

Optimasi Algoritma *Support Vector Machine* Menggunakan Seleksi Fitur *Particle Swarm Optimization* Pada Analisis Sentimen Terhadap Kebijakan PPKM

Hasan Mubarak¹, Iin Ernawati², Nurul Chamidah³

S1 Informatika / Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

Jl. RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450, Indonesia

hasanm@upnvj.ac.id¹, iinernawati@upnvj.ac.id², nurul.chamidah@upnvj.ac.id³

Abstrak. Twitter merupakan media sosial yang bersifat *micro-blogging* yang memungkinkan penggunaannya untuk mengekspresikan opini tentang berbagai topik dan membahas permasalahan yang terjadi saat ini. Salah satu topik yang sering diperbincangkan masyarakat yaitu penerapan kebijakan PPKM di Indonesia yang menimbulkan prokontra sehingga opini dari masyarakat sangat beragam terutama pengguna twitter. Banyaknya opini maka perlu adanya analisis sentimen. Tujuannya adalah untuk mengetahui opini masyarakat terhadap penerapan PPKM melalui tagar #PPKM. Oleh karena itu, penelitian ini dilakukan proses klasifikasi pada penerapan PPKM menggunakan dua kelas yaitu kelas sentimen positif dan kelas sentimen negatif. Metode yang digunakan dalam mengklasifikasi yakni algoritma *Support Vector Machine* dan algoritma *Particle Swarm Optimization* sebagai seleksi fitur. Pengambilan data dimulai tanggal 1 Juli – 30 Agustus 2021. Hasil evaluasi klasifikasi menggunakan *confusion matrix* didapatkan nilai akurasi 79.77%, recall 69.04%, dan 85.29% pada data tanpa PSO. Sedangkan pada data menggunakan PSO didapatkan nilai akurasi sebesar 87.08%, recall 76.83%, dan Precision 94.03%.

Kata Kunci : Twitter, PPKM, *Support Vector Machine* (SVM), *Particle Swarm Optimization* (PSO).

1 Pendahuluan

Perkembangan ilmu pengetahuan dan teknologi saat ini sangat berkembang pesat salah satunya yaitu perkembangan internet. Semakin berkembangnya internet sekarang ini banyak sekali platform-platform yang berkembang khususnya media sosial yang mengakibatkan meningkatnya penggunaan internet dan media sosial di dunia termasuk Indonesia. Menurut data *We Are Social Hootsuite* pada Januari 2021, 73% merupakan jumlah persentase pengguna internet dari jumlah penduduk di Indonesia. Data tersebut menunjukkan bahwa penggunaan internet meningkat sebesar 9% dari tahun sebelumnya [1].

Twitter menjadi salah satu media sosial yang sangat populer di Indonesia. Dimana jumlah pengguna aktif media sosial twitter di Indonesia mencapai 63% dari jumlah penduduk di Indonesia [1]. Twitter merupakan platform media sosial yang bekerja secara realtime, kepopuleran twitter ini memungkinkan penggunaannya memanfaatkan untuk mengekspresikan opini atau perasaan mereka mengenai banyak isu atau suatu permasalahan. Ekspresi yang diberikan dapat berupa opini positif maupun negatif melalui bentuk *tweet*.

Pada saat ini, twitter tidak hanya digunakan penggunaannya untuk membagikan konten pribadi melainkan mencari suatu informasi baik itu tentang bisnis, hiburan, politik dan lainnya. Namun, twitter dapat digunakan sebagai media menyampaikan aspirasi dan pendapat pada suatu permasalahan yang sedang terjadi, topik perbincangan masyarakat yaitu peraturan pemerintah mengenai Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM). PPKM adalah peraturan pemerintah untuk pencegahan dan penanggulangan COVID-19 yang disahkan sejak tahun 2021. Kebijakan ini berlaku di beberapa wilayah Indonesia yaitu pulau Jawa dan Bali [2]. Sebelum diberlakukannya PPKM pemerintah telah memberlakukan sebuah kebijakan yaitu Pembatas Sosial Berskala Besar (PSBB). Kebijakan pemerintah ini menimbulkan prokontra pada masyarakat dimana peraturan pemerintah berdampak terhadap perekonomian dan tidak efektif menurunnya kasus harian COVID-19 karena tidak patuh masyarakat terhadap kebijakan yang sudah ditetapkan.

