

Analisis Sentimen Masyarakat Indonesia Terhadap Dampak Penurunan Global Sebagai Akibat Resesi di Twitter

Stephen Aprius Sutresno*

Program Studi Sistem Informasi, Fakultas Teknik, Universitas Katolik Indonesia Atma Jaya, Jakarta, Indonesia

Email: stephen.sutresno@atmajaya.ac.id

Email Penulis Korespondensi: stephen.sutresno@atmajaya.ac.id

Submitted: 15/02/2023; Accepted: 31/03/2023; Published: 31/03/2023

Abstrak—Resesi memiliki arti yaitu penurunan signifikan di dalam aktivitas ekonomi dan tersebar di seluruh ekonomi dalam kondisi terbesar lebih dari beberapa bulan. Pemerintah dan masyarakat merupakan kunci peranan penting dalam kondisi ini, dibutuhkan perhatian khusus dan strategi yang tepat dalam melawan resesi ekonomi pada tahun 2023. Diperlukan analisa untuk mengetahui faktor-faktor apa saja yang dapat menyebabkan resesi, selain itu juga tanggapan mengenai dampak negatif yang menjadi kekhawatiran masyarakat luas. Akhir-akhir ini banyak sekali opini masyarakat mengenai resesi yang akan terjadi di tahun 2023 khususnya di negara Indonesia. Opini masyarakat ini bisa dijumpai di berbagai media sosial salah satunya adalah Twitter. Berdasarkan permasalahan tersebut, maka dilakukan analisis sentimen terhadap tweet untuk mencari informasi dari polaritas positif atau negatifnya opini tersebut menggunakan metode Naive Bayes dan Support Vector Machine (SVM) guna memilih metode yang lebih efektif dalam studi kasus untuk menentukan prediksi sentimen. Tahapan yang dilakukan terdiri dari pengumpulan data, pengolahan data, pembobotan data, proses klasifikasi, evaluasi dan validasi, serta hasil dan pembahasan. Digunakan teknik web scrapping dan setelah melalui tahapan cleaning data didapatkan total 780 data tweet. Hasil uji klasifikasi menunjukkan metode SVM memiliki tingkat akurasi lebih besar dengan persentase 79,5% dibandingkan metode Naive Bayes dengan persentase 72,5%. Hasil prediksi menggunakan metode SVM juga menunjukkan sejumlah 144 sentimen positif dan 636 negatif. Dilihat dari Wordcloud yang dibentuk maka dapat diasumsikan bahwa masyarakat khawatir dengan adanya kondisi ekonomi mereka, salah satunya adalah tidak stabilnya harga minyak yang dapat memicu terjadinya resesi.

Kata Kunci: Resesi; Analisis Sentimen; Naive Bayes; Support Vector Machine

Abstract—A recession is a significant reduction in economic activity and is spread across the economy at its greatest for more than a few months. The government and society play an important key role in this condition. Special attention and the right strategy are needed in fighting the economic recession in 2023. An analysis is required to find out what factors can cause a recession, as well as responses to the negative impacts that are a wide concern community. Recently, there has been a lot of public opinion regarding the recession that will occur in 2023, especially in Indonesia, on various social media such as Twitter. Based on these problems, sentiment analysis was carried out on tweets to obtain information on the positive or negative polarity of these opinions using the Naive Bayes and Support Vector Machine (SVM) methods to choose a more effective way in case studies to determine sentiment predictions. The steps are taken consist of data collection, processing data, weighting data, classification process, evaluation, validation, and results and discussion. The web scraping technique was used, and after going through the data cleaning stages, a total of 780 tweet data was obtained. The results of the classification test show that the SVM method has a greater accuracy rate with a proportion of 79.5% compared to the Naive Bayes method with a proportion of 72.5%. The SVM method's prediction results also show several 144 positive and 636 negative sentiments. Judging from the Wordcloud that was formed, it can be assumed that people are worried about their economic conditions, one of which is the unstable oil price which can trigger a recession.

Keywords: Recession; Sentiment Analysis; Naive Bayes; Support Vector Machine

1. PENDAHULUAN

Mendengar kata resesi sudah tidak lagi asing bagi hampir semua orang. Resesi memiliki arti yaitu penurunan signifikan di dalam aktivitas ekonomi dan tersebar di seluruh ekonomi dalam kondisi terbesar lebih dari beberapa bulan, selain itu juga dapat dilihat di PDB Rill, pendapatan Rill, Lapangan Kerja, Produksi Industri, maupun Penjualan Grosir-Eceran [1]. Resesi sendiri berawal dari pelemahan ekonomi dalam negeri di berbagai negara di dunia, dimana semakin besarnya ketergantungan ekonomi satu negara pada perekonomian global, maka semakin cepat pula terjadinya resesi di negara tersebut [2].

Pemerintah dan masyarakat merupakan kunci peranan penting dalam kondisi ini, dibutuhkan perhatian khusus dan strategi yang tepat dalam melawan resesi ekonomi pada tahun 2023. Diperlukan analisa untuk mengetahui faktor-faktor apa saja yang dapat menyebabkan resesi, selain itu juga tanggapan mengenai dampak negatif yang menjadi kekhawatiran masyarakat luas. Beberapa faktor penyebab terjadinya resesi ekonomi secara umum antara lain adanya guncangan atau persaingan ekonomi secara tiba-tiba, hutang yang berlebihan, gelembung aset investasi berlebihan di pasar saham, terlalu banyak inflasi atau deflasi, dan juga perubahan teknologi penemuan baru yang berguna untuk peningkatan produktivitas dan perekonomian dalam jangka waktu yang panjang [1, 3].

