

Analisis Sentimen Publik Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Skala Mikro Menggunakan Algoritma Support Vector Machine (Studi Kasus: Twitter)

Renas Madya Pradhana¹⁾ M.J. Dewiyani Sunarto^{1 2)} Julianto Lemantara^{2 3)}

Program Studi/Jurusan Sistem Informasi

Universitas Dinamika Surabaya

Jl. Raya Kedung Baru 98 Surabaya, 60298

Email : 1)17410100180@dinamika.ac.id, 2)dewiyani@dinamika.ac.id, 3)julianto@dinamika.ac.id

Abstract: *All countries in the world are experiencing difficult times due to the Covid-19 pandemic. Crises are emerging in the health and economic sectors, including Indonesia. The president requested that this crisis be addressed immediately and should be resolved simultaneously. Various efforts have been made by the government such as conducting PPKM Mikro. PPKM Mikro policy raises the pros-cons of society because of its impact on the economy and is ineffective in reducing daily cases of Covid-19. Public responses and opinions are conveyed through various media, more than 63.6% of total social media users in Indonesia have a Twitter account. Based on this background, this study hopes to use data from Twitter to understand the response and perception of Indonesian people to PPKM Mikro policy, from the prediction results, which can be a benchmark or evaluation material for the government. This study conducted a sentiment analysis by dividing people's responses into positive and negative sentiments using the Term Frequency Inverse Document Frequency algorithm for weighting and the Support Vector Machine method for classification. The results of the research that has been done showed the validation results using k-fold cross-validation on the support vector machine has an average cross-validation score of 96.42%. The test results showed an accuracy of 97.13%, precision of 97.58%, and recall of 99.15%. The overall percentage of sentiment analysis is for negative 12.8% and for positive 87.2%.*

Keywords: *Sentiment Analysis, PPKM Mikro, Term Frequency-Inverse Document Frequency, Support Vector Machine, Twitter.*

Presiden Joko Widodo mengatakan bahwa dalam satu tahun terakhir, semua negara di dunia mengalami masa-masa sulit akibat pandemi Covid-19. Krisis muncul di sektor kesehatan dan ekonomi, termasuk Indonesia. Presiden juga meminta agar krisis ini segera ditangani dan harus diselesaikan bersamaan (Farisa, 2021). Berbagai upaya telah dilakukan pemerintah untuk mencegah penyebaran kasus Corona (COVID-19). Mulai dari penerapan Pembatasan Sosial Berskala Besar (PSBB) hingga Penerapan Pembatasan Kegiatan Masyarakat (PPKM) bahkan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) skala mikro atau PPKM mikro. Kebijakan mulai dari PSBB hingga PPKM mikro menimbulkan pro-kontra dari masyarakat karena dampaknya terhadap perekonomian dan tidak efektif menurunkan kasus harian Covid-19. Berdasarkan data Satuan Tugas Penanganan Covid-19 pada Selasa (26/1/2021) pukul 12.00 WIB, per 2 Maret 2020 kasus pertama diumumkan,

Indonesia memiliki 1.012.350 kasus Covid-19. (Farisa, 2021). Selain itu, dibandingkan dengan enam bulan sebelumnya, kepercayaan konsumen terhadap pendapatan saat ini juga melemah karena penurunan pendapatan reguler seperti upah, remunerasi, dan omset (Elena, 2021). Masyarakat yang tinggal di Jawa-Bali juga terbagi. Meski pendapatan menurun, sekitar 44% warga di Bali, Jawa, memilih untuk menerima PPKM Mikro secara ketat. Walaupun risiko tertular Covid-19 meningkat, 46% warga tetap memilih untuk menghentikan PPKM Mikro (Makdori, 2021).

Kegiatan PPKM mikro tersebut haruslah mempertimbangkan segala aspek, mulai dari aspek untuk menekan angka kenaikan positif virus corona (Covid-19), hingga aspek untuk membuat roda perputaran ekonomi seperti pada sektor sektor pertanian, perdagangan, manufaktur, transportasi, dan jasa keuangan tetap berjalan. Semua aspek tersebut haruslah dipertimbangkan secara terperinci agar rencana

PPKM mikro dapat berjalan dengan baik dan terhindar dari hal-hal yang justru akan merugikan seperti peningkatan kasus aktif corona sampai aktivitas perekonomian terhenti yang menyebabkan kehilangan pekerjaan dan pendapatan terutama pada masyarakat menengah bawah yang berpenghasilan tidak tetap dan bekerja di sektor informal. Kegiatan PPKM mikro tersebut juga haruslah mempertimbangkan berbagai masukan, diantaranya adalah dengan melihat bagaimana respon dan opini masyarakat terhadap PPKM mikro. Hal ini sangat penting karena pemerintah dapat mempertimbangkan hal tersebut saat menyikapi sikap publik.

