

Klasifikasi Berita Hoaks Berbahasa Indonesia Menggunakan TF-IDF dengan Jumlah Fitur Tinggi dan Algoritma *K-Nearest Neighbors* sebagai Model *Baseline*

Indonesian Hoax News Classification Using High-Dimensional TF-IDF Features and the K-Nearest Neighbors Algorithm as a Baseline Model

Moh. Dimas Ubaidillah, Ahmad Yusril Ihsan Sya'bana, Laili Cahyani

Pendidikan Informatika, Universitas Trunojoyo Madura, Bangkalan, Indonesia

230631100007@student.trunojoyo.ac.id, 230631100008@student.trunojoyo.ac.id, laili.cahyani@trunojoyo.ac.id

Notes

Submitted: 11-12-2025

Revised: 05-01-2026

Accepted: 16-01-2026

Published: 30-01-2026

How to cite: M. D. Ubaidillah, A. Y. I. Sya'bana, and L. Cahyani, "Klasifikasi Berita Hoaks Berbahasa Indonesia Menggunakan TF-IDF dengan Jumlah Fitur Tinggi dan Algoritma K-Nearest Neighbors sebagai Model Baseline," *Journal Software, Hardware and Information Technology (SHiFT)*, vol. 6, no. 1, pp. 69–79, Jan. 2026, doi: 10.24252/shift.v6i1.251.

DOI: 10.24252/shift.v6i1.251

Copyright 2026© the Author(s)

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).



ABSTRACT

This study aims to develop an automatic classification model to distinguish fake news and valid news in Indonesian-language texts using the Term Frequency–Inverse Document Frequency (TF-IDF) method and the K-Nearest Neighbors (KNN) algorithm. This research is positioned as a baseline study to evaluate the effectiveness of using TF-IDF representations with a relatively large number of features in the KNN algorithm, as studies that specifically address this aspect in Indonesian fake news classification remain limited. The dataset consists of 500 news documents that undergo preprocessing stages, including text cleaning, character normalization, punctuation removal, and stopword elimination. Numerical document representations are generated using TF-IDF with a maximum of 5,000 features. The dataset is then divided into 400 training samples and 100 testing samples before training the KNN model using standard parameters. The evaluation results show that the proposed model achieves an accuracy of 78%, with balanced precision, recall, and f1-score across both classes. The confusion matrix indicates that most documents are correctly classified, although some misclassifications still occur due to lexical and structural similarities between fake news and valid news texts. In the context of fake news detection, this level of accuracy is considered sufficiently effective as an early detection support system, as the model is able to limit False Negative errors to a relatively low level, allowing the majority of fake news to be successfully identified. These findings demonstrate that the combination of TF-IDF and KNN represents a simple, efficient, and viable baseline approach for developing Indonesian fake news detection systems.

Keywords: fake news, KNN, news classification, text mining, TF-IDF.

ABSTRAK

Penelitian ini bertujuan mengembangkan model klasifikasi otomatis untuk membedakan berita hoaks dan berita valid pada teks berbahasa Indonesia menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan algoritma K-Nearest Neighbors (KNN). Penelitian ini difokuskan sebagai studi baseline untuk mengevaluasi efektivitas penggunaan representasi TF-IDF dengan jumlah fitur yang relatif besar pada algoritma KNN, mengingat masih terbatas penelitian yang membahas aspek tersebut secara spesifik pada klasifikasi berita hoaks berbahasa Indonesia. Dataset yang digunakan terdiri dari 500 dokumen berita yang melalui proses pra-proses meliputi pembersihan teks, normalisasi karakter, penghapusan tanda baca, serta penghilangan stop-words. Representasi numerik dokumen diperoleh melalui TF-IDF dengan jumlah fitur maksimum 5.000. Dataset kemudian dibagi menjadi 400 data latih dan 100 data uji sebelum dilakukan pelatihan model KNN dengan parameter standar. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 78% dengan nilai precision, recall, dan f1-score yang seimbang pada kedua kelas. Confusion matrix memperlihatkan bahwa sebagian besar dokumen berhasil diklasifikasikan dengan benar, meskipun masih terdapat kesalahan klasifikasi yang disebabkan oleh kemiripan kosakata dan struktur kalimat antara berita hoaks dan berita valid. Dalam konteks deteksi berita hoaks, tingkat akurasi tersebut dinilai cukup efektif sebagai sistem pendukung awal (early detection system), karena model mampu menekan kesalahan False Negative pada tingkat yang relatif terbatas, sehingga mayoritas berita hoaks dapat teridentifikasi. Temuan penelitian ini menunjukkan bahwa kombinasi TF-IDF dan KNN merupakan pendekatan yang sederhana, efisien, dan layak digunakan sebagai model baseline dalam pengembangan sistem deteksi berita hoaks berbahasa Indonesia.

