testing many data, 2504 data, and the accuracy value is 81%, with an average value of 79% precision, 63% recall, and 57% f1-score.

**Keywords:** Metaverse, Support Vector Machine, TF-IDF, Twitter

## **ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP METAVERSE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

### **Abstrak**

Metaverse ramai diperbincangkan tepat pada 28 oktober 2021 ketika Mark Zuckerberg mengumumkan pergantian nama perusahaan facebook menjadi meta. Mengenai konsep Metaverse yang canggih, tidak menutup kemungkinan menjadikan sebuah pro dan kontra antara masyarakat Indonesia. Informasi tersebut tersebar luas di berbagai awak media, salah satunya media sosial twitter. cuitan dari media sosial dapat digunakan untuk memperoleh berupa data maupun informasi sebagai bahan untuk penelitian pada sentimen analisis. Penelitian bertujuan untuk mengetahui mengenai tanggapan masyarakat Indonesia terhadap Metaverse. Bahan data yang diperoleh dari cuitan twitter dengan kata kunci atau query Metaverse. penelitian yang dilakukan Algoritma Support Vector Machine pada penelitian dengan pembobotan kata TF-IDF, pengujian yang dilakukan dengan hasil pembagian kelas sentimen positif 70.69%, netral 15.85%, negatif 13.46% akan menghasilkan nilai akurasi, precision, recall, dan f1-score. Menggunakan variatif terhadap data yaitu, data latih 90% dan data uji 10%. Dihilangkan Akurasi tinggi berada pada pengujian banyak data 2504 data, dan nilai akurasi 81%, dengan nilai rata-rata pada precision 79%, recall 63%, f1-score 57%.

**Kata kunci:** Metaverse, Support Vector Machine, TF-IDF, Twitter

### **1. PENDAHULUAN**

Kemajuan pada teknologi komunikasi dan informasi sangat pesat dan berperan penting bagi kehidupan manusia dalam memenuhi kebutuhan yang harus dipenuhi. Kemajuan teknologi yang memanfaatkan internet berdampak terhadap pola hidup manusia, seperti cara berinteraksi dengan sesama, bertukar informasi, hingga mencari informasi yang sedang ramai dibicarakan

masyarakat. Dengan adanya keterbatasan media konvensional seperti, radio, koran, tv. Untuk mengutarakan pendapat. Tidak menutup media sosial jadi pengganti atau alternatif yang tepat untuk berekspresi [1].

Media sosial twitter dapat digunakan untuk mendapatkan data dan informasi sebagai bahan untuk penelitian sentimen analisis, yang diambil dari cuitan twitter. Dari cuitan (tweet) dapat dijadikan bahan data

untuk penelitian sentimen. Seperti kabar *Metaverse* yang sedang ramai setelah Mark Zuckerberg mengumumkan mengenai perubahan nama induk perusahaan facebook menjadi Meta Platforms Inc atau meta Tepat pada 28 Oktober 2021 [2]. Mark Zuckerberg sebagai salah satu pencetus dalam pengembangan teknologi *Metaverse* setelah Neal Stephenson yang menciptakan Istilah *Metaverse* dalam Snow Crash, novel bergenre cyberpunk yang diterbitkan pada 1992 [3]. Digadag-gadag *Metaverse* merupakan dunia yang melampaui semesta secara virtual yang saling terhubung. pengguna akan direpresentasikan menjadi masuk kedalam ruang virtual melalui avatar, karena avatar dapat menjadi representasi pengguna dengan apa yang diinginkan dengan dunia virtual untuk berinteraksi [3]. seperti main *game*, *meeting*, transaksi ekonomi, wisata, konser dan masih banyak aktifitas lainnya.

Topik perbincangan masyarakat Indonesia pada saat ini mengenai *Metaverse* yang digadag-gadag merupakan dunia virtual sangat canggih dan cukup menguntungkan bagi para pelaku bisnis, investor dan masyarakat biasa pada umumnya seperti halnya perusahaan yang mendirikan metaverse dengan versinya masing-masing [4]. Mengenai konsep *Metaverse* yang sangat canggih, tidak menutup kemungkinan menjadikan sebuah pro dan kontra antara masyarakat Indonesia mengenai *Metaverse*.

