

## Model Logistic Regression dan TF-IDF untuk Deteksi Dini Berita Palsu Berbahasa Indonesia

Abdul Hamid Arribathi<sup>1</sup>, Sunan Reihan Jungjunan<sup>2</sup>, Muhammad Adrian Dwiharyanto<sup>\*3</sup>

<sup>1</sup> Program Magister Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Raharja

<sup>2,3</sup> Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Raharja

e-mail: [1abdulhamid@raharja.info](mailto:abdulhamid@raharja.info), [2sunan@raharja.info](mailto:sunan@raharja.info), [\\*3muhammad.adrian@raharja.info](mailto:muhammad.adrian@raharja.info)

### Abstrak

Penyebaran berita palsu, terutama di platform media sosial, menjadi tantangan signifikan akibat derasnya aliran informasi digital. Mayoritas sistem deteksi berita palsu saat ini berfokus pada bahasa Inggris, sehingga kurang efektif untuk diterapkan dalam konteks bahasa Indonesia. Peneliti merancang sistem deteksi hoaks berbasis web dengan memanfaatkan kombinasi metode TF-IDF dan algoritma klasifikasi Logistic Regression.. Dataset berita asli dan palsu diambil dari Kaggle, kemudian diproses melalui tahap normalisasi, penghilangan stopwords, dan stemming. Hasil vektorisasi TF-IDF digunakan untuk melatih model klasifikasi biner. Sistem ini memungkinkan input berupa teks langsung maupun URL, serta memberikan hasil klasifikasi secara real-time. Evaluasi sistem menunjukkan akurasi tinggi dan potensi besar dalam meningkatkan literasi digital masyarakat. Hasil ini menunjukkan pendekatan ringan namun efektif untuk mitigasi penyebaran hoax berbahasa Indonesia..

**Kata kunci**—Deteksi hoax, Bahasa Indonesia, TF-IDF, *Logistic Regression*, Literasi Digital

### Abstract

*The spread of false information, particularly on social media platforms, has become a significant challenge due to the rapid flow of digital content. Most existing fake news detection systems are primarily designed for the English language, making them less effective when applied to Indonesian contexts. This study proposes a web-based hoax detection system that combines the Term Frequency-Inverse Document Frequency (TF-IDF) method with the Logistic Regression classification algorithm. The dataset, consisting of both real and fake news articles, was obtained from Kaggle and processed through several stages including normalization, stopword removal, and stemming. The TF-IDF vectorization results were then used to train a binary classification model. The system allows for user input either in the form of raw text or a URL, and delivers real-time classification results. Evaluation of the system indicates a high level of accuracy and strong potential in improving public digital literacy. These findings demonstrate a lightweight yet effective approach to mitigating the spread of hoaxes in the Indonesian language.*

**Keywords**—Hoax detection, Indonesian language, TF-IDF, *Logistic Regression*, digital literacy

## 1. PENDAHULUAN

Pertumbuhan teknologi digital yang sangat cepat telah membuka akses tak terbatas terhadap pertukaran informasi. Namun demikian, kemudahan ini turut memicu tingginya risiko penyebaran informasi yang tidak valid atau hoaks. Hoaks, yang dapat didefinisikan sebagai konten yang sengaja dibuat salah dengan tujuan menipu publik [1], telah menjadi tantangan signifikan di era digital. Fenomena yang sering disebut sebagai "infodemi" ini [1] tidak hanya terbatas pada teks, tetapi juga mencakup konten manipulatif canggih seperti deepfake [2]. Di Indonesia sendiri, media sosial menjadi saluran utama penyebaran hoaks, yang berimplikasi negatif terhadap opini publik, ketertiban sosial, dan aspek kesehatan. Platform media sosial telah memfasilitasi penyebaran berita palsu karena kemudahan akses, efektivitas biaya, serta kecepatan dalam berbagi informasi [3]. Fenomena ini sejalan dengan tren global di mana platform media sosial menjadi kanal dominan untuk penyebaran disinformasi [1]. Mengingat laju produksi konten digital yang masif setiap harinya, verifikasi secara manual oleh tenaga ahli menjadi tidak lagi praktis dan efisien. Kondisi ini mendorong urgensi pengembangan sebuah sistem yang mampu melakukan deteksi secara otomatis [4].

