

ANALISIS SENTIMEN PADA PEMBERLAKUAN PEMBATASAN KEGIATAN MASYARAKAT MENGGUNAKAN ALGORITMA DBSCAN

1st Ferdian Wiharmariyansyah
Universitas Buana Perjuangan Karawang
Karawang, Indonesia
Ferdianwiharmariyansyah@mhs.ubpkara
wang.ac.id & 082297697169

2nd Anis Fitri Nur Masruriyah
Universitas Buana Perjuangan Karawang
Karawang, Indonesia
anis.masruriyah@ubpkarawang.ac.id

3rd Hilda Yulia Novita, M.Kom
Universitas Buana Perjuangan Karawang
Karawang, Indonesia
hilda.yulia@karawang.ac.id

Abstract— Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) adalah kebijakan yang dibuat oleh pemerintah untuk menghadapi pandemi *Corona Virus* (COVID-19). PPKM berfungsi untuk mengontrol kegiatan masyarakat sesuai dengan protokol kesehatan, di antaranya adalah menjaga jarak sejauh dua meter, memakai masker, mencuci tangan dan lain-lain dalam pencegahan penyebaran COVID-19. PPKM di Indonesia diberlakukan sejak 11 Januari 2021 di berbagai wilayah di Indonesia. Dampak dari PPKM yang dirasakan oleh masyarakat salah satunya kesulitan mencari pekerjaan, berkurangnya mata pencaharian masyarakat, membatasi kegiatan berkerumun, adanya adaptasi kebiasaan baru. Selain itu, pelaku UMKM selain mengalami penurunan omset juga memperoleh sanksi pidana denda di tengah PPKM darurat oleh pemerintah setempat. Pemerintah menerapkan PPKM demi menanggulangi penyebaran virus COVID-19. Akan tetapi, penerapan PPKM mengundang pro dan kontra di masyarakat. Terlihat pada opini masyarakat yang tersebar di postingan media sosial pengguna twitter. Pengguna jejaring media sosial twitter yang membahas mengenai COVID-19. Analisis sentimen di sosial media twitter dapat melihat pola polarisasi. Hasil yang didapatkan dievaluasi menggunakan *Sum of Square Error*.

Kata kunci — covid-19, ppkm, twitter, sentimen, analisis, text mining

I. PENDAHULUAN

Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) adalah kebijakan dibuat oleh pemerintah untuk menghadapi pandemi *Corona Virus* (COVID-19)[1]. PPKM berfungsi untuk mengontrol kegiatan masyarakat sesuai dengan protokol kesehatan, di antaranya adalah menjaga jarak sejauh dua meter, memakai masker, mencuci tangan dan lain-lain dalam pencegahan penyebaran COVID-19[2]. Adanya PPKM menimbulkan peraturan baru seperti pada sektor pendidikan, transportasi, perdagangan dan lain – lain[3]. PPKM memiliki tingkatan level yang berbeda setiap daerah yang di antaranya level satu sampai dengan level empat. Hingga saat ini PPKM menuai tanggapan positif dan negatif dari masyarakat di dunia maya. Oleh sebab itu menyebabkan berbagai opini disampaikan masyarakat oleh pemberlakuan pembatasan kegiatan masyarakat di media sosial twitter.

Media sosial twitter merupakan perkembangan teknologi yang menggunakan internet untuk komunikasi, saling berbagi informasi dan membentuk jaringan secara *online*. Pengguna sosial media yang umum dipakai bagi masyarakat sebab perubahan teknologi berkembang sangat pesat[4]. Sosial media tempat media komunikasi namun saja dipakai mendapatkan informasi sebagai tempat masyarakat untuk berpendapat[5]. Oleh karena itu pemberlakuan pembatasan kegiatan masyarakat mengadakan sebagian besar masyarakat khawatir lantaran banyak informasi beredar di sosial media khususnya twitter bahwa pemberlakuan pembatasan kegiatan masyarakat akan di perpanjang menjadi level 3. Oleh karena itu, pada penelitian tersebut ingin menganalisis jumlah respon positif, negatif atau netral yang habis disampaikan masyarakat terhadap pemberlakuan pembatasan kegiatan masyarakat yang terdapat pada di media sosial *twitter* menggunakan metode *Algorithm Tidytext*[6]. Solusi yang diberikan yaitu penelitian tersebut memberikan gambar visualisasi data yang didapatkan melalui Bahasa Pemrograman R atau Rstudio.