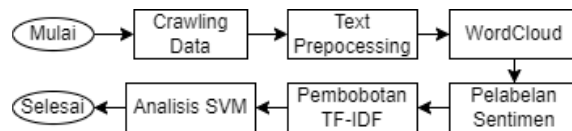
Tujuan penelitian yaitu untuk mengetahui tanggapan atau opini masyarakat indonesia mengenai metaverse dan membantu para investor, pelaku usaha, dan masyarakat pada umumnya dalam melihat pasar metaverse. Maka dari itu peneliti akan melakukan penelitian analisis sentimen untuk mengetahui tanggapan masyarakat Indonesia terhadap metaverse, apakah masyarakat Indonesia lebih banyak memberikan pendapat bersifat negatif, netral atau positif. Karena analisis sentimen yaitu bidang studi atau metode untuk menganalisis tanggapan atau opini [5]. Mengingat data yang diperoleh tidak sedikit maka bantuan klasifikasi algoritma *Support Vector Machine* kernel *Linear*. *Dataset* tweet akan diproses atau diolah untuk menghasilkan klasifikasi untuk mengukur akurasi mengenai *dataset Metaverse* pada tweet media sosial twitter.

Beberapa penelitian yang membahas Analisis Sentimen memakai Algoritma *Support Vector Machine* banyak digunakan. Menurut Penelitian Aulia Rahman Mengenai opini tentang lockdown di jakarta dari twitter, memakai Algoritma SVM dan pembobotan kata menggunakan TF-IDF Hasil sentimen analisis terhadap 2000 dat dengan hasil nilai 74% accuracy, 75% precision, 92% recall dan 83% *F1-Score* [6]. Penelitian Dedi Darwis dalam Analisis Sentimen data twitter tentang pemberantasan korupsi

di Indonesia menggunakan Algoritma SVM dan pembobotan TF-IDF sebagai metodenya. Dengan data 1890 dan 3846 term nilai Akurasi 82%, *precision* 90%, *Recall* 88%, *f1-score* 89% [7]. penelitian Rian Tineger mengenai Analisis Sentimen mengenai Layanan pada Indihome dari Twitter menggunakan algoritma Klasifikasi SVM dengan 280 data tweet dengan evaluasi *Confusion Matrix*, menghasilkan 87% akurasi , 86% *Precision*, 95% *Recall* , 90% *f1-score* [8]. Penelitian Amar P. Natasuwarna mengenai Analisis Sentimen keberlanjutan belajar daring dengan 200 data tweet memakai metode SVM dan pemilihan Fitur TF-IDF evaluasi tertinggi dihasilkan pada 8-Fold Cross Validation dengan hasil accuracy 86,00%, precision,38%, dan recall 85,02% [9]. penelitian ini merupakan penelitian terkait mengenai Analisis Sentimen menggunakan Algoritma SVM dengan pembobotan kata menggunakan TF-IDF.

## 2. METODE PENELITIAN

Menjelaskan tahapan-tahapan yang digunakan penulis dalam melakukan penelitian. Adapun alur metode penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Metode Penelitian

metode penelitian merupakan sebuah alur untuk melakukan tahapan dalam penelitian yang akan dilakukan seperti yang ditunjukkan pada gambar 1.

### 2.1 Crawling

*Crawling* atau pengumpulan data yaitu sebagai pengumpulan data untuk dataset [6]. Data yang diperoleh yaitu dari cuitan Twitter dengan *query Metaverse*, menggunakan Twitter yang terkoneksi dengan Twitter Application Programming Interface (API), untuk mendapatkan akses diperlukan *keys* dan *token*. Untuk proses *crawling* data memakai python sebagai bahasa pemrograman dengan bantuan *library* yaitu *tweepy*.

### 2.2 Text Preprocessing

*Text Preprocessing* yaitu proses untuk meningkatkan kualitas pada teks atau penyeleksian data text agar menghilangkan *noise* [10]. data menjadi lebih terstruktur dan seragam dengan melalui tahapan proses sebagai berikut :

- Cleaning*, untuk menghilangkan tanda baca, *username*, *link*, *hashtag* dan tanda baca.
- Case Folding*, yaitu untuk menyeragamkan kata yang terdapat pada *dataset* untuk menjadi huruf kecil.
- Stemming*, dalam penelitian untuk mengganti kalimat dalam *dataset* menjadi kata dasar

- dengan meniadakan kata tambahan atau imbuhan.
- d. *Tokenizing*, dalam penelitian berfungsi dalam pemotongan pada kalimat yang berada pada *dataset* jadi perkata.
- e. *Text Normalization*, yaitu untuk mengganti kata singkat jadi kata baku.
- f. *Stopwords*, meniadakan kata yang tidak berfungsi tapi seringkali terdapat pada tweet.

