

Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm

Aldiansyah Putra^{1*}, Dede Haeirudin, Hasna Khairunnisa, Retnani Latifah

¹Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jakarta, Jl. Cempaka Putih2 Tengah No.27, Jakarta,10510

*Corresponding Author : 2018470006@ftumj.ac.id

Abstrak

Pemerintah menerapkan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) demi menanggulangi penyebaran virus covid-19 yang mengundang pro dan kontra. Analisis sentimen di media sosial twitter dapat menjadi cara untuk melihat pola polarisasi terhadap isu tersebut. Metode *Support Vector Machine* adalah metode yang digunakan pada penelitian kali ini dimana merupakan metode yang cukup sukses digunakan untuk analisis sentimen. Dari hasil pemodelan menggunakan 3000 data diketahui akurasi adalah 64%, tidak terlalu tinggi jika dibandingkan dengan analisis sentimen yang telah dilakukan di penelitian-penelitian sebelumnya. Namun, algoritma SVM masih mampu mengenali kata-kata penting untuk sentimen negatif seperti 'tolak PPKM', 'pelan-pelan' dan 'rumah makan'. Metode SVM dapat memprediksi 2000 tweet menjadi 57%-60% berlabel positif dan sisanya negatif, dimana hasilnya tidak jauh berbeda dengan data *training*.

Kata kunci: PPKM, Support Vector Machine, Klasifikasi, Sentimen Analisis

Abstract

The government implemented a Community Activity Restrictions (PPKM) to prevent the spread of the COVID-19 virus, which result in pros and cons. Implementing a sentiment analysis using Twitter dataset could be a way to find the polarization pattern of the issue. One method that has been widely and successfully used in sentiment analysis is Support Vector Machine. From the experiment using 3000 Twitter dataset, the accuracy of the method is 64%, compared to the result of previous research it is considered not as high. However, the SVM algorithm is still able to identify important words for negative sentiment such as 'tolak PPKM', 'pelan-pelan' and 'rumah makan'. This method successfully predict another 2000 Twitter dataset into positive label (57%-60%) and negative label, in which the distribution is similar to the *training* data.

Keywords : PPKM, Support Vector Machine, Classification

PENDAHULUAN

Selama masa pandemi, penggunaan media sosial meningkat karena masyarakat memanfaatkan platform tersebut untuk mendapatkan informasi dan berinteraksi. Selain itu, masyarakat juga menggunakan media sosial untuk menyampaikan opini mengenai apa yang sedang terjadi saat ini. Media sosial yang saat ini banyak digunakan oleh masyarakat yaitu Twitter. Menurut *We Are Social* jumlah pengguna sosial media pada Januari 2021, mengalami peningkatan 13,2% menjadi 4,2 miliar pengguna dari

periode yang sama tahun sebelumnya (Kemp, 2021).

Salah satu fitur utama Twitter yang digunakan oleh masyarakat adalah tweet yang dapat digunakan untuk memberikan informasi maupun untuk menyatakan pemikiran dan pendapat seseorang mengenai sesuatu hal atau topik. Pendapat masyarakat adalah menjadi hal yang sangat memberi pengaruh terhadap perilaku manusia. Oleh karena itu dengan mempelajari pendapat masyarakat, maka perilaku masyarakat mengenai suatu topik dapat diketahui dan dilakukan tindakan yang

sesuai (Fairuz, Ramadhani and Tanjung, 2021).

Salah satu topik yang sempat menjadi banyak perbincangan adalah mengenai kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yang merupakan dampak dari kenaikan kasus positif akibat munculnya varian baru yaitu versi Delta. Varian ini diketahui lebih cepat menular dibandingkan dengan varian lainnya, sehingga peningkatan jumlah pasien positif dan pasien meninggal di berbagai wilayah di Indonesia meningkat dengan sangat cepat dalam waktu singkat.

