**Klasifikasi Sentimen Opini Publik Terhadap Tayangan Televisi Pada Media Sosial Twitter**

Rahajeng Febri Shafiyah  
Informatika  
Universitas Muhammadiyah MalangMalang, Indonesia  
rahajengfebri@webmail.umm.ac.id

Umi Nursyafika  
InformatikaUniversitas Muhammadiyah Malang  
Malang, Indonesia umyns@webmail.umm.ac.id

Zahra Sabilla Usman  
InfomatikaUniversitas Muhammadiyah Malang  
Malang, Indonesia zahrasabilla11@webmail.umm.ac.id

*Abstract*—Penelitian ini menyelidiki sentimen publik terhadap acara TV populer di Indonesia, seperti “Hitam Putih” di TransTV, “Mata Najwa” di MetroTV, “Indonesia Lawyers Club” di TvOne, dan “Kick Andy” di MetroTV, dengan menggunakan data Twitter. Tujuannya adalah untuk memberikan wawasan bagi produsen konten dan penyiar dalam meningkatkan keterlibatan penonton dan relevansi program.Penelitian ini menggunakan algoritma klasifikasi Naive Bayes untuk analisis sentimen pada data tekstual, seperti tweet. Naive Bayes dipilih karena kesederhanaan, kemudahan implementasi, dan efisiensinya dalam kategorisasi teks, mengkategorikan opini sebagai positif atau negatif berdasarkan probabilitas kemunculan kata. Metode ekstraksi fitur yang digunakan meliputi TF-IDF, Bag of Words, 2-grams, dan 3-grams. Untuk TF-IDF, akurasi yang diperoleh adalah 84%, Bag of Words mencapai akurasi 88%, 2-grams mendapatkan akurasi 57%, dan 3-grams memperoleh akurasi 51%.Penelitian dilakukan dengan tujuan mengevaluasi analisis sentimen terhadap reaksi individu tentang program televisi dengan menggunakan teknik Naive Bayes menghasilkan tingkat akurasi yang tinggi dalam klasifikasi sentimen dibandingkan dengan metodologi alternatif yang diperiksa.

Keywords—Analisis sentimen, Acara TV Indonesia, Twitter, Naive Bayes

# Introduction

Di era evolusi media sosial saat ini, pandangan publik tentang acara televisi telah menjadi fokus signifikan dalam analisis pasar dan pengambilan keputusan untuk produsen konten dan penyiar. Penelitian tentang sentimen publik terhadap tayangan televisi menawarkan wawasan berharga tentang preferensi dan reaksi penonton terhadap berbagai program TV. Sebuah studi kasus dilakukan untuk memeriksa opini publik tentang acara televisi populer seperti “Hitam Putih” di TransTV, “Mata Najwa” di MetroTV, “Indonesia Lawyers Club” di TvOne, dan “Kick Andy” di MetroTV. Studi ini menggunakan data dari tweet yang mengekspresikan opini publik di acara TV, termasuk angka retweet dan sentimen. Teknik analisis sentimen diterapkan untuk mengkategorikan tweet ini sebagai positif dan negatif. Hasil analisis sentimen ini dianalisis untuk mendapatkan pemahaman yang lebih dalam tentang respons penonton terhadap tayangan TV tertentu. Studi ini bertujuan untuk memberikan wawasan bagi industri televisi untuk memahami preferensi dan persepsi penonton, serta menawarkan rekomendasi untuk meningkatkan konten dan keterlibatan penonton. Dengan memahami opini publik dengan lebih baik, produser konten dan penyiar dapat meningkatkan kualitas dan relevansi program TV mereka untuk lebih melibatkan dan memuaskan audiens mereka.

