
PERBANDINGAN ALGORITMA NAÏVE BAYES DAN RANDOM FOREST UNTUK PREDIKSI BERITA HOAX PADA SOSIAL MEDIA

Andifa Nur Wicaksono
Mahasiswa STMIK ASIA Malang
e-mail: andifawicaksono@gmail.com

ABSTRAK

Penyebaran berita hoax di media sosial telah menjadi ancaman serius yang dapat memicu konflik sosial dan ketidakstabilan. Penelitian ini bertujuan untuk membandingkan performa dua algoritma machine learning, yaitu Naïve Bayes dan Random Forest, dalam mengklasifikasikan berita hoax. Metodologi penelitian mengikuti tahapan Knowledge Discovery in Database (KDD) yang meliputi pengumpulan data, preprocessing (cleaning, case folding, stopword removal, tokenizing, dan stemming), serta ekstraksi fitur menggunakan TF-IDF. Dataset yang digunakan berjumlah 8.400 data yang terdiri dari berita hoax dan non-hoax. Hasil penelitian menunjukkan bahwa algoritma Random Forest memberikan tingkat akurasi yang lebih tinggi dibandingkan Naïve Bayes. Implementasi ini diharapkan dapat menjadi referensi dalam pengembangan sistem deteksi dini berita palsu di platform digital.

Kata kunci: Berita Hoax, Sosial Media, Naive Bayes, Random Forest, TF-IDF

ABSTRACT

The spread of hoax news on social media has become a serious threat that can trigger social conflict and instability. This study aims to compare the performance of two machine learning algorithms, Naïve Bayes and Random Forest, in classifying hoax news. The research methodology follows the stages of Knowledge Discovery in Database (KDD) which includes data collection, preprocessing (cleaning, case folding, stopword removal, tokenizing, and stemming), and feature extraction using TF-IDF. The dataset used consists of 8,400 data points of hoax and non-hoax news. The results show that the Random Forest algorithm provides a higher level of accuracy compared to Naïve Bayes. This implementation is expected to serve as a reference for developing early detection systems for fake news on digital platforms

Keywords: Hoax News, Social Media, Naive Bayes, Random Forest, TF-IDF

PENDAHULUAN

Media sosial telah mengubah cara masyarakat berinteraksi dan mengonsumsi informasi. Namun, sisi negatif dari kemudahan ini adalah masifnya penyebaran berita bohong atau hoax. Berdasarkan dokumen tugas akhir, masalah utama yang diangkat adalah bagaimana teknologi *Machine Learning* dapat membantu membedakan berita asli dan hoax secara otomatis. Penelitian ini difokuskan pada perbandingan algoritma Naïve Bayes yang berbasis probabilitas dengan Random Forest yang berbasis *ensemble tree*.

KAJIAN TEORI

1. Machine Learning

Machine learning (ML) merupakan cabang dari kecerdasan buatan yang berfokus pada pengembangan algoritma dan model yang memungkinkan komputer untuk belajar dari data dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit. ML telah menjadi pendorong utama inovasi di berbagai bidang, mulai dari kesehatan, keuangan, transportasi, hingga industri kreatif (Puppala, 2025).

2. Knowledge Discovery In Database (KDD)

Knowledge Discovery in Databases (KDD) adalah proses sistematis untuk menemukan pengetahuan yang valid, baru, dan dapat dipahami dari data dalam jumlah besar. KDD terdiri dari beberapa tahapan utama, yaitu seleksi data, pra-pemrosesan, transformasi, data mining, dan interpretasi hasil (Llatas dkk., 2024)

3. Naive Bayes

Naïve Bayes banyak digunakan karena memiliki struktur sederhana, komputasi yang cepat dan efisien pada data berdimensi tinggi, serta tetap memberikan kinerja yang baik meskipun asumsi independensi atribut tidak sepenuhnya terpenuhi (Pajila dkk., 2023))