Kata kunci: deteksi berita hoaks, klasifikasi berita, KNN, penambahan teks, TF-IDF.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang semakin pesat telah mengubah cara masyarakat dalam memperoleh dan menyebarkan informasi, khususnya melalui media digital dan media sosial. Informasi dapat diakses dan disebarkan dengan sangat cepat tanpa adanya proses verifikasi yang jelas. Kondisi tersebut berdampak pada meningkatnya penyebaran berita hoaks di tengah masyarakat [1]. Berita hoaks sering kali mengandung informasi yang tidak dapat dipertanggungjawabkan kebenarannya dan berpotensi menimbulkan kesalahpahaman, keresahan publik, hingga konflik sosial. Di Indonesia, penyebaran hoaks banyak terjadi pada isu

politik, kesehatan, bencana, dan sosial, sehingga diperlukan upaya yang serius untuk mengatasi permasalahan tersebut agar masyarakat tidak mudah terpengaruh oleh informasi palsu.

Hal ini mendorong perlunya suatu sistem yang mampu melakukan klasifikasi berita hoaks dan berita valid secara otomatis. Sistem klasifikasi merupakan metode yang digunakan untuk memprediksi kategori suatu data berdasarkan karakteristik tertentu [2]. Dalam penelitian ini, penulis membangun suatu sistem klasifikasi berita hoaks menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan algoritma K-Nearest Neighbors (KNN) untuk membantu mengidentifikasi berita hoaks secara tepat. KNN digunakan karena mampu mengklasifikasikan data berdasarkan kedekatan jarak antar data dalam ruang fitur, sehingga data yang memiliki kemiripan tinggi akan dikelompokkan ke dalam kelas yang sama [3]. Sementara itu, TF-IDF digunakan karena merupakan salah satu metode *text-processing* yang sederhana, relevan, dan efektif dalam memberikan bobot pada kata-kata penting di dalam dokumen berita [4], [5].

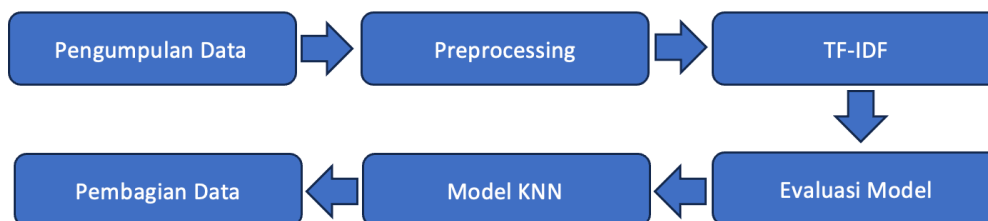
Sebelumnya telah dikembangkan beberapa penelitian serupa. Sani dkk. membandingkan algoritma *Naive Bayes Classifier* dan *Support Vector Machine* (SVM) dengan ekstraksi fitur TF-IDF pada berita hoaks berbahasa Indonesia dan memperoleh akurasi terbaik sebesar 93,1% untuk *Complement Naive Bayes* serta 96,5% untuk SVM dengan kernel *sigmoid* [6]. Hendriyanto dan Sari menerapkan algoritma KNN dengan pembobotan TF-IDF untuk mengklasifikasikan judul berita hoaks dan memperoleh akurasi tertinggi sebesar 93,33% pada nilai $k = 3$ [7]. Sebaliknya, Wulandari dan Kustiyono melaporkan bahwa *Naive Bayes* berbasis TF-IDF mencapai akurasi 95,74%, jauh di atas KNN yang hanya mencapai 48,97% pada klasifikasi berita hoaks program Makan Bergizi Gratis [8]. Penelitian lain membandingkan KNN dan *Naive Bayes* dengan optimasi *Randomized Search* untuk klasifikasi berita hoaks [9], serta menerapkan KNN pada klasifikasi kategori berita daring dengan akurasi sebesar 91% [10]. Dari beberapa referensi tersebut diperoleh bahwa metode klasifikasi berbasis *text mining* dan *machine learning* cukup efektif dan relevan dalam mendeteksi berita hoaks [1]. Namun, hasil yang dilaporkan untuk algoritma KNN sangat bervariasi, yang menunjukkan bahwa performanya sensitif terhadap pemilihan fitur dan kemiripan struktur bahasa antar berita yang dapat menyebabkan terjadinya kesalahan klasifikasi. Selain itu, masih terbatas penelitian yang secara khusus mengkaji pengaruh penggunaan jumlah fitur TF-IDF yang besar terhadap performa algoritma KNN pada klasifikasi berita hoaks berbahasa Indonesia. Oleh karena itu, diharapkan dengan adanya penelitian ini dapat dihasilkan model klasifikasi berita hoaks yang lebih akurat dan efektif, sehingga masyarakat dapat lebih mudah membedakan berita yang benar dan berita palsu serta terhindar dari dampak negatif penyebaran hoaks di masa depan.

