



## **Analisis sentimen kenaikan harga kebutuhan pokok dimedia sosial youtube menggunakan algoritma *support vector machine***

**Rudy Asrianto\*<sup>1</sup>, Melda Herwinanda<sup>2</sup>**

Email: <sup>1</sup>rudyasrianto@umri.ac.id, <sup>1</sup>180402038@student.umri.ac.id

<sup>1,2</sup>Sistem Informasi, Fakultas Ilmu Komputer, Universitas Muhammadiyah Riau

Diterima: 3 Desember 2022 | Direvisi: 20 Desember 2022 | Disetujui: 21 Desember 2022

©2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,  
Universitas Muhammadiyah Riau, Indonesia

### **Abstrak**

Kebebasan setiap orang dalam mengemukakan pendapat, tercurahkan dalam suatu wadah yang sekarang ini disebut sebagai media sosial. Youtube merupakan media sosial berbasis *web streaming video* atau situs untuk mengunggah dan menonton video. Memasuki pekan kedua bulan Januari tahun 2022, terjadi kenaikan harga bahan kebutuhan pokok diumumkan di beberapa media berita online. Didalam forum diskusi channel media sosial Youtube telah dibahas mengenai survei kepuasan masyarakat, terhadap kepemimpinan presiden Jokowi-Ma`ruf Amin yang memperoleh nilai tingkat kepuasan 66-73 %. Diluar sana masyarakat umum merasa tidak puas karena kenaikan harga, dan mempertanyakan hasil survei tersebut dan kepada siapa survei tersebut dilakukan. Konten video ini mendapatkan jumlah viewers atau penonton sebanyak ratusan ribu hingga jutaan lebih, dan dibanjiri hingga puluhan ribu komentar oleh netizen. Penelitian ini menggunakan algoritma *Support Vector Machine* dalam mengklasifikasikan label komentar positif dan label komentar negatif, serta menggunakan teknik penyeimbangan jumlah label pada data yaitu SMOTE. Dari pengujian keempat model klasifikasi SVM, didapatkan SVM linear dengan SMOTE mendapatkan tingkat keakuratan terbaik dengan nilai akurasi tertinggi, yaitu 86.33%, diikuti dengan nilai presisi 75%, nilai recall 66.67%, dan nilai f1-score 70.59%. Pengujian juga dilakukan pada data validasi (data uji baru diluar dari dataset), didapatkan nilai akurasi 72%.

**Kata kunci:** Analisis Sentimen, *Text Mining*, Klasifikasi, *Support Vector Machine*, *Confusion Matrix*.

### ***Sentiment analysis of rising prices for basic necessities on YouTube social media using the support vector machine algorithm***

#### **Abstract**

*Everyone's freedom to express opinions, is poured out in a container which is now known as social media. Youtube is a web-based social media streaming video or site for uploading and watching videos. Entering the second week of January 2022, an increase in the price of basic commodities was announced in several online news media. In the discussion forum on the Youtube social media channel, a survey of people's satisfaction has been discussed regarding the leadership of President Jokowi -Ma'ruf Amin who received a satisfaction score of 66-73%. Out there the general public is dissatisfied with the price increase, and is questioning the results of the survey and to whom the survey was conducted. This video content gets hundreds of thousands to millions of viewers or viewers, and is flooded with tens of thousands of comments by netizens. This study uses the Support Vector Machine algorithm in classifying positive comment labels and negative comment labels, and uses a balancing technique for the number of labels on the data, namely SMOTE. From the test of the four SVM classifications, it was found that linear SVM with SMOTE got the best level of accuracy with the highest accuracy value, namely 86.33%, followed by a precision value of 66.67%, a recall value of 75%, and an f1-score value of 70.59%. Tests were also carried out on data validation (new test data outside of the dataset), the accuracy value was 72%.*

**Keywords:** *Sentiment Analysis, Text Mining, Classification, Support Vector Machine, Confusion Matrix.*

## 1. PENDAHULUAN

Kebebasan setiap orang dalam mengemukakan pendapat, tercurahkan dalam suatu wadah yang sekarang ini disebut sebagai media sosial. Media sosial adalah sebuah *platform* yang membutuhkan koneksi internet, berfungsi menghubungkan pengguna komputer untuk dapat saling berkomunikasi dengan pengguna komputer lain dalam dunia digital. Salah satu media sosial dengan pengguna terbanyak di dunia yaitu Youtube. Youtube merupakan media sosial berbasis *web streaming video* atau situs untuk mengunggah dan menonton video [1]. Memasuki pekan kedua bulan Januari tahun 2022, terjadi kenaikan harga bahan kebutuhan pokok diumumkan diberbagai media berita online. Lembaga survei menyatakan bahwa telah melakukan survei kepuasan masyarakat, terhadap kepemimpinan presiden Jokowi-Ma`ruf Amin yang memperoleh nilai tingkat kepuasan 66-73%. Diluar sana masyarakat umum merasa tidak puas karena kenaikan harga, dan mempertanyakan hasil survei tersebut dan kepada siapa survei tersebut dilakukan. Ada kecurigaan masyarakat bahwa lembaga survei dibayar oleh pemerintah untuk menggiring opini public[2].

Permasalahan survei ini menjadi bahan diskusi oleh beberapa narasumber seperti tokoh partai politik, pengamat politik, pakar ekonomi, hingga *influencer*, pada sebuah konten video di channel media sosial Youtube nasional seperti TVOneNews, KompasTV, dan Detikcom. Konten video ini mendapatkan jumlah *viewers* atau penonton sebanyak ratusan ribu hingga jutaan lebih, dan dibanjiri hingga puluhan ribu komentar oleh netizen dalam merespon konten pada video tersebut. Perbedaan pendapat terjadi antara pihak pro pemerintah dengan pihak oposisi (pihak yang berlawanan), membahas mengenai survei yang dinilai bahwa masyarakat merasa puas dengan pemerintahan sekarang ini. Dalam survei harus memiliki indikator yang jelas, bukan hanya pertanyaan umum saja puas atau tidak. Hasil survei dimanfaatkan oleh partai untuk pencalonan pejabat publik, presiden atau gubernur. Hal-hal sensitive tidak bisa dilakukan secara kuantitatif, harus jelas untuk apa, tujuannya apa dan siapa yang membayar [3].

