



Detection of Hoax News Using TF-IDF Vectorizer and Multinomial Naïve Bayes and Passive Aggressive

Rizky Adrian

Department of Informatics, Dinamika Bangsa University, Jambi, Indonesia
E-mail: adrian070070@gmail.com

Musaddam

Department of Informatics, Dinamika Bangsa University, Jambi, Indonesia
E-mail: musaddam00@gmail.com
*Corresponding Author

Muhammad Ikhsan

Department of Informatics, Dinamika Bangsa University, Jambi, Indonesia
E-mail: oppofyouth45@gmail.com

M. Riza Pahlevi. B

Department of Informatics, Dinamika Bangsa University, Jambi, Indonesia
E-mail: rizapahlevi@unama.ac.id

Received: 18 May, 2024; Accepted: 21 June, 2024; Published: 30 June, 2024

Abstract: The website is a source of information, but not all information is guaranteed to be correct. Some news can be considered hoaxes or not based on facts. This research aims to build a hoax news detection system on English language news websites. The method used involves the multinomial Naïve Bayes and Passive Aggressive approaches. Classification report analysis shows the superiority of the Passive Aggressive Classifier with significant improvements in all evaluation metrics compared to Multinomial Naïve Bayes. The conclusion is based on the characteristics of the dataset, confirming the effectiveness of the Passive Aggressive Classifier in solving the task of classifying fake news in English, with the highest accuracy reaching 93.74%.

Keywords: Hoax; News Website; Tf-idf Vectorizer; Naïve Bayes Multinomial; Passive Aggressive

I. Introduction

Perkembangan teknologi internet saat ini berlangsung dengan sangat pesat, memudahkan masyarakat dalam mengakses internet dan informasi. Akses yang mudah ini memungkinkan penyebaran informasi melalui jaringan sosial maupun media online menjadi semakin luas. Namun, dalam kenyataannya, banyak beredar berita yang mengandung informasi palsu atau hoaks yang tidak dapat dipertanggungjawabkan. Penyebaran berita hoaks ini bertujuan untuk meyakinkan pembaca agar terpengaruh oleh narasi yang dibangun dalam berita tersebut [1].

Berita hoaks memiliki karakteristik yang khas, seperti penggunaan huruf kapital secara keseluruhan, adanya garis bawah, banyak tanda seru, ajakan untuk disebar, ketiadaan tanggal yang dapat diverifikasi, serta alamat URL yang tidak valid atau mencurigakan. Dikemas dengan sensasionalisme, berita hoaks ini membentuk persepsi di kalangan masyarakat yang sulit dikendalikan, sehingga menimbulkan kekhawatiran akan potensi konflik. Dampaknya sangat signifikan terutama dalam bidang politik, agama, budaya, dan berbagai aspek kehidupan lainnya, merusak tatanan sosial dan mempengaruhi stabilitas masyarakat [7],[8].

Website dan blog terkenal sebagai media penyebaran berita, yang dapat memiliki sudut pandang positif, negatif, atau netral. Artikel berita yang tersebar di situs-situs ini sering kali dapat diarahkan oleh penulisnya ke arah tipuan atau informasi yang tidak benar. Informasi palsu dan menyesatkan ini berbahaya karena dapat menyesatkan persepsi manusia dengan menyampaikan informasi yang tidak benar seolah-olah itu adalah kebenaran, sehingga berpotensi membawa dampak negatif pada pemikiran dan pandangan masyarakat [2].

