Analisis Sentimen Masyarakat Twitter Terhadap Kebijakan Pemerintah Dalam Menaikkan Harga Bahan Bakar Minyak Dengan Menggunakan Metode Support Vector Machine

Haifa' Syadza 'Adilah[1], Ronggo Alit[2]

[1,2] Teknik Informatika, Fakultas Teknik, Universitas Negeri Surabaya
[1]haifa.19106@mhs.unesa.ac.id
[2]ronggoalit@unesa.ac.id

Abstrak— **Penelitian ini bertujuan utama untuk mengambil data tweet terkait kebijakan kenaikan harga bahan bakar minyak pemerintah, menganalisis sentimen positif dan negatif menggunakan metode Support Vector Machine, dan mengukur akurasi hasil pemrosesan sentimen tersebut. Langkah-langkah penelitian meliputi studi literatur, analisis kebutuhan, pengumpulan data, perancangan, evaluasi dan pengujian model, serta analisis model dengan metode Support Vector Machine. Proses pengambilan data menggunakan ekstensi Google Colab dan Twitter API dengan kata kunci "bbm naik" berhasil mengumpulkan 493 tweet dari Februari 2023 hingga Maret 2023. Dari data tersebut, 244 tweet memiliki sentimen positif, sedangkan 249 memiliki sentimen negatif, dengan distribusi yang hampir seimbang. Selanjutnya, data diproses dengan pembobotan TF-IDF dan dibagi menjadi data latih (80%) dan data uji (20%). Implementasi metode Support Vector Machine menggunakan kernel linear dengan perbandingan 8:2 menghasilkan akurasi tertinggi sebesar 82.88%, Precision 83.83%, F1-Score 83.83%, dan Recall 83.83% juga mencapai Dalam hasil prediksi, terdapat 43 True Negative (TN), 9 False Negative (FN), 8 False Positive (FP), dan 39 True Positive (TP). Ini menggambarkan performa model dalam mengklasifikasikan sentimen positif dan negatif terhadap kebijakan kenaikan harga bahan bakar minyak pemerintah. Penelitian ini memberikan wawasan tentang pandangan publik terhadap kebijakan tersebut, dengan dukungan dari analisis sentimen yang dapat digunakan untuk pengambilan keputusan lebih lanjut dalam konteks kebijakan publik.**

Kata Kunci*— Twitter API, harga BBM, tweet sentimen positif-negatif, *Support Vector Machine.

I. PENDAHULUAN

Pada awal bulan Februari 2023, Pemerintah Indonesia mengumumkan langkah untuk meningkatkan harga bahan bakar minyak sebagai respons terhadap dampak dari konflik di Rusia dan Ukraina. Peristiwa ini telah menyebabkan kenaikan tajam dalam harga minyak dunia, yang berimbas signifikan terutama pada Indonesia. Efeknya terasa dalam melemahnya kondisi ekonomi, termasuk dampak yang dirasakan oleh masyarakat [1]. Kenaikan harga minyak global juga berperan dalam menaikkan harga minyak mentah di Indonesia. Dasar asumsi harga minyak mentah di Indonesia telah diatur dalam kesepakatan antara Menteri Energi dan Sumber Daya Mineral (ESDM) dan Komisi VII DPR RI, yang didasarkan pada Rancangan Anggaran dan Belanja Negara (RAPBN) tahun 2023, yakni sebesar USD95 per barel. Pada tahun 2022, angka tersebut sebelumnya lebih rendah, mencapai USD63 per barel menurut Anggaran Pendapatan Belanja Negara (APBN) tahun 2022. Penetapan harga baru ini mengalami kenaikan sebesar USD5 per barel dari usulan sebelumnya, sebagaimana diumumkan melalui siaran pers Kementerian Energi dan Sumber Daya Mineral RI pada tanggal 9 September 2022, nomor 343.Pers/04/SJI/2022. Sejak tahun 2020 hingga 2022, terjadi kenaikan yang signifikan dalam harga bahan bakar minyak, seperti pada gambar berikut.

Gambar 1. Kenaikan harga bahan bakar minyak 2020-2022

Sumber : Pertamina

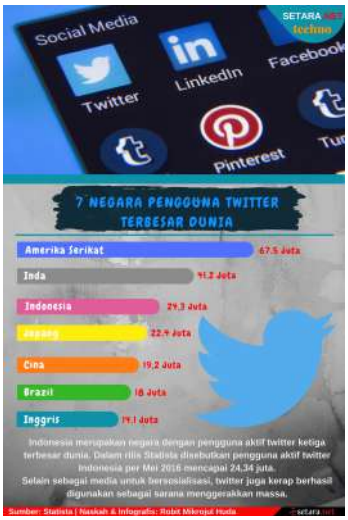
Kenaikan harga bahan bakar minyak tahun 2020 hingga tahun 2022 mencapai kenaikan yang signifikan. Kenaikan tersebut telah dijelaskan pada gambar berikut. Terdapat tiga alasan bahwa harga bahan bakar minyak mengalami kenaikan meliputi : (1) Terjadi disparitas antara konsumsi bahan bakar minyak dan produksi minyak mentah di Indonesia, (2) Indonesia masih perlu mengimpor bahan bakar minyak, dan (3) Pengeluaran subsidi semakin meningkat. Perubahan harga bahan bakar minyak memiliki pengaruh yang lebih besar terhadap tingkat inflasi pada sektor makanan, transportasi, komunikasi, dan sektor lainnya [2]. Kenaikan harga bahan bakar minyak berpengaruh besar terhadap masyarakat, menyebabkan kenaikan harga bahan pangan dan pakaian [3]. dampak kenaikan bahan bakar minyak berpengaruh secara langsung pada masyarakat miskin [4]. Di tahun 2022 telah mengalami dampak kenaikan bahan bakar minyak, dampak kenaikan tersebut diterangkan pada gambar berikut.

Gambar 2. Dampak kenaikan BBM 2022

Sumber : Pertamina

Pada tahun 2022 mengalami dampak kenaikan bahan bakar minyak, meliputi : (1) lonjakan inflasi sebesar 1.26%, pada bulan Agustus mengalami peningkatan inflasi sebesar 4.69%, pada Bulan September mengalami peningkatan inflasi 5.95%, (2) Suku bunga acuan naik menjadi 4.75%, pada bulan Agustus suku bunga acuan BI 3.75%, bulan September 4.25%, bulan Oktober 4.75%. dan (3) Tekanan pada nilai tukar rupiah terhadap USD, tanggal 31 Agustus sebesar Rp14.853, tanggal 31 September sebesar Rp15.232, tanggal Oktober sebesar Rp15.588. Kenaikan harga bahan bakar minyak menimbulkan banyak perbedaan pendapat di media sosial, terutama di media sosial Twitter.

Platform sosial media Twitter merupakan salah satu sumber data real-time paling besar dan paling berharga untuk mengakui pemahaman masyarakat Indonesia. Pada periode pertama tahun 2017, total pengguna Twitter di seluruh dunia mencapai 328 juta. Di Indonesia saja, jumlah penggunanya mencapai 24,2 juta pada bulan Mei 2016. [9]. Pesan-pesan di Twitter disebut dengan *tweet*, sehingga tweets dapat menjadi sarana bagi penduduk Indonesia, untuk mengungkapkan opininya atau berbagi informasi kepada individu lain. Banyak orang dewasa dan remaja dari berbagai latar belakang profesional menggunakan platform sosial medial ini. Twitter juga menjadi tempat masyarakat menyuarakan keluh kesah, argumentasi, kritik dan sarannya dalam berbagai aspek termasuk latar belakang sosial, politik, dan teknologi. Pertanyaan yang kerap ditanyakan ialah tentang kenaikan harga bahan bakar minyak di Indonesia. Oleh karena itu, setelah pengumuman mengenai kenaikan harga bahan bakar minyak, berbagai tweet dengan hashtag (#bbmnaik) membanjiri feed Twitter setiap hari. Adanya berbagai pandangan masyarakat mengenai kenaikan harga BBM menjadi topik penelitian yang menarik. Pendekatan psikoanalisis sangat sesuai untuk mengartikan tanggapan masyarakat terhadap kenaikan harga bahan bakar minyak. Di bawah ini adalah daftar tujuh pengguna Twitter dengan jumlah pengikut terbesar di dunia, seperti yang terlihat dalam Gambar 3.



