



Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine

Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine

**Rizky Rahman Salam¹, Muhammad Fajri Jamil², Yusril Ibrahim³,
Rahmaddeni^{4*}, Soni⁵, Herianto⁶**

^{1,2,3,4}Program Studi Teknik Informatika, STMIK AMIK Riau, Indonesia

⁵Program Studi Teknik Informatika, Universitas Muhammadiyah Riau, Indonesia

⁶Program Studi Sistem Informasi, Universitas Hang Tuah Pekanbaru, Indonesia

E-Mail: ¹2010031802080@sar.ac.id, ²2010031802059@sar.ac.id, ³2010031802092@sar.ac.id,
⁴rahmaddeni@sar.ac.id, ⁵soni@umri.ac.id, ⁶herianto.sy@gmail.com

Received Jan 10th 2023; Revised Mar 05th 2023; Accepted Apr 22th 2023
Corresponding Author: Rahmaddeni

Abstract

Fuel oil (BBM) is one of the basic needs of society. However, the high price of fuel can create an economic burden for the poor. To overcome this, the government implemented the Direct Cash Assistance (BLT) program as a form of assistance for people who experience economic imbalance. The purpose of this research is to analyze people's sentiment towards the Direct Cash Assistance (BLT) Fuel Oil (BBM) program. This research uses scraping data collection techniques, which is taking data from social media Instagram. The amount used was 356 data. The classification process used is based on the Support Vector Machine (SVM) learning model and evaluation with confusion matrix. From the calculation results, it can be seen that the sentiment classification process using the SVM method obtained an accuracy rate of 85.98%, an average precision value of 82.25%, an average recall value of 66.35%, and an average f-measure value of 73.44%. The results obtained show that negative sentiment is more than positive sentiment, with a percentage of 78.61% and 21.34% respectively. From the sentiment analysis conducted, it was found that negative sentiments were the most prevalent, indicating that people were dissatisfied with the BBM Direct cash assistance (BLT) program. In response to the dominant negative sentiment, it is necessary to implement a structured distribution and data collection strategy so that the level of community disappointment can be minimized.

Keywords: Economic Burden, Fuel, Government, Sentiment, Support Vector Machine

Abstrak

Bahan bakar minyak (BBM) merupakan salah satu kebutuhan pokok masyarakat. Namun, harga BBM yang tinggi dapat menyebabkan beban ekonomi bagi masyarakat yang tidak mampu. Dalam rangka mengatasi masalah ini, pemerintah telah menerapkan program Bantuan Langsung Tunai (BLT) sebagai bentuk bantuan bagi masyarakat yang mengalami ketidakseimbangan ekonomi. Tujuan dari penelitian ini adalah untuk menganalisis sentimen masyarakat terhadap program Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM). Penelitian ini menggunakan teknik pengumpulan data *scraping*, yaitu mengambil data dari media sosial Instagram. Jumlah yang digunakan sebanyak 356 data. Proses klasifikasi yang digunakan berdasarkan model pembelajaran dari Support Vector Machine (SVM) dan evaluasi dengan confusion matrix. Dari hasil perhitungan, terlihat bahwa proses klasifikasi sentimen menggunakan metode SVM didapatkan tingkat *accuracy* 85,98%, rata-rata nilai *precision* 82,25%, nilai rata-rata *recall* 66,35%, dan nilai rata-rata *f-measure* 73,44%. Hasil yang diperoleh menunjukkan bahwa sentimen negatif lebih banyak daripada sentimen positif, dengan masing-masing persentase 78.61% dan 21.34%. Dari analisis sentimen yang dilakukan, ditemukan bahwa sentimen negatif adalah yang paling banyak muncul, hal ini menunjukkan bahwa masyarakat tidak puas dengan bantuan langsung tunai BBM. Sebagai respon terhadap sentimen negatif yang dominan, perlu diterapkan strategi untuk melakukan pemerataan bantuan langsung tunai dan pendata'an yang terstruktur agar tingkat kekecewaan masyarakat dapat diminimalisir.

Kata Kunci: Beban Ekonomi, BBM, Pemerintah, *Sentiment*, *Support Vector Machine*

1. PENDAHULUAN

Kenaikan harga Bahan Bakar Minyak (BBM) di Indonesia merupakan masalah yang telah lama dihadapi oleh masyarakat dan pemerintah. Bahan bakar minyak (BBM) adalah salah satu kebutuhan yang tidak dapat dipisahkan dalam kehidupan masyarakat. Minyak merupakan bahan baku utama bagi perekonomian [1], sehingga kenaikan harga BBM akan berdampak langsung terhadap harga-harga barang dan jasa lainnya. Pemerintah selalu berusaha untuk menstabilkan harga BBM dengan mengeluarkan kebijakan-kebijakan yang bertujuan untuk mengurangi beban pemerintah dalam hal subsidi BBM. Salah satu penyebab utama kenaikan harga BBM di Indonesia adalah melonjaknya harga minyak dunia. Harga minyak dunia sangat dipengaruhi oleh faktor-faktor seperti permintaan dan penawaran, konflik politik, dan perubahan cuaca. Selain itu, kondisi ekonomi global juga mempengaruhi harga minyak dunia. Saat ekonomi dunia sedang mengalami krisis, permintaan minyak akan menurun sehingga harga minyak pun akan turun. Namun saat ekonomi dunia sedang tumbuh, permintaan minyak akan meningkat sehingga harga minyak akan naik.

