

## KAJIAN OPINI PUBLIK TERHADAP TAYANGAN CLASH OF CHAMPIONS MENGUNAKAN SUPPORT VECTOR MACHINE

Rahmat Hidayat, Ibnu Daqiqil Id<sup>✉</sup>, Fahira P. Amanda, M. Fadhil, Daffa Ayu Mumtaza,  
Mhd. Aqshal Maulana, Abyan Abdullah Tamir, Imam Firmansyah

FMIPA, Universitas Riau, Pekanbaru, Indonesia

Email: [ibnu.daqiqil@lecturer.unri.ac.id](mailto:ibnu.daqiqil@lecturer.unri.ac.id)

### ABSTRACT

*This study examines audience responses to the educational program Clash of Champions produced by Ruangguru and inspired by the South Korean format University War. Through its interactive academic competition concept, Clash of Champions has succeeded in attracting public interest, triggering various reactions on social media, and providing a basis for conducting opinion studies to understand audience views. This study uses the Support Vector Machine (SVM) algorithm to classify netizen comments on platform X (Twitter) into positive, negative, and neutral sentiments. The data processing process includes data collection through scraping, text preprocessing, and word weighting using the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm. Model evaluation with accuracy, precision, recall, and F1-score metrics shows various audience opinions, with the majority showing positive sentiments towards the educational innovation of the program. This model achieves an accuracy of 79%, precision of 79%, recall of 95%, and F1-score of 86% for the positive sentiment classification. This indicates a strong appreciation from the public for the educational value of this program. This study is expected to provide insight for educational content developers to create educational and interesting programs.*

**Keyword:** Public Opinion, Sentiment Analysis, Clash Of Champions, Support Vector Machine.

### ABSTRAK

*Penelitian ini mengkaji tanggapan penonton terhadap program edukatif Clash of Champions yang diproduksi oleh Ruangguru dan terinspirasi dari format Korea Selatan University War. Melalui konsep kompetisi akademik interaktifnya, Clash of Champions berhasil menarik minat publik, memicu berbagai reaksi di media sosial, dan memberikan dasar untuk melakukan kajian opini guna memahami pandangan penonton. Penelitian ini menggunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan komentar warganet pada platform X (Twitter) menjadi sentimen positif, negatif, dan netral. Proses pengolahan data meliputi pengumpulan data melalui scraping, text preprocessing, dan pembobotan kata menggunakan algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Evaluasi model dengan metrik akurasi, precision, recall, dan F1-score menunjukkan berbagai opini penonton, dengan mayoritas menunjukkan sentimen positif terhadap inovasi edukatif program tersebut. Model ini mencapai akurasi sebesar 79%, precision 79%, recall 95%, dan F1-score 86% untuk klasifikasi sentimen positif. Hal tersebut mengindikasikan apresiasi yang kuat dari publik terhadap nilai edukatif program ini. Penelitian ini diharapkan dapat memberikan wawasan bagi pengembang konten pendidikan untuk menciptakan program yang edukatif dan menarik.*

**Kata Kunci:** Opini Publik, Sentimen Analisis, Clash Of Champions, Support Vector Machine.

### PENDAHULUAN

Perkembangan teknologi di bidang pendidikan telah membuka peluang baru untuk menciptakan metode belajar yang lebih interaktif dan fleksibel. Salah satu contoh nyata dari inovasi ini adalah Ruangguru. Ruangguru hadir untuk meningkatkan pendidikan dan berkolaborasi dengan berbagai pihak untuk mewujudkan pendidikan Indonesia yang lebih efektif dan efisien dalam dunia digital (Haris et al., 2020). Salah satu program unggulan Ruangguru ialah *Clash of Champions*, mengadopsi format reality show Korea Selatan yaitu *University War* sehingga sukses menyita

perhatian penonton. Strategi promosi yang intensif melalui media sosial, menghasilkan peningkatan keterlibatan sebesar 640% dan peningkatan liputan media konvensional sebesar 102%. Faktor kunci keberhasilannya mencakup profil peserta yang berprestasi dan relevansi acara tersebut dengan budaya populer.

*Clash of Champions* dirancang sebagai kompetisi akademik untuk meningkatkan minat belajar siswa melalui pendekatan yang menyenangkan dan interaktif (Helmiyah & Verdian, 2024). Pemilihan tayangan *Clash of Champions* sebagai objek kajian

salah satunya karena tingginya peminat acara tersebut. Tayangan ini memiliki banyak data untuk dipelajari karena jumlah ulasan, komentar, dan tanggapan yang relatif tinggi di media sosial. Akibatnya, analisis sentimen dapat digunakan untuk mengetahui lebih banyak tentang bagaimana tanggapan dan opini dari penonton. *Clash of Champions* berbeda dari program pendidikan lainnya karena menggunakan format *reality show University War* Korea Selatan dan melibatkan peserta dari mahasiswa terbaik dari berbagai universitas.