Meskipun banyak penelitian telah dilakukan dalam bidang deteksi hoax, sebagian besar masih fokus pada bahasa asing dan belum sepenuhnya efektif untuk berita berbahasa Indonesia.. Keterbatasan ini menjadi masalah umum, karena sebagian besar riset dan perangkat canggih cenderung berfokus pada bahasa Inggris, sehingga mengurangi reliabilitas sistem untuk bahasa lain seperti bahasa Spanyol dan, dalam konteks ini, bahasa Indonesia [5]. Keterbatasan ini sejalan dengan temuan dalam tinjauan literatur sistematis yang merekomendasikan adanya perluasan studi deteksi berita palsu untuk diterapkan pada berbagai bahasa selain Inggris [6]. Bahasa Indonesia dikategorikan sebagai *low-resource language*, sehingga ketersediaan data dan alat bantu untuk deteksi hoax masih terbatas

Penelitian ini mengusulkan solusi berbasis machine learning untuk mendeteksi berita palsu dalam bahasa Indonesia, menggunakan metode TF-IDF sebagai representasi fitur teks dan Logistic Regression untuk klasifikasi biner. Pendekatan ini dipilih karena *Logistic Regression* merupakan metode yang efektif dan telah teruji untuk tugas klasifikasi biner, terutama karena kemampuannya dalam memberikan estimasi probabilitas yang dapat diinterpretasi dan bekerja dengan baik pada data tidak seimbang maupun kejadian langka [7].

Pendekatan yang menggabungkan TF-IDF dengan Logistic Regression telah terbukti menjadi metode yang efektif dan berhasil mencapai akurasi tinggi dalam berbagai penelitian sejenis untuk mendeteksi berita palsu [8]. Pendekatan ini dipilih karena Logistic Regression merupakan metode standar yang telah teruji untuk tugas klasifikasi biner [9], sementara TF-IDF adalah teknik yang efektif untuk merepresentasikan fitur tekstual dalam tugas klasifikasi teks [10]. Diharapkan, melalui pendekatan ini, prediksi yang dihasilkan oleh sistem memiliki tingkat akurasi tinggi dan dapat diakses secara luas oleh publik.

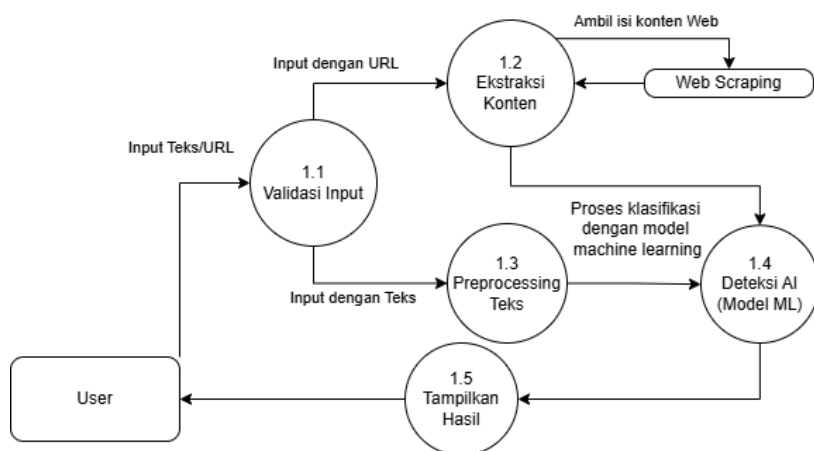
## 2. METODE PENELITIAN

Sebuah sistem deteksi berita palsu berbasis web dibangun dalam penelitian ini menggunakan pendekatan kuantitatif eksperimental. Sistem ini dirancang dengan menggabungkan metode vektorisasi Term Frequency-Inverse Document Frequency (TF-IDF) dan algoritma klasifikasi Logistic Regression. Tahapan penelitian mencakup perancangan arsitektur sistem, pengumpulan dan pra-pemrosesan dataset, pelatihan model, serta

implementasi pada antarmuka pengguna berbasis web. Agar model prediktif yang dihasilkan dapat diakses dan dimanfaatkan secara luas oleh publik, sistem ini diimplementasikan ke dalam sebuah server web. Pendekatan ini telah terbukti efektif untuk diseminasi hasil penelitian di berbagai bidang lain [11]. Data diperoleh dari platform Kaggle yang berisi kumpulan berita berbahasa Indonesia yang telah diberi label sebagai “asli” atau “palsu”. Pengujian sistem dilakukan untuk mengevaluasi akurasi prediksi model dalam membedakan antara berita asli dan berita palsu.

