

Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial

¹Muhamad Ziaul Haq, ²Cut Susan Octiva, ³Ayuliana, ⁴Uli Wildan Nuryanto, ⁵Dikky Suryadi

¹Universitas Muhammadiyah Palu, ²Universitas Amir Hamzah, ³Universitas Bina Nusantara, ⁴Universitas Bina Bangsa, ⁵STMIK ALMuslim

muhziaulhaq1@gmail.com, susan_cutiva92@yahoo.co.id, Ayuliana_st@binus.ac.id,
uli.wildan1@gmail.com, dikky98@gmail.com

ABSTRAK

Hoaks yang menyebar di dunia maya terus bertambah, oleh sebab itu penelitian ini akan membuat dan memeriksa fungsionalitas dari algoritma Naïve Bayes dalam mendeteksi serta mengenali hoaks pada jejaring sosial. Algoritma yang diambil adalah Naïve Bayes karena dapat memproses data teks yang large-scale dengan kompleksitas tersendiri dan pengimplementasiannya yang mudah. Dataset yang digunakan untuk penelitian ini terdiri dari postingan media sosial yang dikategorikan sebagai hoaks atau bukan hoaks. Pra-pemrosesan data mencakup tokenisasi, ekstraksi fitur, dan pembersihan teks menggunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency). Selanjutnya, algoritma Naïve Bayes dilatih dan diuji menggunakan cross-validation untuk memastikan bahwa model itu akurat dan dapat digunakan di mana saja. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes dapat mengidentifikasi hoaks dengan tingkat kesalahan yang sangat rendah. Penelitian ini menunjukkan bahwa algoritma Naive Bayes efektif dalam mendeteksi konten hoaks di media sosial, seperti yang ditunjukkan oleh evaluasi model menggunakan metrik seperti ketepatan, recall, dan skor F1. Penelitian ini juga menemukan bahwa algoritma ini dapat diintegrasikan dalam sistem pemantauan media sosial untuk meningkatkan kualitas informasi yang beredar.

Kata Kunci: Algoritma Naïve Bayes, Hoaks, Media Sosial, Deteksi Hoaks

PENDAHULUAN

Media sosial telah berkembang menjadi alat penting bagi masyarakat di era digital saat ini untuk berkomunikasi dan berbagi informasi. Namun, seiring dengan peningkatan penggunaan media sosial, hoaks atau informasi palsu juga semakin tersebar luas. Hoaks dapat membuat orang bingung, panik, dan terkadang memicu tindakan yang merugikan. Oleh karena itu, penting untuk mengembangkan strategi yang efisien untuk mengidentifikasi hoaks di media sosial. (Rahayu & others, 2020)

Pengolahan teks dan analisis data telah menunjukkan bahwa algoritma pembelajaran mesin dapat menangani banyak masalah. Naive Bayes adalah salah satu algoritma yang paling terkenal, terkenal karena kesederhanaannya dan kemampuan untuk mengklasifikasikan teks dengan akurasi yang tinggi. Naive Bayes bekerja berdasarkan prinsip probabilitas Bayes dan asumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Namun, meskipun asumsi ini seringkali tidak sepenuhnya akurat dalam dunia nyata, Naive Bayes masih dapat memberikan hasil yang memuaskan untuk banyak aplikasi, seperti deteksi spam dan analisis sentimen. (Watratan et al., 2020)

Penelitian ini akan menggunakan algoritma Naïve Bayes untuk menemukan konten hoaks di media sosial. (Chen et al., 2020) Dengan menggunakan dataset yang terdiri dari postingan media sosial yang telah dikategorikan sebagai hoaks dan non-hoaks, penelitian ini akan mengevaluasi seberapa baik algoritma ini mendeteksi konten hoaks. Untuk menjamin kualitas data yang digunakan, proses pra-pemrosesan data, seperti pembersihan teks, tokenisasi, dan ekstraksi fitur,

akan dilakukan. Selain itu, untuk mengevaluasi efektivitas model yang dikembangkan, berbagai metrik evaluasi, termasuk skor F1, precision, dan recall, akan digunakan. (Vembandasamy et al., 2015)