Berdasarkan permasalahan yang terjadi, maka menjadi dasar peneliti untuk melakukan analisis sentimen terhadap kebijakan PPKM melalui *tweet* di media sosial twitter yang nantinya dapat digunakan untuk melakukan evaluasi dan identifikasi sebuah opini masyarakat baik itu positif ataupun negatif. Penelitian mengenai analisis sentimen telah banyak digunakan menggunakan berbagai macam metode seperti Support Vector Machine (SVM), Naïve Bayes (NB), dan lain-lain. Perbandingan metode klasifikasi untuk analisis sentimen menunjukkan bahwa algoritma Support Vector Machine memperoleh performa yang lebih baik dibandingkan dengan metode klasifikasi lainnya [3]. Namun algoritma Support Vector Machine memiliki kelemahan dalam melakukan komputasi data dengan jumlah besar, hal ini dapat menyebabkan kinerja klasifikasi akurasi yang didapatkan tidak optimal [4].

Sehingga pada penelitian ini akan menggunakan algoritma *Support Vector Machine* (SVM) dan algoritma *Particle Swarm Optimization* sebagai seleksi fitur untuk dapat mengoptimalkan nilai akurasi serta menganalisis sentimen untuk mengetahui tanggapan dari masyarakat Indonesia terkait kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat PPKM.

2 Metodologi Penelitian

2.1 Analisis Sentimen

Analisis sentimen adalah studi komputasional dari opini atau emosi seseorang yang diekspresikan kedalam sebuah kalimat berdasarkan suatu entitas sebagai contoh permasalahan kebijakan atau layanan. Selain itu, analisis sentimen bertujuan untuk mengekstraksi, memahami, dan mengolah data menjadi suatu informasi yang bermanfaat informasi tersebut dapat berupa sentimen positif, negatif atau netral [5].

2.2 Text Mining

Text mining merupakan tahap melakukan penggalian informasi dari suatu data teks untuk mendapat suatu informasi yang berguna dalam pengambilan keputusan. *Text mining* memiliki tujuan utama yaitu untuk menganalisis dan menemukan pola yang menarik dalam data teks [6].

2.3 Fleiss Kappa

Fleiss kappa merupakan metode untuk mengukur kesepakatan pelabelan yang dilakukan oleh bantuan lebih dari dua annotator [7]. Fleiss kappa bertujuan untuk mengetahui validasi atau tidak pelabelan yang telah dilakukan oleh annotator [8]. Persamaan untuk menghitung fleiss kappa sebagai berikut:

$$Fleiss\ Kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

Keterangan:

P_o = Proporsi frekuensi pengamatan.

P_e = Peluang kesepakatan antar pengamat

2.4 Pra-proses Data

Pra-proses data merupakan tahap melakukan pembersihan data yang telah didapat melalui teknik *crawling* dengan API (*Application Programming Interface*). Tahap pra-proses data mencakup kegiatan membangun data tidak terstruktur menjadi terstruktur dan membersihkan noise agar siap untuk diolah ke tahap selanjutnya sehingga data tersebut dapat digunakan dengan baik [9]. Berikut tahapan pra-proses data:

a) Case Folding

Case folding merupakan tahapan mengubah setiap kata yang terdapat di dalam dokumen menjadi huruf kecil atau lowercase. Tujuan dilakukan case folding untuk mengubah semua bentuk format kata di setiap dokumen menjadi sama [10].

b) Cleaning

Cleaning adalah tahapan proses pembersihan data dengan menghapus karakter-karakter seperti tanda baca, hastag, URL dan lain-lain. Pembersihan data bertujuan untuk mengurangi noise [10].

c) Normalization

Normalization atau normalisasi adalah memperbaiki kata-kata tweet yang menggunakan bahasa gaul dan sehari-hari menjadi bahasa formal sesuai dengan Kamus Besar Bahasa Indonesia (KBBI)[11].

d) Tokenization

Tokenization adalah proses memecahkan kalimat menjadi potongan kata-kata yang biasa disebut token [9]

e) Stopword Removal

Stopword removal merupakan tahap menghapus kata yang tidak memiliki makna seperti kata hubung dan kata lainnya yang terdapat pada daftar stopwords [12].

f) Stemming

Stemming adalah proses transformasi kata yang memiliki imbuhan didepan maupun dibelakang menjadi kata dasar (root) [13].