4. Random Forest

Random forest adalah metode pembelajaran *ensemble* untuk klasifikasi atau *regresi* yang beroperasi menggunakan menciptakan banyak pohon keputusan selama proses training dan menaruh hasil berupa mode kelas (klasifikasi) atau prediksi rata-rata (regresi) pohon individu (Ramadhan & Adhinata, 2021)

5. Term Frequency - Inverse Document Frequency (TF-IDF)

TF-IDF adalah teknik pembobotan kata yang mengukur seberapa penting suatu istilah dalam suatu dokumen relatif terhadap seluruh kumpulan dokumen. Metode ini merupakan pengembangan dari model *bag-of-words* yang hanya menghitung frekuensi kata, sehingga TF-IDF memberikan bobot yang lebih informatif dengan memperhitungkan tingkat unik kata dalam keseluruhan dokumen. Secara umum, TF-IDF bekerja dengan mengalikan dua komponen utama yaitu *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) (Ionendri dkk., 2025).

6. FlowChart

Flowchart adalah alat visualisasi yang digunakan untuk menggambarkan alur proses analisis data secara sistematis, mulai dari pengumpulan data, *preprocessing*, pemilihan fitur, pemodelan, evaluasi, hingga interpretasi hasil. *Flowchart* membantu dalam memahami dan mengkomunikasikan tahapan analisis data kepada berbagai pemangku kepentingan, baik teknis maupun non-teknis (Ghritlahare, 2025).

7. Berita Hoax

Hoax didefinisikan sebagai informasi palsu atau menyesatkan yang sengaja dibuat agar tampak benar untuk mempengaruhi opini atau tindakan masyarakat, memiliki ciri sumber tidak jelas dan bahasa ambigu, dapat bermotif politis atau komersial, serta menimbulkan dampak negatif, dan dalam konteks media sosial menyebar cepat sehingga memerlukan verifikasi dan literasi media yang baik untuk mengendalikannya (Adam dkk., 2025).

PEMBAHASAN

Dalam penelitian ini digunakan dua algoritma utama, yaitu *Naïve Bayes* dan *Random Forest*. Kedua algoritma ini dipilih karena memiliki karakteristik, pendekatan analitis, serta keunggulan masing-masing dalam menangani permasalahan prediksi berbasis teks. Khususnya dalam konteks klasifikasi teks berita *hoax* pada sosial media yang memiliki sifat dinamis, tidak terstruktur, dan mengandung beragam variasi bahasa pengguna.

Dataset yang digunakan dalam penelitian ini bersumber dari dua dataset berbeda, yaitu dataset berita *hoax* yang diperoleh dari Kaggle oleh Muhammad Ghazi Muharam dan dataset berita non-*hoax* yang disusun oleh Iqbal Maulana. Kedua dataset tersebut tersedia dalam format *Comma Separated Values* (CSV).

Tahap perancangan permodelan klasifikasi bertujuan untuk membangun model yang mampu mengklasifikasikan berita ke dalam dua kelas, yaitu *hoax* dan *non-hoax*, berdasarkan fitur teks yang telah diekstraksi menggunakan metode TF-IDF. Algoritma klasifikasi yang digunakan pada

penelitian ini adalah *Naïve Bayes*, dengan varian *Multinomial Naïve Bayes*.

Pada penelitian ini, dataset berita *hoax* dan *non-hoax* yang telah melalui tahap ekstraksi fitur menggunakan metode TF-IDF digunakan sebagai masukan pada model *Random Forest*. Setiap nilai TF-IDF dianggap sebagai atribut dalam proses pembentukan pohon keputusan. Proses pelatihan dilakukan dengan membangun sejumlah pohon keputusan berdasarkan data latih yang tersedia.

Hasil prediksi dari kedua model dianalisis menggunakan *Confusion Matrix*, yang menggambarkan hubungan antara hasil prediksi model dan data aktual. *Confusion Matrix* terdiri dari empat komponen utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Komponen-komponen ini digunakan sebagai dasar dalam menghitung metrik evaluasi kinerja model.