Studi sebelumnya yang dilakukan oleh (Wilim & Oetama, 2021) meneliti persepsi sentiment Indonesia Lawyers Club menggunakan K-Nearest Neighbor (KNN) dan Decision Tree. Temuan dari penelitian ini mengindikasikan variasi dalam pendekatan analisis sentiment yang digunakan untuk mengeksplorasi respons penonton terhadap program televisi tertentu. Selain itu, ada penelitian lain yang telah dilakukan sebelumnya oleh (Farasalsabila et al, 2024) yang meneliti sentiment tayangan televisi di Twitter menggunakan SVM. Dalam peneltian sebelumnya yaitu untuk SVM mendapat akurasi sebesar 80%, sedangkan untuk penelitian KNN mendapatkan sebesar 76.94%, dan untuk decision tree sebanyak 70.28%. Perbedaan dalam metode analisis sentiment ini menunjukkan variasi dalam pendekatan penelitian terhadap topik yang serupa, dan hasilnya memberikan wawasan yang beragam tentang persepsi dan preferensi penonton terhadap acara televisi. Mengingat fenomena penelitian analisis sentimen yang telah dilakukan sebelumnya, maka dilakukan upaya untuk mengembangkan sistem untuk melakukan analisis opini publik terhadap review program televisi dan film di media sosial.Data media sosial dikumpulkan dan diproses menggunakan algoritma Naive Bayes untuk mengklasifikasikan opini menjadi positif, negatif.

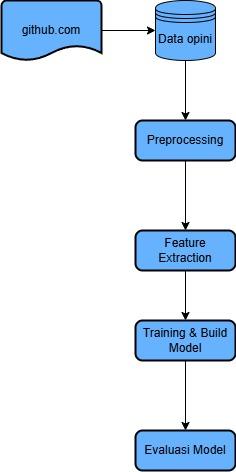
Algoritma Naive Bayes digunakan dalam penelitian ini untuk analisis sentimen pada data tekstual, seperti dataset tweet. Algoritma ini dipilih karena kesederhanaannya, kemudahan implementasi, dan efisiensi dalam kategorisasi teks. Dalam analisis sentimen, setiap kata dipandang sebagai fitur. Naive Bayes, dengan akurasi 84%, mengkategorikan teks ke dalam kelas sentimen positif atau negatif berdasarkan probabilitas kemunculan kata. Meskipun asumsi independensi fitur dalam Naive Bayes cukup kuat, asumsi ini dianggap masuk akal dalam konteks teks, karena kata-kata sering menunjukkan hubungan independen. Naive Bayes telah terbukti efektif dalam berbagai skenario klasifikasi teks, termasuk analisis sentimen. Dalam ranah analisis sentimen, setiap kata dalam teks dipandang sebagai fitur, dengan Naive Bayes digunakan untuk mengkategorikan teks ke dalam kelas sentimen yang berbeda yaitu positif dan negatif dengan menilai probabilitas kemunculan kata [3].

# Metode

1. Data Opini

Penelitian ini menggunakan kumpulan data berasal dari Twitter tentang tayangan televisi pada tahun 2017. Dataset khusus ini tersedia untuk umum dan bersumber dari situs web github.com, dapat diakses melalui URL berikut: [datasets](https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/blob/master/dataset_tweet_sentimen_tayangan_tv.csv.)

Data yang digunakan dalam penelitian ini yaitu sentiment pengguna Twitter terhadap tayangan televisi. Data tersebut terdiri dari kumpulan tweet yang diambil dari platform twitter, yang mencakup berbagai komentar dan pendapat pengguna tentang tanyangan televisi tersebut. Dataset ini mencakup beberapa variable, termasuk kolom sentiment yang memuat kategori sentiment yaitu positif dan negatif, acara tv yaitu Hitamputih TransTV, Indonesia Lawyers Club TvOne, Kick Andy MetroTV, dan Mata Najwa MetroTV, Jumlah retweet pada setiap tweetnya dan teks tweet yang berisi teks dari setiap tweetnya. Dalam analisis ini, variable sentiment akan menjadi variable target yang ingin diprediksi, sementara variable lainnya akan digunakan sebagai fitur-fitur dalam klasifikasi sentimen.



**Gambar 1. Arsitektur Sistem**

1. Preprocessing

Pada tahap ini dilakukkan preprocessing data teks, dimulai dengan pencarian missing value untuk memastikan kelengkapan dataset. Setelah itu, dilakukan penghapusan tanda baca karena tanda baca sering tidak memberikan informasi yang signifikan dalam analisis teks. Normalisasi kata singkat dilakukan untuk mengubah kata-kata informal atau tidak baku menjadi bentuk yang lebih jelas dan standar untuk memudahkan analisis. Pada tahap ini juga mencakup tokenisasi dan stemming. Untuk algoritma stemming yang digunakan adalah menggunakan algoritma stemmer Bahasa Indonesia. Tahap selanjutnya adalah dilakukan proses ektraksi fitur menggunakan n-gram pada text tweet. Proses ini dilakukan untuk menguraikan kata menjadi sejumlah n karakter sebagai satuan term dan menghitung probabilits kemunculan tiap-tiap karakter.