Dari permasalahan tersebut, pemerintah telah mengambil tindakan yang sesuai dengan cara menerapkan program Bantuan Langsung Tunai (BLT), sebagai bentuk solusi bagi masyarakat yang mengalami ketidakseimbangan ekonomi. Menurut [2], secara umum program bantuan tunai dapat dibedakan menjadi dua jenis, yaitu program yang diberikan dengan syarat tertentu (CCT) dan program yang diberikan tanpa syarat (UCT). Kedua jenis program ini digunakan untuk mengatasi masalah kemiskinan dan meningkatkan kesejahteraan masyarakat, namun dengan pendekatan yang berbeda. Program bantuan langsung tunai (BLT) Bahan Bakar Minyak (BBM) termasuk dalam jenis program bantuan tunai tidak bersyarat. BLT adalah dana tunai yang diberikan untuk menjaga daya beli rumah tangga yang miskin dan rentan dari dampak kenaikan harga akibat kenaikan harga BBM [3]. Masyarakat dapat tetap membeli bahan bakar yang diperlukan tanpa terlalu terpengaruh oleh kenaikan harga bahan bakar. BLT disalurkan melalui berbagai tahap dan rangkaian proses untuk diterima dengan baik oleh masyarakat [4].

Program Bantuan Langsung Tunai pertama kali diusulkan oleh pemerintah Brasil dan kemudian diadopsi oleh negara lain. Setiap negara menerapkan mekanisme dan besaran dana yang berbeda-beda sesuai kebijakan pemerintah masing-masing [5]. Negara Indonesia sendiri, proses pembagian dana BLT didasarkan pada data masyarakat miskin yang diterima dari BPS oleh Departemen Sosial, kemudian data tersebut dikirim ke Departemen Dalam Negeri untuk divalidasi oleh Gubernur dan timnya [6]. Ketika proses validasi selesai, kartu dalam bentuk kupon akan diterbitkan untuk penerima dana yang digunakan selama setiap periode pembagian BLT. Program ini mendapat banyak kritikan dari masyarakat, karena dalam pelaksanaannya masih banyak terjadi masalah seperti tidak meratanya pembagian BLT BBM, dan program ini seolah-olah dibuat sebagai pengganti subsidi BBM karena harga BBM yang naik. Kebijakan menaikkan harga BBM menuai kontroversi di seluruh media sosial [7]. Kontroversi ini juga diperkuat oleh berbagai aksi unjuk rasa yang dilakukan oleh beberapa elemen masyarakat yang menentang kebijakan ini. Hal ini dapat dilihat dari maraknya perdebatan dan opini-opini yang muncul di berbagai media sosial seperti *Facebook*, *Twitter*, dan *Instagram*.