Berdasarkan kajian terhadap penelitian terdahulu, sebagian besar studi klasifikasi berita hoaks berbahasa Indonesia masih berfokus pada perbandingan algoritma klasifikasi tanpa mengkaji secara spesifik pengaruh jumlah fitur TF-IDF yang digunakan terhadap performa model, khususnya pada algoritma K-Nearest Neighbors (KNN). Beberapa penelitian menggunakan jumlah fitur yang relatif terbatas tanpa memberikan justifikasi eksplisit terkait pemilihan dimensi fitur.

Selain itu, masih terbatas penelitian yang menempatkan kombinasi TF-IDF dan KNN sebagai model baseline yang dievaluasi secara mendalam melalui analisis kesalahan klasifikasi, terutama dalam konteks False Negative yang berisiko tinggi pada deteksi hoaks. Oleh karena itu, penelitian ini berfokus pada evaluasi efektivitas penggunaan jumlah fitur TF-IDF yang relatif besar (5.000 fitur) terhadap performa algoritma KNN dalam klasifikasi berita hoaks berbahasa Indonesia, serta menganalisis pola kesalahan klasifikasi yang dihasilkan.

2. METODE PENELITIAN

Dalam penelitian ini digunakan metode *text mining* untuk melakukan pengolahan konten berita yang akan dijadikan sebagai dasar pembobotan dalam proses klasifikasi berita hoaks dan berita valid [2]. Pada tahap implementasi, metode Term Frequency–Inverse Document Frequency (TF-IDF) digunakan untuk melakukan pembobotan nilai kata pada setiap dokumen berita, sedangkan algoritma K-Nearest Neighbors (KNN) digunakan untuk menentukan kelas berita berdasarkan tingkat kemiripan antar dokumen dalam ruang fitur.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Tahapan pertama yang dilakukan dalam penelitian ini adalah pengumpulan data dengan menggunakan metode dokumentasi sebagai sumber data utama. Dataset yang digunakan merupakan data sekunder berupa kumpulan berita berbahasa Indonesia yang telah diberi label hoaks dan valid. Data diperoleh dari hasil kompilasi berbagai sumber berita daring serta dataset terbuka yang telah melalui proses pelabelan sebelumnya, sehingga data yang digunakan telah memiliki kelas yang jelas dan siap untuk dianalisis. Dataset utama yang digunakan dalam penelitian ini berjumlah 500 data berita yang terdiri dari dua kelas, yaitu berita hoaks dan berita valid. Meskipun jumlah dataset yang digunakan relatif terbatas, data tersebut dipandang memadai untuk mengevaluasi performa awal dan kestabilan model klasifikasi sebagai studi baseline. Seluruh data disimpan dalam file berformat CSV, kemudian dikonversi dan diolah menggunakan library Pandas pada bahasa pemrograman Python [11] sebelum dilakukan tahap praproses dan analisis lebih lanjut.

2.2. Tahapan Preprocessing

Tahapan *preprocessing* merupakan tahap awal pengolahan data sebelum dilakukan proses ekstraksi fitur dan klasifikasi. Data berita yang telah dikumpulkan akan diolah dan diproses terlebih dahulu agar dapat digunakan dalam sistem klasifikasi berita hoaks. Dataset yang semula berupa teks mentah akan diubah ke dalam bentuk yang lebih terstruktur dan bersih sehingga dapat direpresentasikan dalam bentuk vektor menggunakan metode TF-IDF. Proses *preprocessing* ini dilakukan dengan memanfaatkan library NLTK [12]. Adapun tahapan *preprocessing* yang dilakukan adalah sebagai berikut:

Seleksi Fitur

Seleksi fitur dilakukan untuk memastikan bahwa data teks yang digunakan benar-benar mengandung informasi yang relevan dengan proses klasifikasi berita hoaks. Pada tahap ini, hanya kolom berita yang digunakan sebagai fitur utama, sedangkan kolom label digunakan sebagai target klasifikasi. Seleksi fitur ini bertujuan untuk meningkatkan kualitas data serta mengurangi atribut yang tidak diperlukan dalam proses analisis.