Untuk mengetahui penilaian setiap sentimen netizen dari banyaknya jumlah komentar yang ada, dapat dikategorikan menjadi komentar positif dan komentar negatif [4]. Opini masyarakat yang telah dibangun dalam bentuk komentar pada Youtube, dapat dijadikan sebagai data untuk diolah. Pantauan dari penulis terhadap komentar pada konten video yang sudah disebutkan pada paragraf sebelumnya, hingga kini masih terus bertambah. Dengan mengkategorikan banyaknya jumlah komentar yang ada, membutuhkan waktu yang cukup lama dalam menganalisis sentimen pada komentar netizen tersebut. Analisis sentimen merupakan proses pengenalan dan pemahaman teks berupa opini terhadap subyek permasalahan, dalam mendefinisikan kategori penyampaian kata-kata tersebut ke arah positif atau ke arah negatif [5]. Analisis sentimen dapat dilakukan dengan pendekatan *Supervised Machine Learning*, yaitu teknik pembelajaran mesin yang digunakan untuk mengetahui hubungan antara atribut *input* dengan atribut *target*. Data komentar netizen Youtube dapat dijadikan sebagai *input* pada *Machine Learning* guna mempelajari pola kata dari komentar positif dan komentar negatif. Beberapa penerapan algoritma *Machine Learning* untuk klasifikasi teks diantaranya *Logistic Regression*, *Bernoulli Naive Bayes*, *Support Vector Machine*, dan *Random Forest Classifier* [6].

Penelitian analisis sentimen oleh [3] dengan studi kasus tanggapan masyarakat terhadap pemberlakuan Pembatasan Sosial Berskala Besar di Twitter, menggunakan algoritma *Support Vector Machine* dan teknik *Web Scraping* pada pengumpulan datanya. Lalu dengan pengujian *K-Fold Cross Validation* menghasilkan keakuratan tertinggi dengan nilai 82.07% pada *fold* 9. *Support Vector Machine* merupakan salah satu algoritma *supervised machine learning* dengan akurasi terbaik dalam bidang klasifikasi *text* yang digunakan dalam berbagai macam kasus untuk sentiment klasifikasi. Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan pada *svm* sangat bergantung pada fungsi *kernel* dan *parameter* yang digunakan [5].

Penelitian lain [4] juga membuktikan bahwa algoritma *Support Vector Machine* menghasilkan nilai akurasi terbaik sebesar 96.68% dari algoritma *Machine Learning* lainnya (BM25 + KNN dan Naive Bayes), yang telah diujikan pada analisis sentimen *tweet positif* dan *tweet negatif*, oleh netizen media sosial Twitter terkait pemindahan Ibu Kota Indonesia. Kemudian [1] melakukan analisis sentimen dengan tema keluhan program vaksinasi Covid-19 di Indonesia, pada komentar di media sosial Youtube, menggunakan *Support Vector Machine* dan fitur ekstraksi pembobotan kata TF- IDF, menghasilkan tingkat keakuratan menghasilkan tingkat akurasi rata-rata 76.33%.

Dari uraian penjelasan latar belakang dan beberapa referensi penelitian terdahulu yang telah disebutkan diatas, penulis mengusulkan analisis sentimen berupa komentar masyarakat pada konten video di media sosial Youtube. Algoritma yang digunakan dalam penelitian ini menggunakan *Support Vector Machine*. Pengklasifikasian komentar terbagi menjadi dua label atau kelas, yaitu komentar positif dan komentar negatif. Saat pemberian label pada komentar, penulis tidak memperhitungkan kesamaan jumlah data yang diperlukan untuk komentar positif dan komentar negatif, kemudian penulis menemukan ketidakseimbangan jumlah kelas pada data. Oleh karena itu penulis juga mengusulkan menggunakan teknik penyeimbangan data SMOTE (*Synthetic Minority Oversampling Technique*), yang merujuk pada referensi [7] dan [8] untuk optimalisasi perbandingan jumlah label pada data sebelum dilakukan tahap klasifikasi. Penelitian ini menguji algoritma *Support Vector Machine* dalam mengetahui hasil klasifikasi sentimen komentar pengguna, dan mendapatkan hasil akurasi terbaik, pada konten video di media sosial Youtube bertajuk kaitan kenaikan harga kebutuhan pokok dengan survei kinerja pemerintahan di Indonesia [9].

## 2. METODE PENELITIAN

Penelitian analisis sentimen berupa komentar pada media sosial Youtube menggunakan algoritma *Support Vector Machine*, dilakukan dengan alur kerja penelitian yang terstruktur secara garis besar. Langkah-langkah tersebut dimulai dari identifikasi masalah dan solusi, pengambilan data, pemrosesan data, pengklasifikasian SVM, pengujian, evaluasi, dan analisis sentimen. Tahapan penelitian yang diusulkan dapat dilihat pada gambar 1. dibawah ini.

### 2.1 Identifikasi Masalah

Diluar sana masyarakat umum merasa tidak puas karena kenaikan harga, dan mempertanyakan hasil survei terhadap kepemimpinan presiden Jokowi-Ma'ruf Amin yang memperoleh nilai tingkat kepuasan 66-73 %. Permasalahan survei ini menjadi bahan diskusi oleh beberapa narasumber seperti tokoh partai politik, pengamat politik, pakar ekonomi, hingga *influencer*, pada sebuah konten video di channel media sosial Youtube nasional seperti TVOneNews, KompasTV, dan Detikcom. Penulis melakukan studi literatur mengenai pengkategorian sebuah komentar menggunakan *Text Mining*, dan *Support Vector Machine*. Penulis mengusulkan bahwa diperlukannya analisis sentimen dengan pendekatan *Machine Learning*, untuk mengetahui klasifikasi sentimen komentar positif, dan sentimen komentar negatif pada komentar di media sosial Youtube terhadap sebuah konten video yang membahas masalah naiknya harga kebutuhan pokok disaat lembaga survei mengumumkan tingkat kepuasan kinerja pemerintahan Indonesia saat ini.

