p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

Identifikasi Beritas Palsu Pada Situs Berita Di Indonesia Menggunakan Model *Naive Bayes*

Nabil Akhdan[1], Popy Meilina[1]*, Rully Mujiastuti[1], Hendra[1]
[1]Teknik Informatika, Universitas Muhammadiyah Jakarta, Jakarta Pusat, 10510
**Corresponding Author* : popy.meilila@umj.ac.id

Abstrak

Penyebaran berita palsu di internet menjadi tantangan utama di era digital. Penelitian ini bertujuan membangun model identifikasi berita palsu pada situs berita di Indonesia menggunakan algoritma Naive Bayes Classifier. Dataset diperoleh dari Kaggle dan GitHub yang berisi berita dengan label fakta atau palsu. Tahapan penelitian mencakup text preprocessing, pembobotan TF-IDF, dan klasifikasi. Hasil evaluasi menggunakan data latih dan uji (skenario 80:20) menunjukkan model mampu mendeteksi berita palsu dengan akurasi sebesar 89%, precision 89%, dan recall 89%. Sistem yang dikembangkan diharapkan dapat membantu masyarakat mengenali berita palsu serta meningkatkan literasi digital.
Kata kunci: Berita Palsu, Identifikasi Berita, Naive Bayes, Pemrosesan Teks, TF-IDF.

Abstract

The spread of fake news on the internet has become a major challenge in the digital era. This study aims to develop a fake news identification model for news websites in Indonesia using the Naive Bayes Classifier algorithm. The dataset was obtained from Kaggle and GitHub, containing news articles labeled as either factual or fake. The research includes several stages such as text preprocessing, TF-IDF weighting, and classification. The evaluation results, based on an 80:20 train-test split, show that the model achieved 89% accuracy, 89% precision, and 89% recall. This system is expected to assist the public in identifying fake news more easily and contribute to improving digital literacy.
Keywords : Fake News, Naive Bayes, News Identification, Text Processing, TF-ID

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

PENDAHULUAN

Kemajuan teknologi saat ini memungkinkan masyarakat untuk mengakses informasi dengan cepat. Informasi tersebut dapat diperoleh dari berbagai platform digital, seperti situs web, aplikasi berita, media sosial, dan lainnya. Perkembangan internet berperan besar dalam mempercepat penyebaran informasi di berbagai kanal. Bahkan, kemajuan internet memungkinkan komunikasi antar individu di lokasi berbeda terjadi secara instan. Oleh karena itu, tidak mengherankan jika suatu peristiwa yang baru terjadi dapat menyebar ke seluruh dunia hanya dalam hitungan menit. Namun, seiring dengan meningkatnya kecepatan penyebaran informasi, jumlah berita dari sumber yang tidak terpercaya atau belum terverifikasi juga semakin bertambah. Informasi yang disampaikan bisa saja mengalami distorsi, dilebih-lebihkan, atau bahkan sepenuhnya tidak benar. Jenis berita yang beredar seperti ini dikenal sebagai hoaks atau berita palsu (Amal et al., 2023).

Di era digital yang berkembang pesat ini, di mana setiap orang dapat berperan sebagai sumber informasi, membedakan antara fakta dan fiksi menjadi tantangan yang semakin rumit. Media sosial yang memiliki kecepatan dan jangkauan luas memungkinkan berita palsu menyebar dengan mudah dan memengaruhi persepsi publik. Hoaks tidak hanya menyebar melalui situs web dan media sosial dengan kecepatan yang mengkhawatirkan, tetapi juga berpotensi memanipulasi opini publik, merusak reputasi, dan bahkan memengaruhi hasil pemilu. Situasi ini menjadi ancaman bagi keberlanjutan informasi yang akurat dan tepercaya di tengah masyarakat. Oleh karena itu, upaya untuk mengatasi penyebaran berita palsu di media sosial menjadi semakin mendesak (Raza, 2024).