2.5 Pembobotan TF-IDF

Pembobotan TF-IDF adalah proses ekstraksi fitur dengan memberikan bobot pada setiap masing-masing kata berdasarkan frekuensi munculnya kata yang terdapat di dalam dokumen. Tahap TF-IDF bertujuan untuk mengetahui seberapa penting sebuah kata mewakili sebuah kalimat [14]. Tahapan pembobotan TF-IDF dengan persamaan (2) dan (3) sebagai berikut:

Persamaan menghitung bobot TF-IDF

$$W_{t,d} = W_{tf,t,d} \times idf_t \quad (2)$$

Persamaan menghitung IDF

$$idf_t = \log_{10} \frac{N}{df_t} \quad (3)$$

Keterangan:

$W_{tf,t,d}$ = Bobot kata dalam setiap dokumen.

N = Jumlah keseluruhan dokumen.

$tf_{t,d}$ = Jumlah muncul kata t dalam dokumen d.

df_t = Jumlah dokumen mengandung term.

idf_t = Bobot *invers* setiap kata dari berbagai dokumen.

$W_{t,d}$ = Pembobotan TF-IDF setiap term pada dokumen.

2.6 Particle Swarm Optimization

Menurut (Kiranyaz, Ince, and Gabbouj 2014) Particle swarm Optimization (PSO) diperkenalkan oleh Kennedy dan Eberhart pada tahun 1995 sebagai proses pencarian dan optimasi stokastik berbasis populasi. Particle Swarm Optimization ini dapat disimulasikan seperti kawanan burung atau koloni semut. Metode ini tidak memiliki operator seperti crossover atau mutasi, baris dalam metric disebut particle. Setiap particle akan bergerak dengan kecepatan, setiap pembaruan kecepatan dan posisi nantinya berdasarkan pbest (lokal best) dan gbest (global terbaik) [15]. Particle Swarm Optimization sering digunakan dalam penelitian sebagai metode seleksi fitur, tujuannya untuk mengoptimasi nilai akurasi pada klasifikasi dengan menggunakan persamaan sebagai berikut:

Persamaan memperbarui kecepatan:

$$V_i(t) = V_i(t-1) + c_1 r_1 (X_i^L - X_i(t-1)) + c_2 r_2 (X^G - X_i(t-1)) \quad (4)$$

Persamaan memperbarui posisi:

$$X_i(t) = V_i(t) + X_i(t - 1) \quad (5)$$

Keterangan:

X_i^L = Local best

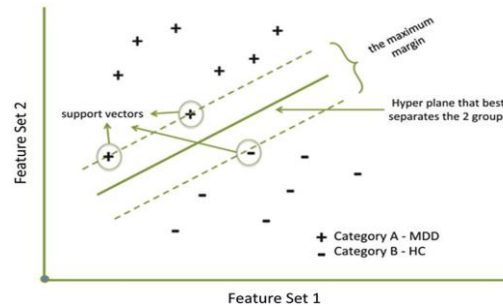
$C_1 C_2$ = Learning factor

X^G = Global best

$r_1 r_2$ = Bilangan random

2.7 Support Vector Machine

Support Vector Machine didasarkan pada teori pembelajaran statistika oleh Vapnik et al mengusulkan metode baru atas dasar jumlah sampel dalam informasi untuk mendapatkan yang terbaik hasil klasifikasi. Algoritma Support Vector Machine memaksimalkan jarak (margin) antar dua kelas untuk mendapatkan hyperplane terbaik [16].



Gambar. 1.Pencarian Hyperplane

Pada Gambar. 1 merupakan ilustrasi data linear klasifikasi support vector machine dengan dua kelas. Support Vector Machine adalah algoritma supervised learning yang digunakan untuk klasifikasi yang bekerja mencari hyperplane dengan margin terbesar. Dalam menentukan hyperplane data dapat mempertimbangkan titik data menggunakan $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$ dimana $y_n = 1/-1$ sebuah konstanta yang menunjukkan kelas yang titik x_n miliki dan $n =$ jumlah sampel, dapat didefinisikan persamaan:

