



Analisis Sentimen Opini Publik Tentang Vaksin Booster Menggunakan Metode Support Vector Machine dan *firefly Algorithm*

Rifky Pujianto¹, Dadang Yusup², Tesa Nur Padilah³

^{1,2,3}Program Studi Informatika, Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang

Abstract

Received: 20 November 2022
Revised: 22 November 2022
Accepted: 25 November 2022

Corona virus is a viral epidemic that has hit the whole world. Indonesia is one of the countries affected by the virus so the government is trying to prevent the spread of the corona virus so that not many people are exposed to the corona virus. One of the government's efforts to prevent the spread of this virus is to create a free vaccination program. With this program the spread of the corona virus can be slowed down. In January 2022 the government issued the newest vaccine, namely the booster vaccine. The emergence of booster vaccines is a little doubted by the public because of the many hoax news about this booster vaccine. This study aims to analyze public opinion sentiment towards the emergence of booster vaccines. The data used for this analysis comes from public opinion that comes from Twitter. Data was retrieved using the crawling method. This study used the support vector machine (SVM) method as a process for classifying public opinion and the firefly algorithm as an optimization of the SVM parameters. There are 3 class labels used for classifiers, namely positive, negative, and neutral. A lot of data is used after the preprocessing process as much as 2223 data which is then split as much as 80% training data and 20% testing before entering the classification stage using the SVM method. The results of the classification using SVM produce an accuracy of 85% in the default parameters and after optimization using the firefly algorithm produce an accuracy of 86% with parameters $C = 1.0-3.0$, $\sigma = 0.1-1.0$.

Keywords: *Sentiment analysis, firefly algorithm, Support Vector Machine, twitter, Corona Virus*

(*) Corresponding Author: rifkypujianto7@gmail.com

How to Cite: Pujianto, R., Yusup, D., & Padilah, T. (2022). Analisis Sentimen Opini Publik Tentang Vaksin Booster Menggunakan Metode Support Vector Machine dan *firefly Algorithm*. *Jurnal Ilmiah Wahana Pendidikan*, 8(23), 363-373. <https://doi.org/10.5281/zenodo.7397891>

PENDAHULUAN

Virus corona atau biasa disebut covid-19 adalah sebuah wabah virus yang melanda seluruh dunia. Kasus virus corona pertama kali ditemukan di kota Wuhan, China. *World Health Organization* (WHO) menduga virus ini bersal dari hewan yang tertular oleh manusia dan sifat dari virus ini dapat menular melalui kontak fisik sehingga penyebaran virus ini sangat cepat sehingga kasus virus corona membeludak ke seluruh dunia. Kemudian WHO menetapkan covid-19 sebagai pandemi pada 9 Maret 2020 (Karyono, Rohadin, & Devia, 2020).

Di Indonesia, kasus virus corona pertama kali ditemukan di kota Depok. Kasus ini langsung ditindak langsung oleh pemerintah agar memperlambat penyebaran kasus virus corona di Indonesia. Berbagai upaya telah dilakukan oleh pemerintah Indonesia untuk menangani kasus agar tidak terus bertambah.



Vaksinasi adalah salah satu cara untuk mengurangi bertambahnya angka kasus virus corona. Vaksin berfungsi sebagai membuat sistem kekebalan tubuh seseorang untuk mencegah peyebaran virus tertentu dengan cara dimasukan kedalam tubuh, salah satunya vaksi virus corona (Covid-19). Meskipun vasksin tidak melindungi 100% dari ancaman virus corona, vaksin bisa memperkecil kemungkinan seseorang terkena virus corona (Zulaikha, Eliyan, & A'yun, 2021).

Pada awal Januari 2022 pemerintah mengumumkan vaksin dosis ketiga atau biasa disebut sebagai vaksin booster. Vaksin booster merupakan vaksin lanjutan setelah melakukan vaksin kesatu dan kedua. Kemunculan vaksin booster ini, banyak mayarakat yang ragu dalam melakukan vaksinasi vaksin booster. Keraguan dalam masyarakat ini berasal dari banyaknya berita-berita negatif dari vaksin booster.