II. DATA DAN METODE

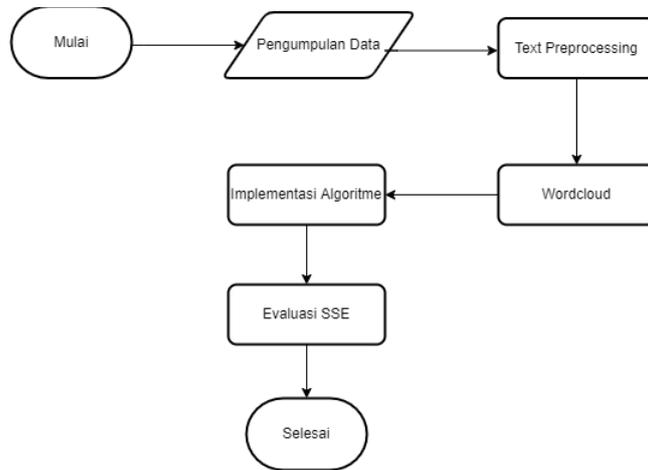
A. Bahan dan Peralatan

Data digunakan untuk melaksanakan sentimen analisis pada penelitian data twitter bulan Januari 2022 yang memuat informasi kebijakan pemerintah terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM). Selain itu, terdapat bahan yang dijadikan.

- 1) Perangkat Keras
 - Laptop dengan spesifikasi AMD A4-6210 APU with AMD Radeon R3 Graphics , 4 GB RAM.
- 2) Perangkat Lunak
 - Rstudio
 - *Microsoft Office*

B. Prosedur Penelitian

Tahapan diawali menggunakan analisis data perlu menerima informasi lalu pengumpulan data untuk memperoleh data dari *twitter*. Selanjutnya, data *twitter* akan diproses menggunakan *text preprocessing* buat melakukan pengelompokkan kata. Setelah itu data di klasifikasikan menggunakan algoritma DBSCAN mencapai langkah akhir yaitu Evaluasi. Prosedur penelitian.



Gambar 1 Prosedur Penelitian

1. Analisis data dilakukan untuk menemukan dan memperoleh sentimen positif, negatif dan netral.
2. Pengumpulan data diperoleh sejak 05 januari 2022 dengan lebih dari 1200 data.
3. *text preprocessing* suatu proses pengubahan bentuk data yang belum terstruktur menjadi data yang terstruktur sesuai dengan kebutuhan *clustering* menggunakan DBSCAN.
4. Algoritma DBSCAN metode *clustering* yang berbasis kepadatan (*density-based*) dengan prinsip pengelompokkan data untuk mencari kata tertentu dari sebuah dokumen
5. Pengujian yaitu untuk mengetahui semua fungsi berkerja dengan baik atau tidak, tujuan dilakukannya pengujian ini mendapatkan tingkat akurasi menggunakan *Sum Square of Error* (SSE) dari hasil prediksi yang akan diperoleh

C. Akuisisi Data

Data didapatkan dari media sosial twitter lalu diakuisi data pokok referensi waktu mengadakan proses analisis sentimen memiliki 1249 data. Data PPKM merupakan tahapan dengan sistem mencari kata sering muncul, sentimen positif, negatif dan netral. Berikut data awal PPKM telah dimasukkan dalam sistem.