Hal ini mengakibatkan banyaknya ragam berita yang disajikan di web portal berita, sehingga sangat sulit untuk melakukan filtrasi dan klarifikasi atas kebenaran dan validitas pemberitaan. Persaingan antar website portal berita untuk mendapatkan lebih banyak akses di dunia maya dan berada di halaman pertama mesin pencari dengan kata kunci apa pun mendorong mereka untuk menggunakan isu-isu yang sedang naik daun, bahkan yang belum terbukti kebenarannya dan cenderung hoaks. Berita-berita ini sering kali disajikan dengan gaya bahasa yang menggigit sehingga membuat banyak orang mempercayainya. Untuk mengetahui apakah sebuah berita mengandung informasi hoaks atau tidak, salah satu teknik yang dapat digunakan adalah text mining. Text mining adalah metode yang dapat digunakan untuk mengklasifikasi data, dengan mencari pola-pola menarik dalam koleksi data teks [3].

Penelitian sebelumnya menyoroti permasalahan klasifikasi berita palsu dengan membandingkan metode Naïve Bayes dan C4.5 pada media sosial, menunjukkan bahwa Naïve Bayes memiliki tingkat akurasi lebih tinggi dibandingkan C4.5 dalam mendeteksi berita palsu. Penelitian lain menggunakan metode Passive Aggressive dan TF-IDF pada media sosial, khususnya Twitter, berhasil menciptakan model machine learning dengan tingkat akurasi sebesar 82.44%. Dalam penelitian ini, solusi yang diusulkan adalah menggunakan metode Multinomial Naïve Bayes dengan Passive Aggressive dan pembobotan TF-IDF untuk melihat performansi dari kedua metode tersebut dalam mendeteksi berita hoaks. Kontribusi penelitian ini adalah memberikan analisis komparatif yang lebih mendalam antara metode Multinomial Naïve Bayes dan Passive Aggressive dengan TF-IDF dalam konteks deteksi berita hoaks, serta memberikan wawasan yang dapat digunakan untuk meningkatkan akurasi deteksi berita palsu. Sistematika penulisan pada penelitian ini adalah sebagai berikut: bagian kedua berisi metode yang diusulkan, bagian ketiga mencakup hasil dan diskusi, serta bagian terakhir menyajikan kesimpulan.

2. Research Method

2.1. Experiment Setup

Ada beberapa tahapan penelitian yang harus dilakukan untuk mencapai tujuan penelitian ini. Oleh karena itu, perlu dirancang alur penelitian yang ditunjukkan pada Fig 1

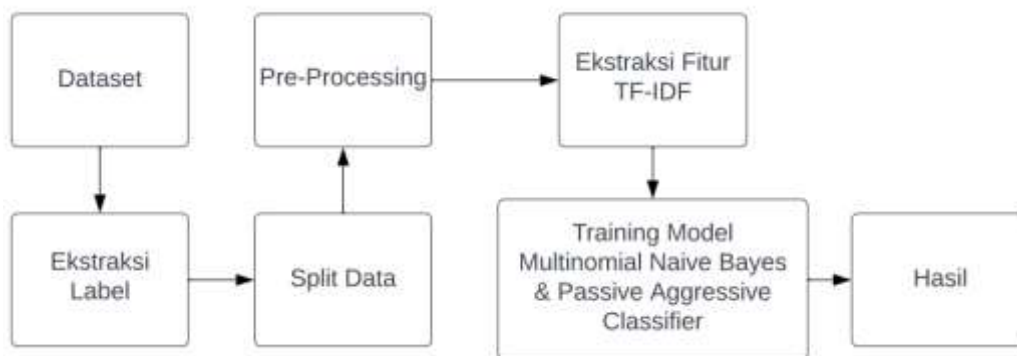


Fig.1. Experiment setup

Berikut adalah tahapan yang dilakukan dalam penelitian ini, yang dapat dibagi menjadi tiga fase:

- Melakukan ekstraksi fitur pada dataset dan membagi dataset menjadi dua bagian, yaitu data training dan data testing.
- Selanjutnya, menggunakan metode Multinomial Naïve Bayes dan Passive Aggressive untuk proses klasifikasi
- Terakhir, melakukan pengujian dengan menggunakan data testing dan menghitung metris keberhasilan seperti akurasi, presisi, dan recall