1. Feature Extraction

Pada tahap ini, dilakukan ekstraksi fitur dari data yang digunakan untuk analisis sentimen. Sebelum prosedur ekstraksi fitur, data awalnya di splitting menjadi dua segmen dengan rasio 80:20, di mana 80% data dialokasikan untuk tujuan pelatihan dan 20% sisanya ditujukan untuk pengujian. Metodologi ekstraksi fitur yang digunakan dalam penyelidikan ini mencakup berbagai pendekatan, seperti TF-IDF (Term Frequency-Inverse Document Frequency), bag of word, 2 gram, 3 gram dan Word2vec. Bag of word adalah kerangka kerja yang mempelajari setiap kosakata dari teks, dengan menghitung banyaknya kemunculan setiap kata[4]. Konsep N-Gram berkaitan dengan sekumpulan kata yang berasal dari kalimat tertentu. Memanfaatkan teknik N-gram memungkinkan untuk menghasilkan kata atau karakter. N-Gram melibatkan pertimbangan kata-kata sebelumnya dan berikutnya. Klasifikasi N-Gram dapat didasarkan pada jumlah segmen kata atau substring yang dihasilkan, seperti Unigram, Bigram, Trigram, dan sebagainya, sesuai dengan nilai 'n' dalam N-Gram [5]. Pada tahun 2003, Bengio dkk. (Bengioet al., 2003) menyajikan konsep penyematan kata, yang merupakan fungsi parameter yang bertanggung jawab untuk mengubah setiap kata menjadi vektor dengan jumlah dimensi yang tinggi[6]. Word2vec, sebuah algoritma untuk penyematan kata, menetapkan vektor untuk setiap kata dalam teks. Word2vec bergantung pada konteks linguistik lokal untuk operasinya. Pemahaman semantik yang diperoleh untuk kata tertentu dibentuk oleh kata-kata yang mengelilinginya[7]. Penggabungan metodologi ini diantisipasi untuk meningkatkan kapasitas model untuk menangkap beragam aspek data teks, akibatnya mengarah ke analisis sentimen yang lebih tepat dan berwawasan luas.

1. Naïve Bayes

Langkah klasifikasi dilakukan dengan memanfaatkan metode Gaussian Naïve Bayes, sebuah pendekatan yang menggunakan Teorema Bayes untuk menghitung probabilitas kelas dari suatu data [8]. Proses klasifikasi terdiri dari dua tahap, yaitu tahap training dan testing terhadap model yang telah disiapkan. Pembagian dataset dilakukan dengan metode split, di mana data dibagi antara data training dan data testing. Pembagian ini dilakukan secara acak dari total keseluruhan data yang tersedia [9]. Data training digunakan untuk melatih model klasifikasi, sementara data testing digunakan untuk mengevaluasi performa model [10]. Perhitungan kelompok probabilitas Naïve Bayes menggunakan pendekatan algoritma Bayes dengan menggunakan 2 persamaan [11].

-)

Keterangan:

= Probabilitas terhadap kelas y

= Nilai kelas-

= Nilai *variance*

= *Mean* kelas y

= Phi 3.14

Keterangan:

= Data yang tidak diketahui kelasnya

= Hipotesis dari data X yang merupakan kelas spesifik

= Probabilitas hipotesis H berdasarkan kondisi X (posteriori probability)