Tabel 1. Data awal PPKM

No	KALIMAT
1.	“WAW BALIKPAPAN MASIH PPKM LEVEL 3 JELANG RAMADHAN, SHOLAT TARAWIH DI MASJID DIIZINKAN?”
2.	“@KEVINKAMALKK PPKM 3 ATAU 4 GITU VIN BIAR JALANAN SEPI WKWK”
3.	“ANGGOTA KORAMIL 07/NGRAMPAL MELAKSANAKAN PATROLI PPKM LEVEL III DI WILAYAH KECAMATAN NGRAMPAL HTTPS://T.CO/RTPWQ7ZBHJ ”
4.	“MEDAN - BRIPKA DR AFRIADI AMIN S.PSI. M.KOM.I DAN BRIPKA SURYANDA MEMBERIKAN HIMBAUAN KEPADA MASYARAKAT AGAR TETAP... HTTPS://T.CO/CKZVGFTS0M ”
5.	“@ROYBERKELANASFC @DENNYSIREGAR7 GILIRAN PPKM DIRUMAH SAJA PADA TERIAK-TERIAK,GAK USAH SOK BERAGAMA ENTE.APA YG HARU... HTTPS://T.CO/WSLHEARN8J ”
.	.
.	.
.	.
1245.	“RT @SAMARTEMARAM: GAMPANG, PPKM LEVEL 4 LAGI AJA, DIJAMIN BIRU ITU LANGIT JAKARTA.”
1246.	“KETAHUI INFORMASI PEMBAGIAN #PPKM DI WILAYAH JABAR BERDASARKAN LEVEL 4, 3, 2 DAN 1 DI #PIKODATA”
1247.	“KETAHUI INFORMASI PEMBAGIAN #PPKM DI WILAYAH JABAR BERDASARKAN LEVEL 4, 3, DAN 2 DI”
1248.	” HALO SOBAT SEHAT PANDEMI COVID-19 DI TANAH AIR SAAT INI MASIH TERKENDALI PADA LEVEL YANG RENDAH. MESKIPUN DEMIKIAN...”
1249.	“12 KASUS COVID-19, BULUNGAN KEMBALI STATUS PPKM LEVEL II #PPKM #COVID19 #BULUNGAN #KALTARA #BERITAKALTARA... HTTPS://T.CO/TYIAUVPGY8 ”

D. Text Preprocessing

Tahap *text processing* berguna merubah data tekstual yang bukan terstruktur kedalam data terstruktur disimpan dalam basis data. Tahapan pada text mining merupakan text preprocessing berniat memperoleh bacaan menjadi data terstruktur. Tahapan yang dilakukan diantaranya:

- 1) *Case folding*
Proses mengubah huruf seluruh huruf menjadi kecil.
- 2) *Tokenizing*
Tokenizing menghilangkan angka, tanda baca dan karakter lain yang tidak memiliki pengaruh terhadap pemrosesan teks.
- 3) *Filtering*
Filtering yang digunakan untuk membuang kata tidak penting dan mengambil kata penting.
- 4) *TF-IDF*
TF-IDF tahapan terakhir dalam proses *text processing* untuk melakukan pembobotan kata Hasil didapat akan proses implementasi algoritme mendapat data pencilan yang terdapat pada dokumen.

E. Algoritma DBSCAN

Algoritma DBSCAN metode *clustering* berbasis kerapatan (*density-based*) dengan prinsip pengelompokkan data untuk mencari kata tertentu dari sebuah dokumen. DBSCAN berguna untuk melakukan *clustering* data banyak mengandung *noise*. DBSCAN setiap titik sebuah *cluster*, *neighborhood* merupakan radius yang menyimpan setidaknya jumlah minimum poin, yaitu kerapatan *neighborhood* perlu melebihi sebagian *threshold* ditetapkan. Algoritma DBSCAN memerlukan dua parameter masukan buat mampu melaksanakan *clustering* yaitu :