## **2.2 Pengambilan Data**

### **2.2.1 Penelusuran Data**

Data komentar diambil pada tiga *link* atau alamat situs konten video di channel media sosial nasional Youtube, dan menggunakan izin layanan dari Youtube Data API v3, untuk mengakses atribut data statistik dari video tersebut seperti, nama pengomentor, isi komentar, tanggal komentar, jumlah like, dan jumlah balasan komentar. Jumlah data komentar dari seluruh *link* yang telah didapat cukup banyak yaitu 5088 baris data yang belum memiliki label. Data tersebut didapatkan dengan teknik *web scraping*. *Web Scraping* atau *Scraping Technique* sebuah teknik yang digunakan untuk mengambil atau mengekstrak sebuah data atau informasi tertentu dari suatu website secara spesifik [10]. *Web Scraping* merupakan proses pengambilan data semi terstruktur dari sebuah web. Umumnya, dalam *web scraping* ada dua cara untuk proses pengambilan datanya, yaitu dengan cara manual dan otomatis [11].

### **2.2.2 Pelabelan Data**

Penulis menggunakan sejumlah 734 baris data komentar yang diketikkan oleh netizen Youtube, dengan jarak waktu yang beragam antara bulan Februari 2022 hingga bulan Mei 2022. Penulis melakukan pelabelan secara manual dengan menambahkan kolom baru yaitu *category*, sesuai kemampuan dan pemahaman dalam mengkategorikan sentimen pada komentar. Dua jenis label yang digunakan yaitu label bernilai 0 untuk sentimen komentar negatif, dan label bernilai 1 untuk sentimen komentar positif. [12] Penulis memvalidasikan data komentar ini agar dapat dilanjutkan ke penelitian analisis sentimen, dengan mendatangi seorang ahli bahasa yang berkompeten pada bidang komunikasi di media sosial.

## **2.3 Pemrosesan Data**

### **2.3.1 Text Preprocessing**

*Text preprocessing* berfungsi mengubah data yang tidak terstruktur menjadi terstruktur. *Text preprocessing* adalah pembersihan, dan penyederhanaan teks agar dapat diproses lebih lanjut [13]. Data komentar yang telah diperoleh merupakan *raw data* atau data mentah, yang perlu dilakukan tahapan *text preprocessing* terlebih dahulu. Tahapan ini meliputi *case folding*, *cleaning*, *tokenization*, *stopword*, dan *stemming*. Tahapan *case folding* merubah semua karakter yang terdeteksi terdapat huruf kapital akan menjadi huruf kecil. Tahapan *cleaning* menemukan dan menghapus objek karakter dalam teks yang tidak diperlukan seperti angka, simbol dan tanda baca, hashtag #, link URL, emoticon atau ekspresi wajah, dan karakter lainnya. Tahapan *tokenization* bertugas memisahkan setiap kata dalam kalimat untuk masing-masing baris data komentar, kemudian menyatukannya kembali untuk diproses pada tahap *preprocessing* teks berikutnya. Tahapan *normalization* merubah menggunakan bahasa indonesia yang tidak baku menjadi kata baku dan salah pengetikkan kata. Tahapan *stopword* menemukan dan menghapus kata-kata yang tidak memiliki makna penting dalam sebuah kalimat, seperti kata penghubung, kata keterangan, kata ganti orang atau benda, dan kata depan. Tahapan *stemming* menghilangkan imbuhan pada kata yang memiliki awalan, akhiran, serta gabungan awalan dan akhiran.

### **2.3.2 Fitur Ekstraksi**

Fitur ekstraksi yang digunakan dalam penelitian ini adalah pembobotan kata TF-IDF (*Term Frequency-Inverse Document Frequency*). TF (*Term Frequency*) adalah frekuensi kemunculan *term* pada dataset dan IDF adalah sebuah pembobotan bagaimana *term* itu didistribusikan secara luas pada seluruh dataset yang bersangkutan [14]. Bobot setiap kata menjadi informasi penting pada masing-masing baris data komentar yang berkaitan langsung dengan labelnya, kemudian menjadi bahan belajar saat proses *training* untuk algoritma *Support Vector Machine*.

### **2.3.3 Balancing Data**

Dilakukan pengecekan jumlah label komentar positif dan komentar negatif pada kolom *category* yang terdapat pada data latih. Jika label komentar positif lebih banyak dari label komentar negatif, atau sebaliknya maka terjadi ketidakseimbangan jumlah label pada data komentar tersebut. Metode penyeimbangan jumlah label yaitu SMOTE (*Synthetic Minority Oversampling Technique*). SMOTE merupakan teknik yang dimanfaatkan untuk mengatasi masalah *class imbalance problem* (CIP). SMOTE bekerja dengan memodifikasi dataset yang tidak seimbang dengan cara membuat data sintetik baru dari kelas minoritas dengan tujuan meningkatkan kinerja dari metode klasifikasi [7].

## **2.4 Klasifikasi Sentimen dengan SVM**

Alur proses klasifikasi secara keseluruhan terdapat pada penelitian analisis sentimen ini diuraikan sebagai berikut.

1. Mempersiapkan *raw data* komentar dengan nama file *raw\_data\_sentimen.csv*.
2. Melakukan tahap *text preprocessing* pada *raw data* komentar dan menghasilkan *clean dataset*.
3. Membagi dataset menjadi dua yaitu data latih dengan proporsi 90% dan data uji dengan proporsi 10%.

- Memberikan nilai bobot untuk setiap kata dengan fitur ekstraksi TF-IDF pada data *input* (isi komentar), kemudian mentransformasikan data setiap kata dalam komentar pada data latih dan data uji menjadi vektor-vektor bobot kata.
- Melakukan penyeimbangan jumlah label pada data latih dengan SMOTE.
- Membuat model klasifikasi SVM dengan konfigurasi parameter tuning kernel linear, rbf gaussian, polynomial, dan sigmoid, dan menggunakan data latih sebagai bahan *input* untuk pelatihan model.
- Menggunakan data uji sebagai *input* untuk pengujian model klasifikasi SVM dalam mengkategorikan label sentimen.
- Mengukur performa model klasifikasi SVM dengan evaluasi *confusion matrix*.
- Mengklasifikasi data komentar baru dengan model klasifikasi SVM yang memiliki tingkat akurasi terbaik.

