

Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19

(Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic)

Brian Laurensz¹, Eko Sedyono²

Abstract—Coronavirus has become a global pandemic and has spread almost all over the world, including Indonesia. Many negative impacts resulted from the spread of COVID-19 in Indonesia, so the government made vaccination measures to reduce the rate of spread of COVID-19. Responses from the public to vaccination measures are quite diverse on social media Twitter. Some are supportive and some disagree. The purpose of this study is to find out how people's sentiment towards vaccination measures. The data used 845 tweets, using two keywords, "vaksinmerahputih" and "vaksinsinovac." The data is then divided into 253 training data and 592 testing data. The classification will use the SVM and Naïve Bayes methods. The classification result of the Naïve Bayes method received an average accuracy of 85.59%, while SMV of 84.41%. Sentiment results on Naïve Bayes method with keyword "vaksinsinovac" gets positive sentiment of 66% and negative sentiment of 34%, while "vaksinmerahputih" obtains 89% and 11% for positive and negative sentiment, respectively. SVM method with keyword "vaksinsinovac" gets 96% positive and 4% negative, while "vaksinmerahputih" obtains 98% positive and 2% negative. It can be concluded that the results of public sentiment towards vaccination measures received a positive response.

Intisari—Virus corona telah menjadi pandemi dunia dan sudah menyebar hampir ke seluruh dunia, termasuk Indonesia. Banyak dampak negatif yang diakibatkan oleh penyebaran COVID-19 di Indonesia, sehingga pemerintah mengambil tindakan vaksinasi agar dapat menekan tingkat penyebaran COVID-19. Tanggapan dari masyarakat terhadap tindakan vaksinasi cukup beragam di media sosial Twitter, ada yang mendukung dan ada juga yang tidak setuju. Makalah ini bertujuan untuk mengetahui sentimen masyarakat terhadap tindakan vaksinasi. Data yang digunakan sebanyak 845 tweet, dengan menggunakan dua kata kunci, yaitu "vaksinmerahputih" dan "vaksinsinovac". Data kemudian dibagi menjadi 253 data latih dan 592 data uji. Klasifikasi dilakukan menggunakan metode SVM dan Naïve Bayes. Hasil klasifikasi dari metode Naïve Bayes mendapatkan rata-rata akurasi 85,59%, sedangkan SMV sebesar 84,41%. Hasil sentimen pada metode Naïve Bayes dengan kata kunci "vaksinsinovac" mendapatkan sentimen positif 66% dan negatif 34%, sedangkan "vaksinmerahputih" memperoleh sentimen positif 89% dan negatif 11%. Metode SVM dengan kata kunci "vaksinsinovac" mendapatkan sentimen positif 96% dan negatif 4%, sedangkan "vaksinmerahputih" mendapatkan sentimen positif 98% dan

negatif 2%. Dapat disimpulkan bahwa hasil sentimen masyarakat terhadap tindakan vaksinasi mendapat respons yang positif.

Kata Kunci—Virus Corona, Indonesia, Vaksin, Sentimen, SVM, Naïve Bayes.

I. PENDAHULUAN

Virus corona mulai menjadi wabah pada bulan November-Desember 2019 di kota Wuhan, China. Virus ini merupakan salah satu virus yang sangat berbahaya karena tingkat penyebarannya yang tinggi sehingga meluas dengan cepat ke seluruh dunia. Menurut catatan WHO, pada tahun 2020 sudah banyak laporan dari berbagai negara yang mengonfirmasi terjangkit virus corona atau COVID-19 [1].

Di Indonesia pertama kali terdeteksi adanya warga yang terjangkit virus corona pada tanggal 2 maret 2020, yang terjadi pada dua orang warga Depok, Jawa Barat [2]. Semenjak saat itu, dari catatan satgas pemulihan COVID-19, diketahui semakin banyak kasus yang terkonfirmasi positif dari bulan ke bulan. Kondisi penyebaran virus yang semakin meluas di Indonesia ini menimbulkan masalah bukan hanya di bidang kesehatan saja, tetapi juga ekonomi, pendidikan, dan lainnya, yang ikut terkena dampak

Upaya pemerintah Indonesia sudah dilakukan untuk menekan tingkat penyebaran virus corona supaya dampak negatif yang ditimbulkan dapat dikendalikan, di antaranya dengan melakukan tindakan vaksinasi. Terdapat dua jenis vaksin yang dipakai oleh pemerintah Indonesia, yaitu vaksin Sinovac dan vaksin merah-putih. Penggunaan kedua vaksin ini mendapatkan berbagai macam tanggapan dan pendapat dari masyarakat. Pendapat yang disampaikan ada yang baik dan membangun, tetapi ada juga yang bertentangan dan menolak. Media sosial Twitter menjadi salah satu tempat masyarakat dapat dengan bebas menyampaikan pendapat [3].