2.2. Dataset

Penelitian ini memanfaatkan dataset yang diambil dari Kaggle.com dengan ukuran sebanyak 6335 baris dan terdiri dari 4 kolom utama, yaitu 'Unnamed: 0', 'title', 'text', dan 'label'. Dalam konteks dataset ini, kolom pertama ('Unnamed: 0') berfungsi sebagai identifier untuk setiap berita, sementara kolom kedua ('title') dan ketiga ('text') berisi informasi mengenai judul dan teks berita. Kolom keempat ('label') memuat label yang menandakan apakah suatu berita dianggap Real atau Fake. Penampakan dari dataset tersebut dapat dilihat pada Fig 2.

unnamed: 0		title	text	label
0	8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello...	FAKE
1	10294	Watch The Exact Moment Paul Ryan Committed Poi...	Google Pinterest Digg LinkedIn Reddit Stumbleu...	FAKE
2	3600	Kerry to go to Paris in gesture of sympathy	U.S. Secretary of State John F. Kerry said Mon...	REAL
3	10142	Bernie supporters on Twitter erupt in anger ag...	— Kaydee King (@KaydeeKing) November 9, 2016 T...	FAKE
4	875	The Battle of New York: Why This Primary Matters	It's primary day in New York and front-runners...	REAL
5	6903	Tehran, USA	'In'm not an immigrant, but my grandparents ...	FAKE
6	7341	Girl Horrified At What She Watches Boyfriend D...	Share This Baylee Luciani (left). Screenshot o...	FAKE
7	95	'Britain's Schindler' Dies at 106	A Czech stockbroker who saved more than 650 Je...	REAL
8	4669	Fact check: Trump and Clinton at the 'commande...	Hillary Clinton and Donald Trump made some ina...	REAL
9	2909	Iran reportedly makes new push for uranium con...	Iranian negotiators reportedly have made a las...	REAL

Fig.2. Dataset

2.3. Pre-Processing

Sering kali beberapa kata yang umum digunakan tampaknya bernilai sedikit dalam membantu memilih dokumen yang sesuai dengan kebutuhan pengguna (Christopher, Raghavan and Schütze 2009). Kata-kata ini disebut dengan stopword. Cara umum untuk menentukan daftar *stopword* adalah mengurutkan berdasarkan frekuensi kemunculan istilah kata pada kumpulan dokumen. Kemudian diambil kata-kata atau istilah yang paling sering muncul dan dikumpulkan menjadi *stopword list*. Anggota dari *stopword list* ini akan dibuang dari teks pada pengindeksan kata saat pelatihan dan pengujian. Pada tahap ini akan dilakukan proses *stopword*, proses ini bertujuan menghilangkan kata-kata yang tidak relevan dari kolom judul berita. Untuk prosesnya digunakan modul *stopword* dari pustaka *scikit-learn* dengan modul *TfidfVectorizer* dalam python untuk mengambil *stopwords* dalam bahasa Inggris [3]

2.4. Feature Extraction

Setelah pra-pemrosesan selesai, langkah selanjutnya melibatkan pengkodean atau transformasi teks ke dalam format numerik. Salah satu metode untuk mencapai hal ini adalah melalui penggunaan *TfidfVectorizer*, yang dirumuskan menggunakan persamaan berikut:

$$TF - IDF = TF \log \left(\frac{n}{DF} \right) \quad (1)$$

Persamaan ini menghitung frekuensi suatu istilah (kata) dalam suatu dokumen menggunakan TF, jumlah dokumen yang mengandung istilah (kata) menggunakan DF, dan N jumlah seluruh dokumen. proses mengubah teks asli menjadi representasi vektor menggunakan metode TF-IDF (Term Frekuensi-Inverse Document Frekuensi). TF-IDF adalah teknik yang digunakan untuk memberikan bobot pada kata-kata dalam dokumen berdasarkan seberapa sering kata tersebut muncul dalam dokumen dan seberapa umum kata tersebut di seluruh koleksi dokumen. [5]