Tahap awal dalam pemrosesan data adalah *preprocessing* data *cleaning*, yang diimplementasikan menggunakan pustaka (*library*) *Pandas* sebagai instrumen utama dalam manipulasi data terstruktur. Langkah pertama yang krusial dalam tahapan ini adalah memastikan keseimbangan proporsi antara dataset kategori *hoax* dan kategori *non-hoax*. Dataset yang sudah diolah menjadi Dataset gabungan yang memiliki jumlah 8400 baris dengan 4200 sampel data dilabeli sebagai *hoax* dan 4200 data dilabeli sebagai data *non-hoax*. Langkah berikutnya adalah *Preprocessing Stopword Removal* dilakukan dengan menggunakan dataset *stopword* bahasa Indonesia. *Stopword* ini disimpan pada file *csv* yang kemudian dilakukan pengecekan data teks pada dataset. Jika terdapat kata pada *stopword* dan dataset sama maka kata tersebut akan dihapus dari dataset. Setelah itu dilakukan *preprocessing* tokenisasi dilakukan dengan melakukan pemotongan kalimat pada spasi. Jadi kalimat akan di bagi berdasarkan spasi sebagai pembeda pada setiap kalimat, Kemudian *preprocessing Stemming* Kolom data yang dilakukan proses *stemming* adalah kolom *judul_clean* dan *narasi_clean*. Kolom *judul_clean* dan *narasi_clean* ini adalah kolom yang digunakan untuk menampung hasil dari proses *cleaning*, *Case Folding* dan *Stopword Removal*.

Tahap berikutnya adalah membagi dataset yang sudah dilakukan *Preprocessing Stemming* data menjadi dua bagian yaitu data latih dan data uji. Penelitian ini menggunakan perbandingan 70 : 30 yang berarti 70% data digunakan untuk data latih dan 20% data digunakan untuk data uji. Kemudian dilakukan proses TF-IDF yang menghasilkan matriks numerik yang kemudian digunakan oleh algoritma *Naïve Bayes* dan *Random Forest*.

Tahap Berikut nya Adalah Permodelan klasifikasi *Naïve Bayes* menggunakan library Scikit-Learn, dengan fokus pada model *Multinomial Naïve Bayes*. Kode dan parameter yang digunakan untuk melakukan permodelan *Naïve Bayes* ditunjukkan pada kode berikut

```
# =====
# Model 1: Naïve Bayes
# =====
nb_model = MultinomialNB(
    alpha = 5.0, # Tingkatkan alpha
    dari 0.1 ke 1.0 atau 2.0
    fit_prior = True
)
nb_model.fit(X_train_tfidf, y_train)

y_pred_nb = nb_model.predict(X_test_tfidf)
y_pred_nb_train =
nb_model.predict(X_train_tfidf)
```

Setelah dilakukan pelatihan menggunakan model *Naïve Bayes* didapatkan skor akurasi training sebesar 91% dan akurasi tes sebesar 88%. Kemudian didapatkan skor *precision* 95%, *recall* 81%, *f1 score* 87% untuk label *hoax* dan skor *precision* 83%, *recall* 96%, *f1 score* 89% untuk label *non-hoax*. Yang bisa di lihat pada Gambar 1.

```
--- PERFORMANCE: NAIVE BAYES ---
Train Accuracy: 0.91
Test Accuracy : 0.88
```

	precision	recall	f1-score	support
hoax	0.95	0.81	0.87	1260
non_hoax	0.83	0.96	0.89	1260
accuracy			0.88	2520
macro avg	0.89	0.88	0.88	2520
weighted avg	0.89	0.88	0.88	2520