Akhir-akhir ini banyak sekali opini masyarakat mengenai resesi yang akan terjadi di tahun 2023 khususnya di negara Indonesia. Opini masyarakat ini bisa dijumpai di berbagai media sosial salah satunya adalah Twitter. Twitter adalah media sosial yang digunakan oleh semua kalangan untuk berkomunikasi tanpa batas dengan siapapun [4]. Hingga saat ini di tahun 2023 terdapat 368,4 juta akun Twitter aktif dari total 1,3 miliar [5]. Banyaknya pengguna Twitter yang menyampaikan opini-opini mengenai topik resesi, maka dapat dilakukan analisa terhadap *tweet* untuk mencari informasi dari polaritas positif atau negatifnya opini tersebut. Penambahan informasi ini membutuhkan

teknik analisis yang tepat sehingga dapat menghasilkan informasi yang berguna ke berbagai pihak yaitu dengan penggunaan analisis sentimen [6].

Salah satu penelitian sebelumnya mengenai analisis sentimen masyarakat terhadap tindakan vaksinasi dalam upaya mengatasi pandemi *COVID-19*. Berlatar belakang dari masyarakat yang mendukung atau bahkan tidak setuju terhadap tindakan vaksinasi yang dilakukan oleh pemerintah dilihat dari tanggapan mereka di medsos Twitter, sehingga untuk mengetahui sentimen masyarakat dilakukan analisis menggunakan uji klasifikasi dengan metode *Naive Bayes* dan SVM. Hasil dari penelitian menunjukkan penggunaan metode *Naive Bayes* dengan kata kunci pertama “vaksinsinovac” mendapatkan nilai sentimen positif 66% dan negatif 34%, kata kunci kedua “vaksinmerahputih” mendapatkan nilai sentimen positif 89% dan negatif 11%, sedangkan metode SVM dengan kata kunci pertama mendapatkan nilai sentimen positif 96% dan negatif 4%, kata kunci kedua mendapatkan nilai sentimen positif 98% dan negatif 2%. Selain dari hasil tersebut juga dilihat bahwa metode *Naive Bayes* mendapatkan nilai rata-rata akurasi lebih besar dengan persentase 85,59%, sedangkan SVM sebesar 84,41% [7].

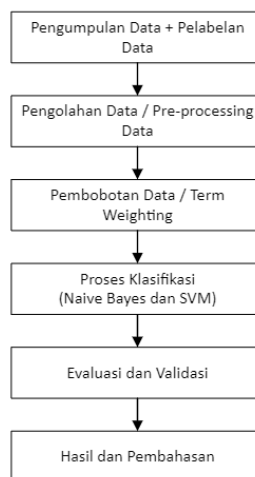
Penelitian lainnya mengenai analisis sentimen kuliah *online* pasca *COVID-19*. Berawal dari cepatnya penyebaran virus, maka pemerintah memberikan status penguncian daerah dan memberlakukan kuliah *online* atau daring. Melihat postingan opini dan pendapat masyarakat pada Twitter sebagai instrumen untuk pengukuran tingkat kepuasan masyarakat mengenai belajar *online*, maka dilakukan sentimen analisis menggunakan algoritma *Naive Bayes* dan SVM. Hasil dari penelitian menunjukkan performa dari algoritma *Naive Bayes* sebesar akurasi 81,20%, waktu 9,00 detik, *recall* 79,60%, presisi 79,40%, sedangkan performa dari algoritma SVM sebesar akurasi 85%, waktu 31,60 detik, *recall* 84%, presisi 83,60% [8]. Hasil tersebut menunjukkan bahwa algoritma SVM memiliki tingkat akurasi lebih tinggi dari *Naive Bayes* namun membutuhkan waktu yang lebih lama.

Beberapa penelitian yang sudah dilakukan sebelumnya memiliki persamaan dalam metode yang digunakan dalam analisis sentimen yaitu metode *Naive Bayes* dan SVM, namun memiliki hasil dan kesimpulan yang cukup berbeda dari segi akurasi. Pada penelitian ini mengadaptasi metode yang digunakan pada penelitian sebelumnya yaitu menggunakan metode *Naive Bayes* dan SVM untuk uji klasifikasi, namun pada penelitian ini akan membuktikan untuk mencari metode mana yang lebih efektif dan memiliki performa lebih tinggi pada data yang digunakan pada penelitian. Perbedaan yang terlihat dengan penelitian sebelumnya adalah hasil uji klasifikasi metode yang terbaik akan dilanjutkan untuk digunakan dalam penentuan prediksi sentimen pada data yang belum diketahui jenis sentimennya. Berdasarkan hal tersebut dan latar belakang masalah yang terjadi, dilakukan penelitian tentang analisis sentimen masyarakat Indonesia terhadap dampak penurunan global sebagai akibat resesi di Twitter menggunakan metode *Naive Bayes* dan *Support Vector Machine* guna memilih metode yang lebih efektif dalam studi kasus untuk menentukan prediksi sentimen.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Terdapat beberapa tahapan yang dilakukan dalam penelitian untuk mendapatkan hasil yang baik. Tahapan terdiri dari pengumpulan data, pengolahan data, pembobotan data, proses klasifikasi, evaluasi dan validasi, serta hasil dan pembahasan. Tahapan-tahapan tersebut lebih jelasnya dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