### 2.1 Arsitektur Sistem

Sistem terdiri dari frontend berbasis HTML, CSS, dan JavaScript, serta backend menggunakan Python Flask yang mengelola pemrosesan model dan API. Arsitektur ini memungkinkan input teks maupun URL artikel berita.



Gambar 1 Diagram Alur Data

Gambar 1 menjelaskan alur data pada sistem yang dirancang. Proses diawali ketika **User** (entitas eksternal) berinteraksi dengan antarmuka dan memberikan masukan.

- **Proses 1.1 Validasi Input:** Sistem akan memvalidasi masukan dari pengguna, yang dapat berupa teks berita langsung atau sebuah URL.
- **Proses 1.2 Ekstraksi Konten:** Jika masukan berupa URL, proses ini akan aktif. Dengan bantuan teknik Web Scraping, sistem mengambil konten teks dari URL target. Hasilnya adalah teks artikel yang siap diproses.
- **Proses 1.3 Preprocessing Teks:** Teks yang berasal dari input langsung maupun hasil ekstraksi URL akan melewati tahap pra-pemrosesan. Tahap ini meliputi normalisasi, pembersihan, dan stemming untuk menyiapkan teks sebelum dianalisis.
- **Proses 1.4 Deteksi AI (Model ML):** Teks yang telah bersih kemudian diumpankan ke model machine learning (Logistic Regression) yang sudah dilatih. Model akan melakukan klasifikasi untuk menentukan apakah berita tersebut terindikasi "Valid" atau "Hoax".

- **Proses 1.5 Tampilkan Hasil:** Hasil klasifikasi beserta tingkat keyakinan (probabilitas) kemudian dikirimkan kembali ke antarmuka untuk ditampilkan kepada User.

### 2.1.1 Teknologi dan Pustaka Pendukung

Sistem ini dibangun menggunakan bahasa pemrograman **Python** dengan dukungan dari beberapa pustaka (libraries) utama yang memiliki peran spesifik dalam alur kerja sistem:

- **Flask:** Digunakan sebagai kerangka kerja (web framework) untuk membangun sisi backend aplikasi, mengelola routing, dan menyediakan API untuk komunikasi dengan frontend.

- **Scikit-learn:** Pustaka inti untuk implementasi machine learning. Kelas *TfidfVectorizer* digunakan untuk ekstraksi fitur dan *Logistic Regression* untuk membangun dan melatih model klasifikasi.

- **Pandas & Numpy:** Digunakan untuk manipulasi data, seperti membaca dataset dari file CSV, mengelola struktur data dalam bentuk DataFrame, dan melakukan operasi numerik.

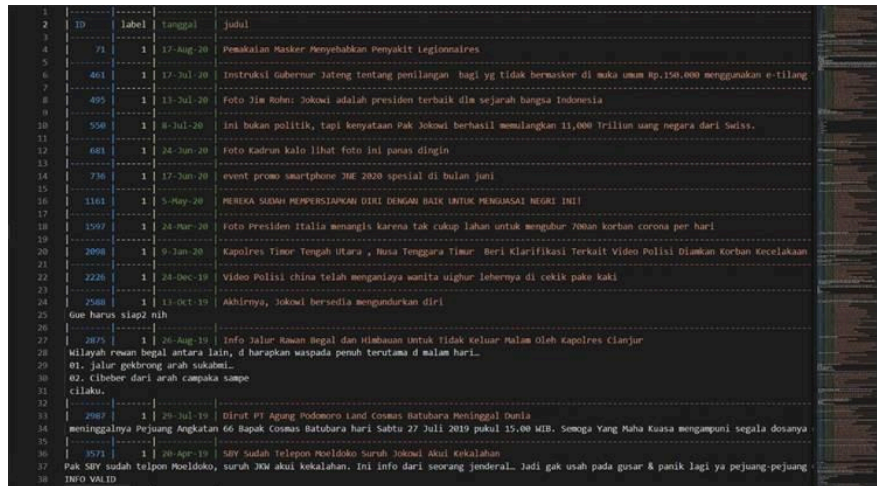
- **Sastrawi:** Pustaka khusus untuk pemrosesan teks Bahasa Indonesia, terutama digunakan untuk proses *stemming* (mengubah kata ke bentuk dasarnya) dan penghapusan *stopword*.