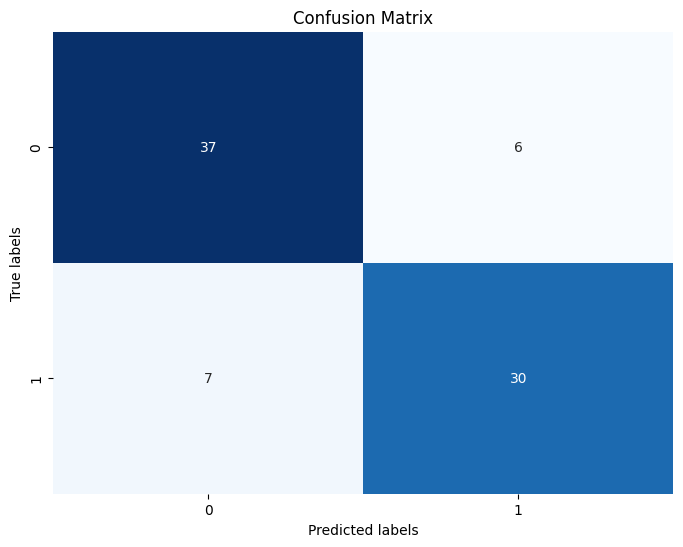
) = Probabilitas hipotesis H (prior probability)

= Probabilitas X berdasarkan kondisi dari hipotesis H

= Probabilitas dari X

# Hasil

## Evaluasi Model

Evaluasi dilakukan untuk menguji hasil klasifikasi dengan mengukur nilai kinerja yang dilakukan oleh sistem. Parameter pengujian yang digunakan untuk mengevaluasi adalah akurasi atau kebenaran tingkat proses klasifikasi yang perhitungannya dari confusion matriks.

1. TF-IDF

Istilah pembobotan adalah melibatkan penetapan nilai numerik untuk setiap istilah yang ada dalam dokumen. Skema bobot istilah dipengaruhi oleh tiga faktor utama: frekuensi istilah (TF), frekuensi dokumen terbalik (IDF), dan normalisasi.

Pada penelitian ini, TF-IDF digunakan sebagai pembobot fitur metode.

Keterangan:

: jumlah dokumen dalam koleksi

: frekuensi istilah

: frekuensi dokumen terbalik.

Dalam konteks penyelidikan ini, TF-IDF diimplementasikan sebagai metodologi untuk membobot karakteristik fitur.

**Tabel 1. Klasifikasi Report TF-IDF**

|  | precision | recall | f1-score | support |
| --- | --- | --- | --- | --- |
| negative | 0.84 | 0.86 | 0.85 | 43 |
| positive | 0.83 | 0.81 | 0.82 | 37 |
| accuracy |  |  | 0.84 | 80 |
| macro avg | 0.84 | 0.84 | 0.85 | 80 |
| weighted avg | 0.84 | 0.84 | 0.85 | 80 |

Berdasarkan Tabel 1, akurasi yang diperoleh adalah sebesar 84%. Nilai presisi untuk label negatif adalah 84%, sedangkan untuk label positif adalah 83%. Recall untuk label negatif adalah 86%, sedangkan untuk label positif adalah 81%. F1-score yang dicapai adalah sebesar 85% untuk label negatif, sedangkan label positif mendapat sebesar 82%. Perbandingan antara data prediksi dan data aktual ditunjukkan pada gambar 2.

**Gambar 2. Confusion Matrix TF-IDF**

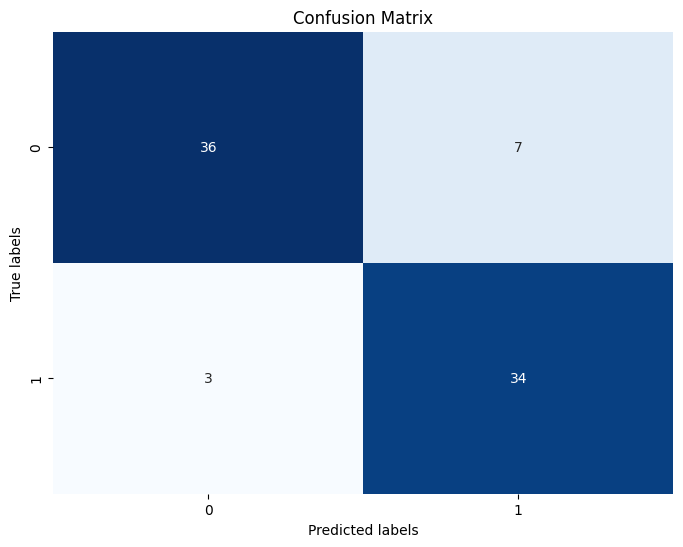
Berdasarkan Gambar 2, Hasil confusion matrix menunjukan bahwa pada terdapat 37 data yang terklasifikasi dengan benar sebagai negatif dan 30 yang terklasifikasi benar positif , 6 data yang seharusnya positif namun salah terklasifikasi sebagai negatif, dan terdapat 7 data yang seharusnya negatif namun salah terklasifikasi sebagai positif.