Cleaning Text

Tahap *cleaning text* dilakukan untuk membersihkan teks dari karakter yang tidak diperlukan. Proses ini meliputi *case folding* (mengubah seluruh huruf menjadi huruf kecil), penghapusan angka, URL, simbol, serta tanda baca menggunakan *library* NLTK. Setelah dilakukan proses *cleaning*, akan dihasilkan kumpulan kata yang lebih bersih dan seragam yang merepresentasikan isi berita, sehingga dapat digunakan pada tahap pelatihan (*training*) model klasifikasi.

Stopword Removal

Tahap *stopword removal* bertujuan untuk menghilangkan kata-kata umum dalam Bahasa Indonesia yang tidak memiliki pengaruh besar terhadap proses klasifikasi, seperti “dan”, “yang”, “di”, dan “ke”. Penghapusan *stopword* dilakukan menggunakan daftar *stopword* Bahasa Indonesia dari *library* NLTK. Proses ini bertujuan untuk mengurangi dimensi data serta meningkatkan relevansi kata yang digunakan sebagai fitur dalam proses ekstraksi TF-IDF [13].

2.3. Ekstraksi Fitur Menggunakan TF-IDF

Metode Term Frequency–Inverse Document Frequency (TF-IDF) digunakan dalam penelitian ini untuk mengubah data teks hasil preprocessing menjadi representasi vektor numerik. Representasi numerik tersebut digunakan sebagai fitur masukan pada algoritma klasifikasi K-Nearest Neighbors (KNN). TF-IDF memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya di dalam suatu dokumen dan di seluruh kumpulan dokumen [4], [14].

Secara umum, pembobotan TF-IDF terdiri dari dua komponen utama, yaitu Term Frequency (TF) dan Inverse Document Frequency (IDF). Nilai TF menggambarkan seberapa sering sebuah kata muncul dalam sebuah dokumen [2] dan dirumuskan sebagai berikut:

$$TF = \frac{\text{jumlah kemunculan kata dalam dokumen}}{\text{jumlah seluruh kata dalam dokumen}} \quad (1)$$

Sementara itu, komponen IDF digunakan untuk memberikan penalti terhadap kata-kata yang terlalu sering muncul dalam banyak dokumen [4]. Nilai IDF dihitung menggunakan rumus:

$$IDF = \log \frac{N}{n} \quad (2)$$

Bobot akhir setiap kata dalam dokumen dihitung menggunakan gabungan komponen TF dan IDF sebagaimana ditunjukkan dalam rumus:

$$w_{ij} = tf_{ij} \times \log \left(\frac{N + 1}{df_i} + 1 \right) \quad (3)$$

Rumus ini memungkinkan TF-IDF memberikan bobot tinggi pada kata-kata yang muncul sering dalam satu dokumen, tetapi jarang muncul di dokumen lain, sehingga lebih informatif dalam proses klasifikasi [5].

Pada penelitian ini, proses pembobotan dilakukan menggunakan *library* TfidfVectorizer dari Scikit-Learn [15]. Untuk menjaga kualitas fitur sekaligus mengurangi dimensi data, jumlah fitur dibatasi menggunakan parameter `max_features = 5000`. Pembatasan ini memastikan bahwa hanya 5.000 kata yang paling representatif dan paling sering muncul dalam dataset yang digunakan sebagai fitur.

Hasil pembobotan TF-IDF berupa sebuah matriks vektor berdimensi tinggi, di mana setiap baris merepresentasikan sebuah dokumen berita, dan setiap kolom merepresentasikan bobot TF-IDF dari suatu kata. Matriks ini kemudian disimpan dalam file Excel (tfidf_output.xlsx) untuk keperluan dokumentasi, analisis lanjutan, serta digunakan pada tahap pembagian data latih dan data uji sebelum proses pelatihan model KNN.

