

ANALISIS SENTIMEN PENGGUNA SOSIAL MEDIA TWITTER/X TERHADAP ACARA CLASH OF CHAMPIONS MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES

Salsabila Azzahra¹, Zulfa Kusuma¹, Dian Eka Ratnawati², Nanang Yudi Setiawan³

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya

Email: ¹salsaaakusuma@ub.ac.id, ²dian_ilkom@ub.ac.id, ³nanang@ub.ac.id

Abstrak

Pemanfaatan media sosial sebagai sumber data untuk memahami persepsi publik menjadi aspek penting bagi banyak organisasi, termasuk Ruangguru dalam menyelenggarakan acara *Clash of Champions*. Analisis sentimen digunakan untuk mengidentifikasi persepsi publik terhadap acara *Clash of Champions*. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap acara *Clash of Champions* menggunakan metode *Multinomial Naïve Bayes* dan melakukan evaluasi melalui *SWOT Analysis*. Metode *Multinomial Naïve Bayes* digunakan untuk mengklasifikasikan sentimen, sementara analisis SWOT dilakukan berdasarkan data yang telah melalui proses *topic modelling* menggunakan *Latent Dirichlet Allocation (LDA)*. Penelitian ini juga bertujuan untuk mengukur performa algoritma *Multinomial Naïve Bayes* dalam klasifikasi sentimen serta mengidentifikasi kekuatan, kelemahan, peluang, dan ancaman yang muncul dari hasil analisis SWOT. Hasil implementasi algoritma menunjukkan performa yang baik, dengan tingkat akurasi mencapai 97%. Analisis SWOT mengungkapkan beberapa temuan utama, yaitu kekuatan pada konten acara yang interaktif, kelemahan pada durasi acara yang terlalu pendek dan kurangnya variasi tantangan, peluang pada meningkatnya minat publik terhadap acara edukatif, serta ancaman berupa ekspektasi publik yang tinggi akibat lonjakan popularitas. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi Ruangguru untuk meningkatkan kualitas acara di masa mendatang dan menjadi panduan bagi organisasi lain yang berencana menyelenggarakan acara serupa.

Kata kunci: Analisis Sentimen, *Multinomial Naïve Bayes*, Analisis SWOT, *Clash of Champions*, Ruangguru

Abstract

The utilization of social media as a data source to understand public perception has become a crucial aspect for many organizations, including Ruangguru in organizing the Clash of Champions event. Sentiment analysis is used to identify public perceptions of the Clash of Champions event. This study aims to analyze public sentiment towards the Clash of Champions event using the Multinomial Naïve Bayes method and to evaluate it through SWOT Analysis. The Multinomial Naïve Bayes method is applied to classify sentiment, while SWOT Analysis is conducted based on data processed through topic modeling using Latent Dirichlet Allocation (LDA). This study also aims to evaluate the performance of the Multinomial Naïve Bayes algorithm in sentiment classification and to identify the strengths, weaknesses, opportunities, and threats revealed from the SWOT analysis. The implementation of the algorithm demonstrates good performance, achieving an accuracy rate of 97%. The SWOT analysis reveals several key findings, including strengths in the interactive event content, weaknesses in the short event duration and limited challenge variety, opportunities in the growing public interest in educational events, and threats from high public expectations due to its surge in popularity. The findings of this study are expected to provide insights for Ruangguru to improve the quality of its events in the future and to serve as a guide for other organizations planning to host similar events.

Keywords: Sentiment Analysis, *Multinomial Naïve Bayes*, SWOT Analysis, *Clash of Champions*, Ruangguru

1. PENDAHULUAN

Pandemi Covid-19 telah memengaruhi

beberapa aspek kehidupan manusia di penjuru dunia, termasuk sektor pendidikan. Kementerian Pendidikan dan Kebudayaan Republik Indonesia

mempublikasikan Surat Edaran Nomor 4 Tahun 2020 yang berisi panduan kebijakan pendidikan selama masa darurat akibat pandemi COVID-19. Kebijakan ini mengatur apabila kegiatan belajar mengajar dilaksanakan dari rumah dengan memanfaatkan metode pembelajaran daring (Dewi, 2020). Pembelajaran daring pada masa pandemi COVID-19 berpeluang memunculkan learning loss pada anak-anak karena sering kali dianggap kurang menarik dan monoton. Learning loss adalah fenomena penurunan keterampilan dan pengetahuan akademis pada siswa (Andriani et al., 2021). Tim peneliti dari Program Inovasi untuk Anak Sekolah Indonesia (INOVASI) bersama Badan Standar, Kurikulum, dan Asesmen Pendidikan (BSKAP) Kemendikbudristek melaksanakan studi terkait hasil belajar siswa selama masa pandemi di sekolah-sekolah mitra. Hasil studi tersebut menunjukkan bahwa setelah satu tahun pandemi, siswa mengalami penurunan hasil belajar yang setara dengan kehilangan 6 bulan pembelajaran literasi serta 5 bulan pembelajaran numerasi jika dibandingkan dengan periode sebelum pandemi. Kondisi ini dapat memengaruhi kemampuan siswa dalam meraih kesuksesan akademis di masa depan (Sukoco et al., 2023).