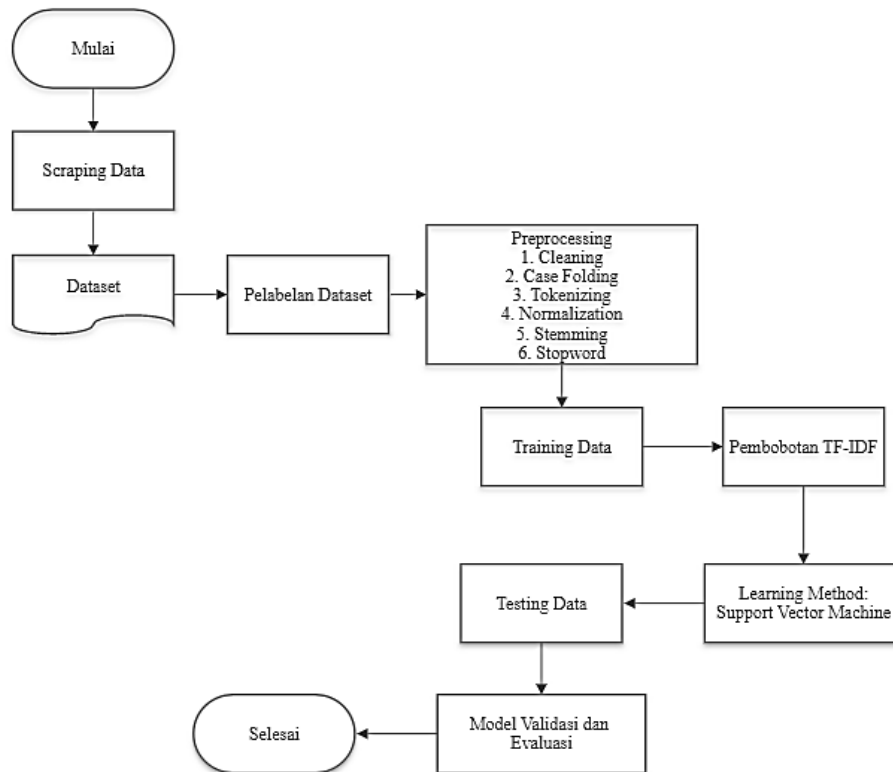
Analisis sentimen merupakan teknik yang digunakan untuk mengevaluasi persepsi masyarakat terhadap suatu entitas. Dengan meningkatnya penggunaan media sosial, analisis sentimen menjadi semakin penting karena memungkinkan untuk mengidentifikasi opini publik yang positif, negatif, atau netral. Analisis sentimen digunakan dalam berbagai bidang seperti pemasaran, politik, dan perusahaan untuk mengevaluasi sikap dan emosi masyarakat terhadap produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atribut lainnya [8]. Penelitian analisis sentimen telah berkembang dengan cepat dan telah ada pendekatan-pendekatan baru yang diluncurkan. *Machine Learning-based sentiment analysis* telah menjadi pendekatan populer yang dianggap efektif dalam menganalisis data [9]. Salah satu metode Machine Learning yang digunakan dalam analisis sentimen adalah Support Vector Machine (SVM). SVM dapat diartikan sebagai proses menemukan batas pemisah yang optimal antara dua kelas di dalam input space yang dikenal sebagai hyperplane [10]. SVM digunakan untuk mempelajari pola dalam data dan menentukan kategori sentimen yang sesuai dengan data tersebut. Selain itu, teknologi Natural Language Processing (NLP) juga digunakan untuk mengekstrak informasi dari teks yang sangat bermanfaat dalam analisis sentimen. Dengan menggabungkan metode Machine Learning dan NLP, analisis sentimen menjadi lebih akurat dan efektif dalam menangkap persepsi masyarakat terhadap suatu peristiwa atau produk.

Pada penelitian sebelumnya, [9] melakukan analisis sentimen mengenai opini publik terhadap kebijakan pemerintah Indonesia dalam penanganan covid-19. Penelitian tersebut menggunakan algoritma Support Vector Machine dengan menggunakan poly kernel. Hasilnya yaitu algoritma SVM dengan menggunakan poly kernel, dapat digunakan sebagai algoritma cerdas untuk memperbaiki sentimen di twitter untuk data baru dengan cepat dan akurat. Selain itu, penelitian lain mendukung hasil yang menyatakan bahwa SVM dan *Naïve Bayes* memiliki kinerja yang baik untuk mengklasifikasikan opini [11]. Selain itu [7] juga melakukan analisis sentimen mengenai opini publik terhadap bantuan subsidi upah (BSU) pada kenaikan harga BBM. Penelitian tersebut menggunakan algoritma *Naïve Bayes*. Hasilnya yaitu tanggapan masyarakat Indonesia pada kata kunci BBM didominasi oleh sentimen negatif, sedangkan pada kata kunci BSU didominasi oleh sentimen positif. Namun pada penelitian tersebut masih dirasa kurang, karena hanya menggunakan algoritma *Naïve Bayes*, dan hanya menggunakan satu sumber dataset, yaitu *Twitter*.

Berdasarkan permasalahan yang telah disebutkan diatas, maka penting untuk dilakukan penelitian analisa sentimen terhadap opini masyarakat tentang program Bantuan Langsung Tunai Bahan Bakar Minyak menggunakan metode Support Vector Machine (SVM) dan dengan sumber data yang berbeda, yaitu *Instagram*. Metode ini digunakan untuk mengidentifikasi sentimen positif dan negatif dari opini masyarakat. Dengan melakukan analisa sentimen, diharapkan dapat memberikan gambaran tentang pandangan masyarakat terhadap program Bantuan Langsung Tunai Bahan Bakar Minyak dan memberikan masukan bagi pemerintah dalam menyusun kebijakan yang lebih baik dan adil.

2. METODOLOGI PENELITIAN

Teknik pengumpulan data yang dilakukan pada penelitian ini yaitu *Web Scraping*, mengambil data dari media sosial *Instagram*. Jumlah yang digunakan sebanyak 356 data. *Flowchat* penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Scraping Data

Scraping Data atau *Web Scraping* merupakan teknik untuk mengambil data dari sebuah situs web dengan cara yang efektif dan cepat [12]. Data yang diambil berupa teks, atau informasi dalam halaman web. Setelah data diperoleh melalui *web scraping* lalu data tersebut akan digunakan untuk keperluan analisis sentimen, dengan mengumpulkan komentar atau ulasan dari berbagai sumber dan menganalisis sentimen dari komentar tersebut.