2.5. Classification

2.5.1 Multinomial Naïve Bayes

Naïve Bayes merupakan sebuah metoda klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ada beberapa macam model pada Naïve Bayes, pada penelitian ini penulis menggunakan *Naïve Bayes Multinomial*. Model multinomial memperhitungkan *frekuensi* setiap kata yang muncul pada dokumen. Misal terdapat dokumen d dan himpunan kelas c . Untuk memperhitungkan kelas dari dokumen d , maka dapat dihitung dengan rumus

$$P(c|\text{term dokumen } d) = P(c) \times P(t_1|c) \times P(t_2|c) \times P(t_3|c) \times \dots \times P(t_n|c) \quad (2)$$

$P(c|\text{term dokumen } d)$ adalah probabilitas suatu dokumen termasuk kelas c . t_n yaitu kata dokumen d ke- n . $P(t_n|c)$ merupakan probabilitas kata ke- n dengan diketahui kelas c . Sedangkan $P(c)$ adalah probabilitas prior dari kelas c . Untuk mencari nilai $P(c)$ dapat dihitung dengan rumus:

$$p(c) = \frac{N_c}{N} \quad (3)$$

$P(c)$ adalah jumlah kelas c pada seluruh dokumen, dan N jumlah seluruh dokumen. Sementara rumus Multinomial yang digunakan dengan pembobotan kata TF-IDF adalah sebagai berikut:

$$p(tn/c) = \frac{Mct+1}{(\sum W' \in V W' ct) + B'} \quad (4)$$

$P(tn|c)$ adalah probabilitas kata ke- n dengan diketahui kelas c . Wct adalah nilai pembobotan TF-IDF atau W dari term t di kategori c . $(\sum W' \in V W' ct)$ merupakan jumlah total W dari keseluruhan term yang berada di kategori c . (B') yaitu jumlah W kata unik (nilai idf tidak dikali dengan tf) pada seluruh dokumen [6].

2.5.2 Passive Aggressive

Algoritma *Passive Aggressive* umumnya digunakan dalam pembelajaran besar dan tugas klasifikasi. Ini efektif untuk pembelajaran online dan diaplikasikan secara inkremental pada dataset. Model yang menggunakan algoritma ini dapat terus belajar saat menerima data baru, cocok untuk sistem yang memproses aliran data kontinu. Keuntungan utama adalah penggunaan memori yang lebih efisien.

Algoritma *Passive Aggressive* digunakan untuk mengklasifikasikan data dengan membuat penyesuaian iteratif pada vektor bobot. Jika terjadi kesalahan klasifikasi, algoritma memperbarui bobot dengan agresif; jika tidak, bobot tetap.

Rumus dasar algoritma ini adalah:

$$w_new = w_old + (loss / \|x\|^2) * y * x \quad (5)$$

Persamaan ini menunjukkan proses memperbarui vektor bobot (w_new) setelah menganalisis data pelatihan. Vektor bobot sebelum pembaruan ditandai sebagai w_old . Kerugian atau kesalahan yang merupakan hasil dari klasifikasi yang salah diwakili oleh "loss." Selain itu, " $\|x\|^2$ " melambangkan kuadrat norma Euclidean dari vektor fitur " x ." " y " merujuk pada label kelas yang sesuai untuk sampel data, sedangkan " x " menunjukkan vektor fitur dari sampel data [5].

2.6. Confusion Matrix dan Classification Report

Berbagai tindakan evaluasi digunakan untuk mengevaluasi keakuratan klasifikasi algoritma dalam mendeteksi berita palsu. Pada bagian ini, metrik ukuran yang paling sering digunakan (Confusion Matrix) untuk mendeteksi berita palsu telah digunakan. Melalui rumusan ini sebagai tugas klasifikasi, dimungkinkan untuk menentukan ukuran-ukuran yang dimiliki matriks konfusi seperti di bawah ini [10],[2].