Dampak learning loss dapat diatasi dengan memfasilitasi pembelajaran daring yang lebih menarik. Ruangguru adalah salah satu perusahaan teknologi paling besar di Indonesia dengan fokus terhadap sektor pendidikan. Instansi ini berdiri pada tahun 2014 oleh Iman Usman bersama Belva Devara, dan hingga saat ini telah memiliki lebih dari 6 juta pengguna. Ruangguru mampu menciptakan pengalaman belajar yang menarik dan efektif dengan menyediakan fitur interaktif, seperti video pembelajaran, kuis, serta forum diskusi online (Manuhuruapon et al., 2024). Salah satu fitur terbaru Ruangguru sebagai respons pembelajaran daring selama pandemi adalah Fitur Adapto. Fitur ini dirancang untuk membantu siswa meningkatkan hasil belajarnya sesuai cara belajar dan kecepatan dalam memahami materi masing-masing individu (Khanifah & Pribadi, 2022).

Ruangguru berinovasi dengan menyelenggarakan acara Clash of Champions pada tahun 2024. Clash of Champions merupakan game show kompetisi yang menghadirkan mahasiswa berprestasi dari universitas terkemuka di Indonesia maupun luar negeri. Berdasarkan data Ruangguru yang dipresentasikan oleh Iman Usman sebagai Co-

Founder Ruangguru pada acara IDEAFEST 2024, acara Clash of Champions berhasil mendapatkan 100 juta tayangan dalam 11 episodenya. Hype acara ini terjadi sejak awal Ruangguru mengungkap cast yang akan berkompetisi. Hype tersebut dibuktikan dengan lebih dari 245 ribu postingan pengguna TikTok menggunakan hastag #ruangguru dan #clashofchampion. Iman Usman juga menyebutkan lebih dari 500 ribu konten sosial media yang memeriahkan acara ini dan cast yang berkompetisi.

Clash of Champions berhasil mendongkrak engagement media sosial dan aplikasi Ruangguru. Berdasarkan data Rezky Yayang Yakhamid dalam Medium, menunjukkan bahwa acara ini memperkuat branding Ruangguru di berbagai platform. Instagram Ruangguru memiliki 1.406.245 pengikut pada 28 Juni 2024 sebelum tayangnya Clash of Champions, kemudian meningkat menjadi 1.759.324 pada 17 Agustus 2024 saat grand final Clash of Champions. Total subscriber Ruangguru pada YouTube meningkat dari 1.120.000 sebelum tayangnya Clash of Champions menjadi 1.910.000 dalam satu minggu setelah episode pertama pada 29 Juni 2024. Unduhan aplikasi Ruangguru juga meningkat dari 23.099.650 menjadi 23.922.389 per 17 Agustus 2024. Indeks penelusuran terkait "Ruangguru" dan "Clash of Champions Ruangguru" di Google juga mengalami lonjakan dari di bawah 20 menjadi 41 pada 29 Juni dan 74 pada 30 Juni 2024 (Rezky Yayang, 2023). Data Ruangguru yang dipresentasikan oleh Iman Usman sebagai Co-Founder Ruangguru pada acara IDEAFEST 2024 memperkuat suksesnya acara Clash of Champions. Data tersebut menyatakan setelah tayangan Clash of Champions, learning motivation meningkat sebesar 20% dan tekad untuk mengejar impian meningkat sebesar 28%.

Ramainya pembahasan tentang acara ini berhasil mendapatkan berbagai respons dari publik melalui media sosial setelah dirilis dua episode pertama pada 29 dan 30 Juli 2024. Masyarakat mengungkapkan respons dan pandangan mereka tentang acara *Clash of Champions* melalui berbagai *platform* media sosial, termasuk Twitter/X. Platform ini menyediakan ruang komunikasi digital bagi pengguna untuk berbagi opini dan tanggapan mengenai isu-isu yang sedang ramai diperbincangkan (Emeraldien et al., 2019). Para pengguna dapat mengekspresikan pendapat mereka dalam bentuk teks, gambar, atau video.

Twitter/X memainkan peran penting dalam menampung persepsi publik di Indonesia, termasuk mengenai acara *Clash of Champions*. Besarnya antusiasme terhadap acara ini menciptakan beragam persepsi di kalangan pengguna Twitter/X, mulai dari pendapat positif hingga kritik. Twitter/X menjadi platform yang ideal untuk analisis sentimen dalam acara *Clash of Champions* karena popularitasnya sebagai sumber data yang kaya akan opini.