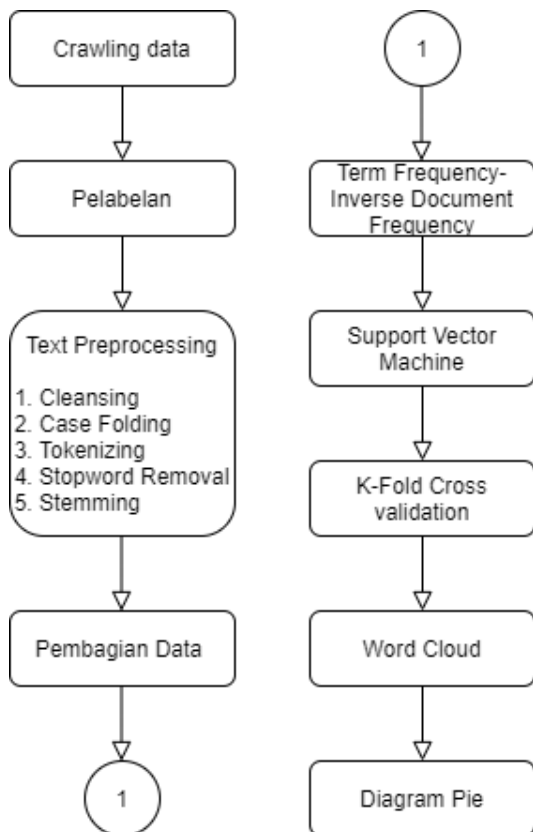
Tanggapan dan opini publik disampaikan melalui berbagai media. Salah satu media sosial yang paling banyak digunakan oleh masyarakat Indonesia adalah media sosial twitter, lebih dari 63,6% dari total pengguna media sosial di Indonesia memiliki akun Twitter (Kemp, 2021). Oleh sebab itu Twitter dapat digunakan secara efektif menjadi sumber penting opini publik karena berbasis teks untuk bahan penelitian terutama analisis sentimen. Dalam banyak penelitian sebelumnya, data twitter juga telah digunakan untuk menganalisis tanggapan dan opini publik. Misalnya penelitian yang dilakukan untuk melihat sentimen analisis terhadap tokoh publik (Pamungkas, 2018), sentimen analisis terhadap bom bunuh diri di Surabaya 13 Mei 2018 (Ihsan et al., 2019), perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen Twitter (Nasution & Hayaty, 2019), dan penelitian yang dilakukan untuk melihat sentimen analisis publik terhadap Joko Widodo terhadap wabah Covid-19 (Hikmawan et al., 2020), perbandingan akurasi metode Naïve Bayes sebesar 84.58%, Support Vector Machine sebesar 92.93%, dan k-NN sebesar 83.70%. Hasil precision dari Naïve Bayes sebesar 82.14%, SVM sebesar 95.70% dan k-NN sebesar 80.66%. Juga hasil recall dari Naïve Bayes sebesar 85.82%, SVM sebesar 89.17%, dan k-NN sebesar 84.13%.

Berdasarkan latar belakang tersebut, penelitian ini berharap dapat menggunakan data dari twitter untuk memahami respon dan persepsi masyarakat Indonesia terhadap kebijakan PPKM mikro, apakah masyarakat lebih banyak beranggapan positif atau negatif sehingga dari hasil prediksi tersebut dapat menjadi tolak ukur atau bahan evaluasi bagi pemerintah agar dapat memperpanjang atau

