

Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan *Support Vector Machine* dan *Naive Bayes*

Policy Sentiment Analysis Face-to-face Learning Using Supports Vector and Naive Bayes Engines

Muhammad Siddik Hasibuan^{1*}, Akbar Serdano²

^{1,2}Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara

Jl. Lapangan Golf, Desa Durian Jangak, Pancur Batu, Kabupaten Deli Serdang 20353, Indonesia

*email: muhammadsiddik@uinsu.ac.id

ABSTRAK

DOI;
10.30595/jrst.v6i2.15145

Histori Artikel:

Diajukan:
25/09/2022

Diterima:
21/11/2022

Diterbitkan:
25/11/2022

Analisis sentimen kebijakan pembelajaran tatap muka selama pandemi dapat dianalisis menggunakan data dari tweet di platform media Twitter. Dari hasil crawling data diperoleh 152 data, selanjutnya dilakukan proses pre-processing untuk memberikan data dari simbol, dan link serta mengubah kata menjadi bobot kata yang dapat dipahami dengan menggunakan teknik TF-IDF. Hasil pembobotan kata ini dianalisis pada kelas positif dan negatif, label positif terkumpul 96 dan negatif 16. Polaritas kelas hasil dari 108 data yang diketahui masyarakat memberikan respon positif terhadap kebijakan pemerintah untuk melakukan pembelajaran tatap muka. Pengujian model *Support Vector Machine* (SVM) dan *Naive Bayes* (NB) dari 108 data diperoleh hasil akurasi SVM sebesar 88,09% dan NB 75,92% pada penelitian ini SVM lebih baik dari NB.

Kata Kunci: Analisis Sentimen, *Twitter*, TF-IDF, SVM, NB

ABSTRACT

Sentiment analysis of face-to-face learning policies during a pandemic can be analyzed using data from tweets on the Twitter media platform. From the results of data crawling, 152 data were obtained, then the pre-processing process was carried out to provide data from symbols, and links and to convert words into word weights that could be understood using the TF-IDF technique. The results of this word weighting were analyzed in positive and negative classes, positive labels collected 96 and negative 16. The class polarity of 108 data known to the public gave a positive response to government policies to conduct face-to-face learning. Testing the Support Vector Machine SVM and Naive Bayes (NB) models from 108 data obtained SVM accuracy results of 88.09% and NB 75.92% in this study SVM is better than NB.

Keywords: Sentiment Analysis, *Twitter*, TF-IDF, SVM, NB

1. PENDAHULUAN

Tahun 2020 akan dikenang dalam sejarah umat manusia sebagai tahun pandemi global yang disebabkan oleh COVID-19. Peringatan kesehatan membuat apa yang kita miliki sampai saat itu dikenal sebagai "normal" tidak lagi berlaku di semua bidang kehidupan kita. Dalam pendidikan, tindakan sehari-hari yang sederhana untuk pergi ke kelompok

menjadi "tinggal di rumah" di depan komputer, dengan konsekuensi stres dan kelelahan teknologi. Situasi ini mengubah cara kita memandang dan melihat pendidikan, cara mengajar, cara belajar, dan cara mengevaluasi perolehan kompetensi dan kemampuan siswa (Dios & Charlo, 2021).