1. Bag Of Word

**Tabel 2. Klasifikasi Report Bag Of Word**

|  | precision | recall | f1-score | support |
| --- | --- | --- | --- | --- |
| negative | 0.92 | 0.84 | 0.88 | 43 |
| positive | 0.83 | 0.92 | 0.87 | 37 |
| accuracy |  |  | 0.88 | 80 |
| macro avg | 0.88 | 0.88 | 0.87 | 80 |
| weighted avg | 0.88 | 0.88 | 0.88 | 80 |

Berdasarkan Tabel 2, akurasi yang diperoleh adalah sebesar 88­­%. Nilai presisi untuk label negatif adalah 92%, sedangkan untuk label positif adalah 83%. Recall untuk label negatif adalah 84%, sedangkan untuk label positif adalah 92%. F1-score yang dicapai adalah sebesar 88% untuk label negaif, sedangkan positif sebesar 87%. Perbandingan antara data prediksi dan data aktual ditunjukkan pada gambar 3.



**Gambar 3. Confusion Matrix Bag of Word**

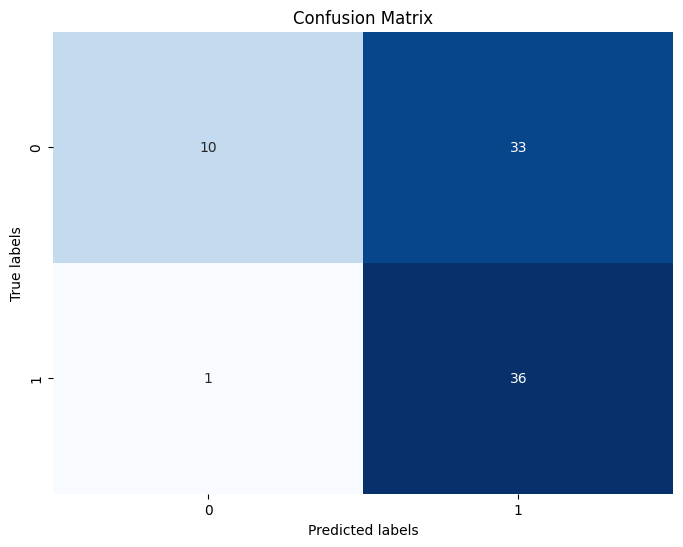
Berdasarkan Gambar 3, Hasil confusion matrix menunjukan bahwa pada terdapat 36 data yang terklasifikasi dengan benar sebagai negatif , 34 data yang terklasifikasi dengan benar sebagai positif, 7 data yang seharusnya negatif namun salah terklasifikasi sebagai positif, dan terdapat 3 data yang seharusnya positif namun salah terklasifikasi sebagai negatif

c) 2-Grams

**Tabel 3. Klasifikasi Report 2-Grams**

|  | precision | recall | f1-score | support |
| --- | --- | --- | --- | --- |
| negative | 0.91 | 0.23 | 0.37 | 43 |
| positive | 0.52 | 0.97 | 0.68 | 37 |
| accuracy |  |  | 0.57 | 80 |
| macro avg | 0.72 | 0.60 | 0.52 | 80 |
| weighted avg | 0.73 | 0.57 | 0.51 | 80 |

Berdasarkan Tabel 3, akurasi yang diperoleh adalah sebesar 57%. Nilai presisi untuk label negatif adalah 91%, sedangkan untuk label positif adalah 52%. Recall untuk label negatif adalah 23%, sedangkan untuk label positif adalah 97%. F1-score untuk label negatif adalah sebesar 37%. Sedangkan F1-Score untuk label positif adalah 68%. Perbandingan antara data prediksi dan data aktual ditunjukkan pada gambar 4.