$$w \cdot x + b = 0 \quad (6)$$

Dimana b adalah scalar dan w adalah vector p-dimensi. Pada kelas positif dapat diinisiasi dengan $y_n=1$, sehingga dapat didefinisikan bidang pembatas pertama dengan persamaan:

$$w \cdot x + b = 1 \quad (7)$$

Sedangkan pada kelas negatif dapat diinisiasi dengan $y_n=-1$, sehingga dapat didefinisikan bidang pembatas kedua dengan persamaan:

$$w \cdot x + b = -1 \quad (8)$$

Jika data dipisahkan secara linear kita dapat memaksimalkan jarak mereka dengan $\frac{2}{|w|}$ atau dengan meminimalkan $|w|^2$. Maka pencarian hyperplane terbaik dengan memaksimalkan kedua bidang pembatas dapat dipresentasikan dalam bentuk persamaan:

$$y_i(w \cdot x_i + b) \geq 1 \quad (9)$$

Dalam hal teknik klasifikasi support vector machine dapat menggunakan data linear atau non-linear. Untuk menyelesaikan permasalahan non-linear, algoritma support vector machine memanfaatkan fungsi kernel. Namun, penggunaan fungsi kernel dapat menyebabkan over fitting, oleh karena itu digunakan soft margin dan konstanta C untuk mengontrol antara margin dan error. Fungsi kernel yang digunakan yaitu Linear dan Radial Basis Function (RBF) dengan persamaan kernel:

Persamaan kernel Linear:

$$K(x_t, x_{tn}) = x_t \quad (10)$$

Persamaan kernel Radial Basis Function:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (11)$$

2.8 Evaluasi

Evaluasi dibutuhkan dalam mengetahui seberapa baik kinerja algoritma yang digunakan. Salah satu metode evaluasi yaitu confusion matrix, metode ini akan membuat visualisasi dalam bentuk tabel untuk mengevaluasi kinerja dari klasifikasi [17].

Tabel. 1. Tabel Confusion Matrix

		Aktual	
		Positif	Negatif
Prediksi	Positif	True Positive (TP)	False Positive (FP)
	Negatif	False Negative (FN)	True Negative (TN)

Pada tabel. 1 terdapat 4 label dalam mempresentasikan hasil proses klasifikasi yaitu:

- True Positive (TP) : Data aktual positif yang diprediksi positif.
- True Negative (TN) : Data aktual negatif yang diprediksi negatif.
- False Positive (FP) : Data aktual negatif tetapi yang diprediksi data positif.
- False Negative (FN) : Data aktual positif tetapi yang diprediksi data negatif.

Berdasarkan keempat kategori tersebut kita dapat mengetahui nilai accuracy, recall dan precision. Sehingga kita mengetahui seberapa bagus model yang dibuat dan kinerja klasifikasi.

Persamaan untuk menghitung nilai accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

Persamaan untuk menghitung nilai recall:

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

Persamaan untuk menghitung nilai precision:

$$Precision = \frac{TP}{TP + FP} \quad (14)$$

3 Hasil dan Pembahasan

Pada penelitian ini menggunakan data tweet pengguna pada media sosial twitter yang didapatkan dengan melakukan teknik crawling. Dalam melakukan crawling media sosial twitter telah menyediakan Application Program Interface (API). Data yang digunakan sebanyak 886 data, mengenai kebijakan PPKM diambil berdasarkan kata kunci #PPKM dalam rentang waktu 1 Juli – 30 Agustus 2021. Berikut data mentah hasil crawling.