Presiden Indonesia Jokowi telah mengeluarkan kebijakan pembatasan sosial

berskala besar (PSBB). Sosial pembatasan mempengaruhi ruang publik, termasuk kantor, sekolah atau kampus. Dalam strategi ini melarang pembelajaran tatap muka dan memaksa semua lembaga pendidikan untuk menghentikan proses pembelajaran tatap muka dan menggantinya dengan pembelajaran online (Watrianthos, 2020). Dalam pendidikan muncul kesulitan, seperti ketidaksiapan teknologi, media yang digunakan atau aspek psikologis siswa. Sementara itu, masa depan akan pendidikan yang tidak merata antara penduduk dan wilayah di Indonesia dalam jangka panjang (Syah, 2020). Indonesia dan negara-negara lain di dunia telah hidup selama 2 (dua) tahun selama pandemi Covid-19, semua aktivitas di luar rumah & semuanya dilakukan tanpa banyak kerumunan manusia. Oleh karena itu, cara lain untuk tetap melakukan aktivitas seperti aktivitas kerja hingga sekolah dilakukan sesuai dengan tempat tinggalnya (Samsir et al., 2021). Hal ini dilakukan untuk mencegah munculnya klaster baru Covid. Kebijakan pengetatan protokol kesehatan juga tidak efisien, sehingga pemberian vaksin dirasa mampu menekan munculnya masalah baru penyebaran COVID-19 (Kemenkes, 2021). Jumlah infeksi covid yang semakin menurun membuat pemerintah melonggarkan kegiatan di luar rumah, beberapa daerah sudah menerapkan total kehadiran pegawai sebesar 75% dari total jumlah pegawai. Tidak hanya perkantoran, sekolah juga sudah mulai melakukan pembelajaran tatap muka (PTM) sesuai peraturan dengan empat menteri nomor 03/KB/2021, yaitu menteri pendidikan dan kebudayaan, menteri agama, menteri kesehatan, dan menteri dalam negeri. Yang dimaksud PTM pada peraturan tiga menteri tersebut adalah belajar langsung di sekolah atau luring. Ada beberapa syarat yang harus disiapkan bagi tempat penyelenggara pendidikan untuk melaksanakan PTM yaitu wajib vaksin. Namun, pendidik bagi belum divaksinasi dan dapat melaksanakan PTM terbatas harus mengikuti protokol kesehatan yang ketat dan mendapatkan izin dari pemerintah daerahnya. Kegiatan PTM yang dioptimalkan pelaksanaannya banyak menuai kontroversi, seperti takut tertular, tidak divaksinasi, atau tidak terjaminnya protokol kesehatan di sekolah atau universitas. Kontroversi tersebut membuat informasi PTM menjadi sebuah pemikiran, pemikiran ini berkembang dimasyarakat. Pemikiran atau pendapat ini banyak dipublikasikan di media sosial, dan menelusuri pemikiran ini bisa menjadi keputusan yang bisa diimplementasikan dengan pembelajaran mesin. Sosial media menjadi hal paling populer, Twitter merupakan sosial media yang sering dipakai orang agar

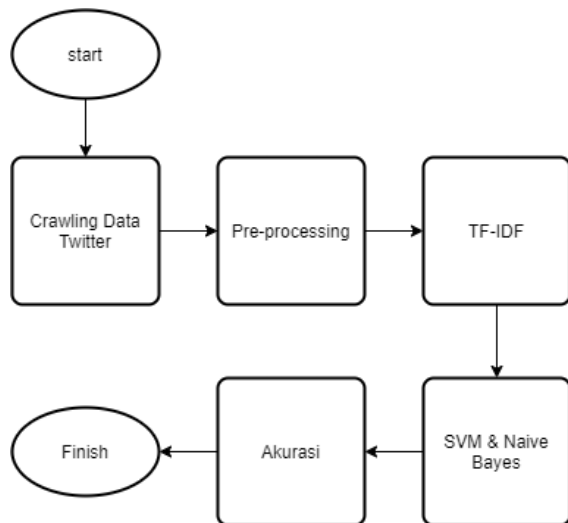
memberikan sentimen (Pilehvar & Camacho-Collados, 2020). Saat ini, era internet telah mengubah cara orang mengekspresikan pandangan dan pendapat mereka. Sekarang sebagian besar dilakukan melalui posting blog, forum online, situs ulasan produk, media sosial, dll. Saat ini, jutaan orang menggunakan situs jejaring sosial seperti Facebook, Twitter, Google Plus, dll. untuk mengekspresikan emosi, pendapat, dan berbagi pandangan tentang kehidupan sehari-hari mereka. Melalui komunitas online, kami mendapatkan media interaktif di mana konsumen menginformasikan dan mempengaruhi orang lain melalui forum. Media sosial menghasilkan sejumlah besar data kaya sentimen dalam bentuk tweet, pembaruan status, posting blog, komentar, ulasan, dll. (Zhang et al., 2011). Media Twitter sering menjadi analisis perkembangan informasi seperti opini, kritik, fakta, atau saran untuk sesuatu (Myslin et al., 2013). Opini yang ada di media sosial dapat dianalisis menjadi data yang dapat menghasilkan tweet positif, negatif, atau netral yang disebut dengan analisis sentimen (Kurniawan & Apriliani, 2020). Oleh karena itu, penelitian ini mengambil data dari kicauan orang-orang di platform Twitter (Maynard et al., 2017).