Analisis sentimen adalah proses memahami dan mengolah data tekstual secara otomatis untuk mendapatkan informasi. Analisis sentimen dilakukan untuk mengkaji opini terhadap suatu subjek dan objek (misalnya individu, organisasi ataupun produk) dalam sebuah kumpulan data (Hasri & Alita, 2022). Dengan memanfaatkan algoritma dan metode statistik seperti *Support Vector Machine*, analisis sentimen dapat menghasilkan wawasan yang mendalam mengenai bagaimana penonton merespons tayangan tertentu, dalam hal ini tayangan *Clash of Champions*.

Berdasarkan penelitian analisis sentimen yang telah ada, terdapat beberapa perbedaan dalam hal algoritma, pemilihan objek serta penambahan fitur pembobotan kata yang digunakan. Pada penelitian terdahulu digunakan algoritma *K-Means Clustering* dan mendapatkan nilai sentimen negatif sebanyak 15,09%, netral 51,75%, dan positif 33,15% (Larasati et al., 2022). Kemudian pada penelitian lain digunakan algoritma *Naïve Bayes Classifier* dalam melakukan klasifikasi sentimen berdasarkan komentar pengunjung restoran. Hasil pengelompokan review kepuasan pengunjung menggunakan algoritma naïve bayes memberikan nilai akurasi sebesar 73% (Permadi, 2020). Selanjutnya pada penelitian yang melakukan klasifikasi dengan algoritma *K-Nearest Neighbor*, total akurasi yang didapatkan sebesar 89,4% dengan hasil klasifikasi pada bulan Februari 2020 pada daerah DKI Jakarta dari total 500 tweets menghasilkan sejumlah sentimen netral sebanyak 426 (85,2%), sentimen negatif sebanyak 71 (14,2%), dan sentimen positif sebanyak 3 (0,6%) (Alrajak et al., 2020).

Penelitian tentang analisis sentimen yang menggunakan berbagai algoritma telah banyak dilakukan, tetapi masih ada perbedaan dalam pendidikan berbasis realitas seperti *Clash of Champions*. Penelitian terdahulu yang menggunakan algoritma *Clustering K-Means* tidak mempelajari algoritma yang lebih canggih karena hanya memahami sedikit tentang sentimen penonton. Selain itu, studi yang lain pada tahun 2020, *K-Nearest Neighbor*, menunjukkan akurasi yang tinggi, tetapi lebih berfokus

pada konteks media sosial daripada tayangan pendidikan. Penelitian lebih lanjut diperlukan karena implementasi algoritma yang lebih efisien seperti *Support Vector Machine* (SVM) dalam analisis sentimen untuk tayangan pendidikan. Dengan menggunakan analisis sentimen yang lebih mendalam terhadap reaksi penonton *Clash of Champions*, penelitian ini bertujuan untuk mengisi celah tersebut dan memberikan gambaran yang lebih luas.

Penelitian ini bertujuan untuk mempresentasikan perbedaan cara pandang penonton dari sebuah tayangan *reality show*, terutama *Clash of Champions*. Dalam penelitian ini, algoritma *Support Vector Machine* (SVM), digunakan untuk mendapatkan klasifikasi pemahaman yang lebih akurat tentang reaksi penonton dalam melihat tayangan tersebut. Penelitian ini memberikan pemahaman yang lebih baik tentang pendapat warganet, yang dapat membantu pendidik dan pengembang konten membuat program yang lebih menarik dan relevan.

## TINJAUAN PUSTAKA

### Analisis Sentimen di Media Sosial

Analisis sentimen merupakan teknik mengidentifikasi isi dari dataset teks berupa opini untuk mendapatkan informasi tentang pandangan yang biasanya ditulis di media sosial yang bersifat positif, negatif, dan netral (Syah & Witanti, 2022). Studi terdahulu menunjukkan bahwa analisis sentimen sangat bermanfaat dalam memberikan gambaran tentang pandangan publik terhadap program pendidikan online.

Saat ini, penggunaan analisis sentimen sangat berkembang pesat dan memberikan banyak keuntungan di berbagai aspek, salah satunya dalam mengkaji opini public terhadap tayangan edukatif seperti *Clash of Champions* di twitter.

### *Support Vector Machine*

Untuk mendapatkan informasi dari sekumpulan data teks diperlukan pengklasifikasian teks menjadi beberapa kategori dari data yang sudah dimiliki (Oryza Habibie Rahman et al., 2021). Pengklasifikasian teks memiliki beberapa metode, salah satunya adalah *Support Vector Machine*.

