JUTIK : Jurnal Teknologi Informasi dan Komputer
Oktober-2025, Vol. 11, No.2, Hal.231-242
ISSN(P): 2442-241X; ISSN(E): 2528-5211

OPINI PUBLIK TERHADAP ISU KEASLIAN IJAZAH PADA *PLATFORM* YOUTUBE DENGAN NAÏVE BAYES, KNN, DAN SMOTE

Agnes Anastasia Putri[1*], Christian Bautista[2], Hafiz Irsyad[3], Abdul Rahman[4]
Universitas Multi Data Palembang, Palembang, Sumatera Selatan, Indonesia[1]
Email*: agnesanastasiaputri_2226250013@mhs.mdp.ac.id
Universitas Multi Data Palembang, Palembang, Sumatera Selatan, Indonesia[2]
Email: christianbautista_2226250002@mhs.mdp.ac.id
Universitas Multi Data Palembang, Palembang, Sumatera Selatan, Indonesia[3]
Email: hafizirsyad@mdp.ac.id
Universitas Multi Data Palembang, Palembang, Sumatera Selatan, Indonesia[4]
Email: arahman@mdp.ac.id

ABSTRAK

Di era digital, opini publik menyebar secara masif dan instan melalui berbagai platform media sosial. Salah satu isu yang sempat memicu perhatian dan perdebatan luas di ruang digital adalah terkait keaslian ijazah Presiden Joko Widodo. Isu ini memancing beragam reaksi, baik dukungan, kecaman, maupun sikap netral dari warganet, khususnya melalui kolom komentar di YouTube. Penelitian ini menganalisis sentimen masyarakat terhadap isu tersebut menggunakan pendekatan *machine learning* dengan algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN), serta teknik penyeimbangan data SMOTE. Sebanyak 1.000 komentar dianalisis dan diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Empat skenario pengujian dilakukan, yaitu: KNN, KNN dengan SMOTE, *Naïve Bayes*, dan *Naïve Bayes* dengan SMOTE dengan perbandingan performa yang diuji untuk melihat efektivitas masing-masing dalam mengklasifikasikan opini digital. Hasil pengujian menunjukkan kombinasi *Naïve Bayes* dan SMOTE memberikan performa terbaik dengan akurasi, presisi, *recall*, dan *F1-score* sebesar 73%. Sebaliknya, model dengan performa terburuk adalah KNN dengan SMOTE, yang hanya memperoleh akurasi sebesar 27%, presisi 53%, *recall* 34%, dan *F1-score* 15%. Penelitian ini menegaskan pentingnya pemilihan algoritma dan strategi penanganan data dalam klasifikasi opini digital, serta dapat menjadi dasar pengembangan sistem analisis sentimen yang andal di masa depan.
Kata kunci: klasifikasi sentimen, KNN, *Naïve Bayes*, opini digital, SMOTE

ABSTRACT

In the digital era, public opinion spreads massively and instantly through various social media platforms. One issue that has sparked widespread attention and debate in the digital space is regarding the authenticity of President Joko Widodo's diploma. This issue has provoked various reactions, both support, criticism, and neutral attitudes from netizens, especially through the comments column on YouTube. This study analyzes public sentiment towards the issue using a machine learning approach with the Naïve Bayes and K-Nearest Neighbors (KNN) algorithms, as well as the SMOTE data balancing technique. A total of 1,000 comments were analyzed and classified into three sentiment categories, namely positive, negative, and neutral. Four test scenarios were carried out, namely: KNN, KNN with SMOTE, Naïve Bayes, and Naïve Bayes with SMOTE with a

Submitted: 21 Mei 2025
Accepted: 15 September 2025
Published: 10 Oktober 2025

performance comparison tested to see the effectiveness of each in classifying digital opinion. The test results showed that the combination of Naïve Bayes and SMOTE provided the best performance with accuracy, precision, recall, and F1-score of 73%. In contrast, the worst performing model is KNN with SMOTE, which only achieves 27% accuracy, 53% precision, 34% recall, and 15% F1-score. This study emphasizes the importance of algorithm selection and data handling strategies in digital opinion classification, and can be the basis for developing a reliable sentiment analysis system in the future.
Keywords: digital opinion, KNN, Naïve Bayes, sentiment classification, SMOTE

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah telah mengubah cara masyarakat menyampaikan opini dan merespons isu publik [1]. Media sosial dan forum daring kini menjadi ruang diskusi terbuka, memungkinkan siapa saja berpendapat tanpa hambatan [2]. Salah satu topik yang menarik perhatian publik di ranah digital adalah dugaan terkait keaslian ijazah Presiden Republik Indonesia ke-7, Joko Widodo. Isu tersebut memicu berbagai tanggapan dari masyarakat, mulai dari dukungan, keraguan, hingga tudingan. Hal ini menunjukkan betapa pentingnya analisis opini digital dalam menangkap persepsi publik terhadap isu-isu sensitif yang memiliki dampak besar.