- True Positive (TP) : sebagai ketepatan prediksi, actual hoax terdeteksi sebagai hoax
- True Negative (TN) : sebagai aktual bukan hoax terdeteksi sebagai bukan hoax
- False Positive (FP) : Sebagai jumlah aktual hoax yang terdeteksi sebagai bukan hoax
- False Negative (FN) : sebagai kesalahan prediksi, dimana aktual bukan hoax dideteksi sebagai hoax.

Berikut adalah nilai-nilai yang digunakan:

- Accuracy : melakukan deteksi bukan hoax dengan benar sebagai presentase dari jumlah total prediksi yang benar.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (6)$$

- Precession : merupakan angka prediksi benar positif dibandingkan dengan nilai keseluruhan hasil yang diprediksi positif.

$$\text{Precession} = \frac{TP}{TP+FP} \quad (7)$$

- Recall : melakukan deteksi sebagian besar bukan hoax tetapi akan ada disertai kesalahan dalam klasifikasi data sehingga ada beberapa hoax terdeteksi sebagai bukan hoax.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

- F1-Score : merupakan perbandingan rata-rata dari precision dan recall yang dibobotkan

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{9}$$

2.7. Enviroment Setup

Penelitian ini dilaksanakan pada sebuah laptop dengan spesifikasi tertentu, yaitu menggunakan sistem operasi Windows 11, dilengkapi dengan prosesor 12th Gen Intel(R) Core(TM) i7-12700H 2.70 GHz serta RAM berkapasitas 16 GB. Selain itu, alat-alat analisis yang diperlukan dalam penelitian meliputi Python, Scikit-learn, Google Colab.

3. Results and Discussion

3.1. Hoax News Prediction Results

Analisis hasil dataset dilakukan melalui pemanfaatan *confusion matrix* serta *classification report*. Proses ini melibatkan 6 kali pengujian pada dua algoritma, yakni Multinomial Naïve Bayes dan Passive Aggressive. Dibawah ini merupakan tampilan *confusion matrix* dan *classification report* yang diperoleh dari pengujian tersebut yang dapat di lihat di Tabel 1, Tabel 2, Fig.3, Fig 4.

Tabel 1. forest Classification report Multinomial Naïve Bayes dengan TF-IDF

Steps	Accuracy	Recall	Precision	F1-score
1	85.65%	85.65%	87.51%	85.40%
2	85.65%	85.65%	87.51%	85.40%
3	85.65%	85.65%	87.51%	85.40%
4	85.65%	85.65%	87.51%	85.40%
5	85.65%	85.65%	87.51%	85.40%
6	85.65%	85.65%	87.51%	85.40%