Analisis sentimen merupakan teknik yang efektif untuk memahami dan mengelompokkan opini publik menjadi sentimen positif, netral, atau negatif (Hasri & Alita, 2022). Teknik ini memberikan pemahaman mendalam tentang persepsi atau emosi yang disampaikan melalui teks, seperti ulasan, komentar, atau postingan di media sosial. Proses analisis sentimen melibatkan beberapa tahapan penting seperti pembersihan data, tokenisasi, dan analisis teks untuk mengelompokkan sentimen dalam kategori positif, netral, dan negatif (Amrullah et al., 2020). Analisis yang lebih mendalam mengenai sentimen yang ada diperlukan untuk memahami bagaimana publik merespons acara *Clash of Champions* secara keseluruhan.

Metode Machine Learning digunakan untuk mengklasifikasikan sentimen pada pengguna sosial media Twitter/X terhadap acara *Clash of Champions*. Metode Multinomial Naïve Bayes (MNB) digunakan dalam penelitian ini untuk mengklasifikasikan sentimen publik ke dalam kategori positif dan negatif. Pemilihan metode ini didasarkan pada karakteristik Multinomial Naïve Bayes yang sangat cocok untuk menganalisis data diskrit, seperti jumlah atau frekuensi kemunculan kata dalam dokumen teks (Ernayanti et al., 2023). Metode ini bekerja dengan prinsip probabilistik sederhana yang menghitung kemungkinan suatu sentimen berdasarkan distribusi kata dalam data. Metode ini dipilih karena efisiensinya, akurasi yang tinggi dalam klasifikasi teks, serta kesesuaian dengan karakteristik data Twitter/X yang berupa teks pendek dengan variasi kata yang besar.

Metode Topic Modelling dengan pendekatan Latent Dirichlet Allocation (LDA) diterapkan dalam penelitian ini untuk mengidentifikasi dan mengelompokkan topik-topik utama yang sering dibahas oleh pengguna Twitter/X terkait acara *Clash of Champions* (Puspita et al., 2024). LDA digunakan untuk menganalisis pola distribusi kata dalam data teks besar, sehingga mampu mengungkapkan topik-topik tersembunyi yang menjadi fokus utama

dalam persepsi publik (Jelita, 2024). Pendekatan ini memberikan pemahaman terstruktur mengenai berbagai isu yang dibahas, baik positif dan negatif. Analisis ini membantu memahami persepsi publik terhadap acara, mulai dari antusiasme, kritik, hingga ekspektasi pengguna media sosial. Hasil pengelompokan data berdasarkan topik dominan memberikan wawasan mendalam tentang konten persepsi publik.

Beberapa penelitian sebelumnya telah membahas analisis sentimen dengan beragam fokus dan metode yang digunakan. Penelitian pertama menganalisis sentimen pengguna aplikasi Ruangguru di Google Play Store dengan memanfaatkan Multinomial Naïve Bayes Classifier yang dipadukan dengan metode Levenshtein Distance untuk meningkatkan akurasi dan kualitas data. Hasil penelitian mengungkapkan bahwa model tersebut berhasil mencapai rata-rata akurasi sebesar 88,20% (Mubaroroh et al., 2022). Penelitian berikutnya menganalisis sentimen masyarakat terhadap aplikasi PeduliLindungi di Google Play Store dengan menerapkan Naïve Bayes Classification yang dikombinasikan dengan Information Gain sebagai metode seleksi fitur demi menaikkan tingkat model. Hasil penelitian ini menunjukkan bahwa nilai akurasi yang dicapai adalah 95,1%, dihitung berdasarkan rata-rata F1-Score menggunakan metode 10-fold cross validation (Syafitri Kustanto et al., 2022). Penelitian lain menganalisis sentimen komentar pengguna aplikasi Tokopedia di Google Play Store dengan metode Naïve Bayes. Hasilnya menunjukkan bahwa model berhasil mencapai akurasi sebesar 97,13%, dengan precision sebesar 1 serta recall sebesar 95,49% (Apriani & Gustian, 2019).

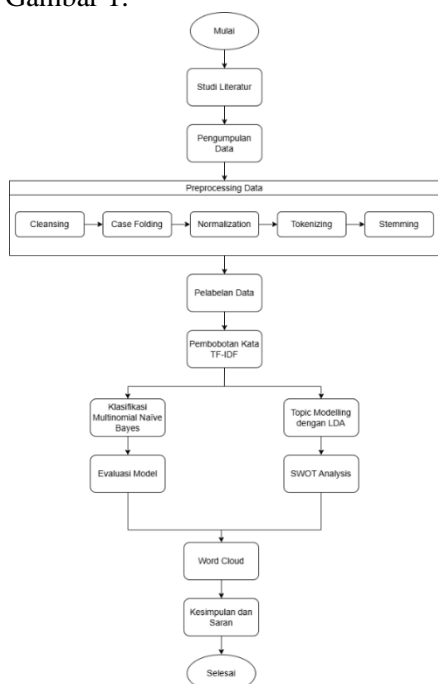
Penelitian ini memiliki perbedaan yang signifikan dibandingkan dengan penelitian-penelitian sebelumnya, terutama dalam pendekatan, fokus, atau metode yang digunakan untuk menganalisis data. Penelitian sebelumnya lebih berfokus pada analisis sentimen terhadap aplikasi di Google Play Store, seperti Ruangguru, PeduliLindungi, dan Tokopedia. Metode pada beberapa penelitian sebelumnya mencakup Multinomial Naïve Bayes Classifier dengan Levenshtein Distance, Naïve Bayes Classification dengan Information Gain, serta Naïve Bayes sederhana. Fokus penelitian ini adalah analisis sentimen terhadap acara *Clash of Champions* yang diselenggarakan oleh Ruangguru, topik yang belum pernah dibahas sebelumnya. Metode dalam penelitian ini adalah