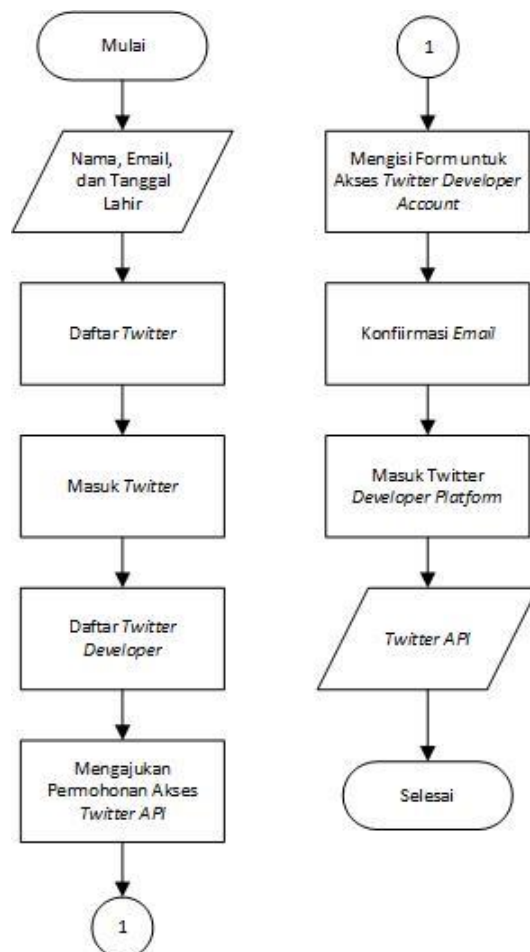
tidak kebijakan mengenai pemberlakuan pembatasan kegiatan masyarakat skala mikro sehingga dapat dilakukan penanganan ke evaluasi yang lebih baik. Untuk mengatasi masalah tersebut, studi ini akan melakukan analisis sentimen dengan membagi respon masyarakat menjadi sentimen positif dan negatif, serta menggunakan metode *Support Vector Machine* untuk mengklasifikasikan sentimen kebijakan PPKM mikro, dan menggunakan algoritma *Term Frequency Inverse Document Frequency* (TF-IDF) untuk pembobotan.

METODE PENELITIAN

Dalam penelitian ini untuk melakukan analisis sentimen publik terhadap kebijakan pemberlakuan pembatasan kegiatan masyarakat skala mikro menggunakan algoritma *support vector machine* pada twitter menggunakan *software* seperti Rapid Miner, Microsoft Excel 2016 dan Python dengan menggunakan Google Colaboratory. Rapid Miner digunakan untuk proses *crawling* data. Microsoft Excel 2016 digunakan untuk mengolah hasil *crawling* data twitter untuk pelabelan. *Python* digunakan untuk lima tahapan penelitian yang dilakukan dalam penelitian ini mulai dari tahap *text preprocessing*, hasil *crawling* data dibagi menjadi data *training* dan data *testing*. Data *training* yang telah dilabeli positif dan negative akan digunakan untuk melatih algoritma, sedangkan data *testing* akan digunakan untuk menguji kinerja algoritma yang dilatih sebelumnya saat menentukan data baru yang belum pernah terlihat sebelumnya. Lalu dilakukan pembobotan *term frequency-inverse document frequency*, klasifikasi menggunakan algoritma *support vector machine*, validasi *k-fold cross validation*, visualisasi data dengan *word cloud*, dan visualisasi data dengan diagram pie.



Gambar 1. Diagram Alir Klasifikasi Analisis Sentimen



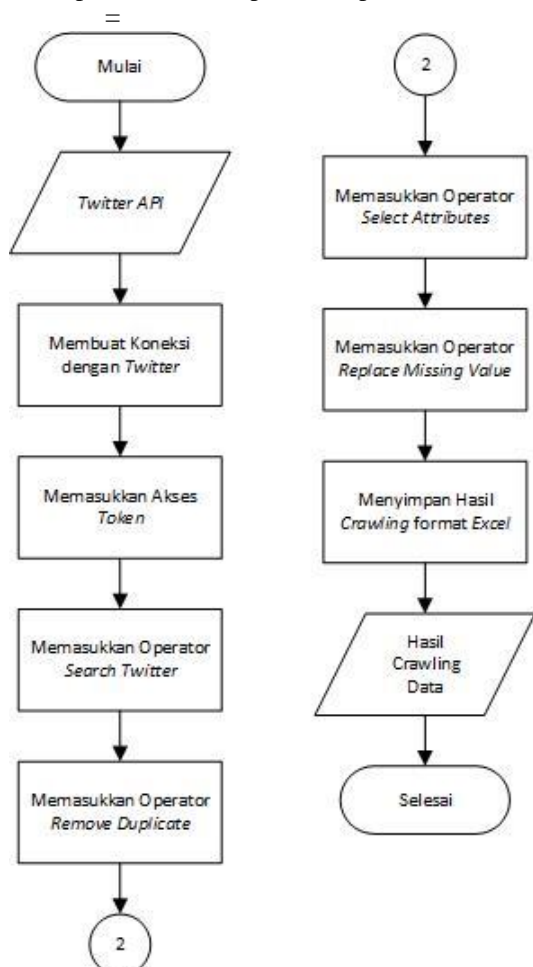
Gambar 1. Diagram Alir Mendapatkan Twitter API