2.2 Analisis Sentimen

Analisis sentimen adalah suatu metode yang digunakan dalam mengartikan serta memahami opini untuk melihat sentimen yang terkandung di dalamnya untuk menjadi sebuah informasi yang bernilai positif, negatif, atau netral [9, 10]. Secara garis besar, tahapan proses dalam analisis sentimen adalah memahami, mengekstrak dan mengolah data

tekstual dengan tujuan membentuk sekumpulan informasi yang bermakna dan berarti [11]. Analisis sentimen ini dapat dilakukan menggunakan pendekatan text mining yang menjadi tindakan tepat untuk dapat memahami dan mengenali maksud dari setiap teks opini yang didapatkan [12].

2.3 Text Mining

Text mining adalah salah satu cabang teknik dari data mining yang populer dan dipakai banyak orang daripada *data mining* itu sendiri, karena 80% dari tiap perusahaan pastinya memiliki dokumen informasi dalam bentuk teks. *Text mining* merupakan teknik penggalian atau pencarian data teks yang sebelumnya tidak diketahui sehingga menghasilkan pengetahuan baru [13]. Terdapat beberapa model yang dapat digunakan di dalam *text mining*, salah satunya adalah klasifikasi. Terdapat beberapa metode yang dapat digunakan untuk model klasifikasi, yaitu metode *Decision Tree* berupa aturan “jika-maka”, metode *Naive Bayes* dan *Support Vector Machine* berupa formula matematis [14]. Metode *text mining* yang digunakan pada penelitian ini adalah *Naive Bayes* dan *Support Vector Machine* yang bertujuan untuk memprediksi hasil sentimen dari setiap data *tweet*.

2.4 Naive Bayes

Naive Bayes termasuk dalam salah satu metode *machine learning* yang penerapannya menggunakan perhitungan probabilitas dengan menggunakan konsep dasar *Teorema Bayes* [15]. Bentuk umum *Teorema Bayes* dapat dilihat pada persamaan (1) :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Keterangan:

$P(H|X)$: Probabilitas posterior bersyarat suatu hipotesis H terjadi jika diberikan *evidence*/bukti X terjadi

$P(X|H)$: Probabilitas sebuah *evidence* X terjadi akan mempengaruhi hipotesis H

$P(H)$: Probabilitas awal (priori) hipotesis H terjadi tanpa memandang *evidence* apapun

$P(X)$: Probabilitas awal (priori) *evidence* X terjadi tanpa memandang hipotesis/*evidence* yang lain

Metode ini memiliki kelebihan pada jumlah data latih yang digunakan cukup kecil atau sedikit sehingga dalam perhitungan dapat lebih cepat dan efisien. Namun metode ini juga memiliki kelemahan dalam hal akurasi yang cukup rendah, serta membutuhkan waktu perhitungan lebih jika terjadi kesalahan dalam memilih fitur [16, 17].

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah salah satu metode pembelajaran berdasarkan ruang hipotesis terkait fungsi-fungsi linear, kemudian dilatih menggunakan algoritma pembelajaran berdasarkan teori optimasi [18]. Analisa dilakukan pertama kali dengan proses pengubahan teks ke dalam bentuk vektor, dilanjutkan proses pembobotan menggunakan *Term Frequency Inverse Document Frequency* (TF-IDF) [19]. SVM memiliki kelebihan untuk mengidentifikasi *hyperlane* secara terpisah, sehingga hal ini dapat memaksimalkan *margin* pada kelas yang berbeda. Namun SVM juga memiliki kelemahan, dimana kesamaan fitur sangat berpengaruh signifikan terhadap tingkat akurasinya [20]. SVM memiliki cara kerja dengan menggunakan beberapa titik data terpilih yang bisa disebut sebagai *support vector* untuk membentuk model yang akan digunakan dalam proses klasifikasi [21]. Bentuk umum dari SVM dapat dilihat pada persamaan (2) :

$$f(x) = w \cdot x + b \quad (2)$$

Keterangan:

w : Parameter *hyperlane* yang dicari (garis tegak lurus antara *hyperlane* dan titik *support vector*)