Opinion mining atau analisis opini adalah cabang NLP (Natural Language Processing) yang menitikberatkan pada identifikasi dan klasifikasi sentimen dalam teks [3]. Dalam kasus politik seperti isu ijazah Presiden ke-7 Jokowi, teknik ini berguna untuk memetakan persepsi publik dan tren opini [4]. Salah satu tantangan utama dalam opini *mining* adalah ketidakseimbangan kelas data, di mana sentimen yang dominan (netral atau positif) lebih banyak muncul dibandingkan opini minoritas (negatif atau ekstrem). Ketimpangan ini dapat menghambat kinerja model klasifikasi, sehingga mengurangi akurasi dan keandalannya dalam menangkap keragaman opini yang ada [5].

Penelitian yang dilakukan oleh [6] membahas mengenai perbandingan antara *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) dalam dalam klasifikasi diagnosis penyakit diabetes melitus menggunakan data dari Kaggle. Dengan pendekatan *Knowledge Discovery in Database* (KDD), data diproses melalui tahap seleksi, pembersihan, transformasi, dan pemodelan. Hasilnya menunjukkan bahwa *Naïve Bayes* memiliki akurasi lebih tinggi (hingga 80%) dibandingkan KNN (maksimal 75%), serta presisi yang lebih baik, sementara KNN unggul dalam *recall*. Penelitian ini memberikan wawasan penting tentang efektivitas kedua algoritma dalam klasifikasi medis berbasis data.

Selanjutnya, penelitian yang dilakukan oleh [7] menerapkan metode *Support Vector Machine* (SVM) dan teknik SMOTE untuk mengklasifikasikan sentimen publik terhadap Polisi Republik Indonesia berdasarkan komentar YouTube. Pada penelitian ini digunakan 1.000 data dan menunjukkan bahwa penggunaan SMOTE membantu mengatasi ketidakseimbangan data dan meningkatkan akurasi model. Akurasi tertinggi sebesar 91% dicapai pada rasio data latih dan uji 70:30 setelah penerapan SMOTE, dibandingkan 84,67% tanpa SMOTE.

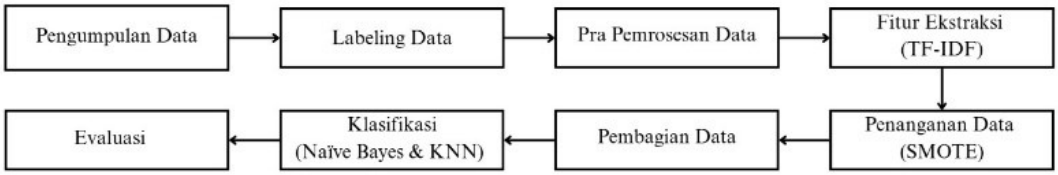
Berdasarkan masalah dan penelitian terdahulu, penelitian ini menggunakan dua algoritma klasifikasi, yaitu *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) karena sederhana dan handal dalam klasifikasi teks [8][9]. Selain itu, penelitian ini juga memanfaatkan teknik SMOTE sebagai solusi untuk mengatasi ketidakseimbangan dalam distribusi data [10]. SMOTE bekerja membuat sampel sintetis di kelas minoritas agar distribusi lebih seimbang dan model belajar lebih adil [11]. Penelitian ini bertujuan untuk

menguji dan membandingkan performa *Naïve Bayes* dan KNN dalam mengklasifikasikan opini digital terhadap isu ijazah Presiden Jokowi, baik sebelum maupun sesudah penerapan SMOTE.

Secara keseluruhan, penelitian ini merupakan studi pertama yang secara khusus menganalisis opini publik terhadap isu keaslian ijazah Presiden Jokowi melalui komentar YouTube dengan pendekatan klasifikasi Naïve Bayes, KNN, dan teknik penyeimbangan data SMOTE. Dengan fokus pada pemodelan dan analisis pola opini menggunakan pembelajaran mesin, penelitian ini memberikan kontribusi baru dalam memahami dinamika opini politik digital, tanpa bertujuan membuktikan kebenaran isu tersebut.

2. METODE

Penelitian ini menerapkan pendekatan eksperimental dengan langkah-langkah yang terencana secara sistematis. Rangkaian metodologi disusun secara runtut untuk menjamin akurasi dan efektivitas dalam proses klasifikasi opini digital yang dilakukan. Pada Gambar 1, dapat dilihat tahapan penelitian yang terdiri dari beberapa tahapan yang akan dilakukan dalam penelitian ini.