1	Presiden @Jokowi: "Saya juga minta kepada seluruh gubernur, bupati, dan wali kota baik yang berada di Pulau Jawa dan Pulau Bali maupun di luar Pulau Jawa, semuanya untuk terus turun ke bawah, mengecek
2	Pemerintah Provinsi DKI Jakarta akan memberikan sanksi kepada pelaku usaha non-esensial dan kritikal yang masih mempekerjakan pegawainya saat PPKM Darurat berjalan.
3	Lepaskan sejenak semua ego, kepentingan golongan, kepentingan pribadi dan mari bersama-sama kt tuntaskan misi menghadapi pandemi ini
4	"Pesantren harus di gencarkan sosialisasi PPKM darurat karena mereka sangat rentan. Jelaskan poin-poin nya dengan detail, terutama protokol kesehatan" #gusmuhammad
5	Jangan sampai ada gelombang PHK baru pada saat PPKM darurat ini.
6	Masa PPKM, masyarakat berlomba untuk vaksin. Dimana sertifikat vaksin adalah syarat utama dalam mobilitas baik melalui udara, laut dan darat. Banyak pihak penyelenggara transportasi mulai membuka layanan
7	Bertarung juga harus mentaati protokol kesehatan ??
8	Semoga kita semua bisa bertahan dan berjuang untuk Indonesia sembuh!
9	mungkin akun ini kecil, tapi selama #ppkm saya akan membantu merepost setiap usaha mencari bantuan atau promo usaha teman teman yang ditag ke saya, disertai doa tulus, semoga yang sakit lekas sembuh dan
10	mengisi #ppkm hari ke4 dgn kerja sembari masak tumis sawi jamur dan goreng jamur kacang tepung

Gambar 2. Data Hasil Crawling

Data tweet akan dilakukan pelabelan menjadi dua kelas yaitu kelas positif dan kelas negatif dengan bantuan 3 annotator, hasil pelabelan dapat dilihat pada tabel 2 sebagai berikut:

Tabel. 2. Contoh Pelabelan Manual

Data	Annotator1	Annotator2	Annotator3	Hasil Akhir
Presiden @jokowi : "Saya juga minta kepada seluruh gubernur, bupati, dan wali kota baik yang berada di Pulau Jawa dan Pulau Bali maupun di luar Pulau Jawa, semuanya untuk terus turun ke bawah, mengecek lapangan & kesediaan alkes." #AyoVaksin #ppkm #6M https://t.co/ENGJUFY7II Semoga kita semua bisa bertahan dan berjuang untuk Indonesia sembuh! Ayo semangat sembuh dengan di rumah aja! #PPKMDarurat #PPKM #luhutbinsarpandjaitan #JabarJuara #covid_19 #IndonesiaKolaps #Ibas #Pejabat	Positif	Negatif	Positif	Positif
Tolong dengarkan kami rakyat bawah pak @jokowi @aniesbaswedan #suararakyat #PPKM #COVID19 #keadilan https://t.co/e6GC5rsdZT	Negatif	Negatif	Negatif	Negatif

Hasil akhir pelabelan ini disepakati dengan pemungutan suara mayoritas pada kelas sentimen dari 3 orang annotator. Untuk mengukur kekonsisten dari 3 annotator menggunakan fleiss kappa [6] dengan persamaan (1). Selanjutnya akan dilakukan pra-proses data untuk menghilangkan noise.

Tabel. 3. Sebelum dan Sesudah Pra-Proses Data

Sebelum Pra-Proses Data	Sesudah Pra-Proses Data
Semoga kita semua bisa bertahan dan berjuang untuk Indonesia sembuh! Ayo semangat sembuh dengan di rumah aja! #PPKMDarurat #PPKM #luhutbinsarpandjaitan #JabarJuara #covid_19 #IndonesiaKolaps #Ibas #Pejabat	moga tahan juang indonesia sembuh ayo semangat sembuh rumah
indonesia.... oh indonesia #ppkm pelan pelan kita mati peraturan di buat hanya untuk rakyat kecil tidak berlaku bagi pejabat https://t.co/k1uylyvvmj saat ini gue bener bener bosan dengan situasi pandemi ini, apalagi sekarang ppkm yang mengharuskan untuk dirumah aja. padahal sebenarnya pengen banget hangout sekalian hunting lama gak aplod sosmed hasil hunting gue karena ppkm ini. #ppkm #ppkmdiperpanjang #camermalang https://t.co/ondaeyhseg	indonesia indonesia pelan pelan mati atur rakyat laku jabat jenuh situasi pandemi ppkm harus rumah jalan foto unggah sosial media hasil foto ppkm

Setelah dilakukan pra-proses data, tahap selanjutnya akan dilakukan pembobotan TF-IDF terhadap setiap kata didalam dokumen tweet menggunakan persamaan (2) dan (3).