Kebijakan PPKM diberlakukan berdasarkan Instruksi Menteri Dalam Negeri no.15 tahun 2021 tentang Pemberlakuan Pembatasan Kegiatan Masyarakat Darurat Corona Virus Disease 2019 di Wilayah Jawa dan Bali (KEMENDAGRI, 2021). Kebijakan ini dibagi menjadi beberapa level dari level 1 – 4, tergantung kondisi pandemi yang terjadi di Indonesia. Di masa awal varian delta menyebar di Indonesia, sekitar bulan Juni-Agustus 2021, pemerintah menetapkan untuk menerapkan kebijakan PPKM level 4. Efek dari pemberlakuan kebijakan tersebut adalah banyak kegiatan masyarakat yang dibatasi sehingga berdampak pada perekonomian masyarakat. Salah satu dampaknya adalah turunnya omset penjualan dan terkendalanya distribusi bahan baku (Rizal, Afrianti and Abdurahman, 2021).

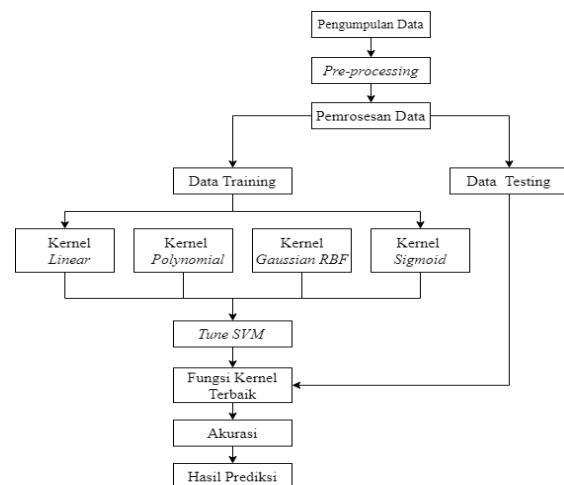
Dampak terhadap beberapa aspek kehidupan ini, menyebabkan volume percakapan mengenai PPKM juga meningkat. Baik itu mengenai berbagai informasi terkait PPKM seperti hal-hal apa yang perlu diperhatikan oleh masyarakat selama PPKM level 4 berlangsung maupun pendapat masyarakat yang mendukung dan menolak kebijakan tersebut. Untuk dapat mengetahui pola polarisasi pendapat masyarakat terhadap PPKM, maka dapat dilakukan sentimen analisis yang juga dikenal dengan opinion mining. Sentimen analisis salah satunya dapat digunakan untuk mengetahui opini publik yang masih terkait dengan isu, kejadian, dan peristiwa yang terjadi di masyarakat (Fairuz et al., 2021).

Analisis sentimen untuk melihat pola polarisasi masyarakat terhadap suatu kasus sudah cukup sering dilakukan antara lain

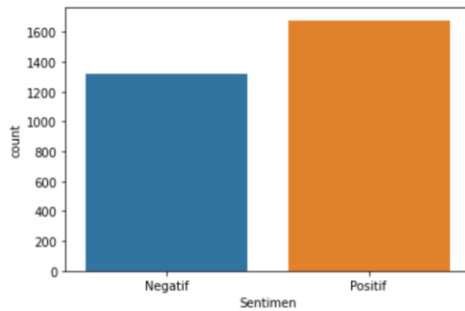
digunakan untuk mengetahui polarisasi mengenai pilkada 2020 (Latifah, Adharani, et al., 2021), mengenai COVID-19 (Fairuz, Ramadhani and Tanjung, 2021), mengenai kebijakan Pembatasan Sosial Berskala Besar atau PSBB yang diterapkan di awal pandemi (Adrian et al., 2021), maupun mengenai vaksin COVID-19 dan proses vaksinasinya (Aditama et al., 2020; Rachman and Pramana, 2020; Harun and Ananda, 2021; Pratama et al., 2021; Pristiyono et al., 2021).