### 2.5 Pengujian

Skenario pengujian sentimen pada data uji dilakukan pada model klasifikasi SVM dengan SMOTE dan SVM tanpa SMOTE, yang sudah melalui proses *training* pada data latih, menggunakan keempat konfigurasi *parameter kernel* yang berbeda. Peran masing-masing model menentukan target label sentimen komentar positif atau sentimen komentar negatif, dari *input* setiap baris data komentar pada data uji.

Tabel 2. Skenario Pengujian Klasifikasi

Klasifikasi	SVM dengan SMOTE	SVM tanpa SMOTE
	<i>Linear</i>	<i>Linear</i>
	<i>Polynomial</i>	<i>Polynomial</i>
Kernel	<i>RF Gaussian</i>	<i>RF Gaussian</i>
	<i>Sigmoid</i>	<i>Sigmoid</i>

### 2.6 Evaluasi

Pengukuran performa hasil pengujian dari klasifikasi SVM terhadap analisis sentimen komentar pada data uji, menggunakan penghitungan evaluasi *confusion matrix*. *Confusion matrix* menampilkan tabel berdimensi 2x2 dikarenakan terdapat 2 jenis label yang diklasifikasikan yaitu komentar positif dan komentar negatif. Evaluasi ini menampilkan banyaknya jumlah baris data komentar pada perbandingan label antara data aktual dengan data hasil prediksinya. Kemudian menentukan dan menghitung terlebih dahulu jumlah kondisi *true positive*, *false positive*, *false negative*, dan *true negative* disetiap baris data komentar yang telah diujikan. Keempat kondisi tersebut dijelaskan sebagai berikut [15][16]:

- True Positive* (TP) adalah kondisi dimana model data yang terprediksi dengan benar (*True*) dan jawaban aktual juga benar (*True*).
- False Positive* (FP) adalah kondisi dimana model data yang terprediksi benar (*True*). Namun jawaban aktual adalah salah (*False*).
- False Negative* (FN) adalah kondisi dimana model data yang terprediksi salah (*False*). Namun jawaban aktual adalah benar (*True*).
- True Negative* (TN) adalah kondisi dimana model data yang terprediksi salah (*False*) dan jawaban aktual juga salah (*False*).

Setelah itu dapat dilakukan penghitungan nilai akurasi, presisi, *recall*, dan *f1-score*, untuk melihat performa model klasifikasi SVM dengan SMOTE dan model klasifikasi SVM tanpa SMOTE.

### 2.7 Analisis Sentimen

Tahapan ini melakukan klasifikasi sentimen pada data validasi atau data komentar baru diluar dari *dataset* yang labelnya sudah divalidasi oleh ahli bahasa. Data komentar baru berjumlah 50 baris data akan dikategorikan menggunakan model klasifikasi SVM, yang memperoleh nilai akurasi terbaik dari hasil pengujian sebelumnya. Hasil akhir dari klasifikasi pada data komentar baru yaitu penghitungan persentase jumlah sentimen yang mengarah ke komentar negatif, persentase jumlah sentimen yang mengarah ke komentar positif, dan persentase kesalahan klasifikasi atau *error classification* untuk komentar yang mendapatkan hasil sentimen tidak sesuai kategori pelabelan. Kemudian memberikan kesimpulan secara keseluruhan dari hasil analisis sentimen pada data komentar tersebut.

## 3. HASIL DAN PEMBAHASAN

Data komentar yang telah dipersiapkan diberi nama *file raw\_data\_sentimen.csv* dan *new\_data\_sentimen.csv*, dalam format Microsoft Excel yang sudah tervalidasi. File diunggah ke media penyimpanan Google Drive, tersimpan di dalam *folder* baru yang telah dibuat dan diberi nama Dataset.

ID	Name	Comment	Time	Likes	Reply Count	Category
0	Yani Mulyana	Puas hutang membengkak ekonomi pakir	2022-05-23T15:15:47Z	0	0	0
1	Henry Suryani	Surenyya Cita ke Palisok, ke daerah2 Rakyat menengki kebawah... AMM	2022-04-24T19:29:10Z	0	0	1
2	Muhammad Rudi	Burhan Burhan...	2022-04-23T04:30:10Z	0	0	1
3	zul ishak anagar	Saya sudah puas kok di kepenjamban pakir jawa jadi dengan hormat jgn di perpanjang lagi yea...ken saya sudah puas.	2022-04-20T16:52:47Z	0	0	0
4	Suprpto Malang	Ki aja uda build yg comen merasa tdk puas ni aj...sune apaan d	2022-04-19T16:05:13Z	0	0	0
5	Wawan Gunawan	Pasti bu indibayar mahal.	2022-04-19T14:39:57Z	0	0	0

Gambar 2. Raw Data Komentar

### 3.2 Preprocessing Data

Proses *preprocessing* teks data komentar yang terdapat pada kolom *comment* sebanyak 734 baris data, dilakukan dengan susunan kode-kode python yang terlebih dahulu harus memasang library Sastrawi, kemudian mengimpor library *re* (*regular expression*), dan membaca *file new\_kamusalay.csv* serta *file stopword\_indonesia.csv*. Lalu dibuatkan fungsi-fungsi untuk masing-masing tahap preprocessing tersebut mulai dari *case folding* hingga *stemming*. Hasil dari tahapan *preprocessing* ini diperoleh dataset komentar yang telah *clean* dan disimpan dengan nama *file clean\_dataset\_sentimen.csv*, kemudian diletakkan pada media penyimpanan Google Drive di dalam folder Dataset.

	Comment	Category
0	puas hutang bengkak ekonomi pahit	0
1	survei coba pelosok daerah rakyat tengah	1
2	burhan burhan	1
3	puas pimpin jokowi hormat panjang puas	0
4	bukti komen puas survei	0
...	...	...
729	kriminalisasi demokrasi hidup kuasa kekal cata...	1
730	kriminalisasi demokrasi nafsu kuasa	1
731	orang indonesia suka diskon promo	1
732	responden responden dewan wakil rakyat buzzerr...	0
733	jokowi	1

734 rows × 2 columns

Gambar 3. Dataset Komentar

### 3.3 Pembagian Dataset

Untuk jumlah data latih ditentukan sebanyak 90% dari dataset dan sisanya 10% ditentukan untuk jumlah data uji, sehingga dari 721 baris data pada dataset komentar, telah terbagi 648 data latih dan 73 data uji. Proses pada saat pembagian dataset ini dilakukan secara random atau acak. Masing-masing data latih dan data uji berisikan baris data pada kolom *comment* yang diberi tanda “x”, dan berhubungan langsung dengan kolom *category* yang diberi tanda “y”.