kombinasi antara Multinomial Naïve Bayes serta analisis SWOT untuk mengevaluasi kekuatan (strengths), kelemahan (weaknesses), peluang (opportunities), serta ancaman (threats) dalam acara tersebut. Perbedaan konteks penelitian dari sisi metodologi yang digunakan diharapkan menjadi perspektif baru dalam melakukan penelitian.

Kontribusi penelitian ini terletak pada analisis sentimen yang dilengkapi dengan SWOT analysis untuk meningkatkan pemahaman terhadap persepsi publik terhadap acara Clash of Champions yang diselenggarakan oleh Ruangguru. Metode Multinomial Naïve Bayes diterapkan untuk mengklasifikasikan sentimen dari persepsi publik di media sosial Twitter/X menjadi kategori sentimen positif dan negatif. Topic modelling dengan Latent Dirichlet Allocation (LDA) dilakukan untuk mengidentifikasi topik yang muncul dalam data. Hasil dari topic modelling tersebut kemudian dianalisis lebih lanjut menggunakan SWOT analysis untuk menentukan kekuatan, kelemahan, peluang, dan ancaman. Kombinasi kedua metode ini diharapkan dapat membantu Ruangguru meningkatkan kualitas acara di masa mendatang dan memberikan pedoman bagi organisasi lain yang ingin menyelenggarakan acara serupa.

2. METODE PENELITIAN

Tahapan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1. Pengumpulan Data

Data persepsi publik terhadap acara *Clash of Champions* Ruangguru dikumpulkan menggunakan teknik *data crawling* menggunakan *library TweetHarvest*. Total data yang berhasil dikumpulkan sebanyak 1.173 *tweet*, yang diambil dalam rentang waktu dari bulan Juni hingga bulan Oktober.

2.2. Preprocessing Data

Tahap *preprocessing text* mencakup serangkaian langkah yang dilakukan sebelum analisis untuk memastikan data dalam kondisi bersih dan siap diolah. Tahap *preprocessing* dalam penelitian ini mencakup lima langkah utama, yakni *cleansing*, *case folding*, *normalization*, *tokenizing*, serta *stemming*. Setiap tahapannya akan dilakukan proses yang berbeda, sehingga data yang dihasilkan adalah data bersih dan terstruktur sesuai dengan kebutuhan penelitian (Khairunnisa et al., 2021).

2.4. Pelabelan Data

Proses pelabelan data dilakukan untuk mengkategorikan data persepsi menjadi dua label, yakni positif serta negatif. Pelabelan dilakukan menggunakan kamus *lexicon* dengan bantuan *library InSet*. Setelah proses pelabelan menggunakan *lexicon*, hasilnya divalidasi oleh aktor pelabelan data. Untuk menghindari bias. Validasi dilakukan oleh tiga aktor berbeda, sehingga memastikan hasil pelabelan lebih objektif dan akurat.

2.3. Pembobotan TF-IDF

Pembobotan TF-IDF diterapkan untuk mengidentifikasi kata-kata penting pada sekumpulan dokumen dengan mempertimbangkan frekuensi kemunculan kata (term). Proses ini bertujuan dalam menilai seberapa penting sebuah term pada suatu dokumen daripada seluruh dokumen (Septian et al., 2019). Pembobotan TF-IDF dapat dihitung menggunakan Persamaan 1.

$$TF - IDF = TF(t, d) \times \log \left(\frac{N}{DF(t)} \right) \quad (1)$$

Keterangan:

$TF(t, d)$: Frekuensi relatif kata t dalam dokumen d .

N : Jumlah total dokumen dalam koleksi.

$DF(t)$: Jumlah dokumen yang mengandung term t .

2.5. Multinomial Naïve Bayes

Naïve Bayes merupakan metode klasifikasi berbasis probabilitas yang efektif untuk analisis teks. Dalam penelitian ini diterapkan variasi *Multinomial Naïve Bayes*, yang lebih sesuai bagi data dengan distribusi kategori seperti frekuensi kata. Algoritma ini dirancang khusus untuk menangani tugas-tugas klasifikasi teks, seperti analisis sentimen dan pengelompokan dokumen. Perhitungan dalam *Multinomial Naïve Bayes* meliputi perhitungan nilai probabilitas kategori (*prior*), menghitung artibut (*likelihood*), dan *posterior* untuk data uji (Ernayanti et al., 2023). Perhitungan Multinomial Naïve Bayes dilakukan menggunakan persamaan 2, 3 dan 4.