x : Titik data masukan SVM

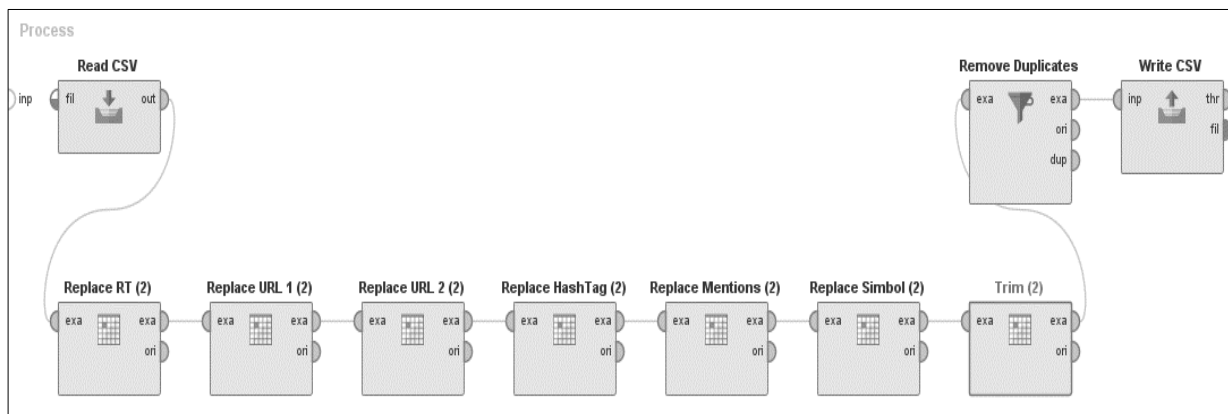
b : Parameter *hyperlane* yang dicari (nilai bias)

2.6 Tahapan Pengumpulan Data

Sumber data didapatkan dari Twitter dengan kata kunci “resesi (turun OR penurunan) until:2023-01-08 since:2022-01-01”, dapat diartikan mengambil semua data *tweet* yang mengandung kata “resesi” dan mengandung kata “turun” atau “penurunan” dari tanggal 1 Januari 2022 s/d 8 Januari 2023. Pengumpulan data dilakukan menggunakan teknik *web scrapping* dengan memanfaatkan RESTful API Twitter dan *tools* PHP yang menghasilkan total data sejumlah 824 *tweet*. *Web scrapping* merupakan teknik pengambilan data dari berbagai sumber *website* pada jaringan internet. Sumber data yang diambil dari halaman *web* berisi bahasa *markup* atau dalam bentuk lainnya, kemudian dianalisa dan dipilih sebagian data penting untuk diambil sesuai keperluan [21].

2.7 Tahapan Pre-processing Data

Dilakukan pemilahan data dengan tahapan *cleaning* dan *remove duplicates* data seperti gambar 2 menggunakan *tools* RapidMiner Studio, dimana dari semula 824 data menjadi 780 data yang akan digunakan untuk proses selanjutnya.



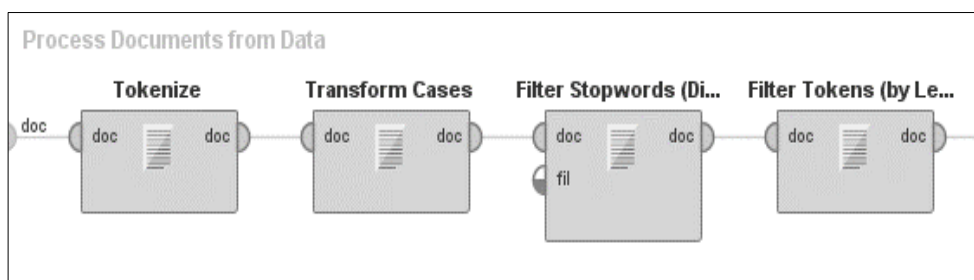
Gambar 2. Cleaning dan Remove Duplicates Data

Tahapan *cleaning* adalah tahapan untuk menghilangkan kata-kata yang tidak terpakai seperti *Retweet*, URL, *HashTag*, *Mentions*, simbol-simbol dan spasi di awal akhir kalimat. Tahapan *remove duplicates data* digunakan untuk menghapus data *tweet* yang identik dengan lainnya. Contoh hasil *cleaning* ini dapat dilihat di tabel 1.

Tabel 1. Hasil *Cleaning Data*

Data No	Sebelum <i>Cleaning</i>	Setelah <i>Cleaning</i>
22	#DPR Masuk Masa Reses hingga 14 Maret, Puan Ingatkan Anggota Turun ke Dapil - https://t.co/WPcIliBPma - Nasional https://t.co/WPcIliBPma #PuanMaharani https://t.co/UKW6rjtB2Z	Masuk Masa Reses hingga 14 Maret Puan Ingatkan Anggota Turun ke Dapil Nasional
44	@psi_id @JihanMila0 Jakarta barat daerah Tegal Alur kpn boss? Anggota PSI yg dapil Jakbar kpn reses turun ke daerah tegal alur..minyak msh susah dan harga msh mahal	Jakarta barat daerah Tegal Alur kpn boss Anggota PSI yg dapil Jakbar kpn reses turun ke daerah tegal alurminyak msh susah dan harga msh mahal
182	Harga Minyak Turun 3% Khawatir Ekonomi AS Resesi #LengkapCepatBeritanya #Keuangan #Ekonomi #Ekonomi #EkonomiIndonesia . https://t.co/T4uZ7BifqW	Harga Minyak Turun 3 Khawatir Ekonomi AS Resesi
458	WTI turun itu sebenarnya 'too good to be true' , entah itu signal RESESI atau memang sedang KOREKSI.	WTI turun itu sebenarnya too good to be true entah itu signal RESESI atau memang sedang KOREKSI

Proses selanjutnya adalah pemberian atau pelabelan setiap data *tweet* menjadi sentimen negatif atau positif. Sejumlah 200 *tweet* diberi label dan 580 *tweet* tidak diberi label, data yang diberi label disebut sebagai data latih, sedangkan data yang tidak diberi label disebut sebagai data uji. Data *tweet* yang telah dikumpulkan dan diberi label, dilanjutkan dengan melakukan pengolahan/*pre-processing data* dengan beberapa tahapan yaitu *tokenizing*, *case folding*, *filter stopwords*, *filter tokens* yang dapat dilihat pada gambar 3.