Gambar 3 Data Pengguna Twitter

Sumber : Setara.net (2018)

Salah satu penelitian yang menggunakan media sosial Twitter untuk analisis adalah Kurniasih, dalam penelitiannya metode yang digunakan adalah metode Naive Bayes menggunakan data sebesar 795 *tweet* dengan kata kunci Bahan Bakar Minyak dan kata kunci BSU. Yang masing-masing terdiri 263 data latih dan 532 data uji. Hasil akurasi yang dihasilkan dari metode Naive Bayes adalah 82,64% dan 92,89%, bisa dikatakan bahwa hasil sikap masyarakat terhadap BSU dinilai positif, dan harga kenaikan Bahan Bakar Minyak dinilai negatif [1]. Pada awal bulan Desember, Risa Wati melakukan analisis mengenai PPKM dengan metode *Support Vector Machine*. Penelitian tersebut menghasilkan akurasi sebesar 86% [5]. Sementara itu Siti Nurhasanah melakukan penelitian terkait Penerapan Social Distancing dengan menggunakan metode *Support Vector Machine*. Penelitian ini menghasilkan akurasi sebesar 67,00% [6]

Pada beberapa penelitian sebelumnya yang menyatakan bahwa metode *Support Vector Machine* lebih baik dan terbukti mendapatkan akurasi yang lebih tinggi dibandingkan metode KNN [7] dan Decision Tree [8]. Oleh karena itu, peneliti merasa tertarik untuk membuat penelitian yang berjudul "Analisis Sentimen Masyarakat Twitter Terhadap Kebijakan Pemerintah Dalam Menaikkan Harga Bahan Bakar Minyak Dengan Menggunakan Metode *Support Vector Machine*". Adapun tujuan penelitian melupiti: (1) mengimplementasi Twitter API untuk pengambilan data tweet terhadap kebijakan pemerintah dalam menaikkan harga bahan bakar minyak, (2) menganalisa data tweet sentimen positif dan sentimen negatif terhadap kebijakan pemerintah dalam menaikkan harga bahan bakar minyak dengan menggunakan metode *Support Vector Machine.*

II. METODOLOGI PENELITIAN

Metode yang diterapkan dalam analisis ini ialah melakukan uji coba ataupun eksperimen terhadap studi kasus implementasi kebijakan pemerintah dalam menaikkan harga bahan bakar minyak. Sementara itu, model yang diaplikasikan dalam analisis adalah *Support Vector Machine.*

Gambar 4. Diagram alur penelitian

Berdasarkan pada Gambar 4 dapat diketahui prosedur-prosedur yang akan diterapkan dalam rangkaian penelitian, yaitu sebagai berikut : Tahap Studi Literatur merupakan langkah awal dalam pelaksanaan penelitian di mana peneliti akan mengumpulkan berbagai data melalui pencarian jurnal, literatur, makalah, buku, dengan sumber-sumber dari internet yang cocok dengan menggunakan materi penulisan, khususnya dalam analisis sentimen dengan menggunakan metode Support Vector Machine. Selanjutnya, dalam Tahap Analisis Kebutuhan, akan dilakukan analisis mendalam untuk menentukan komponen-komponen yang diperlukan dalam proses analisis sentimen menggunakan metode *Support Vector Machine*. Terakhir, dilakukan proses Pengumpulan Data, dilakukan melalui teknik *web scraping* pada platform Twitter dengan menggunakan API Twitter. Data yang diambil berupa tweet yang berkaitan dengan kebijakan pemerintah dalam menaikkan harga bahan bakar minyak pada rentang waktu bulan Februari 2023 sampai dengan bulan Maret 2023. Teknik web scraping telah banyak digunakan dalam penelitian analisis sentimen pada media sosial. Selain itu, Twitter API telah terbukti dapat memberikan data yang akurat dan terpercaya dalam penelitian analisis sentimen. Sebelum pengambilan data dilakukan, dilakukan pula pemilihan kata kunci yang relevan dengan topik penelitian untuk memastikan bahwa data yang diambil sesuai dengan tujuan penelitian. Kata kunci (query) “bbm naik”. Berdasarkan kata kunci tersebut data akan di kumpulkan dan di jadikan bahan dalam perancangan sistem. Perancangan, dalam perancangan terdiri dari desain alur pengambilan data, desain alur sistem dan metode, serta rancangan pemrograman. Pengujian model dan evaluasi, Dalam tahap ini, digunakan performa hasil klasifikasi algoritma *Support Vector Machine* pada analisis sentimen. Evaluasi ini melibatkan berbagai perhitungan seperti akurasi, presisi, *recall*, dan F1-score. Untuk menghasilkan nilai-nilai ini, perlu dibuat matriks konfusi yang nantinya akan digunakan sebagai evaluasi terhadap hasil klasifikasi dataset menggunakan algoritma *Support Vector Machine.*

A. Studi Literatur

Tahapan awal dalam menjalankan penelitian ialah dengan melakukan studi literatur, dimana peneliti menghimpun data melalui pencarian berbagai jurnal, literatur, makalah, buku, serta sumber-sumber daring sebagai rujukan yang berhubungan dengan bahan penulisan, khususnya dalam konteks analisis sentimen dengan menggunakan metode *Support Vector Machine.*

B. Analisis Kebutuhan

Tahapan ini yaitu menganalisa kebutuhan untuk menentkan komponen-komponen yang dibutuhkan dalam sentimen mengunakan metode *Support Vector Machine*. Pada penelitian analisis sentimen ini dibutuhkan sebuah sistem yang dapat mendukung agar pelaksanaan penelitian analisis sentimen ini berjalan dengan lancar. Adapun kebutuhan yang digunakan dalam sistem dalah sebagai berikut :

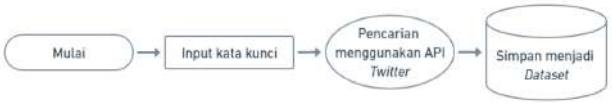
1. Spesifikasi Perangkat Keras *(Hardware)*
   Perangkat Keras *(Hardware)* yang ang dipergunakan dalam studi ini adalah sebuah laptop dengan detail spesifikasi sebagai berikut
   a. *Processor* : Intel Core i7-7500U CPU 2.70GHz 2.90GHz
   a. RAM : 8.00 GB
   b. *System Type* : 64-bit *operating system*
2. Spesifikasi Perangkat Lunak *(Software)*
   Perangkat Lunak *(Software)* yang diterapkan dalam penelitian ini mencakup hal-hal berikut :
   a. Windows 10 64bit
      Adalah Sistem operasi yang diterapkan pada Laptop Asus A456U adalah Windows 10 versi 64-bit.
   a. Google Chrome dan Mozilla Firefox
      Mesin yang digunakan sebagai peramban web sumber terbuka
   b. Google Colab
      Perangkat yang digunakan untuk pengambilan data pada Twitter *Application Programming Interface* (API) berupa komentar atau tanggapan masyarakat dengan menggunakan kata kunci “bbm naik”
   c. Microsoft Exel
      Digunakan untuk melakukan penyaringan dan seleksi data tweet yang telah diambil melalui Google Colab.
   d. Python
      Merupakan alat atau aplikasi yang digunakan dalam proses pembuatan dan perancangan program.

C. Pengumpulan Data

Pengumpulan data dalam analisis sentimen ini menggunakan teknik scraping data pada platform Twitter dengan menggunakan API Twitter. Data yang diambil berupa *tweet* yang berkaitan dengan kebijakan pemerintah dalam menaikkan harga bahan bakar minyak pada rentang waktu bulan Februari 2023 sampai dengan bulan Maret 2023. Teknik *scraping* telah banyak digunakan dalam penelitian analisis sentimen pada media sosial. Selain itu, Twitter API telah terbukti dapat memberikan data yang akurat dan terpercaya dalam penelitian analisis sentimen. Sebelum pengambilan data dilakukan, dilakukan pula pemilihan kata kunci yang relevan dengan topik penelitian untuk memastikan bahwa data yang diambil sesuai dengan tujuan penelitian. Kata kunci *(query)* "bbm naik". Berdasarkan kata kunci tersebut data akan di kumpulkan dan di jadikan bahan dalam perancangan sistem.

Data yang dianalisis dalam penelitian ini meliputi server Twitter. Informasi tweet ini diperoleh dengan menggunakan fungsi yang disediakan oleh Twitter API yang berguna untuk pengambilan data tweet langsung dari server Twitter. Data yang dipulihkan kemudian dikumpulkan dalam format file CSV.

Dalam proses pengumpulan *tweet*, penelitian ini menggunakan kata kunci "bbm naik" sebagai patokan. Data *tweet* berasal dari pengguna Twitter yang memposting tweet terkait kata kunci tersebut, dan ini dilakukan melalui penggunaan Twitter API. Setelah berhasil diambil, data *tweet* dari server Twitter akan disimpan dalam file csv. Berikut gambaran proses pemulihan data melalui scraping data.



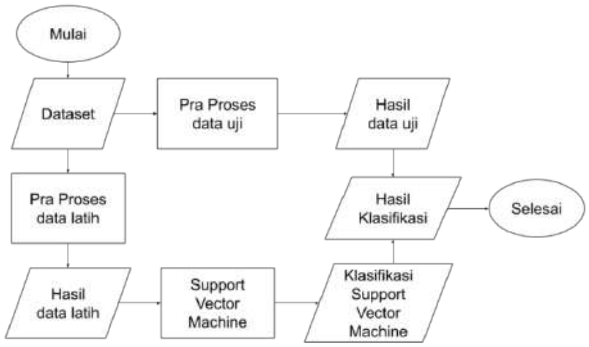
Gambar 5. Alur scraping data

D. Perancangan

Pada tahapan ini, terdapat beberapa langkah yang dilaksanakan, seperti merancang alur pengambilan data, merancang alur sistem dan metode, serta merancang alur pemrograman.