Namun, publikasi mengenai sentimen analisis terhadap PPKM masih belum banyak dilakukan. Dimana salah satu hasil penelitian mengenai analisis sentimen PPKM dengan 1000 data Twitter menghasilkan bahwa 99% masyarakat memiliki sentimen positif (Krisdiyanto, 2021). Topik ini cukup menarik untuk dilakukan untuk melihat perbandingan pola polarisasi terkait PPKM apakah konsisten atau tidak. Oleh karena itu dilakukan penelitian mengenai analisis sentimen terhadap PPKM dengan menggunakan data Twitter dengan menggunakan *Support Vector Machine* (SVM).

METODE PENELITIAN



Gambar 1. Tahap-Tahapan Penelitian
Adapun tahap-tahap penelitian ini dilakukan seperti terlihat pada gambar 1.



Gambar 2. Persebaran Sentimen pada Data *Training*

Dalam penelitian ini yang dilakukan pertama kali adalah pengumpulan data Twitter. Pengumpulan data dilakukan dengan cara melakukan *crawling* menggunakan *tweepy* yang merupakan API pada *python*. Data dikumpulkan selama satu minggu dari 16 juli sampai 24 juli 2021. Jumlah data keseluruhan yang diambil adalah 5000 tweet dengan kata kunci yang digunakan adalah "PPKM".

Dari data tersebut, diambil 3000 tweet yang kemudian diberi label secara manual menjadi dua kelas sentimen yaitu *positive* dan *negative*. Persebaran datanya dapat dilihat pada gambar 2, dimana selisih data dari keduanya tidak terlalu besar yaitu sekitar 300 data. Data ini digunakan sebagai data *training* untuk membuat model dengan menggunakan SVM. Sisa 2000 data akan digunakan untuk melakukan prediksi secara otomatis dengan menggunakan model dengan evaluasi terbaik.

Didalam tahapan *pre-processing* terdapat *cleaning* data yang bertujuan untuk menghapus atau membersihkan karakter yang tidak dibutuhkan seperti tanda baca, *mention*, *hashtag*, *link* dan lain-lain serta merubah huruf menjadi huruf kecil (*casefolding*). Selain itu juga dilakukan *stopword removal* menggunakan *list* dari penelitian Tala (Tala, 2003) dan *stemming*, yaitu merubah kata menjadi bentuk dasarnya atau akar katanya, menggunakan *library* sastrawi. Setelah dilakukan *pre-processing*, data akan dibagi menjadi data *training* dan *testing* dimana masing-masing datanya berjumlah 80% untuk data *training* dan 20% untuk data *testin*.

Pemodelan data *training* untuk penelitian ini dilakukan dengan menggunakan SVM. Dalam penggunaan SVM ini yang merupakan tahapan atau proses *learning* dengan menggunakan ruang hipotesis yang berupa

kumpulan fungsi linear dalam ruang fitur berdimensi tinggi dan juga menerapkan *learning* bias yang berdasar pada konsep *learning* statistik yang dilatih berdasarkan algoritma pembelajaran atau *learning* (Jumeilah, 2017). Masalah yang ada pada algoritma SVM yaitu bagaimana cara untuk memisahkan dua kelas dengan suatu fungsi yang didapat berdasarkan data *training* yang tersedia. Secara sederhana, SVM adalah proses mencari *hyperplane* yang paling baik dengan tujuan dapat memisahkan dua buah kelas (Hasibuan, Mukid and Prahutama, 2017). Dalam penggunaannya algoritma SVM menggunakan beberapa titik data yang terpilih dan kontribusi serta disebut sebagai *Support vector machine* yang bertujuan untuk membentuk model yang digunakan sebagai klasifikasi (Rofiqoh, Perdana and Fauzi, 2017).

SVM awalnya digunakan untuk menyelesaikan permasalahan secara linear, namun kemudian dikembangkan untuk bisa digunakan pada permasalahan non linear dengan menggunakan fungsi yang dapat memisahkan secara non linear yaitu berdasarkan penggunaan kernel yang dinamakan kernel trick pada dataset yang berdimensi tinggi (Pratama and Trilaksono, 2015). Kernel yang bisa digunakan antara lain adalah fungsi linear, *polynomial*, *gaussian RBF* dan *sigmoid*.