Gambar 1. Tahapan penelitian

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data yang diperoleh dari hasil scraping dari komentar Youtube [12] dalam bentuk file .csv. Proses *scraping* ini dibantu dengan website "Apify" [13]. Dari hasil *scraping* dikumpulkan data sebanyak 1.000 data komentar yang nantinya akan dibantu dengan dataset berisikan data negative, positif, dan netral yang didapatkan dari github "ID-OpinionWords" yang dipublish oleh "masdevid" [14].

Labelling Data

Langkah selanjutnya adalah proses pelabelan data, yaitu memberikan label sentimen pada setiap komentar dalam bentuk kategori positif, negatif, atau netral. Pelabelan komentar dilakukan secara objektif menggunakan kamus sentimen *ID-OpinionWords* hasil adaptasi dari Liu's Opinion Lexicon oleh Wahid & Azhari. Label positif atau negatif diberikan berdasarkan kecocokan kata dalam komentar dengan kamus tersebut, sedangkan komentar tanpa kecenderungan kuat diberi label netral. Karena pelabelan berbasis leksikon otomatis, proses ini tidak melibatkan annotator manusia, sehingga jumlah annotator, *inter-annotator agreement*, dan *quality control* manual tidak relevan. Pelabelan ini bertujuan menyiapkan data untuk pelatihan model klasifikasi sentimen.

Pra Pemprosesan Data

Pada tahap ini, semua komentar akan diubah menjadi huruf kecil. Karakter yang merupakan non-huruf akan dihapus, melakukan pemisahan agar teks menjadi kata (tokenisasi). Selanjutkan akan dilakukan proses untuk menghapus kata-kata umum yang

tidak penting (*stopword*), dan akhirnya diterapkan proses *stemming* menggunakan pustaka Sastrawi (*library* Python) agar bentuk kata diubah menjadi kata dasar. Hasil pra-pemrosesan ini akan disimpan dalam kolom baru.

Fitur Ekstraksi

Langkah berikutnya adalah ekstraksi fitur menggunakan metode TF-IDF, yaitu metode pembobotan kata berdasarkan frekuensi kemunculannya dalam dokumen dan keseluruhan kumpulan dokumen. Metode ini menggabungkan TF (frekuensi kata dalam dokumen) dan IDF (jarangnya suatu kata muncul di seluruh dokumen) [15]. Setiap komentar dibersihkan dan diubah menjadi vektor numerik, lalu ditampilkan dalam bentuk *DataFrame*. Setelah itu, data dibagi menjadi data latih dan data uji menggunakan *train_test_split* .

Penanganan Data & Klasifikasi

Teknik SMOTE (*Synthetic Minority Oversampling Technique*) diterapkan untuk menyeimbangkan jumlah data tiap kelas sebelum pelatihan. Dua model klasifikasi yang digunakan adalah *Naïve Bayes* dan *K-Nearest Neighbors* (K-NN). *Naïve Bayes* adalah metode probabilistik yang memprediksi kemungkinan suatu kejadian berdasarkan data masa lalu [16], sedangkan K-NN mengklasifikasikan data uji berdasarkan mayoritas label dari k tetangga terdekatnya menggunakan ukuran jarak tertentu [17][18].

Dalam eksperimen, model pertama akan dilatih menggunakan data tanpa SMOTE, lalu dibandingkan dengan model kedua yang dilatih menggunakan SMOTE. Pada algoritma KNN, nilai k (jumlah tetangga terdekat) akan dievaluasi secara menyeluruh menggunakan *cross-validation* untuk mendapatkan akurasi terbaik, baik dengan maupun tanpa SMOTE.

Evaluasi

Evaluasi performa model klasifikasi sentimen komentar YouTube menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix*. Metrik ini mengukur ketepatan prediksi model dan matriks kebingungan divisualisasikan dengan *heatmap* untuk memudahkan pemahaman hasil.

Confusion matrix adalah matriks yang menunjukkan jumlah prediksi benar dan salah model pada tiap kelas, terdiri dari prediksi positif yang benar (*True Positive*), prediksi positif yang salah (*False Positive*), prediksi negatif yang benar (*False Positive*) dan prediksi negatif yang salah (*False Negative*) [19].