Adapun tujuan penelitian ini adalah untuk mengulas sentimen yang muncul dari masyarakat di platform Twitter terkait keputusan pemerintah dalam menaikkan harga bahan bakar minyak. Metodologi yang diterapkan dalam penelitian ini mengandalkan *Support Vector Machine*, sebuah algoritma pembelajaran mesin yang berguna dalam klasifikasi serta regresi. Dalam kerangka penelitian ini, *Support Vector Machine* diterapkan untuk mengklasifikasikan sentimen yang muncul dari masyarakat Twitter terkait langkah-langkah pemerintah. Data dikumpulkan melalui penggunaan API Twitter dan kemudian dianalisis menggunakan bahasa pemrograman Python. Metode *Support Vector Machine* dimanfaatkan untuk mengklasifikasikan sentimen dari data yang terhimpun. Selain itu, untuk mengevaluasi keakuratan sistem yang dikembangkan, dilakukan pengujian menggunakan algoritma *Support Vector Machine.* Dengan pemanfaatan sistem ini, diharapkan akan tercipta pemahaman yang lebih komprehensif mengenai pandangan masyarakat terhadap langkah-langkah pemerintah, yang pada gilirannya dapat memberikan panduan bagi pemerintah dalam mengambil keputusan yang tepat terkait regulasi harga bahan bakar minyak.

1. Alur Proses Sistem
   Langkah analisis sentimen menggunakan metode *Support Vector Machine* diawali dengan entri data, meliputi data latih dan data uji. Data ini kemudian mengalami pra-pemrosesan hingga lolos tahap klasifikasi.

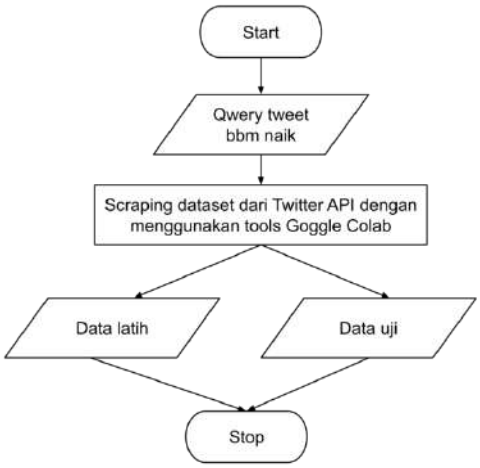


Gambar 6. Alur proses sistem

Gambar diatas menujukkan di atas menggambarkan urutan tindakan yang dijalankan dalam penelitian ini. Berikut ini adalah penjabaran dari tiap-tiap proses yang dilakukan :

a. Input Data
b. Pre Proses Program
c. Pembobotan Kata
d. Penerapan Klasifikasi Metode Support Vector Machine
e. Proses Penginputan Pengujian Data Uji

2. Proses Dataset
   Dataset dalam teks bahasa Indonesia diambil dari Twitter. Data diambil untuk penelitian ini menggunakan kata kunci "bbm naik". Alur proses pengambilan dataset seperti Gambar 7



Gambar 7. Alur dataset

Dataset dari hasil scraping ini akan dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih diklasifikasikan menggunakan *Support Vector Machine* dengan label sentimen negatif dan negatif. Data uji nantinya akan digunakan pada saat evaluasi untuk menentukan keakuratan data pada sistem. Pengambilan data latih dan data uji dilakukan dengan waktu pengambilan yang berbeda

3. *Pre-Processing*
Tahap *Pre-processing* ini bertujuan untuk menghapus gangguan dan konten yang tidak sesuai, serta merapikan bentuk kata-kata guna mengurangi jumlahnya, agar hasil pada langkah berikutnya menjadi lebih akurat. Hasil dari tahap ini akan menyediakan data yang siap digunakan dalam langkah-langkah berikutnya.. Pemrosesan data meliputi beberapa tahap yaitu :
   a. Pada proses *Case Folding* adalah untuk mengubah hutuf yang bercampur menjadi huruf kecil
   b. Pada proses *Cleaning* data adalah untuk membersihkan data tweet dari komponen yang tidak di inginkan
   c. Pada proses *Tokenization* adalah untuk merubah kalimat menjadi token-token atau potongan kata tunggal
   d. Pada proses *Stopword Removal* yaitu untuk menghapus kata-kata dan jumlah kemunculan yang banyak dalam teks
   e. Pada prses *Stemming* adalah proses yang betujuan untuk menghilangkan imbuhan dari kata yang tidak diperlukan kemudian mengembalikan ke kata dasar

4. Pembobotan TF-IDF
Metode pembobotan yang diintegrasikan dari *term frequency* dan *inverse dokument frequency* dengan menggunakan rumus :

$$W = tf \times idf$$

Metode ini digunakan untuk mencari representasi nilai dari suatu kumpulan data.

5. Klasifikasi *Support Vector Machine*
Dalam proses penerapan *Support Vector Machine* ini, terdapat dua tahapan utama: pelatihan *(training)* dan pengujian *(testing)*. Proses pelatihan melibatkan penggunaan dataset untuk mengajari mesin, di mana mesin diberikan pengetahuan dari sejumlah dataset yang sudah diberi label sebelumnya. Selanjutnya, dalam proses pengujian, mesin diuji menggunakan data baru untuk mengevaluasi akurasi dan hasil klasifikasinya. Proses pemodelan *Support Vector Machine* ini memerlukan langkah-langkah berikut:
   a. Pemilihan fungsi kernel yang akan digunakan pada SVM Multi Kelas. Dalam konteks ini, digunakan fungsi kernel Radial Basis Function (RBF) sebagai pendekatan permodelan hyperplane. Parameter C dan Gamma juga dimasukkan dalam permodelan.
   b. Menetukan jumlah permodelan k-fold cross validation untuk menghasilkan model terbaik dengan menggunakan parameter C dan Gamma.
   c. Menentukan nilai optimal untuk parameter C dan Gamma yang akan digunakan dalam permodelan hyperplane SVM, sehingga dapat menghasilkan tingkat akurasi yang tinggi saat diujikan.

Support Vector Machine Linear

Dalam *linear Support Vector Machine* pemisah merupakan fungsi *linear.* Setiap data latih dinyatakan oleh (x_i, y_i) dengan i = 1, 2, …,N, dan;

$x_i = \{x_{i1}, x_{i2}, \ldots\ldots, x_{iq}\}^T$ adalah atribut set untuk data latih kelas ke-I untuk yi ∈ {-1, +1} menyatakan label kelas. *Hyperlane* klasifikasi *Suport Vector Machine,* dinotasikan :

$$w.x_i + b = 0 \quad (1)$$

w dan b adalah parameter model $w.x_i$ adalah *inner product* antara w dan x_i. Data x_i yang masuk kedalam kelas -1 adalah data yang memenuhi pertidaksamaan berikut :

$$w.x_i + b \leq -1 \quad (2)$$

Sementara data x_i yang masuk kedalam kelas +1 adalah data yang memenuhi pertidaksamaan berikut :

$$w.x_i + b \geq +1 \quad (3)$$

Jika data dala kelas -1 (misal x_a) bertempat di *hyperlane,* maka untuk data kelas -1 dinotasikan :

$$w.x_a + b = 0 \quad (4)$$

Sementara kelas +1 (misal x_b) akan memenuhi persamaan :

$$w.x_b + b = 0 \quad (5)$$

Dengan mereduksi persamaan (5) menjadi (4), diperoleh:

x_a-x_b adalah vektor paralel pada posisi hyperplane dan diarahkan jauh ke x_b

Label -1 untuk kelas pertama dan +1 untuk kelas kedua, maka untuk memprediksi seluruh data uji gunakan rumus:

$$y = \begin{cases} +1, jika\ w.z + b > 0 \\ -1, jika\ w.z + b < 0 \end{cases} \quad (6)$$

Hyperlane untuk kelas -1 adalah data *support vector* yang menggunakan rumus :

$$w.x_a + b = -1 \quad (7)$$

Sementara kelas +1 adalah data *support vector* yang menggunakan rumus :

$$w.x_b + b = +1 \quad (8)$$

Dengan demikian margin dapat dihitung dengan mengurangi persamaan (8) dengan (7) didapatkan :

$$w.(x_b - x_a) = 2 \quad (9)$$

Margin hyperlane diberikan oleh jarak antara dua *hyperlane* dari dua kelas tersebut. Notasi diatas diringkas menjadi :

$$\|w\| x\ d = 2\ atau\ d = \frac{2}{\|d\|} \quad (10)$$

E. Evaluasi dan Pengujian Model

Pada Proses ini akan dilakukan pengukuran kinerja dari hasil klasifikasi yang diberikan oleh algoritma *Support Vector Machine* dalam analisis sentimen. Proses evaluasi ini melibatkan beberapa penghitungan, termasuk akurasi, presisi, *recall, dan f1-score*. Untuk memperoleh nilai-nilai dari penghitungan tersebut, penting untuk menyusun matriks kebingungan *(confusion matrix)* yang berperan sebagai alat evaluasi hasil klasifikasi dataset menggunakan algoritma Support Vector Machine. Pada tahap ini, perhitungan performa dari hasil klasifikasi oleh algoritma *Support Vector Machine* dalam analisis sentimen akan dijalankan. Evaluasi ini melibatkan beberapa perhitungan seperti akurasi, presisi, recall, dan skor F1. Untuk mendapatkan nilai-nilai dari perhitungan tersebut, proses pembuatan matriks kebingungan *(confusion matrix)* perlu dilakukan sebagai bagian dari evaluasi hasil klasifikasi dataset menggunakan algoritma *Support Vector Machine*. Tabel evaluasi ini diperlihatkan dalam tabel berikut.