Crawling Data

Teknik *crawling* data yang digunakan adalah teknik *crawling* data menggunakan Rapid miner. Data awal yang dibutuhkan pada penelitian ini adalah API Twitter, dengan melakukan login ke API Twitter Authorization. Untuk mengakses API Twitter memerlukan akun Twitter, aplikasi, dan token akses. Langkah pertama untuk mendapatkan API Twitter adalah mendaftar akun menggunakan nama, *email* dan tanggal lahir pada Twitter dan masuk kedalam Twitter menggunakan akun tersebut. Setelah dapat masuk kedalam Twitter maka buka *website* Twitter Developer untuk melakukan pendaftaran dan ajukan permohonan akses API Twitter dengan mengisi form untuk akses Twitter Developer Account. Lakukan konfirmasi email dan masuk ke Twitter Developer Platform, maka API Twitter berhasil didapatkan dan digunakan untuk mengambil data-data atau cuitan Twitter dengan kata kunci *ppkm* mikro yang dengan menggunakan Bahasa Indonesia.

Pengambilan data *tweet* pada Twitter menggunakan Rapid Miner adalah membuat koneksi dengan *connection type* yang dipilih yaitu Twitter, lalu *request access token* dengan membuka *url* yang telah disediakan dan pilih *authorize app* agar memberikan otorisasi Rapid Miner Social Media Extension untuk mengakses akun. Setelah mendapat otorisasi maka akan mendapatkan *code* 6 digit, *copy code* tersebut kedalam kolom yang telah disediakan di dalam Rapid Miner, Maka *access token* otomatis terisi. Setelah koneksi dapat tersambung maka langkah selanjutnya adalah memasukkan operator *search twitter* kedalam *process*, pada operator *search twitter* terdapat beberapa parameter yang harus diisi, untuk parameter *connection source* pilih *repository*, parameter *connection entry* pilih koneksi yang sudah dibuat, parameter *query* masukkan “*ppkm mikro*”, parameter *result type* pilih *recent or popular*, parameter *limit* atur

9999, parameter *language* masukkan "id", dan untuk parameter *until* pilih 13 April 2021.



Gambar 3. Diagram Alir Crawling Data

Pelabelan

Hasil *crawling* data berbentuk file excel yang didapatkan dari aplikasi Rapid Miner, dilakukan pelabelan data secara manual dengan bantuan sukarelawan. Pada tahapan ini dilakukan pemberian label negatif dan angka label positif.

Text Preprocessing

Hasil *crawling* data tersebut akan dimasukkan pada tahap *text preprocessing*. Namun sebelum melakukan tahapan tersebut terlebih dahulu memasukkan atau membaca data dengan format *excel* kedalam python menggunakan *library pandas*.

Cleansing

Proses *cleansing* menggunakan *library re* untuk menghilangkan komponen tertentu pada

cuitan seperti tanda baca, angka, *url*, *mention*, dan *hashtag*.

Case folding

Proses *cleansing* menggunakan *library re* dan *function remove punctuation* untuk penyatuan format huruf kecil dari semua teks dalam dokumen menjadi huruf kecil (*lowercase*).

Tokenizing

Proses *tokenizing* menggunakan *library nltk* untuk pemisahan atau membagi teks berupa kalimat pada dokumen menjadi *token* atau *term*.

Stopword Removal atau Filtering

Proses *stopword removal* atau *filtering* menggunakan *library sastrawi* untuk menghapus kata-kata sesuai dengan kata-kata yang terdapat dalam *stopword* atau *stoplist*.

Stemming

Proses *stemming* menggunakan *library sastrawi* untuk mengubah kata menjadi bentuk kata dasar.

Pembagian Data

Hasil *crawling* data Twitter dilakukan pembagian menjadi data training sebesar 70% dan data testing sebesar 30%, setelah dibagi maka melakukan pelabelan sentimen positif dan negatif.

Pembobotan Term Frequency-Inverse Document Frequency

Metode TF-IDF menurut Abdul (dikutip dalam Amrizal, 2018) merupakan metode untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval*. Pembobotan TF-IDF menggunakan *library tfidfvectorizer* untuk mengubah data yang berupa kata menjadi data numerik dan memiliki bobot, juga untuk menyaring data yang akan diproses. Algoritma TF-IDF digunakan rumus terhadap kata kunci dengan rumus yaitu:

$$W_{dt} = t_{fdt} * I_{dft} \tag{1}$$

Dimana:

W_{dt} = bobot dokumen ke- *d* terhadap kata ke- *t*
t_{fdt} = banyaknya kata yang dicari pada sebuah dokumen

I_{dft} = Inversed Document Frequency ($\log(N/df)$)

N = total dokumen

df = banyak dokumen yang mengandung kata yang dicari.