2.4. Pembagian Data

Dataset yang digunakan dalam penelitian ini berjumlah 500 data berita yang terdiri dari dua kelas, yaitu berita hoaks dan berita valid. Data tersebut kemudian dibagi menjadi dua bagian, yaitu data latih (*training data*) sebanyak 400 data dan data uji (*testing data*) sebanyak 100 data. Proses pembagian data dilakukan menggunakan fungsi `train_test_split` dengan teknik stratified sampling (*stratify*) [15], sehingga proporsi jumlah data pada masing-masing kelas tetap seimbang baik pada data latih maupun data uji. Pembagian data ini bertujuan untuk memastikan bahwa model dilatih dan diuji secara adil serta mampu menghasilkan performa yang representatif.

2.5. Pelatihan Model KNN

Pada tahap ini dilakukan pelatihan model menggunakan algoritma K-Nearest Neighbors (KNN) yang diimplementasikan melalui pustaka Scikit-Learn [15]. KNN merupakan metode klasifikasi berbasis kedekatan jarak antar data dalam ruang fitur [3]. Setiap dokumen hasil ekstraksi TF-IDF direpresentasikan sebagai vektor berdimensi tinggi, sehingga proses klasifikasi dilakukan dengan membandingkan jarak antara dokumen yang akan diprediksi dengan seluruh dokumen lain dalam dataset.

Perhitungan kedekatan antar dokumen dilakukan menggunakan jarak Euclidean [16], yang dirumuskan sebagai berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

Di mana x merupakan vektor fitur dokumen yang akan diklasifikasikan, y adalah vektor fitur dokumen pembanding, dan n adalah jumlah fitur TF-IDF yang digunakan (sebanyak 5.000 fitur).

Setelah seluruh jarak dihitung, algoritma memilih k tetangga terdekat dengan jarak paling kecil:

$$Neighbors = \arg \min_k (d(x, y)) \quad (5)$$

Pada penelitian ini, nilai k yang digunakan adalah $k = 5$, yang merupakan parameter standar dan memberikan performa stabil pada dataset berukuran sedang. Dengan pendekatan ini, dokumen berita yang memiliki tingkat kemiripan tertinggi dalam ruang fitur TF-IDF akan dikelompokkan ke dalam kelas yang sama, sehingga algoritma KNN mampu melakukan klasifikasi berdasarkan pola kemiripan antar dokumen.

Meskipun nilai $k = 5$ digunakan sebagai parameter utama, penelitian ini juga melakukan pengujian awal (*preliminary testing*) dengan beberapa variasi nilai k yang umum digunakan, yaitu $k = 3$, $k = 5$, dan $k = 7$, untuk mengamati kestabilan performa model. Hasil pengujian awal

menunjukkan bahwa nilai $k = 5$ memberikan keseimbangan terbaik antara nilai *precision*, *recall*, dan *f1-score* pada kedua kelas.

Pengujian tersebut dilakukan secara terbatas dan tidak menggunakan metode *hyperparameter tuning* yang komprehensif seperti *grid search* atau *cross-validation*, karena fokus utama penelitian ini adalah mengevaluasi efektivitas kombinasi metode TF-IDF dan KNN dalam klasifikasi berita hoaks sebagai model baseline. Oleh karena itu, nilai $k = 5$ dipilih sebagai parameter yang paling representatif untuk dataset yang digunakan.

2.6. Evaluasi Model

Evaluasi terhadap performa model klasifikasi dilakukan untuk mengetahui tingkat keberhasilan sistem dalam mengklasifikasikan berita hoaks dan berita valid. Metode evaluasi yang digunakan dalam penelitian ini meliputi *confusion matrix*, nilai *precision*, *recall*, *f1-score*, serta akurasi keseluruhan (accuracy) [17]. *Confusion matrix* digunakan untuk melihat jumlah prediksi yang benar dan salah pada masing-masing kelas, sedangkan *precision*, *recall*, dan *f1-score* digunakan untuk mengukur kualitas hasil klasifikasi secara lebih detail. Nilai akurasi digunakan untuk mengetahui persentase ketepatan model secara keseluruhan dalam mengklasifikasikan dua kelas secara seimbang.

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini terdiri dari 500 data berita berbahasa Indonesia yang terbagi ke dalam dua kelas, yaitu berita hoaks dan berita valid. Seluruh data berhasil dikompilasi tanpa duplikasi maupun data kosong, sehingga layak digunakan dalam tahap pre-processing dan klasifikasi.