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	TP	FP
Negatif	FN	TN

Tabel 1. *Confusiin matrix*

F. Penyusunan Laporan

Tahap ini berfokus pada penyusunan laporan yang berfungsi sebagai dokumen referensi yang memfasilitasi pemahaman dan pengembangan oleh individu lain. Penelitian mengenai analisis sentimen masyarakat di platform Twitter terkait kebijakan pemerintah tentang kenaikan harga bahan bakar minyak, yang menerapkan metode *Support Vector Machine,* memerlukan langkah-langkah sistematis dalam menyusun laporan yang bersifat akademis dan terstruktur. Langkah-langkah ini melibatkan aspek-aspek seperti penentuan judul penelitian, tujuan penelitian, oerumusan maslaah, metodologi penelitian, kerangka teori, pengumpulan data, analisis data, diskusi hasil penelitian, simpulan, dan rekomendasi. Tahapan pertama adalah memilih judul penelitian yang mencerminkan substansi permasalahan penelitian. Setelah itu, fokus beralih pada perumusan masalah dan tujuan penelitian yang jelas dan terstruktur. Kerangka teori digunakan untuk menghubungkan teori dengan hasil penelitian yang diharapkan. Selanjutnya, metodologi penelitian dirancang untuk menjawab pertanyaan penelitian melalui penerapan metode *Support Vector Machine*, dengan data yang dikumpulkan melalui teknik *scraping* dan API Twitter. Data yang diperoleh diolah dan dianalisis menggunakan alat pengolahan data. Hasil analisis disajikan secara mendalam melalui pembahasan dan ditutup dengan simpulan serta rekomendasi bagi pengembangan penelitian selanjutnya. Keseluruhan proses ini harus dijalani secara terstruktur, sistematis, dan dengan pendekatan ilmiah untuk menghasilkan laporan penelitian yang bermutu dan berdaya guna untuk kemajuan dalam bidang ilmu pengetahuan dan teknologi.

III. HASIL DAN PEMBAHASAN

A. *Hasil Pelabelan Data*

Pengambilan data dilakukan dengan menggunakan bantuan ekstensi Google Colab. Proses pengumpulan data teks dari unggahan di media sosial Twitter tentang kebijakan pemerintah terkait kenaikan harga bahan bakar minyak direalisasikan melalui teknik scraping dengan menggunakan API Twitter. Pengambilan data dilakukan berdasarkan kata kunci "bbm naik" dan berhasil mengumpulkan sebanyak 493 data *tweet.* Data ini diambil dari rentang waktu Februari 2023 hingga Maret 2023.

Langkah pertama dalam proses scraping data dimulai dengan menyisipkan token API Twitter yang telah dibuat sebelumnya. Selanjutnya, dibuatlah kata kunci (query) untuk mengidentifikasi data yang akan diambil, serta ditentukan tanggal, bulan, dan tahun untuk pengambilan data. Hasilnya, data dapat diunduh dan diakses.

Data yang diambil terdiri dari *tweet* berbahasa Indonesia yang merupakan asli dari penulisnya. Data ini dapat dilihat melalui tampilan sistem dan juga bisa diunduh dalam format csv. Gambar tentang hasil *scraping* data ini dapat ditemukan dalam Gambar 8.

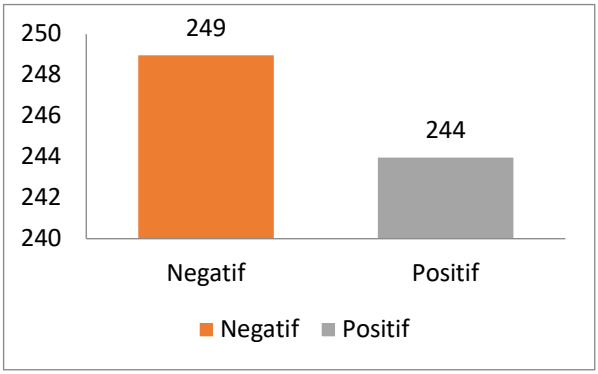
Gambar 8. Hasil scraping data

B. *Hasil Pelabelan Data*

Data yang telah terkumpul melalui proses *scraping* kemudian diberi label untuk memisahkan antara sentimen positif dan sentimen negatif. Dalam rangka mempermudah proses klasifikasi dalam sistem, sentimen positif diberi tanda dengan label 1 dan sentimen negatif diberi tanda dengan label 0. Gambaran mengenai data yang telah diberi label ini tersedia dalam Gambar 9.

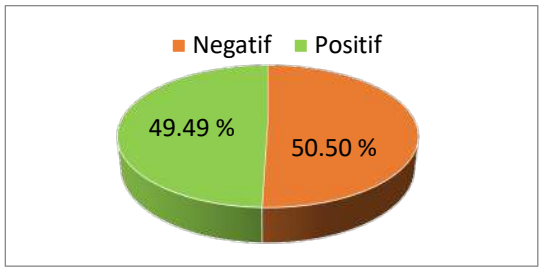
Gambar 9. Hasi labelling data

sentimen positif sebanyak 244 data, dan hasil label sentimen negatif sebanyak 249 data *tweet*. Visualisasi dari klasifikasi data tersebut dapat dilihat pada Gambar 10



Gambar 10. Visualisasi pembagian data

Gambar **7**. Visualisasi pembagian data jika dipresentasikan, data positif sebesar 49.49% dan data negatif sebesar 50.50% seperti yang terdapat pada Gambar 11



Gambar 11. Presentase pembagian data

C. *Library Yang Digunakan*

Library yang digunakan dalam penelitian ini dapat dilihat dalam Tabel 2 berikut.

Tabel 2. Library yang digunakan

No.	Library	Fungsi
1.	Sastrawi	Untuk mengurangi katakata infleksi dalam Bahasa Indonesia ke bentuk dasarnya
2.	Pandas	Untuk mengelola dataset yaitu dengan membaca, menambah, dan mengubahnya.
3.	NLTK	*Platform* untuk membangun program agar bekerja dengan data bahasa manusia.
4.	Numpy	Menyediakan fungsi yang dapat digunakan untuk proses pengolahan atau penyimpanan data sehingga 34 dapat dengan melakukan modifikasi terhadap data untuk pelatihan dan pengujian
5.	Matplotlib	Untuk memvisualisasikan data
6.	Stopwords	Untuk mengidentifikasi dan menghapus kata-kata penghubung yang umum dan memiliki sedikit nilai informasi dalam teks atau dokumen saat melakukan pemrosesan bahasa alami (NLP)
7.	Re	Untuk bekerja dengan ekspresi regular (regex). Regex adalah urutan karakter yang membentuk pola pencarian.
8.	Twepy	Digunakan untuk mengakses dan berinteraksi dengan API Twitter menggunakan bahasa pemrograman Python.
9.	Streamlit	Membuat aplikasi *web* untuk *machine learning* dan *data science*

D. *Hasil Pre-Processing Data*

Proses awal dalam mengelola data teks yang telah diperoleh dalam bentuk aslinya sebelum data tersebut diselidiki lebih jauh. *Pre-pengolahan* dilakukan dengan maksud untuk menghapus gangguan, mengungkap fitur-fitur, menyesuaikan data asli dengan kebutuhan, dan mengubah skala data. Tahap pre-pengolahan memiliki beberapa langkah yang termasuk di dalamnya:

a. Cleaning

Tahapan *Cleanin*g dibuat untuk menghilangkan kata-kata yang tidak diperlukan seperti mention (@), hashtag (#), link, tanda baca dan angka. Selain itu, setiap baris data akan diubah menjadi spasi, menghilangkan spasi di kiri dan kanan teks, dan menghilangkan tanda baca apa pun.

b. Case Folding

Langkah berikutnya adalah Tahapan penyelarasan Huruf, yang melibatkan proses mengubah huruf yang

bercampur antara huruf kecil dan huruf kapital menjadi huruf kecil semuanya.

c. *Stopword Removal, Stemming, tokenizing*

Pada tahapan proses *Tokenizing* yaitu dilakukan dengan tujuan mengubah kalimat menjadi potongan-potongan kata tunggal atau token.

Pada tahapan proses *Stopwords Removal* yaitu dilakukan dengan tujuan menghilangkan kata-kata umum yang dianggap tidak bermakna dan muncul dalam jumlah yang besar dalam teks. Hasilnya diperlihatkan pada Tabel 7 contoh *stopwords* dalam bahasa Indonesia yang akan dihapus adalah "kita, untuk, lebih, yang, dan, dalam".

Pada tahapan proses *Stemming* dilakukan proses yang memiliki tujuan untuk menghilangkan afiks dari kata yang tidak diperlukan, dan selanjutnya mengembalikannya ke bentuk dasarnya.