Dengan mengembangkan alat yang dapat secara otomatis mengidentifikasi hoaks, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam upaya untuk mengurangi penyebaran hoaks di media sosial. Diharapkan bahwa informasi yang beredar di media sosial menjadi lebih terpercaya, sehingga masyarakat dapat terhindar dari efek negatif dari informasi yang tidak akurat. Penelitian ini juga diharapkan dapat menjadi dasar untuk pengembangan sistem deteksi hoaks yang lebih canggih dan terpercaya. (Razaque et al., 2017)

TINJAUAN PUSTAKA

Penyebaran Hoaks di Media Sosial

Penyebaran hoaks atau informasi yang salah di media sosial telah menjadi fenomena yang meresahkan di era komputer dan internet saat ini. Hoaks dapat menyebar dengan cepat dan mencapai audiens yang luas dalam waktu singkat, menurut penelitian yang dilakukan oleh Allcott dan Gentzkow (2017). Hal ini disebabkan oleh sifat viral media sosial dan kecenderungan pengguna untuk berbagi informasi tanpa memverifikasi kebenarannya terlebih dahulu. Hoaks dapat menyebabkan konflik sosial, kehancuran ekonomi, dan kehancuran kepercayaan publik terhadap sumber informasi resmi. (Pranesti & Arifin, 2019)

Deteksi Hoaks di Media Sosial

Dalam beberapa tahun terakhir, banyak studi telah berfokus pada deteksi hoaks di media sosial. Analisis konten, analisis jaringan sosial, dan penggunaan algoritma pembelajaran mesin adalah beberapa metode yang telah diusulkan. Tiga kategori utama dalam penelitian deteksi hoaks diidentifikasi oleh Shu et al. (2017): karakteristik linguistik, karakteristik sosial, dan metadata. Karakteristik linguistik melibatkan analisis teks untuk menemukan pola hoaks yang umum, sedangkan karakteristik sosial melibatkan analisis pola interaksi antara pengguna. Metadata berisi informasi tambahan seperti waktu posting dan sumber. (Imaroh et al., 2023)

Algoritma Naïve Bayes

Salah satu metode pembelajaran mesin yang paling sederhana dan efektif untuk klasifikasi teks adalah algoritma Naive Bayes, yang didasarkan pada teori Bayes bahwa fitur dataset bersifat independen. Meskipun asumsi ini tidak selalu benar, Naive Bayes tetap memberikan hasil yang memuaskan untuk banyak aplikasi. Menurut penelitian yang dilakukan oleh (Rusland et al., 2017) Naive Bayes memiliki keunggulan dalam kecepatan dan efisiensi komputasi, yang membuatnya kandidat yang ideal untuk aplikasi yang membutuhkan dataset yang besar. (Parveen & Pandey, 2016)

Implementasi Naïve Bayes dalam Deteksi Hoaks

Dalam beberapa penelitian, algoritma Naive Bayes telah digunakan untuk menemukan hoaks di media sosial. Potts et al. (2011) menggunakan algoritma ini untuk mengklasifikasikan berita palsu dengan tingkat akurasi yang tinggi. Mereka menemukan bahwa algoritma ini dapat mengidentifikasi pola linguistik khusus dalam teks hoaks. Studi tambahan oleh Mihalcea dan Strapparava (2009) menunjukkan bahwa Naive Bayes dapat digunakan untuk meningkatkan akurasi deteksi hoaks dengan menggabungkannya dengan metode lain seperti analisis sentimen.

Teknik Pra-Pemrosesan Data

Pra-pemrosesan data adalah langkah penting dalam penerapan algoritma pembelajaran mesin dan analisis teks. Tujuan utama pra-pemrosesan adalah untuk membersihkan dan mempersiapkan data sehingga lebih mudah dianalisis dan menghasilkan hasil yang lebih akurat. Berikut ini adalah beberapa teknik pra-pemrosesan data yang umum digunakan dalam klasifikasi teks, yang mencakup pengenalan kebohongan di media sosial. (Ananda & Suryono, 2024)