<p>List Data Latih X:</p> <pre> 608      jokowi tambah wkwk rakyat kebal 155      survei otak atik 159      that bullshit 225      tertawa takut dosa 659      oposisi parlemen oposisi amen ..... 709      sewa survei dana survei 439      puas mundur rakyat miskin gara orang urus nega... 174      hasil survei sumber 122      survei baca erti tuju gada survei Name: Comment, dtype: object  Label Data Latih Y: [0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 0 0 0 0  1 0 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 1 1 0 0 1 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0  0 0 1 1 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0  0 1 0 1 0 1 0 0 1 0 0 0 1 0 0 0 0 1 1 0 1 1 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1  0 1 0 1 0 0 0 0 0 1 0 0 0 0 0 1 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0  1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 0 1 0 1  0 1 0 0 1 0 1 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 1 1  0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1  1 0 0 0 0 0 0 1 1 0 0 0 1 1 0 0 1 1 0 0 0 0 0 0 1 0 1 1 1 0 1 1 0 1  0 0 0 0 1 0 0 1 1 0 0 0 1 1 0 0 0 1 1 0 0 0 0 0 0 1 0 1 1 1 0 0 1 0  1 0 0 0 1 0 0 1 0 1 0 1 0 1 0 1 1 0 0 0 1 1 0 0 1 0 0 1 1 0 1 0 1 1 0  0 1 1 0 0 0 1 0 0 1 0 0 0 1 1 1 0 1 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 1  1 0 0 0 0 1 0 1 0 1 0 0 0 1 1 0 0 0 1 0 0 0 0 1 1 1 0 0 1 1 0 1 1 1  0 1 0 1 1 1 1 0 0 0 1 0 0 0 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 0 0 1 1  0 0 0 1 1 1 0 1 1 0 1 1 0 0 0 0 1 1 0 1 1 1 0 0 1 0 0 0 1 0 0 1 0 1 1 0  0 0 0 0 1 0 0 0 1 0 0 1 1 1 1 0 0 0 1 0 1 0 1 1 1 0 0 0 0 1 1 0 0  1 0 0 0 0 1 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1  0 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1  0 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0]</pre> <p>Jumlah Data Latih X: 648</p>	<p>List Data Uji X:</p> <pre> 408      puas minyak goreng mahal harga kedelai impor m... 124      lembaga survei bayar 285      stasiun televisi netral pihak sepet televisi 385      suruh survei negara atur kacau balau segi ekon... 476      surpay citra ..... 419      puas partai adil sejahtera keluarga kadrurnisty 539      puas coba kalang warung kopi coba undang undan... 697      survei nyesar rakyat lelah kibal hasil survei ... 184      metoda survei survei mana kacau balau puas 620      survei abal abal mayoritas rakyat tengah bilan... Name: Comment, dtype: object  Label Data Uji Y: [1 0 1 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 1 1 0 0  0 1 0 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]</pre> <p>Jumlah Data Uji X: 73</p>
--	---

Gambar 5. Data Uji

### 3.4 Pembobotan Kata TF-IDF

Tahap pembobotan kata TF-IDF akan dijadikan sebagai *learning feature* atau model pembelajaran bagi *Support Vector Machine* untuk analisis sentimen, dalam mengenali karakter setiap kata yang telah diberikan nilai bobot. Proses pembobotan ini awalnya diterapkan pada data latih sebagai acuannya, yang kemudian akan ditransformasikan juga ke data uji. Pertama-tama dilakukan tokenisasi untuk membentuk *dictionary* atau kamus kata pada data latih. Didapatkan sebanyak 1419 kata-kata unik (token) yang ditemukan, dan telah disusun berdasarkan abjad awal hingga akhir.

Pembobotan kata TF-IDF dimulai dari menghitung kemunculan jumlah term atau kata (TF), lalu menghitung nilai IDF kata tersebut, dan mengalikan nilai TF setiap kata dengan nilai IDFnya, yang menghasilkan nilai bobot setiap kata untuk setiap baris data komentar pada data latih dan data uji.

<b>List Pembobotan Data Latih:</b>	<b>List Pembobotan Data Uji:</b>
(0, 1404) 0.4578813617088546	(0, 1035) 0.13443309184324323
(0, 1269) 0.5030977587922838	(0, 800) 0.24783135927469618
(0, 1064) 0.2595420178207289	(0, 719) 0.6021264104701354
(0, 575) 0.6003428600049602	(0, 579) 0.35387756200161546
(0, 529) 0.33085298113276107	(0, 469) 0.4091545850982383
(1, 1250) 0.2770189620348741	(0, 412) 0.27024665452576996
(1, 913) 0.5982280738251716	(0, 383) 0.27334102241571157
(1, 64) 0.7519199866744771	(0, 270) 0.34287865897299047
(646, 1250) 0.3388413467544464	(1, 1250) 0.36068166513817
(646, 1235) 0.7707465034835681	(1, 684) 0.6395920893027294
(646, 419) 0.5395705413544064	(1, 107) 0.6788451191063588
(647, 1335) 0.3932390806899685	(71, 1250) 0.5297163217508536
(647, 1250) 0.37231805249949596	(71, 1035) 0.2362078110837505
(647, 351) 0.5052964299224484	(71, 739) 0.5483137113375262
(647, 339) 0.4750883134264193	(71, 545) 0.6024604239392305
(647, 71) 0.4750883134264193	(72, 1291) 0.3876103960175339
Shape Data Latih X Setelah di TFIDF-kan: (648, 1419)	(72, 1250) 0.14280166970104619
	(72, 1064) 0.16757288378486151
	(72, 1018) 0.23678718882671745
	(72, 760) 0.36443789903439333
	(72, 448) 0.3016517568097666
	(72, 151) 0.2724578632072854
	(72, 27) 0.3876103960175339
	(72, 0) 0.5449157264145708
	Shape Data Uji X Setelah di TFIDF-kan: (73, 1419)