- **Newspaper3k & BeautifulSoup:** Kombinasi pustaka ini digunakan dalam fitur ekstraksi konten dari URL. *Newspaper3k* untuk mengekstrak artikel utama dan *BeautifulSoup* sebagai alternatif (fallback) untuk scraping konten dari struktur HTML.

- **Pickle:** Digunakan untuk serialisasi (menyimpan) model machine learning yang telah dilatih dan memuatnya kembali saat aplikasi dijalankan.

### 2.2 Dataset dan Preprocessing

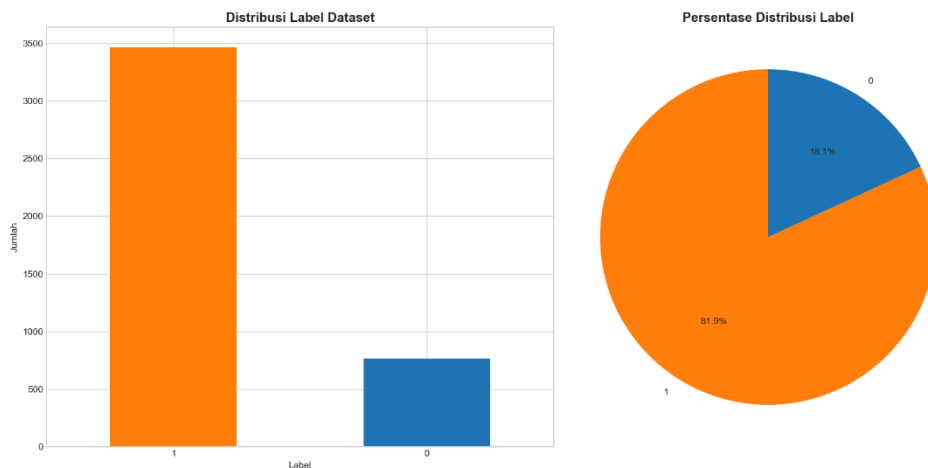
Data yang digunakan bersumber dari Kaggle, dengan dua label utama: berita asli dan hoaks. Sebelum digunakan dalam pelatihan model, data ini menjalani serangkaian tahap pra-proses, meliputi normalisasi huruf, penghapusan angka dan karakter khusus, stopword removal, dan stemming dengan pustaka Sastrawi. Sastrawi untuk stemming Bahasa Indonesia merupakan praktik yang umum divalidasi dalam berbagai penelitian analisis sentimen dan klasifikasi teks berbahasa Indonesia [12]. Tahapan ini merupakan praktik standar dalam Natural Language Processing (NLP) untuk membersihkan dan mempersiapkan data teks sebelum diolah oleh model machine learning [10].



ID	Label	Tanggal	Judul
71	1	17-Aug-20	Pemakaian Masker Menyebabkan Penyakit Legionnaires
461	1	17-Jul-20	Instruksi Gubernur Jateng tentang penilangan bagi yg tidak bermasker di muka umum Rp.150.000 menggunakan e-tilang
495	1	13-Jul-20	Foto Jim Rube: Jokowi adalah presiden terbaik dlm sejarah bangsa Indonesia
598	1	8-Jul-20	Ini bukan politik, tapi kenyataan Pak Jokowi berhasil memulangkan 11.000 Triliun uang negara dari Swiss.
681	1	24-Jun-20	Foto Kadruan kalo lihat foto ini panas dingin
798	1	17-Jun-20	event promo smartphone ZMI 2020 spesial di bulan juni
1161	1	5-May-20	PEREKA SEMAH MEMPERSIAPKAN DIRI DENGAN BAIK UNTUK MENGHASI NEGRI INI!
1597	1	28-Mar-20	Foto Presiden Italia semangis karena tak cukup lahan untuk mengubur 7000 korban corona per hari
2008	1	9-Jan-20	Kapolres Timor Tengah Utara, Nusa Tenggara Timur Beri Klarifikasi Terkait Video Polisi Dimakan Korban Kecelakaan
2226	1	28-Dec-19	Video Polisi china telah menganiaya wanita uighur lehernya di cekik pake kaki
2588	1	13-Oct-19	Ahlianya, Jokowi bersedia mengabdikan diri
2875	1	26-Aug-19	Info Jalur Raman Begal dan Hibunan untuk tidak keluar malam oleh Kapolres Cianjur
2982	1	29-Jul-19	Hilayah pemam begal antara lain, di harapkan waspada penuh terutama di malam hari.
3571	1	20-Apr-19	SBY Sudah telepon Poeldoko Suruh Jokowi Akui Kekalahan