Twitter adalah media sosial yang sangat populer di Indonesia karena kemudahannya dalam menyampaikan pendapat dan juga kebebasan dalam menyampaikan opini tentang sesuatu hal (Giovani, Ardiyansyah, Haryanti, Kurniawati dan Gata, 2020). Berbagai macam cara untuk menganalisis pendapat atau opini masyarakat di media sosial twiter, salah satunya adalah analisis sentimen (Laurensz dan Sedyono, 2021).

Analisis sentimen atau option mining adalah pengaplikasian dari ilmu komputasi linguistik, text analysis, dan natural language processing yang digunakan untuk mengekstrak subjektifitas dari sebuah sumber text (Husada & Paramita, 2021). Analisis sentimen saat ini dibagi menjadi penggunaan klasifikasi pembelajaran mesin dan metode klasifikasi berdasarkan aturan, metode pembelajaran mesin menggunakan kata emosi sebagai klasifikasi fitur, dan kamus emosi dapat digunakan untuk mewujudkan pemilihan karakteristik sentiment dengan cepat dan efisien. Metode pengklasifikasian yang umum digunakan adalah Naïve Bayes, (NB), Support Vector Machine (SVM) dan Entropi Maksimum (EM) (Isnain et al, 2021).

Support Verctor Machine (SVM) merupakan metode klasifikasi yang memiliki tingkat akurasi yang sangat tinggi. Prinsip dari kerja metode SVM adalah dengan cara ruang pemisah untuk data set yang paling optimal (Sasongko, 2016). SVM punya keunggulan dalam memisahkan kata data non-linear berdimensi tinggi. Dalam SVM terdapat kernel yang mempunyai fungsi sebagai pemisah input non-linear berdimensi tinggi dan hal tersebut membuat SVM menjadi istimewa dengan menggunakan fungsi kernelnya (Anjasmoros, Istiadi, & Marisa, 2020).

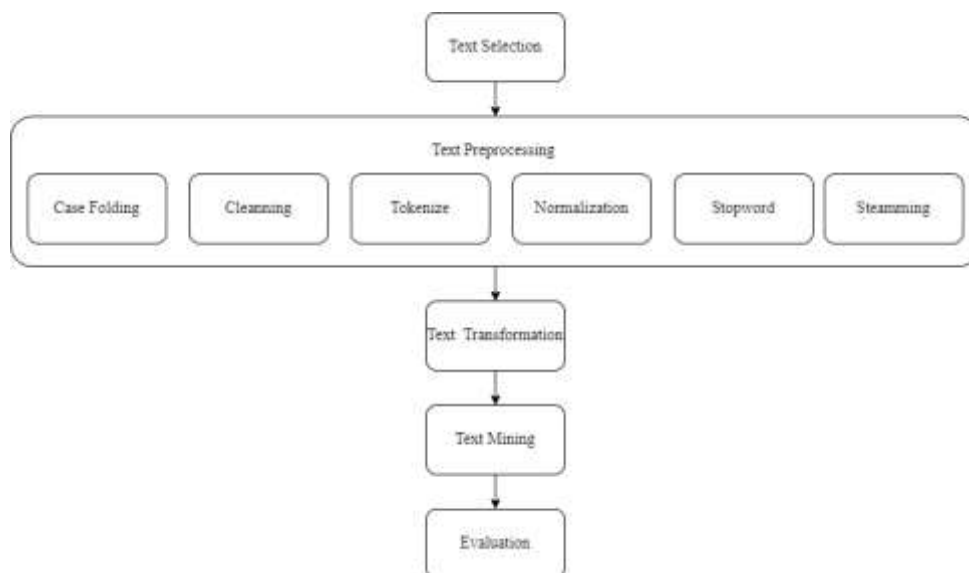
Penelitian yang dapat membuktikan bahwa SVM merupakan metode yang memiliki tingkat akurasi sangat tinggi yaitu penelitian yang dilakukan oleh Pamungkas & Kharisudin (2021) yang meneliti opini masyarakat tentang pademi covid-19 di media sosial twitter dengan menggunakan tiga metode yakni SVM, naïve bayes dan K-nearest neighbor (KNN) hasil dari klasifikasi metode SVM sebesar 90,1% dibandingkan dengan metode naïve bayes yang mendapatkan akurasi sebesar 79,2% dan KNN sebesar 62,11%. Penelitian Novantika & sugiman (2022) yang berjudul "Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression" SVM mendapatkan akurasi yang lebih tinggi dibandingkan dengan Logistic Regression. Akurasi yang didapatkan dari metode SVM lebih baik yaitu sebesar

