Analisis Sentimen Dampak Putusan MK Batas Usia Minimum Capres-Cawapres dengan SVM, Naïve Bayes, dan KNN

Aldi Lukmana, Galih Hendro Martono, Neny Sulistianingsih

Universitas Bumigora, Mataram, Indonesia

e-mail: [1]aldilukmana1313@gmail.com,

Abstrak

Mahkamah Konstitusi (MK) berperan penting dalam menegakkan konstitusi, termasuk menetapkan batas usia minimum pencalonan Presiden dan Wakil Presiden. Putusan ini memicu beragam reaksi di media sosial, mulai dari dukungan hingga penolakan yang dinilai politis. Penelitian ini bertujuan menganalisis sentimen publik terhadap putusan tersebut menggunakan Support Vector Machine (SVM), Naïve Bayes (NB), dan K-Nearest Neighbors (KNN), serta membandingkan kinerjanya berdasarkan akurasi, presisi, recall, dan F1-score. Penelitian dilakukan melalui enam tahap: ***(1)*** *Business Understanding – menentukan kebutuhan, tujuan, dan pengumpulan data;* ***(2)*** *Data Understanding – mengumpulkan, mendeskripsikan, dan mengevaluasi kualitas data;* ***(3)*** *Data Preparation – membersihkan, memilih, dan mentransformasi data;* ***(4)*** *Modelling – menerapkan algoritma SVM, NB, dan KNN;* ***(5)*** *Evaluation – mengukur kinerja model menggunakan confusion matrix; serta* ***(6)*** *Deployment – menyusun laporan dan dokumentasi hasil analisis. Data diambil dari media sosial X dan YouTube, diolah menggunakan teknik text mining dan machine learning. Hasil menunjukkan SVM dan KNN memiliki akurasi tertinggi, masing-masing 89,5%, sedangkan NB mencapai 88,5%, sehingga SVM dan KNN dinilai lebih efektif dalam menganalisis sentimen publik terhadap putusan MK.*

Kata kunci: *Analisis Sentimen, Mahkamah Konstitusi, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors.*

Abstract

The Constitutional Court plays a crucial role in upholding the constitution, including setting the minimum age requirement for presidential and vice-presidential candidates. This decision sparked diverse reactions on social media, ranging from support to politically charged opposition. This study aims to analyze public sentiment toward the ruling using Support Vector Machine (SVM), Naïve Bayes (NB), and K-Nearest Neighbors (KNN), and to compare their performance based on accuracy, precision, recall, and F1-score. The research followed six stages: Business Understanding, defining needs, objectives, and data collection; Data Understanding, gathering, describing, and evaluating data quality; Data Preparation, cleaning, selecting, and transforming data; Modeling, applying the SVM, NB, and KNN algorithms; Evaluation, assessing model performance using a confusion matrix; and Deployment, preparing reports and documenting the analysis results. The data were collected from social media platforms X and YouTube and processed using text mining and machine learning techniques. The results indicate that SVM and KNN achieved the highest accuracy, while NB followed closely, suggesting that SVM and KNN are more effective for analyzing public sentiment regarding the Constitutional Court's decision.

Keywords: Sentiment Analysis, Constitutional Court, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors.

1. Pendahuluan

Mahkamah Konstitusi (MK) merupakan lembaga peradilan yang memiliki kewenangan untuk menguji undang-undang terhadap Undang-Undang Dasar (UUD) serta menangani sengketa hasil pemilu [1 Keputusan yang dikeluarkan oleh MK sering kali memiliki dampak luas terhadap sistem hukum dan politik di Indonesia. Salah satu putusan MK yang menuai banyak perhatian publik adalah keputusan terkait batas usia minimum pencalonan Presiden dan Wakil Presiden. Keputusan ini memicu berbagai reaksi dari berbagai kalangan, termasuk akademisi, politisi, aktivis, serta masyarakat umum [2]. Di era digital, media sosial seperti X, Facebook, dan Instagram menjadi ruang diskusi publik yang memungkinkan masyarakat mengekspresikan pendapat secara cepat [3]. Sentimen terhadap putusan MK

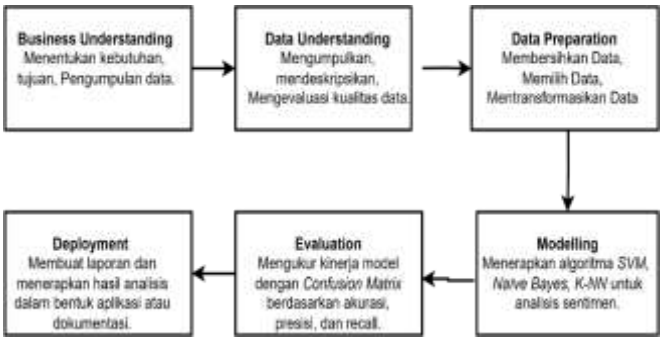
tersebut bervariasi: sebagian mendukung karena memberi peluang bagi pemimpin muda, sementara lainnya menolak dengan alasan politis. Analisis sentimen diperlukan untuk memahami persepsi publik secara objektif dan mengevaluasi dampak sosial kebijakan.