Akurasi adalah ukuran yang menunjukkan seberapa tepat model dalam melakukan prediksi secara keseluruhan. Nilai akurasi yang tinggi mencerminkan kemampuan model dalam menghasilkan prediksi yang benar dengan baik [19]. Untuk mendapatkan nilai *accuracy* digunakan persamaan (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+TN}\ x\ 100\% \quad (1)$$

Precision adalah perbandingan antara jumlah prediksi positif yang tepat (*True Positive*) dengan seluruh prediksi yang dikategorikan sebagai positif [19]. Untuk mendapatkan nilai *precision* digunakan persamaan (2).

$$Precision = \frac{TP}{TP+FP}\ x\ 100\% \quad (2)$$

Recall mengukur proporsi data positif yang berhasil terdeteksi dengan benar oleh model, yaitu perbandingan antara *true positive* (TP) dan seluruh data yang sebenarnya positif [19]. Untuk mendapatkan nilai *recall* digunakan persamaan (3).

$$Recall = \frac{TP}{TP+FN}\ x\ 100\% \quad (3)$$

F1-Score menghitung rata-rata harmonis dari presisi dan recall untuk menilai performa keseluruhan model [19]. Untuk mendapatkan nilai *F1-Score* digunakan persamaan (4).

$$F1 - score = 2x\frac{Precision\ x\ Recall}{(Precision + Recall)} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Berdasarkan penelitian yang dilakukan, akan disajikan hasil evaluasi performa model klasifikasi sentimen komentar YouTube beserta pembahasan terkait interpretasi metrik dan visualisasi yang diperoleh. Analisis ini bertujuan untuk memahami kekuatan dan kelemahan model dalam mengklasifikasi sentimen dengan akurat.

Labelling Data

Berdasarkan data *scraping* yang telah didapatkan, dilakukan pelabelan sentimen pada total 1.000 komentar yang tersedia, terdiri dari 246 komentar dengan sentimen positif, 489 komentar dengan sentimen negatif, dan 265 komentar dengan sentimen netral.

Pra Pemprosesan Data

Setelah melakukan pelabelan, data akan dilanjutkan ke tahap pemprosesan. Tahap ini dimulai dengan melakukan *case folding* pada data. Tabel 1 merupakan hasil dari proses *case folding* pada beberapa contoh komentar.

Tabel 1. Hasil *case folding*

Sebelum	dan para anda semua sa'atnya skrg siapa yg jujur siapa yg bohong dan penjilat
Sesudah	dan para anda semua sa'atnya skrg siapa yg jujur siapa yg bohong dan penjilat

Selanjutnya, akan dilakukan proses tokenisasi terhadap data. Tabel 2 yang merupakan hasil tokenisasi dari beberapa komentar setelah proses *case folding*.

Tabel 2. Hasil tokenisasi

Sebelum	dan para anda semua sa'atnya skrg siapa yg jujur siapa yg bohong dan penjilat
Sesudah	['dan', 'para', 'anda', 'semua', 'saatnya', 'skrg', 'siapa', 'yg', 'jujur', 'siapa', 'yg', 'bohong', 'dan', 'penjilat']

Setelah proses tokenisasi selesai, tahap selanjutnya adalah penghilangan *stopword*. Tabel 3 merupakan hasil setelah dilakukan proses penghilangan *stopword* pada token-token yang telah dihasilkan.

Tabel 3. Hasil *stopword*

Sebelum	['dan', 'para', 'anda', 'semua', 'saatnya', 'skrg', 'siapa', 'yg', 'jujur', 'siapa', 'yg', 'bohong', 'dan', 'penjilat']
Sesudah	['semua', 'saatnya', 'skrg', 'siapa', 'yg', 'jujur', 'siapa', 'yg', 'bohong', 'penjilat']

Langkah terakhir dalam pra pemrosesan adalah proses *stemming*. Proses ini bertujuan untuk menyamakan variasi kata agar analisis menjadi lebih efektif dan konsisten. Tabel 4 adalah hasil dari proses *stemming* pada data komentar.

Tabel 4. Hasil *stemming*

Sebelum	['semua', 'saatnya', 'skrg', 'siapa', 'yg', 'jujur', 'siapa', 'yg', 'bohong', 'penjilat']
Sesudah	[‘semua’, ‘saat’, ‘skrg’, ‘siapa’, ‘yg’, ‘jujur’, ‘siapa’, ‘yg’, ‘bohong’, ‘jilat’]

Ekstraksi TF-IDF

Pada tahap ekstraksi fitur, dilakukan perhitungan bobot setiap kata atau term dalam dokumen menggunakan metode TF-IDF. Proses ini menghasilkan representasi numerik dari kata-kata yang terdapat dalam setiap dokumen, sehingga memudahkan dalam proses analisis selanjutnya. Tabel 5 menunjukkan hasil ekstraksi TF-IDF, di mana setiap kolom berisi *term* hasil pemrosesan teks, dan setiap baris (D1, D2, D3, D4, dan D5) merepresentasikan dokumen yang berisi kumpulan term dari kalimat-kalimat yang dianalisis. Nilai pada tabel menunjukkan bobot pentingnya suatu *term* dalam dokumen tertentu.