Hasil *Pre-Processing* dari *Case Folding*, *Cleaning*, *Stopwords Removal, Stemming, Tokenizing* dapat ditampilkan pada Tabel 3

Tabel 3. Hasil pre-processing

Tweet Asli	
Kenaikan harga BBM mengingatkan kita untuk lebih menghargai energi yang kita gunakan setiap hari. Mari kita jadi lebih hemat dan bijak dalam menggunakannya.	
Tahap *Pre-Processing*	**Hasil**
Cleaning	Kenaikan harga BBM mengingatkan kita untuk lebih menghargai energi yang kita gunakan setiap hari Mari kita jadi lebih hemat dan bijak dalam menggunakannya
Case Folding	kenaikan harga bbm mengingatkan kita untuk lebih menghargai energi yang kita gunakan setiap hari. mari kita jadi lebih hemat dan bijak dalam menggunakannya
Tokenizing	[kenaikan, harga, bbm, mengingatkan, kita, untuk, lebih, menghargai, energi yang, kita, gunakan, setiap, hari, mari, kita, jadi, lebih, hemat, dan, bijak, dalam, menggunakannya]
Stopwords Removal	[kenaikan, harga, bbm, mengingatkan, menghargai, energi, gunakan, setiap, hari, mari, jadi, hemat, bijak, menggunakannya]
Stemming	[naik, harga, bbm, ingat, hargai, energi, guna, setiap, hari, mari, jadi, hemat, bijak, guna]
Hasil Akhir	
harga bbm ingat hargai energi guna setiap hari mari jadi hemat bijak guna	

E. *Word Cloud*

Hasil *pre-processing* dapat ditampilkan sebagai wordcloud. Wordcloud akan menampilkan kata berdasarkan frekuensi kemunculan kata tersebut, semakin besar gambar kata tersebut maka semakin sering kata tersebut dimanipulasi secara manual. Berikut tampilan *wordcloud* hasil pengambilan *tweet* tentang kebijakan pemerintah untuk meningkatkan harga bahan bakar minyak,seperti yang ditunjukkan pada Gambar 12

Gambar 12. Wordcloud

F. *Hasil Pengujian*

Data yang telah diolah kemudian dibagi menjadi data latih dan data uji. Dalam setiap penelitian menggunakan data latih dan data uji. Jadi jumlah data latihnya 80%, data ujinya 20%

Klasifikasi dengan menerapkan metode *Support Vector Machine* pada penelitian ini menggunakan skala poin 10. Dengan data latih dan data uji diberi bobot menggunakan bobot TF-IDF.

Langkah selanjutnya adalah menerapkan metode pembobotan TF-IDF. Metode pembobotan ini menggabungkan frekuensi term dan invers frekuensi dokumen dengan menggunakan rumus W=tf*idf. Tujuan dari metode ini adalah untuk menghitung representasi nilai dari sekumpulan nilai tertentu. Di bawah ini adalah contoh data yang akan dihitung menggunakan pembobotan TF-IDF.. Seperti yang ditunjukkan dalam Tabel 4

Tabel 4. Data atau document

Data A	harga bbm ingat hargai energi guna setiap hari mari jadi hemat bijak guna
Data B	harga bbm naik semangat tinggi guna momen mencari alternatif energi ramah lingkungan

Tabel 5. Contoh hasil perhitungan TF-IDF

Kata	Count		TF		IDF	TF*IDF	
	A	B	A	B		A	B
Harga	1	1	0.07	0.07	0	0	0
Bbm	1	2	0.07	0.15	0	0	0
Ingat	1	0	0.07	0	0.69	0.05	0
Hargai	1	0	0.07	0	0.69	0.05	0
Energi	1	1	0.07	0.07	0	0	0
Guna	2	1	0.15	0.07	0	0	0
Setiap	1	0	0.07	0	0.69	0.05	0
Hari	1	0	0.07	0	0.69	0.05	0
Mari	1	0	0.07	0	0.69	0.05	0
Jadi	1	0	0.07	0	0.69	0.05	0
Hemat	1	0	0.07	0	0.69	0.05	0
Bijak	1	0	0.07	0	0.69	0.05	0
Naik	0	1	0	0.07	0.69	0	0.05
Semangat	0	1	0	0.07	0.69	0	0.05
Tinggi	0	1	0	0.07	0.69	0	0.05
Momen	0	1	0	0.07	0.69	0	0.05
Mencari	0	1	0	0.07	0.69	0	0.05
Altermatif	0	1	0	0.07	0.69	0	0.05
Ramah	0	1	0	0.07	0.69	0	0.05

lingkungan	0	1	0	0.07	0.69	0	0.05

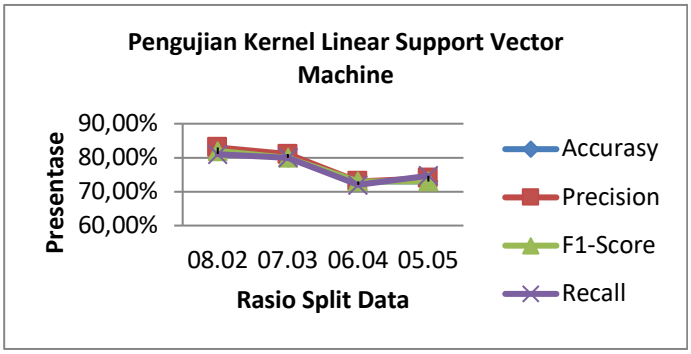
Kemudian dilakukan klasifikasi menggunakan metode *Support Vector Machine.* untuk klasifikasi. Pada rangkaian penelitian ini, kumpulan data akan diuji melalui proses dekomposisi data, yaitu kumpulan data akan dibagi menjadi data pelatihan *(train)* dan data pengujian *(test).* Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk menguji performa model.

1. Kenel *Linear*
   Pada Kernel *Linear* dilakukan sebanyak 4 kali dengan skala 10 untuk mencari hasil dengan ketelitian terbaik. Tabel berikut menunjukkan hasil klasifikasi Support Vector Machine menggunakan kernel linier.

Tabel 6. Hasil pengujian kernel *linear Support Vector Machine*

No	Rasio	Accuracy	Precision	F1-Score	Recall
1.	8:2	82%	83%	82%	81%
2.	7:3	80 %	81%	80%	80%
3.	6:4	73%	73%	73%	72%
4.	5:5	73%	74%	73%	72%

Grafik merepresentasikan hasil *metode Support Vector Machine* dengan pengujian data dapat ditampilkan pada Gambar 13



Gambar 13. Grafik metode *Support Vector Machine* pengujian split data kernel linear

Berdasarkan hasil pengujian pada Tabel 7 dapat disimpulkan bahwa hasil akurasi terbaik untuk kernel *linear* terletak pada perbandingan 8:2 menghasilkan akurasi 82%, akurasi 83%, skor F1 83%, dan recall 83%. Tabel berikut menunjukkan hasil evaluasi kernel linear

Tabel 7. Evaluasi hasil kernel *linear*

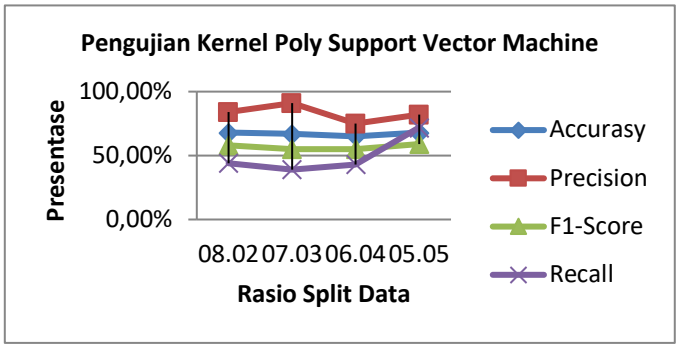
Metrik Evaluasi	Nilai (%)
Akurasi	82
Precision	83
F1-Score	82
Recall	81

1. Kernel Poly
   Pada Kernel *Poly* dilakukan sebanyak 4 kali dengan skala 10 untuk mencari hasil yang paling akurat. Tabel berikut menunjukkan hasil klasifikasi *Support Vector Machine* menggunakan kernel poly

Tabel 8. Pengujian kernel *poly Support Vector Machine*

No	Rasio	Accuracy	Precision	F1-Score	Recall
1.	8:2	68%	84%	58%	44%
2.	7:3	67%	91%	55%	39%
3.	6:4	65%	75%	55%	43%
4.	5:5	68%	82%	59%	45%

Grafik merepresentasikan hasil *metode Support Vector Machine* dengan pengujian data dapat ditampilkan pada Gambar 14



Gambar 14. Grafik metode *Support Vector Machine* pengujian split data kernel poly

Berdasarkan hasil pengujian pada Tabel 9 dapat disimpulkan bahwa hasil akurasi terbaik untuk kernel *poly* terletak pada perbandingan 8:2 menghasilkan akurasi sebesar 68%, *Precision* sebesar 84%, *F1-Score* sebesar 58%, dan *Recall* sebesar 44%. Tabel berikut menunjukkan evaluasi hasil dari kernel *poly*

Tabel 9. Evaluasi hasil kernel poly

Metrik Evaluasi	Nilai (%)
Akurasi	68
Precision	84
F1-Score	58
Recall	44

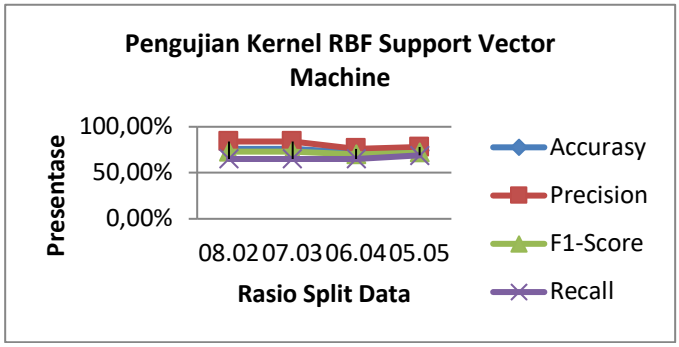
2. Kernel RBF
   Pada kernel RBF dilakukan sebanyak 4 kali dengan skala 10 untuk mencari hasil yang paling akurat.