*Support Vector Machine* merupakan model pembelajaran statistika yang memiliki hasil lebih baik daripada metode lainnya (Darwis et al., 2020). Metode ini diciptakan oleh Guyon Boser untuk melakukan klasifikasi dengan teknik regresi. Pengoptimalan akurasi data pada *Support Vector Machine* ini dilakukan dengan cara menghindari overfit pada data secara otomatis (Zoellanda A.Tane et al., 2019).

Penelitian sebelumnya tentang implementasi twitter sentimen analisis untuk review film menggunakan algoritma *Support Vector Machine* menunjukkan bahwa memudahkan penelitian dalam mengklasikasikan dengan melihat opini positif, negatif, dan netral (Faisal Rahutomo et al., 2018).

### **Clash of Champions**

*Clash of Champions* merupakan acara pertandingan yang dikembangkan oleh Ruangguru dan hadir dalam inovasi baru dalam tayangan edukatif di Indonesia. *Clash of Champions* merupakan versi Indonesia yang diadaptasi dari program reality show Korea selatan, *University War* (Assalamah, 2024).

Tayangan ini menggabungkan beberapa peserta dari berbagai universitas-universitas Indonesia untuk bertanding dalam satu ajang yang mampu menarik perhatian publik dari berbagai kalangan, mulai dari kalangan muda hingga tua. Penelitian ini dilakukan untuk mengklasifikasikan pendapat/opini dari penonton yang bervariasi. Pengklasifikasian sentimen terbagi menjadi tiga, yaitu positif, negatif, dan netral.

### **Algoritma Term Frequency Inverse Document Frequency (TF IDF)**

Algoritma TF-IDF adalah sebuah metode yang digunakan untuk menghitung bobot setiap kata yang paling sering muncul dalam suatu kumpulan data. TF-IDF berfungsi untuk menentukan seberapa penting sebuah kata dalam sebuah dokumen, sehingga proses ini menjadi lebih efisien, mudah, dan mampu menghasilkan akurasi tinggi (Putri Gabriella, 2023)

Metode ini menggabungkan dua konsep utama dalam pembobotan, yaitu frekuensi kata dalam dokumen tertentu dan kebalikan dari frekuensi dokumen yang memuat kata tersebut. Semakin sering suatu kata muncul dalam dokumen tertentu, semakin penting kata tersebut di dalam konteks dokumen tersebut. Dua komponen inti dari TF-IDF adalah TF (*Term Frequency*), yang mengukur seberapa sering kata muncul dalam dokumen, dan IDF (*Inverse Document Frequency*), yang menghitung kebalikan dari frekuensi kemunculan kata tersebut di seluruh dokumen (Thomas & Rumaisa, 2022).

### **Klasifikasi**

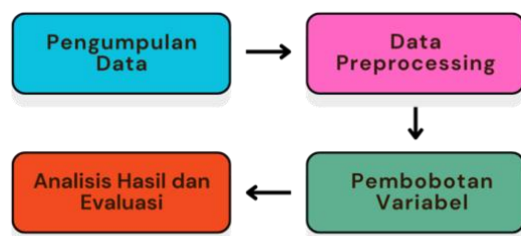
Klasifikasi adalah salah satu teknik utama dalam data mining yang berfungsi untuk mengelompokkan data ke dalam kategori atau kelas tertentu berdasarkan atribut-atribut yang dimilikinya. Teknik ini bertujuan membangun model yang dapat memprediksi kategori atau label dari data yang belum diketahui. Klasifikasi memiliki kegunaan yang besar dalam membantu proses

pengambilan keputusan serta analisis prediktif di berbagai sektor (Setiawan et al., 2023)

Ada banyak jenis algoritma umum dalam klasifikasi dalam memprediksi kategori data seperti Naive Bayes classifier, Random Forest, Decision Tree, K-Nearest Neighbor (KNN) dan yang digunakan pada penelitian ini adalah algoritma Support Vector Machine (Purnamawati et al., 2020).

### **METODE PENELITIAN**

Penelitian ini memiliki empat tahapan utama yang dilakukan secara berurutan seperti pada Gambar 1 yaitu dimulai dari pengumpulan data, tahap *preprocessing*, pembobotan variabel menggunakan TF-IDF, serta evaluasi model, Penjelasan lebih lanjut terkait tahapan dalam penelitian dijelaskan pada bab ini.



Gambar 1. Langkah-langkah Penelitian

### **Pengumpulan Data**

Data yang digunakan dalam penelitian ini diambil dari *platform* media sosial Twitter atau yang saat ini dikenal dengan nama X menggunakan teknik *scraping*, teknik ini digunakan untuk mendapatkan sebuah tweet yang relevan terhadap topik yang diambil (Ghozali et al., 2023a). Pengumpulan data dilakukan dengan rentang waktu tertentu dengan mengambil momentum saat tayangan Clash of Champions sedang hangat dibicarakan dikalangan masyarakat terutama di dunia maya.