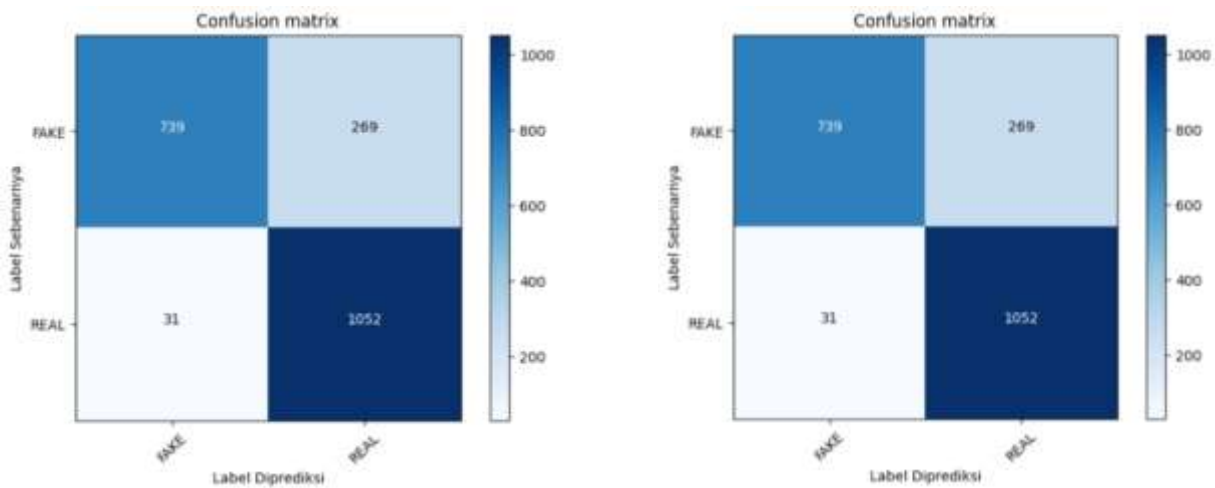


Fig.3. Confusion Matrix Multinomial Naïve Bayes dengan TF-IDF

Tabel 2. Classification report Passive Aggressive dengan TF-IDF

Step	Accuracy	Recall	Precision	F1-score
1	93.35%	93.35%	93.38%	93.35%
2	93.45%	93.45%	93.49%	93.45%
3	93.59%	93.59%	93.62%	93.59%
4	93.74%	93.74%	93.77%	93.74%
5	93.74%	93.74%	93.77%	93.74%
6	93.64%	93.64%	93.69%	93.64%

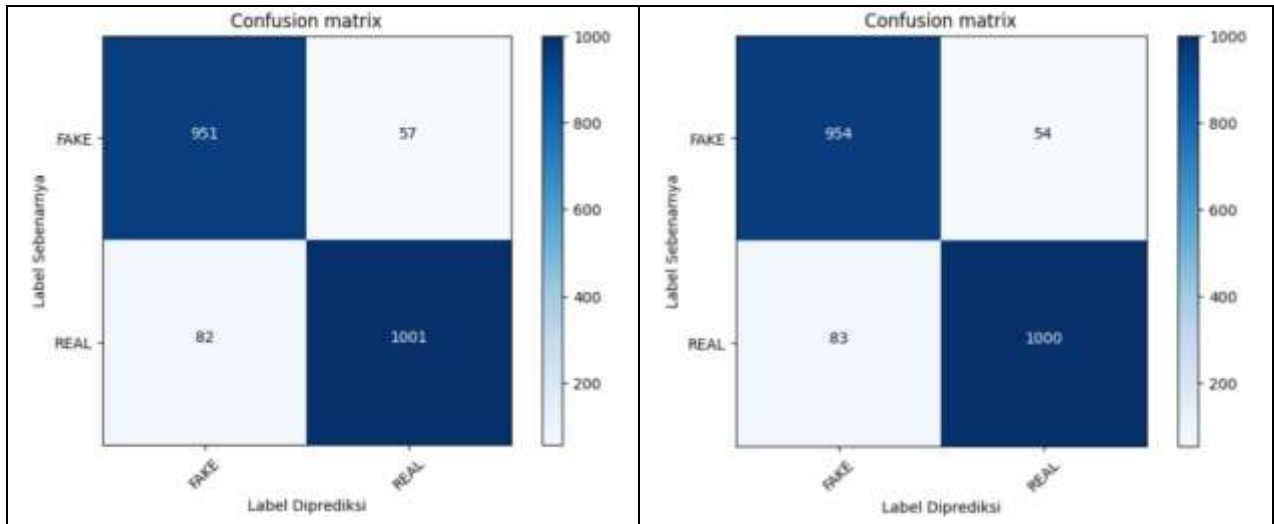


Fig.4. Confusion Matrix Passive Aggressive dengan TF-IDF

Dari hasil classification report untuk Multinomial Naïve Bayes, terlihat bahwa model tersebut memiliki stabilitas performa yang tinggi pada dataset tertentu, dengan accuracy yang konsisten pada 85.65% di setiap langkah pengujian. *Recall* yang stabil pada 85.65% menunjukkan kemampuan model untuk mengidentifikasi sebagian besar sampel positif secara konsisten, sedangkan *precision* yang tetap pada 87.51% mengindikasikan ketepatan model dalam memberi label positif. *F1-score* yang konstan pada 85.40% mencerminkan nilai harmonik antara *recall* dan *precision*.