1. *Eps*, radius menentukan batas daerah *neighborhood* dari titik (*Eps-neighborhood*)
2. *MinPts*, memiliki minimum poin yang harus ada *Eps-neighborhood*

Urutan algoritma DBSCAN memiliki 5 langkah yaitu :

1. Memiliki point p
2. Memiliki semua point *density reachable*
3. P adalah *core point* cluster terjadi
4. Jika p border *point*, bukan menjadi hubungan *desity-reachable* dari p dan DBSCAN akan bakal menghadiri *point* selanjutnya dari database.
5. Lanjutan proses sampai semua *point* telah diproses

III. HASIL DAN PEMBAHASAN

Penelitian ini memiliki data *text pre processing* mendapatkan perolehan *clustering* memerlukan algoritma DBSCAN terdapat hasil sentimen positif, negatif dan netral.

A. Hasil Text Pre processing

Data diolah menggunakan *pre processing* diperoleh atas proses implementasi twitter kedalam program mempunyai 1249 data. Tahapan text preprocessing diantaranya *To lower*, *Tokenizing*, *filtering*. Hasil text preprocessing mendapatkan data pada tabel 2.

Tabel 2. Hasil *text preprocessing*

No	Proses	Sebelum	Sesudah
1	<i>To lower</i>	“Waw Balikpapan Masih PPKM Level 3 Jelang Ramadhan, Sholat Tarawih di Masjid Diizinkan” “Ini mau PPKM lagi ya? Minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya. Huhuhu”	“waw Balikpapan masih ppkm level 3 jelang ramadhan, sholat tarawih di masjid diizinkan” "ini mau ppkm lagi ya? minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya. huhuhu"

No	Proses	Sebelum	Sesudah
2	Tokenizing	<p>“Waw Balikpapan Masih PPKM Level 3 Jelang Ramadhan, Sholat Tarawih di Masjid Diizinkan”</p> <p>“Ini mau PPKM lagi ya? Minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya. Huhuhu”</p>	<p>“waw balikpapan masih ppkm level 3 jelang ramadhan sholat tarawih di masjid diizinkan”</p> <p>“ini mau ppkm lagi ya minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya huhuhu”</p>
3	Filtering	<p>“Waw Balikpapan Masih PPKM Level 3 Jelang Ramadhan, Sholat Tarawih di Masjid Diizinkan”</p> <p>“Ini mau PPKM lagi ya? Minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya. Huhuhu”</p>	<p>“waw balikpapan masih ppkm level jelang ramadhan sholat tarawih di masjid diizinkan”</p> <p>“ini mau ppkm lagi ya minggu depan tiket kereta lokal gak tersedia dan masih disuruh nunggu kabar selanjutnya huhuhu”</p>

B. Implementasi Algoritme

Data yang diperoleh implementasi algoritma merupakan data twitter pada bulan Januari 2022 dalam melakukan analisis sentimen. Contoh data dipakai mendapatkan klasifikasi algoritme DBSCAN berjumlah 1249 data. Tahapan menggunakan tidy text dengan sepuluh sampel data pada gambar 2.

no	keterangan
1	waw balikpapan masih ppkm level jelang ramadhan sholat t...
2	ppkm atau gitu vin biar jalanan sepi wkwk
3	anggota koramil ngrampal melaksanakan patroli ppkm level...
4	medan briпка dr afriadi amin spsi mkomi dan briпка suryan...
5	giliran ppkm dirumah saja pada teriakteriakgak usah sok ber...
6	alhamdulillah jadi sekarang udah boleh berkumpul tanpa di...
7	pelaksanaan yustisi dalam rangka pemberlakuan ppkm level...
8	rt ppkm level menjelang ramadhan tidak diperbolehkan ngo...
9	mandatori vaksin covid id masker lockdown ppkm dan sunti...
10	anggota sabhara polsek jatinegara bergabung dgn pilar di p...