Berdasarkan data dari Kementerian Komunikasi dan Informatika (Kominfo, 2023), hingga Mei 2023, Tim AIS Direktorat Jenderal Aplikasi Informatika telah mengidentifikasi sebanyak 11.642 konten hoaks sejak Agustus 2018. Dari jumlah tersebut, kategori hoaks yang paling banyak ditemukan adalah seputar kesehatan dengan 2.287 kasus, disusul oleh 2.111 konten terkait pemerintahan, 1.938 konten yang berhubungan dengan penipuan, serta 1.373 konten bertemakan politik. Penelusuran yang dilakukan Tim AIS juga mengungkapkan bahwa jumlah total hoaks, disinformasi, dan misinformasi yang tersebar dari Januari hingga Mei 2023 mengalami peningkatan dibandingkan dengan periode yang sama pada tahun sebelumnya. Hoaks yang berkaitan dengan pemerintahan, seperti penggunaan nama pejabat publik secara tidak sah dan modus penipuan, menjadi yang paling dominan, meskipun isu kesehatan, terutama yang berkaitan dengan pandemi Covid-19, masih ditemukan.

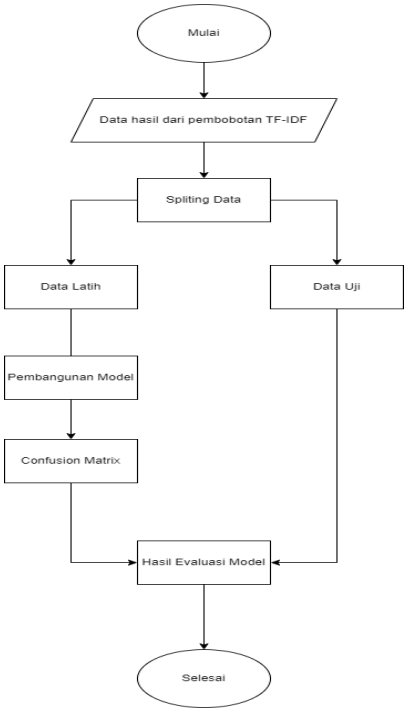
Menurut Hanifah (2023), Tim AIS dibentuk pada Januari 2018 dengan tujuan mendeteksi, mengidentifikasi, dan memverifikasi konten yang beredar di dunia maya di Indonesia. Konten yang diawasi mencakup hoaks, terorisme, radikalisme, pornografi, perjudian, serta berbagai jenis konten negatif lainnya. Saat ini, Tim AIS memiliki 100 personel yang bekerja dengan dukungan mesin AIS, yang beroperasi tanpa henti selama 24 jam sehari dan 7 hari dalam seminggu. Mayoritas hoaks yang terdeteksi berkaitan dengan isu politik, kesehatan, dan agama.

Kemajuan pesat dalam teknologi informasi telah menyebabkan penyebaran informasi di internet menjadi semakin tidak terkendali, termasuk beredarnya dokumen yang mengandung informasi palsu. Di Indonesia, perkembangan teknologi informasi terus meningkat seiring dengan bertambahnya jumlah pengguna internet. Berdasarkan survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2017, jumlah pengguna internet di Indonesia telah mencapai 143,26 juta orang, yang setara dengan sekitar 54,56% dari total populasi sebanyak 262 juta jiwa (Mustofa & Mahfudh, 2019).

Saat ini, salah satu metode yang dapat digunakan untuk mengidentifikasi berita palsu adalah dengan mencari klarifikasi terkait berita tersebut. Jika tidak ditemukan sumber klarifikasi, maka berita tersebut dapat dicurigai sebagai berita palsu. Namun, proses identifikasi berita palsu secara manual membutuhkan waktu yang cukup lama, sementara penyebaran berita palsu terjadi dengan sangat cepat setiap harinya.

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

Pemrosesan teks dan analisis data telah membuktikan bahwa algoritma pembelajaran mesin mampu menyelesaikan berbagai permasalahan. Salah satu algoritma yang paling dikenal adalah Naive Bayes, yang terkenal karena kesederhanaannya serta kemampuannya dalam mengklasifikasikan teks dengan tingkat akurasi yang tinggi. Naive Bayes didasarkan pada prinsip probabilitas Bayes dengan asumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Meskipun dalam praktiknya asumsi ini tidak



selalu sepenuhnya akurat, algoritma Naive Bayes tetap mampu memberikan hasil yang memuaskan dalam berbagai aplikasi, seperti deteksi spam dan analisis sentimen (Haq et al., 2024).

