



# Analisis Sentimen Ulasan Aplikasi Wetv Untuk Peningkatan Layanan Menggunakan Metode Support Vector Machine

Rezky Abdillah\*, Elin Haerani, Reski Mai Candra

Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru  
Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

Email: <sup>1,\*</sup>11950111734@students.uin-suska.ac.id, <sup>2</sup>elin.haerani@uin-suska.ac.id, <sup>3</sup>reski.candra@uin-suska.ac.id

Email Penulis Korespondensi: 11950111734@students.uin-suska.ac.id

Submitted: 09/04/2023; Accepted: 30/04/2023; Published: 30/04/2023

**Abstrak**—Wetv ialah sebuah media streaming online yang berjalan sejak 2019. Wetv memiliki banyak review pengguna dari berbagai aplikasi. Peringkat terdiri dari positif, netral dan negatif. Respon tersebut digunakan untuk mengetahui sentimen dengan menggunakan metode klasifikasi support vector machine. Studi ini mengambil 12.000 komentar dari Google Play Store, penelitian ini menggunakan preprocessing yaitu, cleaning, case folding, tokenizing, normalisasi, stopword removal, dan steaming, selanjutnya ke tahap TF-IDF dan hasil akhir dilakukan pengujian dengan confusion matrix dengan program python, hasil akurasi tertinggi dari proses pengujian memperoleh dengan accuracy 0,76%, precision 0,77%, recall 0,79%, dan f1-score 0,78, pada dataset 90% data latih dan 10% data uji. Berdasarkan hasil penelitian diketahui metode Support Vector Machine baik dalam proses identifikasi tanggapan negatif pada WeTV.

**Kata Kunci:** Analisis Sentimen; Confusion Matrix; Python; TF-IDF; WeTV

**Abstract**—Wetv is an online streaming media that has been running since 2019. Wetv has many user reviews from various applications. The rating consists of positive, neutral and negative. The response is used to determine sentiment by using the support vector machine classification method. This study took 12,000 comments from the Google Play Store, this study used preprocessing namely, cleaning, case folding, tokenizing, normalization, stopword removal, and steaming, then to the TF-IDF stage and the final results were tested with a fusion matrix with the Python program, the score results highest from the acquisition test process with accuracy of 0.76%, precision of 0.77%, recall of 0.79%, and f1 score of 0.78, in a dataset of 90% training data and 10% test data. Based on the research results of the Support Vector Machine method which is known to be good in the process of requesting negative responses on WeTV.

**Keywords:** Sentiment Analysis; Confusion Matrix; Python; TF-IDF; WeTV

## 1. PENDAHULUAN

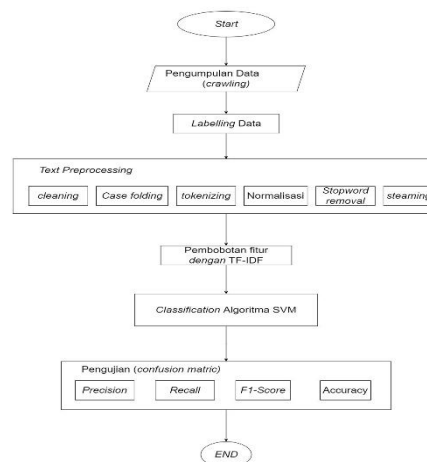
Merebaknya COVID-19, ramai negara seluruh dunia kebijakan Pembatasan Sosial Berskala Besar (PSBB) dilaksanakan, termasuk Indonesia. Upaya tersebut dilakukan untuk menghambat tentang penyebaran virus Covid-19. Akibat pemberlakuan PSBB, banyak perusahaan yang mengalami harus menghentikan sementara operasional. Salah satunya adalah penutupan pemutaran film (bioskop) karena penutupan pusat perbelanjaan yang menimbulkan kekecewaan dikalangan penggemar film. Di sisi lain, meskipun penyebaran virus Covid-19 membawa banyak dampak buruk, ada sisi positif seperti peningkatan penggunaan aplikasi streaming film yang disebabkan meningkatnya penggunaan internet sebagai bagian dari penerapan PSBB yang didukung, yang mengharuskan masyarakat beraktivitas di rumah [1]. Perkembangan teknologi secara mendasar telah mengubah cara industri media beroperasi, dengan konsumen dan platform media bergerak secara online [2]. Banyak orang beralih dari televisi ke layanan streaming digital. Kemajuan internet yang pesat di dunia menyebabkan banyak perusahaan dan layanan berpindah dari offline ke online, termasuk media streaming. Menurut We Are Social, 64% penduduk Indonesia terhubung dengan internet dari total 272 juta penduduk Indonesia. Semakin cepat pengguna internet, semakin menuntut pengguna media sosial. Fakta ini dibuktikan dengan 59% atau 160 juta penduduk Indonesia yang aktif menggunakan jejaring sosial [3]. Banyaknya perangkat lunak yang tersedia saat ini mencerminkan pesatnya perkembangan teknologi dan internet. Berdasarkan informasi dari asosiasi penyelenggara jasa internet Indonesia, Indonesia akan memiliki sekitar 196,7 juta pengguna internet pada akhir tahun 2020. Aplikasi yang tersedia dapat diunduh dari Play Store di smart phone. Aplikasi global meningkat 8,7% pada kuartal pertama tahun 2021 dibandingkan periode yang sama tahun 2020, dengan Google Play Store menjadi produsen teratas dengan 3,8 miliar unduhan [4]. Setiap aplikasi memiliki kelebihan dan kekurangannya masing-masing, yang dapat menimbulkan reaksi yang berbeda seperti kepuasan atau kekecewaan dari para pengguna aplikasi tersebut. Di bagian Komentar Tinjauan Aplikasi, seperti merasa puas atau kecewa. Bagian kolom komentar ulasan aplikasi adalah tempat memberikan respon tentang kepuasan dan ketidakpuasan user terhadap aplikasi. Kepuasan pengguna dapat disebut sebagai penilaian positif karena memiliki efek positif mudah digunakan. Sebaliknya, tanggapan yang mengecewakan dari pengguna dapat disebut sebagai review negatif karena pengguna mendapatkan sesuatu yang tidak sesuai dengan keinginannya. Dengan adanya review dari aplikasi yang teliti oleh Andry Novantika dan Sugiman pada tahun 2022 menjadikan informasi tambahan bagi perusahaan dalam mempengaruhi asumsi dan keputusan pengguna dalam menggunakan aplikasi yang diinginkan, perusahaan memperoleh informasi tambahan tentang pengaruh asumsi dan keputusan pengguna saat menggunakan aplikasi yang diinginkan [5]. Analisis sentimen adalah bidang penelitian yang berkembang yang berada di persimpangan bidang seperti penambahan data, pemrosesan bahasa alami (NLP), dan pembelajaran machine learning. Fokus utamanya adalah