$$\hat{P}(c) = \frac{N_c}{N'} \quad (2)$$

$$\hat{P}(t|c) = \frac{w_{ct}+1}{(\sum_w w_{ct}+B')} \quad (3)$$

$$C_{map} = arg \max_{c \in C} \hat{P}(c) \times \prod_{k=1}^n \hat{P}(t_k|c) \quad (4)$$

Keterangan:

- C : Kumpulan semua kelas.
- $\hat{P}(c)$: Probabilitas *prior* dari kelas c .
- $\hat{P}(t_k|c)$: Probabilitas kata t_k (fitur ke- k) muncul di kelas c .
- n : Jumlah kata dalam dokumen.
- N_c : Jumlah dokumen yang berada dalam kelas c .
- N' : Total jumlah dokumen dalam data pelatihan.
- w_{ct} : Jumlah kehadiran kata t dalam dokumen kelas c .
- $\sum_w w_{ct}$: Total kehadiran seluruh kata dalam kelas c .
- B' : Jumlah seluruh kata unik dalam korpus

2.6. Confusion Matrix

Confusion matrix ialah tabel yang biasa diterapkan dalam menganalisis akurasi suatu metode klasifikasi dalam memprediksi kelas data, dengan menampilkan jumlah prediksi yang benar serta salah untuk masing-masing kelas (Gifari et al., 2022). Tabel *confusion matrix* disajikan dalam Tabel 1.

Tabel 1. Confusion Matrix

	Prediksi Positif	Prediksi Negatif
Aktual Positif	True Positive (TP)	False Negative (FN)
Aktual Negatif	False Positive (FP)	True Negative (TN)

	True Positive (TP)	False Negative (FN)
Aktual Positif	True Positive (TP)	False Negative (FN)
Aktual Negatif	False Positive (FP)	True Negative (TN)

Precision ialah metrik yang diterapkan dalam pengukuran terhadap tingkat ketepatan prediksi positif yang dilakukan oleh model.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Accuracy ialah metrik yang diterapkan dalam melakukan pengukuran terhadap tingkat ketepatan model dalam memprediksi nilai (positif maupun negatif) dibandingkan dengan keseluruhan data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Recall merupakan metrik yang diterapkan dalam melakukan pengukuran terhadap sejauh mana model mampu mendeteksi semua data positif yang sesungguhnya.

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

F-measure merupakan metrik yang dihitung sebagai *harmonic mean* dari nilai *precision* dan *recall* untuk menyeimbangkan kedua metrik tersebut dalam satu nilai kinerja model.

$$F - measure = 2 \times \left(\frac{precision \times recall}{precision + recall} \right) \quad (8)$$

2.7. Topic Modelling dengan Latent Dirichlet Allocation

Topic Modeling dengan *Latent Dirichlet Allocation* (LDA) model probabilistik yang berfungsi untuk mengidentifikasi dan mempelajari topik-topik dalam dokumen serta mengelompokkan kata-kata yang berhubungan terhadap topik khusus. Metode ini akan mengelompokkan data berdasarkan kesamaan topik (Syaifuddin et al., 2021). Hasil dari proses ini adalah daftar topik dengan data yang memiliki kesamaan tema atau topik didalamnya. Daftar topik tersebut selanjutnya akan dilakukan untuk analisis SWOT.

2.8. SWOT Analysis

Analisis SWOT adalah metode yang diterapkan dalam melakukan identifikasi terhadap berbagai faktor secara sistematis guna merumuskan strategi perusahaan. Metode ini mengevaluasi kekuatan (*strengths*), kelemahan (*weaknesses*), peluang (*opportunities*), serta ancaman (*threats*) yang memengaruhi sebuah proyek atau bisnis. Kekuatan dan kelemahan bersumber dari faktor internal perusahaan,

seperti sumber daya, kapasitas produksi, dan keahlian karyawan, sementara peluang serta ancaman lebih berhubungan dengan faktor eksternal, seperti kondisi pasar, persaingan, serta perubahan regulasi atau tren industri (Mashuri & Nurjannah, 2020).

2.9. Word Cloud

Word cloud ialah representasi visual teks yang sering muncul pada dataset. Proses ini dilakukan untuk memudahkan dalam memahami pola dan tren dari kumpulan dokumen yang besar dengan menampilkan kata-kata berdasarkan frekuensinya. Kata yang lebih sering muncul dalam dokumen akan dipaparkan dengan ukuran yang lebih besar, menyajikan visualisasi yang jelas mengenai frekuensi kata melalui representasi grafis (Kabir et al., 2020).

3. HASIL DAN PEMBAHASAN

3.1. Pelabelan Data

Pelabelan data dilakukan dengan menggunakan kamus *lexicon* dari *library InSet*. Jenis pelabelan yang dilakukan adalah pelabelan sentimen positif dan negatif. Hasil pelabelan menggunakan kamus *lexicon* dari 1.173 data disajikan pada Tabel 2.