3.2. Preprocessing

Tahapan preprocessing dilakukan untuk meningkatkan kualitas data teks sebelum proses ekstraksi fitur dan klasifikasi. Hasil preprocessing menunjukkan bahwa teks berita menjadi lebih bersih dan terstruktur, dengan pengurangan noise berupa angka, simbol, tanda baca, serta kata-kata umum yang kurang informatif. Proses ini menghasilkan representasi teks yang lebih ringkas dan relevan, sehingga mendukung efektivitas pembobotan TF-IDF dalam merepresentasikan karakteristik berita hoaks dan berita valid.

3.3. Pembobotan TF-IDF

Setelah melalui tahap preprocessing, data teks diproses menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk menghasilkan representasi numerik dokumen. Pembobotan TF-IDF menghasilkan vektor fitur yang merepresentasikan tingkat kepentingan kata dalam masing-masing dokumen berita. Jumlah fitur dibatasi hingga 5.000 kata yang paling representatif untuk menjaga keseimbangan antara kualitas fitur dan kompleksitas model. Hasil pembobotan ini digunakan sebagai dasar dalam proses pembagian data serta pelatihan model klasifikasi.

3.4. Hasil Pembagian Data dan Pelatihan Model KNN

Dataset kemudian dibagi menjadi dua bagian, yaitu 400 data latih dan 100 data uji, dengan menggunakan teknik *stratified split* untuk memastikan distribusi kelas hoaks dan valid tetap seimbang pada kedua subset. Pembagian ini bertujuan agar model dilatih dan diuji secara adil serta menghasilkan performa yang representatif.

Pelatihan model dilakukan menggunakan algoritma K-Nearest Neighbors (KNN) dengan nilai parameter $k = 5$. Pemilihan nilai k ini memberikan kestabilan performa pada dataset berukuran menengah dan mampu mengelompokkan dokumen berdasarkan tingkat kemiripan dalam ruang fitur TF-IDF. Proses pelatihan berjalan dengan baik tanpa kendala, yang menunjukkan bahwa representasi fitur hasil ekstraksi TF-IDF kompatibel dengan algoritma KNN dan siap digunakan pada tahap evaluasi performa model.

3.5. Hasil Evaluasi Model

Berdasarkan hasil pengujian yang dilakukan terhadap 100 data uji, model K-Nearest Neighbors (KNN) dengan nilai $k = 5$ menunjukkan performa klasifikasi yang cukup stabil. Proses evaluasi dilakukan menggunakan confusion matrix dan classification report untuk mengetahui seberapa baik model dapat membedakan berita hoaks dan berita valid berdasarkan representasi fitur TF-IDF. Hasil evaluasi ini menjadi dasar untuk menilai akurasi sistem serta kualitas prediksi yang dihasilkan oleh model.

Confusion Matrix

Hasil evaluasi kinerja model K-Nearest Neighbors (KNN) terhadap data uji disajikan menggunakan confusion matrix, yang bertujuan untuk menunjukkan perbandingan antara kelas aktual dan kelas hasil prediksi model. Confusion matrix hasil klasifikasi ditunjukkan pada Tabel 1.

Tabel 1. Confusion Matrix Hasil Klasifikasi Berita Hoaks dan Valid

Kelas Aktual	Prediksi Hoaks	Prediksi Valid
Hoaks	40	10
Valid	12	38

Interpretasi:

- 40 berita hoaks diklasifikasikan benar sebagai hoaks
- 10 berita hoaks diklasifikasikan salah sebagai valid
- 38 berita valid diklasifikasikan benar sebagai valid
- 12 berita valid salah diklasifikasikan sebagai hoaks

Dari data ini terlihat bahwa model sedikit lebih baik dalam mendeteksi hoaks dibandingkan mendeteksi berita valid.

Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang relatif baik dalam mendeteksi berita hoaks dibandingkan mendeteksi berita valid. Kesalahan klasifikasi berupa berita hoaks yang diprediksi sebagai berita valid (*false negative*) mengindikasikan bahwa beberapa berita hoaks memiliki karakteristik linguistik yang menyerupai berita valid, seperti penggunaan kosakata formal dan struktur kalimat informatif. Kondisi ini menyebabkan nilai bobot TF-IDF antar dokumen menjadi saling berdekatan, sehingga algoritma KNN cenderung mengelompokkan dokumen ke kelas tetangga terdekat yang dominan.