Banyak metode analisis yang dapat digunakan untuk menganalisis pendapat masyarakat berdasarkan informasi yang ada pada media sosial semacam Twitter. Salah satu di antaranya adalah metode analisis sentimen. Metode analisis sentimen merupakan salah satu metode untuk menganalisis data yang didapatkan dari internet sehingga dapat diketahui polaritas dari data tersebut [4]. Dengan menggunakan analisis sentimen, polaritas dari opini yang ada dapat dikumpulkan, sehingga akan dapat digunakan untuk memprediksi suasana publik atau gambaran perasaan netizen bersifat negatif atau positif [4].

Beberapa penelitian tentang klasifikasi sentimen pada konten media sudah dilakukan sebelumnya. Dalam sebuah penelitian, dilakukan prediksi penyebaran COVID-19 di

^{1,2} Program Studi Magister Sistem Informasi, Universitas Kristen Satya Wacana, Jl. Dr. O. Notohamidjodjo, Kota Salatiga, INDONESIA, 50715 (e-mail: Brianlaurensz@gmail.com, Ekosed1@yahoo.com)

Indonesia dengan menggunakan algoritme *Naïve Bayes* [1]. Data yang digunakan dikumpulkan secara kuantitatif dengan fokus pada penggunaan tabel, angka, dan grafik. Algoritme yang digunakan untuk melakukan analisis adalah *Naïve Bayes*, dengan data yang didapatkan akan diproses untuk menghasilkan nilai probabilitas yang berbeda-beda pada setiap kriteria *class*. Nilai probabilitas yang didapatkan digunakan untuk memprediksi tingkat penyebaran COVID-19. Hasilnya menunjukkan bahwa tingkat keakuratan klasifikasi sentimen dengan menggunakan *Naïve Bayes* mendapatkan persentase akurasi 48,48%, dengan enam belas data berhasil diklasifikasi dengan benar dari total 33 data yang digunakan. Hasil dari penelitian ini digunakan untuk melihat tingkat penyebaran COVID-19 di Indonesia [1].

Dalam penelitian lain, dilakukan analisis sentimen terhadap pengguna jasa transportasi daring (*online*) di Indonesia dengan menggunakan metode *Support Vector Machine* (SVM) berbasis *Particle Swarm Optimization* (PSO) [4]. Data yang digunakan sebanyak 1.852 *tweet* dan dibagi menjadi data latih (*training*) dan data uji (*testing*). Tujuan dari penelitian ini yaitu untuk mengetahui sentimen pengguna transportasi daring di Indonesia serta membandingkan tingkat akurasi dari SVM dan SVM-PSO. Penggunaan nilai *default k-fold* atau 10 *k-fold* dalam penelitian ini mendapatkan hasil akurasi SVM 95,46% serta nilai AUC 0,979, sedangkan untuk akurasi, SVM-PSO mendapatkan nilai akurasi sebesar 96,04% serta AUC 0,993. Algoritme SVM berbasis PSO dalam penelitian ini mempunyai tingkat akurasi yang lebih baik dari algoritme SVM biasa. Hasil dari penelitian ini digunakan untuk melihat tanggapan masyarakat Indonesia terhadap pelayanan yang didapat dari jasa transportasi daring [4].

Telah dilakukan juga penelitian yang membahas tentang analisis sentimen publik terhadap kebijakan yang dibuat oleh presiden Joko Widodo untuk menghadapi wabah COVID-19 [5]. Dalam penelitian ini, data diambil menggunakan metode *web scraping* dengan kata kunci “Jokowi” dan “covid”. Proses klasifikasi data dalam penelitian ini menggunakan metode *Naïve Bayes*, *k-NN*, dan SVM. Hasil dari penelitian ini menunjukkan bahwa SVM mempunyai rata-rata nilai terbaik dibandingkan dengan metode lain yang digunakan dalam penelitian ini, dengan tingkat persentase akurasi sebesar 84,58%, *recall* 85,82%, dan presisi 82,14%. Tujuan penelitian ini yaitu untuk mengklasifikasikan sentimen dari media sosial Twitter sesuai dengan kata kunci yang dipakai. Hasil dari penelitian ini digunakan sebagai indikator keberhasilan kebijakan yang dibuat oleh pemerintah [5].

Beberapa penelitian yang sudah dilakukan sebelumnya dapat membuktikan bahwa untuk mengetahui sentimen masyarakat terhadap sebuah masalah yang terjadi dan menjadi topik pembicaraan di sosial media, dapat dilakukan analisis sentimen menggunakan *machine learning*. Berdasarkan hal tersebut dan latar belakang masalah yang terjadi, dilakukan penelitian tentang analisis sentimen masyarakat di media sosial Twitter terhadap tindakan vaksinasi yang dibuat oleh pemerintah. Proses klasifikasi dalam makalah ini menggunakan metode *Naïve Bayes* dan SVM.

