

ANALISIS SENTIMEN KEBIJAKAN PEMBELAJARAN TATAP MUKA SELAMA PANDEMI COVID-19 MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*

Candra Bella Vista¹, Osa Mahanani Sihono², Annisa Taufika Firdausi³

^{1,2,3} Jurusan Teknologi Informasi, Politeknik Negeri Malang

¹bellavista@polinema.ac.id, ²1841720066@gmail.com, ³annisa.taufika@polinema.ac.id

Abstrak

Sebagai upaya memulihkan pembelajaran setelah pandemi covid-19, Kemendikbudristek mengeluarkan kebijakan memperbolehkan pembelajaran secara tatap muka di semester genap tahun ajaran 2022 secara terbatas. Keadaan tersebut tidak bertahan lama karena setelah kegiatan pembelajaran tatap muka dilaksanakan, angka covid naik kembali. Sehingga banyak opini pro dan kontra terkait pembelajaran tatap muka di tengah kenaikan kasus covid dan kemunculan berbagai varian virus covid. *Twitter* sebagai salah satu media sosial yang paling banyak digunakan oleh masyarakat Indonesia untuk menyampaikan opini. Opini mengenai pembelajaran tatap muka selama pandemi covid-19, sempat menduduki *trending topic* di Indonesia dalam beberapa waktu. Hal ini membuka peluang untuk dilakukan analisis sentimen terkait pembelajaran tatap muka (ptm). Pada penelitian ini penulis melakukan analisis sentimen pada data *Twitter* terkait pembelajaran tatap muka (ptm) menggunakan metode Support Vector Machine (SVM) dan fitur TF-IDF untuk mengekstraksi sebuah data opini untuk melihat kecenderungan kalimat tersebut bernilai positif, negatif, atau netral. Penerapan tahap *preprocessing* menggunakan *cleaning*, *case folding*, *tokenizing*, dan *stopword removal*. Pengujian dilakukan untuk mencari nilai akurasi, *precision*, *recall*, dan *f-measure* berdasarkan rasio data latih dan data. Rasio yang digunakan sebagai perbandingan adalah 60:40, 70:30, 80:20, dan 90:10. Hasil pengujian menunjukkan bahwa menggunakan perbandingan data latih dan data uji 90:10 mendapatkan akurasi tertinggi sebesar 73%.

Kata kunci : *support vector machine*, analisis sentimen, pembelajaran tatap muka

1. Pendahuluan

Covid-19 belum juga sepenuhnya hilang hingga saat ini meskipun vaksin sudah banyak digunakan oleh umum. Wabah yang belum jelas kapan berakhirnya ini menyebabkan banyak permasalahan, salah satunya dalam dunia pendidikan. Efek pandemi yang mengkhawatirkan membuat pemerintah mengambil keputusan untuk meliburkan sekolah dan pembelajaran dilaksanakan secara daring atau jarak jauh. Hasil penelitian menunjukkan bahwa pembelajaran jarak jauh menyebabkan masalah pada kognitif anak adalah anak-anak menjadi malas untuk belajar mandiri di rumah dikarenakan merasa kesepian, karena biasanya saat di sekolah bisa belajar bersama-sama dengan guru dan teman-temannya namun kali ini harus belajar mandiri di rumah mereka masing-masing (Royani, Nazurty, & Hamidah, 2021). Tanggal 21 Desember 2021, Kemendikbudristek menerbitkan Surat Keputusan Bersama (SKB) mengenai penyelenggaraan pembelajaran di masa pandemi covid-19 bahwa pemerintah berupaya memulihkan pembelajaran dengan kembali membuka sekolah secara tatap muka di semester genap tahun ajaran 2022 secara terbatas.

(kemendikbud, 2022). Keadaan tersebut tidak bertahan lama karena setelah sekolah diperbolehkan untuk menyelenggarakan kegiatan pembelajaran tatap muka, angka covid tiba-tiba naik kembali.