### 2.3 Word Cloud

*Word Cloud* yaitu bentuk visual dari keberadaan kata. makin banyak istilah yang sering muncul dalam dokumen yang dianalisis, maka dari itu kata yang muncul pada gambar tentu semakin besar [11]. *Word Cloud* akan diproses setelah tahapan preprocessing selesai.

### 2.4 Labeling Sentimen

*Labeling Sentimen* pada tahap ini Bertujuan untuk mengetahui pendapat masyarakat Indonesia mengenai *Metaverse* pelabelan atau *Labeling* yaitu proses pemberian terhadap kelas berdasarkan dari ciri atau karakteristik yang terkandung dalam dokumen ataupun kalimat [12]. penelitian pada pelabelan sentimen dilakukan dengan membagi kelas menjadi 3 kelas Sentimen yaitu Positif, Netral, Negatif. apabila kalimat bernilai >0 berada dalam klasifikasi kelas positif, apabila kalimat bernilai == 0 berada dalam klasifikasi kelas netral, lalu apabila kalimat bernilai < 0 berada dalam klasifikasi kelas negatif. Penulis membuat kamus positif begitupun juga negatif dengan bahasa Indonesia dan membuat sistem menggunakan pemrograman python. Maka ketika kata yang ada pada dokumen mengandung kata yang ada dikamus positif maka dalam sistem akan terdeteksi positif, begitupun dengan kata negatif lain dengan netral, netral akan terdeteksi jika kata yang ada dalam dokumen tidak ada dalam kamus positif dan negatif maka akan terdeteksi netral.

### 2.5 Pembobotan TF-IDF

Tahap kelima Pembobotan terhadap *Term Frequency-Inverse Document Frequency* (TF-IDF), yakni gabungan dua proses dari TF dan IDF. TF-IDF digunakan saat ingin merubah atau transformasi data teks menjadi vektor (angka).TF untuk mengetahui jumlah pada suatu kata yang sering muncul pada kalimat atau bahasan, IDF mengukur seberapa penting suatu kata di dokumen. Berikut merupakan rumus pembobotan kata TF-IDF.

$$idf_t = \log \frac{N}{df_t} \tag{1}$$

$$w_{t,d} = tf_{t,d} \times idf_t \tag{2} \text{ [13].}$$

### 2.6 Klasifikasi Support Vector Machine

Klasifikasi SVM merupakan algoritma klasifikasi terhadap data linear maupun non linear, terutama pada masalah nonlinear dengan menggunakan konsep kernel ke dalam ruang berdimensi yang lebih tinggi. Klasifikasi (classification) digunakan pada data baru untuk mengelompokkan jenis objek. Klasifikasi termasuk pada model supervised. Pada persoalan klasifikasi sampel data yang dimiliki dan memprediksi beberapa kelas yang ada berdasarkan sampel yang ada [14]. pada metode kernel untuk klasifikasi pada data yang tidak terklasifikasi dengan linear, kernel liner yang paling sering digunakan untuk klasifikasi karena fungsi kernelnya paling sederhana dan kernel linear bisa digunakan pada klasifikasi teks [15]. Dengan menggunakan evaluasi *Confusion Matrix* adalah tahap pengujian yaitu untuk menghitung dan dapat digunakan untuk menghasilkan matrik evaluasi seperti akurasi, *precision*, *Recall* dan F1-score dihasilkan semenjak proses pelabelan dan klasifikasi *Support Vector Machine* pada proses testing data [16].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Crawling

Setelah data tweet sudah diperoleh dengan teknik *crawling* data sebanyak 3096 data tweet. terlihat pada gambar 2.

	user	tanggal	tweet
0	Darmin A. Pella	11/03/2022 23.37	20 Desain teknologi tersebut membuat tubuh dan...
1	LinAprilia	14/03/2022 04.40	Tantangan besar berikutnya u/ dekade mendatang...
2	aA@alettAa	15/03/2022 03.41	Kayaknya lu yang gak tau konsep grup aespa kay...
3	AKSAN NUGROHO	15/03/2022 03.31	Kalau kota-kota lain merasa ketinggalan saat i...
4	Muslimin B. Putra	15/03/2022 03.09	Indeks SPBE (Sistem Pemerintahan Berbasis Elek...
...	...	...	...
95	Darmin A. Pella	11/03/2022 23.33	2 Dunia virtual immersif bermakna seseorang mer...
96	Darmin A. Pella	11/03/2022 23.33	1 Kita mundur sedikit. Apa itu metaverse? Meta...
97	AKSAN NUGROHO	11/03/2022 16.48	@SupirPete2 "Menurutnya jika seseorang tidak p...
98	sahretech_offical	11/03/2022 16.47	Apa itu metaverse. Cara Kerja, Kelebihan dan K...
99	Kang sariawan	11/03/2022 16.38	@memefess Dunia sudah menyambut kehadiran meta...