Tabel 2. Distribusi Sentimen

Positif	Negatif
1.010	163

3.2 Evaluasi Model

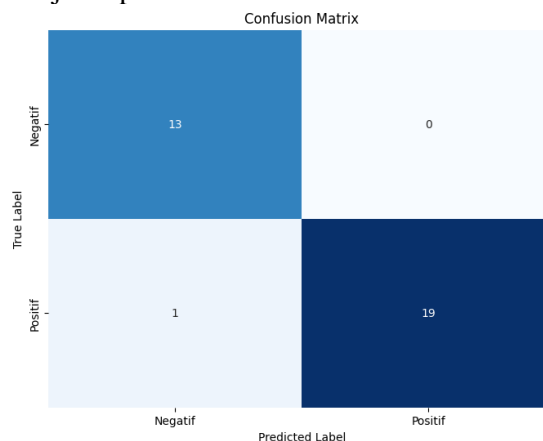
Pengujian klasifikasi *Multinomial Naïve Bayes* melibatkan perbandingan persentase data latih dan data uji dengan rasio yang berbeda. Pengujian ini dilakukan dengan 10 iterasi kemudian diambil nilai rata-ratanya. Hasil pengujian persentase disajikan pada Tabel 3.

Tabel 3. Hasil Pengujian Persentase

Pengujian Ke-i	Data Latih : Data Uji	
	90%:10%	80%:20%
1	82%	83%
2	91%	83%
3	85%	86%
4	97%	80%
5	85%	80%
6	82%	82%
7	88%	83%
8	82%	80%
9	88%	82%
10	85%	89%
Mean	86.5%	82.8%

3.3. Confusion Matrix

Pengujian performa algoritma *Multinomial Naïve Bayes* dilakukan untuk mengevaluasi model. Pengujian ini dilakukan dengan *confusion matrix 2x2* yang terdiri dari kelas prediksi dan kelas aktual. Data latih dan data uji dibagi dengan perbandingan 90%:10%, dengan jumlah data latih sebanyak 293 data dan jumlah data uji sebanyak 33 data. Hasil *confusion matrix* disajikan pada Gambar 2.



Gambar 2. Confusion Matrix Multinomial Naïve Bayes

Model berhasil mengklasifikasikan seluruh 13 data negatif sebagai *true negative*. Selain itu, model mengklasifikasikan 19 dari 20 data positif sebagai *true positive* dan 1 data sebagai *false negative*.

3.4. Topic Modelling dengan Latent Dirichlet Allocation (LDA)

Data yang dimiliki selanjutnya dianalisis untuk mengidentifikasi pola-pola tematik yang mungkin berkaitan dengan topik penelitian. Proses analisis dilakukan menggunakan metode *Latent Dirichlet Allocation (LDA)*, yang memanfaatkan pendekatan probabilistik untuk mengelompokkan kata-kata ke dalam sejumlah topik berdasarkan distribusi kata dalam dokumen. Hasil *topic modelling* menggunakan *Latent Dirichlet Allocation (LDA)* disajikan pada Tabel 4.

Tabel 4. Hasil Topic Modelling LDA

Nomor Topik	Kata Kunci dan Bobot	Nama Topik
0	0.074 "of" + 0.073 "clash" + 0.073 "champions" + 0.064 "seru" + 0.050 "banget" + 0.038 "menonton" +	Antusiasme dan Serunya Menonton Clash of Champions

	0.017 “juga” + 0.013 ”yang” + 0.010 “ini” + 0.008 “episode”	
Topik 1	0.016 “tidak” + 0.016 “dan” + 0.014 “of” + 0.013 “clash” + 0.013 “champions” + 0.013 “dari” + 0.009 “di” + 0.008 “saya” + 0.008 “yang” + 0.007 “kalau”	Kritik dan Evaluasi terhadap Aspek Acara
Topik 2	0.029 “champions” + 0.029 “of” + 0.029 “clash” + 0.023 “saya” + 0.022 “yang” + 0.017 “menonton” + 0.015 “tidak” + 0.013 “ini” + 0.012 “di” + 0.011 “tapi”	Pengalaman Pribadi dan Inspirasi dari Acara
Topik 3	0.039 “of” + 0.037 “champions” + 0.037 “clash” + 0.019 “yang” + 0.016 “banget” + 0.015 “saya” + 0.015 “sudah” + 0.015 “dan” + 0.012 “seru” + 0.012 “orang”	Apresiasi dan Dampak Positif terhadap Penonton
Topik 4	0.013 “yang” + 0.008 “anak” + 0.006 “dan” + 0.006 “jadi” + 0.005 “itu” + 0.005 “of” + 0.004 “beasiswa” + 0.004 “males” + 0.004 “kesempatan” + 0.004 “show”	Motivasi dan Harapan terhadap Generasi Muda
Topik 5	0.035 “champions” + 0.035 “of” + 0.035 “clash” + 0.026 “saya” + 0.021 “tidak” + 0.020 “yang” + 0.016 “menonton” + 0.013	Peluang dan Kritik terhadap Format Acara