Untuk mengatasi permasalahan tersebut, dikembangkan sebuah website yang mampu mengklasifikasikan suatu berita sebagai berita palsu atau berita fakta. Website ini dibangun dengan menerapkan algoritma Naïve Bayes. Diharapkan, penggunaan website ini dapat membantu masyarakat dalam membedakan berita palsu dari berita yang valid. Data yang digunakan dalam penelitian ini diperoleh dari platform Kaggle dan GitHub, mengingat keberhasilan algoritma Naïve Bayes sangat bergantung pada kualitas dataset yang digunakan dalam proses pelatihan dan evaluasi model.

METODE PENELITIAN

Metodologi penelitian identifikasi berit *Cross Stadar Industry For Data Mining (CRISP-DM)* ini terdiri dari, *Business Understanding*, *Data Understanding, Data Preparation*, *Modeling*, *Evaluation*, dan implementasi. Sebagai berikut:

1. Pengumpulan Data

Dalam penelitian ini, data yang digunakan berasal dari dataset yang tersedia di platform Kaggle dan GitHub, mencakup periode tahun 2015 hingga 2020. Setelah semua dataset digabungkan, diperoleh total 13.499 data. Dataset ini disimpan dalam format .csv dan terdiri dari kolom berita serta tagging. Label data diklasifikasikan ke dalam dua kategori, yaitu 1 untuk berita palsu (false) dan 0 untuk berita fakta (true). Dari total dataset yang terkumpul, 5.697 data dikategorikan sebagai berita fakta, sementara 7.802 data merupakan berita palsu.

2. *Pra-Processing* Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset yang tersedia di platform Kaggle, kemudian diolah lebih lanjut untuk digunakan dalam perhitungan TF-IDF. Perhitungan TF-IDF bertujuan untuk memberikan bobot pada setiap kata dalam dokumen serta membantu dalam pemilihan kata atau fitur yang memiliki nilai informasi tinggi dalam teks. Adapun tahapan pra-pemrosesan, yaitu *cleansing, case folding, tokenizing, stopwords, steamming.*

3. Pembobotan Kata TF-IDF

Pembobotan Term-Frequency Inverse Document Frequency (TF-IDF) adalah metode yang digunakan untuk memberikan bobot pada kata-kata dalam sebuah dokumen guna mencari data dengan tingkat relevansi tinggi. Tahap ini merupakan lanjutan dari proses preprocessing, di mana dalam pembobotan kata dilakukan beberapa perhitungan, seperti:

- Menghitung Document Frequency (DF), yaitu jumlah dokumen yang mengandung kata tertentu.
- Menghitung Inverse Document Frequency (IDF), yaitu perhitungan untuk menurunkan bobot kata yang sering muncul di banyak dokumen dan meningkatkan bobot kata yang jarang

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

muncul.

- Menghitung TF-IDF, yaitu kombinasi antara TF (Term Frequency) dan IDF (Inverse Document Frequency).

Pada tahap ini, proses perhitungan TF dilakukan dengan menghitung frekuensi kemunculan suatu kata dalam dokumen, sedangkan IDF dihitung setelah mendapatkan nilai DF. Metode ini diterapkan setelah proses labelling dan preprocessing selesai, dengan tujuan mengukur probabilitas kemunculan setiap kata dalam kumpulan data yang digunakan.

4. *Modeling*

Setelah proses pembobotan kata selesai, langkah berikutnya adalah klasifikasi menggunakan metode Naïve Bayes. Proses klasifikasi ini terdiri dari empat tahap utama, yaitu:

1. Menggunakan data hasil pembobotan TF-IDF sebagai input untuk model.
2. Membagi data (Splitting Data) ke dalam set pelatihan dan pengujian.
3. Membangun model dengan menerapkan algoritma Naïve Bayes.
4. Evaluasi model menggunakan confusion matrix untuk mengukur kinerja klasifikasi.

Pada gambar 1 yang disertakan, terdapat flowchart yang menggambarkan alur implementasi metode Naïve Bayes dalam penelitian ini.

5. *Evaluation*

Gambar 1. Flowchart Implemntasi Naive Bayes

Pada tahap evaluasi, kinerja model Naïve Bayes dalam mengklasifikasikan berita palsu dan berita fakta diukur menggunakan Confusion Matrix serta metrik Accuracy, Precision, Recall, dan F1-Score.