Twitter merupakan salah satu media sosial yang paling banyak digunakan masyarakat Indonesia untuk menyampaikan opini terkait berbagai hal. Indonesia tercatat sebagai peringkat 5 dunia kategori pengguna *Twitter* paling aktif di dunia (Annur, 2022). Kelebihan *Twitter* dibanding media sosial lainnya adalah penyebaran informasinya yang cepat sehingga hal ini memudahkan masyarakat untuk mendapatkan informasi secara *up to date* karena topik berita selalu *update* setiap saat. (Solihin, Awaliyah, & Shofa, 2021). Opini mengenai pembelajaran tatap muka selama pandemi covid-19, sempat menduduki *trending topic* di Indonesia dalam beberapa waktu. Banyak masyarakat yang beropini atau mengeluarkan pendapat mereka mengenai terselenggaranya kegiatan pembelajaran tatap muka di tengah angka covid yang melonjak lagi di sosial media. Hal ini membuka peluang untuk dilakukan analisis sentimen terkait pembelajaran tatap muka. Analisis sentimen merupakan metode yang digunakan untuk mengekstraksi sebuah data opini untuk melihat kecenderungan kalimat tersebut

bernilai positif, negatif, atau netral (Sari & Wibowo, 2019).

Umumnya sebuah kebijakan pemerintah akan selalu mendapatkan respon pro dan kontra dari masyarakat. Dikarenakan masyarakat akan menanggung dampak dari kebijakan pemerintah secara langsung. Dengan adanya kemajuan teknologi dan perkembangan media sosial, memudahkan pemerintah untuk memantau bagaimana tanggapan masyarakat terhadap kebijakan yang akan ataupun yang telah dikeluarkan. Salah satu cara untuk memudahkan untuk mengetahui kecondongan tanggapan masyarakat adalah dengan menggunakan analisis sentiment berdasarkan opini warga di media sosial *twitter*. Analisis sentimen terkait kebijakan pembelajaran tatap muka dapat digunakan oleh sarana pemerintah ataupun pemangku kepentingan untuk mengevaluasi bagaimana kecenderungan respon masyarakat terkait kebijakan yang akan diambil selanjutnya. Penelitian terkait analisis sentimen terhadap kebijakan pemerintah telah dilakukan sebelumnya untuk mengkaji kebijakan pemindahan ibukota RI (Sa'rony, 2019) dan analisis sentiment terkait penetapan kebijakan bencana nasional pandemi covid-19 (Fitri, 2020).

Berbagai metode telah diterapkan untuk mendapatkan akurasi terbaik dalam analisis sentimen. Penulis melakukan studi pustaka terhadap beberapa metode yang telah digunakan pada penelitian terdahulu, yaitu algoritma *K-Nearest Neighbor* (Sari, 2020) untuk analisis sentimen pelanggan obyek wisata Dunia Fantasi, *Naïve Bayes Classifier* (Saputra dkk., 2019) untuk analisis sentimen komentar channel video pelayanan pemerintah di Youtube, (Rozi dkk., 2018) analisis sentimen pelayanan samsat Kota Malang, dan *Support Vector Machine (SVM)* (Alhaq dkk, 2021) analisis sentimen *marketplace* Bukalapak. Penelitian terdahulu menyatakan bahwa algoritma SVM menghasilkan akurasi yang paling optimal dibandingkan *Naïve Bayes* dan KNN dalam analisis sentiment terhadap pandemi covid-19 di Indonesia (Pamungkas & Kharisudin, 2021). Kinerja SVM juga memberikan hasil lebih baik dibanding algoritma *Naïve Bayes* dalam menganalisis sentimen *review film* (Indrayuni, 2018). Sehingga berdasarkan pemaparan di atas penulis akan menggunakan metode *Support Vector Machine* sebagai algoritma pengklasifikasian sentimen positif, negatif, dan netral pada opini kebijakan pembelajaran tatap muka selama pandemi.

Sehingga berdasarkan latar belakang di atas penelitian ini bertujuan melakukan analisis sentimen pengguna *Twitter* terhadap kebijakan pembelajaran tatap muka selama pandemi menggunakan metode *Support Vector Machine (SVM)*. Hasil analisis sentimen yang dilakukan diharapkan dapat menjadi

masukan bagi pemangku kepentingan dalam merumuskan kebijakan pembelajaran tatap muka kedepannya.

2. Metode Penelitian

2.1 Sumber Data

Data yang digunakan adalah data dari *Twitter* yang dikumpulkan dalam kurun waktu 1 Januari 2022 sampai dengan 5 Agustus 2022. Proses pengumpulan data dilakukan dengan *crawling* data *tweet* memanfaatkan fasilitas *Twitter API*. Data yang diambil adalah data *tweet* yang berkaitan dengan beberapa kata kunci, yaitu pembelajaran tatap muka, ptm, ptm covid, ptm omicron. Jumlah data terkumpul sebanyak 1694 data.