2.2 Data Preprocessing

Data preprocessing adalah proses pengolahan data awal yang digunakan untuk mengubah data mentah yang diperoleh dari berbagai sumber menjadi informasi yang lebih bersih dan dapat digunakan untuk analisis lebih lanjut [13]. Ada beberapa tahap yang dilakukan dalam data preprocessing, yaitu:

1. Cleaning

Proses pembersihan data bertujuan untuk menemukan dan memperbaiki atau menghapus data yang tidak valid atau tidak berguna dari suatu kumpulan data [14]. Setelah melalui proses cleaning data, data akan siap untuk digunakan dalam tahap selanjutnya dari data preprocessing, yaitu pengolahan dan pengintegrasian data. Namun, dalam beberapa kasus, cleaning data dapat dilakukan secara terus menerus sesuai dengan kebutuhan analisis atau pemodelan yang dilakukan.

2. Case folding

Case folding atau case normalization adalah proses pada text preprocessing yang dilakukan untuk mengubah semua huruf pada teks ke dalam huruf kecil atau huruf besar yang sama, sehingga memudahkan dalam proses perbandingan dan analisis teks [15]. Proses case folding juga dapat digunakan untuk mengatasi masalah penulisan yang tidak konsisten dalam data, seperti penggunaan huruf besar dan kecil yang tidak sesuai standar atau ada kesalahan ejaan. Dengan mengubah semua karakter menjadi huruf kecil, konsistensi dalam data akan diperbaiki sehingga lebih mudah untuk diproses dan dianalisis.

3. Tokenizing

Merupakan suatu tahapan untuk memecah suatu kumpulan teks menjadi sebuah kata. Analisis kata-kata dalam teks menjadi penting karena dengan melakukan tokenisasi, makna dari teks dapat dengan mudah ditentukan [16]. Dalam pemrosesan bahasa alami, tokenizing sering digunakan untuk memecah teks menjadi unit-unit yang lebih kecil dan mudah diolah oleh mesin. Beberapa algoritma tokenizing yang umum digunakan adalah pemotongan karakter, pemotongan kata, dan pemotongan frase.

4. Normalization

Teknik normalisasi yaitu mengubah data menjadi bentuk yang lebih standar [17]. Normalisasi sering digunakan untuk mengubah teks menjadi bentuk yang lebih standar sebelum dianalisis. Normalisasi dapat membantu meningkatkan akurasi dalam analisis teks dan membuat data lebih mudah diolah.

5. Stemming

Proses melakukan reduksi kata atau token ke dasar bentuk kata [18]. Tujuannya adalah untuk mengurangi kata-kata yang berbeda menjadi bentuk kata dasar yang sama, sehingga lebih mudah untuk mengelompokkan atau menganalisis kata-kata tersebut. Proses ini sering digunakan dalam natural language processing (NLP) dan information retrieval (IR).

6. Stopwords

Kata yang tidak memiliki makna atau tidak efektif akan dihilangkan dengan menggunakan stopwords [19]. Stopwords biasanya diabaikan dalam proses pemrosesan teks karena dianggap tidak berguna dan hanya menambah ukuran file tanpa memberikan informasi yang bermanfaat. Namun, ada juga kasus di mana stopwords perlu dipertahankan karena memiliki makna penting dalam konteks tertentu.

2.3 TF-IDF

TF-IDF (term frequency-inverse document frequency) adalah metode yang digunakan untuk mengevaluasi pentingnya sebuah kata dalam suatu dokumen terhadap koleksi dokumen lain. Fitur TF-IDF untuk optimasi dalam analisis sentimen. Dengan menggabungkan ekstraksi fitur TF-IDF dan algoritma stochastic gradient descent, analisis sentimen dapat mengklasifikasikan teks dalam bahasa Indonesia dengan tepat menurut sentimen positif dan negatif[20].

TF-IDF menganggap bahwa kata-kata yang sering muncul dalam suatu dokumen akan memiliki nilai yang lebih tinggi daripada kata-kata yang jarang muncul. Jika kata tersebut muncul dalam banyak dokumen, nilai TF-IDF-nya akan menurun karena dianggap kurang spesifik. Hasil dari perhitungan TF-IDF dapat digunakan untuk berbagai tujuan seperti analisis teks, pengelompokan dokumen, dan pembuatan indeks. Perhitungan ini dapat dilakukan secara manual atau menggunakan alat bantu yang tersedia. Rumus TF-IDF dapat dilihat pada Persamaan 1.

$$w_{t,d} = w_{tf} \times idf_t \quad (1)$$

$w_{t,d}$ adalah banyaknya bobot kata disetiap dokumen, $tf_{t,d}$ adalah jumlah kemunculan term dalam dokumen, N adalah jumlah dokumen keseluruhan, dan df_t adalah jumlah dokumen yang mengandung term.