Dalam beberapa tahun terakhir, model Natural Language Processing (NLP) telah mencapai kesuksesan fenomenal dalam tugas linguistik dan semantik seperti klasifikasi teks, terjemahan mesin, sistem dialog kognitif, pencarian informasi melalui Natural Language Understanding (NLU), dan Natural Language Generation (NLG) (Singh & Mahmood, 2021). Penyematan kata yang sadar konteks dapat digunakan untuk tugas pelabelan berurutan seperti penandaan bagian dari ucapan (POS) dan pengenalan entitas bernama (NER), dan penyematan kalimat dapat digunakan untuk tugas tingkat kalimat, seperti analisis sentimen dan parafrase klasifikasi (Zhou et al., 2020). Kajian ini menggali pengetahuan tentang sentimen publik tentang kebijakan pemerintah yang menerapkan pembelajaran tatap muka di masa pandemi. Review yang dijadikan data adalah postingan atau tweet dari media sosial Twitter, nantinya data tweet ini akan menjadi dataset. Dataset ini akan diukur menggunakan dua algoritma klasifikasi, yaitu SVM dan Naive Bayes. Agar hasil dari model yang diuji dapat diketahui, maka akan dilakukan pengukuran performansi dari masing-masing algoritma SVM dan Naive Bayes

2. METODE PENELITIAN

Proses penelitian yang dilakukan terhadap kata kunci dengan query "*Pembelajaran*

"Tatap Muka" adapun tahapan proses penelitian seperti gambar 1 dibawah ini.



Gambar 1. Alur Metode Penelitian

Gambar 1 merupakan alur proses processing test yang dimulai dari tahap *crawling*, *labeling* (negatif/positif), *case folding*, *tokenizing*, serta *stemming* (Soufyane et al., 2021). Pada tahapan pre-prosesing data proses yang dilakukan adalah membersihkan data seperti menghapus duplikasi text, menghapus tanda baca, link, simbol, dan menyeragamkan bentuk kata menjadi *lower case*. Setelah proses pre-processing text, lalu data di pisah pisah atau tokenizing dengan metode TF-IDF, proseds TF-IDF membantu untuk meningkatkan akurasi model (Amidror, 2018). Hasil dari proses tokenizing tersebut selanjutnya di proses menggunakan algoritma SVM dan Naive Bayes. Hasil proses model sentimen ini berupa akurasi model SVM dan Naive Bayes. Adapun proses tersebut dapat digambarkan melalui gambar berikut 2. Relevansi untuk setiap kata dalam kumpulan data kami dihitung menggunakan frekuensi dokumen terbalik (tf-idf), yang membebani frekuensi istilah dengan keseluruhan frekuensi istilah dan mengurangi pentingnya istilah yang sangat umum. 100 terbanyak kata-kata relevan yang digunakan dipilih berdasarkan nilai tf-idf mereka untuk analisis lebih dalam (Giunti et al., 2021)

3. HASIL DAN PEMBAHASAN

A. Analisis Data

Dari hasil penarikan data (Crawling) twitter diperoleh sebanyak 152 data dengan query "pembelajaran tatap muka" adapun data dapat dilihat pada tabel 1.