2.2 Praproses Data

Data yang terkumpul akan dilakukan praproses terlebih dahulu. Praproses data bertujuan untuk menyeragamkan dan mengurangi volume kata yang tidak mempengaruhi klasifikasi (Widowati & Sadikin, 2020). Praproses data pada penelitian ini meliputi *cleansing*, *case folding*, *tokenizing*, dan *stopword removal*.

Tahapan *cleansing* adalah proses membersihkan data *tweet* untuk menghilangkan karakter yang tidak diperlukan. *Cleansing* bertujuan untuk mengurangi *noise* pada *dataset*. Proses *cleansing* akan menghilangkan tanda baca, *link url*, *username*, ikon emoji, *hashtag*, ataupun email.

Praproses data bertujuan untuk menyeragamkan data. *Tweet* yang terkumpul memiliki format huruf yang masih belum seragam, sehingga diperlukan proses *case folding*. *Case folding* adalah praproses data untuk menyeragamkan data menjadi huruf kecil.

Tahapan selanjutnya dari praproses data adalah *tokenizing*. Pada proses tokenisasi, akan dilakukan pemecahan kalimat menjadi kata-kata yang memiliki makna. Setelah itu dilakukan proses *stopword removal*. Proses ini akan menghilangkan kata-kata umum yang tidak begitu penting dalam proses klasifikasi. Contoh *stopword* adalah kata di, ke, yang, itu, oleh, terhadap, dan lain sebagainya.

2.3 Pelabelan Data

Untuk melakukan analisis sentiment menggunakan metode SVM diperlukan suatu data pelatihan. Pada penelitian data pelatihan yang digunakan akan dilabeli secara manual. Aturan pelabelan mengikuti penelitian sebelumnya, dimana kalimat akan dilabeli berdasarkan jumlah kata positif dan negatif di dalam kalimat tersebut (Santoso & Nugroho, 2019). Setiap kalimat akan dihitung skornya, skor diperoleh menurut persamaan 1.

$$\text{Skor} = \sum \text{kata positif} - \sum \text{kata negatif} \quad (1)$$

Jika dalam kalimat yang dilabeli mempunyai skor > 0 akan diklasifikasikan ke dalam kelas positif, jika kalimat mempunyai skor = 0 akan diklasifikasikan ke dalam kelas netral, sedangkan apabila kalimat mempunyai skor < 0 diklasifikasikan ke dalam kelas negatif

2.4 TF-IDF

Metode *Terms Frequency – Inverse Document Frequency* (TF-IDF) merupakan salah satu metode yang tepat untuk digunakan dalam pencarian kata di ditiap dokumen. Metode ini menggabungkan dua konsep untuk perhitungan bobot dengan frekuensi kemunculan sebuah kata di dalam sebuah dokumen tertentu dan inverse frekuensi dokumen yang mengandung kata tersebut (Luthfi & Lhaksamana, 2020). Tabel 1 menunjukkan pembobotan dari setiap kata pada setiap dokumen (tf).

Tabel 1. Pembobotan tf

Term	D11	D2	D3	D4	tf
doa	1	0	0	0	1
semoga	1	1	0	0	2
senin	1	0	0	0	1
ptm	1	0	1	0	2
stress	1	0	0	0	1
rumah	1	0	0	0	1
daring	0	1	0	0	1
jumat	0	1	0	0	1
temu	0	1	0	1	2
tatap	0	1	0	1	2
muka	0	1	0	1	2
batas	0	1	0	0	1
amin	0	1	0	0	1
besok	0	0	1	0	1
padahal	0	0	1	0	1
kena	0	0	1	0	1
covid	0	0	1	0	1
dilema	0	0	0	1	1

Setelah mendapatkan nilai tf, selanjutnya adalah menghitung nilai idf dengan rumus seperti pada persamaan 2.

$$idf = \log \left(\frac{\text{jumlah dokumen}}{tf} \right) \quad (2)$$

Hasil perhitungan idf ditunjukkan pada Tabel 2.