Keterangan:

N	= Jumlah keseluruhan teks dokumen
w_{tf}	= Bobot kata
$tf_{t,d}$	= Jumlah kata atau term yang muncul dalam dokumen
df_t	= Jumlah dokumen mengandung suatu kata atau term
idf_t	= Bobot inverse
$w_{t,d}$	= Pembobotan TF-IDF

2.4 Klasifikasi menggunakan Support Vector Machine

Support Vector Machine (SVM) adalah algoritma klasifikasi yang digunakan untuk membuat garis pemisah (hyperplane) antara dua atau lebih kelas. Proses klasifikasi yang digunakan berdasarkan model pembelajaran dari SVM dan mengevaluasi dengan confusion matrix disajikan pada Tabel 1.

Tabel 1. *Confusion Matrix*

Komentar	Prediksi Data	
	Negatif	Positif
Negatif	TN	FN
Positif	FP	TP

Confusion matrix adalah tabel yang digunakan untuk mengevaluasi kinerja algoritme klasifikasi [21]. Nilai Accuracy, precision, recall, dan f-measure dapat dihitung setelah mengetahui nilai-nilai dalam confusion matrix [22].

1. Accuracy mengukur seberapa baik hasil klasifikasi yang dilakukan oleh suatu model sesuai dengan nilai yang sebenarnya. Ini menunjukkan persentase dari klasifikasi yang benar yang dilakukan oleh model.

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \times 100 \% \quad (1)$$

2. Precision mengukur seberapa baik model dalam mengidentifikasi sampel positif yang sebenarnya. Ini mengukur rasio dari sampel positif yang benar terhadap jumlah sampel positif yang diidentifikasi oleh model.

$$Precision = \frac{TP}{TP+FP} \times 100 \% \quad (2)$$

3. Recall mengukur seberapa baik model dalam menemukan sampel yang sebenarnya positif. Ini mengukur rasio dari sampel positif yang benar yang ditemukan oleh model terhadap jumlah sampel positif yang sebenarnya.

$$Recall = \frac{TP}{TP+FN} \times 100 \% \quad (3)$$

4. Nilai *f-measure* didapatkan dengan memadukan nilai *recall* dan *precision*.

$$f - measure = \frac{2 \times recall \times precision}{recall + precision} \times 100 \% \quad (4)$$

Keterangan:

- TN = Kelas kata negatif terprediksi negatif
- FN = Kelas kata negatif terprediksi kata positif
- FP = Kelas kata positif terprediksi negatif
- TP = Kelas kata prediksi benar bernilai positif

2.5 Validasi

Pada tahapan validasi, ditujukan untuk mengevaluasi kinerja model dengan menggunakan data yang tidak digunakan dalam data latih. Metode yang digunakan dalam validasi yaitu *Stratified k-fold cross validation*. Kemudian dalam melakukan tahapan proses evaluasi SVM, dapat menggunakan metrik seperti akurasi, *precision*, *recall*, dan *f-measure*. Kedua proses ini sangat penting dalam proses pembuatan model SVM untuk memastikan bahwa model dapat digunakan dengan baik pada data baru dan dapat memberikan hasil yang baik.

3. HASIL DAN ANALISIS

Data training diambil dari Instagram Page Kementerian Sosial Republik Indonesia dengan cara scraping komentar. Didapatkan sebanyak 356 data yang terdiri dari 76 data sentimen positif dan 280 data sentimen negatif. Data tersebut kemudian diolah sehingga diperoleh nilai *Accuracy*, *recall*, *precision*, dan *f-measure*. Penelitian ini menggunakan sampel data yang sudah diberikan label pada Tabel 1.

Tabel 1. Sampel Data Sentimen Positif dan Negatif

No	Text	Label
1.	Lebih baik turunkan harga drpd bantuan yg tidak merata yg dpet org2 terdekat perangkat desa terus, kasian yg benar2 membutuhkan tp yg dpet malah lebih berkecukupan	Negatif
2.	Semoga bantuan merata khususnya saudara di Indonesia Timur yang masih kesulitan beras	Positif
3.	Terbaiklah pakde	Positif
4.	Percuma KTP dki dan indonesia klo ga pernah kedata data nya, dan yg dpt biasa nya yg sudah mampu	Negatif
5.	Bantuan Langsung Tewas	Negatif
6.	Sangat bermanfaat. Jazakumullahu khayr wa barakallahu fiik	Positif
7.	Bius sementara. Dampak ke depannya lebih menyengsarakan daripada bantuan yg sesaat	Negatif
8.	Pak, gak usahlah banyak bantuan pak, harga sembako dan bbm aja diturunin pak. Plis pak, kumohon turunkan harga sembako dan bbm.	Negatif
9.	Semua naikkkkkk , masyarakat menengah ke bawah pada jatuh	Negatif
...
10.	Biar nampak prorakyat,pdahal apa2 udh d naikin	Negatif

3.1 Pre-Processing Data

Dalam proses persiapan data atau *pre-processing*, terdapat lima langkah yang dilakukan, yaitu *cleaning data*, *casefolding*, *normalisasi*, *stemming*, dan *stopword removal*. Output yang dihasilkan dapat dilihat pada Tabel 2.