Fokus pada makalah ini yaitu melakukan perbandingan hasil analisis sentimen dari vaksin Sinovac dan vaksin merah-putih. Jika sentimen yang dihasilkan positif, tindakan vaksinasi yang dibuat dinilai sangat tepat oleh masyarakat, dan apabila sentimen negatif, berarti masyarakat tidak setuju atau banyak yang menolak vaksinasi tersebut. Dalam makalah ini dibandingkan tingkat akurasi dari kedua algoritme yang digunakan. Pengujian validasi dilakukan dengan nilai 10 *k-fold cross validation*.

II. KONTEN UTAMA

A. Analisis Sentimen

Analisis sentimen atau *opinion mining* merupakan cabang ilmu dari *data mining* yang biasanya digunakan untuk menganalisis data tekstual berupa sebuah opini yang mengandung polaritas sehingga nantinya menghasilkan sebuah informasi yang memiliki nilai positif, negatif, atau netral [4], [6], [7]. Di dalam proses klasifikasi terdapat tiga metode yang digunakan, yaitu *lexicon base*, *hybrid approach*, dan *machine learning* [4], [7]. Metode *machine learning* digunakan dalam makalah ini karena dengan metode ini dapat diprediksi polaritas sentimen dari data yang digunakan [4], [7], [8].

B. Support Vector Machine (SVM)

SVM adalah teknik *supervised learning* yang mempunyai tingkat akurasi dan kualitas yang baik sehingga membuatnya menjadi sangat diminati di antara algoritme yang lainnya. Akan tetapi, untuk implementasinya diperlukan tahap pelatihan *sequential training* dan harus melalui proses pengujian [4], [9].

Proses analisis akan dimulai dengan mengubah data teks yang ada menjadi sebuah data vektor, kemudian akan dikombinasikan dengan menggunakan *Term Frequency Inverse Document Frequency* (TF-IDF) untuk dilakukan pembobotan [7]. Kelebihan metode SVM adalah kemampuannya mengidentifikasi *hyperplane* yang terpisah sehingga bisa memaksimalkan margin dari kelas yang berbeda. Metode SVM juga memiliki kekurangan, yaitu pada masalah yang mempunyai fitur yang sama dapat memengaruhi tingkat akurasi secara signifikan [10]–[12].

C. Naïve Bayes

Naïve Bayes adalah metode probabilitas yang pertama kali diperkenalkan oleh seorang ilmuwan asal Inggris, yaitu Thomas Bayes. *Naïve Bayes* digunakan untuk memprediksi peluang di masa depan dengan menggunakan pengalaman yang ada di masa lalu [7], [6], [13]. Penerapan *Naïve Bayes* ke dalam analisis sentimen harus mempunyai dua proses penting, yaitu pelatihan dan pengujian [4], [7]. Kelebihan *Naïve Bayes* terletak pada kecilnya jumlah data latih yang digunakan sehingga perhitungan dapat dilakukan lebih cepat dan efisien [1]. Adapun kelemahan *Naïve Bayes* yaitu jika pada pemilihan fitur terjadi kesalahan, tingkat akurasi akan menurun dan waktu perhitungan akan semakin bertambah [5].

D. Pre-processing

Data teks yang diambil dari media sosial Twitter terkadang memberikan tantangan tersendiri karena terdapatnya kata yang

tidak baku, penggunaan bahasa daerah, atau singkatan yang tidak ada di dalam Kamus Besar Bahasa Indonesia (KBBI) [7]. Untuk membuat data menjadi lebih terstruktur, diperlukan *pre-processing*. Proses ini dilakukan untuk membuat sebanyak mungkin bahasa yang tidak baku menjadi bahasa baku dan untuk membersihkan data dari *noise* yang ada sehingga data menjadi lebih terstruktur sehingga siap untuk dianalisis [4], [13].

Terdapat beberapa proses dalam *pre-processing*, yaitu sebagai berikut.

1) *Transform Cases*: Kata-kata yang tidak berhubungan akan diubah. Misalnya, kata yang mengandung huruf besar akan diubah semuanya menjadi huruf kecil atau *lower case*.

2) *Tokenize*: Dalam proses ini, tanda baca, karakter khusus, simbol, atau karakter yang bukan termasuk huruf akan dihapus

3) *Filter Token (By Length)*: Kata-kata yang lebih pendek dari empat huruf atau lebih dari 25 huruf akan dihapus, misalnya 'tdk', 'yg', dan 'gan'.

4) *Stopword Removal*: Kata atau konjungsi yang tidak relevan, misalnya tetapi, dengan, untuk, yang, dan kaya sambung lainnya, akan dihapus

5) *Stemming*: Kata yang berimbuhan, seperti mem-, me-, meny-, meng-, per-, dan ber-, akan diseleksi dan diubah menjadi kata dasar.