87,02% dengan kernel linear dibandingkan dengan metode Logistic Regression yang mendapatkan akurasi 85,17%.

Penelitian Isnain, Sakti, Alita, & Marga (2021) yang meneliti analisis sentimen publik terhadap kebijakan lockdown pemerintah dengan menggunakan metode Support Verctor Machine (SVM) dengan menghasilkan nilai Accuracy sebesar 74 %, Precision sebesar 75%, Recall sebesar 92% dan F1-Score sebesar 83%. Namun penelitian tersebut dilakukan tanpa menggunakan optimasi sehingga peneliti menyarankan penggunaan metode firefly sebagai metode untuk optimasi agar akurasi yang didapatkan akan tinggi. Oleh karena itu penelitian ini akan melakukan pengabungan 2 metode yaitu metode Support Vector Machine (SVM) dengan metode firefly untuk meningkatkan akurasi dari metode Support Vector Machine (SVM).

METODOLOGI PENELITIAN

Metode yang digunakan untuk penelitian ini adalah *Knowledge Discovery in Database (KDD)*. Tahapan metode terdiri dari: *text Selection, text processing, text transformation, text mining, evaluation*.



Gambar 1. Rancangan Penelitian

1. Text Selection

Pada tahap ini dilakukan pengambilan data penelitian menggunakan teknik crawling dengan menggunakan pemograman pyhton. Data yang diambil merupakan opini masyarakat yang berasal dari media sosial twitter dari 20 Juli 2022 sampai dengan 20 Agustus 2022. data akan diberikan label secara manual yaitu dengan tiga label positif, netral, dan negatif dengan bantuan validasi oleh ahli bahasa.

2. Text Preprocessing

data yang didapatkan akan diolah menjadi sebuah text yang akan menjadi data yang berkualitas. Adapun tahapan dari *text processing* yaitu *Case folding, Cleaning, Tokenized, Normalization, Stopword, dan Steaming*.

3. Text Transformation

Text transformation adalah tahapan untuk mengontrol kapitalisasi text dan untuk merepresentasikan dokumennya dengan 2 cara yaitu *bag of word* dan *vector space*. Peneliti akan melakukan tahap ini dengan TF-IDF.

4. Text Mining

Tahapan *text mining* merupakan proses klasifikasi data. Proses klasifikasi data menggunakan metode *Support Vector Machine* (SVM) dan metode *firefly* dengan menggunakan *kernel* RBF pendekatan *k-fold validation*. Untuk data yang akan digunakan pada skenario proses klasifikasi adalah 80% data *training* dan 20% data *testing*.

5. Evaluation

Setelah melakukan klasifikasi akan dilakukan perhitungan akurasi dari FA-SVM. *Confusion Matrix* merupakan tools yang bisa membantu dalam pengukuran performa dalam klasifikasi. Hasil dari *Confusion Matrix* adalah *accuracy*, *precision*, *recall*, dan *f-score*.