Sementara itu, Passive Aggressive Classifier menunjukkan peningkatan performa yang signifikan. *Accuracy*, *recall*, *precision*, dan *F1-score* meningkat seiring dengan langkah pengujian, mencapai nilai tertinggi pada langkah terakhir dengan nilai masing-masing 93.74%, 93.74%, 93.77%, dan 93.64%. Dengan demikian, dapat disimpulkan bahwa Passive Aggressive Classifier dengan TF-IDF menunjukkan performa yang lebih baik dibandingkan dengan Multinomial Naïve Bayes pada dataset ini, dengan peningkatan yang konsisten pada semua metrik evaluasi. Tetapi, perlu diingat bahwa kesimpulan ini bergantung pada karakteristik dan sifat dataset yang digunakan.

3.2. Web page

Sistem ini akan diimplementasikan dalam sebuah website. Website ini akan dibangun menggunakan bahasa pemrograman Python dan Streamlit, sebuah library Python yang memungkinkan pembuatan web app yang ditampilkan seperti pada Fig 5 [3].



Fig.5. Halaman Web

4. Conclusion

Penelitian ini menyimpulkan bahwa Passive Aggressive Classifier dengan TF-IDF menunjukkan performa yang lebih superior dibandingkan dengan Multinomial Naïve Bayes dalam konteks klasifikasi pada dataset yang diuji. Hasil analisis menggunakan classification report menunjukkan bahwa Multinomial Naïve Bayes memiliki stabilitas performa yang baik, sementara Passive Aggressive Classifier menonjol dengan peningkatan yang signifikan pada semua metrik evaluasi, termasuk akurasi, recall, precision, dan F1-score. Kesimpulan ini mengindikasikan bahwa model Passive Aggressive Classifier lebih cocok untuk penanganan tugas klasifikasi dalam kerangka penelitian ini. Oleh karena itu, dalam konteks penelitian ini, *Passive Aggressive Classifier* dengan TF-IDF merupakan pilihan yang lebih efektif dalam mendeteksi dan mengklasifikasikan berita hoax berbahasa Inggris.