Teknik *text mining* dan *machine learning* banyak digunakan untuk klasifikasi sentimen, di antaranya *Support Vector Machine* (SVM), *Naïve Bayes* (NB), dan *K-Nearest Neighbors* (KNN) [4]. Penelitian inibertujuan menganalisis sentimen publik terhadap putusan MK dengan ketiga metode tersebut, membandingkan kinerjanya, dan menentukan metode yang paling akurat dalam klasifikasi sentimen.

2. Metode Penelitian

Penelitian ini melibatkan beberapa tahapan utama. Tahap pertama adalah pengumpulan data, di mana data dikumpulkan dari media sosial X dan platform youtube menggunakan teknik web scraping. Data dikumpulkan dari periode bulan april 2023 hingga bulan april 2025, data dikumpulkan dalam bentuk file .csv, data yang berhasil dikumpulkan sebanyak 1233 dari platform media youtube dan dari prlatform sosial media X sebanyak 1900 data. Setelah data terkumpul, tahap berikutnya adalah preprocessing data yang mencakup beberapa langkah penting seperti cleaning data, case folding, tokenizing, stopwords removal, dan stemming untuk memastikan data dalam format yang siap digunakan. Setelah preprocessing selesai, data kemudian dilabeli menggunakan proses TF-IDF untuk analisis sentimen. Proses pelabelan ini bertujuan untuk mengklasifikasikan data ke dalam kategori sentimen tertentu. Selanjutnya, dilakukan pembuatan model klasifikasi dengan menerapkan algoritma Support Vector Machine (SVM), Naïve Bayes (NB), dan K-Nearest Neighbors (KNN).

Tahap akhir dari penelitian ini adalah evaluasi model. Model yang telah dibangun dievaluasi berdasarkan beberapa metrik performa utama, yaitu akurasi, presisi, recall, dan F1-score. Evaluasi ini bertujuan untuk menentukan efektivitas model dalam mengklasifikasikan sentimen dengan tingkat keakuratan yang optimal[5], [6]. Berikut ditunjukan metodologi penelitian pada gambar 1.



Gambar 1. Metodologi Penelitian

2.1 Business Understanding

Penelitian ini bertujuan menganalisis sentimen publik terhadap keputusan Mahkamah Konstitusi mengenai batas usia minimum pencalonan presiden dan wakil presiden. Reaksi publik yang beragam, baik positif maupun negatif, dikumpulkan dari media sosial X dan YouTube. Analisis sentimen dilakukan untuk memetakan persepsi masyarakat dan menilai dampak potensial keputusan tersebut terhadap dinamika sosial-politik [12].

Support Vector Machine (SVM) telah lama menjadi salah satu algoritma andalan dalam analisis sentimen. Penelitian awal oleh [13], menunjukkan keunggulan SVM dalam mengklasifikasikan teks berdimensi tinggi, sementara [14], membuktikan bahwa SVM mampu mengungguli metode berbasis kamus dalam klasifikasi polaritas ulasan film. Pengembangan lebih lanjut oleh [15]varian NB-SVM yang menggabungkan kekuatan pembobotan Naïve Bayes dan margin maksimal SVM untuk meningkatkan akurasi. Studi lain seperti [16], [17] juga menunjukkan bahwa SVM dapat dioptimalkan dengan representasi vektor kata berlabel sentimen maupun penambahan fitur linguistik, sehingga tetap menjadi pilihan kuat untuk berbagai jenis data teks, termasuk media sosial.

Naïve Bayes (NB) dikenal sebagai metode probabilistik sederhana namun efektif. Penelitian [18], dan [19] menjadikannya baseline kuat dalam analisis sentimen dokumen. Penelitian [20] dan [21], menunjukkan bahwa varian NB, seperti multinomial dan bernoulli, memberikan perbedaan signifikan dalam kinerja klasifikasi teks. Dalam konteks media sosial, [21] membangun korpus Twitter secara otomatis dan membuktikan efektivitas NB untuk data mikroblog, sementara tinjauan [22] menegaskan keunggulannya dalam hal kecepatan dan efisiensi pada dataset besar. Namun, asumsi independensi fitur menjadi kelemahan yang dapat memengaruhi akurasi pada data teks yang kaya konteks.