Tabel berikut menunjukkan hasil klasifikasi *Support Vector Machine* menggunakan kernel RBF

Tabel 10. Pengujian kernel RBF *Support Vector Machine*

No	Rasio	Accuracy	Precision	F1-Score	Recall
1.	8:2	76%	84%	73%	65%
2.	7:3	76%	84%	73%	65%
3.	6:4	72%	76%	70%	65%
4.	5:5	74%	78%	72%	69%

Grafik merepresentasikan hasil *metode Support Vector Machine* dengan pengujian data dapat ditampilkan pada Gambar 15



Gambar 15. Grafik metode Support Vector Machine pengujian split data kernel RBF

Berdasarkan hasil pengujian pada Tabel 11 dapat disimpulkan bahwa hasil akurasi terbaik untuk kernel RBF terletak pada perbandingan 8:2 menghasilkan akurasi sebesar 76%, *Precision* sebesar 84%, *F1-Score* sebesar 73%, dan *Recall* sebesar 65%. Tabel berikut menunjukkan evaluasi hasil dari kernel RBF

Tabel 11. Evaluasi hasil kernel RBF

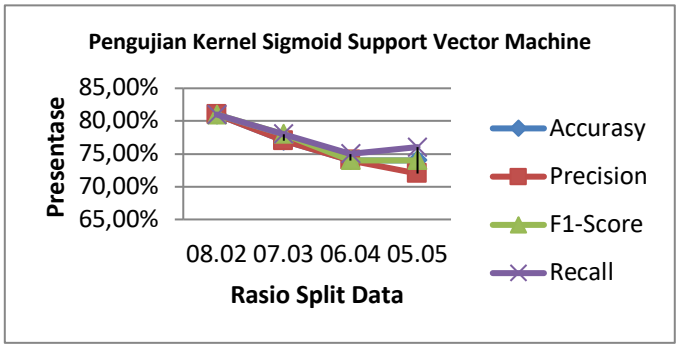
Metrik Evaluasi	Nilai (%)
Akurasi	76
Precision	84
F1-Score	73
Recall	65

3. Kernel Sigmoid
   Pada kernel sigmoid dilakukan sebanyak 4 kali menggunakan skala 10 untuk mencari hasil akurasi akurasi terbaik. Berikut ini tabel hasil klasifikasi *Support Vector Machine* menggunakan kernel *sigmoid*

Tabel 12. . Pengujian kernel RBF Support Vector Machine

No	Rasio	Accuracy	Precision	F1-Score	Recall
1.	8:2	81%	81%	81%	81%
2.	7:3	77%	77%	78%	78%
3.	6:4	74%	74%	74%	75%
4.	5:5	74%	72%	74%	76%

Grafik merepresentasikan hasil *metode Support Vector Machine* dengan pengujian data dapat ditampilkan pada Gambar 16



Gambar 16. Grafik metode Support Vector Machine pengujian split data kernel sigmoid

Berdasarkan hasil pengujian pada Tabel 13 dapat disimpulkan bahwa hasil akurasi terbaik untuk kernel sigmoid terletak pada perbandingan 8:2 menghasilkan akurasi sebesar 81%, *Precision* sebesar 81%, *F1-Score* sebesar 81%, dan *Recall* sebesar 81%. Tabel berikut menunjukkan evaluasi hasil dari kernel *sigmoid*

Tabel 13. Evaluasi hasil kernel sigmoid

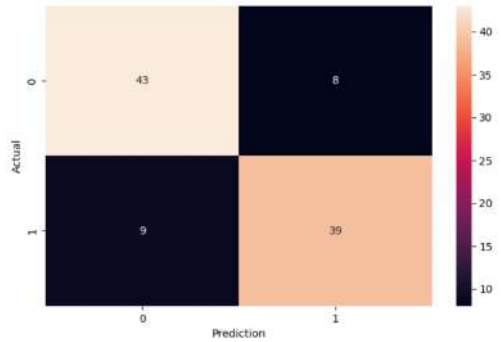
Metrik Evaluasi	Nilai (%)
Akurasi	81
Precision	81
F1-Score	81
Recall	81

Dari hasil pengujian diatas dapat disimpulkan bahwa diantara 4 kernel, hasil akurasi terbaik dihasilkan oleh kernel *linear* adalah dengan perbandingan 8:2 menghasilkan akurasi sebesar 82%, *Precision* sebesar 83%, *F1-Score* sebesar 83%, dan *Recall* sebesar 83%.

G. *Confusion Matrix*

Setelah semua tahap telah dilakukan, proses yang terakhir adalah pengukuran keefektifan model yang telah dibuat dan diuji. Nilai akurasi yang didapatkan sebelumnya akan diukur menggunakan matrix confusion. Hasil pengukuran keefektifan metode yang telah dibangun adalah sebagai berikut :

1. Kernel *Linear*



Gambar 17. *Confusion matrix* analisis sentimen kernel linear model *Support Vector Machine*

Pada gambar diatas adalah prediksi confusion matrix untuk analisis sentimen dengan model algoritma Mesin Vector Pendukung menggunakan kernel linear menunjukkan bahwa sisi kiri menggambarkan Nilai Aktual dan bagian bawah adalah Hasil Prediksi, sehingga dapat disimpulkan terdapat 43 *True Negative (TN),* 9 False Negative (FN), 8 *False Positive (FP),* dan 39 *True Positive (TP).* Pada langkah ini, evaluasi hasil dilakukan menggunakan matriks konfusi. Matriks Konfusi akan menilai performa sistem dalam proses klasifikasi dengan menghitung akurasi, presisi, serta recall untuk mengetahui hasil terbaik. Berikut hasil dari *confusion matrix*

Tabel 14. Data *confusion matrix* kernel *linear*

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	(TP) 39	(FN) 9
Negatif	(FP) 8	(TN) 43

Dari data *Confusion Matrix* diatas, hasil akurasi, presisi, recall kernel *linear* dapat dihitung sebagai berikut :

$$Akurasi = \frac{TP + TN}{(TP + FP + FN + TN)} = \frac{39 + 43}{(39 + 8 + 8 + 43)} = 0.82\%$$

$$Recall = \frac{TP}{\mathrm{TP + FN}} = \frac{39}{39 + 9} = 0.81\%$$

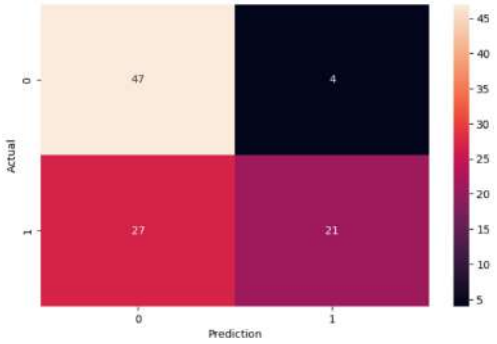
$$Presisi = \frac{TP}{\mathrm{TP + FP}} = \frac{39}{39 + 8} = 0.83\%$$

Berdasarkan perhitungan evaluasi hasil menggunakan Confusion Matrix diatas sama dengan hasil perhitungan yang dilakukan menggunakan program python, maka dapat disimpulkan bahwa kernel *linear* memiliki akurasi 82%, recall sebesar 81%, dan presisi sebesar 83%

```
SVM test Accuracy:  0.8282828282828283
              precision    recall  f1-score   support

           0       0.83      0.84      0.83        51
           1       0.83      0.81      0.82        48

    accuracy                           0.83        99
   macro avg       0.83      0.83      0.83        99
weighted avg       0.83      0.83      0.83        99
```

Gambar 18. Hasil pengujian kernel linear menggunakan *Support Vector Machine*

2. Kernel Poly



Gambar 19. *Confusion matrix* analisis sentimen kernel *poly* model *Support Vector Machine*

Pada gambar diatas adalah prediksi confusion matrix untuk analisis sentimen menggunakan model algoritma *Support Vector Machine* dengan kernel polinomial menunjukkan bahwa sisi kiri menggambarkan Nilai Aktual dan bagian bawah adalah Hasil Prediksi, sehingga dapat diindikasikan terdapat 47 True Negative (TN), 27 False Negative (FN), 4 False Positive (FP), dan 21 True Positive (TP). Pada langkah ini, evaluasi hasil dilakukan menggunakan matriks konfusi. Matriks Konfusi akan menilai performa sistem dalam proses klasifikasi dengan menghitung akurasi, presisi, serta recall untuk mendapatkan hasil terbaik.. Berikut hasil *confusion matrix.*