Tabel. 4. Perhitungan TF-IDF

Term	Dokumen			DF	IDF	TF – IDF		
	D1	D2	D3			D1	D2	D3
moga	1	0	0	1	0.477121255	0.477121255	0	0
tahan	1	0	0	1	0.477121255	0.477121255	0	0
juang	1	0	0	1	0.477121255	0.477121255	0	0

indonesia	1	2	0	2	0.176091259	0.176091259	0.352182518	0
sembuh	1	0	0	1	0.477121255	0.477121255	0	0
ayo	1	0	0	1	0.477121255	0.477121255	0	0
semangat	1	0	0	1	0.477121255	0.477121255	0	0
rumah	1	0	1	2	0.176091259	0.176091259	0	0.176091259
pelan	0	2	0	1	0.477121255	0	0.954242509	0
mati	0	1	0	1	0.477121255	0	0.477121255	0
atur	0	1	0	1	0.477121255	0	0.477121255	0
rakyat	0	1	0	1	0.477121255	0	0.477121255	0
laku	0	1	0	1	0.477121255	0	0.477121255	0
jabat	0	1	0	1	0.477121255	0	0.477121255	0
jenuh	0	0	1	1	0.477121255	0	0	0.477121255
situasi	0	0	1	1	0.477121255	0	0	0.477121255
pandemi	0	0	1	1	0.477121255	0	0	0.477121255
ppkm	0	0	2	1	0.477121255	0	0	0.954242509
harus	0	0	1	1	0.477121255	0	0	0.477121255
jalan	0	0	1	1	0.477121255	0	0	0.477121255
foto	0	0	2	1	0.477121255	0	0	0.954242509
unggah	0	0	1	1	0.477121255	0	0	0.477121255
sosial	0	0	1	1	0.477121255	0	0	0.477121255
media	0	0	1	1	0.477121255	0	0	0.477121255
hasil	0	0	1	1	0.477121255	0	0	0.477121255

Tabel 4 diatas merupakan contoh sampel 3 dokumen dari 886 data untuk melakukan perhitungan TF-IDF. Setelah dilakukan pembobotan menghasilkan 1815 fitur dengan nilai bobot masing-masing.

Setelah dilakukan pra-prsoes dan pembobotan TF IDF akan dilakukan klasifikasi menggunakan algoritma support vector machine. Dalam penelitian ini akan melakukan klasifikasi menggunakan algoritma support vector machine dan particle swarm optimization sebagai seleksi fitur dan klasifikasi menggunakan support vector machine tanpa seleksi fitur. Tahapan proses seleksi fitur particle swarm optimization akan dilakukan beberapa skenario untuk mendapatkan hasil yang optimal dengan mengatur iterasi secara konstan dengan selisih 50 setiap skenarionya dimulai dari iterasi 50 hingga iterasi 1000. Sebelumnya data sebanyak 886 dokumen akan dilakukan pembagian data dengan rasio 80% data latih dan 20% data uji.

Tabel. 5. Pembagian Data dan Hasil Seleksi Fitur

	Data Latih	Data Uji	Jumlah Data	Jumlah Fitur	
SVM Tanpa PSO	708	178	886	1815	
Menggunakan PSO	SVM + PSO (Iterasi=50)	708	178	886	941
	SVM + PSO (Iterasi=100)	708	178	886	982
	SVM + PSO (Iterasi=150)	708	178	886	940
	SVM + PSO (Iterasi=200)	708	178	886	955
	SVM + PSO (Iterasi=250)	708	178	886	926
	SVM + PSO (Iterasi=300)	708	178	886	925
	SVM + PSO (Iterasi=350)	708	178	886	904
	SVM + PSO (Iterasi=400)	708	178	886	940
	SVM + PSO (Iterasi=450)	708	178	886	937
	SVM + PSO (Iterasi=500)	708	178	886	956
	SVM + PSO (Iterasi=550)	708	178	886	915
	SVM + PSO (Iterasi=600)	708	178	886	960
	SVM + PSO (Iterasi=650)	708	178	886	959
	SVM + PSO (Iterasi=700)	708	178	886	938
	SVM + PSO (Iterasi=750)	708	178	886	921
	SVM + PSO (Iterasi=800)	708	178	886	956
	SVM + PSO (Iterasi=850)	708	178	886	924
	SVM + PSO (Iterasi=900)	708	178	886	910
	SVM + PSO (Iterasi=950)	708	178	886	919
	SVM + PSO (Iterasi=1000)	708	178	886	934

Setelah dilakukan pembagian data lalu ke tahap klasifikasi. Pada proses klasifikasi support vector machine menggunakan kernel RBF. Hasil klasifikasi algoritma SVM dan SVM+PSO dapat dilihat pada tabel 6.