K-Nearest Neighbor (KNN) merupakan metode non-parametrik yang menentukan kelas

berdasarkan kedekatan jarak dengan tetangga terdekat. Konsep dasarnya diperkenalkan oleh *Cover dan Hart (1967)*, dengan pengembangan seperti Weighted-KNN oleh [23] yang mampu meningkatkan akurasi melalui pembobotan fitur. Dalam analisis sentimen, KNN terbukti kompetitif sebagaimana dilaporkan pada studi komparatif terbaru yang membandingkannya dengan SVM, NB, dan Decision Tree. Penerapan KNN pada analisis ulasan marketplace dan Twitter menunjukkan akurasi tinggi pada fitur sederhana seperti TF-IDF, meskipun kelemahan utama tetap pada kebutuhan komputasi yang tinggi saat jumlah data latih membesar.

2.2 Data Understanding

Data diperoleh melalui scraping komentar dari media sosial X dan YouTube menggunakan *tweet-harvest* dan YouTube Data API v3. Dataset berisi atribut seperti nama pengguna, isi komentar, dan waktu publikasi. Evaluasi kualitas data dilakukan dengan memeriksa duplikasi, *missing values*, dan inkonsistensi.

2.3 Data Preparation

Tahapan preprocessing meliputi:

- *Cleaning*: menghapus tanda baca, angka, simbol, dan elemen tak relevan.
- *Case Folding*: mengubah semua teks menjadi huruf kecil.
- Tokenisasi: memisahkan teks menjadi kata-kata.
- *Stopword* Removal: menghapus kata umum yang tidak bermakna analitis.
- *Stemming*: mengubah kata ke bentuk dasar.
- TF-IDF: memberi bobot kata berdasarkan frekuensi relatif terhadap dokumen.
- TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah metode yang digunakan untuk membobot kata-kata dalam dokumen [24]. Kata yang lebih sering muncul dalam satu dokumen tapi jarang muncul di dokumen lain akan diberi bobot lebih tinggi. Rumus TF-IDF ditunjukkan pada Persamaan (1)–(3).

$$TF - IDF(t,d) = TF(t,d) \times IDF(t) \quad (1)$$

$$TF(t,d) = \frac{\text{jumlah kemunculan term } t \text{ dalam } d}{\text{jumlah total kata dalam } d} \quad (2)$$

$$IDF(t) = \log \frac{N}{df(t)} \quad (3)$$

2.4 Modeling

Dalam penelitian ini digunakan tiga algoritma utama untuk melakukan klasifikasi sentimen, yaitu *Support Vector Machine* (SVM*), Naïve Bayes* (NB), dan *K-Nearest Neighbor* (KNN).

1. Support Vector Machine (SVM), bekerja dengan mencari *hyperplane* optimal yang memisahkan data ke dalam kelas sentimen berbeda secara maksimal. Metode ini sangat efektif dalam menangani data berdimensi tinggi, seperti teks yang telah diubah menjadi representasi numerik melalui teknik pembobotan seperti TF-IDF [13]. Keunggulan SVM terletak pada kemampuannya menjaga margin pemisah yang besar, sehingga mampu meminimalkan kesalahan klasifikasi dan mengurangi risiko *overfitting*.
2. Naïve Bayes (NB), merupakan algoritma berbasis probabilitas yang memanfaatkan Teorema Bayes dengan asumsi independensi antar fitur [18]. Pada analisis sentimen, NB menghitung probabilitas sebuah dokumen termasuk dalam kategori positif, negatif, atau netral berdasarkan distribusi kata yang terdapat di dalamnya [25]. Algoritma ini dikenal sederhana, cepat, dan tetap memberikan performa yang kompetitif meskipun data latih terbatas.
3. K-Nearest Neighbor (KNN), adalah metode non-parametrik yang menentukan label suatu data berdasarkan mayoritas kelas dari *k* tetangga terdekat [26]. Ukuran kedekatan antar data biasanya dihitung menggunakan *Euclidean Distance* [27]. Dalam analisis sentimen, KNN membandingkan vektor fitur dari dokumen uji dengan seluruh dokumen latih, kemudian menetapkan kelas sesuai dominasi tetangga terdekat tersebut.

3.5 Evaluasi Model

Evaluasi klasifikasi didasarkan pada pengujian pada hal-hal yang tepat dan tidak pantas. Dari hasil klasifikasi, jenis model optimal dipilih menggunakan validasi ini. Evaluasi penelitian ini menggunakan matriks konfusi [28]. Model algoritma klasifikasi dapat memprediksi matriks konfusi, yang merupakan informasi hasil klasifikasi data nyata. Jumlah TP dibandingkan dengan jumlah catatan positif menggunakan tabel matriks, sedangkan jumlah TN dibandingkan dengan jumlah catatan negatif. Tabel klasifikasi biner yang menunjukkan data yang digunakan bersifat *True Positives (*TP*), False Positives (*FP*), False Negatives*

*(*FN*), dan True Negatives (*TN*).* Rumus untuk menghitung *confusion matriks* dapat dilihat pada tabel 2.