Tabel 5. Hasil ekstraksi TF-IDF

No	abal	abis	abius	abu	acara	acau	…	yup	zaenal	zaman
D1	0	0	0	0	0	0	…	0	0	0
D2	0	0	0	0	0	0.393	…	0	0	0
D3	0	0	0	0	0	0	…	0	0	0
D4	0	0	0	0	0	0	…	0	0	0
D5	0	0	0	0	0	0	…	0	0	0

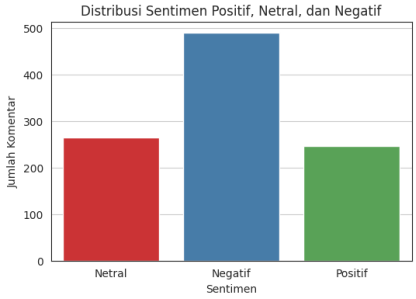
Tabel 5 menunjukkan hasil ekstraksi fitur menggunakan metode TF-IDF pada lima dokumen (D1, D2, D3, D4, dan D5) yang diperoleh dengan menggunakan *TfidfVectorizer* pada data teks yang telah dibersihkan. Sebagian besar nilai dalam tabel bernilai 0 karena *term* yang ditampilkan tidak muncul dalam dokumen-dokumen tersebut,

sehingga bobot TF-IDF-nya bernilai 0. Sebaliknya, jika sebuah *term* muncul dalam dokumen, maka akan menghasilkan nilai TF-IDF lebih dari 0, seperti pada D2 untuk *term* "acau" yang memiliki nilai 0.393, menandakan bahwa *term* tersebut muncul dan memiliki tingkat kepentingan tertentu dalam dokumen tersebut dibandingkan dengan keseluruhan korpus.

Model Klasifikasi

Dalam model klasifikasi yang digunakan terdapat empat skenario berbeda, yaitu KNN, KNN dengan SMOTE, *Naïve Bayes*, dan *Naïve Bayes* dengan SMOTE. Setiap skenario ini bertujuan untuk menguji performa model dalam mengklasifikasikan data, dengan atau tanpa teknik penyeimbangan data menggunakan SMOTE.

Pertama, metode KNN digunakan sebagai skenario klasifikasi dalam penelitian ini, dengan penerapan *10-fold cross validation* untuk menjaga kestabilan hasil. Pengujian dilakukan terhadap nilai k dari 3 hingga 10 untuk menemukan parameter terbaik. Sebelum data diseimbangkan dengan SMOTE, Gambar 2 menunjukkan distribusi awal sentimen guna menggambarkan ketidakseimbangan kelas. Tabel 6 menyajikan hasil kinerja model pada skenario pertama dengan metode KNN.

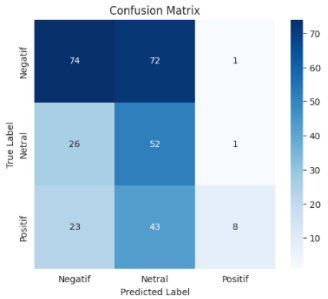


Gambar 2. Distribusi sentimen positif, netral, dan negatif KNN sebelum SMOTE

Tabel 6. Hasil performa KNN

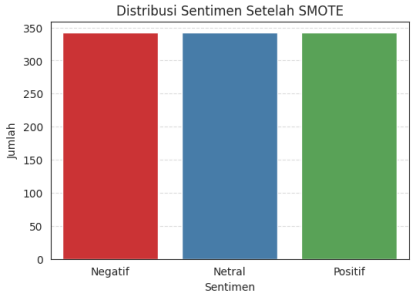
K	*Accuracy*	*Precision*	*Recall*	*F-1 Score*
3	0.273	0.414	0.273	0.131
4	0.288	0.402	0.288	0.154
5	0.282	0.262	0.282	0.146
6	0.311	0.313	0.311	0.196
7	0.298	0.219	0.298	0.181
8	0.337	0.441	0.337	0.261
9	0.365	0.527	0.365	0.311
10	0.402	0.545	0.402	0.364

Berdasarkan nilai akurasi yang diperoleh untuk setiap parameter KNN pada Tabel 6, parameter dengan performa terbaik adalah k dengan nilai 10, yang menghasilkan akurasi sebesar 40,2%. Selanjutnya, *confusion matrix* dihitung untuk parameter tersebut dan ditampilkan pada Gambar 2.



Gambar 2. *Confusion matrix* KNN

Pada skenario kedua, dilakukan optimasi dengan menerapkan SMOTE untuk menyeimbangkan distribusi kelas sentimen. Gambar 3 menunjukkan distribusi sentimen setelah SMOTE. Pengujian model tetap menggunakan *10-fold cross validation* dan variasi nilai k dari 3 hingga 10 pada metode KNN.