Acknowledgment

Penelitian ini didukung oleh Universitas Dinamika Bangsa, Jambi, Indonesia

References

- [1] Rozi, Faizal Nur, and Dwi Harini Sulistyawati. "KLASIFIKASI BERITA HOAX PILPRES MENGGUNAKAN METODE MODIFIED K-NEAREST NEIGHBOR DAN PEMBOBOTAN MENGGUNAKAN TF-IDF." *KONVERGENSI* 15, no. 1 (October 1, 2019). <https://doi.org/10.30996/konv.v15i1.2828>.
- [2] Rahutomo, Faisal, Ingrid Yanuar Risca Pratiwi, and Diana Mayangsari Ramadhani. "Eksperimen Naïve Bayes Pada Deteksi Berita Hoax Berbahasa Indonesia." *JURNAL PENELITIAN KOMUNIKASI DAN OPINI PUBLIK* 23, no. 1 (July 1, 2019). <https://doi.org/10.33299/jpkop.23.1.1805>.
- [3] Nabila Rizqi Amalia Nur Asri, Rendra Soekarta, Muhammad Yusuf "DeteksiTeksBerita Hoax Bahan Bakar Minyak (BBM) Menggunakan Metode K-Nearest Neighbor"
- [4] Handika, Yola Tri, Sarjon Defit, and Gunadi Widi Nurcahyo. "TEXT MINING DALAM MEMBANDINGKAN METODE NAÏVE BAYES DENGAN C.45 DALAM MENGIDENTIFIKASI BERITA HOAX PADA MEDIA SOSIAL." *Rang Teknik Journal* 5, no. 1 (January 6, 2022): 116–23. <https://doi.org/10.31869/rtj.v5i1.2855>.
- [5] Lazuardi, Maulana Fajar, Renaldy Hiunarto, Kareena Putri Ramadhani, Noviandi Noviandi, Riya Widayanti, and Muhamad Hadi Arfian. "Hoax News Detection Using Passive Aggressive Classifier And TfidfVectorizer." *JURNAL TEKNIK*
- [6] Lazuardi, Maulana Fajar, Renaldy Hiunarto, Kareena Putri Ramadhani, Noviandi Noviandi, Riya Widayanti, and Muhamad Hadi Arfian. "Hoax News Detection Using Passive Aggressive Classifier And TfidfVectorizer." *JURNAL TEKNIK INFORMATIKA* 16, no. 2 (December 22, 2023): 185–93. <https://doi.org/10.15408/jti.v16i2.34084>.
- [7] Information Security and Networking Research Group (In FORSNET), Faculty of Information Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia., Sy. Yuliani, Shahrin Sahib, Information Security and Networking Research Group (InFORSNET), Faculty of Information Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia., Mohd Faizal Bin Abdollah, Information Security and Networking Research Group (InFORSNET), Faculty of Information Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia., Fariska Z. Ruskanda, and School of Electrical Engineering and Informatics, School of Electrical Engineering and Informatics, Indonesia. "Hoax News Classification Using Machine Learning Algorithms." *International Journal of Engineering and Advanced Technology* 9, no. 2 (December 30, 2019): 3938–44. <https://doi.org/10.35940/ijeat.B3753.129219>.
- [8] Safitri, Intan Anggreani, and Teddy Dyatmika. "Pentingnya Literasi Bahaya Hoax Kepada Masyarakat Desa Sidorejo di Era Globalisasi," 2021.
- [9] Wirawan, Rio, and M Bayu Wibisono. "PERANCANGAN APLIKASI SOSIAL MEDIA UNTUK ANALISIS BERITA HOAX," 2017.
- [10] Department of Computer Science, College of Science, Al-NahrainUniversity, Baghdad, Iraq, Reham Jehad, Suhad A.Yousif, and Department of Computer Science, College of Science, Al-NahrainUniversity, Baghdad, Iraq. "Fake News Classification Using Random Forest and Decision Tree (J48)." *Al-Nahrain Journal of Science* 23, no. 4 (December 1, 2020): 49–55. <https://doi.org/10.22401/ANJS.23.4.09>.

Authors' Profiles



Muhammad Ikhsan. lahir di Jambi, Indonesia. Saat ini dia tengah menempuh Program Sarjanah Informatika di Universitas Dinamika Bangsa, Indonesia. Dia memiliki fokus penelitian pada machine learning, Algoritma, dan software engineering



Rizky Adrian. lahir di Jambi, Indonesia. Saat ini dia tengah menempuh Program Sarjanah Informatika di Universitas Dinamika Bangsa, Indonesia. Dia memiliki fokus penelitian pada machine learning dan software engineering



Musaddam. lahir di Jambi, Indonesia. Saat ini dia tengah menempuh Program Sarjana Informatika di Universitas Dinamika Bangsa, Indonesia. Minat penelitian pada machine learning dan



M. Riza Pahlevi. B meraih gelar Sarjana di bidang Teknik Informatika dari STIKOM Dinamika Bangsa, Indonesia, dan gelar M. Kom di bidang Ilmu Komputer dari Universitas Budi Luhur Jakarta, Indonesia. Saat ini beliau adalah dosen di Fakultas Ilmu Komputer, Universitas Dinamika Bangsa, Indonesia. Minat penelitiannya meliputi Data Mining, Machine Learning, dan Sistem Informasi