Pembagian Data	Data Latih	Data Uji
80% data latih, 20% data uji	12.483	3.121
70% data latih, 30% data uji	10.922	4.682
60% data latih, 40% data uji	9.362	6.242

Gambar 2. Hasil Split Data Latih dan Data Uji

Gambar 2 merupakan hasil tiga skenario dari pembagian data dengan 80% data latih 20% data uji, 70% data latih 30% data uji, dan 60% data latih 40% data uji

6. *Deployment*

Setelah model dibangun, selanjutnya adalah mendeploy model menggunakan Flask. Proses ini melibatkan mengintegrasikan model Naive Bayes ke dalam aplikasi web menggunakan framework Flask. Ini memungkinkan pengguna untuk berinteraksi dengan model melalui antarmuka web.

HASIL DAN PEMBAHASAN

Dalam penelitian ini, proses identifikasi berita palsu dilakukan menggunakan model Naïve Bayes. Untuk memperoleh hasil dengan tingkat akurasi terbaik, serangkaian uji coba telah dilakukan dengan menerapkan algoritma Naïve Bayes pada dataset yang terdiri dari 13.499 data berita palsu dan fakta.

Pengujian dilakukan dengan membagi data ke dalam tiga skenario:

- 80% data latih dan 20% data uji
- 70% data latih dan 30% data uji
- 60% data latih dan 40% data uji

Dari hasil pengujian, skenario 80:20 menghasilkan performa model terbaik. Implementasi mencakup proses pelatihan model, evaluasi kinerja, serta analisis hasil yang diperoleh dari model tersebut.

Pembagian Data	Data Latih	Data Uji	Accuracy
80% data latih, 20% data uji	12.483	3.121	89%
70% data latih, 30% data uji	10.922	4.682	88%
60% data latih, 40% data uji	9.362	6.242	88%

Gambar 3. Hasil Pengujian Split Data

Gambar 3 menampilkan hasil uji coba pembagian data yang dilakukan dengan membagi data latih dan data uji untuk menentukan skenario yang menghasilkan akurasi paling optimal. Berdasarkan hasil percobaan, skenario terbaik diperoleh dengan perbandingan 80:20, di mana 80% digunakan sebagai data latih dan 20% sebagai data uji, menghasilkan akurasi sebesar 89%.

Setelah proses pelatihan model selesai,

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

evaluasi kinerja dilakukan menggunakan berbagai metrik, termasuk akurasi, classification report, precision, recall, dan F1-score.

Model Naïve Bayes telah dilatih dan dievaluasi menggunakan skenario pembagian data 80:20 untuk pelatihan dan pengujian. Berikut adalah hasil evaluasi yang memberikan gambaran mengenai kinerja model berdasarkan metrik yang digunakan.

Classification Report	Precision	Recall	F1- Score	Support
0 (Fakta)	0.89	0.89	0.89	1561
1 (Palsu)	0.89	0.89	0.89	1560
Accuracy			0.89	3121
Macro Avg	0.89	0.89	0.89	3121
Weighted Avg	0.89	0.89	0.89	3121

Gambar 4. Hasil Classification Report test_size 0.2

Gambar 4 menampilkan hasil evaluasi model Naïve Bayes dengan pembagian data 80% untuk pelatihan dan 20% untuk pengujian. Model ini mencapai akurasi sebesar 89%, dengan nilai Precision, Recall, dan F1-Score masing-masing sebesar 0,89 untuk kelas berita fakta dan berita palsu.

Selain itu, nilai macro average dan weighted average yang sama menunjukkan bahwa model beroperasi secara seimbang tanpa adanya bias terhadap salah satu kelas. Hasil ini mengindikasikan bahwa model memiliki kinerja yang baik dalam mengklasifikasikan berita fakta dan berita palsu.

Classification Report	Precision	Recall	F1- Score	Support
0 (Fakta)	0.89	0.87	0.88	2341
1 (Palsu)	0.87	0.89	0.88	2341
Accuracy			0.88	4682
Macro Avg	0.88	0.88	0.88	4682
Weighted Avg	0.88	0.88	0.88	4682

Gambar 5. Hasil Classification Report test_size 0.3

Gambar 5 menyajikan hasil evaluasi model Naïve Bayes dengan pembagian data 70% untuk pelatihan dan 30% untuk pengujian. Model ini mencapai akurasi sebesar 88%, dengan nilai Precision dan Recall yang seimbang antara kelas berita fakta (0,89 dan 0,87) serta berita palsu (0,87 dan 0,89).