Gambar 3. *Pre-processing Data*

Tokenize adalah proses untuk mengubah kalimat menjadi per kata untuk dapat dilakukan proses selanjutnya. *Case folding* adalah proses untuk mengubah teks menjadi *upper case* atau *lower case*, pada penelitian ini menggunakan *lower case* yang bertujuan untuk membuat semua karakter menjadi huruf kecil. *Filter stopwords* adalah proses untuk membuang semua kata yang tidak diperlukan atau tidak memiliki makna berarti, daftar kata ini sudah dipersiapkan dan disimpan dalam *file* ekstensi *.csv*. *Filter tokens* adalah proses untuk menghilangkan kata yang di luar batasan minimum atau maksimum karakter, pada penelitian menggunakan batasan 4 s/d 25 karakter.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Pada penelitian ini, proses analisis selengkapnya dilakukan menggunakan *tools* RapidMiner Studio. Data yang digunakan dalam penelitian terdapat 780 data *tweet* yang sudah melalui proses *pre-processing data*. Proses berikutnya adalah pembobotan data dengan menggunakan perhitungan *Term Frequency Inverse Document Frequency* (TF-IDF). Teknik perhitungan TF-IDF sebenarnya tidak terlepas dari tahapan *pre-processing* sebelumnya, dimana hasil dari *pre-processing* akan dilanjutkan dengan perhitungan frekuensi munculnya kata dalam dokumen yang menunjukkan pentingnya kata itu di dalam sebuah dokumen [22]. Metode ini dilakukan dengan mengalikan nilai TF dan IDF berdasarkan hasil *query*-nya dengan tahapan seperti berikut:

- Menentukan TF (*Term Frequency*) terdiri dari Q, D1, D2, D3

$$Q = \text{Query}$$

$$D1 = \text{Dokumen 1}$$

$$D2 = \text{Dokumen 2}$$

$$D3 = \text{Dokumen 3}$$

- Menghitung df (data frekuensi)

$$df = Q + D1 + D2 + D3$$

- Menghitung $Idf = \log(D/df)$

- Menentukan W (pembobotan kata) terdiri dari Q, D1, D2, D3

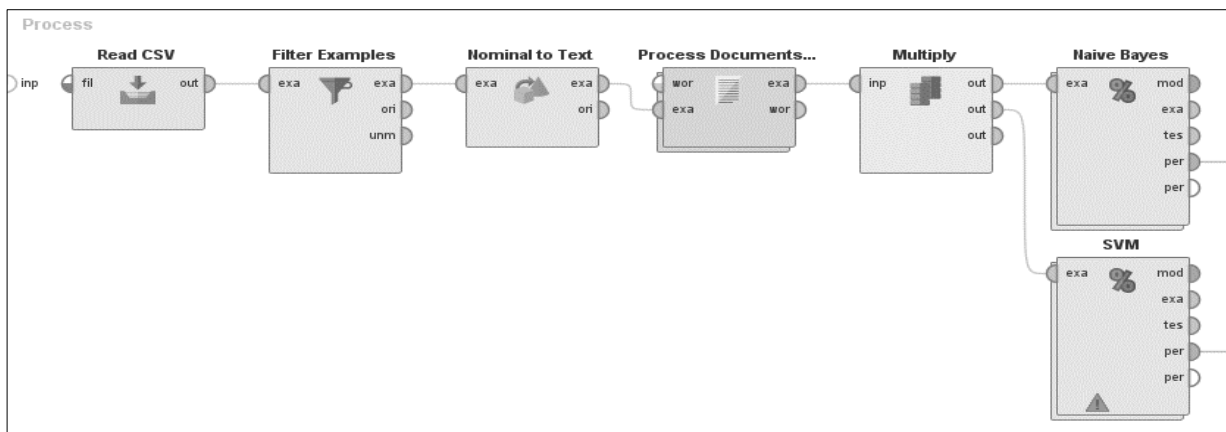
$$Q = \log(D/df) \times Q$$

Output dari perhitungan TF-IDF dapat dilihat pada tabel 2.

Tabel 2. Hasil Pembobotan Kata dengan Teknik TF-IDF

Data No	Teks	abis	acara	acuan
454	penurunan inflasi sentimen positif pasar saham penurunan inflasi indikasi resesi kenaikan suku bunga acuan cmiw	0	0	0,356
457	jadwal suku bunga inflasi turunturun inflasi resesi hyperinflasi coming	0	0	0
458	analisis menilai kenaikan suku bunga acuan instrumen investasi menarik emas faktor menahan penurunan harga emas resesi ketegangan geopolitik ulasanhargaemas	0	0	0,235
464	bursa efek indonesia rabu pagi turun dipicu kekhawatiran resesi akibat kebijakan moneter agresif bank sentral global	0	0	0
466	abis perusahaan jogja pergi penghasilan kota jogjanya turun penghasil pajaknya hilang pengangguran harga wisatawan main kesana belanja murah puter haluan	0,211	0	0

Pembobotan kata dengan teknik TF-IDF yang terlihat pada tabel 2 mendapatkan total 2.984 kata dengan *range* nilai antara 0 s/d 1. Nilai 0 berarti kata tersebut tidak muncul sama sekali di dalam teks, sedangkan nilai 1 berarti kata tersebut muncul keseluruhan dalam teks atau berarti teks hanya terdiri dari kata tersebut. Hasil pembobotan kata selanjutnya digunakan untuk pengujian klasifikasi menggunakan metode *Naive Bayes* dan SVM seperti terlihat di gambar 4.