Gambar 2 Dataset dengan format CSV

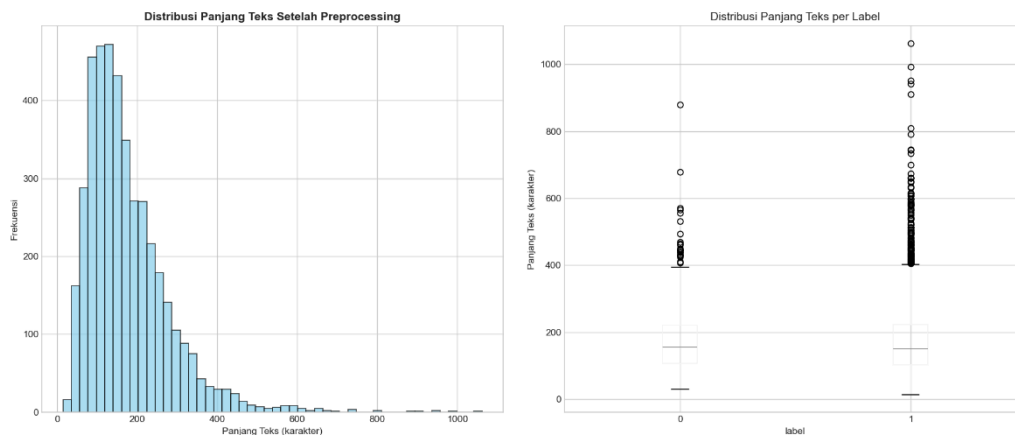
Dataset yang digunakan bersumber dari Kaggle, terdiri dari 4.231 artikel berita berbahasa Indonesia dengan label "Valid" (label 1) dan "Hoax" (label 0). Seperti yang divisualisasikan pada Gambar 3, terdapat ketidakseimbangan kelas (class imbalance) yang signifikan pada dataset ini. Dari total data, 81.9% merupakan berita valid (3.465 data), sedangkan hanya 18.1% merupakan berita hoax (766 data). Distribusi data yang tidak proporsional ini menjadi tantangan utama, sebab model machine learning yang dilatih pada data semacam ini akan memiliki kecenderungan untuk lebih memihak kelas mayoritas, yang berisiko menurunkan performanya jika tidak ada penanganan khusus. Masalah ketidakseimbangan kelas ini merupakan salah satu tantangan utama dalam deteksi berita palsu, di mana model seringkali kesulitan mengidentifikasi kelas minoritas dengan benar [13]. Oleh karena itu, diperlukan pendekatan khusus untuk menanganinya.



Gambar 3 Distribusi Label pada Dataset Penelitian

Setelah data melewati seluruh tahap preprocessing, dilakukan analisis eksplorasi untuk

memahami karakteristik panjang teks. Histogram pada Gambar 4 (kiri) menunjukkan bahwa sebagian besar artikel berita memiliki panjang antara 100 hingga 300 karakter setelah dibersihkan. Sementara itu, box plot (kanan) membandingkan distribusi panjang teks antara kedua label. Terlihat bahwa berita valid (label 1) memiliki rentang panjang teks yang sedikit lebih lebar dan lebih banyak outlier (artikel yang sangat panjang) dibandingkan dengan berita hoax (label 0). Perbedaan ini mengindikasikan bahwa panjang teks dapat menjadi salah satu fitur implisit yang membantu model dalam proses klasifikasi.