untuk mengekstraksi sentimen dari sebuah kalimat berdasarkan kontennya [6]. Klasifikasi merupakan suatu teknik untuk menilai objek data serta mengelompokkan objek berdasarkan atribut, untuk mempermudah dalam pengklasifikasian data diperlukan suatu sistem menggunakan metode text mining [7]. Teks adalah alat yang berguna di tangan para ahli untuk mengekstraksi informasi yang diperlukan dari organisasi, suplemen teks dapat digunakan diberbagai industri dan area aplikasi, termasuk pengambilan keputusan, analisis sentimen, dan lainnya [8]. Tujuan dari analisis sentimen adalah untuk menafsirkan dan menghasilkan emosi yang disampaikan dalam data tekstual [9]. Berdasarkan hasil pengujian yang diteliti oleh Naraswati dkk pada tahun 2021, dapat disimpulkan bahwa masyarakat lebih banyak banyak sentimen negatif terkait penanganan COVID-19, terutama PSBB, kewajiban masker dan jam malam yaitu 72,72 persen. Tiga kata paling umum dalam arti positif adalah demo, jakarta, dan kerja. Pada saat yang sama ada perasaan negatif terhadap Jakarta, protes dan rakyat [10]. Dari penelitian mengklasifikasi sentimen masyarakat melalui media sosial twitter terhadap tatanan baru New Normal menggunakan metode Support Vector Machine yang dianalisis oleh Alfredo Gormantara pada tahun 2022. Pengujian dilakukan dengan menggunakan data latih dan data uji berupa data twitter yang diambil menggunakan aplikasi rapidminer memberikan hasil accuracy 76.50%, recall 90.91% dan precision 70.80% [11]. Berdasarkan hasil pengujian yang dilakukan oleh Fadhilah Dwi Ananda dan Yoga Pristyanto pada tahun 2020 menggunakan kernel linear dan kernel RBF dengan 3 skenario pengujian pada setiap kernel, dapat diambil kesimpulan sebagai berikut. Algoritma SVM menggunakan dengan kernel linear maupun kernel RBF memiliki hasil kinerja evaluasi baik dari sisi akurasi, presisi dan recall yang relatif sama. Sehingga dapat dikatakan bahwa algoritma SVM baik dengan kernel RBF maupun Linear sama-sama dapat digunakan dengan baik dalam menentukan sentimen pengguna internet Biznet. Selain itu dengan 3 skenario pengujian dengan jumlah data yang berbeda algoritma SVM baik dengan kernel RBF maupun Linear sama-sama konsisten kinerjanya [12]. Berdasarkan hasil penelitian dan pengujian metode Support Vector Machine yang dilakukan oleh Isnain dkk pada tahun 2021, untuk mengklasifikasi kebijakan lockdown pemerintah Jakarta dengan tweet berbahasa Indonesia menggunakan ekstraksi fitur TF-IDF didapat kesimpulan bahwa penggunaan metode Support Vector Machine tingkat nilai Accuracy sebesar 74%, Precision sebesar 75%, Recall sebesar 92% dan F1-Score sebesar 83%. Yang artinya cukup baik dalam penelitian mengenai kebijakan lockdown pemerintah Jakarta [13]. Temuan studi yang diteliti oleh Rima Tamara Aldisa dan Pandu Maulana pada tahun 2022, menunjukkan bahwa kinerja model SVM lebih bagus dibandingkan dengan dua model lainnya dalam rata-rata. Selain itu, jumlah emosi pada setiap dataset telah dihitung dan ditemukan bahwa terdapat 2248 emosi positif dan 751 emosi negatif [14].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian melibatkan beberapa tahapan, yaitu pengumpulan data, preprocessing data, ekstraksi fitur, pelabelan data, pembagian data, pembentukan model, evaluasi model, dan interpretasi hasil. Dalam tahapan pengumpulan ulasan aplikasi yang akan dianalisis sentimennya. Setelah itu, data akan diproses melalui tahapan preprocessing dan ekstraksi fitur. Data kemudian akan dilabeli dan dibagi menjadi data latih dan data uji. Setelah itu, kan membentuk model SVM dan mengevaluasi kinerjanya. Terakhir, hasil yang diperoleh dapat digunakan sebagai umpan balik bagi pengembang aplikasi untuk meningkatkan kualitas dan kepuasan pengguna, pada gambar 1.