Tabel 15. Data *confusion matrix* kernel *poly*

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	(TP) 21	(FN) 27
Negatif	(FP) 4	(TN) 47

Dari data Confusion Matrix diatas, hasil akurasi, presisi, recall kernel poly dapat dihitung sebagai berikut :

$$Akurasi = \frac{TP + TN}{(TP + FP + FN + TN)} = \frac{21 + 47}{(21 + 4 + 27 + 47)} = 68\%$$

$$Recall = \frac{TP}{\text{TP} + \text{FN}} = \frac{21}{21 + 27} = 44\%$$

$$Presisi = \frac{TP}{\text{TP} + \text{FP}} = \frac{21}{21 + 4} = 84\%$$

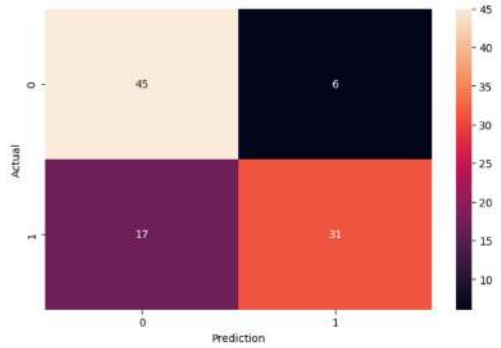
Berdasarkan perhitungan evaluasi hasil menggunakan Confusion Matrix diatas sama dengan hasil perhitungan yang dilakukan menggunakan program python, maka dapat disimpulkan bahwa kernel poly memiliki akurasi 68%, recall sebesar 44%, dan presisi sebesar 84%

SVM test Accuracy: 0.6868686868686869

	precision	recall	f1-score	support
0	0.64	0.92	0.75	51
1	0.84	0.44	0.58	48
accuracy			0.69	99
macro avg	0.74	0.68	0.66	99
weighted avg	0.73	0.69	0.67	99

Gambar 20. Hasil pengujian kernel poly menggunakan Support Vector Machine

3. Kernel RBF



Gambar 21. Confusion matrix analisis sentimen kernel RBF model Support Vector Machine

Pada Gambar diatas prediksi matrix confusion untuk analisis sentimen menggunakan model algoritma Mesin Vector Pendukung dengan kernel RBF menunjukkan bahwa sisi kiri menggambarkan Nilai Aktual dan bagian bawah adalah Hasil Prediksi, sehingga dapat disimpulkan terdapat 45 True Negative (TN), 17 False Negative (FN), 6 False Positive (FP), dan 31 True Positive (TP). Pada langkah ini, evaluasi hasil dilakukan menggunakan matriks konfusi. Matriks Konfusi akan menilai performa sistem dalam proses klasifikasi dengan menghitung akurasi, presisi, serta recall untuk mengetahui hasil terbaik. Berikut hasil confusion matrix.

Tabel 16. Data confusion matrix kernel RBF

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	(TP) 31	(FN) 17
Negatif	(FP) 6	(TN) 45

Dari data Confusion Matrix diatas, hasil akurasi, presisi, recall kernel RBF dapat dihitung sebagai berikut :

$$Akurasi = \frac{TP + TN}{(TP + FP + FN + TN)} = \frac{31 + 45}{(31 + 6 + 17 + 45)} = 76\%$$

$$Recall = \frac{TP}{\text{TP} + \text{FN}} = \frac{31}{31 + 17} = 65\%$$

$$Presisi = \frac{TP}{\text{TP} + \text{FP}} = \frac{31}{31 + 6} = 84\%$$

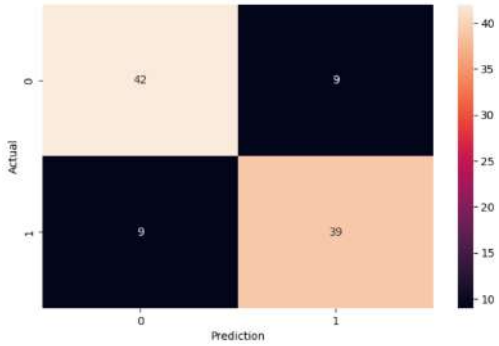
Berdasarkan perhitungan evaluasi hasil menggunakan Confusion Matrix diatas sama dengan hasil perhitungan yang dilakukan menggunakan program python, maka dapat disimpulkan bahwa kernel RBF memiliki akurasi sebesar 76%, recall 65% dan presisi 84%

SVM test Accuracy: 0.7676767676767676

	precision	recall	f1-score	support
0	0.73	0.88	0.80	51
1	0.84	0.65	0.73	48
accuracy			0.77	99
macro avg	0.78	0.76	0.76	99
weighted avg	0.78	0.77	0.76	99

Gambar 22. Hasil pengujian kernel RBF menggunakan Support Vector Machine

4. Kernel Sigmoid



Gambar 23. Confusion matrix analisis sentimen kernel sigmoid model Support Vector Machine

Pada Gambar diatas prediksi matrix confusion untuk analisis sentimen menggunakan model algoritma Mesin Vector Pendukung dengan kernel sigmoid menyatakan bahwa sisi kiri menggambarkan Nilai Aktual dan bagian bawah adalah Hasil Prediksi,

sehingga dapat diindikasikan terdapat 42 True Negative (TN), 9 False Negative (FN), 9 False Positive (FP), dan 39 True Positive (TP). Pada tahapan ini, evaluasi hasil dilakukan menggunakan matriks konfusi. Matriks Konfusi akan menilai performa sistem dalam proses klasifikasi dengan menghitung akurasi, presisi, serta recall untuk mendapatkan hasil terbaik. Berikut hasil confusion matrix.

Tabel 17. Data confusion matrix kernel sigmoid

Nilai Aktual	Nilai Prediksi	
	Positif	Negatif
Positif	(TP) 39	(FN) 9
Negatif	(FP) 9	(TN) 42

Dari data Confusion Matrix diatas, hasil akurasi, presisi, recall kernel sigmoid dapat dihitung sebagai berikut :

$$Akurasi = \frac{TP + TN}{(TP + FP + FN + TN)} = \frac{39 + 42}{(39 + 9 + 9 + 42)} = 81\%$$

$$Recall = \frac{TP}{\mathrm{TP + FN}} = \frac{39}{39 + 9} = 81\%$$

$$Recall = \frac{TP}{\mathrm{TP + FP}} = \frac{39}{39 + 9} = 81\%$$

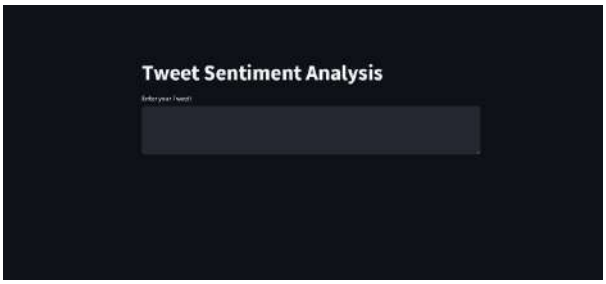
Berdasarkan perhitungan evaluasi hasil menggunakan Confusion Matrix diatas sama dengan hasil perhitungan yang dilakukan menggunakan program python, maka dapat disimpulkan bahwa kernel sigmoid memiliki akurasi sebesar 81%, recall 81% dan presisi sebesar 81%

```
SVM test Accuracy:  0.8181818181818182
              precision    recall  f1-score   support

           0       0.82      0.82      0.82        51
           1       0.81      0.81      0.81        48

    accuracy                           0.82        99
   macro avg       0.82      0.82      0.82        99
weighted avg       0.82      0.82      0.82        99
```

Gambar 24. Hasil pengujian kernel sigmoid menggunakan Support Vector Machine

H. *Tampilan Aplikasi Web*

Berikut adalah tampilan dari implementasi sistem analisis sentimen opini terhadap kebijakan pemerintah dalam menaikkan bahan bakar minyak dengan Support Vector Machine dengan model yang telah dihasilkan dari tahap pelatihan. Implementasi sistem ini berbentuk web yang dibangun dengan menggunakan library pyhton Streamlit.