6) *Generate n-Grams*: Untuk menyelesaikan masalah dalam proses klasifikasi, *n-grams* akan digunakan. Kesalahan yang biasanya terjadi adalah penggunaan *term* tunggal. Contohnya, kata 'buruk' termasuk dalam sentimen negatif, tetapi jika disandingkan dengan kata negasi akan menjadi sentimen positif, misalnya 'tidak buruk'.

E. Term Weighting

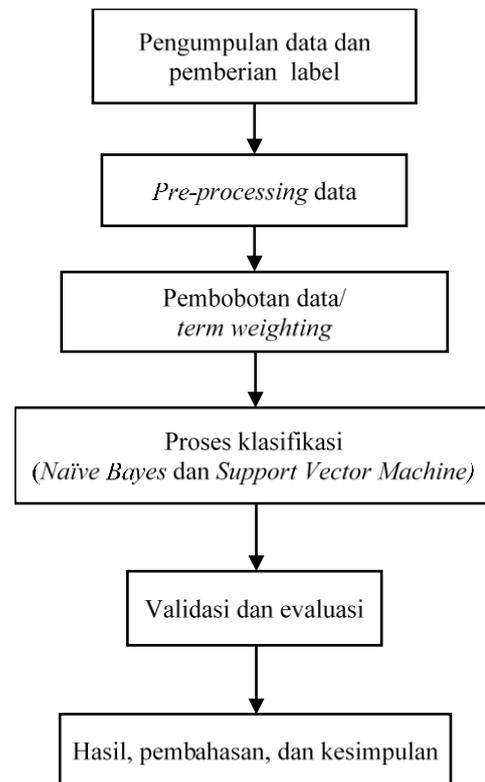
Terhadap kata-kata yang sudah melalui *pre-processing* dan sudah berbentuk data numerik akan dilakukan pembobotan [5]. *Term weighting* merupakan tahap pembobotan untuk setiap kata sehingga kemampuan dalam proses analisis sentimen menjadi meningkat [4], [7]. Dalam makalah ini, pembobotan dilakukan menggunakan TF-IDF. Dalam metode ini, pembobotan mengombinasikan frekuensi kata dan frekuensi dokumen [5], [7].

III. METODOLOGI

Pada makalah ini, ada beberapa tahapan yang perlu dilakukan agar diperoleh hasil yang terbaik. Tahap-tahap ini terdiri atas pengumpulan data dan pelabelan data, kemudian dilanjutkan dengan *pre-processing*. Data yang sudah melalui *pre-processing* akan diklasifikasi dengan menggunakan algoritme *Naïve Bayes* dan SVM. Proses pengujian akan dilakukan dengan metode *k-Fold Cross Validation*. Gbr. 1 menunjukkan tahapan-tahapan yang dilakukan.

A. Pengumpulan Data

Prose pengumpulan data menggunakan teknik *web scraping* dengan tools Octoparse. Data diambil dari bulan November



Gbr. 1 Tahapan penelitian.

2020 sampai Februari 2021, berdasarkan kata kunci "vaksinSinovac" dan "vaksinMerahputih". Data yang berhasil dikumpulkan sebanyak 845 *tweet* dari setiap kata kunci. Data *tweet* yang dipakai ini merupakan campuran dari *tweet* berita dan *tweet* tanggapan masyarakat. Data tersebut kemudian dipersiapkan untuk dibagi ke dalam data latih sebanyak 253 dan data uji sebanyak 592 *tweet*. Selanjutnya, data yang sudah dikumpulkan diberi label secara manual menurut ahli bahasa menjadi sentimen negatif atau positif.

B. Pre-processing

Data-data yang sudah dikumpulkan dan dibagi menjadi data uji dan latih akan melalui tahap *pre-processing* menggunakan tools RapidMinerStudio. Proses ini bertujuan agar data dapat diklasifikasi dan untuk mempermudah proses analisis.

C. Pembobotan TF/IDF

Tahap selanjutnya adalah melakukan pembobotan terhadap setiap kata berdasarkan frekuensi dari *term* atau istilah yang muncul pada dokumen. Pada makalah ini, pembobotan dilakukan dengan metode TF-IDF.

D. Klasifikasi

Proses klasifikasi dilakukan menggunakan algoritme *Naïve Bayes* dan SVM. Proses ini dilakukan setelah data yang digunakan melalui proses pembobotan.