Klasifikasi Support Vector Machine

Support Vector Machine (SVM) Metode klasifikasi yang digunakan Pembelajaran mesin (*supervised learning*) adalah memprediksi kategori berdasarkan model atau pola hasil dari proses *training*. Untuk klasifikasi Temukan *hyperplane* atau bidang pemisah (*decision boundary*) dengan kelas lain, dalam hal ini bidang pemisah Berperan dalam memisahkan *tweet* sentimen positif (ditandai sebagai +1) *tweet* negatif (ditandai sebagai -1). *Hyperplane* yang baik bisa didapatkan dengan memaksimalkan jarak *margin*. Berikut, diasumsikan jika pada kelas -1 dan +1 dapat terpisah secara sempurna oleh *hyperplane*, yang dimana dapat diberikan sebuah definisi :

$$w_i \cdot x_i + b = 0 \tag{2}$$

Jika w_i berada di kelas +1, maka dapat dituliskan sebuah definisi berikut ini :

$$w_i \cdot x_i + b \geq +1 \tag{3}$$

Sedangkan, Jika w_i berada di kelas -1, maka dapat dituliskan sebuah definisi berikut ini :

$$w_i \cdot x_i + b \leq -1 \tag{4}$$

Pada permasalahan dalam pengklasifikasian SVM *non-linear* ada beberapa fungsi *kernel* yang umum digunakan yaitu:

a. *Kernel Linear*

$$K(x,y) = x \cdot y \tag{5}$$

Proses klasifikasi *support vector machine* menggunakan *library scikit-learn*.

Validasi K-fold Cross Validation

Teknik validasi yang digunakan dalam penelitian ini adalah *k-fold cross-validation*. Pembagian data dilakukan sehingga ukuran setiap *subset (fold)* sama. Proses *k-fold cross-validation* menghasilkan nilai yang disebut *performance value*. Evaluasi performansi adalah menguji hasil klasifikasi dengan mengukur nilai kinerja dari sistem yang dibuat. Evaluasi di atas menggunakan parameter uji berupa akurasi yang perhitungannya diperoleh dari tabel *confusion matrix* (Anjasmoros, 2020). Sama seperti proses klasifikasi *support vector machine*, validasi *k-fold cross validation* juga menggunakan *library scikit-learn*.

Tabel 1 Confusion Matrix

Dokumen	Kelas Aktual	
	Negative	Positive

Kelas Prediksi	True Negative (TN)	False Positive (FP)
	False Negative (FN)	True Positive (TP)

Keterangan :

- a. True Positive (TP) = jumlah dokumen aktual yang *positive* dan diprediksi *positive*.
- b. False Positive (FP) = jumlah dokumen aktual yang *negative* dan diprediksi *positive*.
- c. False Negative (FN) = jumlah dokumen aktual yang *positive* dan diprediksi *negative*.
- d. True Negative (TN) = Jumlah dokumen aktual yang *negative* dan diprediksi *negative*.

Berdasarkan *matrix confusion*, maka dapat dihitung nilai *accuracy*, *precision*, dan *recall*. *Accuracy* menggambarkan seberapa akurat model dalam mengklasifikasikan dengan benar dengan rumus sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

Precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model. dengan rumus sebagai berikut:

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

Recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.dengan rumus sebagai berikut:

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

Visualisasi Data dengan Word Cloud

Pada tahap visualisasi data dengan *word cloud* menggunakan *library wordcloud*, dan akan menghasilkan kata yang paling sering muncul dalam text yang telah dianalisis sebelumnya.

Visualisasi Data dengan Diagram Pie

Pada tahap visualisasi dengan diagram *pie* menggunakan *library matplotlib* guna menunjukkan persentase sentimen positif dan negatif yang diberikan oleh penggunaanya berbentuk lingkaran.