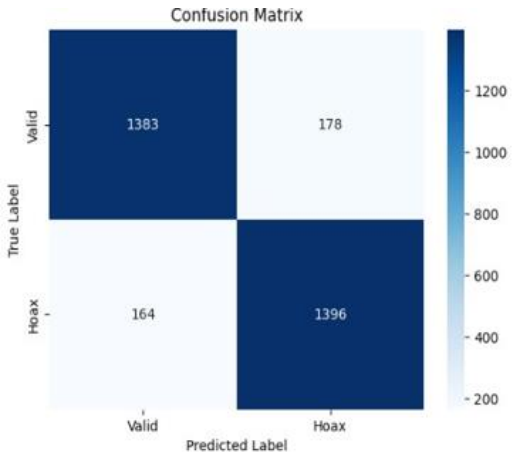
Selain itu, nilai F1-Score sebesar 0,88 menunjukkan keseimbangan antara precision dan recall. Hasil ini sedikit lebih rendah dibandingkan dengan skenario pembagian 80:20, yang mengindikasikan bahwa jumlah data latih yang lebih sedikit dapat berdampak pada penurunan kinerja model.

Classification Report	Precision	Recall	F1- Score	Support
0 (Fakta)	0.88	0.87	0.88	3121
1 (Palsu)	0.87	0.89	0.88	3121
Accuracy			0.88	6242
Macro Avg	0.88	0.88	0.88	6242
Weighted Avg	0.88	0.88	0.88	6242

Gambar 6. Hasil Classification Report test_size 0.4

Gambar 6 menampilkan hasil evaluasi model Naïve Bayes dengan pembagian data 60% untuk pelatihan dan 40% untuk pengujian. Model ini memperoleh akurasi sebesar 88%, dengan Precision dan Recall yang hampir seimbang antara kelas berita fakta (0,88 dan 0,87) serta berita palsu (0,87 dan 0,89).

Selain itu, nilai F1-Score sebesar 0,88 menunjukkan bahwa model tetap bekerja dengan cukup baik meskipun jumlah data latih lebih sedikit dibandingkan skenario sebelumnya. Namun, hasil ini juga menunjukkan bahwa mengurangi jumlah data latih tidak memberikan peningkatan performa yang signifikan, sehingga penggunaan lebih banyak data latih tetap menjadi faktor penting dalam meningkatkan akurasi model.



Gambar 7. Confussion Matrix

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

Gambar 7 menampilkan hasil evaluasi model pada data uji setelah dilakukan pembagian data dengan skenario terbaik, yaitu 80:20. Dalam penelitian ini, beberapa skenario pembagian data telah diuji, yaitu 80:20, 70:30, dan 60:40.

Berdasarkan hasil pengujian, skenario 80:20 menghasilkan performa terbaik dibandingkan skenario lainnya. Oleh karena itu, confusion matrix yang ditampilkan dalam gambar merupakan hasil evaluasi dari skenario terbaik ini.

Hasil Tampilan

Bagian ini menyajikan hasil implementasi sistem identifikasi berita palsu menggunakan model Naïve Bayes. Sistem ini dikembangkan dengan memanfaatkan dataset berita yang telah melalui tahap preprocessing dan diuji menggunakan algoritma klasifikasi Naïve Bayes untuk menilai kinerjanya dalam mendeteksi berita palsu.

Gambar 8. Halaman Utama Website

Gambar 8 menampilkan tampilan awal sistem identifikasi berita palsu. Pada halaman utama, pengguna dapat memasukkan teks berita yang ingin diperiksa ke dalam kolom input yang tersedia. Setelah itu, pengguna dapat menekan tombol "Cek Berita" untuk memproses berita tersebut menggunakan model Naïve Bayes.

Gambar 9. Tampilan Hasil Berita Fakta

Gambar 9 menunjukkan hasil ketika sistem mengklasifikasikan suatu berita sebagai berita fakta. Setelah pengguna memasukkan teks berita dan menekan tombol cek, sistem akan menganalisis berita tersebut dan menampilkan hasil klasifikasi. Jika berita tersebut termasuk dalam kategori fakta, maka sistem akan memberikan label Berita Fakta.