Tabel 1. Matriks Confusi

	Prediksi	
	Positive	*Negative*
Positive	*True Positive (TP)*	*False Positive (FP)*
Negative	*False Negative (FN)*	*True Negative (TN)*

Pada Tabel 2 *confusion matrix* di atas merupakan alat yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan cara membandingkan hasil prediksi model terhadap nilai sebenarnya. Tabel ini terdiri dari empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) [29]. True Positive menunjukkan jumlah data yang benar-benar positif dan diprediksi positif oleh model, sedangkan True Negative menunjukkan jumlah data yang benar-benar negatif dan berhasil diprediksi negative [29]. False Positive merupakan kesalahan ketika model memprediksi data sebagai positif padahal sebenarnya negative. Keempat komponen ini menjadi dasar dalam menghitung metrik evaluasi seperti akurasi, presisi, dan recall untuk menilai seberapa baik model dalam melakukan klasifikasi [29].

3. Hasil dan Pembahasan

3.1. Dataset

Data dikumpulkan dari media sosial X dan platform youtube menggunakan teknik web scraping. Data dikumpulkan dari periode bulan april 2023 hingga bulan april 2025, data dikumpulkan dalam bentuk file .csv, data yang berhasil dikumpulkan sebanyak 1233 dari platform media youtube dan dari prlatform sosial media X sebanyak 1900 data.

Tabel 2.Data Sebelum Preprocessing

No	created_at	full_text	username
1	Mon Mar 25 03:43:15 +0000 2024	@abulhasanshalih @silabukayu @aanniinnd @NenkMonica @prabowo Rekam jejak 2x walikota solo 1x gubernur DKI Jakarta 2x presiden RI Oh ya yg ajukan batas usia capres/cawapres itu bukan kelompok Gibran dan yg mengajukan gugatan ke MK hanya fans dan yg pasti UU itu tidak berlaku buat Gibran aj dan yg pasti putusan MK final dan mengikat	mochyus10193354
2	Sat Mar 23 18:27:24 +0000 2024	@kegblgnunfaedh Padahal jelas terbukti loh keterlibatan ketua MK pada kasus meloloskan syarat batas usia capres cawapres dalam UU Pemilu untuk kepentingan salah satu pihak tapi orang ttep aja nyangkal pdhal udah ada putusan MKMK sbgai produk hukumnya	openworld33
3	Sat Mar 23 04:00:35 +0000 2024	inilah yang dapat dijelaskan dari fakta politik lahirnya putusan MK soal batas usia capres-cawapres tuntutan revisi PKPU dan penetapan Gibran sebagai cawapres dari Golkar usai putusan MK. #MengkritikDemokrasi #IndonesiaLebihBaikDenganIslam	brynn_brew69327

3.2. Preprocessing

Tahap preprocessing dilakukan untuk mempersiapkan data sebelum masuk ke proses pembobotan dan klasifikasi. Untuk menggambarkan hasil preprocessing, berikut ditampilkan data sebelum dan data sesudah preprocessing. Pada tahap ini, karakter tidak penting seperti emotikon, tanda baca, simbol, dan URL dihapus. Hasil *preprocessing* data ditunjukan pada tabel 3.

Tabel 3.Hasil Preprocessing

No	Stopword Removal	Stemming	Sentimen
1	['rekam', 'jejak', 'walikota', 'solo', 'gubernur', 'dki', 'jakarta', 'presiden', 'ri', 'ajukan', 'batas', 'usia', 'capres', 'cawapres', 'kelompok', 'gibran', 'mengajukan', 'gugatan', 'mk', 'fans', 'uu', 'berlaku', 'gibran', 'putusan', 'mk', 'final', 'mengikat']	['rekam', 'jejak', 'walikota', 'solo', 'gubernur', 'dki', 'jakarta', 'presiden', 'ri', 'ajuk', 'batas', 'usia', 'capres', 'cawapres', 'kelompok', 'gibran', 'ajuk', 'gugat', 'mk', 'fans', 'uu', 'laku', 'gibran', 'putus', 'mk', 'final', 'ikat']	Netral
2	['jelas', 'terbukti', 'keterlibatan', 'ketua', 'mk', 'kasus', 'meloloskan', 'syarat', 'batas', 'usia', 'capres', 'cawapres', 'uu', 'pemilu', 'kepentingan', 'pihak', 'orang', 'nyangkal', 'putusan', 'mkmk', 'produk', 'hukumnya']	['jelas', 'bukti', 'libat', 'ketua', 'mk', 'kasus', 'lolos', 'syarat', 'batas', 'usia', 'capres', 'cawapres', 'uu', 'pemilu', 'penting', 'pihak', 'orang', 'nyangkal', 'putus', 'mkmk', 'produk', 'hukum']	Negatif
3	['fakta', 'politik', 'lahirnya', 'putusan', 'mk', 'batas', 'usia', 'capres', 'cawapres', 'tuntutan', 'revisi', 'pkpu', 'penetapan', 'gibran', 'cawapres', 'golkar', 'putusan', 'mk']	['fakta', 'politik', 'lahir', 'putus', 'mk', 'batas', 'usia', 'capres', 'cawapres', 'tuntut', 'revisi', 'pkpu', 'tetap', 'gibran', 'cawapres', 'golkar', 'putus', 'mk']	Netral