Tabel. 6. Hasil Evaluasi

	Akurasi	Recall	Precision
SVM	79.78%	69.05%	85.29%
SVM+PSO (Iterasi 50)	80.34%	76.83%	79.75%
SVM+PSO (Iterasi 100)	81.46%	73.17%	84.51%
SVM+PSO (Iterasi 150)	83.15%	75.61%	86.11%
SVM+PSO (Iterasi 200)	83.15%	75.61%	86.11%
SVM+PSO (Iterasi 250)	83.71%	78.05%	85.33%
SVM+PSO (Iterasi 300)	84.27%	76.83%	87.5%
SVM+PSO (Iterasi 350)	84.27%	81.71%	83.75%
SVM+PSO (Iterasi 400)	84.27%	73.17%	90.91%
SVM+PSO (Iterasi 450)	84.83%	75.61%	89.86%
SVM+PSO (Iterasi 500)	83.71%	73.17%	89.85%
SVM+PSO (Iterasi 550)	84.83%	75.61%	89.86%
SVM+PSO (Iterasi 600)	84.27%	78.05%	86.49%
SVM+PSO (Iterasi 650)	83.71%	75.61%	87.32%
SVM+PSO (Iterasi 700)	84.27%	85.37%	81.4%
SVM+PSO (Iterasi 750)	85.39%	81.71%	85.9%
SVM+PSO (Iterasi 800)	84.27%	79.27%	85.53%
SVM+PSO (Iterasi 850)	85.29%	80.49%	86.84%
SVM+PSO (Iterasi 900)	87.08%	76.83%	94.03%
SVM+PSO (Iterasi 950)	85.39%	76.83%	90%
SVM+PSO (Iterasi 1000)	84.83%	74.39%	91.04%

Berdasarkan hasil evaluasi tabel 6, dapat disimpulkan bahwa nilai akurasi terbesar terdapat pada SVM + PSO iterasi 900 dengan jumlah fitur sebanyak 910 didapatkan nilai akurasi 87.08%, recall 76.83%, dan precision 94.03%. Hasil evaluasi berdasarkan nilai tabel 7 confusion matrix.

Tabel. 7. Confusion Matrix SVM+PSO Iterasi 900

		Aktual	
		Positif	Negatif
Prrediksi	Positif	63 (TP)	4 (FP)
	Negatif	19 (FN)	92 (TN)

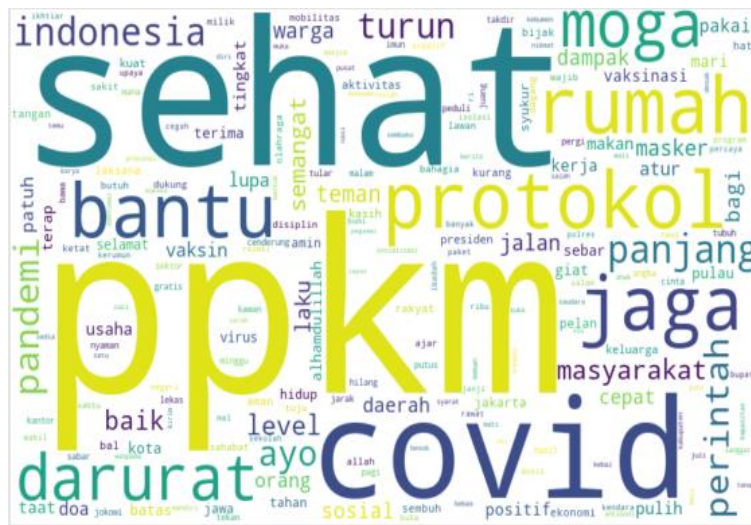
Berdasarkan tabel 7 confusion matrix diperoleh hasil prediksi model dengan data sebanyak 178 data. Memperoleh sebanyak 63 data diprediksi true positif (TP), 4 data diprediksi false positif (FP), 19 data diprediksi false negatif (FN), dan 92 data diprediksi true negatif (TN). Dari tabel x, perhitungan akurasi, recall, dan precision dengan persamaan (12), (13), dan (14) sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{63+92}{63+92+4+19} = \frac{155}{178} = 0.8708$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{63}{63+19} = \frac{63}{82} = 0.7683$$

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{63}{63+4} = \frac{63}{67} = 0.9403$$

Kemudian tahap terakhir penelitian ini melakukan visualisasi terhadap data berdasarkan kelas sentimen positif dan sentimen negatif pada opini pengguna twitter terhadap kebijakan PPKM. Visualisasi yang digunakan berupa wordcloud sentimen positif dan negatif menggunakan library matplotlib.pyplot dan wordcloud pada python.