Kesalahan klasifikasi berupa berita hoaks yang diprediksi sebagai berita valid (*false negative*) menunjukkan bahwa sebagian berita hoaks memiliki karakteristik linguistik yang menyerupai berita valid. Beberapa berita hoaks cenderung menggunakan kosakata formal dan struktur kalimat yang bersifat informatif, sehingga bobot TF-IDF yang dihasilkan mendekati pola berita valid. Selain itu, berita dengan panjang teks yang relatif pendek memiliki keterbatasan variasi kata, sehingga informasi yang tersedia tidak cukup kuat untuk membedakan kelas secara

akurat. Kondisi tersebut menyebabkan jarak antar dokumen dalam ruang fitur menjadi sangat dekat, sehingga algoritma KNN mengelompokkan dokumen ke kelas tetangga terdekat yang dominan.

Classification Report

Untuk mengevaluasi performa model secara lebih rinci, digunakan metrik precision, recall, f1-score, dan support. Hasil evaluasi tersebut disajikan dalam bentuk classification report yang ditampilkan pada Tabel 2.

Tabel 2. Hasil Classification Report Model KNN

Kelas	Precision	Recall	F1-Score	Support
Hoaks	0.77	0.80	0.78	50
Valid	0.79	0.76	0.78	50
Akurasi			0.78	100

Berdasarkan Tabel 2, model klasifikasi berita menggunakan metode TF-IDF dan algoritma K-Nearest Neighbors (KNN) memperoleh akurasi keseluruhan sebesar 78%. Nilai f1-score yang seimbang pada kedua kelas (0.78) menunjukkan bahwa model tidak mengalami bias terhadap salah satu kelas serta mampu menjaga keseimbangan antara precision dan recall dalam proses klasifikasi [17]. Akurasi tersebut berada di antara hasil KNN pada penelitian terdahulu, yaitu 48,97% pada klasifikasi berita hoaks program Makan Bergizi Gratis [8] dan 93,33% pada klasifikasi judul berita hoaks [7]. Perbedaan ini dipengaruhi oleh karakteristik dataset, panjang teks yang diklasifikasikan, serta pengaturan parameter k dan jumlah fitur yang digunakan.

Dalam konteks deteksi berita hoaks, tingkat performa tersebut dapat dikategorikan cukup efektif sebagai sistem pendukung awal (*early detection system*), karena model mampu mengidentifikasi sebagian besar berita hoaks dengan tingkat kesalahan *false negative* yang relatif terbatas. Dengan demikian, kombinasi metode TF-IDF dan algoritma KNN dapat dijadikan sebagai model baseline yang layak untuk pengembangan sistem deteksi berita hoaks berbahasa Indonesia pada penelitian selanjutnya.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, model klasifikasi berita hoaks menggunakan metode TF-IDF dan algoritma K-Nearest Neighbors (KNN) menunjukkan performa yang cukup efektif sebagai model baseline, meskipun jumlah dataset yang digunakan relatif terbatas. Representasi teks berbasis TF-IDF dengan 5.000 fitur mampu memberikan informasi yang relevan bagi algoritma KNN dalam membedakan berita hoaks dan berita valid.

Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 78% dengan nilai precision, recall, dan f1-score yang seimbang pada kedua kelas. Tingkat performa tersebut menunjukkan bahwa sistem mampu melakukan klasifikasi secara cukup tepat dan konsisten terhadap pola data yang ada, serta efektif sebagai sistem pendukung awal dalam deteksi berita hoaks. Meskipun demikian, performa model masih memiliki peluang untuk ditingkatkan dengan memperbesar jumlah dataset, memperluas variasi berita, serta mengintegrasikan metode representasi teks yang lebih kaya secara semantik, seperti Word2Vec, FastText, atau BERT. Dengan demikian, penelitian lanjutan diharapkan dapat menghasilkan model pendeteksi berita hoaks yang lebih akurat dan lebih adaptif terhadap dinamika bahasa dalam berita digital.

PERNYATAAN PENGGUNAAN ARTIFICIAL INTELLIGENCE

DeepSeek dan Gemini dimanfaatkan sebagai alat bantu untuk mendukung penyelesaian aspek teknis penelitian dan penyempurnaan penulisan artikel. Saran terkait pemrograman, pengorganisasian data, dan penyajian informasi diperiksa kembali sebelum digunakan. Penulis memastikan bahwa seluruh hasil yang dilaporkan sesuai dengan data penelitian dan bertanggung jawab penuh atas isi publikasi.

KONTRIBUSI PENULIS

Moh. Dimas Ubaidillah: Analisis Data, Penulisan Artikel.

Ahmad Yusril Ihsan Sya'bana: Kurasi Data, Penulisan Artikel.

Laili Cahyani: Supervisi, Review, Keputusan Akhir Artikel.