Gambar 4. Proses Uji Klasifikasi *Naive Bayes* dan SVM

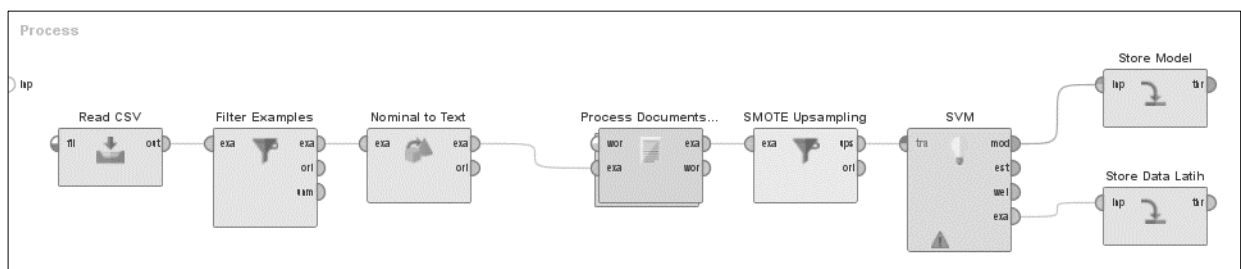
Data diuji dan diklasifikasi dengan metode *k-fold Cross Validation* menggunakan parameter nilai 10 untuk *k-fold*, artinya total dari data yang digunakan akan dibagi menjadi sepuluh bagian secara acak dan akan diulang-ulang sebanyak jumlah kelompok yang sudah ditentukan.

Tabel 3. Persentase Hasil Pengujian Metode Klasifikasi

Metode	Akurasi	Presisi	Recall
Naive Bayes	72,5%	83,8%	79,3%
SVM	79,5%	78,6%	100%

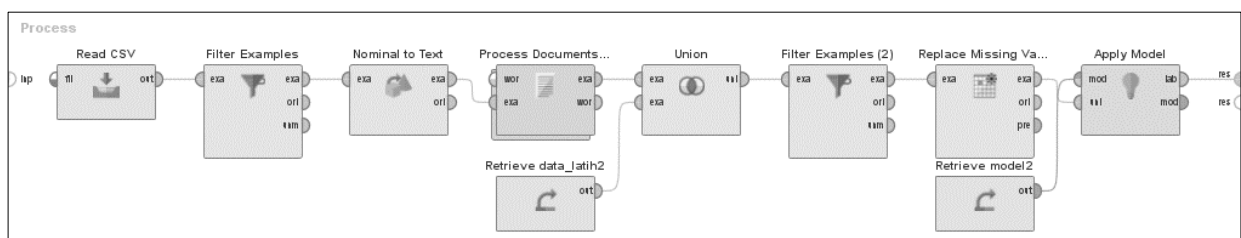
Berdasarkan hasil pengujian pada tabel 3, dapat dilihat bahwa hasil analisis dengan menggunakan metode *Naive Bayes* mendapatkan tingkat akurasi sebesar 72,5%, presisi 83,8%, *recall* 79,3%. Untuk hasil analisis menggunakan metode SVM mendapatkan tingkat akurasi sebesar 79,5%, presisi 78,6%, *recall* 100%. Terlihat bahwa metode SVM memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan metode *Naive Bayes*, sehingga disini akan ditetapkan metode SVM sebagai proses berikutnya.

Selanjutnya 200 data yang sudah diberi label atau bisa disebut sebagai data latih akan dilakukan pemodelan menggunakan metode SVM. Pada proses juga ditambahkan algoritma *SMOTE Upsampling* yang bertujuan untuk mengatasi ketidakseimbangan data dikarenakan jumlah sentimen positif dan negatif yang tidak seimbang di dalam 200 data tersebut. Proses lengkap pelatihan data dapat dilihat di gambar 5.



Gambar 5. Proses Pelatihan Data

Hasil pelatihan data dari gambar 5 akan menghasilkan 2 file yaitu data model dan data latih. Kedua file inilah yang menjadi landasan untuk digunakan dalam proses prediksi sentimen pada data uji.



Gambar 6. Proses Pengujian Data

Hasil dari proses pengujian data di gambar 6 digunakan untuk memprediksi sentimen dengan tingkat *confidence* masing-masing. Prediksi sentimen ini dapat digunakan sebagai acuan dalam melihat sentimen setiap *tweet* dengan tingkat akurasi seperti yang sudah didapatkan sebelumnya yaitu 79,5%. Beberapa contoh data yang dihasilkan dari proses pengujian data tersebut dapat dilihat pada tabel 4.