HASIL DAN PEMBAHASAN

Teknik pengolahan data yang digunakan adalah teknik *crawling* data menggunakan *Rapid miner*. Akses API Twitter digunakan untuk mengambil data atau cuitan Twitter dengan kata kunci *ppkm* mikro pada

tanggal 7 April 2021 hingga 13 April 2021. Jumlah crawling data dari Twitter mengenai cuitan ppk mikro yang dengan menggunakan Bahasa Indonesia sebesar 4066 cuitan. Berikut hasil crawling data rapid miner yang ditampilkan pada Gambar 4.

Row No. ↑	Text	id
1	Aipda Aftoni selaku bhabinkamtibmas dan juga petug...	1382844922...
2	Dengan cara pendekatan persuasif, petugas PPKM Mi...	1382843813...
3	@PEDOMAN_id: Provinsi Prioritas PPKM Mikro Dimint...	1382843110...
4	Petugas PPKM Mikro Kelurahan Purwoyoso memberik...	1382841597...
5	Bersama Dukung PPKM Berbasis Mikro dengan 5M.	1382835971...

Gambar 4. Crawling Data Rapid Miner

Pelabelan

Pada pelabelan dilakukan pemberian label negatif dan angka label positif. pelabelan ini dikerjakan dengan perangkat pembantu yaitu Microsoft Excel 2016. data dimasukkan menggunakan library *pandas*. Berikut adalah hasil tahap pelabelan yang ditampilkan pada Gambar 5.

Text	id	label
Aipda Aftoni selaku bhabinkamtibmas dan juga petugas	1382844922	positif
Dengan cara pendekatan persuasif, petugas PPKM Mikrc	1382843813	positif
@PEDOMAN_id: Provinsi Prioritas PPKM Mikro Diminta	1382843110	positif
Petugas PPKM Mikro Kelurahan Purwoyoso memberik	1382841597	positif
Bersama Dukung PPKM Berbasis Mikro dengan 5M. -M	1382835971	positif

Gambar 5. Pelabelan Positif Negatif

Text Preprocessing

Pada Tahap text preprocessing menggunakan Google Colaboratory Python Sebelum masuk kedalam tahap text preprocessing data terlebih dimasukkan ke dalam Google Colaboratory Python. Setelah dimasukkan data maka menghapus kolom yang tidak digunakan yaitu id dan merubah label positif menjadi 1 dan label negative menjadi 0 untuk proses *term frequency-inverse document frequency*. Berikut adalah data yang akan digunakan pada tahap *text preprocessing* yang ditampilkan pada Gambar 6.

	Text	label
0	Aipda Aftoni selaku bhabinkamtibmas dan juga p...	1
1	Dengan cara pendekatan persuasif, petugas PPKM...	1
2	@PEDOMAN_id: Provinsi Prioritas PPKM Mikro Di...	1
3	Petugas PPKM Mikro Kelurahan Purwoyoso memberi...	1
4	Bersama Dukung PPKM Berbasis Mikro dengan 5M.\...	1

Gambar 6. Data Text Preprocessing

Text preprocessing melakukan pengolahan terhadap data *training* yang bertujuan untuk mengubah data yang

sebelumnya tidak terstruktur menjadi data terstruktur untuk tahap pemrosesan selanjutnya. Pada *python* proses *cleansing* dilakukan menggunakan library *string* dan *re*. Tahapan *text preprocessing* dalam penelitian ini meliputi *cleansing*, *case folding*, *filtering*, *tokenizing*, *stopword removal*, dan *stemming*. Adapun langkah pertama yang dilakukan adalah *cleansing*, berikut adalah hasil tahap *cleansing* yang ditampilkan pada Gambar 7.

	Text	label
0	Aipda Aftoni selaku bhabinkamtibmas dan juga p...	1
1	Dengan cara pendekatan persuasif petugas PPKM...	1
2	id Provinsi Prioritas PPKM Mikro Diminta Se...	1
3	Petugas PPKM Mikro Kelurahan Purwoyoso memberi...	1
4	Bersama Dukung PPKM Berbasis Mikro dengan M ...	1

Gambar 7. Cleansing

Tahap *case folding* bertujuan untuk proses penyatuan format huruf kecil dari semua teks dalam dokumen menjadi huruf kecil (*lowercase*). Berikut adalah hasil tahap *case folding* yang ditampilkan pada Gambar 8.

	Text	label
0	aipda aftoni selaku bhabinkamtibmas dan juga p...	1
1	dengan cara pendekatan persuasif petugas ppkm...	1
2	id provinsi prioritas ppkm mikro diminta se...	1
3	petugas ppkm mikro kelurahan purwoyoso memberi...	1
4	bersama dukung ppkm berbasis mikro dengan m ...	1

Gambar 8. Case Folding

Tahap *tokenizing* bertujuan memisahkan atau membagi teks berupa kalimat pada dokumen menjadi *token* atau *term*. Pada *python* proses *parsing/tokenizing* dilakukan menggunakan library *nlTK*. Berikut adalah hasil tahap *tokenizing* yang ditampilkan pada Gambar 9.