Tabel 1. Hasil Crawling data Twitter

No	Text
1	Untuk meningkatkan prestasi anak, Batalyon Arhanud 15/DBY memfasilitasi putra-putri prajurit yang memiliki minat untuk memperdalam keterampilan berbahasa Inggris di luar jam pembelajaran tatap muka di kelas. 8/2022 https://t.co/ZqtchQ8QdI
2	@iniroseblepink @enginafess ASTAGA TERNYATA SINGKATAN PASS THE MIC TOH.. tdi gue sempet ngiranya pembelajaran tatap muka pls???
3	Pelaksanaan screening PTM (Pembelajaran Tatap Muka) di Institusi Pendidikan Wilayah Kerja Puskesmas Mayangan https://t.co/4lNqjDAINr
4	Terapkan disiplin prokes pada pembelajaran tatap muka. Selain itu disiplin prokes tak bisa di tawar tawar https://t.co/NHn64kGI9x
5	RT @sg2772215: Terapkan Disiplin Prokes Pada Pembelajaran Tatap Muka
6	IKN Satuan Bangsa https://t.co/rChCx37T4
7	Terapkan Disiplin Prokes Pada Pembelajaran Tatap Muka
8	IKN Satuan Bangsa https://t.co/CB00lciCe3
9	@apoycuy Terapkan disiplin prokes pada pembelajaran tatap muka. IKN Satuan Bangsa
10	@apoycuy Terapkan disiplin prokes pada pembelajaran tatap muka. IKN Satuan Bangsa https://t.co/fhFLSXX7NT
11	Terapkan Disiplin Prokes Pada Pembelajaran Tatap Muka
12	IKN Satuan Bangsa https://t.co/rChCx37T4
13	terapkan disiplin prokes pada pembelajaran tatap muka. IKN Satuan Bangsa https://t.co/tJC6lmpkIY
	Terapkan Disiplin Progres Pada Pembelajaran Tatap Muka.
	Para guru dan karyawan sekolah juga memberi teladan untuk selalu menjaga

113 dst
gaya hidup bersih dan sehat serta disiplin prokes

Selanjutnya data pada tabel 1 akan dilakukan tahap pra-processing, tahap ini adalah proses membersihkan data seperti menghilangkan duplikasi text, menghilangkan tanda baca, link, simbol, dan menyeragamkan bentuk kata menjadi lower case.

B. Pra-processing

Proses pra-processing yang dilakukan didapat 108 data yang telah siap diolah. Selanjutnya data tersebut menjadi dasar untuk mengukur kinerja algoritma, adapun data dapat dilihat pada tabel 2.

Tabel 2. Data Pra-processing

TEXT	LABEL
notula acara dialog dekanat halo fmipa kamsis september dilaksanakan acara dialog dekanat daring mengusung tema dinamika adaptasi pembelajaran tatap muka	Positif
telyu tanya matkul koordinator beliau bikin rencana pembelajaran minggu gimana rencana kelas tatap muka asinkron dosen matkul sebenarnya wajib rencana koordinator cont	Positif
garuda kota bogor resmi memberlakukan pelaksanaan pembelajaran tatap muka persen sekolah kegia	Positif
schfess pandemi melandai universitas sekolah melaksanakan pembelajaran tatap muka perasaan kaget takut presentasi tatap muka ikuti webinar link regis	Positif
pigshavew pandemi melandai universitas sekolah melaksanakan pembelajaran tatap muka perasaan kaget takut presentasi tatap muka ikuti webinar link regis	Positif
santi pandemi melandai universitas sekolah melaksanakan pembelajaran tatap muka perasaan kaget takut presentasi tatap muka ikuti webinar link regis	Positif
anduago pandemi melandai universitas sekolah melaksanakan pembelajaran tatap muka perasaan kaget takut presentasi tatap muka ikuti webinar link regis	Positif
pandemi melandai universitas sekolah melaksanakan pembelajaran tatap muka perasaan kaget takut	Positif

presentasi tatap muka ikuti webinar link regis
korantempo seiring pemberlakuan pembelajaran tatap muka penuh Positif
kenaikan covid usia anak
dst

Dari 108 data dikelompokkan berdasarkan polaritas positif dan negatif, adapun hasil polaritas tersebut dapat dilihat pada tabel 3.