Gambar 2. Data awal twitter

no	baris	word
1	3	1 masih
2	4	1 ppkm
3	5	1 level
4	6	1 jelang
5	7	1 ramadhan
6	9	1 tarawih
7	10	1 di
8	11	1 masjid
9	13	2 ppkm
10	14	2 atau

Gambar 3. Data pengelompokan tidy text

Selanjutnya proses *tidytext* yaitu menemukan nilai disetiap kata yang sering muncul pada data yang digunakan. Kata sering muncul terdapat pada “ppkm” dengan nilai sebanyak 902 kata. Terdapat data sampel digunakan sebanyak sepuluh data.

word	n
1 ppkm	902
2 level	392
3 patroli	289
4 di	286
5 dalam	247
6 dan	219
7 giat	216
8 rangka	189
9 koramil	174
10 rt	170

Gambar 4 Frekuensi data tidytext

Setelah dilakukan tahapan seterusnya analisis sentimen mulai masing – masing kata ditunjukkan pada gambar 5.

negative	33	negative
positive	465	positive

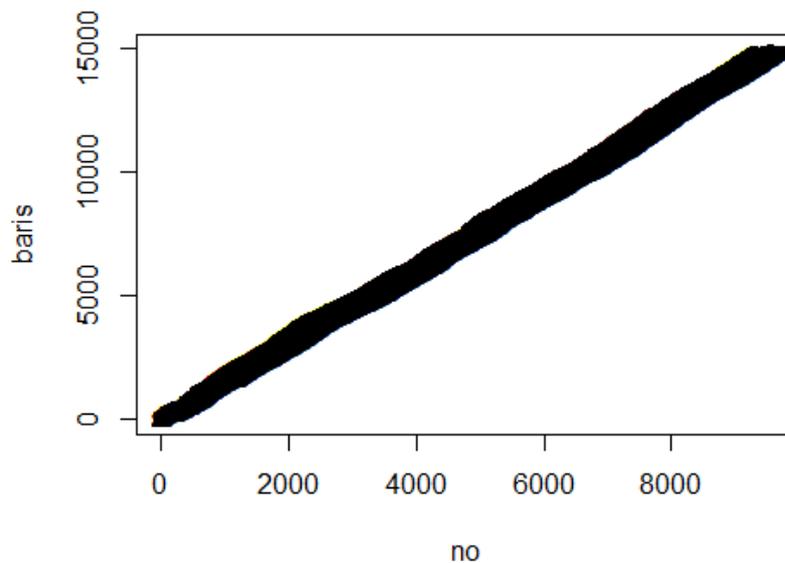
Gambar 5. Jumlah sentimen

Selanjutnya mengelompokkan kata didapatkan hasil TF IDF telah dilakukan pengelompokkan.

no	word	wordTotal	n	tf	idf	tf_idf
1	masih	30	1	0.033333333	5.433722	0.181124067
2	ppkm	902	1	0.001108647	5.433722	0.006024082
3	level	392	1	0.002551020	5.433722	0.013861536
4	jelang	15	1	0.066666667	5.433722	0.362248134
5	ramadhan	30	1	0.033333333	5.433722	0.181124067
6	tarawih	15	1	0.066666667	5.433722	0.362248134
7	di	286	1	0.003496503	5.433722	0.018999028
8	masjid	16	1	0.062500000	5.433722	0.339607625
9	ppkm	902	1	0.001108647	5.433722	0.006024082
10	atau	30	1	0.033333333	5.433722	0.181124067
11	gitu	11	1	0.090909091	5.433722	0.493974728

Gambar 6. Hasil TF IDF

Tahapan berikutnya, ialah proses *clustering* guna mendapati poin pencilan berawal data PPKM.