HASIL DAN PEMBAHASAN

1. Text Selection

Pada proses pengambilan data dalam twitter, hal pertama yang harus dilakukan adalah meminta API untuk mengakses pengambilan data dan untuk mengambil API diperlukan akun *Twitter Developer*. Setelah mendapatkan API data twitter diambil dengan metode *crawling*. Gambar 2 menunjukkan hasil dari *crawling* data yang telah dilakukan dengan *keyword* “vaksin booster”. Data yang didapat sebanyak 2223 data dan kemudian data diberi label positif, netral, dan negatif. Hasil dari pelabelan data yang dilakukan ada 1854 data label positif, 231 data label netral, dan 138 data label negatif. Hasil dari pelabelan bisa dilihat pada gambar 3

2. Text Preprocessing

a. Cleanning

Proses *Cleanning* merupakan proses pembersihan kalimat atau kata pada data yang tidak berpengaruh dalam proses pengklasifikasian seperti *emoticon*, *space*, dan *punctuation*.

	Label	Tweet
0	Netral	akhirnya setelah libur hampir tahun towel tow...
1	Positif	gue udah booster pfizer bulan kemarin aman aj...
2	Positif	astaga wkwk aku vaksin sinovac booster pfizer...
3	Netral	karena prestasi bukan cuma dari sisi akademik...
4	Positif	yang sudah booster harus tetep pake masker da...

Gambar 2. Hasil Proses *Cleanning*

b. *Case Folding*

Proses *case folding* merupakan proses untuk menyeragamkan huruf besar (kapital) menjadi huruf kecil agar pada saat proses klasifikasi data tidak terjadi kesalahan karena perbedaan huruf.

akhirnya setelah libur hampir tahun towel tow...
 gue udah booster pfizer bulan kemarin aman aj...
 astaga wkwk aku vaksin sinovac booster pfizer...
 karena prestasi bukan cuma dari sisi akademik...
 yang sudah booster harus tetep pake masker da...

Gambar 3. Hasil Proses *Case folding*

c. *Tokenized*

Tokenized adalah proses pemotongan kalimat menjadi potongan kata-kata.

	Label	Tweet
0	Netral	[akhirnya, setelah, libur, hampir, tahun, towe...
1	Positif	[gue, udah, booster, pfizer, bulan, kemarin, a...
2	Positif	[astaga, wkwk, aku, vaksin, sinovac, booster, ...
3	Netral	[karena, prestasi, bukan, cuma, dari, sisi, ak...

Gambar 4. Hasil Proses *Tokenized*

d. *Normalization*

Proses *normalization* merupakan proses mengubah kata menjadi kata yang singkatan yang sebenarnya atau kata formal.

	Label	Tweet
0	Netral	[libur, towel, towel, perdana, pandemi, tersel...
1	Positif	[udah, booster, pfizer, kemarin, aman, kayak, ...
2	Positif	[astaga, vaksin, sinovac, booster, pfizer, gaa...
3	Netral	[prestasi, sisi, akademik, nyanyi, melukis, se...
4	Positif	[booster, tetep, pake, masker, jagajarak, boos...

Gambar 5. Hasil Proses *Normalization*

e. *Stopword*

Proses *stopword* merupakan proses penghapusan kata umum yang frekuensi kemunculannya sangat besar dan juga tidak memiliki makna.

	Label	Tweet
0	Netral	[libur, towel, towel, perdana, pandemi, tersel...
1	Positif	[udah, booster, pfizer, kemarin, aman, kayak, ...
2	Positif	[astaga, vaksin, sinovac, booster, pfizer, gaa...
3	Netral	[prestasi, sisi, akademik, nyanyi, melukis, se...
4	Positif	[booster, tetep, pake, masker, jagajarak, boos...

Gambar 6. Hasil Proses *Stopword*

f. Stemming

Proses steamming merupakan proses mengubah kata menjadi bentuk kata dasarnya. Dalam kata bahasa Indonesia, imbuhan dalam katanya di hilangkan.

	Label	Tweet
0	Netral	libur towel towel perdana pandemi selenggara s...
1	Positif	udah booster pfizer kemarin aman kayak vaksin ...
2	Positif	astaga vaksin sinovac booster pfizer gaada efe...
3	Netral	prestasi sisi akademik nyanyi luk sepakbola ba...
4	Positif	booster tetep pake masker jagajarak booster ms...