Dari hasil tersebut terlihat bahwa data telah dibersihkan dan disederhanakan hanya menjadi kata-kata inti yang memiliki makna penting. Proses ini bertujuan untuk mengurangi kompleksitas data dan memfokuskan pada fitur-fitur penting yang akan digunakan dalam pembobotan TF-IDF dan klasifikasi.

3.3. Pembobotan TF-IDF

TF-IDF digunakan untuk mengkonversi data teks menjadi representasi numerik yang dapat diproses oleh model pembelajaran mesin. Proses ini menghasilkan bobot numerik untuk setiap kata

berdasarkan frekuensi kemunculannya dan signifikansinya dalam seluruh dokumen. Visualisasi hasil pembobotan ditunjukkan pada gambar 3.

Gambar 3. Hasil TF-IDF

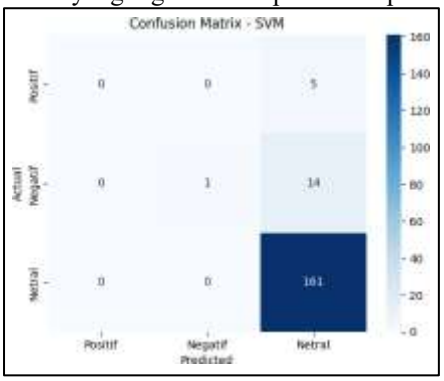
No	Term	TF A	TF B	TF C	TF D	DF	IDF A	IDF B	IDF C	IDF D
1	ajuk	2	0	0	0	1	0.602	0	0	0
2	batas	1	1	1	1	4	0.000	0.000	0.000	0.000
3	biasa	0	0	0	1	1	0	0	0	0.602
…	…	…	…	…	…	…	…	…	…	…
15	gugat	1	0	0	1	2	0.301	0	0	0.301

3.4. Evaluasi Model

Model dievaluasi berdasarkan metrik akurasi, presisi, dan recall dari confusion matrix. Tiga skenario evaluasi dilakukan: (a) menggunakan data original, dan (b) menggunakan data hasil SMOTE.

a. Evaluasi Support Vector Machine (SVM)

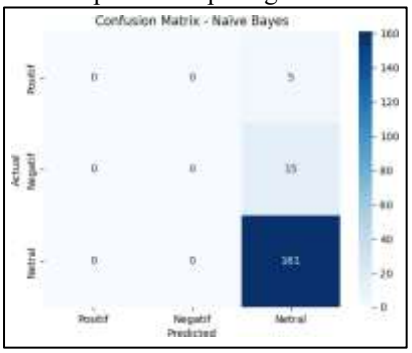
Setelah melalui tahap pelatihan, model *Support Vector Machine (SVM)* dievaluasi untuk mengukur performanya dalam melakukan klasifikasi terhadap data. Hasil *Confusion Matrix* yang menunjukkan hasil evaluasi dari model SVM pada dataset yang digunakan dapat dilihat pada gambar 4.



Gambar 4 Hasil Evaluasi Model *Support Vector Machine*

Hasil evaluasi menggunakan *Confusion Matrix* pada model *Support Vector Machine* (SVM) pada gambar 4 diatas menunjukkan bahwa model cenderung lebih sering mengklasifikasikan data ke dalam kategori netral, sementara kemampuan dalam membedakan kategori positif dan negatif masih kurang optimal. Hasil Evaluasi Naïve Bayes (NB)

Hasil *Confusion Matrix* yang menunjukkan hasil evaluasi model *Naïve Bayes (NB)* dalam melakukan klasifikasi pada dataset yang digunakan dapat dilihat pada gambar 5.