Gambar 25. Tampilan utama aplikasi web

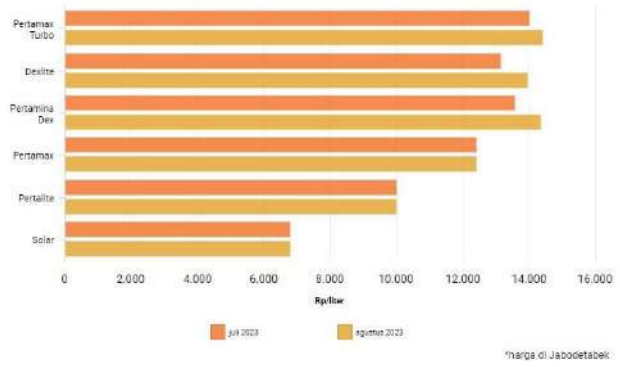
Kemudian pada Gambar 26 dan Gambar 27 menunjukkan hasil prediksi opini dari teks opini yang dimasukkan oleh pengguna

Gambar 26. Hasil prediksi opini positif

Gambar 27. Hasil prediksi opini negatif

I. Rekomendasi Untuk Masyarakat

Dengan merujuk pada Keputusan Menteri (Kepmen) ESDM No. 245.K/MG.01/MEM.M/2022, yang merupakan revisi dari Kepmen No. 62 K/12/MEM/2020 mengenai Formula Harga Dasar dalam Penetapan Harga Jual Eceran untuk Bahan Bakar Minyak Jenis Bensin dan Minyak Solar yang Dijual Melalui SPBU, terjadi peningkatan harga.. Dapat dilihat pada Gambar 28



Gambar 28. Grafik kenaikan bahan bakar minyak bulan Juli – Agustus 2023

Dari data tersebut terdapat cara untuk mengatasi situasi kenaikan bahan bakar minyak ini, berikut beberapa rekomendasi untuk masyarakat :

1. Pengelolaan Anggaran : Tinjau ulang anggaran pribadi atau keluarga Anda. Identifikasi area di mana Anda dapat mengurangi pengeluaran yang tidak penting atau mengurangi konsumsi untuk sementara waktu. Prioritaskan kebutuhan utama seperti makanan, perawatan kesehatan, pendidikan, dan transportasi.
2. Berbagi Perjalanan: Jika memungkinkan, koordinasikan perjalanan dengan teman, keluarga, atau rekan kerja untuk berbagi biaya transportasi. Carpooling atau berbagi kendaraan dapat membantu mengurangi biaya bahan bakar.
3. Pertimbangkan Transportasi Alternatif: Jika jaraknya memungkinkan, pertimbangkan untuk menggunakan transportasi umum, bersepeda, atau berjalan kaki. Ini bukan hanya mengurangi biaya bahan bakar, tetapi juga membantu mengurangi dampak lingkungan.
4. Mengambil Keputusan tentang Kendaraan Efisien Bahan Bakar: Apabila Anda sedang dalam proses mencari kendaraan baru, pertimbangkanlah pilihan kendaraan yang memiliki efisiensi tinggi dalam hal konsumsi bahan bakar. Kendaraan hibrida ataupun listrik mungkin menjadi alternatif yang lebih hemat dalam jangka waktu yang lebih panjang.
5. Perawatan Kendaraan yang Lebih Baik: Pastikan kendaraan Anda dalam kondisi yang baik dengan menjalankan perawatan rutin. Pemeliharaan yang baik, seperti pergantian oli teratur dan pemeriksaan ban, dapat membantu meningkatkan efisiensi bahan bakar.
6. Berkendara dengan Bijak: Praktikkan gaya berkendara yang hemat bahan bakar, seperti menjaga kecepatan konstan, menghindari pengereman mendadak, dan mematikan mesin jika berhenti dalam waktu yang lama.
7. Pertimbangkan Alternatif Energi: Jika memungkinkan, eksplorasi penggunaan energi alternatif seperti gas alam cair (LPG) atau biodiesel, yang mungkin lebih ekonomis daripada bahan bakar konvensional.
8. Rencanakan Perjalanan dengan Cermat: Pertimbangkan untuk merencanakan perjalanan Anda dengan lebih baik. Hindari kemacetan dan perjalanan jarak jauh yang tidak perlu. Gunakan aplikasi navigasi untuk mencari rute tercepat dan menghindari lalu lintas.
9. Prioritaskan Pemanfaatan Energi yang Efisien di Rumah: Selain kendaraan bermotor, pertimbangkan mengurangi penggunaan energi di dalam rumah Anda. Matikan perangkat listrik saat tidak dipakai, gantikan lampu dengan yang lebih efisien dalam penggunaan energi, dan pertimbangkan untuk menggunakan sumber energi terbarukan seperti panel surya.
10. Tetap Positif dan Adaptif: Fluktuasi harga BBM adalah hal yang biasa dalam ekonomi. Penting untuk tetap positif dan beradaptasi dengan perubahan tersebut. Mengelola keuangan dengan bijak dan membuat perubahan dalam gaya hidup jika diperlukan dapat membantu menghadapi tantangan ini.

IV. KESIMPULAN

Berdasarkan hasil yag didapatkan dari penelitian dengan menggunakan metode *Support Vector Machine* pada analisis sentimen di media sosial Twitter dapat ditarik kesimpulan sebagai berikut:

1. Penerapan dalam pengambilan data tweet mengenai kebijakan pemerintah terkait kenaikan harga bahan bakar minyak dilakukan melalui ekstensi Google Colab. Proses pengumpulan data teks tweet dari media sosial Twitter berkaitan dengan kebijakan pemerintah terhadap kenaikan harga bahan bakar minyak dilakukan melalui teknik scraping dengan menggunakan API Twitter dan kata kunci "bbm naik". Jumlah total data tweet yang berhasil diambil sebanyak 493 data dari Februari 2023 hingga Maret 2023.
2. Data tweet yang memiliki sentimen positif dan negatif terkait kebijakan kenaikan harga bahan bakar minyak dianalisis menggunakan metode Mesin Vector Pendukung. Setelah diberi label, hasilnya menunjukkan 244 data sentimen positif dan 249 data sentimen negatif. Dalam presentasi, proporsi data positif adalah 49.49% dan data negatif adalah 50.50%.
3. Setelah melalui proses preprocessing, data dipecah menjadi dua bagian, yaitu data latig dan data uji dengan perbandingan 8:2. Ini berarti data pelatihan terdiri dari 80% dan sisanya, 20%, digunakan sebagai data pengujian.Dalam implementasi metode Mesin Vector Pendukung dengan skala 10 dan pemberian bobot menggunakan metode pembobotan TF-IDF, hasil akurasi terbaik ditemukan pada kernel linear dengan perbandingan 8:2. Hasil akurasi tersebut

mencapai 82%, dengan nilai presisi sebesar 83%, F1-Score sebesar 82%, dan Recall sebesar 81%. Analisis terhadap sentimen dari model algoritma Mesin Vector Pendukung menyatakan bahwa bagian kiri mewakili Nilai Aktual dan bagian bawah menunjukkan Hasil Prediksi, sehingga dapat disimpulkan bahwa terdapat 43 True Negative (TN), 9 False Negative (FN), 8 False Positive (FP), dan 39 True Positive (TP).

V. SARAN

Penelitian tentang analisis sentimen dengan menggunakan dataset unggahan masyarakat Indonesia memiliki peluang besar untuk terus diperluas. Salah satu upaya pengembangan yang dapat dilakukan adalah mencoba metode atau model algoritma yang berbeda, diharapkan dengan menggabungkan dua atau lebih metode untuk meningkatkan akurasi sistem. Selain itu, setiap tahap penelitian dapat diperbaiki, terutama tahap pelabelan, sehingga penelitian ini dapat menghasilkan tingkat akurasi yang lebih tinggi.

REFERENSI

[1] U. Kurniasih and A. T. Suseno, "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)," *J. Media Inform. Budidarma*, vol. 6, no. 4, pp. 2335–2340, 2022, doi: 10.30865/mib.v6i4.4958.

[2] Harunurrasyidh, "Ekonomi pembangunan," vol. 11, no. 2, pp. 29–41, 2013.

[3] G. Rozy Hrp, N. Aslami, and P. Studi Manajemen Fakultas Ekonomi Bisnis Islam, "Analisis Damfak Kebijakan Perubahan Publik Harga BBM terhadap Perekonomian Rakyat Indonesia," *J. Ilmu Komputer, Ekon. dan Manaj.* , vol. 2, no. 1, pp. 1464–1474, 2022.

[4] P. Ir, B. Wirjodirdjo, M. Eng, and K. K. Ummatin, "DAMPAK KEBIJAKAN HARGA BBM TERHADAP KEMISKINAN DI INDONESIA : Sebuah Pendekatan Model Dinamik Kampus ITS Sukolilo Surabaya 60111," *Chart*, 2008.

[5] R. Wati and S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," *J. Tek. Inform. UNIKA St. Thomas*, vol. 06, pp. 240–247, 2021, doi: 10.54367/jtiust.v6i2.1465.

[6] Siti Nurhasanah Nugraha, Tri Rivanie, S. Rahayu, W. Gata, and R. Pebrianto, "Sentimen Analisis Penerapan Social Distancing Menggunakan Feature Selection Pada Algoritma Support Vector Machine," *J. Tek. Komput. AMIK BSI*, 2020, doi: 10.31294/jtk.v4i2.

[7] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor," *UNNES J. Math.*, vol. 10, no. 2, pp. 22–27, 2021, [Online]. Available: http://journal.unnes.ac.id/sju/index.php/ujm.

[8] K. A. Rokhman, B. Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 1–7, 2021, doi: 10.24076/joism.2021v3i1.341.

[9] Huda, R. M. (2017, Juni 11). Indonesia Pengguna Twitter Terbesar Ketiga Dunia. Retrieved from setara.net: https://setara.net/indonesia-pengguna-twitter-terbesar-ketiga-dunia/