**Gambar 4. Confusion Matrix 2-Grams**

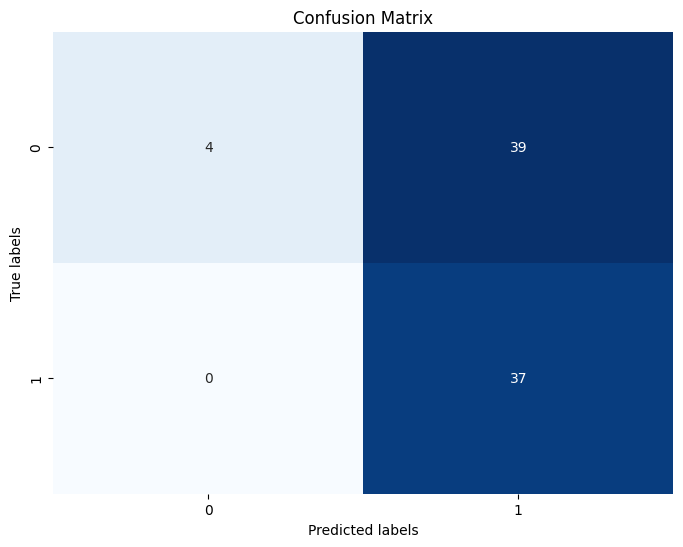
Berdasarkan Gambar 4, Hasil confusion matrix menunjukan bahwa pada terdapat 10 data yang terklasifikasi dengan benar sebagai negatif, 36 data yang terklasifikasi dengan benar sebagai positif, 33 data yang seharusnya positif namun salah terklasifikasi sebagai negatif, dan terdapat 1 data yang seharusnya negatif namun salah terklasifikasi sebagai positif

1. 3-Grams

**Tabel 4. Klasifikasi Report 3-Gram**

|  | precision | recall | f1-score | support |
| --- | --- | --- | --- | --- |
| negative | 1.00 | 0.09 | 0.17 | 43 |
| positive | 0.49 | 1.00 | 0.65 | 37 |
| accuracy |  |  | 0.51 | 80 |
| macro avg | 0.74 | 0.55 | 0.41 | 80 |
| weighted avg | 0.76 | 0.51 | 0.39 | 80 |

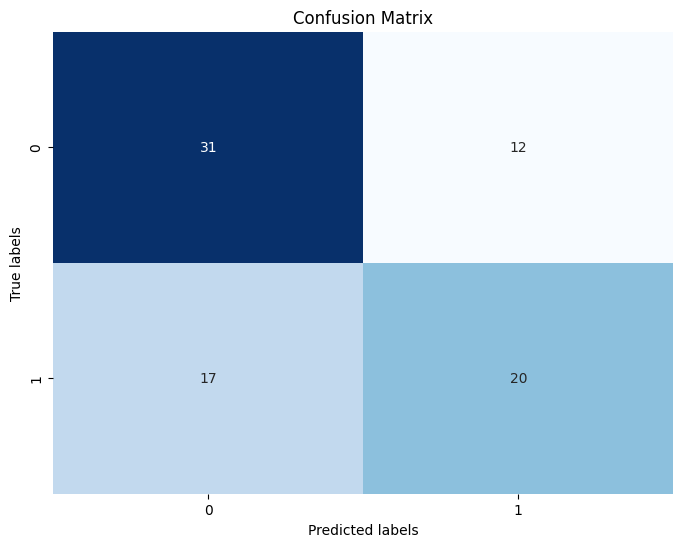
Berdasarkan Tabel 4, akurasi yang diperoleh adalah sebesar 51%. Nilai presisi untuk label negatif adalah 100%, sedangkan untuk label positif adalah 49%. Recall untuk label negatif adalah 9%, sedangkan untuk label positif adalah 100%. F1-score untuk label negatif adalah sebesar 17%. Sedangkan F1-Score untuk label positif adalah 65%. Perbandingan antara data prediksi dan data aktual ditunjukkan pada gambar 5

****

**Gambar 5. Confusion Matrix 3-Gram**

Berdasarkan Gambar 5, Hasil confusion matrix menunjukan bahwa pada terdapat 4 data yang terklasifikasi dengan benar sebagai negatif, 37 data yang terklasifikasi dengan benar sebagai positif, 39 data yang seharusnya positif namun salah terklasifikasi sebagai negative

1. Word2Vec

****Word2vec, sebuah algoritma untuk penyematan kata, menetapkan vektor untuk setiap kata dalam teks. Word2vec bergantung pada konteks linguistik lokal untuk operasinya.