	Text	label
0	[aipda, aftoni, selaku, bhabinkamtibmas, dan, ...	1
1	[dengan, cara, pendekatan, persuasif, petugas,...	1
2	[id, provinsi, prioritas, ppkm, mikro, diminta...	1
3	[petugas, ppkm, mikro, kelurahan, purwoyoso, m...	1
4	[bersama, dukung, ppkm, berbasis, mikro, denga...	1

Gambar 9. Hasil Tahap Tokenizing

Tahap *stopword removal* bertujuan menghapus kata-kata yang tidak penting dalam deskripsi dengan memeriksa apakah kata-kata hasil *parsing* dalam deskripsi yang termasuk

dalam daftar kata tidak penting (*stopword*) Pada *python* proses *filtering* dilakukan menggunakan library *Sastrawi* dengan fungsi *StopWordRemover*. Dengan menggunakan daftar *Stopwords*, maka setiap kata dalam kumpulan akan dicocokkan dengan kata yang sudah ada. Berikut adalah hasil tahap *stopword removal* yang ditampilkan pada Gambar 10.

	Text	label
0	aipda aftoni bhabin kamtibmas petugas pphk mikr...	1
1	pendekatan persuasif petugas pphk mikro mengaj...	1
2	id provinsi prioritas pphk mikro serius pemben...	1
3	petugas pphk mikro kelurahan purwoyoso edukasi...	1
4	dukung pphk berbasis mikro mencuci tangan mema...	1

Gambar 10. *Stopword Removal*

Tahap *stemming* bertujuan untuk mengubah bentuk kata menjadi akar kata dasar sesuai dengan struktur morfologi bahasa Indonesia yang baik dan benar. Pada *python* proses *stemming* dilakukan menggunakan library *Sastrawi* dengan fungsi *Stemmer*. Berikut adalah hasil tahap *stemming* yang ditampilkan pada Gambar 11.

	Text	label
0	aipda aftoni bhabin kamtibmas tugas pphk mikro ...	1
1	dekat persuasif tugas pphk mikro ajak masyarak...	1
2	id provinsi prioritas pphk mikro serius bentuk...	1
3	tugas pphk mikro lurah purwoyoso edukasi proto...	1
4	dukung pphk bas mikro cuci tangan pakai masker...	1

Gambar 11. *Stemming*

Hasil *crawling* data dari Twitter mengenai cuitan pphk mikro dengan menggunakan Bahasa Indonesia sebesar 4066 cuitan dilakukan pembagian data dengan membagi data menjadi data *training* sebanyak 70% dan data *testing* sebanyak 30%. Pada tahap ini dapat diketahui data *training* memiliki data sebesar 2846 cuitan yang ditampilkan pada Gambar 12. Data *testing* sebesar 1.220 cuitan yang ditampilkan pada Gambar 13.

2804	gubernur keluar surat edar kab kota sumsel lak...
1482	ka jawat aman kap kasihan bpk drs petrus santo...
1311	pphk mikro kabupaten bogor positif sembuh aktif
3447	gusmenteri grafik positif covid pphk mikro des...
3560	giat tengah laku pphk skala mikro haramain gus...
	...
1594	tugas pphk mikro polsek gayamsari himbauan dag...
2448	efektif pphk mikro tuk halau sebar covid haramain
942	bhabin kamtibmas desa tambak ukir polsek kendit...
874	laksana himbauan sosialisasi protokol sehat pp...
606	patroli posko pphk mikro kel wonorejo surabaya

Gambar 12. *Data Training*

2522	polres pematangsiantar laksana operasi yustis...
3133	kepk ipda supandi pimpin operasi yustisi tni p...
1656	protokol sehat cegah sebar virus covid edukasi...
1290	babinsa koramil tellu siatting serda muh asdar...
3706	polres kebumen polsek jajar rutin gelar operas...
	...
1709	kapolda jatim tinjau laksana pphk mikro dusun ...
585	giat laksana pphk skala mikro wilayah kodim pr...
3159	darnramil rambipuji kapten chb mulyadi konfirm...
2398	giat ops yustisi tega disiplin patuh protokol ...
3962	grafik positif covid pphk mikro desa labur har...