Gambar 4 Analisis Distribusi Panjang Teks Setelah Preprocessing

### 2. 2.1 Representasi Fitur dan Model

Setelah tahap prapemrosesan selesai, teks yang telah dibersihkan kemudian dikonversi ke dalam bentuk numerik menggunakan pendekatan *Term Frequency-Inverse Document Frequency* (TF-IDF). Teknik ini menetapkan bobot pada setiap kata dengan mempertimbangkan seberapa sering kata tersebut muncul dalam suatu dokumen dan seberapa jarang kemunculannya di keseluruhan kumpulan dokumen. Proses penghitungan bobot TF-IDF mengikuti rumus sebagai berikut:

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

di mana:

- $TF(t,d) = \frac{\text{jumlah kemunculan term } t \text{ dalam dokumen } d}{\text{Total term dalam dokumen } d}$
- $IDF(t) = \log \log \left( \frac{N}{df(t)} \right)$
- $N$  menyatakan jumlah keseluruhan dokumen dalam korpus.
- $df(t)$  merupakan jumlah dokumen yang memuat term  $t$ .

Dalam implementasi praktisnya, proses ekstraksi fitur dan pemodelan tidak dilakukan secara manual, melainkan memanfaatkan pustaka (library) machine learning yang telah teruji dalam Python. Vektorisasi *TF-IDF* diimplementasikan menggunakan kelas *TfidfVectorizer* dari pustaka *Scikit-learn*, yang secara otomatis menangani proses tokenisasi, penghitungan frekuensi, dan pembobotan IDF. Selanjutnya, model klasifikasi dibangun menggunakan kelas *Logistic Regression* yang juga berasal dari *Scikit-learn*, yang menangani proses pelatihan dan optimasi parameter secara efisien.

Pada tahap implementasi, seluruh dataset yang terdiri dari 4.231 entri dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian, dengan perbandingan 80:20. Pembagian ini menghasilkan sebanyak 3.384 data untuk pelatihan dan 847 data sisanya digunakan sebagai data pengujian. Proses vektorisasi TF-IDF diterapkan pada data latih dengan membatasi jumlah fitur yang digunakan sebanyak 5.000 kata yang paling sering muncul (*vocabulary size*). Hasilnya adalah sebuah matriks TF-IDF dengan dimensi (3384, 5000), yang kemudian digunakan untuk melatih model klasifikasi *Logistic Regression*.

Vektor hasil transformasi tersebut digunakan sebagai masukan bagi algoritma *Logistic Regression* dalam melakukan klasifikasi biner, yaitu antara berita valid dan hoaks. *Logistic Regression* merupakan salah satu metode yang sering diterapkan pada tugas klasifikasi karena sifatnya yang sederhana namun cukup efektif [9]. Meskipun tergolong sederhana, *Logistic Regression* tetap menjadi pilihan utama dalam klasifikasi biner karena mampu menghasilkan estimasi probabilitas yang dapat diinterpretasi dan bekerja efektif pada data tidak seimbang maupun kejadian langka [7].

Rumus dasar Logistic Regression:

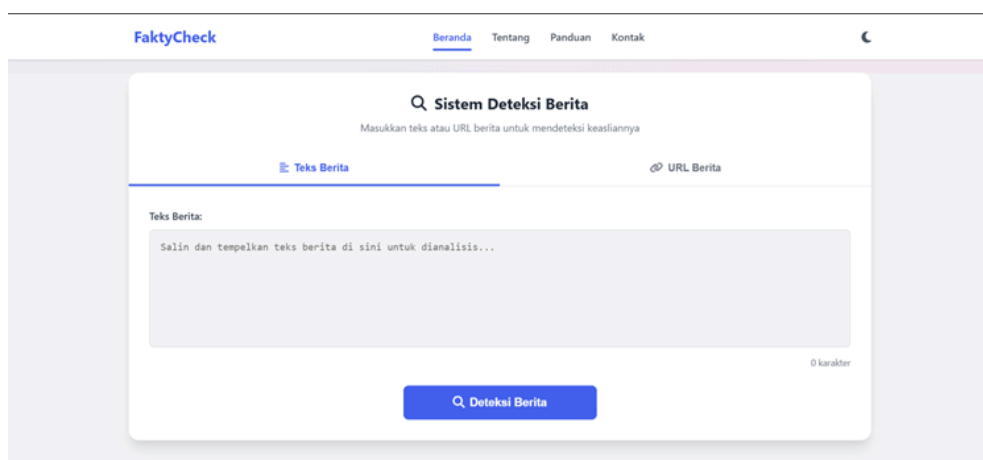
$$P(X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}} \quad (2)$$