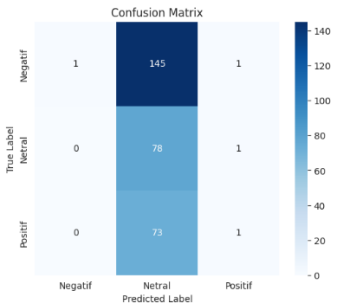
Gambar 3. Distribusi sentimen setelah SMOTE

Berdasarkan pada Gambar 3, setelah diterapkannya metode SMOTE, jumlah data untuk sentimen positif dan negatif menjadi seimbang, masing-masing sebanyak 342 komentar. Dengan distribusi data yang telah seimbang, berikut ditampilkan hasil pelatihan model KNN yang telah dioptimalkan menggunakan SMOTE pada Tabel 7.

Tabel 7. Hasil performa KNN dan SMOTE

K	*Accuracy*	*Precision*	*Recall*	*F-1 Score*
3	0.271	0.372	0.271	0.136
4	0.271	0.397	0.271	0.133
5	0.268	0.250	0.268	0.127
6	0.266	0.189	0.266	0.123
7	0.267	0.251	0.267	0.125
8	0.267	0.243	0.267	0.123
9	0.267	0.243	0.267	0.123
10	0.268	0.231	0.268	0.122

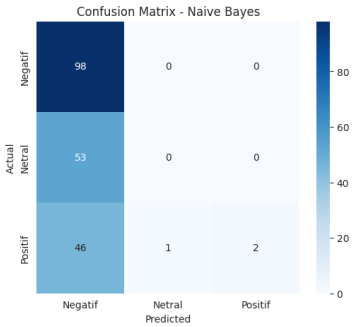
Berdasarkan nilai akurasi dari setiap parameter KNN yang tercantum pada Tabel 7, nilai optimal dalam penerapan metode SMOTE diperoleh ketika parameter k bernilai , dengan akurasi mencapai 27,1%. Selanjutnya, perhitungan *confusion matrix* dilakukan untuk parameter k tersebut dan hasilnya ditampilkan pada Gambar 4.



Gambar 4. *Confusion matrix* KNN dan SMOTE

Gambar 4 menunjukkan bahwa model sangat bias terhadap kelas netral, dengan sebagian besar data dari semua kelas diprediksi sebagai netral, sehingga performa klasifikasi untuk kelas negatif dan positif sangat rendah.

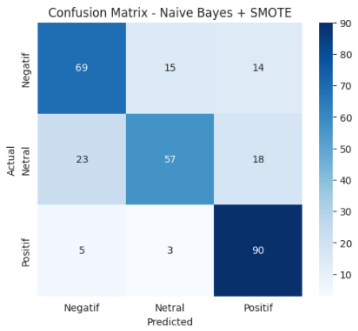
Pada skenario ketiga, model *Multinomial Naïve Bayes* digunakan untuk klasifikasi dengan pembagian data 80% latih dan 20% uji. Model dilatih pada data latih dan digunakan untuk memprediksi sentimen pada data uji. Gambar 5 menampilkan *confusion matrix* dari hasil prediksi model.



Gambar 5. *Confusion matrix Naïve Bayes*

Berdasarkan Gambar 5, model *Naïve Bayes* cenderung memprediksi sebagian besar komentar sebagai negatif. Sebanyak 98 komentar negatif berhasil diklasifikasikan dengan benar, namun seluruh komentar netral dan sebagian besar komentar positif salah diprediksi sebagai negatif. Hal ini menunjukkan ketidakseimbangan model dalam mengenali sentimen, terutama untuk kelas netral dan positif.

Pada skenario keempat, teknik SMOTE diterapkan sebelum pelatihan model *Naïve Bayes* untuk mengatasi ketidakseimbangan kelas. Data hasil *resampling* dibagi menjadi 80% data latih dan 20% data uji. Model kemudian dilatih dan dievaluasi menggunakan data tersebut, dengan hasil prediksi divisualisasikan dalam *confusion matrix* yang dapat dilihat pada Gambar 6.



Gambar 6. *Confusion matrix Naïve Bayes* dan SMOTE

Berdasarkan Gambar 6, model Naive Bayes dengan SMOTE berhasil mengklasifikasikan dengan cukup baik, dengan 69 prediksi benar untuk sentimen negatif, 57 untuk netral, dan 90 untuk positif. Model menunjukkan performa seimbang dalam membedakan ketiga jenis sentimen.