### **Data pre-Processing**

Data Pre-processing dilakukan untuk merubah kata kata atau beberapa kata asing ke Bahasa Indonesia dan membuang beberapa karakter yang tidak perlukan (Ghozali et al., 2023b). Data yang diambil adalah data mentah yang tidak diperlukan pada saat mengelola data. Mencakup Langkah-langkah *case folding*, *tokenizing*, *normalisasi*, *filtering*, dan *stemming*.

Dalam pemrosesan kata menggunakan *Natural Language Processing* (NLP), terdapat tiga langkah utama. Pertama, tokenisasi yang berfungsi untuk memecah kalimat menjadi kumpulan kata. Kedua, penyaringan yang bertujuan untuk menghilangkan kata-kata penghubung dan kata depan dari kumpulan

kata yang diperoleh. Ketiga, *stemming* yang digunakan untuk menghapus imbuhan. *Text mining* adalah proses penggalian pengetahuan dalam dokumen teks yang dapat mengekstraksi wawasan berharga dari teks yang ada (Putri Gabriella, 2023).

**Pembobotan Variabel *Term Frequency Inverse Document Frequency (TF IDF)***

Algoritma TF-IDF adalah sebuah metode yang digunakan untuk menghitung bobot setiap kata yang sering dipakai dalam *information retrieval*. *Term Frequency-Inverse Document Frequency* merupakan ukuran statistik yang berfungsi untuk menilai tingkat pentingnya sebuah kata dalam sebuah dokumen atau dalam kumpulan dokumen. Frekuensi kemunculan kata dalam dokumen tertentu mencerminkan seberapa signifikan kata tersebut di dalam dokumen tersebut, sedangkan frekuensi dokumen yang memuat kata tersebut menunjukkan seberapa sering kata itu muncul secara umum (Putri Gabriella, 2023).

**Evaluasi Model**

Pada tahap evaluasi model, kinerja klasifikasi sentimen menggunakan *Support Vector Machine (SVM)* diukur melalui beberapa metrik evaluasi utama. Akurasi digunakan untuk menghitung persentase prediksi yang benar dari total data uji, sedangkan *classification report* yang mencakup *precision*, *recall*, dan *F1-score* memberikan analisis yang lebih mendalam terkait kemampuan model dalam memprediksi setiap kelas sentimen, seperti positif, netral, dan negatif. Selain itu, untuk memvisualisasikan distribusi kesalahan yang terjadi, digunakan *confusion matrix*. *Confusion matrix* merupakan alat evaluasi yang digunakan dalam metode prediksi dan klasifikasi, yang bertujuan untuk mengukur kinerja model dengan menampilkan hasil klasifikasi yang benar (*True Positive* dan *True Negative*) serta kesalahan klasifikasi (*False Positive* dan *False Negative*) (Han, J., Pei & Kamber, 2012).

**HASIL DAN PEMBAHASAN**

**Pengumpulan Data**

Pengumpulan data penelitian ini dilakukan melalui *crawling* data Twitter yang dilakukan sebanyak dua kali dengan kata kunci “coc ruangguru” dan “*clash of champions*”. Pengumpulan data ini dilakukan dalam rentang waktu yang sama yaitu periode cuitan sejak 23 juni hingga 23 agustus 2024 dengan total data yang dikumpulkan adalah 882 data. Informasi data dari tahapan ini dapat dilihat pada Tabel 1. Tabel 2 menunjukkan contoh hasil dari pengumpulan data yang telah dilakukan.

**Tabel 1.** Hasil Pengumpulan Data

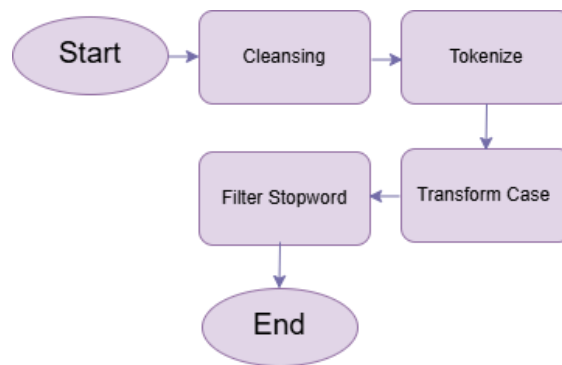
Dataset	Jumlah Data
Coc ruangguru	321
Clash of champions	561

**Tabel 2.** Contoh Data Hasil Crawling

No	Tweet
1	Gw bangga jadi fans COC #ClashofChampions #Ruangguru #CoC
2	Terimakasih kepada kak iman #askIU kak belva dan tim ruangguru yg udh ngasih kesemptn anak <sup>2</sup> hebat coc ini naik panggung
3	@srii hi @tanyarlfs Gk salah woi Ruang Guru nge cast anak anak COC Mereka melek politik ternyata salut

**Preprocessing data**

Tahapan berikutnya adalah preprocessing. Pada proses preprocessing ini terdapat beberapa langkah yang harus dilalui, yaitu cleansing, tokenize, transform case, dan filter stopwords. Setiap langkah tersebut dijelaskan lebih lanjut pada Gambar 2.