Gambar 1 Hasil Training Naive Bayes

Selanjutnya dilakukan Permodelan klasifikasi model *Random Forest* dilakukan menggunakan library Scikit-Learn dengan memanfaatkan modul *RandomForestClassifier*. Kode yang digunakan untuk untuk melakukan permodelan *Random Forest* dapat digunakan kode berikut :

```
# =====
# Model 2: Random Forest
# =====
rf_model = RandomForestClassifier(
    random_state= 42,
    n_jobs=-1,
    n_estimators= 5,
    min_samples_split= 20,
    min_samples_leaf= 5,
    max_samples= 0.2,
    max_depth= None,
    bootstrap= True,
)
rf_model.fit(X_train_tfidf, y_train)

y_pred_rf = rf_model.predict(X_test_tfidf)
y_pred_rf_train =
rf_model.predict(X_train_tfidf)
```

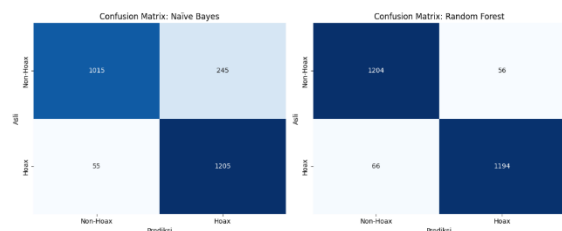
Setelah dilakukan permodelan *Random Forest* didapatkan skor akurasi latih sebesar 96% dan akurasi uji sebesar 95%. Kemudian didapatkan skor *precision* 95%, *recall* 96%, *f1-score* 95% untuk label *hoax* dan didapatkan skor *precision* 96%, *recall* 95%, *f1-score* 95% untuk label *non_hoax*. Hasil skor ini dapat dilihat pada

```
--- PERFORMANCE: RANDOM FOREST ---
Train Accuracy: 0.96
Test Accuracy : 0.95
```

	precision	recall	f1-score	support
hoax	0.95	0.96	0.95	1260
non_hoax	0.96	0.95	0.95	1260
accuracy			0.95	2520
macro avg	0.95	0.95	0.95	2520
weighted avg	0.95	0.95	0.95	2520

Gambar 2 Hasil Skor Training Random Forest

Setelah didapatkan permodelan dari *Naïve Bayes* dan *Random Forest* dilakukan perbandingan prediksi dari kedua model. Perbandingan pertama dilakukan dengan *Confusion matrix*. Yang dapat dilihat pada Gambar 3.



Gambar 3 Confusion Matrik Perbandingan Kedua Model

Dijelaskan bahwa untuk *Naïve Bayes* model berhasil menebak dengan benar 1015 berita yang aslinya *non-hoax* dan terdapat 245 berita *non-hoax* yang salah ditebak model sebagai *hoax*. Kemudian model berhasil menebak dengan benar 1205 berita yang aslinya *hoax* dan terdapat 55 berita *hoax* yang salah ditebak oleh model sebagai *non-hoax*. Setelah itu dijelaskan untuk *Random Forest* model berhasil menebak 1204 berita yang aslinya *non-hoax* dan terdapat 56 berita yang salah ditebak model menjadi *hoax*. Kemudian model berhasil menebak dengan benar 1194 berita yang aslinya *hoax* dan terdapat 66 berita *hoax* yang salah ditebak oleh model sebagai *non-hoax*.

Analisis berikut nya dapat dilakukan menggunakan tabel yang menunjukkan perbandingan *accuracy, precision, recall & F1-score* pada model *Naïve Bayes* dan *Random Forest*. Perbandingan akurasi training testing model *Naïve Bayes* dan *Random Forest* dalam melakukan prediksi berita *hoax* dapat dilihat pada Tabel 1.

Tabel 1 Tabel Perbandingan Skor Klasifikasi

Metrik Evaluasi	Naive Bayes	Random Forest
Accuracy	91%	96%
Precision	95%	95%
Recall	81%	96%
F1-Score	87%	95%

Berdasarkan Tabel 1 diperlihatkan bahwa algoritma *Random Forest* menunjukkan performa yang lebih optimal dibandingkan *Naive Bayes* dengan keunggulan pada nilai *accuracy* sebesar 96%. Meskipun kedua model memiliki tingkat *precision* yang sama yaitu 95%, *Random Forest* jauh lebih efektif dalam meminimalisir kesalahan deteksi, yang dibuktikan dengan nilai *recall* dan *F1-score* yang lebih tinggi secara signifikan.