Gambar 7. Hasil Pembobotan Kata TF-IDF

Hasil pembobotan kata pada data latih diindeks baris pertama, yang bertuliskan komentar dengan kalimat “jokowi tambah wkwk rakyat kebal”, yaitu:

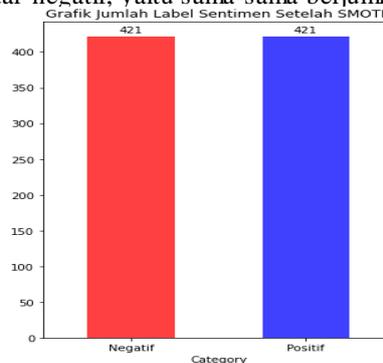
(0, 1404)	0.4578813617088546,	adalah nilai bobot untuk kata wkwk
(0, 1269)	0.5030977587922838,	adalah nilai bobot untuk kata tambah
(0, 1064)	0.2595420178207289,	adalah nilai bobot untuk kata rakyat
(0, 575)	0.6003428600049602,	adalah nilai bobot untuk kata kebal
(0, 529)	0.33085298113276107,	adalah nilai bobot untuk kata jokowi,

Kemudian hasil pembobotan kata pada data uji diindeks baris pertama, yang bertuliskan komentar dengan kalimat “puas minyak goreng mahal harga kedelai impor mahal demo”, yaitu:

(0, 1035)	0.13443309184324323,	adalah nilai bobot untuk kata puas
(0, 800)	0.24783135927469618,	adalah nilai bobot untuk kata minyak
(0, 719)	0.6021264104701354,	adalah nilai bobot untuk kata mahal
(0, 579)	0.35387756200161546,	adalah nilai bobot untuk kata kedelai
(0, 469)	0.4091545850982383,	adalah nilai bobot untuk kata impor
(0, 412)	0.27024665452576996,	adalah nilai bobot untuk kata harga
(0, 383)	0.27334102241571157,	adalah nilai bobot untuk kata goreng
(0, 270)	0.34287865897299047,	adalah nilai bobot untuk kata demo.

### 3.5 Penyeimbangan Jumlah Data Label

Jumlah label komentar negatif lebih banyak dari label komentar positif. Komentar negatif berjumlah 421 baris data, sementara itu komentar positif berjumlah 227 baris data. Selisih diantara keduanya terhitung sebanyak 194 baris data. Setelah itu dilakukan pembuatan data sintetik atau data duplikat, menggunakan teknik SMOTE (*Synthetic Minority Oversampling Technique*), dari data komentar yang memiliki jumlah label paling sedikit, yaitu pada data komentar positif. Dari hasil penyeimbangan data pada label ini, jumlah data latih bertambah menjadi 842 baris data. Hal tersebut terjadi dikarenakan jumlah label komentar positif sudah seimbang dengan jumlah label komentar negatif, yaitu sama-sama berjumlah 421 baris data.



Gambar 8. Jumlah Label Sentimen Sesudah SMOTE

### 3.6 Klasifikasi SVM (*Support Vector Machine*)

Model klasifikasi SVM yang diimplementasikan untuk analisis sentimen komentar positif dan komentar negatif, dibentuk menggunakan empat kernel antara lain linear, polynomial, rbf gaussian, dan sigmoid. Model klasifikasi SVM terkonfigurasi untuk dua skenario pengujian yaitu menggunakan SMOTE dan tanpa SMOTE.

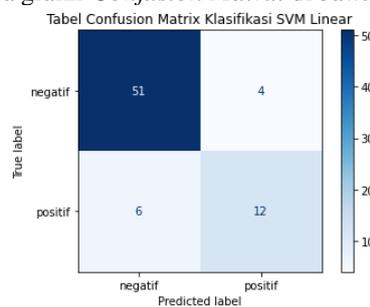
Klasifikasi SVM dengan SMOTE Kernel: linear    Klasifikasi SVM tanpa SMOTE Kernel: linear  
 Klasifikasi SVM dengan SMOTE Kernel: poly    Klasifikasi SVM tanpa SMOTE Kernel: poly  
 Klasifikasi SVM dengan SMOTE Kernel: rbf    Klasifikasi SVM tanpa SMOTE Kernel: rbf  
 Klasifikasi SVM dengan SMOTE Kernel: sigmoid    Klasifikasi SVM tanpa SMOTE Kernel: sigmoid  
 Gambar 9. Konfigurasi Kernel SVM dengan dan tanpa SMOTE

### 3.7 Pengujian Klasifikasi Sentimen

Skenario pengujian pertama pada keempat model klasifikasi SVM dengan SMOTE, dengan data latih berjumlah 842 yang sudah melalui tahap *training*, kemudian dilakukan pengujian menggunakan data uji yang berjumlah 73 baris data komentar sebagai *input*.

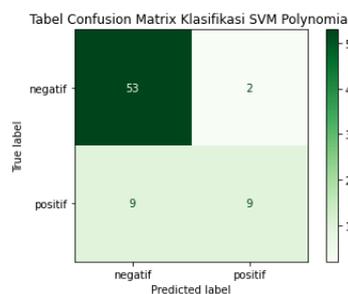
Skenario pengujian kedua pada keempat model klasifikasi SVM tanpa penyeimbangan jumlah label pada data latih. Data latih yang digunakan tetap berjumlah 648 yang sudah melalui tahap *training*. Pengujian dilakukan pada data uji yang berjumlah 73 baris data sebagai *input*.

Jika label hasil klasifikasi SVM dengan SMOTE ini sama dengan label aktual pada data uji, maka dinyatakan benar atau sesuai. Sebaliknya jika tidak sama, maka dinyatakan salah atau tidak sesuai. Adapun hasil pengujian pada keempat model klasifikasi SVM dengan SMOTE, yang ditampilkan pada grafik *Confusion Matrix* di bawah ini.