Gambar 2. Hasil *Crawling* Data

Berikut adalah hasil data tweet yang sudah diperoleh dari media sosial dengan menggunakan teknik *crawling* yang ditunjukkan pada gambar 2.

### 3.2.1 Text Preprocessing

*Preprocessing* dilakukan setelah *crawling* data terlihat pada gambar 2. hasil *preprocessing* pada tabel 1.

Tabel 1. Hasil *Text Preprocessing*

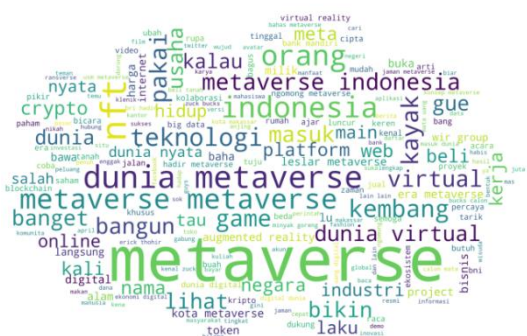
Alur Text Preprocessing	Sebelum	Sesudah
<i>Cleaning</i>	Apakah metaverse itu buruk? Tidak sepenuhnya. Karena metaverse akan membuat dunia pendidikan akan	Apakah metaverse itu buruk Tidak sepenuhnya Karena metaverse akan membuat dunia pendidikan akan

	semakin canggih. https://t.co/E9hAXZfvfj #Metaverse #Church #God #Jawabancom	semakin canggih Metaverse Church God Jawabancom
Case Folding	belajar tentang METAVERSE emg paling the best	belajar tentang metaverse emg paling the best
Stemming	belajar tentang metaverse emg paling the best	ajar tentang metaverse emg paling the best
Tokenizing	ajar tentang metaverse emg paling the best	[ajar,tentang,metaverse ,emg,paling,the,best]
Text Normalization	ajar tentang metaverse emg paling the best	ajar tentang metaverse memang paling the best
Stopwords	apakah metaverse itu buruk tidak sepuh karena metaverse akan buat dunia didik akanmakin canggih metaverse church god awabancom	metaverse buruk sepuh metaverse dunia didik canggih metaverse church god jawabancom

Terlihat hasil dari *Text preprocessing* pada tabel 1. *Dataset* yang sudah melalui tahap *Preprocessing* data sudah seragam, terstruktur dan *noise* pada teks sudah hilang, agar pada saat masuk tahap klasifikasi lebih optimal dalam perhitungan [17].

3.3 Word Cloud

Visualisasi *Word Cloud* berfungsi untuk visualisasi pada *dataset*, untuk mengetahui data yang sering muncul dalam dokumen. Proses *WordCloud* memakai python sebagai bahasa pemrograman dengan bantuan *library* matplotlib. Berikut visualisasi *Word Cloud* dapat dilihat pada gambar 3.



Gambar 3. Visualisasi *Word Cloud*

Dari hasil *Word Cloud* pada gambar 3. Bisa dilihat kata yang berukuran besar, artinya kata yang sering muncul

3.4 Labeling Sentimen

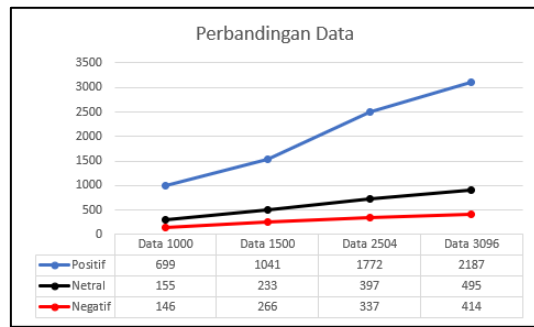
*labeling* sentimen menggunakan bahasa pemrograman python dengan membuat kamus negatif dan positif bahasa Indonesia dan membuat sistem. Berikut hasil *labeling* pada tabel 2.

Tabel 2. Hasil *Labeling*

No	Tweet	Label	Value
1	dakwah aji laku virtual metaverse	positif	2
2	metaverse besok kaya ya	positif	1
3	indonesia kerja tip meta kembang metaverse	positif	2
4	Polisi metaverse	netral	0
5	metaverse jelek	negatif	-1
6	jowo indonesia nfts indonesia metaverse	netral	0
7	harga kripto bitcoin ethereum anjlok uang kripto metaverse	negatif	-1

Proses pada *labeling* bertujuan untuk membagi kelas menjadi 3 bagian kelas yaitu kelas negatif, netral, positif [18]. Dimana dari setiap kalimat akan mengandung *value* dari setiap kata yang mengandung kelas sentimen seperti yang ada pada tabel 2.