Gambar 7. Hasil Proses *Stemming*

Setelah dilakukan *pre-processing*, data tersebut diproses kembali untuk menghapus data tweet yang sama.

3. Text Transformation

Untuk menyelesaikan tahapan text transformation yaitu menggunakan metode TF-IDF untuk menghitung bobot dari setiap kata yang ada dan fungsi dari TF-IDF ini untuk mengurangi waktu dari komputasi.

Proses pertama yang dilakukan pada tahapan ini dilakukan *split* data. *split* menjadi dua menjadi data *training* dan data *testing* untuk melanjutkan tahap selanjutnya. Di penelitian ini, split pada data training sebesar 80% dan data testing 20%. Hasil dari split data yang dilakukan ada sekitar 1778 data *training* dan 445 data *testing*. Kemudian dilanjutkan dengan TF-IDF untuk menghitung bobot setiap kata yang ada dan juga menghitung *term* pada dokumen. Hasil dari TF-IDF yang dilakukan menghasilkan matriks 1778 X 3210. Pada matriks tersebut terdapat 1778 dokumen dari *dataset* dan 3210 *term* yang memiliki nilai bobot kata atau nilai TF-IDF.

	yuspahrudin	yuyuk	zeneca	zenica	zeus	zevian	zombie	zonauang	\
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
...	
1773	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1774	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1775	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1776	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1777	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

	zoster
0	0.0
1	0.0
2	0.0
3	0.0
4	0.0
...	...
1773	0.0
1774	0.0
1775	0.0
1776	0.0
1777	0.0

[1778 rows x 3210 columns]

Gambar 8. Hasil TF-IDF

4. Text Mining

a. Klasifikasi Support Vector Machine (SVM)

Pada proses ini dilakukan pengklasifikasi sentiment terhadap data tweet setelah melalui tahap transformation menggunakan algoritma support vector machine (SVM). Tujuan dari algoritma SVM ini mencari hyperlane sebagai pemisah antara data ke dalam kelas.

Proses pengklasifikasian dilakukan dengan data yang telah dibagi atau split sebelumnya yaitu 80% data Training dan 20% data testing dengan menggunakan kernel RBF (Radial Basis Function) dengan parameter $C = 1$ dan $\gamma = 1$. Gambar 9 Hasil dari proses pengklasifikasian.

```
In [9]: from sklearn.svm import SVC
svclassifier = SVC(kernel='rbf', C=1,gamma=1)
svclassifier.fit(text_tf_Xtrain, y_train)
svclassifier_score = svclassifier.score(text_tf_Xtest, y_test)

print(svclassifier_score)

0.849438202247191
```

Gambar 9. Hasil Klasifikasi SVM

b. Firefly Algorithm

Firefly algorithm merupakan algoritma yang didasarkan dari perilaku kunang-kunang. Pada penelitian ini, algoritma firefly digunakan untuk optimasi dari support vector machine (SVM) dengan mencari parameter. Parameter yang dicari dengan rentang $C = 1.0-3.0$, $\sigma = 1.0-2.0$ dan $C = 1.0-3.0$, $\sigma = 0.1-1.0$.

Jumlah firefly yang akan digunakan sebanyak 10 dan generasi yang akan digunakan sebanyak 10. Untuk konstanta nilai dari $\alpha = 0.2$, $\gamma = 1.0$, dan $\beta_0 = 1.0$, nilai konstanta tersebut. Setelah parameter di tetapkan, untuk mencari parameter yang mendapatkan akurasi tertinggi perlu pendekatan K-fold cross validation. Dalam penelitian ini 10 fold (k=10) yang akan digunakan.