**Gambar 2.** Tahap Preprocessing Data

Proses ini melibatkan tahap pembersihan data yang sebelumnya dijelaskan dalam metodologi penelitian. Text yang sepenuhnya diambil dari tahapan pengumpulan data akan diolah untuk dilakukan pebersihan data agar data yang diolah dapat menjadi informasi yang bermanfaat. Hasil pembersihan karakter non-huruf ini ditampilkan dalam Tabel 3.

**Tabel 3.** Contoh Data Hasil *Cleansing*

Text	Cleansing
Gw bangga jadi fans COC #ClashofChampions	Gw bangga jadi fans COC

#Ruangguru #CoC	ClashofChampions Ruangguru CoC	clashofchampions ruangguru coc	'clashofchampions', 'ruangguru', 'coc']
Terimakasih kepada kak iman #askIU kak belva dan tim ruangguru yg udh ngasih kesemptn anak <sup>2</sup> hebat coc ini naik panggung	Terimakasih kepada kak iman askIU kak belva dan tim ruangguru yg udh ngasih kesemptn anak hebat coc ini naik panggung	terimakasih kepada kak iman askiu kak belva tim ruangguru udh ngasih kesemptn anak hebat coc naik panggung	['terimakasih', 'kepada', 'kak', 'iman', 'askiu', 'kak', 'belva', 'tim', 'ruangguru', 'udh', 'ngasih', 'kesemptn', 'anak', 'hebat', 'coc', 'naik', 'panggung']
@srii_hi @tanyarlfe	tanyarlfe Gk salah woi Ruang Guru nge cast anak anak COC Mereka meleak politik ternyata salut	sriihi tanyarlfe gk salah woi ruang guru nge cast anak anak coc mereka meleak politik ternyata salut	['sriihi', 'tanyarlfe', 'gk', 'salah', 'woi', 'ruang', 'guru', 'nge', 'cast', 'anak', 'anak', 'coc', 'meleak', 'politik', 'salut']

Tahapan berikutnya adalah proses *case folding* yang masih termasuk dalam tahap *cleansing* data yaitu mengubah format kata dalam Tweet agar seragam atau mengubahnya menjadi huruf kecil (*lower case*). Hasil dari tahapan tersebut ditampilkan dalam Tabel 4.

**Tabel 4.** Contoh Data Hasil *Case Folding*

<i>Text</i>	<i>Case Folding</i>
Gw bangga jadi fans COC Clash of Champions Ruangguru CoC	gw bangga jadi fans coc clash of champions ruangguru coc
Terimakasih kepada kak iman askIU kak belva dan tim ruangguru yg udh ngasih kesemptn anak hebat coc ini naik panggung	terimakasih kepada kak iman askiu kak belva dan tim ruangguru yg udh ngasih kesemptn anak hebat coc ini naik panggung
tanyarlfe Gk salah woi Ruang Guru nge cast anak anak COC Mereka meleak politik ternyata salut	tanyarlfe gk salah woi ruang guru nge cast anak anak coc mereka meleak politik ternyata salut

Tahap selanjutnya adalah *tokenize*. *Tokenize* adalah proses membagi teks menjadi bagian token (Alghifari & Juardi, 2021). Pada tahapan ini dilakukan pembagian setiap teks. Contoh data hasil tokenisasi ditunjukkan pada Tabel 5.

**Tabel 5.** Contoh Data Hasil *Tokenize*

<i>Text</i>	<i>Tokenize</i>
gw bangga jadi fans coc	['gw', 'bangga', 'jadi', 'fans', 'coc',

Pada tahap ini dilakukan filter stopwords, yaitu proses menghapus kata-kata menggunakan kamus Bahasa Indonesia. Filter stopwords dilakukan dengan memilih kata-kata yang akan dipakai, serta hanya mempertahankan kata-kata penting dari token yang dihasilkan untuk merepresentasikan dokumen. Contoh tampilan data setelah filter stopwords terdapat pada Tabel 6.