Tabel 2. Hasil proses persiapan data

Fase Pre-processing	Hasil
Kalimat dasar	Lebih baik turunkan harga drpd bantuan yg tidak merata yg dpet org2 terdekat perangkat desa terus, kasian yg benar2 membutuhkan tp yg dpet malah lebih berkecukupan
<i>Cleaning data</i>	Lebih baik turunkan harga drpd bantuan yg tidak merata yg dpet org terdekat perangkat desa terus kasian yg benar membutuhkan tp yg dpet malah lebih berkecukupan
<i>Case Folding</i>	lebih baik turunkan harga drpd bantuan yg tidak merata yg dpet org terdekat perangkat desa terus kasian yg benar membutuhkan tp yg dpet malah lebih berkecukupan
<i>Tokenizing</i>	'lebih', 'baik', 'turunkan', 'harga', 'drpd', 'bantuan', 'yg', 'tidak', 'merata', 'yg', 'dpet', 'org', 'terdekat', 'perangkat', 'desa', 'terus', 'kasian', 'yg', 'benar', 'membutuhkan', 'tp', 'yg', 'dpet', 'malah', 'lebih', 'berkecukupan'
Normalisasi	'lebih' 'baik' 'turunkan' 'harga' 'daripada' 'bantuan' 'yang' 'tidak' 'merata' 'yang' 'dapat' 'orang' 'terdekat' 'perangkat' 'desa' 'terus' 'kasihan' 'yang' 'benar' 'membutuhkan' 'tetapi' 'yang' 'dapat' 'malah' 'lebih' 'berkecukupan'
<i>Stemming word</i>	'lebih' 'baik' 'turun' 'harga' 'daripada' 'bantu' 'yang' 'tidak' 'rata' 'yang' 'dapat' 'orang' 'dekat' 'perangkat' 'desa' 'terus' 'kasihan' 'yang' 'benar' 'butuh' 'tetapi' 'yang' 'dapat' 'malah' 'lebih' 'cukup'
<i>Stopword Removal</i>	'turun' 'harga' 'bantu' 'orang' 'perangkat' 'desa' 'kasihan' 'butuh'

3.2 Pembobotan TF-IDF

Proses vektorisasi dapat digunakan untuk memberikan nilai numerik pada setiap kata dalam dokumen, sehingga memudahkan dalam proses analisis dan klasifikasi. Salah satu metode yang digunakan dalam proses vektorisasi adalah algoritma TF.IDF, yang memungkinkan untuk mengukur relevansi kata dalam suatu dokumen dengan menganalisis sampel data. Tabel 3 menggambarkan output perhitungan TF.IDF, yang menunjukkan nilai numerik dari setiap kata dalam 3 sampel komentar yang diambil.

Tabel 3. Sampel Data yang di Vektorisasi

Dokumen	TF.IDF		
	D1	D2	D3
bikin	0.233	0	0
cemburu	0.233	0	0
sosial	0.233	0	0
semoga	0	0.233	0
bantuan	0	0.233	0
merata	0	0.233	0
saudara	0	0.233	0
indonesia	0	0.233	0
timur	0	0.233	0
kesulitan	0	0.233	0
beras	0	0.233	0
pengalihan	0	0	0.233
isu	0	0	0.233

Dokumen	TF.IDF		
	D1	D2	D3
Total	0.699	1.864	0.466
y	-1	1	-1

Setiap term diberikan nilai bobot melalui proses pembobotan, dan akan digunakan dalam pembentukan vektor untuk klasifikasi SVM.

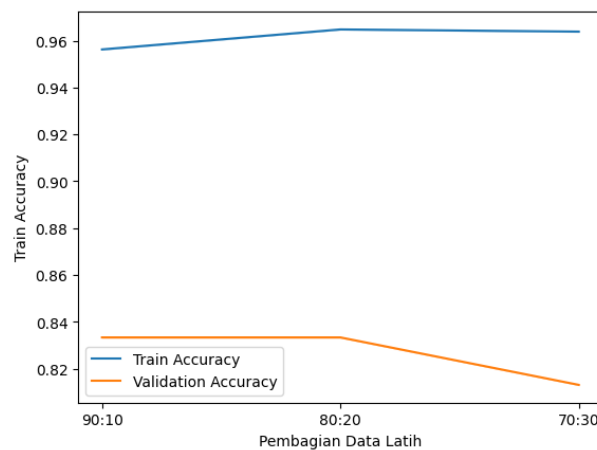
3.3 Validasi dan Evaluasi

Data yang digunakan dibagi menjadi dua bagian, yaitu data yang digunakan untuk melatih model dan data yang digunakan untuk menguji keakuratan model. Pembagian data yang digunakan adalah 70:30, 80:20, dan 90:10. Tabel 4 menunjukkan hasil pengujian perbandingan dari pembagian data yang berbeda.