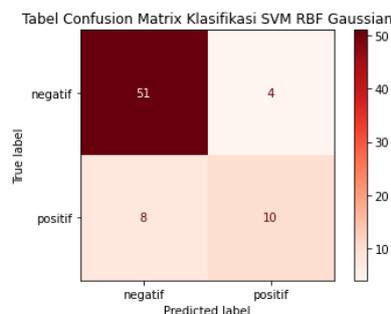
Gambar 10. Confusion Matrix SVM Linear dengan SMOTE

Pada gambar 10 di atas. hasil pengujian klasifikasi SVM Linear dengan SMOTE, didapatkan 12 baris data dengan kondisi *true positive*, 4 baris data dengan kondisi *false positive*, 6 baris data dengan kondisi *false negative*, dan 51 baris data dengan kondisi *true negative*.



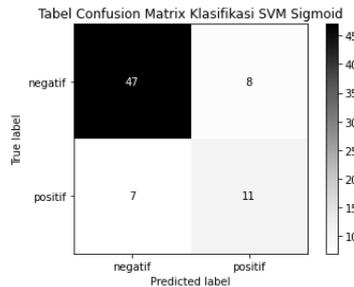
Gambar 11. Confusion Matrix SVM Polynomial dengan SMOTE

Pada gambar 11 di atas. hasil pengujian klasifikasi SVM Polynomial dengan SMOTE, didapatkan didapatkan 9 baris data dengan kondisi *true positive*, 2 baris data dengan kondisi *false positive*, 9 baris data dengan kondisi *false negative*, dan 53 baris data dengan kondisi *true negative*.



Gambar 12. Confusion Matrix SVM RBF Gaussian dengan SMOTE

Pada gambar 12 di atas. hasil pengujian klasifikasi SVM RBF Gaussian dengan SMOTE, didapatkan 10 baris data dengan kondisi *true positive*, 4 baris data dengan kondisi *false positive*, 8 baris data dengan kondisi *false negative*, dan 51 baris data dengan kondisi *true negative*.



Gambar 13. Confusion Matrix SVM Sigmoid dengan SMOTE

Pada gambar 13 di atas. hasil pengujian klasifikasi SVM Sigmoid dengan SMOTE, didapatkan 11 baris data dengan kondisi *true positive*, 8 baris data dengan kondisi *false positive*, 7 baris data dengan kondisi *false negative*, dan 47 baris data dengan kondisi *true negative*.

### 3.8 Evaluasi Confusion Matrix

Untuk mendapatkan hasil evaluasi berupa akurasi, presisi, *recall*, dan *f1-score*, dari grafik *confusion matrix* di atas, dilakukan dengan penghitungan pada salah satu contoh model klasifikasi SVM Linear dengan SMOTE sebagai berikut.

$$\begin{aligned} \text{Akurasi \% (SVM Linear)} &= \frac{TP+TN}{TP+FP+FN+TN} \\ &= \frac{12+51}{12+6+4+51} = \frac{63}{73} = 0.86301 = 86.30\% \end{aligned}$$

$$\begin{aligned} \text{Recall \% (SVM Linear)} &= \frac{TP}{TP+FN} \\ &= \frac{12}{12+6} = \frac{12}{18} = 0.66667 = 66.67\% \end{aligned}$$

$$\begin{aligned} \text{Presisi \% (SVM Linear)} &= \frac{TP}{TP+FP} \\ &= \frac{12}{12+4} = \frac{12}{16} = 0.75 = 75\% \end{aligned}$$

$$\begin{aligned} \text{F1-Score \% (SVM Linear)} &= 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \\ &= 2 \times \frac{0.75 \times 0.66667}{0.75 + 0.66667} = \frac{1.00005}{1.41667} = 0.70591 = 70.59\% \end{aligned}$$

Kemudian evaluasi performa klasifikasi SVM dengan SMOTE, dan klasifikasi SVM tanpa SMOTE pada kernel linear, polynomial, rbf gaussian, dan sigmoid, ditampilkan pada tabel 3. dan tabel 4. di bawah ini.

Tabel 3. Evaluasi Pengujian Klasifikasi SVM dengan SMOTE

No	Evaluasi	Kernel			
		Linear	Polynomial	RBF Gaussian	Sigmoid
1.	Akurasi	86.30%	84.93%	83.56%	79.45%
2.	Presisi	75.00%	81.82%	71.43%	57.89%
3.	Recall	66.67%	50.00%	55.56%	61.11%
4.	F1-Score	70.59%	62.07%	62.50%	59.46%

Tabel 4. Evaluasi Pengujian Klasifikasi SVM tanpa SMOTE

No	Evaluasi	Kernel			
		Linear	Polynomial	RBF Gaussian	Sigmoid
1.	Akurasi	80.82%	78.08%	80.82%	80.82%
2.	Presisi	70.00%	100.00%	100.00%	75.00%
3.	Recall	38.89%	11.11%	22.22%	33.33%
4.	F1-Score	50.00%	20.00%	36.36%	46.15%

Dari kedua tabel hasil evaluasi pada skenario pengujian klasifikasi, didapatkan performa SVM *linear* dengan SMOTE pada skenario pengujian pertama ini, mendapatkan performa nilai akurasi tertinggi dalam melakukan klasifikasi komentar positif dan komentar negatif yaitu 86.30%. Kemudian diikuti dengan nilai presisi 75%, nilai recall 66.67%, dan nilai f1-score 70.59%.

### 3.9 Analisis Sentimen Data Validasi

Model klasifikasi SVM linear dengan SMOTE telah mengklasifikasikan 50 baris data komentar. Sebanyak 36 baris data yang berhasil terklasifikasi dengan benar. Terdapat 7 baris data hasil klasifikasi komentar positif yang sesuai dengan data validasi atau 19.45% masyarakat memberikan komentar positif. Lalu 29 baris data hasil klasifikasi komentar negatif yang juga sesuai dengan data validasi atau 80.55% masyarakat memberikan komentar negatif. Evaluasi tingkat keakuratan SVM linear dengan SMOTE, terhadap data validasi berdasarkan penghitungan akurasi,

$$\begin{aligned} \text{Akurasi \% (SVM Linear)} &= \frac{TP+TN}{TP+FP+FN+TN} \\ &= \frac{7+29}{7+7+7+29} = \frac{36}{50} = 0.72 = 72\% \end{aligned}$$

maka didapatkan nilai akurasi 72%, dalam mengklasifikasikan label komentar positif dan label komentar negatif pada data validasi.