Sebagian besar data, yaitu 80%, dimanfaatkan dalam tahap pelatihan model, sedangkan 20% lainnya digunakan sebagai data pengujian. Untuk menilai performa model, dilakukan evaluasi menggunakan metrik akurasi, presisi, dan recall.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Tampilan Antarmuka

Sistem memungkinkan pengguna memasukkan teks atau URL artikel. Jika menggunakan URL, sistem melakukan ekstraksi konten dengan pustaka BeautifulSoup.



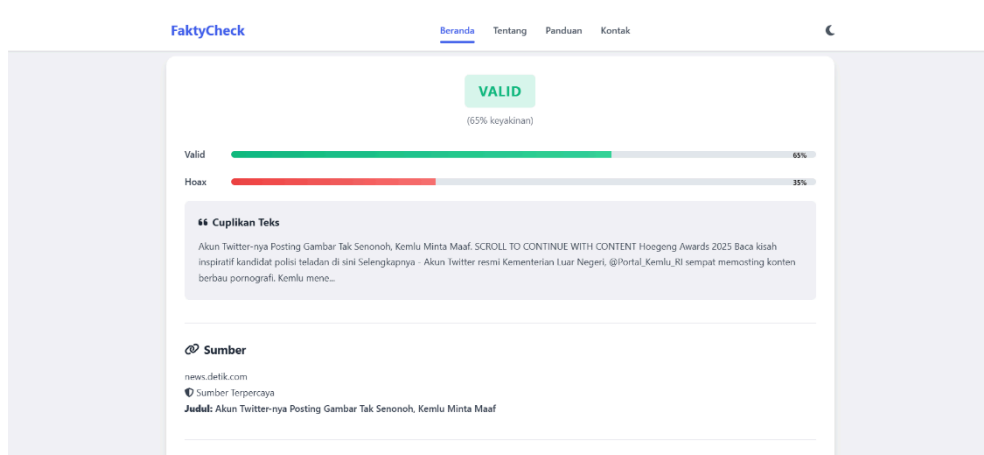
Gambar 5 Antarmuka input URL dan teks

Setelah input dikirimkan, sistem memberikan hasil klasifikasi beserta tingkat keyakinan dalam bentuk persentase dan visual bar chart.

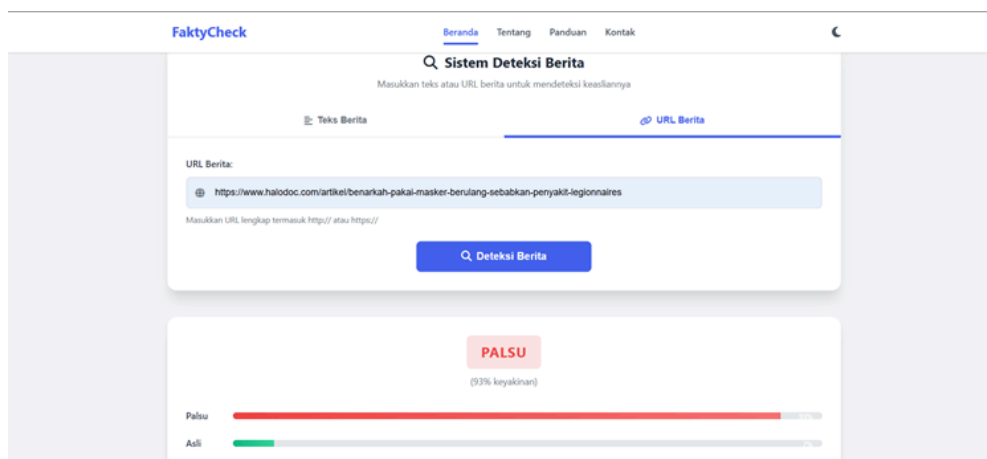
### 3.2 Evaluasi Berita Valid dan Hoax

Untuk menguji kemampuan sistem dalam mengidentifikasi berita yang faktual dan dapat dipercaya, dilakukan pengujian menggunakan sebuah artikel berita dari news.detik.com dengan judul "Akun Twitter-nya Posting Gambar Tak Senonoh, Kemlu Minta Maaf". Artikel ini melaporkan insiden nyata yang terjadi pada akun media sosial resmi milik Kementerian Luar Negeri RI.