Tabel 3. Polaritas Label

No	Label	Absolute Count
1	Positif	96
2	Negatif	16

Pada tabel 3 diketahui data atau sentimen masyarakat terhadap pembelajaran tatap muka bernilai positif. Berarti bisa dilihat dari sebaran data yang dianalisis pembelajaran tatap muka sangat dinanti oleh masyarakat.

C. TF-IDF

Pada proses ini dilakukan pembobotan kata-kata yang sering muncul, proses ini juga sering dilakukan untuk pencarian kata dalam integrasi antar kata. Adapun kata-kata yang muncul dalam proses ini dapat dilihat pada tabel 4.

Tabel 4. Daftar Pembobotan Kata

Word	Attribut Name	Total Occurences	Document Occurences	Positif	Negatif
acara	acara	2	1	2	0
adadikompas	adadikompas	4	4	3	1
adaptasi	adaptasi	1	1	1	0
agama	agama	1	1	1	0
agustus	agustus	1	1	1	0
agustusanbarengti	agustusanbarengti	1	1	0	1
akademik	akademik	2	2	2	0
akibat	akibat	1	1	0	1
aldysuhandadst	aldysuhandadst	2	2	2	0

D. Evaluasi

Proses evaluasi disini untuk mengukur kinerja masing masing algoritma SVM atau NB, evaluasi yang digunakan adalah cross validation. Cross validation atau validasi silang Validasi silang mirip dengan metode subsampling acak berulang, tetapi pengambilan sampel dilakukan sedemikian rupa sehingga tidak ada dua pengujian set tumpang tindih. Dalam *k-fold cross-validation*, *learning set* yang tersedia dipartisi menjadi k subset yang terpisah dengan ukuran yang kira-kira sama (Wong & Yeh, 2020). Di sini, fold mengacu pada jumlah subset yang dihasilkan. Partisi ini dilakukan secara random sampling kasus dari pembelajaran ditetapkan tanpa penggantian. Model dilatih menggunakan k-1 himpunan bagian, yang bersama-sama mewakili himpunan pelatihan (Jung, 2018). Maka modelnya adalah diterapkan ke subset yang tersisa, yang dilambangkan sebagai set validasi, dan kinerjanya diukur. Prosedur ini diulangi sampai masing-masing subset k berfungsi sebagai set validasi, dalam penelitian ini K-Fold yang dipakai adalah 10. Adapun hasil pengujian model dengan algoritma SVM dan NB dapat dilihat dibawah ini.

PerformanceVector:SVM

accuracy: 88.09% +/- 7.45% (micro average: 87.96%)

ConfusionMatrix:

True:	Positif	Negatif
Positif:	91	12
Negatif:	1	4

PerformanceVector:Naïve Bayes

accuracy: 75.91% +/- 7.73% (micro average: 75.93%)

ConfusionMatrix:

True:	Positif	Negatif
Positif:	76	10
Negatif:	16	6

Hasil dari pengujian dapat dilihat pada tabel 5.

Tabel 5. Pengujian *Confusion Matrix*

Model	Akurasi	Recall	Precision
SVM	88.09	98.91	88.35
NB	75.91	82.61	88.37

Pengujian yang telah dilihat dapat diketahui SVM memiliki akurasi 88.09% sedangkan NB 75.91% dapat disimpulkan bahwa dalam penelitian ini SVM memiliki tingkat akurasi yang lebih baik dari NB.

4. KESIMPULAN

Hasil dari analisis yang dilakukan terhadap 108 data didapat antara lain :Data terhadap query “pembelajaran tatap muka” pada platform twitter dari opini masyarakat dinilai positif. Hasil diperoleh adalah 96 positif dan 16 negatif. Evaluasi model dengan algoritma SVM dan NB, didapat nilai akurasi bahwa SVM mendapat nilai 88.09% sedangkan NB mendapat nilai 75.91%.