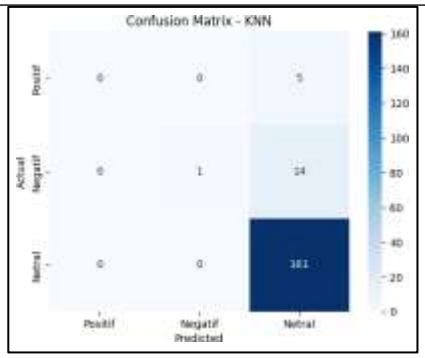


Gambar 5 Hasil Evaluasi Model Naive Bayes

Hasil evaluasi model *Naïve Bayes (NB)* pada gambar 5. diatas menggunakan *Confusion Matrix* menunjukkan bahwa model ini memiliki keterbatasan dalam membedakan kategori positif dan negatif, dengan kecenderungan mengklasifikasikan sebagian besar data ke dalam kategori netral..

b. Hasil Evaluasi K-Nearest Neigboard (KNN)

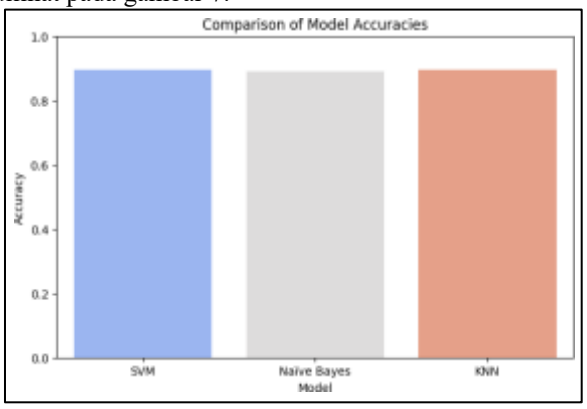
Hasil *Confusion Matrix* yang menunjukkan hasil evaluasi model *K-Nearest Neighbors* (KNN) pada dataset yang digunakan dapat dilihat pada gambar 6.



Gambar 6 Hasil Evaluasi Model KNN

Hasil evaluasi model *K-Nearest Neighbors* (KNN) menggunakan *Confusion Matrix* menunjukkan bahwa model ini memiliki kecenderungan untuk mengklasifikasikan sebagian besar data ke dalam kategori netral. Dari hasil klasifikasi, tidak ada data yang berhasil diprediksi sebagai positif, sementara 5 data positif dan 14 data negatif salah dikategorikan sebagai netral. Model hanya berhasil mengklasifikasikan 1 data negatif dengan benar, sedangkan 161 data netral berhasil diprediksi secara akurat. Hasil ini menunjukkan bahwa model KNN mengalami kesulitan dalam mengenali kategori positif dan negatif, yang mungkin disebabkan oleh distribusi data yang tidak seimbang atau pemilihan parameter K yang kurang optimal.

Setelah dilakukan evaluasi terhadap ketiga metode yang digunakan, yaitu *Support Vector Machine* (SVM*), Naïve Bayes* (NB), dan *K-Nearest Neighbors* (KNN), langkah selanjutnya adalah menganalisis hasil yang diperoleh. Analisis ini bertujuan untuk memahami sejauh mana setiap model mampu mengklasifikasikan data dengan akurasi yang optimal serta mengevaluasi kelebihan dan kekurangan masing-masing metode. Dalam tahap ini, perbandingan antara hasil evaluasi ketiga model akan ditampilkan dalam bentuk visualisasi untuk memudahkan interpretasi. Dengan adanya perbandingan ini, dapat diketahui model mana yang memiliki kinerja terbaik dalam menangani data yang digunakan. Selain itu, analisis ini juga akan memberikan wawasan lebih lanjut mengenai faktor-faktor yang memengaruhi performa model, seperti jumlah data latih, distribusi kelas, serta parameter yang digunakan dalam setiap algoritma. Visualisasi perbandingan hasil akurasi *Support Vector Machine* (SVM), *Naïve Bayes* (NB*),* dan *K-Nearest Neighbors* (KNN) dapat dilihat pada gambar 7.