E. Validasi

Data yang sudah selesai diklasifikasikan divalidasi dengan menggunakan *k-fold Cross Validation*. *K-fold Cross Validation*

TABEL I
SENTIMEN DATA LATIH

Kata Kunci	Positif	Negatif	Jumlah
“Vaksin Sinovac”	171 (68%)	82 (32%)	253 (100%)
“Vaksin merahputih”	227 (90%)	26 (10%)	253 (100%)

TABEL II
CONTOH HASIL *PRE-PROCESSING*

Tahapan <i>Pre-processing</i>	Hasil
Kalimat asli	<p>T1: Kami sangat mendukung langkah pak Erick lakukan membuat vaksin buatan Merah putih agar tidak ketergantungan impor. Semoga ditahun 2022 vaksin terealisasi dengan baik dan cepat VaksinMerahPutih</p> <p>T2: H+1 pasca VaksinSinovac , badan pegel2 kek abis workout. Lengan nyeri bgt di sekitaran bekas suntikan, ternyata ijo</p>
Cases folding	<p>T1: kami sangat mendukung langkah pak erick lakukan membuat vaksin buatan merah putih agar tidak ketergantungan impor. semoga ditahun 2022 vaksin terealisasi dengan baik dan cepat vaksinmerahputih</p> <p>T2: h+1 pasca vaksinsinovac, badan pegel2 kek abis workout. lengan nyeri bgt di sekitaran bekas suntikan, ternyata ijo</p>
Tokenizing	<p>T1: kami sngt mendukung langkah pak erick lakukan membuat vaksin buatan merah putih agar tidak ketergantungan impor semoga ditahun vaksin terealisasi dengan baik dan cepat vaksinmerahputih</p> <p>T2: h pasca vaksinsinovac badan pegel kek abis workout lengan nyeri bgt di sekitaran bekas suntikan ternyata ijo</p>
Filter token (by length)	<p>T1: kami sngt mendukung langkah erick lakukan membuat vaksin buatan merah putih agar tidak ketergantungan impor semoga ditahun vaksin terealisasi dengan baik cepat vaksinmerahputih</p> <p>T2: pasca vaksinsinovac badan pegel abis workout lengan nyeri sekitaran bekas suntikan ternyata</p>
Stopword removal	<p>T1: sngt mendukung langkah erick lakukan vaksin buatan merah putih ketergantungan impor semoga ditahun vaksin terealisasi cepat vaksinmerahputih</p> <p>T2: pasca vaksinsinovac badan pegel abis workout lengan nyeri sekitaran bekas suntikan</p>
Stemming	<p>T1: sngt mendukung langkah erick lakukan vaksin buatan merah putih ketergantungan impor semoga ditahun vaksin terealisasi cepat vaksinmerahputih</p> <p>T2: pasca vaksinsinovac badan pegel abis workout lengan nyeri sekitaran bekas suntikan</p>

merupakan metode untuk memisahkan data secara acak ke dalam K atau “fold”, serta menghilangkan *noise* atau *bias* pada kata sehingga dapat diperoleh tingkat akurasi yang tinggi. *K-fold* berfungsi untuk mengacak data ke dalam beberapa bagian,

TABEL III
CONTOH HASIL *PREDICTION* DATA LATIH

Row	Sentimen	Prediction (Sentimen)	Confidence (Positif)	Confidence (Negatif)
1	Negatif	Negatif	0,428	0,572
2	Negatif	Positif	0,608	0,392
3	Positif	Positif	0,702	0,298
4	Positif	Positif	0,748	0,252
5	Negatif	Positif	0,696	0,304

TABEL IV
PERSENTASE HASIL PENGUJIAN MENGGUNAKAN *NAÏVE BAYES*

Kata Kunci	Akurasi	Presisi	Recall
Vaksin Sinovac	78,68%	68,17%	70,83%
Vaksin Merahputih	92,50%	61,66%	65,00%
Rata-rata	85,59%	64,92%	67,92%

tetapi mekanisme ini biasanya menyebabkan rata-rata akurasi yang didapatkan menjadi tidak konstan [4], [7], [14].

IV. HASIL DAN PEMBAHASAN

Dalam makalah ini, proses analisis dilakukan menggunakan *tools* RapidMiner. Data latih dari setiap kata kunci yang digunakan dibagi secara manual sesuai dengan kelasnya, yaitu positif atau negatif, seperti yang ditunjukkan pada Tabel I. Selanjutnya, proses yang dilakukan adalah *pre-processing*. Tabel II merupakan contoh proses dalam *pre-processing*. Diambil dua *tweet* yang dianggap dapat mewakili proses dalam *pre-processing*. Total data latih yang digunakan yaitu 253 *tweet*.

Setelah melalui *pre-processing*, data latih akan ditentukan sebagai kelas positif atau negatif berdasarkan perhitungan SVM dan *Naïve Bayes*. Jika hasil *confidence* positif lebih besar daripada *confidence* negatif, maka hasil *prediction* dari kalimat tersebut positif. Sebaliknya, jika *confidence* negatif lebih besar daripada *confidence* positif, maka hasil *prediction* menjadi negatif. Pada Tabel III ditampilkan contoh hasil prediksi label data latih. Nantinya, data latih akan menjadi bahan pembelajaran untuk proses klasifikasi dengan menggunakan data uji.