PENUTUP

Hasil pengujian menunjukkan bahwa algoritma *Random Forest* memiliki performa yang lebih unggul dibandingkan *Naive Bayes* dalam melakukan klasifikasi berita *hoax* pada dataset yang digunakan. Hal ini dibuktikan dengan nilai skor *accuracy* 96%, *precision* 95%, *recall* 96%, *F1-score* 96% yang diperoleh *Random Forest* yang lebih unggul dibandingkan dengan *Naive Bayes* dengan skor skor *accuracy* 91%, *precision* 95%, *recall* 81%, *F1-score* 87%. Perbedaan nilai akurasi tersebut menunjukkan bahwa *Random Forest* lebih efektif dalam mengenali pola dan karakteristik data berita *hoax*.

Kemudian untuk Penelitian selanjutnya disarankan untuk menggunakan dataset dengan jumlah data yang lebih besar dan lebih beragam, baik dari segi sumber berita maupun topik, agar model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik.

Penulis mengucapkan terima kasih kepada Ibu Rina Dewi Indahsari, S.Kom., M.Kom selaku dosen pembimbing, serta Institut Teknologi dan Bisnis Asia Malang yang telah memfasilitasi penelitian ini.

DAFTAR PUSTAKA

[1] Puppala, A. (2025). A Comprehensive Overview of Machine Learning Models: Techniques, Applications, and Future Directions. *International Journal For Multidisciplinary Research*, 7. <https://doi.org/10.36948/ijfmr.2025.v07i02.40603> [akses 7 Februari 2026]

[2] Llatas, C., Soust-Verdaguer, B., Torres, L. C., & Cagigas, D. (2024). Application of Knowledge Discovery in Databases (KDD) to environmental, economic, and social indicators used in BIM workflow to support sustainable design. *Journal of Building Engineering*, 91. <https://doi.org/10.1016/j.job.2024.109546> [akses 7 Februari 2026]

[3] Pajila, P. J. B., Sheena, B. G., Gayathri, A., Aswini, J., Nalini, M., & R, S. S. (2023). A Comprehensive Survey on Naive Bayes Algorithm: Advantages, Limitations and Applications. 2023 4th International Conference on Smart Electronics and Communication (ICOSEC), 1228-1234. <https://doi.org/10.1109/ICOSEC58147.2023.10276274> [akses 7 Februari 2026]

[4] Ramadhan, N. G., & Adhinata, F. D. (2021). TEKNIK SMOTE DAN GINI SCORE DALAM KLASIFIKASI KANKER PAYUDARA. *RADIAL: Jurnal Peradaban Sains, Rekayasa Dan Teknologi*, 9(2), 125-134. <https://doi.org/10.37971/radial.v9i2.229> [akses 7 Februari 2026]

[5] Ionendri, Feri Candra, & Afdi Rizal. (2025). News Classification using Natural Language Processing with TF-IDF and Multinomial Naive Bayes. *Journal of Applied Computer Science and Technology*, 6(1), 37-45. <https://doi.org/10.52158/jacost.v6i1.1099> [akses 7 Februari 2026]

[6] Ghritlahare, A. (2025). The Role of Flowcharts in Problem Solving and Process Visualization. <https://doi.org/10.5281/zenodo.15358152> [akses 7 Februari 2026]

[7] Adam, S., Wahyuddin, M., Seftiyani Botutihe, A., Biko, M., & Mohammad, N. (2025). Mengenali dan Melawan Hoax di Media Sosial Facebook. *SOSIAL: Jurnal Ilmiah Pendidikan IPS*, 3(2), 166-181. <https://doi.org/10.62383/sosial.v3i2.777> [akses 7 Februari 2026]