Langkah pertama untuk melakukan optimasi Firefly Algorithm – Support Vector Machine (SVM) adalah melakukan inisiasi parameter yang

dibutuhkan untuk proses pencarian firefly. Inisiasi parameter telah dilakukan sebelumnya, selanjutnya menentukan populasi firefly secara acak agar firefly mentapi titik secara acak.

Setelah ditentukan posisinya, selanjutnya mencari intensitas cahaya dengan cara persamaan (1.1).

$$I(x) = f(x) \tag{1.1}$$

jika ada intensitas cahaya yang kecil dari intensitas cahaya lainnya, maka firefly tersebut akan melakukan perpindahan menuju yang lebih besar. Untuk menghitung perpindahan diperlukan menghitung jarak terlebih dahulu, perhitungan jarak firefly bisa dilakukan dengan persamaan (1.2).

$$r_{i,j} = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \tag{1.2}$$

Setelah mendapatkan jarak dari firefly 1 dan firefly 2, selanjutnya mencari keaktrarifan dari firefly tersebut akan digunakan sebagai acuan untuk perpindahan dari firefly. Untuk menghitung perpindahan pada firefly bisa dilakukan dengan persamaan (1.3). hasil dari perpindahan ini akan dimasukan kedalam metode SVM.

$$x_i = x_i + \beta_0 * \exp(-\gamma r_{ij}^2) * (x_j - x_i) + a * (rand - \frac{1}{2}) \tag{1.3}$$

Setelah melakukan pencarian firefly dari masing-masing parameter, kemungkinan pertama pada C = 1.0-3.0 mendapatkan parameter terbaik 2.92 sedangkan $\sigma = 0.1-1.0$ mendapatkan hasil parameter terbaik 0.5 dan menghasilkan akurasi 86%. Kemungkinan kedua pada C = 1.0-3.0 mendapatkan parameter terbaik 2.32 dan parameter $\sigma = 1.0-2.0$ mendapatkan parameter terbaik 1.07 dan menghasilkan akurasi 86%.

5. Evaluation

Evaluasi terhadap hasil dari optimasi parameter Support Vector Machine (SVM) menggunakan algoritma firefly dan divalidasi modelnya menggunakan confusion matrix. Confusion matrix merupakan tools yang dapat membantu dalam proses mendapatkan model hasil prediksi.

	[[1 0 25]				
	[0 7 42]				
	[0 0 370]]				
	precision	recall	f1-score	support	
Negatif	1.00	0.04	0.07	26	
Netral	1.00	0.14	0.25	49	
Positif	0.85	1.00	0.92	370	
accuracy			0.85	445	
macro avg	0.95	0.39	0.41	445	
weighted avg	0.87	0.85	0.79	445	

Gambar 10. Confusion Matrix SVM

Gambar merupakan hasil confusion matrix dari metode Support Vector Machine dengan menggunakan kernel RBF menghasilkan accuracy 85% dan menghasilkan AA (*true negatif*) = 1, BB (*true netral*) = 7, dan CC (*true positif*) = 370. Hasil confusion matrix dapat dihitung secara manual sebagai berikut

$$\text{Akurasi} = \frac{AA+BB+CC}{AA+AB+AC+BA+BB+BC+CA+CB+CC} = \frac{1+7+370}{445} \approx 0,849$$

Untuk mempermudah mencari data/nilai yang *true positif*, *true negatif*, *false positif* dan, *false negatif* dari hasil confusion matrix 3x3 yaitu dengan cara mengubah ukuran matriks menjadi 2x2 di setiap kelas, hasil dari pengubahan matriks bisa dilihat pada gambar 11.

Data Aktual	Data Prediksi			Data Aktual	Data Prediksi			Data Aktual	Data Prediksi	
	Negatif	Bukan			Netral	Bukan			Positif	Bukan
Negatif	1	25		Netral	7	42		Positif	370	0
Bukan	0	419		Bukan	0	396		Bukan	67	8

Gambar 11. Hasil Confusion Matrix 2x2

Untuk membuktikan *recall*, *precision*, dan *f1-score* dari masing-masing sesuai dengan hasil gambar 11 maka perhitungan diperlukan perhitungan secara manual seperti gambar 12.