Tentukan berat lapisan input ke lapisan tersembunyi menggunakan rumus yang ditentukan dalam Persamaan.

keterangan:

: hidden neuron

: input vector

: nilai *transpose* pada bobot dari *input layer* ke *hidden layer*

Tentukan berat lapisan tersembunyi ke lapisan keluaran melalui rumus yang ditentukan dalam Persamaan.

Keterangan:

: output pada baris j

: *nilai transpose* bobot dari *hidden layer* ke *ouput layer*

**Tabel 5. Klasifikasi Report Word2Vec**

|  | precision | recall | f1-score | support |
| --- | --- | --- | --- | --- |
| negative | 0.65 | 0.72 | 0.68 | 43 |
| positive | 0.62 | 0.54 | 0.58 | 37 |
| accuracy |  |  | 0.64 | 80 |
| macro avg | 0.64 | 0.63 | 0.63 | 80 |
| weighted avg | 0.64 | 0.64 | 0.63 | 80 |

Berdasarkan Tabel 5, akurasi yang diperoleh adalah sebesar 63%. Nilai presisi untuk label negatif adalah 65%, sedangkan untuk label positif adalah 62%. Recall untuk label negatif adalah 72%, sedangkan untuk label positif adalah 54%. F1-score untuk label negatif adalah sebesar 68%. Sedangkan F1-Score untuk label positif adalah 58%. Perbandingan antara data prediksi dan data aktual ditunjukkan pada gambar 6

**Gambar 6. Confusion Matrix Word2Vec**

Berdasarkan Gambar 6, Hasil confusion matrix menunjukan bahwa pada terdapat 31 data yang terklasifikasi dengan benar sebagai negatif, 20 data yang terklasifikasi dengan benar sebagai positif, 12 data yang seharusnya positif namun salah terklasifikasi sebagai negative, dan terdapat 17 data yang seharusnya negatif namun salah terklasifikasi sebagai positif

## Diskusi

**Tabel 6. Perbandingan Akurasi**

|  |  |
| --- | --- |
| Algoritma | Akurasi |
| SVM (Support Vector Machine) | 80% |
| K-Nearest Neighbor (KNN) | 76.94% |
| Decision Tree | 70.28% |
| Naive Bayes | 84% |

Penelitian ini dilakukan untuk menilai kemanjuran proses analisis sentimen. Dataset untuk percobaan ini terdiri dari 400 ulasan, dipisahkan menjadi 200 ulasan positif dan 200 ulasan negatif. Dataset bersumber dari beragam komentar dan posting di platform media sosial online. Selanjutnya, dataset ini akan digunakan untuk melatih dan mengevaluasi algoritma Naïve Bayes. Hasil eksperimen mengungkapkan bahwa algoritma Naive Bayes yang diterapkan mampu mencapai tingkat akurasi 84%

Hasil penelitian yang menunjukkan kinerja yang lebih unggul dibandingkan dengan studi sebelumnya, yang menggunakan metode analisis sentimen alternatif seperti SVM, KNN, dan decision tree. Potensi peningkatan kinerja Naive Bayes dalam penelitian ini mungkin berasal dari berbagai faktor. Terutama, Naive Bayes telah menunjukkan efektivitas dalam analisis teks dan sentimen. Meskipun model ini sederhana, tetapi terbukti kuat ketika dihadapkan dengan data tekstual. Selain itu, pemanfaatan metode pembobotan TF-IDF dalam preprocessing data dapat lebih meningkatkan kinerja model ini dengan meningkatkan kesiapan data sebelum integrasi ke dalam kerangka kerja Naive Bayes. Mempertimbangkan aspek-aspek ini, penelitiaan ini selaras dengan penelitian sebelumnya, menampilkan manfaat kinerja dari teknik analisis sentimen tertentu, terutama Naive Bayes.