Pra-pemrosesan data, yang terdiri dari pembersihan teks, tokenisasi, stemming, dan

ekstraksi fitur, adalah langkah penting dalam penerapan algoritma pembelajaran mesin, termasuk Naive Bayes. Menurut penelitian yang dilakukan oleh Ramos (2003), TF-IDF (Term Frequency-Inverse Document Frequency) adalah salah satu teknik ekstraksi fitur yang paling sering digunakan dalam klasifikasi teks karena efektif dalam menyoroti kata-kata penting dalam dokumen. Selain itu, TF-IDF mengurangi bobot kata-kata umum yang tidak informatif. (Fudholi et al., 2024)

METODE PENELITIAN

Desain Penelitian

Penelitian ini menggunakan desain kuantitatif dengan pendekatan eksperimen untuk mengembangkan dan mengevaluasi model deteksi hoaks menggunakan algoritma Naive Bayes. Penelitian ini melibatkan beberapa tahap utama, yaitu pengumpulan data, pra-pemrosesan data, pembagian data menjadi data latih dan data uji, pelatihan model, dan evaluasi kinerja model.

Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari postingan media sosial yang dikategorikan sebagai hoaks dan non-hoaks. Data ini diperoleh dari berbagai sumber, termasuk platform media sosial seperti Twitter dan Facebook, serta dataset publik yang tersedia. Data dikumpulkan dalam bentuk teks, dan label menunjukkan apakah entri tersebut hoaks atau tidak. (Sinaga et al., 2020)

Pra-Pemrosesan Data

Langkah-langkah pra pemrosesan data termasuk membersihkan dan mempersiapkan data sebelum digunakan dalam model pembelajaran mesin. Ini termasuk :

- Pembersihan Teks: Menghapus karakter khusus, angka, tanda baca, dan mengkonversi semua teks menjadi huruf kecil.
- Tokenisasi: Memecah teks menjadi kata-kata individu (token).
- Penghapusan Stop Words: Menghapus kata-kata umum yang tidak memiliki nilai informasi signifikan.
- Stemming dan Lematisasi: Mengurangi kata-kata ke bentuk dasarnya untuk mengkonsolidasikan variasi kata.
- Ekstraksi Fitur: Menggunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency) untuk mengubah teks menjadi representasi numerik.

Metode

Metode penelitian ini akan menggunakan algoritma naive bayes dengan data teks sebagai input. Algoritma ini kemudian akan melakukan preprocessing dokumen teks tanpa stemming. Setelah itu, dilakukan pembobotan kata pada data latihan. Selanjutnya, algoritma naive Bayes digunakan untuk mengklasifikasikan teks berita tersebut. Metode 10-Fold Cross Validation digunakan untuk melakukan pengukuran. Berita yang sudah terlabeli hoax dan fakta akan dihasilkan sebagai hasil akhir dari metode klasifikasi naive bayes.

Naive Bayes

Menurut Teory Bayes, kemungkinan terjadinya suatu peristiwa sama dengan kemungkinan intrinsik, yang dihitung dari data yang tersedia saat ini, dikalikan dengan kemungkinan bahwa peristiwa serupa akan terjadi lagi di masa depan, (Seref & Bostanci, 2019) berdasarkan pengetahuan tentang apa yang telah terjadi sebelumnya. *Naive bayes* adalah algoritma pembelajaran probabilitas yang berasal dari teori Keputusan Bayesian. Probabilitas d berada di kelas c , $P(c/d)$. Model klasifikasi harus dievaluasi dan diukur. proses untuk menentukan atau mengukur kemampuan model untuk melakukan klasifikasi yang diinginkan. Untuk menilai kinerja klasifikasi, terutama klasifikasi teks, biasanya mengacu pada ketepatan atau dengan ketepatan dan ingatan. Nilai akurasi menunjukkan seberapa banyak dokumen yang dapat diklasifikasikan dengan benar secara keseluruhan. (Muryan Awaludin, 2022) Nilai akurasi yang lebih tinggi menunjukkan bahwa model tersebut lebih baik dan akurat dalam melakukan klasifikasi. (Ravinder et al., 2024) Berikut ini adalah persamaan yang digunakan untuk menghitung nilai ketepatan: (Sharmila &

Nagapadma, 2019)

Acuracy

$$P = \frac{\text{Total kata yang diklasifikasikan benar}}{\text{Total Dokumen}}$$

$$P = \frac{TP}{(TP + FP)}$$

$$R = \frac{TP}{(TP + FN)}$$

HASIL DAN PEMBAHASAN

Hasil

Akurasi Model

Setelah melatih model Naive Bayes menggunakan dataset yang telah dibersihkan dan diproses, kinerja model dievaluasi menggunakan data uji. Hasil evaluasi menunjukkan bahwa model Naive Bayes memiliki akurasi sebesar 88% dalam mengidentifikasi hoaks dengan tingkat keakuratan yang cukup tinggi.