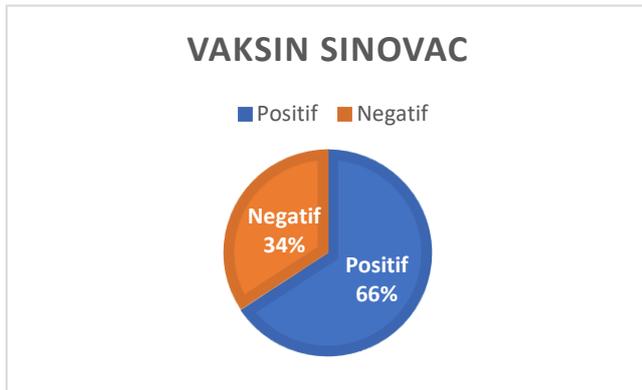
A. Analisis Data Menggunakan *Naïve Bayes*

Data latih yang sudah melewati *pre-processing* akan menjadi bahan pembelajaran untuk proses klasifikasi data uji menggunakan algoritme *Naïve Bayes*. Pengujian dengan *Naïve Bayes* menggunakan nilai 10 *k-fold*, dengan total dari data yang digunakan akan dibagi menjadi sepuluh bagian secara acak dan akan diulang-ulang sebanyak jumlah kelompok yang sudah ditentukan. Hasil dari pengujian data latih menggunakan *Naïve Bayes* diperlihatkan pada Tabel IV.

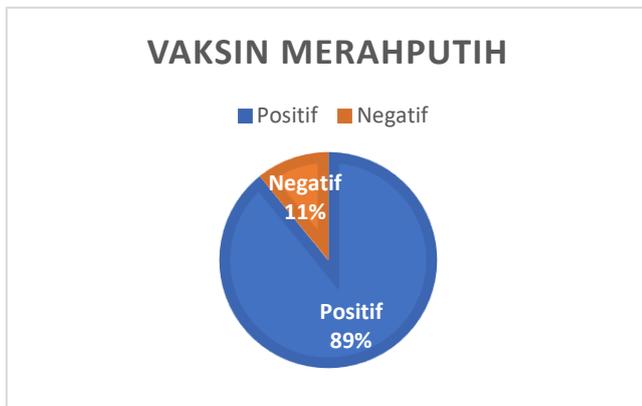
Berdasarkan hasil pengujian pada Tabel IV, dapat dilihat bahwa analisis menggunakan metode *Naïve Bayes* pada kata kunci “Vaksin Sinovac” mendapatkan tingkat akurasi sebesar 78,68%, presisi 68,17%, dan *recall* 70,83%, sedangkan pada kata kunci “Vaksin Merahputih”, diperoleh tingkat akurasi sebesar 92,50%, presisi 61,66%, dan *recall* 65,00%. Untuk nilai rata-rata tingkat akurasi dengan metode *Naïve Bayes*, diperoleh persentase sebesar 85,59%.

TABEL V
PERSENTASE HASIL PENGUJIAN MENGGUNAKAN SVM

<i>Kata Kunci</i>	<i>Akurasi</i>	<i>Presisi</i>	<i>Recall</i>
Vaksin Sinovac	77,05%	99,05%	32,56%
Vaksin Merahputih	91,77%	97,14%	28,93%
Rata-rata	84,41%	98,10%	30,75%



Gbr. 2 Hasil analisis sentimen menggunakan *Naïve Bayes* pada kata kunci “Vaksin Sinovac”.



Gbr. 3 Hasil analisis sentimen menggunakan *Naïve Bayes* pada kata kunci “Vaksin Merahputih”.

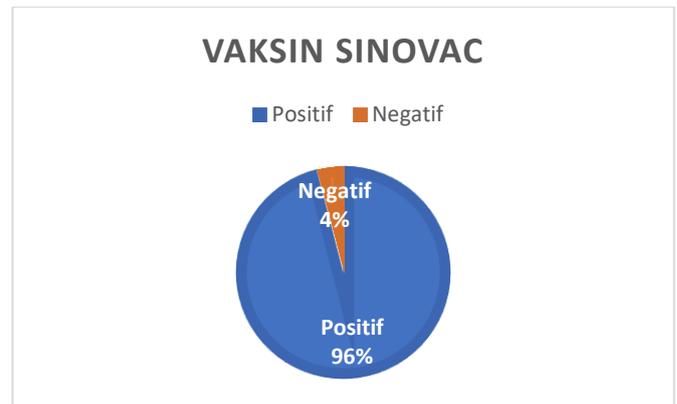
B. Analisis Data Menggunakan SVM

Pengujian dengan SVM menggunakan data latih dan proses pengujian yang sama dengan yang digunakan pada metode *Naïve Bayes*. Sebelumnya, data yang akan diproses diubah dari polinomial ke binomial agar dapat dianalisis dengan metode SVM. Hasil pengujian menggunakan SVM ditunjukkan pada Tabel V.