PENDANAAN

Penelitian ini tidak menerima hibah khusus dari lembaga pendanaan mana pun, baik sektor publik, komersial, maupun nirlaba.

KONFLIK KEPENTINGAN

Penulis menyatakan tidak terdapat konflik kepentingan terkait publikasi artikel ini.

DAFTAR PUSTAKA

- [1] M. D. Desriansyah, I. U. Sari, and Z. Zulfahmi, "Analisis Efektivitas Algoritma Machine Learning dalam Deteksi Hoaks: Pada Berita Digital Berbahasa Indonesia," *JISKA*, vol. 3, no. 2, pp. 63–69, Jun. 2025, doi: 10.47233/jiska.v3i1.2024.
- [2] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008.
- [3] P. K. Syriopoulos, N. G. Kalampalikis, S. B. Kotsiantis, and M. N. Vrahatis, "kNN classification: A review," *Ann. Math. Artif. Intell.*, vol. 93, no. 1, pp. 43–75, Feb. 2025, doi: 10.1007/s10472-023-09882-x.
- [4] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Inf. Process. Manage.*, vol. 24, no. 5, pp. 513–523, 1988, doi: 10.1016/0306-4573(88)90021-0.
- [5] M. Bounabi, K. El Moutaouakil, and K. Satori, "Text classification using fuzzy TF-IDF and machine learning models," in *Proc. 4th Int. Conf. Big Data and Internet of Things (BDIoT'19)*, Rabat, Morocco, Oct. 2019, pp. 1–6, doi: 10.1145/3372938.3372956.
- [6] R. R. Sani, Y. A. Pratiwi, S. Winarno, E. D. Udayanti, and F. Alzami, "Analisis Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Berita Hoax pada Berita Online Indonesia," *Jurnal Masyarakat Informatika (JMASIF)*, vol. 13, no. 2, pp. 85–98, Nov. 2022, doi: 10.14710/jmasif.13.2.47983.
- [7] M. D. Hendriyanto and B. N. Sari, "Penerapan Algoritma K-Nearest Neighbor dalam Klasifikasi Judul Berita Hoax," *Jurnal Ilmiah Informatika (JIF)*, vol. 10, no. 02, pp. 80–84, Sep. 2022, doi: 10.33884/jif.v10i02.5477.
- [8] D. A. Nur'Faradila, L. Magdalena, and M. Febima, 'Penerapan Naïve Bayes Classifier Untuk Klasifikasi Postingan Berita Hoaks Di Instagram Cirebon Saber Hoaks', *MJI*, vol. 16, no. 2, p. 243, Dec. 2024, doi: 10.35194/mji.v16i2.4616.
- [9] B. Christopher, F. Arisandy, and M. I. Fahmi, "Analisis Komparasi Metode K-Nearest Neighbor dan Naïve Bayes Classifier Berbasis Optimasi Randomized Search dalam Klasifikasi Berita Hoaks," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 12, no. 2, pp. 272–283, Jun. 2025, doi: 10.35957/jatisi.v12i2.11371.

- [10] A. T. Saputra, Y. A. Mustofa, Y. Lasena, and I. S. K. Idris, "Klasifikasi Berita Online pada Liputan 6 Menggunakan Metode KNN," *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, vol. 4, no. 1, pp. 58–66, 2025.
- [11] W. McKinney, "Data structures for statistical computing in Python," in *Proc. 9th Python in Science Conf. (SciPy 2010)*, Austin, TX, USA, 2010, pp. 56–61, doi: 10.25080/Majora-92bf1922-00a.
- [12] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol, CA, USA: O'Reilly Media, 2009.
- [13] S. Sarica and J. Luo, "Stopwords in technical language processing," *PLoS ONE*, vol. 16, no. 8, p. e0254937, Aug. 2021, doi: 10.1371/journal.pone.0254937.
- [14] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 1, pp. 84–93, 2023, doi: 10.34010/jamika.v13i1.9424.
- [15] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [16] D. Setiawan, A. Muhammad, and A. Firizkiansah, "Pengklasifikasian Dokumen Teks Bahasa Indonesia berbasis Vektor Space Model dengan Menggunakan Metode k-Nearest Neighbor (k-NN) dan Euclidean Distance," *JIKOMTI: Jurnal Ilmiah Ilmu Komputer dan Teknologi Informatika*, vol. 1, no. 1, pp. 30–37, Dec. 2024.
- [17] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, 2009, doi: 10.1016/j.ipm.2009.03.002.