Tabel 4. Hasil Uji / Prediksi Sentimen

Data No	Teks	Sentimen Prediksi	Confidence (Positif)	Confidence (Negatif)
250	minyak penurunan terburuk terbebani kekhawatiran resesi	negatif	47,5%	52,5%
374	gelap artian ekonominya perbanyak tabungan dana darurat kalo udah resesi inflasi harga emas harga saham turun	negatif	44,9%	55,1%
432	pendapatan perusahaan turun fedex bahas resesi global	negatif	35,7%	64,3%
433	danny ismaya gelar reses aspirasi penurunan harga menggema arena reses	positif	59,4%	40,6%
559	ancaman resesi global pelaku usaha tekstil ekspor terhambat penurunan	positif	50,3%	49,7%

Didapatkan hasil dari proses akhir adalah 144 sentimen positif dan 636 sentimen negatif. Terlihat ada beberapa sentimen prediksi yang tidak sesuai seperti pada tabel 4 data no 559, seharusnya dilihat dari kalimat teksnya masuk di dalam sentimen negatif, namun untuk prediksinya adalah sentimen positif. Hal ini bisa dilihat dari tingkat *confidence* yang memiliki selisih sedikit antara *confidence* positif 50,3% dan *confidence* negatif 49,7%.

Penelitian ini juga menghasilkan *Wordcloud* untuk menganalisa kata-kata apa saja yang sering keluar atau muncul di dalam teks. *Wordcloud* sendiri memiliki arti representasi visual dari data teks dan digunakan untuk

memvisualisasikan suatu bentuk teks secara bebas. *Wordcloud* digunakan untuk menampilkan daftar kata yang memiliki frekuensi terbanyak, dimana kata dengan frekuensi tertinggi memiliki makna bahwa kata tersebut lebih menonjol atau lebih sering digunakan oleh orang dibandingkan dengan kata-kata yang lainnya [23]. *Wordcloud* dalam penelitian di-generate menggunakan *tools* RapidMiner Studio dengan menghilangkan kata-kata yang sudah digunakan sebagai kata kunci pencarian di awal pengambilan data seperti kata "resesi", "turun" dan "penurunan". Data juga diberi batasan maksimal sebanyak 30 kata teratas sesuai peringkat frekuensi kemunculan yang hasilnya dapat dilihat seperti pada gambar 7.



Gambar 7. *Wordcloud* Hasil Analisa

Dari total 30 kata teratas yang sering muncul, didapatkan bahwa peringkat 1 adalah kata "ekonomi" dengan frekuensi kemunculan sebanyak 252 kali, disusul oleh kata "harga" sebanyak 241 kali, kata "minyak" sebanyak 169 kali, dan kata "kekhawatiran" sebanyak 137 kali.

3.2 Pembahasan

Penelitian ini menunjukkan kemiripan dengan penelitian yang dilakukan sebelumnya mengenai analisis sentimen kuliah *online* pasca *COVID-19* [8]. Hasil dari uji klasifikasi menunjukkan bahwa metode SVM memiliki performa lebih baik dari *Naive Bayes* terutama mengenai tingkat akurasi. Hal ini berlawanan dengan hasil penelitian lainnya mengenai analisis sentimen masyarakat terhadap tindakan vaksinasi dalam upaya mengatasi pandemi *COVID-19* [7], yang menunjukkan metode *Naive Bayes* memiliki nilai akurasi lebih besar dari SVM.

Dapat diketahui bahwa pemilihan data yang tepat menjadi peranan yang penting dan berpengaruh terhadap hasil performanya, karena data adalah variabel utama yang digunakan dalam proses perhitungan uji performa metode tersebut. Faktor yang berpengaruh terhadap data adalah dari segi kuantitas maupun kualitas, sehingga tahapan *pre-processing* termasuk faktor pendukung untuk meningkatkan performa metode yang digunakan.

4. KESIMPULAN

Berdasarkan hasil penelitian sentimen masyarakat Indonesia terhadap dampak penurunan global sebagai akibat resesi dari media sosial Twitter berjumlah 780 data *tweet*, hasil uji klasifikasi menunjukkan metode SVM memiliki tingkat akurasi lebih tinggi dari metode *Naive Bayes* yaitu sebesar 79,5%. Hasil prediksi menggunakan metode SVM menunjukkan sejumlah 144 sentimen positif dan 636 negatif, maka dapat disimpulkan masyarakat Indonesia terhadap dampak penurunan global sebagai akibat resesi dominan beropini secara negatif. Hal ini dapat dilihat dari *Wordcloud* yang terbentuk untuk kata-kata yang sering dilontarkan oleh masyarakat seperti ekonomi, harga, minyak, kekhawatiran, dll. Dapat diasumsikan bahwa masyarakat khawatir dengan adanya kondisi ekonomi mereka, salah satunya adalah tidak stabilnya harga minyak yang dapat memicu terjadinya resesi. Berikut contoh salah satu *tweet* dari masyarakat "Harga minyak dunia kan lagi turun harga bbm jangan dinaikan kami sedang kesulitan dimasa resesi". Opini lainnya juga banyak yang menyinggung negara agar dapat turun langsung ke masyarakatnya seperti yang ada di salah satu *tweet* berikut "Lucu nya pejabat hari ini Ketika datangnya masa pemilu mereka datang dan turun ke tengah tengah rakyat mengumpulkan masa sebanyak banyaknya untuk dapat mendengar keluhan dan dijadikan aspirasi Begitu terpilih dan pulang di saat reses hanya kelompok/komunitas yang ditemuiny".