Evaluasi

Dalam proses pelatihan model klasifikasi, dilakukan dua jenis skenario pelatihan karena adanya ketidakseimbangan pada data. Ketidakseimbangan ini dikhawatirkan dapat mempengaruhi kinerja model yang dihasilkan. Oleh sebab itu, disiapkan skenario jenis kedua dengan menerapkan teknik resampling, khususnya metode *oversampling* menggunakan SMOTE, untuk mengatasi permasalahan tersebut.

Model klasifikasi seperti KNN dan *Naïve Bayes* menunjukkan tingkat kesalahan prediksi tinggi pada kelas sentimen negatif dan netral akibat ketidakseimbangan data. Ketidakseimbangan ini membuat model bias terhadap kelas mayoritas. Untuk mengatasinya, diterapkan metode SMOTE sebagai teknik *oversampling* guna menyeimbangkan distribusi data dan meningkatkan akurasi pada semua kelas. Perbandingan akurasi sebelum dan sesudah SMOTE ditampilkan pada Tabel 8.

Tabel 8. Hasil performa semua model

Performance	Model KNN	Model KNN+SMOTE	Model *Naïve Bayes*	Model *Naïve Bayes* + SMOTE
Accuracy (Macro)	45%	27%	50%	73%
Precision (Macro)	57%	53%	50%	73%
Recall (Macro)	42%	34%	35%	73%
F1-Score (Macro)	39%	15%	25%	73%

Berdasarkan Tabel 8, model *Naïve Bayes* dengan SMOTE menunjukkan performa terbaik dengan akurasi, presisi, *recall*, dan *F1-score* (*Macro*) sekitar 73%. SMOTE secara signifikan meningkatkan kinerja *Naïve Bayes* dibanding tanpa SMOTE yang hanya mencapai *F1-score* 25%. Namun, penerapan SMOTE pada KNN justru menurunkan performa drastis (akurasi 27%, *F1-score* 15%) karena KNN sensitif terhadap *noise* dan data sintetis dari SMOTE yang tidak selalu mewakili pola asli, sehingga mengganggu perhitungan jarak dan menyebabkan bias ke kelas mayoritas. Ini menunjukkan efektivitas SMOTE tergantung pada algoritma dan paling optimal untuk Naïve Bayes.

Berdasarkan komentar mengenai ijazah Presiden Jokowi yang telah dikumpulkan, berikut merupakan visualisasi sentimen publik terhadap ijazah Presiden Jokowi.

Gambar 7. Visualisasi Sentimen

Gambar 7 menampilkan *word cloud* untuk masing-masing sentimen. *Word cloud* positif berisi kata-kata seperti "orang", "ugm", dan "asli" yang mencerminkan pandangan mendukung. *Word cloud* negatif didominasi kata seperti "yg", "bohong", dan "salah" yang menunjukkan kritik. Sementara itu, *word cloud* netral memuat kata seperti "ugm", "lulus" dan "asli" yang bersifat informatif tanpa muatan emosional.

4. KESIMPULAN DAN SARAN

Hasil menunjukkan bahwa kombinasi *Naïve Bayes* dengan SMOTE memberikan performa terbaik, dengan nilai akurasi, presisi, *recall*, dan *F1-score (Macro)* masing-masing sebesar 73%. Di sisi lain, KNN bekerja cukup baik secara umum tanpa SMOTE, namun performanya justru menurun ketika SMOTE diterapkan. Kondisi ini menegaskan bahwa efektivitas teknik penyeimbangan data sangat bergantung pada algoritma yang digunakan, sehingga pemilihan model dan metode penyeimbangan harus disesuaikan dengan karakteristik data dan algoritma yang dipakai.

Saran bagi peneliti berikutnya dianjurkan untuk mengeksplorasi beragam teknik dan metode dalam upaya meningkatkan kinerja model, termasuk metode penyeimbangan data dan penggabungan algoritma. Pengujian yang lebih komprehensif dapat membantu menemukan pendekatan yang lebih efektif dan aplikatif pada berbagai jenis data serta permasalahan klasifikasi.

5. DAFTAR PUSTAKA

[1] F. S. Pratiwi, "Peran Komunikasi Digital dalam Pembentukan Opini Publik : Studi Kasus Media Sosial," pp. 293–315, 2024, doi: 10.30589/proceedings.2024.1059

[2] D. Al Mustaqim, F. Abdul Hakim, H. Atfalina, and A. Fatakh, "Peran Media Sosial Sebagai Sarana Partisipasi Warganet Dalam Mewujudukan Keadilan dan Akuntabilitas Penegakan Hukum di Indonesia," *J. Multidiscip. Res. Dev.*, vol. 1, no. 1, pp. 53–66, 2024, doi: 10.56916/jmrd.v1i1.655