**Gambar 1.** Tahapan Penelitian

### 2.2. Pengumpulan Data

Data diperoleh melalui tinjauan aplikasi WeTV dengan Play Store. Data diambil dari komentar pengguna WeTV di Play Store menggunakan teknik scraping dengan Python bernama Google Play Scraper [15]. Hasilnya



merupakan 12000 komentar dari bulan Juli 2019 sampai dengan Maret 2023. Komentar yang terhimpun terbagi kedalam tiga jenis sentimen, yakni positif, netral, dan negatif.

### 2.3 Pembobotan Term Frequency-Inverse Document Frequency

TF-IDF merupakan metode untuk mencari bobot suatu kata dalam dokumen kunci di setiap kategori dan mencari kata kunci yang hampir mirip dengan kategori yang tersedia. Pembobotan ini menggunakan teknik Term Frequency and Inverse Document Frequency (TF-IDF) untuk mengekstraksi ciri dari suatu teks dengan cara menggabungkan metode TF dan IDF [16].

$$w_{ij} = tf(i, j) \times idf \tag{1}$$

Rumus mencari nilai idf sebagai berikut:

$$idf = \log\left(\frac{n}{df_i}\right) \tag{2}$$

Dengan:

- W(i,j) : berat dokumen j
- Tf(i, j) : Frekuensi istilah i dalam dokumen j
- Idf : i nilai kata IDF
- N : jumlah seluruh dokumen
- DF : Jumlah dokumen yang mengandung kata i

### 2.4 Support Vector Machine

Support Vector Machine ialah teknik prediktif yang bagus untuk klasifikasi dan regresi. menggunakan algoritma Support Vector Machine untuk mengklasifikasikan teks menggunakan bobot indeks istilah sebagai fitur oleh Thorsten Joachim. Boser, Guyon dan Vapnik menyukai pembelajaran SVM sejak tahun 1992 . SVM adalah metode yang dapat memecahkan masalah linier dan nonlinier serta menggunakan teknik kernel untuk mengubah data menjadi dimensi yang lebih tinggi. Algoritma SVM banyak digunakan dalam analisis sentimen sebagai alat untuk mengklasifikasikan teks ke dalam kategori positif, negatif, dan netral [17].

### 2.5 Pengujian

Menghitung keakuratan klasifikasi dengan Confusion Matrix ialah metode yang dipakai guna mengevaluasi model klasifikasi agar dapat menghitung kebenaran atau kesalahan dari objek tersebut. Suatu matriks akan dibuat dari perkiraan yang dibandingkan dengan kelas asli, yang berisi informasi aktual dan nilai prediksi klasifikasi. Setelah sistem berhasil mengklasifikasikan ulasan, diperlukan pengukuran untuk menentukan seberapa valid atau tepat klasifikasi yang telah dibuat oleh sistem.[18]

Tabel 1. Confusion Matrix

		Kelas Hasil Prediksi		
		Positif	Netral	Negatif
Actual Class	Positif	True Positif	False Positif	False Positif Negatif
	Netral	False Netral Positif	False Netral Negatif	True Netral
	Negatif	False Negatif Positif	True Negatif	False Negatif Netral

Confusion matrix adalah penjumlahan yang mendapatkan 3 output, yaitu accuracy, recall, dan precision.

$$Akurasi = \frac{(TP+TN)}{TP+TN+FP+FN} \times 100\% \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{6}$$

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan Data

Pengumpulan data yang dilakukan menggunakan metode crawling dari aplikasi wetv menggunakan google-play-scraper [19]. Metode scraping ini mengumpulkan data dari web dan mengubahnya menjadi format CSV. Data yang dikumpulkan oleh survei online Google Play berisi variable, tanggal ulasan, nama pengguna, dan rating. Berdasarkan penilaian tersebut, informasi yang terkandung dalam kedua aplikasi tersebut dikaji secara deskriptif atau secara umum.