Tabel 4. Hasil Perbandingan Data Latih dan Data Uji

Splitting Data	SVM			
	Accuracy	Precision	Recall	F-Measure
70 : 30	85.98%	82.25%	66.35%	73.44%
80 : 20	84.72%	84.09%	65.78%	73.81%
90 : 10	80.56%	71.93%	65.17%	68.38%

Dari pengujian diatas, dapat dilihat bahwa model dengan pembagian data 70:30 mendapatkan skor *Accuracy* tertinggi, yaitu sebesar 85.98%. Diikuti oleh model dengan pembagian data 80:20 dengan skor *Accuracy* 84.72%. Model dengan pembagian data 90:10 mendapat skor *Accuracy* terendah, yaitu 80.56%.



Gambar 3. Hasil Validasi Stratified K-Fold

Berdasarkan hasil validasi menggunakan tiga model (90:10, 80:20, dan 70:30), model dengan pembagian data 70:30 memiliki performa yang lebih baik dibandingkan dengan model dengan pembagian data 90:10 dan 80:20.

4. KESIMPULAN

Setelah dilakukan tahapan-tahapan dan pengujian terhadap data sentimen bantuan langsung tunai, didapatkan kesimpulan bahwa metode *Support Vector Machine* dapat digunakan untuk analisis sentimen melalui data tekstual. Hasil yang diperoleh menunjukkan bahwa sentimen negatif lebih banyak daripada sentimen positif, dengan masing-masing persentase 78.61% dan 21.34%. Nilai *accuracy* yang diperoleh yaitu 85.98%, rata-rata nilai *precision* 82.25%, rata-rata nilai *recall* 66.35% rata-rata nilai dan *f-measure* 73.44%. Dari analisis sentimen yang dilakukan, ditemukan bahwa sentimen negatif adalah yang paling banyak muncul, hal ini menunjukkan bahwa masyarakat tidak puas dengan bantuan langsung tunai BBM. Sebagai respon terhadap sentimen negatif yang dominan, perlu diterapkan strategi untuk melakukan pemerataan bantuan langsung tunai dan pendataan yang terstruktur agar tingkat kekecewaan masyarakat dapat diminimalisir. Sebagai saran dan masukan bagi penelitian selanjutnya yaitu untuk dapat meningkatkan kuantitas data penelitian, penambahan kata baru pada kamus, mengganggu singkatan bahasa tidak formal, mengidentifikasi *emoticon* sebagai indikator sentimen, dan menyingkirkan komentar yang tidak berisi pendapat seseorang.