Gambar 13. *Data Testing*

Pembobotan Term Frequency-Inverse Document Frequency

Tahap pembobotan *term frequency-inverse document frequency* melakukan pengolahan terhadap data *training* yang bertujuan untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval*. sebuah ukuran statistik yang digunakan untuk mengevaluasi seberapa penting sebuah kata di dalam sebuah dokumen atau dalam sekelompok kata. Untuk dokumen tunggal tiap kalimat dianggap sebagai dokumen. Frekuensi dokumen yang mengandung kata tersebut menunjukkan seberapa umum kata tersebut. Kata yang memiliki bobot akan muncul pada hasil pembobotan *term frequency-inverse document frequency* tiap kalimat. Bobot kata semakin besar jika sering muncul dalam suatu dokumen dan semakin kecil jika muncul dalam banyak dokumen. Pembobotan data pada *python* menggunakan library *sklearn*. Berikut adalah hasil dari pembobotan TF-IDF yang ditampilkan pada Gambar 14.

(0, 1870)	0.058772674052201754
(0, 2781)	0.3565982761355653
(0, 2433)	0.05875203042496681
(0, 1607)	0.18808155619493622
(0, 3054)	0.36430151732755905
(0, 1535)	0.21794016871704716
(0, 1256)	0.21633313957436345
(0, 737)	0.3965849774531933
(0, 3080)	0.3965849774531933

Gambar 14. Hasil dari Pembobotan TF-IDF

Klasifikasi Support Vector Machine

Support Vector Machine (SVM) Metode klasifikasi yang digunakan pembelajaran mesin (*supervised learning*) adalah memprediksi kategori berdasarkan model atau pola hasil dari proses *training*. Data *training* dan label data *training* digunakan sebagai parameter untuk

- 167–176.
<https://doi.org/10.31599/jki.v20i2.117>
- Ihsan, M., Roza, E., & Widodo, E. (2019). Analisis Sentimen Twitter terhadap Bom Bunuh Diri di Surabaya 13 Mei 2018 menggunakan Pendekatan Support Vector Machine. *Prisma 2 (2019): 416-426*, 2, 416–426.
- Mochamad Tri Anjasmoros, Istidadi, dan F. M. (2020). *Analisis sentimen aplikasi go-jek menggunakan metode svm dan nbc (studi kasus: komentar pada play store) 1*. *Ciastech*, 489–498.
- Pamungkas, I. T. S. A. (2018). Analisis Sentimen Terhadap Tokoh Publik Menggunakan Algoritma Support Vector Machine (Svm). *Log!K@*, 8(1), 69–79.
- Elena, M. (2021, 02 08). *Dampak PPKM Berlanjut, Masyarakat Kian Pesimis Lapangan Kerja dan Penghasilan Pulih*. Retrieved 03 17, 2021, from *Bisnis.com*: <https://ekonomi.bisnis.com/read/20210208/9/1353653/dampak-ppkm-berlanjut-masyarakat-kian-pesimis-lapangan-kerja-dan-penghasilan-pulih>
- Farisa, F. C. (2021, 01 27). *1 Juta Kasus Covid-19 dan Respons Pemerintah...* Retrieved 03 17, 2021, from *KOMPAS.com*: <https://nasional.kompas.com/read/2021/01/27/06405851/1-juta-kasus-covid-19-dan-respons-pemerintah?page=all>
- Farisa, F. C. (2021, 01 27). *Jokowi: Kesehatan dan Ekonomi Sama Penting, Harus Diselesaikan Bersamaan*. (www.kompas.com) Retrieved 03 15, 2021, from *KOMPAS.com*: <https://nasional.kompas.com/read/2021/01/27/13050811/jokowi-kesehatan-dan-ekonomi-sama-penting-harus-diselesaikan-bersamaan>
- Kemp, S. (2021, February 11). *Digital 2021: Indonesia*. Retrieved Maret 03, 2021, from *DATAREPORTAL*: <https://datareportal.com/reports/digital-2021-indonesia?rq=indonesia>
- Makdori, Y. (2021, 03 23). *Survei SMRC: Masyarakat Terbelah soal PPKM Mikro, Ada yang Pro dan Kontra*. Retrieved 04 20, 2021, from *Liputan6.com*: <https://www.liputan6.com/news/read/4513857/survei-smrc-masyarakat-terbelah-soal-ppkm-mikro-ada-yang-pro-dan-kontra>
- Nasution, M. R., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *Jurnal Informatika*, 226-235.