**Tabel 6.** Contoh Data Hasil *filter Stopwords*

<i>Text</i>	<i>Filter Stopwords</i>
['gw', 'bangga', 'jadi', 'fans', 'coc', 'clashofchampions', 'ruangguru', 'coc']	['bangga', 'fans', 'coc', 'clashofchampions', 'ruangguru', 'coc']
['terimakasih', 'kepada', 'kak', 'iman', 'askiu', 'kak', 'belva', 'tim', 'ruangguru', 'udh', 'ngasih', 'kesemptn', 'anak', 'hebat', 'coc', 'naik', 'panggung']	['terimakasih', 'tim', 'ruangguru', 'ngasih', 'kesemptn', 'anak', 'hebat', 'coc', 'panggung']
['sriihi', 'tanyarlfe', 'gk', 'salah', 'woi', 'ruang', 'guru', 'nge', 'cast', 'anak', 'anak', 'coc', 'meleak', 'politik', 'salut']	['salah', 'ruang', 'guru', 'cast', 'anak', 'anak', 'coc', 'politik', 'salut']

**Term Frequency Inverse Document Frequency (TF IDF)**

Setelah melalui proses *preprocessing*, data teks akan dikonversi menjadi vektor numerik menggunakan metode *Term Frequency-Inverse Document Frequency*

(TF-IDF). Metode ini menentukan bobot setiap kata berdasarkan frekuensinya dalam satu dokumen dibandingkan dengan seluruh dokumen lain dalam kumpulan data, sehingga menunjukkan pentingnya kata tersebut dalam konteks analisis.

Vektor hasil dari proses TF-IDF ini kemudian akan dioptimalkan dan diklasifikasikan dengan bantuan algoritma *Support Vector Machine* (SVM). SVM adalah algoritma pembelajaran mesin yang sangat efektif dalam melakukan klasifikasi, terutama untuk masalah binary classification seperti analisis sentimen. Algoritma ini mencari hyperplane terbaik yang mampu memisahkan kelas positif, netral dan negatif dengan jarak maksimum. Dengan begitu, SVM akan menghasilkan prediksi kategori sentimen teks berdasarkan nilai terbesar yang diperoleh dari keputusan hyperplane yang telah terbentuk. Gambar 3 menunjukkan perhitungan dari TF-IDF.

term	TF	TF-IDF
nebak	0.0625	0.37810834095199736
kereeen	0.0625	0.37810834095199736
penonton	0.0625	0.33478664216700077
one	0.0625	0.32084017020986266
shakiraaa	0.0625	0.29981065542103685
day	0.0625	0.2715615851870958
banget	0.125	0.2659700148688305
clashofchampions	0.0625	0.24078180486848366
menang	0.0625	0.2254616512414221
selamat	0.0625	0.18782568859428342
keren	0.0625	0.16765235157784272
acara	0.0625	0.14677696756996653
ruangguru	0.125	0.14213482118698822
coc	0.0625	0.04902849642263258

Gambar 3. Perhitungan TF-IDF

**Evaluasi Model**

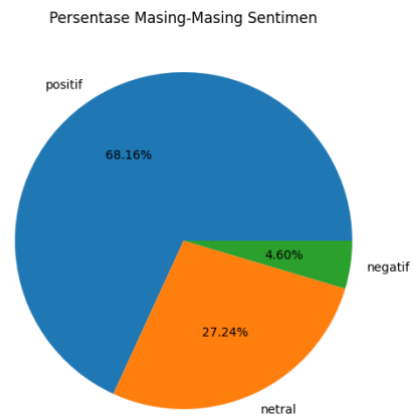
Tahap Selanjutnya ialah evaluasi model, akurasi digunakan untuk menghitung persentase prediksi yang benar dari total data uji, berikut table hasil evaluasi yang mencakup *precision*, *recall*, dan *F1-score*. Tabel 7 menunjukkan ringkasan dari hasil evaluasi secara keseluruhan.

**Tabel 7. Hasil Evaluasi**

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Positif	79 %	95 %	86%
<i>Acuracy</i>			79%

Berdasarkan hasil evaluasi ini, model *Support Vector Machine* (SVM) menunjukkan hasil optimal dalam mendeteksi sentimen positif pada topik *Clash of Champions*. Pada kategori ini, model mencapai *precision* sebesar 79%, artinya 79% dari semua prediksi yang dikategorikan sebagai sentimen positif memang benar, sementara *recall* yang mencapai 95% menunjukkan bahwa model mampu mengenali

sebagian besar data positif yang ada. *F1-score* sebesar 86% mengindikasikan keseimbangan antara *precision* dan *recall*, menegaskan kinerja yang sangat baik dalam mendeteksi sentimen positif.



Gambar 4. Persentase Sentimen

Keberhasilan model dalam mendeteksi sentimen positif ini mengindikasikan bahwa tanggapan publik terhadap *Clash of Champions* secara umum bersifat positif. Hal ini menunjukkan adanya apresiasi dari penonton terhadap sifat acara yang edukatif, sehingga topik ini menarik perhatian dengan dominasi sentimen yang mendukung. Tingginya jumlah respon positif mencerminkan antusiasme masyarakat, yang tidak hanya menganggap acara ini relevan, tetapi juga bermanfaat. Pada Gambar 4 menunjukkan persentase sentimen penonton *Clash of Champions* dengan sentimen positif sebesar 68,16%, sentimen bernada negatif hanya 4,6%, dan sentimen bersikap netral sebesar 27,24%. Sementara itu, perkumpulan kata-kata pada sentimen *Clash of Champions* ini dapat dilihat pada Gambar 5.