Gambar 7. Hasil Perbandingan Akurasi

Berdasarkan hasil perbandingan akurasi dari ketiga metode yang digunakan, yaitu *Support Vector Machine* (SVM), *Naïve Bayes* (NB), dan *K-Nearest Neighbors* (KNN), dapat dilihat bahwa ketiga model memiliki performa yang hampir setara. Ketiganya menunjukkan tingkat akurasi yang cukup tinggi, berkisar di atas 85%, yang menandakan bahwa model dapat mengklasifikasikan data dengan baik. Dari hasil visualisasi, metode *Support Vector Machine (SVM)* dan *K-Nearest Neighbors* (KNN) menunjukkan tingkat akurasi yang lebih tinggi yaitu 89,5% dibandingkan dengan *Naïve Bayes* dan sebesar 88,5% Dengan demikian**,** *Support Vector Machine (SVM)* dan *K-Nearest Neighbors* (KNN) dapat dianggap sebagai metode terbaik dalam analisis sentimen publik terhadap keputusan MK. Hal ini menunjukkan bahwa setiap model memiliki keunggulan masing-masing dalam menangani data teks yang digunakan dalam penelitian ini. Hasil analisis sentimen menunjukkan bahwa opini publik terhadap keputusan Mahkamah Konstitusi cenderung netral**.** Faktor utama yang mempengaruhi sentimen publik meliputi opini yang lebih bersifat informatif dan tidak menunjukkan sikap yang kuat. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pemangku kebijakan dalam memahami opini publik dan merancang strategi komunikasi yang lebih efektif terhadap kebijakan politik yang kontroversial.

4. Kesimpulan

Berdasarkan hasil pengujian analisis sentimen publik terhadap keputusan Mahkamah Konstitusi (MK)

mengenai batas usia minimum pencalonan presiden dan wakil presiden dengan metode Support Vector Machine (SVM), Naïve Bayes (NB), dan K-Nearest Neighbors (KNN), diketahui bahwa ketiganya mampu mengklasifikasikan sentimen positif, negatif, dan netral dengan baik. Hasil evaluasi melalui confusion matrix menunjukkan SVM dan KNN memiliki akurasi tertinggi sebesar 89,5%, mengungguli NB yang memperoleh 88,5%. Hal ini mengindikasikan bahwa pendekatan berbasis hyperline pada SVM dan tetangga terdekat pada KNN lebih efektif dalam klasifikasi sentimen dibandingkan perhitungan probabilitas pada NB, meskipun masing-masing metode tetap memiliki keunggulan dan keterbatasan tergantung karakteristik data yang digunakan.

Daftar Pustaka

[1] Z. Setiawan, "Peran Mahkamah Konstitusi dalam Menjaga Stabilitas Hukum di Indonesia," vol. 16, no. 3, hal. 19–25, 2024.

[2] H. Inarni, *Tinjauan Hukum Islamterhadap Putusan Mahkamah Konstitusi Nomor 90/Puu-Xxi Tentang Batas Usia Calon Presiden Dan Wakil Presiden*, no. 3. 2021.

[3] S. Khairunnisa, A. Adiwijaya, dan S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, hal. 406, 2021, doi: 10.30865/mib.v5i2.2835.

[4] M. N. Muttaqin dan I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor," *UNNES J. Math.*, vol. 10, no. 2, hal. 22–27, 2021, [Daring]. Tersedia pada: http://journal.unnes.ac.id/sju/index.php/ujm

[5] M. A. Hasanah, S. Soim, dan A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," vol. 5, no. 2, 2021.

[6] Y. A. Singgalen, "Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT)," *J. Media Inform. Budidarma*, vol. 7, no. 3, hal. 1551, 2023, doi: 10.30865/mib.v7i3.6461.

[7] P. Edastama, A. S. Bist, dan A. Prambudi, "Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm," *Int. J. Cyber IT Serv. Manag.*, vol. 1, no. 2, hal. 159–172, 2021, doi: 10.34306/ijcitsm.v1i2.46.

[8] A. Sukarno Hatta, "Clustering Pada Data Sentimen Penggunaan Transportasi Online Menggunakan Algoritma Spectral Clustering," *e-Proceeding Eng.*, vol. 8, no. 6, hal. 11945, 2021.

[9] M. Soleimani, A. Intezari, dan D. J. Pauleen, "Mitigating cognitive biases in developing ai-assisted recruitment systems: A knowledge-sharing approach," *Int. J. Knowl. Manag.*, vol. 18, no. 1, hal. 1–18, 2022, doi: 10.4018/IJKM.290022.

[10] B. M. Iqbal, K. M. Lhaksmana, dan E. B. Setiawan, "2024 Presidential Election Sentiment Analysis in News Media Using Support Vector Machine," *J. Comput. Syst. Informatics*, vol. 4, no. 2, hal. 397–404, 2023, doi: 10.47065/josyc.v4i2.3051.

[11] H. Yun, "Prediction model of algal blooms using logistic regression and confusion matrix," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 3, hal. 2407–2413, 2021, doi: 10.11591/ijece.v11i3.pp2407-2413.

[12] S. Kumar, J. Thakur, D. Ekka, dan I. Sahu, "Web Scraping Using Python," *Int. J. Adv. Eng. Manag.*, vol. 4, no. 9, hal. 235, 2022, doi: 10.35629/5252-0409235237.