### 3.2 Pelabelan Data

Pelabelan ulasan WeTV dilakukan untuk menentukan kelas dari pelabelan data yang terdiri 3 kelas yaitu positif, negatif dan netral. Berikut contoh ulasan data WeTV:

Tabel 2. Pelabelan Data

Ulasan	Label
Dramanya bagus.	Positif
WeTV memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat. Secara keseluruhan, aplikasi ini masih layak untuk dicoba.	Netral
udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus.	Negatif

### 3.3 Text Preprocessing

Preprocessing adalah bagian dari text mining yang diperlukan untuk menghilangkan noise dari kalimat, data yang diambil adalah review atau ulasan kemudian dilakukan melewati tahapan proses Case folding, cleansing, tokenize, normalisasi, stopword, dan stemming [20].

#### 3.3.1 Case Folding

Case Folding adalah proses yang mengubah segalanya mengonversi karakter teks menjadi huruf kecil dan menghapus tanda baca dan angka [21].

Tabel 3. Case Folding

Text sebelum proses Case Folding	Text sesudah proses Case Folding
Dramanya bagus.	dramanya bagus.
WeTV memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat. Secara keseluruhan, aplikasi ini masih layak untuk dicoba.	wetv memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat. secara keseluruhan, aplikasi ini masih layak untuk dicoba.
udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus.	udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus.

#### 3.3.2 Cleaning

Cleaning adalah menghapus karakter non-alfanumerik untuk mengurangi kebisingan. Merek yang dihapus termasuk mis. tanda baca, simbol seperti tanda @ pada nama pengguna, tag hash (#), emotikon, dan URL halaman webTokenizing [22].

Tabel 4. Cleaning

Text sebelum proses Case Folding	Text sesudah proses Case Folding
dramanya bagus.	dramanya bagus
wetv memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat. secara keseluruhan, aplikasi ini masih layak untuk dicoba.	wetv memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat secara keseluruhan aplikasi ini masih layak untuk dicoba
udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus.	udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus

#### 3.3.3 Tokenizing

Tokenizing adalah langkah dimana teks kalimat dibagi menjadi bagian-bagian per kata sesuai dengan spasi dalam teks [23].

Tabel 5. Tokenizing

Text sebelum proses Tokenizing	Text sesudah proses Tokenizing
dramanya bagus	'dramanya' 'bagus'
wetv memiliki pilihan tontonan yang beragam, tetapi saya mengalami beberapa masalah teknis seperti buffering dan loading yang lambat secara keseluruhan aplikasi ini masih layak untuk dicoba	'wetv' 'memiliki' 'pilihan' 'tontonan' 'yang' 'beragam' 'tetapi' 'saya' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'dan' 'loading' 'yang' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'ini' 'masih' 'layak' 'untuk' 'dicoba'
udah bayar tapi nonton vip masih disuruh bayar lagi apk tidak bagus	'udah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'apk' 'tidak' 'bagus'

### 3.3.4 Normalisasi

Normalisasi ialah langkah-langkah mengganti kata yang tidak baku. Tujuannya adalah memperbaiki kesalahan pada kata, seperti salah ejaan, supaya kata tersebut memiliki makna yang serupa. [24]. Langkah ini penting untuk memperbaiki kata-kata tersebut menjadi kata baku bahasa Indonesia.

**Tabel 6.** Normalisasi

Text sebelum proses normalisasi	Text sesudah proses normalisasi
'dramanya' 'bagus'	'dramanya' 'bagus'
'wetv' 'memiliki' 'pilihan' 'tontonan' 'yang' 'beragam' 'tetapi' 'saya' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'dan' 'loading' 'yang' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'ini' 'masih' 'layak' 'untuk' 'dicoba'	'wetv' 'memiliki' 'pilihan' 'tontonan' 'yang' 'beragam' 'tetapi' 'saya' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'dan' 'loading' 'yang' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'ini' 'masih' 'layak' 'untuk' 'dicoba'
'udah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'apk' 'tidak' 'bagus'	'sudah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'aplikasi' 'tidak' 'bagus'

### 3.3.5 Stopword

Adalah langkah menghilangkan kata-kata yang tak berarti, seperti kata penghubung atau kata penjelas, serta sejenisnya [25].