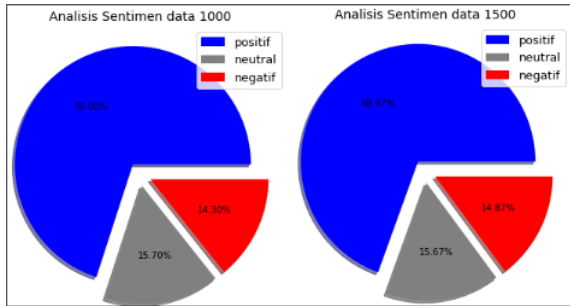
Adapun jumlah perbandingan data labeling dari banyak data yang berbeda terdapat dalam gambar 3.



Gambar 4. Perbandingan Data

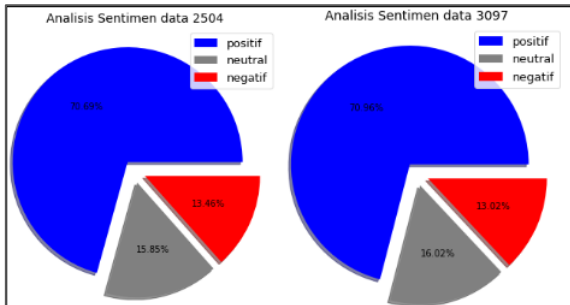
Pada gambar 4. Diagram garis yang digunakan berfungsi untuk menyajikan atau menampilkan dalam perkembangan suatu informasi [19]. Terlihat bahwa kelas yang lebih unggul pada penelitian sentimen analisis *Metaverse* dengan banyak data yang berbeda adalah kelas sentimen positif, peringkat kedua yaitu kelas netral dan ketiga kelas negatif.

Adapun persentase dari hasil labeling dari setiap banyak data pada gambar 5.



Gambar 5. Persentase Hasil Labeling

Terlihat pada gambar 5. Nilai persentasi dari setiap banyak data yang berbeda memiliki tingkat persentasi yang berbeda, dan persentasi yang paling unggul dari banyak data 1000 dan 1500, yaitu berada pada kelas sentimen positif. Lalu dapat dilihat nilai persentasi dengan banyak data yang berbeda pula pada gambar 6.



Gambar 6. Persentase Hasil Labeling

Terlihat pada gambar 6. Nilai persentasi dari banyak data 2504 dan 3096, kelas sentimen positif yang unggul.

Pada gambar 5 dan gambar 6. Di buatkan diagram pie chart fungsinya untuk perbandingan komponen yang menunjukkan persentase atau ukuran dari data set yang dimiliki [20]. Seperti halnya membaca data kelas sentimen *Metaverse*

### 3.4 Pembobotan TF-IDF

Pembobotan kata mengimplementasikan TF-IDF. Yaitu dengan dihitung banyak kata muncul dalam dokumen, dan tahapan pembobotan pada kata dilakukan setelah *preprocessing*. TF-IDF untuk memberi nilai pada *term* lalu nilai *term* untuk input pada proses klasifikasi SVM. Berikut proses pembobotan TF-IDF python sebagai dengan bahasa pemrograman dibantu oleh library Scikit learn, *TfidfVectorizer* pada gambar 6.

```
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf_vect = TfidfVectorizer()
tfidf_vect.fit(df['Tweet'])
train_X_tfidf = tfidf_vect.transform(df_train['Tweet'])
test_X_tfidf = tfidf_vect.transform(df_test['Tweet'])
tfidf_vect |
```

Gambar 7. Script TF-IDF

Terlihat pada gambar 7. Proses pengerjaan *TfidfVectorizer* merupakan proses transformasi

dokumen berbentuk teks menjadi *vector*. Adapun hasil dari pembobotan sebagai berikut.

Tabel 3. Hasil dari TF-IDF

Urutan pada Kata	TF-IDF
(0, 522)	0.20564267326870783
(0, 451)	0.23578017163840045
(0, 315)	0.289278630931714
(1, 4669)	0.2639725496518336
(1, 4372)	0.2782973972241244
(1, 4302)	0.26056160962469144

Seperti yang ada pada Tabel 3. Bahwa angka 0 pada kolom urutan kata (0, 522) yaitu merupakan kalimat atau tweet yang pertama, lalu 522 merupakan kata yang berada dalam urutan 522 dalam urutan pada penggabungan. Adapun penjelasan yang berada di kolom hasil TF-IDF yaitu 0.20564267326870783 merupakan nilai pembobotan terhadap kata dengan TF-IDF.