Gambar 10. Tampilan Hasil Berita Palsu

Gambar 10 menunjukkan hasil ketika sistem mengidentifikasi suatu berita sebagai berita palsu. Setelah pengguna menginput teks berita dan menjalankan analisis, sistem akan memberikan hasil klasifikasi yang menunjukkan bahwa berita tersebut termasuk dalam kategori berita palsu.

Sistem ini dirancang untuk membantu pengguna dalam membedakan berita fakta dan berita palsu dengan cepat dan mudah berdasarkan analisis teks menggunakan model Naïve Bayes.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, maka diperoleh kesimpulan sebagai berikut :

a. Model Naïve Bayes telah digunakan untuk mengklasifikasikan berita sebagai fakta atau palsu. Berdasarkan evaluasi dengan tiga skenario pembagian data (80:20, 70:30, dan 60:40), model menunjukkan kinerja terbaik dengan akurasi 89% pada skenario 80% data latih dan 20% data uji. Hasil ini membuktikan bahwa model Naïve Bayes dapat secara efektif mendeteksi berita palsu dan berita fakta.

b. Implementasi model Naïve Bayes pada sistem berbasis website memungkinkan pengguna untuk dengan mudah mengidentifikasi berita palsu dan berita

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

fakta. Dengan adanya website ini, pengguna dapat lebih cepat mendeteksi berita yang berpotensi hoaks, sehingga dapat membantu mengurangi penyebaran berita palsu yang dapat berdampak negatif di masyarakat.

DAFTAR PUSTAKA

Andryadi, A. A., & Hasri Fatonah, N. (2021). Analisis User Experience Dan User Interface (Ui/Ux) Pada Website Menggunakan Metodegoogle Design Sprint. Jurnal Teknologi Dan Bisnis, 3(2), 137–144. https://doi.org/10.37087/jtb.v3i2.61

Damayanti, C., Triayudi, A., & Sholihati, I. D. (2022). Analisis UI/UX Untuk Perancangan Website Apotek dengan Metode Human Centered Design dan System Usability Scale. Jurnal Media Informatika Budidarma, 6(1), 551.https://doi.org/10.30865/mib.v6i1.3526

Andryadi, A. A., & Hasri Fatonah, N. (2021). Analisis User Experience Dan User Interface (Ui/Ux) Pada Website Menggunakan Metodegoogle Design Sprint. Jurnal Teknologi Dan Bisnis, 3(2), 137–144. https://doi.org/10.37087/jtb.v3i2.61

Haq, M. Z., Octiva, C. S., Ayuliana, A., Nuryanto, U. W., & Suryadi, D. (2024). Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial. Jurnal Minfo Polgan, 13(1), 1079–1084. https://doi.org/10.33395/jmp.v13i1.13937

Rizki As Sidiq, V. A., Triyadi, S., & Pratiwi, W. D. (2022). Analisis Kelengkapan Unsur Berita Detik.com serta Relevansinya sebagai Bahan Ajar. Jurnal Pendidikan Bahasa, 11(2), 240–264. https://doi.org/10.31571/bahasa.v11i2.4202

Raza, M. N. (2024). Sistem Deteksi Berita Hoax Menggunakan Algoritma Naïve Bayes Dan Random Forest Pada Machine Learning. Pondasi: Journal of Applied Science Engineering, 1(2), 43–57. https://journal.alshobar.or.id/index.php/pondasi/article/view/221

Soen, G. I. E., Marlina, M., & Renny, R. (2022). Implementasi Cloud Computing dengan Google Colaboratory pada Aplikasi Pengolah Data Zoom Participants. JITU : Journal Informatic Technology And Communication, 6(1), 24–30. https://doi.org/10.36596/jitu.v6i1.781

Nurohanisah, S., Astuti, R., Basysyar, F. M., Informatika, T., Informasi, S., Informasi, S., Cirebon, K., Palsu, B., & Forest, A. R. (2024). Deteksi berita palsu menggunakan algoritma random forest. 8(1), 422–428.