REFERENCES

- [1] A. S. Zahra, N. Murialti and M. F. Hadi, "Analisis Pengaruh Resesi Ekonomi di Provinsi Riau Tahun 2006-2020," *Economics, Accounting and Business Journal*, vol. 2, no. 1, pp. 141-150, 2022.
- [2] B. H. Miraza, "Seputar Resesi dan Depresi," *Jurnal Ekonomi KIAT*, vol. 30, no. 2, pp. 11-13, 2019.



- [3] H. J. Christanto, "Game Theory Analysis on Marketing Strategy Determination of KAI Access and Traveloka based on Usability of HCI (Human-Computer Interaction)," *Journal of Information Systems and Informatics*, vol. 4, no. 3, pp. 665-672, 2022.
- [4] S. Azeharie and O. Kusuma, "Analisis Penggunaan Twitter Sebagai Media Komunikasi Selebritis di Jakarta," *Jurnal Komunikasi Universitas Tarumanagara*, vol. 6, no. 2, pp. 83-98, 2014.
- [5] M. Ahlgren, "55+ STATISTIK TWITTER, FAKTA & TREN UNTUK 2023," 2023. Website: <https://www.websiterating.com/id/research/twitter-statistics/>. [Diakses 12 Januari 2023].
- [6] T. T. Widowati and M. Sadikin, "Analisa Sentimen Twitter Terhadap Tokoh Publik Dengan Algoritma Naive Bayes dan Support Vector Machine," *Jurnal Teknik Industri, Mesin, Elektro dan Ilmu Komputer*, vol. 11, no. 2, pp. 626-636, 2020.
- [7] B. Laurensz and E. Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 2, pp. 118-123, 2021.
- [8] H. Setiawan, E. Utami and Sudarmawan, "Analisis Sentimen Twitter Kuliah Online Pasca Covid-19 Menggunakan Algoritma Support Vector Machine dan Naive Bayes," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 1, pp. 43-51, 2021.
- [9] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online JD.ID Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi," *Jurnal Teknik Industri, Mesin, Elektro dan Ilmu Komputer*, vol. 10, no. 2, pp. 681-686, 2019.
- [10] V. K. S. Que, A. Iriani and 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, pp. 162-170, 2020.
- [11] D. Ikasari, Y. Fajarwati and Widiastuti, "Analisis Sentimen dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum MRT Jakarta Menggunakan Naive Bayes Classifier," *Jurnal Ilmiah Informatika Komputer*, vol. 25, no. 1, pp. 64-75, 2020.
- [12] H. J. Christanto and Y. A. Singgalen, "Sentiment Analysis of Customer Feedback Reviews Towards Hotel's Products and Services in Labuan Bajo," *Journal of Information Systems and Informatics*, vol. 4, no. 4, pp. 805-822, 2022.
- [13] A. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)," *Jurnal Penelitian Ilmu dan Teknologi Komputer*, vol. 13, no. 1, pp. 66-78, 2021.
- [14] E. K. Putri and T. Setiadi, "Penerapan Text Mining Pada Sistem Klasifikasi Email Spam Menggunakan Naive Bayes," *Jurnal Sarjana Teknik Informatika*, vol. 2, no. 3, pp. 73-83, 2014.
- [15] H. Susana, N. Suarna, Fathurrohman and Kaslani, "Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet," *Jurnal Sistem Informasi dan Teknologi Informasi*, vol. 4, no. 1, pp. 1-8, 2022.
- [16] A. F. Watratan, A. Puspita and 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, pp. 7-14, 2020.
- [17] S. Hikmawan, A. Pardamean and S. N. Khasanah, "Sentimen Analisis Publik Terhadap Joko Widodo Terhadap Wabah Covid-19 Menggunakan Metode Machine Learning," *Jurnal Kajian Ilmiah*, vol. 20, no. 2, pp. 167-176, 2020.
- [18] I. M. Parapat, M. T. Furqon and Sutrisno, "Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, pp. 3163-3169, 2018.
- [19] H. Tuhuteru and 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, pp. 394-401, 2018.
- [20] H. Irsyad, A. Farisi and M. R. Pribadi, "Klasifikasi Opini Masyarakat Terhadap Jasa ISP MyRepublic dengan Naive Bayes," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 8, no. 1, pp. 30-34, 2019.
- [21] Herianto, "Penerapan Text-Mining untuk Mengidentifikasi Pengguna Twitter Terhadap Fenomena Peran DPR RI," *Jurnal Sains & Teknologi*, vol. 8, no. 2, pp. 36-44, 2018.
- [22] M. A. Rofiqi, A. C. Fauzan, A. P. Agustin, A. Agun and H. Di, "Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) untuk Mencari Relevansi Dokumen Berdasarkan Query," *Journal of Computer Science and Applied Informatics*, vol. 1, no. 2, pp. 58-64, 2019.
- [23] Andreyestha, "Analisis Sentimen Masyarakat Terhadap Fenomena Teroris Melalui Twitter di Indonesia," *Jurnal Kajian Ilmiah*, vol. 19, no. 3, pp. 239-247, 2019.