**Tabel 7.** Stopword

Text sebelum proses Stopword	Text sesudah proses Stopword
'dramanya' 'bagus'	'dramanya' 'bagus'
'wetv' 'memiliki' 'pilihan' 'tontonan' 'yang' 'beragam' 'tetapi' 'saya' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'dan' 'loading' 'yang' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'ini' 'masih' 'layak' 'untuk' 'dicoba'	'wetv' 'memiliki' 'pilihan' 'tontonan' 'beragam' 'tetapi' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'loading' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'masih' 'layak' 'dicoba'
'sudah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'aplikasi' 'tidak' 'bagus'	'sudah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'aplikasi' 'tidak' 'bagus'

### 3.3.6 Stemming

Perubahan penghilangan imbuhan. imbuhan dalam bahasa Indonesia meliputi me-, mem-, meng-, di-, dan lain-lain [26].

**Tabel 8.** Stemming

Text sebelum proses Stopword	Text sesudah proses Stopword
'dramanya' 'bagus'	'drama' 'bagus'
'wetv' 'memiliki' 'pilihan' 'tontonan' 'beragam' 'tetapi' 'mengalami' 'beberapa' 'masalah' 'teknis' 'seperti' 'buffering' 'loading' 'lambat' 'secara' 'keseluruhan' 'aplikasi' 'masih' 'layak' 'dicoba'	'wetv' 'milik' 'pilih' 'tonton' 'ragam' 'tapi' 'alami' 'beberapa' 'salah' 'teknis' 'seperti' 'buffering' 'loading' 'lambat' 'cara' 'seluruh' 'aplikasi' 'masih' 'layak' 'coba'
'sudah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'disuruh' 'bayar' 'lagi' 'aplikasi' 'tidak' 'bagus'	'sudah' 'bayar' 'tapi' 'nonton' 'vip' 'masih' 'suruh' 'bayar' 'lagi' 'aplikasi' 'tidak' 'bagus'

### 3.4 TF-IDF

TF-IDF (term frequency-inverse document frequency) adalah metode yang digunakan untuk mengukur seberapa penting suatu kata dalam suatu dokumen atau korpus teks. merupakan jumlah kemunculan setiap kata dalam sebuah dokumen. Semakin ramai kata yang muncul dalam setiap dokumen, semakin tinggi nilai yang diberikan. TF IDF [22].



	TF-IDF	TF
koleksi	0.408618	1
asianya	0.408618	1
thank	0.371446	3
jadul	0.359479	4
tambahin	0.308236	14
mantap	0.304262	15
kendala	0.297101	15
guna	0.240466	46
film	0.151872	313
bagus	0.150330	276

**Gambar 2.** TF-IDF

### 3.5 Klasifikasi

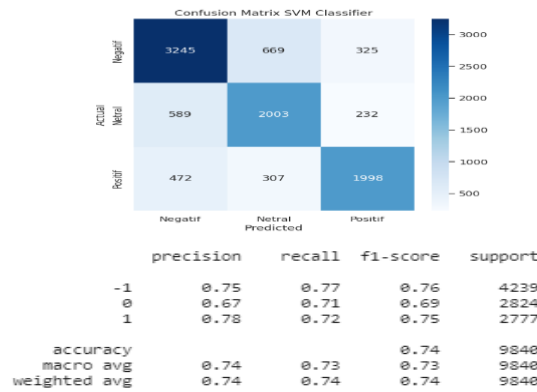
Setelah tahap pra-pemrosesan dan pembobotan, langkah selanjutnya adalah mengklasifikasikan data. Metode yang dipakai untuk klasifikasi ialah support vector machine. [27]. Dengan kernel linear dan dengan teknik One-vs-Rest untuk mengklasifikasikan dataset yang memiliki lebih dari dua kelas (Positif, Negatif, Netral) Hasil klasifikasi menggunakan algoritma SVM memberikan akurasi model pada skenario pertama 60% Data latih dan 40% data uji, dalam skenario kedua 70% data tiha dan 30% data uji, dalam skenario ketiga 80% data latih dan 20% data uji, dan dalam skenario keempat 90% data latih dan 10% data uji. Berikut hasil pengujian menggunakan algoritma SVM:

**Tabel 9.** Hasil Klasifikasi pada SVM

Skenario	Karnel	Accuracy
60:40	Linier	0,74%
70:30	Linier	0,75%
80:20	Linier	0,76%
90:10	Linier	0,76%