Naury, C., Fudholi, D. H., & Hidayatullah, A. F. (2021). Topic Modelling pada Sentimen Terhadap Headline Berita Online Berbahasa Indonesia Menggunakan LDA dan LSTM. Jurnal Media Informatika Budidarma, 5(1), 24. https://doi.org/10.30865/mib.v5i1.2556

Cardova, A., & Hermawan, A. (2023). Implementasi Metode LSTM Untuk Mengklasifikasi Berita Palsu Pada PolitiFact. Jurnal Fasilkom, 13(3), 471–479. https://doi.org/10.37859/jf.v13i3.6175

Amal, I., Pamungkas, E. W., Kom, S., & Kom, M. (2023). Aplikasi Pendeteksi Berita Palsu Bahasa Indonesia Menggunakan Framework Flask dan Streamlit serta Algoritma Machine Learning. 1–18. https://eprints.ums.ac.id/id/eprint/116531%0A https://eprints.ums.ac.id/116531/1/Naskah Publikasi_Ikhlasul Amal.pdf

Ula, M. (2020). Analisa Dan Deteksi Konten Hoax Pada Media Berita Indonesia Menggunakan Machine Learning. Jurnal Teknologi Terapan and Sains 4.0, 1(2), 229. https://doi.org/10.29103/tts.v1i2.3263

Agustina, N., & Hermawati, M. (2021). Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media. Factor Exacta, 14(4), 1979–276. https://doi.org/10.30998/faktorexacta.v14i4.11259

Prasetya, F., & Ferdiansyah, F. (2022). Analisis Data Mining Klasifikasi Berita Hoax COVID 19 Menggunakan Algoritma Naive Bayes. Jurnal Sistem Komputer Dan Informatika (JSON), 4(1), 132. https://doi.org/10.30865/json.v4i1.4852

Komparasi, S., & Dan, M. S. (2024). MODEL XGBOOST PADA KLASIFIKASI RUMAH TANGGA. 5(3), 2266–2283.

Shevira, S., Suarjaya, I. M. A. D., & Buana, P. W. (2022). Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia. JITTER : Jurnal Ilmiah Teknologi Dan Komputer, 3(2),

p - ISSN : 2407 – 1846
e - ISSN : 2460 – 8416
Website : jurnal.umj.ac.id/index.php/semnastek

1074.https://doi.org/10.24843/jtrti.2022.v03.i02.p06

Mustofa, H., & Mahfudh, A. A. (2019). Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes. Walisongo Journal of Information Technology, 1(1), 1. https://doi.org/10.21580/wjit.2019.1.1.3915

Ryansyah, A., & Andayani, S. (2022). Implementasi Algoritma TF-IDF Pada Pengukuran Kesamaan Dokumen. Jurnal Sistem & Teknologi Informasi Komunikasi, 1(1), 58–62. http://pdfbox.apache.org/

Luh, N., & Astariyani, G. (2021). Pengaturan Hoax (Berita Bohong) Dalam Perspektif Perbandingan Hukum Di Indonesia. Jurnal Kertha Semaya, 9(9), 1688–1702.

Sriyano, C. S., & Setiawan, E. B. (2021). Pendeteksian Berita Hoax Menggunakan Naive Bayes Multinomial Pada Twitter dengan Fitur Pembobotan TF-IDF. Repository Telkom University, 8(2), 3396.

Soleman, S. (2021). Pemanfaatan Metode Klasifikasi Naïve Bayes Untuk Pendeteksi Berita Hoax Pada Artikel Berbahasa Indonesia. Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi, 7(2), 83. https://doi.org/10.24014/coreit.v7i2.14290

Santiyasa, I. W., Brahmantha, G. P. A., Supriana, I. W., Kadyanan, I. G. G. A., Suhartana, I. K. G., & Mahendra, I. B. M. (2022). Identification of Hoax Based on Text Mining Using K-Nearest Neighbor Method. JELIKU (Jurnal Elektronik Ilmu Komputer Udayana), 10(2), 217. https://doi.org/10.24843/jlk.2021.v10.i02.p04

Sani, R. R., Pratiwi, Y. A., Winarno, S., Udayanti, E. D., & Alzami, F. (2022). Analisis Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Berita Hoax pada Berita Online Indonesia. Jurnal Masyarakat Informatika, 13(2), 85–98. https://doi.org/10.14710/jmasif.13.2.47983

Ibrahim, M., Bu'ulolo, E., & Lubis, I. (2020). Penerapan Algoritma Naive Bayes Classifier Untuk Mendeteksi Tingkat Krediblitas Hoax News/ Fake News Pada Sosial Media Di Indonesia Berbasis Android (Studi Kasus : Kantor Tribun Medan). Resolusi: Rekayasa Teknik Informatika Dan Informasi, 1(1), 9–17.