# kesimpulan

Penelitian dilakukan dengan tujuan mengevaluasi analisis sentimen terhadap reaksi individu tentang program televisi dengan menggunakan teknik Naive Bayes menghasilkan tingkat akurasi yang tinggi dalam klasifikasi sentimen dibandingkan dengan metodologi alternatif yang diperiksa. Hasilnya menunjukkan bahwa menggunakan pendekatan Naive Bayes untuk analisis sentimen reaksi individu terhadap program televisi menyebabkan tingkat akurasi 84%. Penemuan ini memiliki implikasi signifikan untuk memahami sentimen publik mengenai program televisi. Selain itu, hasil ini berpotensi membentuk dasar untuk penelitian selanjutnya dalam meningkatkan kompleksitas dan ketepatan metode analisis sentimen dalam ranah persepsi televisi.

DAFTAR PUSTAKA

[1] N. N. Wilim and R. S. Oetama, “Sentiment Analysis About Indonesian Lawyers Club Television Program Using K-Nearest Neighbor, Naïve Bayes Classifier, And Decision Tree,” *IJNMT Int. J. New Media Technol.*, vol. 8, no. 1, pp. 50–56, Jun. 2021, doi: 10.31937/ijnmt.v8i1.1965.

[2] F. Farasalsabila, E. Utami, and M. Hanafi, “ANALYSIS OF PUBLIC OPINION ON INDONESIAN TELEVISION SHOWS USING SUPPORT VECTOR MACHINE,” *JURTEKSI J. Teknol. Dan Sist. Inf.*, vol. 10, no. 2, pp. 239–246, Mar. 2024, doi: 10.33330/jurteksi.v10i2.2935.

[3] R. Sari and R. Y. Hayuningtyas, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website,” *Indones. J. Softw. Eng. IJSE*, vol. 5, no. 2, pp. 51–60, Dec. 2019, doi: 10.31294/ijse.v5i2.6957.

[4] H. Utama and A. Masruro, “Analisis Sentimen pada Twitter menggunakan Word Embedding dengan Pendekatan Word2Vec,” *J. Sist. Cerdas*, vol. 5, no. 2, pp. 128–134, Aug. 2022, doi: 10.37396/jsc.v5i2.242.

[5] N. Arifin, U. Enri, and N. Sulistiyowati, “Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification,” *STRING Satuan Tulisan Ris. Dan Inov. Teknol.*, vol. 6, no. 2, p. 129, Dec. 2021, doi: 10.30998/string.v6i2.10133.

[6] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, “A Neural Probabilistic Language Model”.

[7] A. Nurdin, B. Anggo Seno Aji, A. Bustamin, and Z. Abidin, “PERBANDINGAN KINERJA WORD EMBEDDING WORD2VEC, GLOVE, DAN FASTTEXT PADA KLASIFIKASI TEKS,” *J. Tekno Kompak*, vol. 14, no. 2, p. 74, Aug. 2020, doi: 10.33365/jtk.v14i2.732.

[8] S. Mujahidin, B. Prasetio, and M. C. C. Utomo, “Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube Dengan Metode Gaussian naïve bayes,” *Voteteknika Vocat. Tek. Elektron. Dan Inform.*, vol. 10, no. 3, p. 17, Sep. 2022, doi: 10.24036/voteteknika.v10i3.118299.

[9] Fajar Hadi Saputra and Cholifah, “Pengaruh Narasi dalam Konten Vlog Channel Youtube ‘Menjadi Manusia’ Terhadap Sikap dalam Menjaga Kesehatan Mental,” *J. ISIP J. Ilmu Sos. Dan Ilmu Polit.*, vol. 19, no. 1, pp. 11–22, Sep. 2023, doi: 10.36451/jisip.v19i1.66.

[10] A. R. T. Lestari and M. A. Fauzi, “Analisis Sentimen Tentang Opini Pilkada Dki 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Näive Bayes dan Pembobotan Emoji”.

[11] A. Abdurrahim, L. Syafa’ah, and M. Lestandy, “Sentiment analysis of Covid-19 vaccine tweets utilizing Naïve Bayes,” presented at the 1ST INTERNATIONAL CONFERENCE ON TECHNOLOGY, INFORMATICS, AND ENGINEERING, Malang, Indonesia, 2022, p. 020052. doi: 10.1063/5.0094607.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**