Pada penelitian mengenai klasifikasi yang telah dilakukan sebelumnya dengan menggunakan teknik SVM, menunjukkan bahwa SVM mampu memberikan akurasi tertinggi dibandingkan dengan *decision tree*, *naive bayes*, dan *bayesian network* (Jumeilah, 2017). Pada penelitian lain mengenai sentimen analisis vaksin COVID-19 menggunakan data Twitter juga didapatkan hasil bahwa SVM memiliki nilai akurasi dan F1-measure tertinggi dibandingkan enam teknik lain seperti *logistic regression*, *decision tree*, *kNN*, *multinomial naive bayes*, *random forest*, *adaboost classifier* dan *multilayer perceptron* (Latifah, Meilina, et al., 2021).

Penelitian ini menggunakan empat tipe kernel untuk mengetahui fungsi mana yang memberikan evaluasi yang paling optimal. Selain itu juga dilakukan tuning terhadap parameter C yang berperan sebagai

penelitian. Dalam penelitian kali ini teknik yang digunakan yaitu akurasi.

terlihat dari visualisasi pada gambar 4 yang tidak jauh berbeda dengan gambar 2.

HASIL DAN PEMBAHASAN

Berdasarkan hasil pemodelan data training menggunakan SVM dengan empat kernel yang berbeda, diketahui hasilnya adalah seperti terlihat pada tabel 1. Kernel yang memberikan nilai akurasi tertinggi adalah kernel linear dengan 64%. Perbedaan akurasi dari setiap kernel tidak jauh berbeda dan hanya memiliki selisih 1-4% antara satu dengan yang lain.

Tabel 1. Hasil Akurasi SVM dengan Berbagai Kernel

Jenis Kernel	Accuracy
Kernel <i>Linear</i>	0.64
Kernel <i>Gaussian RBF</i>	0.62
Kernel <i>Sigmoid</i>	0.63
Kernel <i>Polynomial</i>	0.60

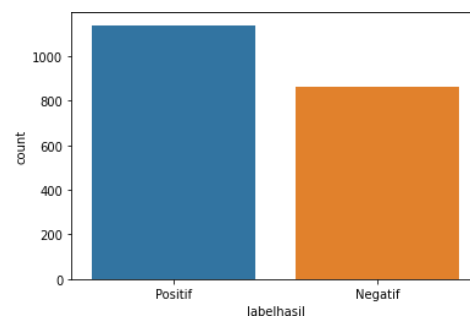
Rendahnya nilai akurasi bisa dikarenakan kata-kata yang digunakan untuk identifikasi pola sentimen pada data *training* terlalu beragam. Isi tweet yang beragam menyebabkan sulit untuk mendapatkan pola yang optimal. Selain itu, setelah melakukan *pre-processing*, token yang digunakan untuk pembuatan model menjadi terlalu sedikit dan berbeda-beda.

Percobaan selanjutnya adalah melakukan tuning terhadap parameter C. Nilai C yang digunakan adalah 0.1, 0.5, 1, 5, 10, 50 dan 100. Hasilnya dapat dilihat pada gambar 3. Nilai akurasi tertinggi diperoleh pada saat C=1 dengan akurasi 64%. Terlihat bahwa setelah C=1, nilai akurasi semakin menurun sebelum kemudian menjadi stabil di angka 60%. Sedangkan sebelum C=1, nilai akurasi semakin meningkat dengan semakin tingginya nilai C.