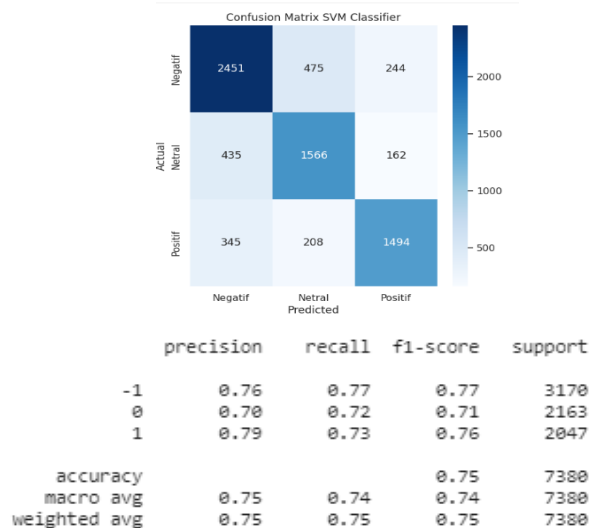
### 3.6 Pengujian

Pengujian menggunakan confusion matrix yang berguna dalam pengecekan dan melihat hasil prediksi algoritma SVM. Untuk hasil penerapan cofusion matrix menampilkan nilai accuracy, precision, recall, dan f1-score sebagai berikut:



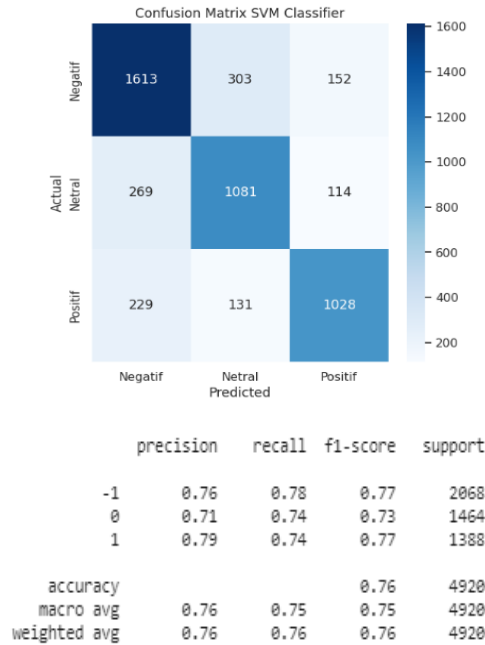
**Gambar 3.** Hasil Confusion Matrix Skenario Pertama

Skenario pertama pada gambar 3, data latih sebanyak 60% dan 40% data uji menghasilkan nilai accuracy sebesar 74%, precision 0,75%, recall 0,77%, dan f1-score 0,76%. Adapun hasil prediksi untuk kelas sentimen negatif sebesar 3245 data, sedangkan true negative sebesar 994 data. untuk kelas netral diprediksi true positif sebesar 2003 data, sedangkan true negative sebesar 821 data. Untuk kelas positif true positif sebesar 1998 data, sedangkan true negatif sebesar 779 data.



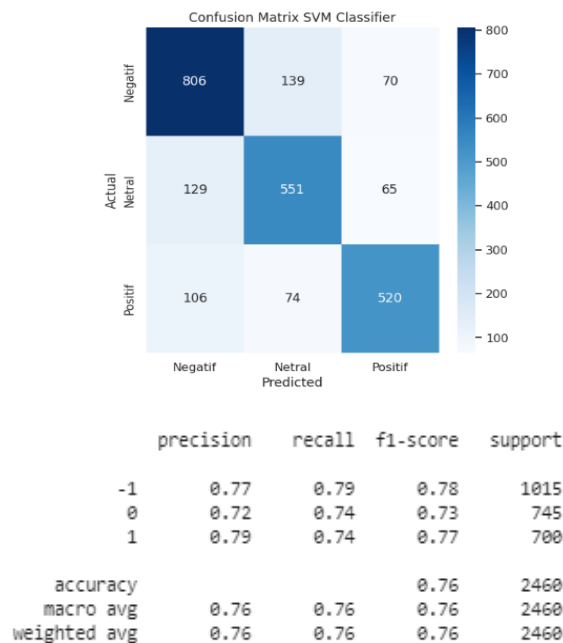
**Gambar 4.** Hasil Confusion Matrix skenario kedua

Skenario kedua pada gambar 4, data latih sebanyak 70% dan 30% data uji menghasilkan nilai accuracy sebesar 0,75%, precision 0,76 %, recall 0,77%, dan f1-score 0,77%. Adapun hasil prediksi untuk kelas sentimen negatif sebesar 2451 data, sedangkan true negative sebesar 597 data. untuk kelas netral diprediksi true positif sebesar 1566 data, sedangkan true negative sebesar 89 data. Untuk kelas positif true positif sebesar 1494 data, sedangkan true negatif sebesar 553 data.