### 3.10 Hasil Analisis

Pada sejumlah 734 baris data komentar untuk penelitian analisis sentimen ini, terdapat 248 komentar positif dan 486 komentar negatif. 66.21% masyarakat terutama pengguna media sosial Youtube memang lebih dominan memberikan sentimen negatif. Hal ini dapat diketahui dari penyampaian sentimen yang kurang baik dengan ekspresi tidak setuju, marah, dan rasa kecewa yang ada pada data komentar, dengan hasil survei yang tidak transparansi dan tidak jelas asalnya.

33.79% masyarakat pengguna Youtube memberikan sentimen positif dalam hal memberikan kritik dan masukan, untuk melakukan survei tidak hanya di wilayah ramai penduduk saja, tetapi juga dilakukan didaerah-daerah terpencil. Komentar positif juga dilontarkan dalam mengapresiasi kinerja kebijakan pemerintah dalam menaikkan harga kebutuhan pokok, dan memberikan saran untuk lebih mempertimbangkan nasib ekonomi rakyat.

Rangkaian hasil penelitian berdasarkan urutan/susunan logis untuk membentuk sebuah cerita. Isinya menunjukkan fakta/data. Dapat menggunakan Tabel dan Angka tetapi tidak menguraikan secara berulang terhadap data yang sama dalam gambar, tabel dan teks. Untuk lebih memperjelas uraian, dapat menggunakan sub judul.

Pembahasan adalah penjelasan dasar, hubungan dan generalisasi yang ditunjukkan oleh hasil. Uraianya menjawab pertanyaan penelitian. Jika ada hasil yang meragukan maka tampilkan secara objektif.

## 4. KESIMPULAN

Hasil penelitian analisis sentimen komentar positif dan komentar negatif, pada konten video dimedia sosial Youtube menggunakan algoritma *Support Vector Machine*, dapat diambil beberapa poin kesimpulan diantaranya:

1. Hasil analisis sentimen dengan algoritma *Support Vector Machine*, menunjukkan bahwa mayoritas masyarakat berpendapat pada hasil kecurangan lembaga survei pada tingkat kepuasan kinerja presiden, yang tidak sebanding dengan kondisi perekonomian di Indonesia yang sedang mengalami krisis kenaikan harga kebutuhan pokok.
2. Pengujian model klasifikasi SVM linear dengan SMOTE mendapatkan tingkat keakuratan terbaik dengan nilai akurasi tertinggi yaitu 86.33%, diikuti dengan nilai presisi 75%, nilai recall 66.67%, dan nilai f1-score 70.59%.
3. SVM kernel linear dengan SMOTE telah berhasil mengklasifikasi data validasi (data uji baru, diluar dari dataset), dengan benar pada sentimen komentar negatif sebanyak 29 baris data, dan sentimen komentar positif sebanyak 7 baris data. Namun kesalahan klasifikasi juga telah terhitung sebanyak 7 baris data pada sentimen komentar negatif aktual dan 7 baris data pada sentimen komentar positif aktual.
4. Evaluasi performa tingkat keakuratan SVM dengan kernel linear berdasarkan penghitungan keakuratan pada data validasi, diperoleh nilai akurasi 72%.

## DAFTAR PUSTAKA

- [1] F. F. Abdulloh and I. R. Pambudi, "Analisis Sentimen Pengguna Youtube Terhadap Program Vaksin Covid-19," *CSRID (ComputerSci.Res.ItsDev. Journal)*, vol. 13, no. 3, p. 141, 2021.
- [2] Ash Shiddicky and Surya Agustian, "Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 pada Media Sosial Twitter menggunakan Metode Logistic Regression," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 3, no. 2, pp. 99–106, 2022.
- [3] H. Tuhuteru, "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berskala Besar Menggunakan Algoritma Support Vector Machine," *Inf. Syst. Dev.*, vol. 5, no. 2, pp. 7–13, 2020.
- [4] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021.
- [5] I. Subagyo, L. D. Yulianto, W. Permadi, and A. W. Dewantara, "Sentiment Analisis Review Film Di IMDB Menggunakan Algoritma SVM," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 8–1, no. 1, pp. 47–56, 2019.
- [6] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, "Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 47–58, 2021.
- [7] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiyari, "Analisis Sentimen pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Bertas SMOTE," *AITI J. Teknol. Inf.*, vol. 18, no. 2, pp. 173–184, 2021.
- [8] Andreyestha and Q. N. Azizah, "Analisa Sentimen Kicauan Twitter Tokopedia Dengan Optimalisasi Data Tidak Seimbang Menggunakan Algoritma SMOTE," *Infotek J. Inform. dan Teknol.*, vol. 5, no. 1, pp. 108–116, 2022.
- [9] Aditya Quantano Surbakti, Regiolina Hayami, and Januar Al Amien, "Analisa Tanggapan Terhadap Pssb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 2, no. 2, pp. 91–97, 2021.
- [10] A. N. Kasanah, M. Muladi, and U. Pujianto, "Penerapan Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Objektivitas Berita Online Menggunakan Algoritma KNN," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 2, pp. 196–201, 2019.
- [11] D. S. Utami and A. Erfina, "Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes Pada dasarnya Indonesia," vol. 6, pp. 418–427, 2022.
- [12] M. O. Ibrohim and I. Budi, "Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter," pp. 46–57, 2019.
- [13] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *Jdmsi*, vol. 2, no. 1, pp. 31–37, 2021.
- [14] F. N. Rozi and D. H. Sulistyawati, "Klasifikasi Berita Hoax Pilpres Menggunakan Metode Modified K-Nearest Neighbor Dan Pembobotan Menggunakan Tf-Idf," *Konvergensi*, vol. 15, no. 1, 2019.
- [15] D. A. Lionovan, L. W. Santoso, and R. Intan, "Klasifikasi Topik dan Analisa Sentimen Terhadap Kuesioner Umpun Balik Universitas Menggunakan

- [16] Metode Long Short-Term Memory,” *J. Infra*, vol. 8, no. 2, pp. 1–6, 2017.  
D. T. Lukmana, S. Subanti, and Y. Susanti, “Analisis Sentimen Terhadap Calon Presiden 2019 Dengan Support Vector Machine Di Twitter,” *Semin. Nas. Penelit. Pendidik. Mat. 2019 UMT*, no. 2002, pp. 154–160, 2019.