[13] M. A. Chandra dan S. S. Bedi, "Survey on SVM and their application in image classification," *Int. J. Inf. Technol.*, vol. 13, no. 5, 2021, doi: 10.1007/s41870-017-0080-1.

[14] Q. Wang, G. Zhu, S. Zhang, K. Li, X. Chen, dan H. Xu, "Extending emotional lexicon for improving the classification accuracy of Chinese film reviews," *Conn. Sci.*, vol. 0091, hal. 1–20, 2020, doi: 10.1080/09540091.2020.1782839.

[15] E. I. Setiawan, P. Tjendika, J. Santoso, F. X. Ferdinandus, Gunawan, dan K. Fujisawa, "Aspect-Based Sentiment Analysis of Healthcare Reviews from Indonesian Hospitals based on Weighted Average Ensemble," *J. Appl. Data Sci.*, vol. 5, no. 4, hal. 1579–1596, 2024, doi: 10.47738/jads.v5i4.328.

[16] H. K. Obayes, F. S. Al-Turaihi, dan K. H. Alhussayni, "Sentiment classification of user's reviews on drugs based on global vectors for word representation and bidirectional long short-term memory recurrent neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 1, hal. 345–353, 2021, doi: 10.11591/ijeecs.v23.i1.pp345-353.

[17] N. S. Mohd Nafis dan S. Awang, "An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification," *IEEE Access*, vol. 9, no. Ml, hal. 52177–52192, 2021, doi: 10.1109/ACCESS.2021.3069001.

[18] I. Wickramasinghe dan H. Kalutarage, "Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation," *Soft Comput.*, vol. 25, no. 3, hal.

2277–2293, 2021, doi: 10.1007/s00500-020-05297-6.

[19] S. S A, "Comparative Study of Naive Bayes, Gaussian Naive Bayes Classifier and Decision Tree Algorithms for Prediction of Heart Diseases," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 3, hal. 475–486, 2021, doi: 10.22214/ijraset.2021.33228.

[20] M. Yusran *et al.*, "Comparison of Multinomial Naïve Bayes and Bernoulli Naïve Bayes on Sentiment Analysis of Kurikulum Merdeka with Query Expansion Ranking," vol. 13, hal. 96–106, 2024.

[21] A. H. Odeh, M. Odeh, H. Odeh, dan N. Odeh, "Using Natural Language Processing for Programming Language Code Classification with Multinomial Naive Bayes," *Rev. d'Intelligence Artif.*, vol. 37, no. 5, hal. 1229–1236, 2023, doi: 10.18280/ria.370515.

[22] A. Adadi, *A survey on data-efficient algorithms in big data era*, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00419-9.

[23] S. Alter, "Understanding artificial intelligence in the context of usage: Contributions and smartness of algorithmic capabilities in work systems," *Int. J. Inf. Manage.*, vol. 67, no. November, 2022, doi: 10.1016/j.ijinfomgt.2021.102392.

[24] Y. Sibaroni, "Perbandingan Pembobotan Fitur TF-IDF dan TF-ABS Dalam Klasifikasi Berita Online Menggunakan Support Vector Machine (SVM) 2 nd," *e-Proceeding Eng.*, vol. 10, no. 3, hal. 3652–3655, 2023.

[25] Haniah Mahmudah, Okkie Puspitorini, Nur Adi Siswandari, Ari Wijayanti, dan Eliya Alfatekha, "Metode Naive Bayes Classifier – Smoothing pada Sensor Smartphone untuk Klasifikasi Aktivitas Pengendara," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 9, no. 3, hal. 268–277, 2020, doi: 10.22146/.v9i3.382.

[26] L. Li, D. Song, R. Ma, X. Qiu, dan X. Huang, "KNN-BERT: Fine-Tuning Pre-Trained Models with KNN Classifier," 2021, doi: https://doi.org/10.48550/arXiv.2110.02523.

[27] S. Suhirman dan H. Wintolo, "System for Determining Public Health Level Using the Agglomerative Hierarchical Clustering Method," *Compiler*, vol. 8, no. 1, hal. 95, 2021, doi: 10.28989/compiler.v8i1.425.

[28] I. Made, S. Ramayu, F. Susanto, dan G. S. Mahendra, "Penerapan Data Mining Dengan Algoritma C4.5 Dalam Pemesanan Obat Guna Meningkatkan Keuntungan Apotek," *Online) SENADA*, vol. 5, hal. 237–245, 2022.

[29] A. M. Elkhatat, K. Elsaid, dan S. Almeer, "Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text," *Int. J. Educ. Integr.*, vol. 19, no. 1, hal. 1–16, 2023, doi: 10.1007/s40979-023-00140-5.