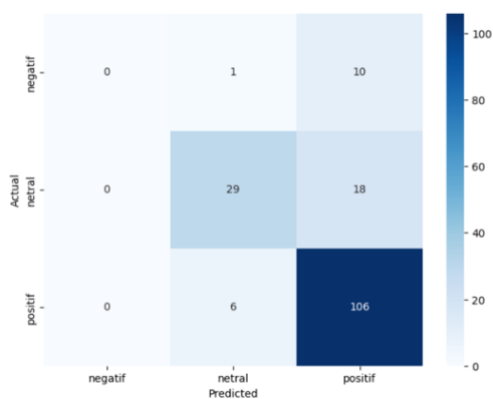


Gambar 5. Kumpulan Sentimen

Secara keseluruhan, persepsi publik terhadap *Clash of Champions* cenderung positif, dengan mayoritas tanggapan memberikan apresiasi yang tinggi terhadap nilai edukatif acara tersebut. Model SVM berhasil mengidentifikasi pola ini dengan akurat, sehingga hasil analisis ini memberikan pemahaman

bahwa acara *Clash of Champions* mendapat respon yang kuat dan didominasi sentimen positif di kalangan.

Confusion matrix berfungsi untuk mengevaluasi performa model klasifikasi dengan menunjukkan distribusi prediksi model terhadap kelas-kelas aktual data (Rosandy, 2016). Matriks ini menampilkan jumlah sampel yang diklasifikasikan dengan benar dan yang salah oleh model, dibagi ke dalam empat elemen utama: *true positives*, *false positives*, *true negatives*, dan *false negatives* (Normawati & Prayogi, 2021). Dengan informasi ini, kita bisa memahami secara spesifik area kekuatan dan kelemahan model dalam mengklasifikasikan setiap kategori.



Gambar 5. Confussion Matrix

Dalam konteks penelitian ini, *confusion matrix* membantu menganalisis seberapa baik model SVM mengenali sentimen positif pada topik *Clash of Champions*. Dari matriks tersebut, terlihat bahwa model berhasil mengklasifikasikan 106 sampel positif secara benar, sedangkan hanya 6 sampel yang salah diklasifikasikan sebagai netral. Dominasi hasil *true positives* pada kategori positif menunjukkan bahwa model memiliki keandalan yang tinggi dalam mendeteksi sentimen positif, sekaligus memperlihatkan bahwa kategori ini memang paling dominan di dalam dataset. Hasil dari *confusion matrix* dapat dilihat pada Gambar 6.

Melalui penggunaan *confusion matrix*, kita dapat menilai bahwa model SVM lebih efektif dalam mengenali sentimen positif dibandingkan kategori lainnya. Matriks ini juga berguna dalam menghitung metrik evaluasi lainnya, seperti *precision*, *recall*, dan *F1-score*, yang memberikan pemahaman mendalam mengenai performa model pada tiap kategori.

## KESIMPULAN

Hasil dari kajian ini menunjukkan bahwa penerapan algoritma *Support Vector Machine* (SVM) dalam analisis opini publik terhadap acara edukatif *Clash of Champions* yang diproduksi oleh

Ruangguru berhasil memberikan pemahaman yang lebih mendalam tentang respons penonton. Dengan memanfaatkan analisis opini publik, penelitian ini mengelompokkan umpan balik penonton ke dalam tiga kategori utama: positif, negatif, dan netral. Hasil klasifikasi menunjukkan bahwa sebagian besar penonton memberikan respons positif, mencerminkan apresiasi terhadap pendekatan inovatif dan interaktif yang diterapkan dalam Pendidikan. Metode pra-pemrosesan yang digunakan, termasuk pembersihan teks, pengubahan huruf kecil, tokenisasi, serta penghapusan kata-kata yang tidak relevan melalui *stopwords*, terbukti efektif dalam menyederhanakan data dari media sosial. Langkah-langkah ini berkontribusi pada peningkatan akurasi model, sehingga SVM dapat mengklasifikasikan opini publik dengan efisien. Secara keseluruhan, kajian ini mengisi kekosongan dalam literatur terkait studi opini publik pada program pendidikan dengan pendekatan kompetisi. Selain itu, penelitian ini memberikan kontribusi bagi pengembangan konten edukatif yang lebih menarik dan relevan, serta mampu meningkatkan keterlibatan penonton dalam proses belajar. Hasil kajian ini diharapkan dapat menjadi referensi bagi pengembang konten dan pendidik dalam merancang program edukatif yang menarik dan memiliki daya tarik kuat bagi kalangan muda. Selain itu, penelitian ini juga menunjukkan potensi penerapan algoritma canggih seperti SVM dalam konteks pendidikan digital, yang diharapkan dapat memperkuat inovasi metode pembelajaran interaktif di Indonesia di masa depan.