**Gambar 5.** Hasil Confusion Matrix Skenario ketiga

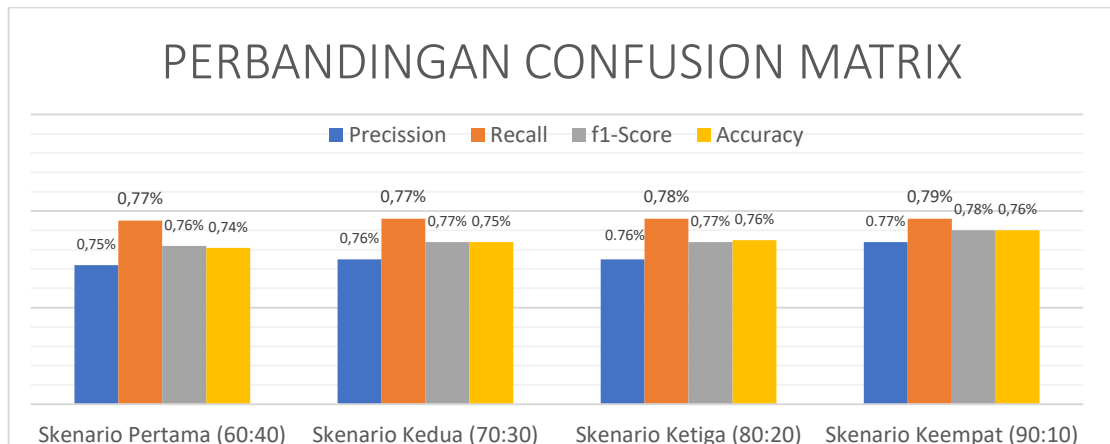
Skenario ketiga pada gambar 5, data latih 80% dan 20% data uji menghasilkan nilai accuracy sebesar 0,76%, precision 0,76%, recall 0,78%, dan f1-score 0,77%. Adapun hasil prediksi untuk kelas sentimen negatif sebesar 1613 data, sedangkan true negative sebesar 455 data. untuk kelas netral diprediksi true positif sebesar 1081 data, sedangkan true negative sebesar 383 data. Untuk kelas positif true positif sebesar 1028 data, sedangkan true negatif sebesar 360 data.



**Gambar 6.** Hasil Confusion Matrix Skenario keempat

Skenario keempat pada gambar 6, dengan pembagian data latih sebanyak 90% dan 10% data uji menghasilkan nilai accuracy sebesar 0,76%, precision 0,77%, recall 0,79%, dan f1-score 0,78%. Adapun hasil prediksi untuk kelas sentimen negatif sebesar 806 data, sedangkan true negative sebesar 209 data. untuk kelas netral diprediksi true positif sebesar 551 data, sedangkan true negative sebesar 65 data. Untuk kelas positif true positif sebesar 520 data, sedangkan true negatif sebesar 74 data.

netral diprediksi true positif sebesar 551 data, sedangkan true negative sebesar 194 data. Untuk kelas positif true positif sebesar 520 data, sedangkan true negatif sebesar 180 data.



**Gambar 7.** Grafik Perbandingan Confusion Matrix

#### 4. KESIMPULAN

Berdasarkan hasil pengujian dengan algoritma Support Vector Machine dalam melakukan klasifikasi sentimen data ulasan aplikasi wetv di play store menghasilkan akurasi yang baik. Skenario dengan model tertinggi didapatkan oleh 3 skenario dengan yang pertama skenario keempat dengan accuracy 0,76%, precision 0,77%, recall 0,79%, dan f1-score 0,78%. Kedua yang tertinggi pada skenario ketiga dengan accuracy 0,76%, precision 0,76%, recall 0,78% dan f1-score 0,77%. Ketiga yang tertinggi pada skenario kedua dengan accuracy 0,75%, precision 0,76%, recall 0,77% dan f1-score 0,77%. Sedangkan skenario yang memiliki nilai akurasi terendah pada skenario pertama dengan nilai accuracy 0,74%, precision 0,75%, recall 0,77% dan f1-score 0,76%. Keluhan pengguna subtitle yang ditampilkan saat menonton video dalam program ini lebih cepat atau terkadang lebih lambat dari yang seharusnya sehingga membuat terputusnya hubungan antara video atau audio yang masuk dengan teks yang ada. Berdasarkan ulasan wetv dapat menjadi saran pengembang agar kualitas aplikasi tetap terjaga dan pengguna tetap terjaga. Selain itu, diharapkan hasil penelitian ini dapat dijadikan sebagai bahan informasi sebagai referensi dalam penyusunan artikel pada aplikasi streaming online.