Gambar 7. Sebaran data PPKM

C. Evaluasi

Setelah melakukan proses tahapan *tidy text* selanjutnya akan diimplementasikan ke dalam algoritme DBSCAN dan tentu dievaluasi menggunakan *Sum of Square Error*. Data digunakan yaitu *tidy text* dengan jumlah data sebanyak 9562 data *tidy* pada gambar 8

	no	baris
1	-1.7225889	-1.7868704
2	-1.7099692	-1.7342985
3	-1.6973495	-1.7295192
4	-1.6847298	-1.6769473
5	-1.6721101	-1.6721680
6	-1.6594904	-1.6195961
7	-1.6468707	-1.6148168
8	-1.6342510	-1.5813619
9	-1.6216313	-1.5765827
10	-1.6090116	-1.5718034

Gambar 8. Data numerik

Hasil diperoleh kemudian dievaluasi menggunakan *cluster SSE* mendapatkan bobot nilai 88.5 % dari pemeriksaan *cluster*. Nilai diperoleh menunjukkan akurasi *cluster* yang sesuai dengan hasil klasifikasi pengelompokan kata.

```
Within cluster sum of squares by cluster:
[1] 718.2891 761.9553 710.5976
(between_SS / total_SS = 88.5 %)
```

Gambar 9. Nilai evaluasi SSE dengan algoritme DBSCAN

IV. KESIMPULAN DAN SARAN

Model yang dihasilkan berbentuk *wordcloud* dan histogram kata yang paling sering muncul pada kata “ppkm” berjumlah 902 kata dan “level” berjumlah 392 kata. Lalu menghasilkan nilai sentimen yang dihasilkan tweet kebijakan pemerintah terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yaitu bersifat positif.

Penelitian pada data yang menggunakan ukuran besar tidak disarankan menggunakan laptop yang memiliki spesifikasi rendah karena akan berpengaruh terhadap pemrosesan algoritme DBSCAN dalam menangani data yang ukuran besar. Data yang memiliki ukuran besar sulit untuk menampilkan visualisasi *wordcloud* maupun dalam bentuk plot.

PENGAKUAN

Naskah ilmiah ini adalah Sebagian dari penelitian Tugas Akhir milik Ferdian Wiharmariyansyah dengan judul Analisis Sentimen Pada Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Algoritma DBSCAN yang dibimbing oleh Ibu Anis Fitri Nur Masruriyah, M.Kom dan Ibu Hilda Yulia Novita, M.Kom.

DAFTAR PUSTAKA

- [1] Windy Rohalidyawati, Rita Rahmawati, And Mustafid , "Segmentasi Pelanggan EMoneydengan Menggunakan Algoritma Dbscan (Density Based Spatial Clustering Applications With Noise)Di Provinsi Dki Jakarta," Jurnal Gaussian, 2020.
- [2] Muhammad Tanzil Furqon And Lailil Muflikhah, "Clustering The Potential Risk Of Tsunami Using Density-Based Spatial Clustering Of Application With Noise (Dbscan)," Journal Of Environmental Engineering & Sustainable Technology, P. 8, 2016
- [3] Devi Putri Isnarwaty And Irhamah, "Text Clustering Pada Akun Twitter Layanan Ekspedisi Jne, J&T, Dan Pos Indonesia Menggunakan Metode Density-Based Spatial Clustering Of Applications With Noise (Dbscan) Dan K-Means," Jurnal Sains Dan Seni Its, 2019
- [4] Nurulhuda Zainuddin, Ali Selamat, And Roliana Ibrahim, "Hybrid Sentiment Classification On Twitter Aspect-Based Sentiment Analysis," Crossmark, 2018.

- [5] Andriyani, L., Gultom, A., Ketiara, K., Dahlan, J. K. H. A., Timur, K. C., Tangerang, K., Dahlan, J. K. H. A., Timur, K. C., & Selatan, K. T. (2021). *Dampak Sosial Ekonomi Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) di Indonesia*.
- [6] Ardianne Luthfika, F., Rima Dias, R., & Nia Annisa Ferani, T. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Jurnal Dinda*.