Dari hasil percobaan diketahui bahwa model SVM dengan akurasi tertinggi adalah dengan menggunakan kernel linear dan C=1. Model ini kemudian digunakan untuk melakukan prediksi sentimen pada 2000 tweet secara otomatis. Hasil prediksi menunjukkan bahwa sebagian besar tweet diklasifikasikan sebagai label positif. Namun perbedaan keduanya tidak terlalu jauh dan cukup merefleksikan data *training* yang digunakan,



Gambar 3. Hasil Perbandingan Penalti C



Gambar 4. Visualisasi Hasil Prediksi Menggunakan SVM

Berdasarkan hasil yang didapat pada penelitian ini berdasarkan hasil prediksi dilakukan visualisasi dengan melihat gambar hasil prediksi untuk setiap sentimen. Dengan diprosesnya data tersebut terdapat kata-kata yang paling sering muncul yaitu untuk sentimen negatif adalah "Tolak PPKM" dan "tolak PPKM". Hal ini berarti beberapa percakapan dengan kata tersebut cenderung bertendensi negatif. Hal tersebut membahas tentang dampak PPKM terhadap masyarakat yang menolak kebijakan karena secara perlahan dapat menyebabkan penurunan ekonomi. Hal ini dapat dilihat pada gambar 5. Kata lain yang cukup nyata terlihat adalah 'rumah makan' yang mengimplikasikan bahwa banyak rumah makan yang terdampak. Berdasarkan pemaparan di atas memiliki tujuan yang sama dengan penelitian yang dilakukan sebelumnya mengenai dampak PPKM terhadap kedai kopi (Rizal, Afrianti and Abdurahman, 2021).

Sedangkan pada tweet yang bertendensi positif, kata-kata yang muncul antara lain

adalah “mendagri pemda” dan ‘salur bansos’. Frase pertama berarti memberi suatu instruksi perintah terhadap masyarakat untuk terus menerapkan protokol kesehatan. Sedangkan “salur bansos” berarti akan dilakukan penyaluran bantuan sosial yang disambut cukup baik oleh masyarakat. Visualisasi dapat dilihat di gambar 6.



Gambar 5. Visualisasi Word Cloud Hasil Prediksi Untuk Sentimen Negatif



Gambar 6. Visualisasi Word Cloud Hasil Prediksi Untuk Sentimen Positif

Dari hasil analisis sentimen diketahui bahwa algoritma memiliki akurasi yang cukup rendah, tidak seperti hasil penelitian-penelitian sebelumnya dimana akurasi bisa mencapai di atas 70% bahkan lebih. Salah satu faktor yang mempengaruhi adalah proses *pre-processing* data teks yang kurang optimal. Dari hasil prediksi menggunakan SVM, tweet yang diberi label positif lebih banyak dibandingkan label negatif sehingga sesuai dengan penelitian sebelumnya. Namun, rasio label positif hanya sekitar 57%-60%, tidak sebesar penelitian sebelumnya yang mencapai 99%. Hal ini juga dipengaruhi dari data *training* yang persebaran labelnya juga cukup seimbang.

KESIMPULAN

Sentimen analisis mengenai PPKM dari data Twitter dengan menggunakan SVM menunjukkan akurasi sebesar 64% dengan kernel Linear dan penalti regularisasi $C = 1$. Dimana dari hasil prediksi terhadap data tweet non label, diketahui persebarannya adalah 57%-60% berlabel positif dan sisanya berlabel negatif. Meski demikian, dari analisis *word cloud* diketahui bahwa algoritma SVM mampu mengenali tweet penolakan PPKM sebagai tweet yang bertendensi negatif dan dapat mengidentifikasi kata-kata penting yang berhubungan tendensi negatif tersebut seperti ‘rumah makan’ dan ‘pelan-pelan’.