Tabel 2. Hasil perhitungan idf

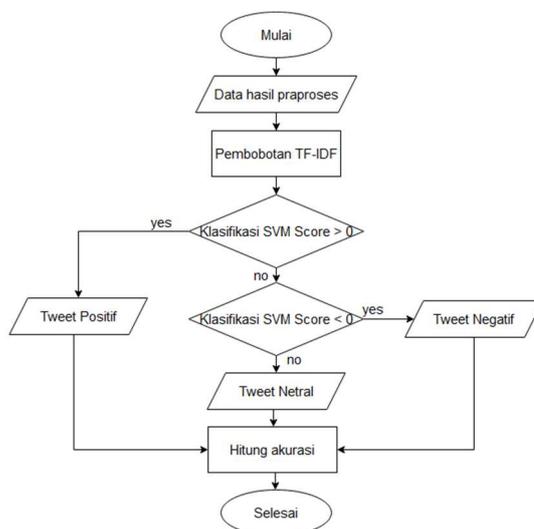
Term	tf	idf
doa	1	0.60205999132
semoga	2	0.30102999566
senin	1	0.60205999132
ptm	2	0.30102999566
stress	1	0.60205999132
rumah	1	0.60205999132
daring	1	0.60205999132
jumat	1	0.60205999132
temu	2	0.30102999566
tatap	2	0.30102999566
muka	2	0.30102999566

batas	1	0.60205999132
amin	1	0.60205999132
besok	1	0.60205999132
padahal	1	0.60205999132
kena	1	0.60205999132
covid	1	0.60205999132
dilema	1	0.60205999132

Perhitungan pembobotan TF-IDF merupakan perkalian yang dilakukan dari pembobotan *terms frequency* dengan *inverse document frequency*.

2.5 SVM

Analisis sentimen menggunakan metode *support vector machine*. Gambar 1 menunjukkan alur kerja algoritma SVM.



Gambar 1. Alur proses SVM

Setelah tahapan pembobotan TF-IDF selesai maka dilanjutkan dengan proses pelatihan menggunakan metode SVM dengan Langkah-langkah berikut:

1. Kernelisasi data latih. Kernel yang digunakan pada penelitian ini adalah kernel *polynomial* seperti pada persamaan 3.
$$K(x, x_j) = (x * x_j + 1)^d \quad (3)$$
2. Berikutnya menghitung matriks *hessian* sesuai persamaan 4.
$$D_{ij} = y_i y_j (K(x_i x_j) + \lambda^2) \quad (4)$$
3. Menghitung *error rate* seperti pada persamaan 5.
$$E_i = \sum_{j=1}^i a_{ij} D_{ij} \quad (5)$$
4. Menghitung nilai δa_i dan a_i menggunakan persamaan 6 dan 7.
$$\delta a_i = \min \{ \max [\gamma(1 - E_i) - a_0, C - a_0] \} \quad (6)$$

$$a_i = a_i + \delta a_i \quad (7)$$
5. Ulangi langkah ke-4 hingga iterasi maksimal terpenuhi atau dengan melihat nilai .
6. Setelah diperoleh nilai *support vector*, dilanjutkan menentukan dan dengan memilih *support vector* masing-masing kelas dengan nilai α terbesar.

- Saputra, P.Y., Subhi, D.H., & Winatama, F. (2019): *Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah di Youtube menggunakan Algoritma Naïve Bayes*. Jurnal Informatika Polinema Vol. 5 No.4.
- Rozi, I.F., Hamdana, E.N., & Alfahmi, M. (2018): *Pengembangan Aplikasi Analisis Sentimen Twitter menggunakan Metode Naïve Bayes Classifier (Studi Kasus SAMSAT Kota Malang)*. Jurnal Informatika Polinema Vol 4, No.2.
- Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J.D. (2021): *Penerapan Metode Support Vector Machine untuk Analisis Sentimen Pengguna Twitter*. JOISM (Jurnal of Information System Management) Vol. 3., No.1.
- Pamungkas, F.S. & Kharisudin, I. (2021): *Analisis Sentimen dengan SVM, Naïve Bayes, dan KNN untuk Studi Tanggapan Masyarakat Indonesia terhadap Pandemi Covid-19 pada Media Sosial Twitter*.
- Indrayuni, E. (2018): *Komparasi Algoritma Naïve Bayes dan Support Vector Machine untuk Analisa Sentimen Review Film*. Jurnal PILAR Nusa Mandiri Vol. 14., No.2.
- Widowati, T. & Sadikin, M. (2020): *Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naïve Bayes dan Support Vector Machine*. Jurnal SIMETRIS Vol. 11., No. 2.
- Santoso, B. & Nugroho, A. (2019): *Analisis Sentimen Calon Presiden Indonesia 2019 berdasarkan Komentar Publik di Facebook*. Jurnal Eksplora Informatika.
- Luthfi, M.F. & Lhaksamana, K.M. (2020): *Implementation of TF-IDF Method and Support Vector Machine Algorithm for Job Applicants Text Classification*. Jurnal Media Informatika Budidarma Vol.4, No.4.