	“banget” + 0.012 “ini” + 0.011 “jadi”	
Topik 6	0.029 “of” + 0.029 “champions” + 0.029 “clash” + 0.026 “tidak” + 0.024 “bisa” + 0.022 “saya” + 0.012 “yang” + 0.010 “dan” + 0.010 “tapi” + 0.009 “seperti”	Kesan terhadap Peserta dan Prestasi Akademik
Topik 7	0.056 “of” + 0.056 “clash” + 0.056 “champions” + 0.047 “banget” + 0.038 “menonton” + 0.028 “keren” + 0.021 “saya” + 0.021 “seru” + 0.014 “tidak” + 0.012 “orang”	Keseruan dan Popularitas Acara di Kalangan Penonton

3.5. SWOT Analysis

Analisis SWOT dilakukan untuk menggali lebih dalam hasil klasifikasi sentimen dan topik yang telah dihasilkan melalui *topic modeling*. Data yang telah diklasifikasikan melalui analisis topik akan dikelompokkan menjadi empat aspek utama, yaitu *strengths* (kekuatan), *weaknesses* (kelemahan), *opportunities* (peluang), dan *threats* (ancaman). Setiap aspek akan dijelaskan berdasarkan hasil analisis dari topik yang telah diidentifikasi.

A. Strengths (Kekuatan)

Aspek *strengths* atau kekuatan dapat dianalisis berdasarkan sentimen positif yang bersifat internal, yaitu faktor yang berasal dari dalam acara dan dapat dikontrol oleh penyelenggara. Hasil analisis *strengths* ditampilkan pada Tabel 5.

Tabel 5. Hasil Analisis *Strengths*

Topik	Penjelasan
Topik 0: “Antusiasme dan Serunya Clash of Champions”	Acara Clash of Champions berhasil menarik perhatian publik dan menciptakan antusiasme tinggi.
Topik 4: “Apresiasi Dampak Positif	Clash of Champions memberikan inspirasi



Gambar 4. Visualisasi Word Cloud Negatif

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil penelitian dan pengujian yang dilakukan, dapat ditarik kesimpulan sebagai berikut:

1. Implementasi metode *Multinomial Naïve Bayes* dalam klasifikasi data sentimen pengguna Twitter/X terhadap acara *Clash of Champions* menunjukkan performa model yang sangat baik. Pengujian dilakukan menggunakan *confusion matrix* menunjukkan bahwa *Multinomial Naïve Bayes* memberikan hasil yang baik. Pengujian dilakukan dengan perbandingan data latih dan data uji sebesar 90%:10% dan 80%:20%. Masing-masing perbandingan dilakukan perulangan uji sebanyak 10 kali. Rata-rata akurasi pada perbandingan data 90%:10% sebesar 85.9%, sedangkan untuk perbandingan 80%:20% sebesar 82.8%. Hasil pengujian terbaik diperoleh pada perbandingan 90%:10% pada pengujian ke-4 dengan akurasi sebesar 97%. Evaluasi model menghasilkan nilai *precision* sebesar 96.5%, *recall* sebesar 97.5%, dan *F-measure* sebesar 96.5%.
2. Berdasarkan metode *Topic Modelling* menggunakan *Latent Dirichlet Allocation* (LDA) didapatkan beberapa topik yang dapat diklasifikasikan dalam aspek SWOT. Hasil analisis menunjukkan bahwa *Strengths* (kekuatan) dari acara *Clash of Champions* meliputi inspirasi yang diberikan kepada penonton, konten edukatif yang menarik, serta antusiasme publik yang tinggi terhadap acara ini. *Weaknesses* (kelemahan) mencakup durasi tayangan yang dirasa terlalu singkat, dominasi tantangan eksakta yang dianggap kurang inklusif, serta visualisasi soal yang kurang jelas bagi audiens. *Opportunities* (peluang) mencakup potensi pengembangan variasi tantangan, peluang kolaborasi dengan

institusi lain, dan kemampuan menjangkau audiens yang lebih luas. *Threats* (ancaman) meliputi ekspektasi publik yang tinggi, risiko kejenuhan jika format tidak diperbarui, serta kritik terhadap elemen teknis seperti tempat dan editing.

4.2. Saran

Berikut adalah beberapa saran yang dapat penulis berikan :

1. Menggunakan data dari data dari berbagai platform media sosial lainnya untuk memberikan cakupan analisis yang lebih luas.
2. Melakukan klasifikasi pada aspek yang lebih spesifik untuk memberikan wawasan yang lebih mendalam.
3. Menggunakan metode analisis lain dalam penelitian mendatang untuk memperdalam hasil analisis.

5. DAFTAR PUSTAKA

Amrullah, A. Z., Sofyan Anas, A., & Hidayat, M. A. J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal*, 2(1), 40–44. <https://doi.org/10.30812/bite.v2i1.804>

Andriani, W., Subandowo, M., Karyono, H., & Gunawan, W. (2021). Learning Loss dalam Pembelajaran Daring di Masa Pandemi Corona. Seminar Nasional Teknologi Pembelajaran. *Edudikara: Jurnal Pendidikan Dan Pembelajaran*, 1(1), 485–501.