### 3.5 Klasifikasi Support Vector Machine

Klasifikasi terhadap penelitian yang dilakukan, data yang didapat dari tweet mengenai *Metaverse* pada media sosial twitter. Diperoleh data sebanyak 3096 data. Tahap klasifikasi SVM ini sudah melalui tahap, *preprocessing*, *Labeling* sentimen, pembobotan terhadap kata menggunakan TF-IDF, lalu data yang akan memasuki pada tahap klasifikasi dibagi 2 atau *split* data menjadi data latih serta data uji. Dan algoritma yang dipakai yakni SVM dengan metode kernel *linear*. Begitupun persentase dari *accuracy* yang didapat dari nilai pelabelan sentimen, dari tiap label mempunyai nilai yaitu nilai bobot yang dihitung. Lalu dilakukannya evaluasi *confusion matrix multiple class*, dalam memperoleh nilai terhadap akurasi, *precision*, *recall* dan *F1-score*.

Data yang diproses pengujian klasifikasi menggunakan *split* data, yaitu 90% pada data latih dan 10% pada data uji dengan banyak data berbeda. Adapun tabel banyak data yang berbeda dengan split data sebagai berikut.

Tabel 4. Perbandingan Split Data

Banyak Data	Data Latih 90%	Data Uji 10%
1000	900	100
1500	1350	150
2504	2253	251
3096	2786	310

Terlihat tabel 4. Data latih berfungsi untuk melatih pada algoritma dan data uji data Pengujian pada penelitian akan dikerjakan dengan 4 kali mengetahui performa pada algoritma yang dipakai [21]. Pengujian pada banyak data berbeda dengan split data sebanyak 90% pada data latih dan 10% pada data uji. Adapun *script* untuk proses Algoritma SVM kernel *linear* pada pengerjaan klasifikasi dan dilakukan evaluasi *confusion matrix* terlihat pada gambar 8.

```

from sklearn.svm import SVC
from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
model = SVC(kernel='linear')
model.fit(train_X_tfidf, train_Y)

predictions_SVM = model.predict(test_X_tfidf)
test_prediction = pd.DataFrame()
test_prediction['Tweet'] = test_X
test_prediction['Label'] = predictions_SVM

# Accuracy, Precision, Recall, f1-score
print (confusion_matrix(test_Y, predictions_SVM))
print (classification_report(test_Y, predictions_SVM))
print ("nilai akurasi dengan SVM adalah :", accuracy_score(test_Y, predictions_SVM)*100)
print ("SVM Recall Score :", recall_score(test_Y, predictions_SVM, average='macro')*100)
print ("SVM Precision Score :", precision_score(test_Y, predictions_SVM, average='macro')*100)
print ("SVM F1 Score :", f1_score(test_Y, predictions_SVM, average='macro')*100)
    
```

Gambar 8. Script pada Klasifikasi SVM

Terlihat pada gambar 7. Peneliti melakukan klasifikasi algoritma SVM menggunakan python yang dimana hasil dari *script* tersebut menghasilkan akurasi dari pengujian banyak data yang digunakan peneliti. Hasil dari klasifikasi yang sudah dilakukan dengan melalui tahapan-tahapan yang sudah dilalui. Berikut hasilnya pada tabel 5.

Tabel 5. Hasil Klasifikasi SVM

Data 1000				
Accuracy	kelas Pada Sentimen	precision	Recall	f1-score
74%	Negatif	1.00	0.13	0.23
	Neutral	1.00	0.40	0.57
	Positif	0.72	1.00	0.84
	Rata-rata	90%	57%	63%
Data 1500				
74%	Negatif	0.60	0.13	0.21
	Neutral	0.67	0.19	0.30
	Positif	0.75	0.98	0.85
	Rata-rata	67%	43%	45%
Data 2504				
81%	Negatif	0.83	0.36	0.51
	Neutral	0.70	0.39	0.50
	Positif	0.82	0.98	0.89
	Rata-rata	79%	63%	57%
Data 3096				
80%	Negatif	0.71	0.38	0.49
	Neutral	0.71	0.40	0.51
	Positif	0.82	0.97	0.89
	Rata-rata	74%	58%	63%