UCAPAN TERIMA KASIH

Terima kasih kepada segenap lembaga penelitian dan pengabdian masyarakat LPPM UIN Sumatera Utara tahun anggaran BOPTN 2022 dengan No Kontrak 221140000057907 atas dukungan terhadap karya penelitian ini dan fasilitas yang diberikan.

DAFTAR PUSTAKA

- Amidror, Y. (2018). The evolution and development of the IDF. In *Routledge Handbook on Israeli Security*. <https://doi.org/10.4324/9781315165196-4>
- Dios, M. T. C., & Charlo, J. C. P. (2021). Face-to-face vs. E-learning models in the covid-19 era: Survey research in a Spanish university. *Education Sciences*, 11(6). <https://doi.org/10.3390/educsci11060293>
- Giunti, G., Claes, M., Zubiete, E. D., Rivera-Romero, O., & Gabarron, E. (2021). Impact of COVID-19 on multiple sclerosis topic discussion on Twitter. In *Public Health and Informatics: Proceedings of MIE 2021*. <https://doi.org/10.3233/SHTI210302>
- Jung, Y. (2018). Multiple predicting K-fold cross-validation for model selection. *Journal of Nonparametric Statistics*, 30(1). <https://doi.org/10.1080/10485252.2017.1404598>
- Kurniawan, r., & apriliani, a. (2020). Analisis sentimen masyarakat terhadap virus corona berdasarkan opini dari twitter berbasis web scraper. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*. <https://doi.org/10.24252/instek.v5i1.13686>
- Maynard, D., Bontcheva, K., & Augenstein, I. (2017). Natural Language Processing for the Semantic Web. In *Synthesis Lectures on the Semantic Web: Theory and Technology* (Vol. 6, Issue 2). <https://doi.org/10.2200/S00741ED1V01Y201611WBE015>

- Myslin, M., Zhu, S. H., Chapman, W., & Conway, M. (2013). Using twitter to examine smoking behavior and perceptions of emerging tobacco products. *Journal of Medical Internet Research*. <https://doi.org/10.2196/jmir.2534>
- Pilehvar, M. T., & Camacho-Collados, J. (2020). Embeddings in Natural Language Processing: Theory and Advances in Vector Representations of Meaning. In *Synthesis Lectures on Human Language Technologies* (Vol. 13, Issue 4). <https://doi.org/10.2200/S01057ED1V01Y202009HLT047>
- Samsir, Ambiyar, Verawardina, U., Edi, F., & Watrianthos, R. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *Jurnal Media Informatika Budidarma*.
- Singh, S., & Mahmood, A. (2021). The NLP Cookbook: Modern Recipes for Transformer Based Deep Learning Architectures. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3077350>
- Soufyane, A., Abdelhakim, B. A., & Ahmed, M. Ben. (2021). An intelligent chatbot using NLP and TF-IDF algorithm for text understanding applied to the medical field. In *Advances in Science, Technology and Innovation*. https://doi.org/10.1007/978-3-030-53440-0_1
- Syah, R. H. (2020). Dampak Covid-19 pada Pendidikan di Indonesia: Sekolah, Keterampilan, dan Proses Pembelajaran. *SALAM: Jurnal Sosial Dan Budaya Syar-I*, 7(5). <https://doi.org/10.15408/sjsbs.v7i5.15314>
- Watrianthos, R. (2020). Analisis Pembelajaran Daring di Era Pandemic Covid-19. *Merdeka Kreatif Di Era Pandemi Covid-19*.
- Wong, T. T., & Yeh, P. Y. (2020). Reliable Accuracy Estimates from k-Fold Cross Validation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8). <https://doi.org/10.1109/TKDE.2019.2912815>
- Zhang, L., Ghosh, R., Dekhil, M., Hsu, M., & Liu, B. (2011). Combining lexicon-based and learning-based methods for twitter sentiment analysis. *HP Laboratories Technical Report*.
- Zhou, M., Duan, N., Liu, S., & Shum, H. Y. (2020). Progress in Neural NLP: Modeling, Learning, and Reasoning. In *Engineering* (Vol. 6, Issue 3). <https://doi.org/10.1016/j.eng.2019.12.014>