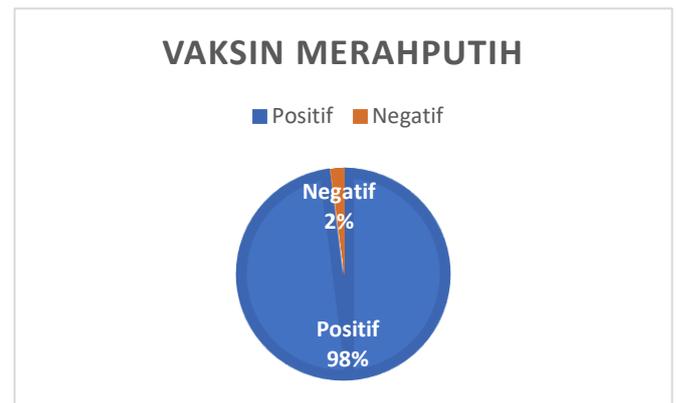
Berdasarkan hasil pengujian pada Tabel V, dapat dilihat bahwa hasil analisis dengan menggunakan metode SVM pada kata kunci “Vaksin Sinovac” mendapatkan tingkat akurasi sebesar 77,05%, presisi 99,05%, dan *recall* 32,56%. Pada kata kunci “Vaksin Merahputih” tingkat akurasi yang didapatkan sebesar 91,77%, presisi 97,14%, dan *recall* 28,93%. Untuk nilai rata-rata tingkat akurasi dengan menggunakan metode SVM, diperoleh persentase sebesar 84,41%.

C. Analisis Sentimen Menggunakan *Naïve Bayes* dan SVM

Hasil analisis sentimen dengan menggunakan *Naïve Bayes* pada kata kunci “Vaksin Sinovac” dan “Vaksin Merahputih”



Gbr. 4 Hasil analisis sentimen menggunakan SVM pada kata kunci “Vaksin Sinovac”.



Gbr. 5 Hasil analisis sentimen menggunakan SVM pada kata kunci “Vaksin Merahputih”.

ditunjukkan pada Gbr. 2 dan Gbr. 3. Gambar tersebut merupakan hasil sentimen yang dilakukan terhadap 592 *tweet* data uji dengan menggunakan metode *Naïve Bayes*. Sentimen positif menjadi sentimen yang paling mendominasi pada kedua kata kunci, dengan persentase 66% pada “vaksinsinovac” dan 89% pada “vaksinmerahputih”, sedangkan untuk sentimen negatif, dihasilkan persentase sebesar 34% pada kata kunci “vaksinsinovac” dan 11% “vaksinmerahputih”.

Selanjutnya, hasil analisis sentimen menggunakan metode SVM diperlihatkan pada Gbr. 4 dan Gbr. 5. Total data yang digunakan sama dengan yang dipakai pada metode *Naïve Bayes*, yaitu 592 *tweet*. Pada metode SVM, sentimen positif yang dihasilkan lebih banyak daripada metode *Naïve Bayes*. Persentase sentimen positif pada kata kunci “vaksinsinovac” adalah sebesar 96% dan sentimen negatif sebesar 4%. Nilai persentase sentimen positif yang didapatkan pada kata kunci “vaksinmerahputih” yaitu sebesar 98% dan sentimen negatif sebesar 2%.

V. KESIMPULAN

Hasil penelitian yang telah dilakukan dapat menunjukkan sentimen masyarakat terhadap tindakan vaksinasi yang dibuat oleh pemerintah Indonesia untuk menekan tingkat penyebaran COVID-19. Hasil analisis sentimen dengan metode *Naïve Bayes* menggunakan kata kunci “vaksinsinovac” mendapatkan nilai sentimen positif sebesar 66% dan negatif 34%, sedangkan

kata kunci “vaksinmerahputih” menghasilkan sentimen positif 89% dan negatif 11%. Hasil sentimen pada metode SVM menggunakan kata kunci “vaksinsinovac” mendapatkan persentase positif sebesar 96% dan negatif 4%, sedangkan persentase sentimen positif pada kata kunci “vaksinmerahputih” sebesar 98% dan negatif 2%.

Dari hasil ini dapat disimpulkan bahwa sentimen positif menjadi sentimen yang paling mendominasi pada kedua metode yang digunakan, dengan kata kunci “vaksinmerahputih” yang paling banyak mendapatkan sentimen positif. Oleh karena itu, dapat dinilai bahwa tindakan vaksinasi yang dibuat mendapat respons positif dari masyarakat. Faktor-faktor yang memengaruhi sentimen positif pada masyarakat adalah kesadaran masyarakat akan dampak yang disebabkan oleh pandemi COVID-19 sehingga tindakan vaksinasi dinilai sangat membantu dalam menekan tingkat penyebaran COVID-19, sedangkan sentimen negatif dihasilkan dari banyaknya berita hoaks yang beredar di masyarakat tentang vaksinasi, terlebih untuk vaksin Sinovac yang merupakan produksi China sehingga membuat masyarakat banyak berpendapat negatif.