Dari hasil pengujian yang sudah peneliti lakukan, yaitu bertujuan untuk mencari terhadap nilai pada akurasi begitupun nilai rata-rata pada *Precision*, *Recall* dan *F1-Score*, dan hasilnya ditemukan variasi nilai. Nilai pada akurasi yang tinggi terdapat pada pengujian klasifikasi dengan banyak data 2504 menghasilkan akurasi sebanyak 81%. nilai rata-rata pada setiap kelas sentimen 79% *Precision*, 63% *Recall*, 57% *F1-Score*, dengan label sentimen positif

70.69%, netral 15.85%, negatif 13.46%. menunjukkan bahwa pada penelitian yang dilakukan jika banyak kata yang terdeteksi dalam kelas sentimen positif dan negatif maka nilai akurasi semakin tinggi, dan jika data pada label sentimen netral semakin kecil maka atau semakin sedikit maka nilai akurasi akan semakin kecil. Untuk menghasilkan akurasi yang tinggi tergantung variasi data yang didapat.

#### 4. DISKUSI

Dari empat penelitian terkait yang sudah dipaparkan pada bagian pendahuluan, peneliti akan melakukan penelitian atau pengamatan yang sama, namun dengan kasus atau objek penelitian yang berbeda dan pengujian dengan banyak data yang berbeda pada algoritma yang dipilih, dengan tujuan sebagai perbandingan.

Para peneliti melakukan penelitian analisis sentimen untuk mengetahui opini atau tanggapan seseorang terhadap topik tertentu, dengan melakukan klasifikasi algoritma untuk mengetahui nilai akurasi mengenai data topik yang dipilih, untuk mengetahui opini atau tanggapan masyarakat Indonesia terutama pada topik *Metaverse* dari tweet media sosial twitter, peneliti melakukan penelitian analisis sentimen menggunakan algoritma *Support Vector Machine* kernel *linear* dan pembobotan kata menggunakan *TF-IDF*.

Analisis sentimen membantu untuk mengetahui tanggapan atau opini masyarakat indonesia mengenai *metaverse* dan membantu para investor, pelaku usaha, dan masyarakat pada umumnya dalam melihat pasar *metaverse*. bisa dibuktikan dengan persentase kelas sentimen yang ada pada gambar 5 dan 6. Dari 4 kali pengujian banyak data yang berbeda, persentase kelas sentimen positif selalu unggul. Artinya bahwa masyarakat Indonesia menyambut baik adanya *metaverse* dengan memberikan tanggapan positif. Dan penelitian analisis sentimen ini, menghasilkan nilai akurasi yang cukup baik dari pengujian algoritma SVM kernel linear, menggunakan pembobotan kata *TF-IDF* dengan evaluasi *Confusion Matrix* yaitu tahap pengujian untuk mendapatkan nilai *precision*, *Recall* dan *F1-score*, untuk lebih jelasnya dapat dilihat pada tabel 5.

#### 5. KESIMPULAN

Kesimpulan dari penelitian yang sudah dilakukan dapat disimpulkan sebagai berikut :

1. *TF-IDF* membantu Algoritma SVM kernel *linear* pada pembobotan pada kata dalam analisis terhadap teks.
2. Pengujian dikerjakan sebanyak 4 kali dengan banyak data yang berbeda, dan hasil akurasi tinggi dihasilkan pada banyak data 2504 dengan akurasi 81% dengan nilai rata-rata pada setiap kelas sentimen 79% *Precision*, 63% *Recall*, 57% *F1-Score*. Dengan jumlah sentimen tweet positif 1772, netral 397, negatif 337.

3. *Text Preprocessing* berpengaruh terhadap klasifikasi untuk meningkatkan akurasi.