## DAFTAR PUSTAKA

- Alghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 9(02), 75–81. <https://doi.org/10.33884/jif.v9i02.3755>
- Alrajak, M. S., Ernawati, I., & Nurlaili, I. (2020). Analisis sentimen terhadap Pelayanan PT PLN di Jakarta pada Twitter dengan Algoritma K-Nearest Neighbor (K-NN). *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 1(2), 110–122.
- Assalamah, T. M. (2024). *Strategi Viral Marketing Melalui Konten Edutainment Clash of Champions by Ruangguru Viral Marketing Strategy Through Edutainment Content Clash of Champions by Ruangguru*. 7(3). <https://doi.org/10.1080/019722409032.....154>
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Eduatic - Scientific Journal of Informatics*

- Education*, 7(1), 1–11.  
<https://doi.org/10.21107/edutic.v7i1.8779>
- Faisal Rahutomo, Saputra, P. Y., & Fidyawan, M. A. (2018). Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine. *Jurnal Informatika Polinema*, 4(2), 93.  
<https://doi.org/10.33795/jip.v4i2.152>
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023a). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1340–1348.  
<https://doi.org/10.30865/klik.v3i6.936>
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023b). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1340–1348.  
<https://doi.org/10.30865/klik.v3i6.936>
- Han, J., Pei, J., & Kamber, M. (2012). *Techniques, Data Mining: Concepts and* (3rd ed.). Morgan Kaufmann.
- Haris, C. A., WA, B. S., & Nasiri, A. (2020). Evaluasi Aplikasi Ruang Guru Menggunakan Model UTAUT2 dan Model Kesuksesan DeLone and McLean. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, 4(1), 1.  
<https://doi.org/10.30872/jurti.v4i1.3043>
- Hasri, C. F., & Alita, D. (2022). Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 3(2), 145–160.  
<https://doi.org/10.33365/jatika.v3i2.2026>
- Helmiyah, S., & Verdian, A. (2024). Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP dan Model BERT: Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP. *Jurnal Pendidikan Rosalia*, 7(2), 138–149.
- Larasati, M. A. Z., Winarsih, N. A. S., Rohman, M. S., & Saraswati, G. W. (2022). Penerapan Metode K-Means Clustering Dalam Menganalisis Sentimen Masyarakat Terhadap K-Popers Pada Twitter. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 201.  
<https://doi.org/10.35889/progresif.v18i2.877>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Oryza Habibie Rahman, Gunawan Abdillah, & Agus Komarudin. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 17–23.  
<https://doi.org/10.29207/resti.v5i1.2700>
- Permadi, V. A. (2020). Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura. *Jurnal Buana Informatika*, 11(2), 141–151.  
<https://doi.org/10.24002/jbi.v11i2.3769>
- Purnamawati, A., Nugroho, W., Putri, D., & Hidayat, W. F. (2020). Deteksi Penyakit Daun pada Tanaman Padi Menggunakan Algoritma Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN. *InfoTekJar: Jurnal Nasional Informatika Dan Teknologi Jaringan*, 5(1), 212–215.
- Putri Gabriella, Y. A. (2023). Optimasi Penerimaan Siswa Baru Dengan Penerapan Algoritma Text Mining Dan Tf-Idf. *Journal of Computing and Informatics Research*, 2(3), 110–117.  
<https://doi.org/10.47065/comforch.v2i3.941>
- Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4.5) Untuk Menganalisa Kelancaran Pembiayaan. *2016*, 2(01), 52–62.
- Setiawan, Z., M., P., A. M., P., A. Y. P., A., M., Y., & S., Wijaya, A. (2023). *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia.
- Syah, H., & Witanti, A. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm). *Jurnal Sistem Informasi Dan Informatika (Simika)*, 5(1), 59–67.  
<https://doi.org/10.47080/simika.v5i1.1411>
- Thomas, V. W. D., & Rumaisa, F. (2022). Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF. *Jurnal Media Informatika Budidarma*, 6(3), 1767.  
<https://doi.org/10.30865/mib.v6i3.4218>
- Zoellanda A.Tane, O., Muslim Lhaksmana, K., & Nhita, F. (2019). Analisis Sentimen pada Twitter Tentang Calon Presiden 2019 Menggunakan Metode SVM (Support Vector Machine). *EProceedings of Engineering*, 6(2), 9716–9725.