Penelitian ini juga menunjukkan bahwa hasil klasifikasi dengan menggunakan metode *Naïve Bayes* mempunyai rata-rata tingkat akurasi lebih besar dengan persentase sebesar 85,59%, sedangkan metode SVM 84,41%. Dari hasil tersebut dapat diasumsikan bahwa metode klasifikasi dengan menggunakan *Naïve Bayes* adalah yang terbaik pada kasus ini.

Pengembangan selanjutnya dari penelitian ini yaitu perlunya dilakukan *pre-processing* yang lebih baik lagi sehingga data yang dianalisis menjadi lebih baik dan akan menambah tingkat akurasi. Selain itu, perlu ditambahkan sumber data dari media sosial yang lain, misalnya Facebook, sehingga data yang didapatkan lebih beragam. Perlu juga dilakukan pengujian dengan algoritme yang lain sehingga dapat dicari algoritme yang tepat untuk menganalisis kasus ini.

REFERENSI

- [1] A.F. Watratan, A. Puspita B., dan D. Moeis, “Implementasi Algoritma Naive Bayes untuk Memprediksi Tingkat Penyebaran Covid-19 di Indonesia,” *Journal of Applied Computer Science and Technology*, Vol. 1, No. 1, hal. 7–14, 2020.
- [2] A. Susilo, C.M. Rumende, C.W. Pitoyo, W.D. Santoso, M. Yulianti, dkk., “Coronavirus Disease 2019: Tinjauan Literatur Terkini,” *Jurnal Penyakit Dalam Indonesia*, Vol. 7, No. 1, hal. 45–67, 2020.
- [3] V.N. Setiawan (2020) “Riset Medsos: Publik Lihat Negatif Kebijakan Pemerintah Atasi Corona,” [Online], <https://katadata.co.id/agungjatmiko/berita/5ea5c764827c4/riset-medsos-publik-lihat-negatif-kebijakan-peemerintah-atasi-corona>, tanggal akses: 30-Sep-2020.
- [4] V.K.S. Que, A. Iriani, dan H.D. Purnomo, “Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, Vol. 9, No. 2, hal. 162–170, 2020.
- [5] S. Hikmawan, A. Pardamean, dan S.N. Khasanah, “Sentimen Analisis Publik terhadap Joko Widodo terhadap Wabah Covid-19 Menggunakan Metode Machine Learning,” *Jurnal Kajian Ilmiah*, Vol. 20, No. 2, hal. 167–176, 2020.
- [6] N.D. Susanti, E. Sedyono, dan I. Sembiring, “Uji Perbandingan Akurasi Analisis Sentimen Pariwisata Menggunakan Algoritma Support Vector Machine dan Naive Bayes,” *Nusantara of Engineering*, Vol. 3, No. 2, hal. 26–33, 2016.
- [7] H. Tuhuteru dan A. Iriani, “Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier,” *Jurnal Informatika: Jurnal Pengembangan IT*, Vol. 3, No. 3, hal. 394–401, 2018.
- [8] A. D’Andrea, F. Ferri, P. Grifoni, dan T. Guzzo, “Approaches, Tools and Applications for Sentiment Analysis Implementation,” *International Journal of Computer Applications*, Vol. 125, No. 3, hal. 26–33, 2015.
- [9] R. Moraes, J.F. Valiati, dan W.P.G. Neto, “Document-level Sentiment Classification: An Empirical Comparison Between SVM and ANN,” *Expert Systems with Applications*, Vol. 40, No. 2, hal. 621–633, 2013.
- [10] N. Yunita, “Analisis Sentimen Berita Artis dengan Menggunakan Algoritma Support Vector Machine dan Particle Swarm Optimization,” *Jurnal Sistem Informasi STMIK Antar Bangsa*, Vol. 5, No. 2, hal. 104–112, 2016.
- [11] J.S. Chou, M.Y. Cheng, Y.W. Wu, dan A.D. Pham, “Optimizing Parameters of Support Vector Machine Using Fast Messy Genetic Algorithm for Dispute Classification,” *Expert Systems with Applications*, Vol. 41, No. 8, hal. 3955–3964, 2014.
- [12] A.S.H. Basari, B. Hussin, I.G.P. Ananta, dan J. Zeniarja, “Opinion Mining of Movie Review Using Hybrid Method of Support Vector Machine and Particle Swarm Optimization,” *Procedia Engineering*, Vol. 53, hal. 453–462, 2013.
- [13] H. Irsyad, A. Farisi, dan M.R. Pribadi, “Klasifikasi Opini Masyarakat terhadap Jasa ISP MyRepublic dengan Naive Bayes,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, Vol. 8, No. 1, hal. 30–34, 2019.
- [14] E. Indrayuni, “Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization,” *Jurnal Evolusi*, Vol. 4, No. 2, hal. 20–27, 2016.