Gambar. 3. Wordcloud Sentimen Positif

Berdasarkan gambar 3 menunjukkan bahwa kata yang sering muncul pada wordcloud sentimen positif yaitu kata “ppkm”, “covid”, “rumah”, “sehat”, dan lain-lain. Dapat disimpulkan sentimen positif mengenai penerapan kebijakan ppkm berkaitan dengan masyarakat dihimbau untuk tetap berada di rumah selama ppkm dengan menjaga kesehatan dan mematuhi protokol kesehatan.

- <https://covid19.go.id/p/ regulasi/instruksi-menteri-dalam-negeri-nomor-42-tahun-2021> (accessed Nov. 20, 2021).
- [3] M. S. Hadna, P. I. Santosa, and W. W. Winarno, “Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter,” *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 57–64, 2016, [Online]. Available: <https://fti.uajy.ac.id/sentika/publikasi/makalah/2016/95.pdf>.
- [4] P. Arsi, R. Wahyudi, and R. Waluyo, “Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021, doi: 10.29207/resti.v5i2.2698.
- [5] Nugraha, F. A., Harai, N. H., & Habibi, R. (2020). Analisis Sentimen Terhadap Pembatasan Sosial Menggunakan Deep Learning (R. M. Awangga (ed.)). Kreatif Industri Nusantara
- [6] C. C. Aggarwal, *Aggarwal - Mining Text Data*. 2012.
- [7] T. R. Nichols, P. M. Wisner, G. Cripe, and L. Gulabchand, “Putting the kappa statistic to use,” *Qual. Assur. J.*, vol. 13, no. 3–4, pp. 57–61, 2010, doi: 10.1002/qaj.481.
- [8] J. A. Zulqornain and P. P. Adikara, “Analisis Sentimen Tanggapan Masyarakat Aplikasi Tiktok Menggunakan Metode Naïve Bayes dan Categorical Proportional Difference (CPD),” vol. 5, no. 7, pp. 2886–2890, 2021.
- [9] A. Saputra, D. Rosiyadi, W. Gata, and S. Husain, “Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization,” *http://jurnal.iaii.or.id J.*, vol. 3, no. 3, pp. 377–382, 2019.
- [10] W. A. Luqyana, I. Cholissodin, and R. S. Perdana, “Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 11, pp. 4704–4713, 2018.
- [11] H. Tuhuteru, “Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berskala Besar Menggunakan Algoritma Support Vector Machine,” *Inf. Syst. Dev.*, vol. 5, no. 2, pp. 7–13, 2020.
- [12] R. D. Himawan and Eliyani, “Perbandingan Akurasi Analisis Sentimen Tweet terhadap Pemerintah Provinsi DKI Jakarta di Masa,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, pp. 58–63, 2021, [Online]. Available: <https://jurnal.untan.ac.id/>.
- [13] D. Wahyudi, T. Susyanto, and D. Nugroho, “Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia,” *J. Ilm. SINUS*, vol. 15, no. 2, pp. 49–56, 2017, doi: 10.30646/sinus.v15i2.305.
- [14] P. P. A. Arsyia Monica Pravina, Imam Cholissodin, “Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 2789–2797, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4793>.
- [15] A. Taufik, “Optimasi Particle Swarm Optimization Sebagai Seleksi Fitur Pada Analisis Sentimen Review Hotel Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes,” *J. Tek. Komput.*, vol. III, no. 2, pp. 40–47, 2017.
- [16] D. A. Pisner and D. M. Schnyer, *Support vector machine*. Elsevier Inc., 2019.
- [17] K. Srinivasa, G. Siddesh, and H. Srinidi, *Network Data Analytics: A Hands-On Approach for Application Development*. 2018.