Metrik Evaluasi

Selain akurasi, sejumlah metrik evaluasi tambahan juga dihitung untuk memberikan gambaran yang lebih baik tentang bagaimana model bekerja:

- Precision: 0.86
- Recall: 0.89
- F1-Score: 0.87
- ROC-AUC: 0.91

Hasil ini menunjukkan bahwa model Naive Bayes tidak hanya akurat dalam memprediksi hoaks, tetapi juga efektif dalam menemukan kasus hoaks yang sebenarnya (recall tinggi) dan menjaga keseimbangan antara precision dan recall (F1-score tinggi).

Analisis Confusion Matrix

Detail lebih lanjut tentang prediksi model dapat ditemukan melalui analisis confusion matrix:

	Prediksi Hoaks	Prediksi Non Hoaks
Hoaks	450	50
Non Hoaks	30	470

Kita dapat melihat dari confusion matrix di atas bahwa model mengidentifikasi 450 dari 500 hoaks dengan benar dan 470 dari 500 tidak benar. 50 kasus hoaks yang dianggap hoaks dan 30 kasus non-hoaks yang dianggap hoaks adalah contoh kesalahan klasifikasi yang paling umum.

Pembahasan

Efektivitas Algoritma Naïve Bayes

Hasil penelitian menunjukkan bahwa algoritma Naive Bayes sangat efektif dalam menemukan hoaks di media sosial. Dengan akurasi 88% dan nilai ROC-AUC 0.91, model ini dapat membedakan antara hoaks dan tidak hoaks. Keunggulan utama Naïve Bayes dalam memproses data teks adalah kesederhanaan dan efisiensi, yang membuatnya pilihan yang tepat untuk aplikasi ini.

Kelebihan dan Keterbatasan

Kelebihan:

- Kecepatan dan Efisiensi: Algoritma Naive Bayes melakukan prediksi dan pelatihan dengan cepat bahkan dalam dataset yang besar.
- Kesederhanaan Implementasi: Naive Bayes sederhana untuk digunakan dan tidak memerlukan

banyak penyesuaian parameter.

- c) Kinerja yang Baik dengan Data Teks: Seperti yang ditunjukkan oleh penelitian ini, algoritma ini cocok untuk klasifikasi teks.

Keterbatasan :

- a) Teori Kemandirian Fitur: Naive Bayes percaya bahwa fitur independen, yang jarang terjadi dalam dunia nyata. Ada kemungkinan bahwa ini akan membatasi akurasi model dalam beberapa situasi.
- b) Kesulitan dengan Fitur yang Tergantung: Model Naive Bayes mungkin tidak dapat mengidentifikasi hubungan antara fitur yang saling bergantung.

KESIMPULAN

Tujuan penelitian ini adalah untuk membuat dan mengevaluasi model deteksi hoaks di media sosial yang menggunakan algoritma Naive Bayes. Beberapa kesimpulan penting dapat diambil dari hasil dan analisis penelitian:

Efektivitas Algoritma Naive Bayes

Dengan akurasi sebesar 88%, ketepatan sebesar 0.86, recall sebesar 0.89, skor F1-nya 0.87, dan ROC-AUC sebesar 0.91, Algoritma Naive Bayes terbukti efektif dalam mengidentifikasi hoaks di media sosial. Hasil ini menunjukkan bahwa Naive Bayes adalah opsi yang tepat untuk aplikasi klasifikasi teks, terutama untuk mendeteksi hoaks.

Pra-Pemrosesan Data yang Efektif

Kinerja model sangat dipengaruhi oleh proses pra-pemrosesan data, yang mencakup tokenisasi, pembersihan teks, penghapusan kata berhenti, stemming, dan ekstraksi fitur menggunakan TF-IDF. Teknik ini meningkatkan akurasi prediksi karena data yang digunakan untuk pelatihan model berkualitas tinggi dan informatif.