#### REFERENCES

- [1] M. F. Al-shufi and A. Erfina, "Sentimen Analisis Mengenai Aplikasi Streaming Film Menggunakan Algoritma Support Vector Machine Di Play Store," *Sismatik*, pp. 156–162, 2021.
- [2] J. Krisiyono, *KONVEGENSI MEDIA Tranformasi Media Komunikasi di Era Digital pada Masyarakat Berjejaring*, 1st ed. Jakarta: Kencana, 2022.
- [3] R. El Husna, R. Wasono, M. Al Haris, U. M. Semarang, and A. Sentimen, "Analisis Sentimen pada Twitter Mengenai Netflix Diblokis Telkom Menggunakan Support Vector Machine," *Semin. Nas. Variansi*, pp. 214–222, 2020.
- [4] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, pp. 1–7, 2022, doi: 10.31539/intecom.v5i1.3708.
- [5] A. Novantika, "Analisis Sentimen Ulasan Pengguna Aplikasi Video Conference Google Meet menggunakan Metode SVM dan Logistic Regression," *Prism. Pros. Semin. Nas. Mat.*, vol. 5, pp. 808–813, 2022, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [6] H. Tuhuteru, "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine," *Inf. Syst. Dev.*, vol. 5, no. 2, pp. 7–13, 2020.
- [7] M. Ramadhan and R. Andarsyah, *KLASIFIKASI TEKS SPAM MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN NAIVE BAYES*. Bandung: Penerbit Buku Pedia, 2022.
- [8] N. Purwati, H. Kurniawan, and S. Karnila, *DATA MINING*. Purwokerto: Zahira Media Publisher, 2021.
- [9] Z. Munawar and dkk, *Big Data Analytics Konsep, Implementasi, dan Aplikasi Terkini*, 1st ed. Bandung: Kaizen Media Publishing, 2023.
- [10] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, p. 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [11] A. Gormantara, "Analisis Sentimen Terhadap New Normal Era di Indonesia pada Twitter Analisis Sentimen Terhadap New Normal Era di Indonesia pada Twitter Menggunakan Metode Support Vector Machine," *Konf. Nas. Ilmu Komput.* 2020, no. July, pp. 1–5, 2020.
- [12] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 2, pp. 407–416, 2021, doi: 10.30812/matrik.v20i2.1130.



- [13] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, “Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm,” *J. Data Min. dan Sist. Inf.*, vol. 2, no. 1, p. 31, 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [14] R. T. Aldisa and P. Maulana, “Analisis Sentimen Opini Masyarakat Terhadap Vaksinasi Booster COVID- 19 Dengan Perbandingan Metode Naive Bayes , Decision Tree dan SVM,” vol. 4, no. 1, pp. 106–109, 2022, doi: 10.47065/bits.v4i1.1581.
- [15] Raksaka Indra Alhaqq, I Made Kurniawan Putra, and Yova Ruldeviyani, “Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 105–113, 2022, doi: 10.22146/jnteti.v11i2.3528.
- [16] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- [17] M. Amrullah and S. Pane, *Analisis Sentimen Masyarakat Terhadap Kebijakan Polisi Tilang Manual Di Indonesia*. Bandung: Penerbit Buku Pedia, 2023.
- [18] A. M. Pravina, I. Cholissodin, and P. P. Adikara, “Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine ( SVM ),” vol. 3, no. 3, pp. 2789–2797, 2019.
- [19] E. B. Santoso and A. Nugroho, “Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook,” *Eksplora Inform.*, vol. 9, no. 1, pp. 60–69, 2019, doi: 10.30864/eksplora.v9i1.254.
- [20] R. Safitri, N. Alfira, D. Tamitiadini, W. Dewi, and N. Febriani, *ANALISIS SENTIMEN: METODE ALTERNATIF PENELITIAN BIG DATA*. Universitas Brawijaya Press, 2021.
- [21] D. P. Santoso and W. Wibowo, “Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Metode Naïve Bayes Classifier pada Situs Google Play Store,” *J. Sains dan Seni ITS*, vol. 11, no. 2, 2022, doi: 10.12962/j23373520.v11i2.72534.
- [22] H. Tuhuteru and A. Iriani, “Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier,” *J. Inform. J. Pengemb. IT*, vol. 3, no. 3, pp. 394–401, 2018, doi: 10.30591/jpit.v3i3.977.
- [23] M. I. Petiwi, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine,” *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 542, 2022, doi: 10.30865/mib.v6i1.3530.
- [24] T. M. Permata Aulia, N. Arifin, and R. Mayasari, “Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinisasi Covid-19,” *SINTECH (Science Inf. Technol. J.)*, vol. 4, no. 2, pp. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [25] A. Erfina, E. S. Basryah, A. Saepulrohman, and D. Lestari, “Analisis Sentimen Aplikasi Pembelajaran Online Di Play Store Pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine (Svm),” *Semin. Nas. Inform.*, vol. 1, no. 1, pp. 145–152, 2020, [Online]. Available: <http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4094>
- [26] 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, doi: 10.25126/jtiik.0813944.
- [27] F. F. Irfani, “Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine,” *JBMI (Jurnal Bisnis, Manajemen, dan Inform.)*, vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.