REFERENSI

- [1] Przekota, G. (2022). *Do High Fuel Prices Pose an Obstacle to Economic Growth? A Study for Poland*. *Energies* 2022, 15(18), 6606; <https://doi.org/10.3390/en15186606>.
- [2] Akbar, Ali, Khomaini Direktorat, Jenderal Perbendaharaan, and Alamat Korespondensi. *INDONESIAN TREASURY REVIEW DAMPAK PROGRAM BANTUAN TUNAI (CASH TRANSFER) TERHADAP KESEJAHTERAAN SUBJEKTIF DI INDONESIA*.
- [3] Jaya Irfan, Lubis Imran, & Sihombing Marrto. (2022). *Simulation of the BLT (Direct Cash Assistance) Distribution Queue With the Exponential Method*. *International Journal Of Health, Engineering And Technology (IJHET)*, 1(2), Page. 49-57.
- [4] Novalia Karina, Fenza Putra Ilham, Wulandari Megia, & Zukhri Nizwan. (2022). *Analysis of the Role of the Jada Bahrin Village Government in the Direct Cash Assistance (BLT) Program During the Covid-19 Pandemic*. *INTERNATIONAL JOURNAL OF BUSINESS, TECHNOLOGY AND ORGANIZATIONAL BEHAVIOR (IJTOB)*, 2(6), 700-708. DOI: <https://doi.org/10.52218/ijbtob.v2i6.244>.
- [5] Paramita Djako, Meyko Panigoro, & Sudirman (2022). *Pengaruh Pemberian Bantuan Langsung Tunai (Blt) Terhadap Kesejahteraan Masyarakat Kelurahan Moodu Kecamatan Kota Timur Kota Gorontalo*. *JAMBURA ECONOMIC EDUCATION JOURNAL*, 4(2), 196-206. <https://ejurnal.ung.ac.id/index.php/jej/article/view/15957/5321>.
- [6] Amran. (2022). *IMPLEMENTASI KEBIJAKAN PENYALURAN DANA BANTUAN LANGSUNG TUNAI (BLT) SECARA MERATA KEPADA MASYARAKAT*. *Jurnal Transparansi Publik (JTP)*, 2 (1), 1-9. DOI: <https://doi.org/10.29103/jtp.v2i1.7715>
- [7] Kurniasih, Ulfa, and Akrim Teguh Suseno. 2022. "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) Pada Kenaikan Harga Bahan Bakar Minyak (BBM)." *JURNAL MEDIA INFORMATIKA BUDIDARMA* 6(4): 2335. <https://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/4958>.
- [8] Bourequat, W., & Mourad, H. (2021). *Sentiment analysis approach for analyzing iPhone release using support vector machine*. *International Journal of Advances in Data and Information Systems*, 2(1), 36-44. DOI: <https://doi.org/10.25008/ijadis.v2i1.1216>.
- [9] Prastyo, Pulung Hendro, Amin Siddiq Sumi, Ade Widyatama Dian, and Adhistya Erna Permasari. 2020. "Tweets Responding to the Indonesian Government's Handling of COVID-19: Sentiment Analysis Using SVM with Normalized Poly Kernel." *Journal of Information Systems Engineering and Business Intelligence* 6(2): 112.
- [10] Denny, Y. R., Permata, E., & Assaat, L. D. (2022). *Classification of diseases of banana plant fusarium wilted banana leaf using support vector machine*. *Gravity: Jurnal Ilmiah Penelitian dan Pembelajaran Fisika*, 8(1). DOI: <http://dx.doi.org/10.30870/gravity.v8i1.15893>.
- [11] W. W. Cholil, F. P. (2022). Comparison of Machine Learning Methods in Sentiment Analysis PeduliLindungi Applications. *2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)* (pp. 276-280). Jakarta: IEEE.
- [12] Niu Q, Kandhro Ali Irfan, Kumar A, Shah S, Hasan M, Ahmed M H, & Liang F (2022). *Web Scraping Tool For Newspapers And Images Data Using Jsonify*. *Journal of Applied Science and Engineering*, 26(4), 465-474.
- [13] Nurkholis Andi, Alita D, & Munandar A (2022). Comparison of Kernel Support Vector Machine Multi-Classin PPKM Sentiment Analysis on Twitter. *JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(2), 227-233. DOI: <https://doi.org/10.29207/resti.v6i2.3906>.
- [14] Maharana (2022). *A review: Data pre-processing and data augmentation techniques*. *Global Transitions Proceedings*, 3(1), Pages 91-99. <https://doi.org/10.1016/j.glt.2022.04.020>.
- [15] Yunanda, G., Nurjanah, D., & Meliana, S. (2022). *Recommendation System from Microsoft News Data using TFIDF and Cosine Similarity Methods*. *Building of Informatics, Technology and Science (BITS)*, 4(1), 277-284. DOI: <https://doi.org/10.47065/bits.v4i1.1670>.
- [16] Razia Sulthana A., Pilani, Jaithunbi A. K., R. M. D., Haritha Harikrishnan, & Vijayakumar Varadarajan (2022). *Sentiment Analysis on Movie Reviews Dataset Using Support Vector Machines and Ensemble Learning*. *International Journal of Information Technology and Web Engineering*, 17(1), 1-23. DOI: 10.4018/IJITWE.311428.
- [17] Abdulkareem (2021). *Realizing an Effective COVID-19 Diagnosis System Based on Machine Learning and IoT in Smart Hospital Environment*. *IEEE INTERNET OF THINGS JOURNAL*, 8(21), 15919-15928. DOI: 10.1109/IJOT.2021.3050775.
- [18] Majid Rahardi, Afrig Aminuddin, Ferian Fauzi Abdulloh, & Rizky Adhi Nugroho (2022). *Sentiment Analysis of Covid-19 Vaccination using Support Vector Machine in Indonesia*. (IJACSA) *International Journal of Advanced Computer Science and Applications*, 13(6), 534-539.
- [19] Kochhar, T. S., & Goyal, G. (2022). *Design and Implementation of Stop Words Removal Method for Punjabi Language Using Finite Automata*. In *Advances in Data Computing, Communication and*

- Security (pp. 89-98). Springer, Singapore. DOI: https://doi.org/10.1007/978-981-16-8403-6_8.
- [20] Antonio Dwi V, Efendi S, & Mawengkang H, (2022). Sentiment analysis for covid-19 in Indonesia on Twitter with TF-IDF featured extraction and stochastic gradient descent. *International Journal of Nonlinear Analysis and Applications*, 13(1), 1367-1373.
- [21] Jin Chen, Ruoyi Jia, & Yaxin Yan (2022). *Component Identification and Prediction of Ancient Glass Products Based on Decision Tree Model and SVM Model*. *Highlights in Science, Engineering and Technology*, Vol 22, 142-149. DOI: <https://doi.org/10.54097/hset.v22i.3304>.
- [22] Desi Musfiroh, Ulfa Khaira, Pradita Eko Prasetyo Utomo, dan Tri Suratno (2021). *Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon*. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 24-33. Journal Homepage: <https://journal.irpi.or.id/index.php/malcom>.