#### DAFTAR PUSTAKA

- [1] w. purbo Onno, *Text mining analisis medsos kekuatan brand & intelegen di internet*, 1st ed. yogyakarta: andi (anggota IKAPI), 2019.
- [2] J. López-Díez, “Metaverse: Year One. Mark Zuckerberg’s video keynote on Meta (October 2021) in the context of previous and prospective studies on metaverses,” *Pensar Public.*, vol. 15, pp. 299–303, 2021.
- [3] K. Laeq and E. Sciences, “Metaverse : Why , How and What,” 2022. [https://www.researchgate.net/publication/358505001\\_Metaverse\\_Why\\_How\\_and\\_What](https://www.researchgate.net/publication/358505001_Metaverse_Why_How_and_What) (accessed Jun. 18, 2022).
- [4] Indrawan Nugroho, “Apa itu Corporate Metaverse?,” *Corporate Innovation Consulting*, 2022. <https://www.cias.co/post/apa-itu-corporate-metaverse> (accessed Jun. 06, 2022).
- [5] A. Muzaki and A. Witanti, “Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier,” *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, 2021, doi: 10.20884/1.jutif.2021.2.2.51.
- [6] A. R. Isnain *et al.*, “SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM,” *Univ. Teknokr. Indones.*, vol. 2, no. 1, pp. 31–37, 2021.
- [7] D. Darwis, E. S. Pratiwi, and A. F. O. Pasaribu, “Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia,” *J. Ilm. Educic Pendidik. dan Inform.*, vol. 7, no. 1, pp. 1–11, 2020.
- [8] R. Tineges, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM),” *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [9] A. P. Natasuwarna, “Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring,” *Techno.Com*, vol. 19, no. 4, pp. 437–448, 2020, doi: 10.33633/tc.v19i4.4044.
- [10] I. M. Agus, W. Putra, A. Susanto, and I. Soesanti, “Ekstraksi Garis Pantai Pada Citra Satelit Landsat dengan Metode Segmentasi dan Deteksi Tepi,” *J. Nas. Pendidik. Tek. Inform.*, vol. 4, pp. 115–120, 2015.
- [11] A. R. Widangsa and A. R. Pratama, “Analisis Sentimen Kebijakan Pendidikan di Masa Pandemi COVID-19 dengan CrowdTangle di Facebook,” *AUTOMATA*, vol. 2, no. 2, 2021.
- [12] M. Priandi and Painem, “Analisis Sentimen Masyarakat Terhadap Pembelajaran Daring di Era Pandemi Covid-19 pada Media Sosial Twitter Menggunakan Ekstraksi Fitur Countvectorizer dan Algoritma K-Nearest Neighbor,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl. Jakarta-Indonesia*, no. September, pp. 311–319, 2021.
- [13] Y. T. Pratama, F. A. Bachtiar, and N. Y. Setiawan, “PARIWISATA PANTAI MALANG SELATAN MENGGUNAKAN TF-IDF DAN SUPPORT VECTOR MACHINE SKRIPSI memperoleh gelar Sarjana Komputer Disusun oleh : Yoga Tika Pratama,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, pp. 6244–6252, 2018.
- [14] N. L. W. S. R. Ginantra *et al.*, *Data mining dan penerapan algoritma*. Yayasan Kita Menulis, 2021.
- [15] M. C. Kirana, N. P. Perkasa, M. Z. Lubis, and M. Fani, “Visualisasi Kualitas Penyebaran Informasi Gempa Bumi di Indonesia Menggunakan Twitter,” *J. Appl. Informatics Comput.*, vol. 3, no. 1, pp. 23–32, 2019, doi: 10.30871/jaic.v0i0.1246.
- [16] A. Gupta, P. Tyagi, T. Choudhury, and M. Shamoan, “Sentiment Analysis Using Support Vector Machine,” *Proc. 4th Int. Conf. Contemp. Comput. Informatics, IC3I 2019*, pp. 49–53, 2019, doi: 10.1109/IC3I46837.2019.9055645.
- [17] S. Gusriani, K. D. K. Wardhani, and M. I. Zul, “Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naive Bayes (Studi Kasus: Facebook Page BerryBenka),” *4th Appl. Bus. Eng. Conf.*, vol. 1, no. 1, pp. 1–7, 2016.
- [18] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, “Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification,” *Sistemasi*, vol. 10, no. 1, p. 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [19] N. Kirana, “Jangan Sampai Salah! Gunakan Grafik Sesuai Fungsinya,” 2021. <http://www.thetastatistik.com/jangan-sampai-salah-gunakan-grafik-sesuai-fungsinya/#:~:text=Grafik Garis,turun dalam kurun waktu tertentu.> (accessed Jul. 02, 2022).
- [20] L. X. Sistem Informasi, “9 JENIS CHART YANG DAPAT ANDA GUNAKAN UNTUK VISUALISASI DATA DALAM PRESENTASI ANDA,” 2021. <https://ldikti12.ristekdikti.go.id/2021/09/01/9-jenis-chart-yang-dapat-anda-gunakan-untuk-visualisasi-data-dalam-presentasi-anda.html#:~:text=Diagram pie dapat digunakan untuk,komponen sebagai>

- persentase dari total. (accessed Jul. 02, 2022).
- [21] KEVIN OLLA, "Belajar Machine Learning dalam Pengolahan Data, Ini Panduannya," 2017.  
<https://www.jagoanhosting.com/blog/belajar-machine-learning-dalam-pengolahan-data-ini-panduannya/> (accessed Jul. 02, 2022).