[3] Aulia A S, M. Rezani Alwi, Subandi, and Muhammad R H, "Analisis Sentimen Data Twitter Terhadap Pelaksanaan Pembelajaran Online Di Indonesia Pada Masa Pandemi Covid-19 Menggunakan Metode Natural Language Processing," *Jursistekni* , vol. 5, no. 3, pp. 319–331, 2023, doi: 10.52005/jursistekni.v5i3.205

[4] A. A. Santoso, "Penetapan Ganjar Pranowo Sebagai Calon Presiden: Studi Analisis Topik pada Reverse Agenda Setting Terkait Kasus Ganjar Pranowo," *JIIP - J. Ilm. Ilmu Pendidik.*, vol. 7, no. 4, pp. 3805–3812, 2024, doi: 10.54371/jiip.v7i4.4140

[5] N. Alfriyanto, B. C. Purnama, and F. K. Hasanah, "Analisis Emosi Terhadap Komentar Video Youtube ' Penyebab Kegagalan Adopsi Sistem Pendidikan Finlandia di Indonesia ' Menggunakan Metode Random Forest," pp. 812–827, 2024. Diperoleh dari https://centive.ittelkom-

pwt.ac.id/index.php/centive/article/view/409
[6] N. M. Putry, "Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Mellitus," *EVOLUSI J. Sains dan Manaj.*, vol. 10, no. 1, 2022, doi: 10.31294/evolusi.v10i1.12514
[7] F. Destiyanti, A. I. Hadiana, and F. R. Umbara, "Penerapan Metode Support Vector Machine dan SMOTE untuk Klasifikasi Sentimen Publik Terhadap Polisi Republik Indonesia," vol. 8, no. 1, pp. 1–15, 2024, doi: 10.26874/jumanji.v8i1.336
[8] A. K. S, H. Hs, and Z. F. K, "Media 2024 Analisis Sentimen Masyarakat Terhadap Pelantikan Kabinet Merah Putih Pada Media Sosial Menggunakan Algoritma Naive Bayes," pp. 1080–1089, 2024. Diperoleh dari https://centive.ittelkom-pwt.ac.id/index.php/centive/article/view/383
[9] A. K. Wardhani *et al.*, "Optimasi Nilai K Pada Algortima K-Nearest," vol. 5, no. 1, pp. 44–51, 2025, doi: 10.54840/jcstech.v5i1.360
[10] I Gusti Ngurah Ady Kusuma, I Made Pradipta, I Made Ari Santosa, and I Komang Dharmendra, "Penanganan Ketidakseimbangan Data Pada Klasifikasi Pengaduan Masyarakat," *J. Teknol. Inf. dan Komput.*, vol. 9, no. 5, pp. 489–496, 2023, doi: 10.36002/jutik.v9i5.2643
[11] M. Azhari, "Analisis Sentimen Opini Publik Program Makan Siang Gratis dengan Random Forest Pada Media X," vol. 6, no. 3, pp. 1932–1942, 2024, doi: 10.47065/bits.v6i3.6423
[12] O. INews, "UGM Klaim Ijazah Jokowi Asli, Roy Suryo: UGM Bohong! | Rakyat Bersuara | 29/04."
[13] Apify, "Web Scraper," Apify Documentation, [Online]. Available: https://apify.com/streamers/youtube-scraper [Accessed: May 8, 2025].
[14] Masdevid, "ID-OpinionWords," GitHub, [Online]. Available: https://github.com/masdevid/ID-OpinionWords [Accessed: May 8, 2025].
[15] P. S. Zalukhu, T. Handhayani, and M. Sitorus, "Analisis Sentimen Terhadap Kenaikan Bbm Di Indonesia Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 8, no. 1, pp. 65–69, 2023, doi: 10.51876/simtek.v8i1.177
[16] Subkhi Mahmasani, "View metadata, citation and similar papers at core.ac.uk," vol. 8, no. 1, pp. 274–282, 2020. Diperoleh dari https://core.ac.uk/reader/212014220
[17] A. D. Adhi Putra, "Sentiment Analysis on User Reviews of the Bibit and Bareksa Application with the KNN Algorithm," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962
[18] R. Sari, "Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi Menggunakan Algoritma K-Nearest Neighbor (K-Nn)," *EVOLUSI J. Sains dan Manaj.*, vol. 8, no. 1, pp. 10–17, 2020, doi: 10.31294/evolusi.v8i1.7371
[19] M. Maulidah, Windu Gata, Rizki Aulianita, and Cucu Ika Agustyaningrum, "Algoritma Klasifikasi Decision Tree Untuk Rekomendasi Buku Berdasarkan Kategori Buku," *E-Bisnis J. Ilm. Ekon. dan Bisnis*, vol. 13, no. 2, pp. 89–96, 2020, doi: 10.51903/e-bisnis.v13i2.251