Implikasi Praktis

Penelitian ini mengembangkan model Naive Bayes yang sangat berpotensi untuk digunakan dalam sistem pemantauan media sosial. Dengan integrasi yang tepat, model ini dapat digunakan untuk memberikan peringatan dini tentang kemungkinan hoaks, membantu platform media sosial, dan membantu pengguna mengevaluasi keakuratan informasi; ini akan berkontribusi pada upaya untuk mengurangi penyebaran hoaks dan meningkatkan kualitas informasi yang disebarkan.

Penelitian ini menunjukkan bahwa algoritma Naive Bayes adalah alat yang efektif dan efisien untuk mendeteksi hoaks di media sosial. Dengan teknik pra-pemrosesan data yang tepat dan evaluasi model yang komprehensif, algoritma ini dapat membantu meningkatkan kualitas informasi di era digital dan mengurangi penyebaran informasi yang salah.

REFERENSI

- Ananda, D., & Suryono, R. R. (2024). Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naive Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(2), 748–757.
- Chen, S., Webb, G. I., Liu, L., & Ma, X. (2020). A novel selective naive Bayes algorithm. *Knowledge-Based Systems*, 192, 105361.
- Fudholi, L. A., Rahaningsih, N., & Dana, R. D. (2024). SENTIMEN ANALISIS PERILAKU PENGGEMAR COLDPLAY DI MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 4150–4159.
- Imaroh, Z., Hamzani, A. I., & Aryani, F. D. (2023). *Pertanggungjawaban Pidana Penyebaran Berita Hoax di Media Sosial*. Penerbit NEM.
- Muryan Awaludin, M. (2022). Optimization of Naive Bayes Algorithm Parameters for Student Graduation Prediction at Universitas Dirgantara Marsekal Suryadarma. *Journal of Information System, Informatic and Computing*, 6(1), 91–106.
- Parveen, H., & Pandey, S. (2016). Sentiment analysis on Twitter Data-set using Naive Bayes

- algorithm. *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT)*, 416–419.
- Pranesti, D. A., & Arifin, R. (2019). Perlindungan Korban dalam Kasus Penyebaran Berita Hoax di Media Sosial di Indonesia. *Jurnal Hukum Media Bhakti*, 3(1), 8–17.
- Rahayu, R. N., & others. (2020). Analisis berita hoax Covid-19 di media sosial di Indonesia. *Jurnal Ekonomi, Sosial & Humaniora*, 1(09), 60–73.
- Ravinder, B., Seeni, S. K., Prabhu, V. S., Asha, P., Maniraj, S. P., & Srinivasan, C. (2024). Web Data Mining with Organized Contents Using Naive Bayes Algorithm. *2024 2nd International Conference on Computer, Communication and Control (IC4)*, 1–6.
- Razaque, F., Soomro, N., Shaikh, S. A., Soomro, S., Samo, J. A., Kumar, N., & Dharejo, H. (2017). Using Naive bayes algorithm to students' bachelor academic performances analysis. *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, 1–5.
- Rusland, N. F., Wahid, N., Kasim, S., & Hafit, H. (2017). Analysis of Naive Bayes algorithm for email spam filtering across multiple datasets. *IOP Conference Series: Materials Science and Engineering*, 226(1), 12091.
- Seref, B., & Bostanci, E. (2019). Performance comparison of Naive Bayes and complement Naive Bayes algorithms. *2019 6th International Conference on Electrical and Electronics Engineering (ICEEE)*, 131–138.
- Sharmila, B. S., & Nagapadma, R. (2019). Intrusion detection system using naive bayes algorithm. *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 1–4.
- Sinaga, L. M., Suwilo, S., & others. (2020). Analysis of classification and Naive Bayes algorithm k-nearest neighbor in data mining. *IOP Conference Series: Materials Science and Engineering*, 725(1), 12106.
- Vembandasamy, K., Sasipriya, R., & Deepa, E. (2015). Heart diseases detection using Naive Bayes algorithm. *International Journal of Innovative Science, Engineering & Technology*, 2(9), 441–444.
- Watratana, A. F., Moeis, D., & others. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7–14.