DAFTAR PUSTAKA

- Aditama, M. I. *et al.* (2020) ‘Analisis Klasifikasi Sentimen Pengguna Media Sosial Twitter Terhadap Pengadaan Vaksin COVID-19’, *Journal Information Engineering and Educational Technology (JIEET)*, 4(2), pp. 90–92.
- Adrian, M. R. *et al.* (2021) ‘Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB’, *Jurnal Informatika Upgris*, 7(1), pp. 6–11. doi: 10.26877/jiu.v7i1.7099.
- Fairuz, A. L., Ramadhani, R. D. and Tanjung, N. A. (2021) ‘Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial’, *Jurnal DINDA*, 1(1), pp. 10–12. Available at: <http://journal.itelkom-pwt.ac.id/index.php/dinda/article/view/180>.
- Harun, A. and Ananda, D. P. (2021) ‘Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decision Tree’, *Indonesia Journal of Machine Learning and Computer Science*, 1(April), pp. 58–63.
- Hasibuan, C. A., Mukid, M. A. and Prahutama, A. (2017) ‘Klasifikasi Diagnosa Penyakit Demam Berdarah Dengue (Dbd) Menggunakan Support Vector Machine (Svm) Berbasis Gui Matlab’, *Jurnal Gaussian*, 6(2), pp. 171–180. Available at: <http://ejournal-s1.undip.ac.id/index.php/gaussian>.

- Jumeilah, F. S. (2017) ‘Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian’, *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 1(1), pp. 19–25. doi: 10.29207/resti.v1i1.11.
- KEMENDAGRI (2021) *Instruksi Menteri Dalam Negeri Nomor 15 Tahun 2021 Tentang Pemberlakuan Pembatasan Kegiatan Masyarakat Darurat Corona Virus Disease (Covid-19) Di Wilayah Jawa Dan Bali*.
- Kemp, S. (2021) *Digital in Indonesia: All the Statistics You Need in 2021 — DataReportal – Global Digital Insights, Datareportal*. Available at: <https://datareportal.com/reports/digital-2021-turkey%0Ahttps://datareportal.com/reports/digital-2021-indonesia?rq=indonesia>.
- Krisdiyanto, T. (2021) ‘Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers’, *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 7(1), pp. 32–37. Available at: <http://ejournal.uin-suska.ac.id/index.php/coreit/article/view/12945>.
- Latifah, R., Meilina, P., et al. (2021) *Laporan Penelitian: Analisis Sentimen dan Deteksi Topik secara Otomatis Mengenai Vaksin COVID-19*.
- Latifah, R., Adharani, Y., et al. (2021) *Laporan Penelitian: Pengembangan Analisis Sentimen Twitter Berbahasa Indonesia Mengenai Pilkada 2020*.
- Pratama, E. E. and Trilaksono, B. R. (2015) ‘Klasifikasi Topik Keluhan Pelanggan Berdasarkan Tweet dengan Menggunakan Penggabungan Feature Hasil’, *Jurnal Edukasi dan Penelitian Informatika*, 1(2), pp. 53–59.
- Pratama, G. Y. et al. (2021) ‘Public Sentiment Towards The COVID 19 Vaccine in Indonesia’, *Turkish Journal of Physiotherapy and Rehabilitation*, 32(3), pp. 6081–6084.
- Pristiyono et al. (2021) ‘Sentiment Analysis of COVID-19 Vaccine in Indonesia using Naïve Bayes Algorithm’, in *IOP Conference Series: Materials Science and Engineering*, p. 012045. doi: 10.1088/1757-899x/1088/1/012045.
- Rachman, F. F. and Pramana, S. (2020) ‘Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter’, *Health Information Management Journal*, 8(2), pp. 100–109. Available at: <https://inohim.esaunggul.ac.id/index.php/INO/article/view/223/175>.
- Rizal, M., Afrianti, R. and Abdurahman, I. (2021) ‘Dampak Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) bagi Pelaku Bisnis Coffe shop pada Masa Pandemi Terdampak COVID-19 di Kabupaten Purwakarta The Impact of the Policy for Implementing Community Activity Restrictions for Coffee Shop Busi’, *Jurnal Inspirasi*, 12(1), pp. 97–105.
- Rofiqoh, U., Perdana, R. S. and Fauzi, M. A. (2017) ‘Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature’, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), pp. 1725–1732. Available at: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>.
- Tala, F. Z. (2003) ‘A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia’, *M.Sc. Thesis, Appendix D*, pp. 39–46.