Apriani, R., & Gustian, D. (2019). Analisis Sentimen Dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia. *Jurnal Rekayasa Teknologi Nusa Putra*, 6(1), 54–62. <https://doi.org/10.52005/rekayasa.v6i1.86>

Dewi, W. A. F. (2020). Dampak COVID-19 terhadap Implementasi Pembelajaran Daring di Sekolah Dasar. *Edukatif: Jurnal Ilmu Pendidikan*, 2(1), 55–61. <https://doi.org/10.31004/edukatif.v2i1.89>

Emeraldien, F. Z., Sunarsono, R. J., & Alit, R. (2019). Twitter Sebagai Platform Komunikasi Di Indonesia. *Jurnal Teknologi Dan Informasi*, XIV(1), 21–30. <https://core.ac.uk/download/pdf/234615318.pdf>

Ernayanti, T., Mustafid, M., Rusgiyono, A., &

- Hakim, A. R. (2023). Penggunaan Seleksi Fitur Chi-Square Dan Algoritma Multinomial Naïve Bayes Untuk Analisis Sentimen Pelanggan Tokopedia. *Jurnal Gaussian*, 11(4), 562–571. <https://doi.org/10.14710/j.gauss.11.4.562-571>
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, 2(1), 36–40. <https://doi.org/10.46229/jifotech.v2i1.330>
- 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>
- Jelita, M. (2024). *Text Mining dengan Topic Modelling LDA dari Pertanyaan Gelar Wicara Literasi Perpustakaan Nasional RI*. 31(3), 253–265. <https://doi.org/10.37014/medpus.v3i1i3.5237>
- Kabir, A. I., Ahmed, K., & Karim, R. (2020). Word Cloud and Sentiment Analysis of Amazon Earphones Reviews with R Programming Language. *Informatica Economica*, 24(4/2020), 55–71. <https://doi.org/10.24818/issn14531305/24.4.2020.05>
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *Jurnal Media Informatika Budidarma*, 5(2), 406. <https://doi.org/10.30865/mib.v5i2.2835>
- Khanifah, A., & Pribadi, F. (2022). Peran Aplikasi Ruang Guru Dalam Meningkatkan Minat Belajar Siswa Saat Pandemi Covid-19. *Joyful Learning Journal*, 11(3), 92–99. <https://doi.org/10.15294/jlj.v11i3.52201>
- Manuhuruapon, J., Toganti, N., & Pangalila, T. (2024). Aplikasi Pendidikan Online “Ruang Guru” Sebagai Peningkatan Minat Belajar Generasi Milenial Dalam Pembelajaran Pendidikan Pancasila dan Kewarganegaraan. *Tumoutou Social Science Journal (TSSJ)*, 1(1), 08–14. <https://jurnal.uisu.ac.id/index.php/Bahastr/article/view/3179>
- Mashuri, & Nurjannah, D. (2020). Analisis SWOT Sebagai Strategi Meningkatkan Daya Saing. *JPS (Jurnal Perbankan Syariah)*, 1(1), 97–112. <https://doi.org/10.46367/jps.v1i1.205>
- Mubaroroh, H. H., Yasin, H., & Rusgiyono, A. (2022). Analisis Sentimen Data Ulasan Aplikasi Ruangguru Pada Situs Google Play Menggunakan Algoritma Naïve Bayes Classifier Dengan Normalisasi Kata Levenshtein Distance. *Jurnal Gaussian*, 11(2), 248–257. <https://doi.org/10.14710/j.gauss.v11i2.35472>
- Puspita, E., Shiddieq, D. F., & Roji, F. F. (2024). Pemodelan Topik pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 481–489. <https://doi.org/10.57152/malcom.v4i2.1204>
- Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *Journal of Intelligent System and Computation*, 1(1), 43–49. <https://doi.org/10.52985/insyst.v1i1.36>
- Sukoco, G. A., Arsendy, S., Purba, R. E., & Zulfa, A. H. (2023). *Bangkit Lebih Kuat: Pemulihan Pembelajaran Pasca Pandemi COVID-19 Studi Kasus INOVASI*. i–29. <http://www.inovasi.or.id>
- Syafitri Kustanto, N., Gusriani, N., & Firdaniza. (2022). Analisis Sentimen dengan Metode Klasifikasi Naïve Bayes Dan Seleksi Fitur Information Gain (Studi Kasus: Ulasan Aplikasi PeduliLindungi). *In Search*, 21(02), 134–144.
- Syaifuddin, A., Harianto, R. A., & Santoso, J. (2021). Analisis Trending Topik untuk Percakapan Media Sosial dengan Menggunakan Topic Modelling Berbasis Algoritme LDA. *Journal of Intelligent System and Computation*, 2(1), 12–19. <https://doi.org/10.52985/insyst.v2i1.150>