<p>Kelas Negatif</p> <p>a. Recall</p> $\text{Recall} = \frac{TP}{TP+FP} = \frac{1}{1+25} = 0,04$ <p>b. Precision</p> $\text{Precision} = \frac{TP}{TP+FN} = \frac{1}{1+0} = 1$ <p>c. F1-score</p> $\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0,08}{1,04} = 0,076$	<p>Kelas Netral</p> <p>a. Recall</p> $\text{Recall} = \frac{TP}{TP+FP} = \frac{7}{7+42} = 0,14$ <p>b. Precision</p> $\text{Precision} = \frac{TP}{TP+FN} = \frac{7}{7+0} = 1$ <p>c. F1-score</p> $\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0,28}{1,14} = 0,245$	<p>Kelas Positif</p> <p>1. Recall</p> $\text{Recall} = \frac{TP}{TP+FP} = \frac{370}{370+0} = 1$ <p>2. Precision</p> $\text{Precision} = \frac{TP}{TP+FN} = \frac{370}{370+67} = 0,85$ <p>3. F1-score</p> $\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 1,7}{1,85} = 0,918$
---	--	--

Gambar 12. Hasil Perhitungan Manual

Dari hasil perhitungan manual recall, precision, f1-score setiap masing-masing kelas jika dibandingkan dengan hasil gambar 10 hasilnya sama. Kemudian dilanjutkan dengan hasil dari evaluasi confusion matrix dari optimasi algoritma *firefly* seperti gambar 13.

	[[3 0 23]				
	[0 11 38]				
	[1 2 367]]				
	precision	recall	f1-score	support	
Negatif	0.75	0.12	0.20	26	
Netral	0.85	0.22	0.35	49	
Positif	0.86	0.99	0.92	370	
accuracy			0.86	445	
macro avg	0.82	0.44	0.49	445	
weighted avg	0.85	0.86	0.82	445	

Gambar 13. Hasil Confusion Matrix FA-SVM

Setelah dioptimasi menggunakan algoritma firefly, accuracy meningkat sebesar 1%. Untuk hasil dari confusion matrix menghasilkan accuracy 86% dan menghasilkan AA (*true negatif*) = 3, BB (*true netral*) = 11, dan CC (*true positif*) = 367. Untuk perhitungan manualnya seperti dibawah.

$$\text{Akurasi} = \frac{AA+BB+CC}{AA+AB+AC+BA+BB+BC+CA+CB+CC} = \frac{2+10+369}{445} \approx 0,856$$

Untuk mencari nilai *true positif*, *true negatif*, *false positif* dan, *false negatif* dari setiap kelas yang sudah dioptimasi. Maka setiap kelas akan dirubah menjadi matriks ukuran 2x2. Hasil perubahan bisa dilihat pada gambar 14.

Data Aktual	Data Prediksi		Data Aktual	Data Prediksi		Data Aktual	Data Prediksi	
	Negatif	Bukan		Netral	Bukan		Positif	Bukan
Negatif	3	23	Netral	11	38	Positif	367	3
Bukan	1	418	Bukan	2	394	Bukan	61	14

Gambar 14. Hasil Confusion Matrix 2x2

Untuk membuktikan *recall*, *precision*, dan *f1-score* dari masing-masing sesuai dengan hasil gambar 13 maka perhitungan diperlukan perhitungan secara manual, seperti gambar 15.

Kelas Negatif	Kelas Netral	Kelas Positif
1) Recall	1) Recall	1) Recall
$Recall = \frac{TP}{TP+FP} = \frac{3}{3+23} = 0,115$	$Recall = \frac{TP}{TP+FP} = \frac{11}{11+38} = 0,22$	$Recall = \frac{TP}{TP+FP} = \frac{367}{367+3} = 0,99$
2) Precision	2) Precision	2) Precision
$Precision = \frac{TP}{TP+FN} = \frac{3}{3+1} = 0,75$	$Precision = \frac{TP}{TP+FN} = \frac{11}{11+2} = 0,85$	$Precision = \frac{TP}{TP+FN} = \frac{367}{367+1} = 0,86$
3) F1-score	3) F1-score	3) F1-score
$F1\text{-score} = \frac{2 \times Precision \times Recall}{Precision+Recall} = \frac{2 \times 0,115 \times 0,75}{0,865} = 0,19$	$F1\text{-score} = \frac{2 \times Precision \times Recall}{Precision+Recall} = \frac{2 \times 0,85 \times 0,22}{1,11} = 0,35$	$F1\text{-score} = \frac{2 \times Precision \times Recall}{Precision+Recall} = \frac{2 \times 0,99 \times 0,86}{1,85} = 0,92$

Gambar 15 Hasil Perhitungan Manual

KESIMPULAN

Berdasarkan hasil dari analisis yang telah dilakukan, maka hasil yang didapatkan adalah sebagai berikut:

1. Hasil dari analisis sentimen opini masyarakat tentang vaksin booster menghasilkan 83,4% data positif, 10,4% data netral dan 6,2% data negatif. Ketepatan dalam penentuan kelas sentimen menggunakan metode *Support Vector Machine* (SVM) sebesar 85%, setelah dioptimasi menggunakan algoritma firefly menghasilkan ketepatan sebesar 86%.

2. Hasil penelitian menunjukkan bahwa algoritma *firefly* dapat menemukan kombinasi parameter SVM yang sesuai untuk meningkatkan akurasi. berdasarkan hasil dari *confusion matrix*. Pada evaluasi SVM menghasilkan akurasi sebesar 85% dengan parameter $C=1$ dan $\gamma = 1$ dan juga menghasilkan. untuk hasil evaluasi SVM dengan optimasi algoritma *firefly* menghasilkan "86%" akurasi dengan parameter $C= 1.0 - 3.0$ dan $\sigma = 0.1 - 1.0$.

DAFTAR PUSTAKA

- Anjasmosros, M. T., Istiadi, I., & Marisa, F. (2020). Analisis sentimen aplikasi gojek menggunakan metode SVM dan NBC (Studi kasus: Komentar pada play store). In *Conference on Innovation and Application of Science and Technology (CIASTECH)*, 3(1), pp. 489-498. Retrieved from <http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/article/view/1905>.
- Giovani, A.P., Ardiansyah., Haryanti, T., Kurniawati, L., Gata, W. (2020). Analisis sentimen aplikasi ruang guru di twitter menggunakan algoritma klasifikasi. *Jurnal TEKNOINFO*, 14(2), 116-124. <https://doi.org/10.33365/jti.v14i2.679>.
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18-26.
- Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), 31-37.
- Karyono., Rohadin., & Indriyani, D. (2020). Penanganan dan pencegahan pandemi wabah virus corona (COVID-19) kabupaten Indramayu. *Jurnal Kolaborasi Resolusi Konflik*, 2, 164-173. doi.org/10.24198/jkrk.v2i2.29127.
- Laurensz, B., & Eko Sedyono. (2021). Analysis of public sentiment on vaccination in efforts to overcome the covid-19 pandemic. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 10(2), 118-123. <https://doi.org/10.22146/jnteti.v10i2.1421>.
- Novantika, A., & Sugiman, S. (2022, February). Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression. In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 5, pp. 808-813).
- Pamungkas, F. S., & Kharisudin, I. (2021, February). Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 4, pp. 628-634).
- Sasongko, T. B. (2016). Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Klasifikasi Jalur Minat SMA). *Jurnal Teknik Informatika dan Sistem Informasi*, 2(2).
- Zulaikha, L. I. (2021). Pemberian Vaksin Sinovac Kepada Masyarakat Untuk Mencegah Penyebaran Covid 19. *Jurnal Pengabdian Kepada Masyarakat*, 2(2), 34-37.