Hasil analisis system menunjukan hasil valid dengan tingkat keyakinan sebesar 65%. Angka keyakinan 65% ini menunjukkan bahwa model berhasil mengenali artikel sebagai berita asli, meskipun nilainya tidak setinggi pada deteksi berita hoax. Fenomena ini kemungkinan dipengaruhi oleh sejumlah faktor, salah satunya adalah gaya bahasa yang digunakan dalam artikel. Misalnya, penggunaan ungkapan yang bersifat sensasional seperti “gambar tak senonoh” dapat memiliki pola yang serupa dengan narasi yang umum ditemukan dalam berita hoaks.



Gambar 6 Visualisasi untuk berita valid



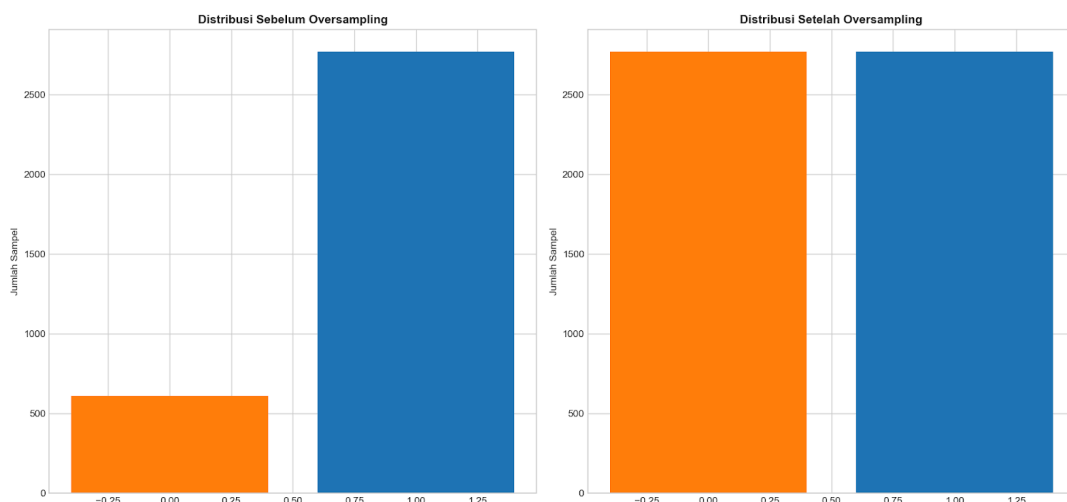
Gambar 7 Visualisasi untuk berita hoax

Hasil evaluasi terhadap berita yang diklasifikasikan sebagai hoax menunjukkan performa yang belum optimal. Hal ini disebabkan oleh ketidakseimbangan distribusi data dalam dataset, di mana jumlah data dengan label hoax hanya mencapai 766, sementara data valid mencapai 3.465. Ketimpangan ini menyebabkan model memiliki kecenderungan yang lebih kuat dalam mengenali berita valid dibandingkan berita hoax. Temuan ini menegaskan bahwa karakteristik dataset, seperti ketidakseimbangan kelas (*class imbalance*), secara signifikan memengaruhi performa model, sebuah pengamatan yang konsisten dengan hasil studi *benchmark* berskala besar yang membandingkan berbagai algoritma klasifikasi [9]. Selain itu, tidak semua berita yang dilabeli sebagai valid benar-benar memiliki isi yang akurat atau dapat dipercaya. Label valid tidak selalu berarti berita tentang suatu kejadian positif. Dalam konteks dataset ini, valid berarti informasi yang disajikan di dalam artikel tersebut akurat dan sesuai fakta. Dalam kasus ini, artikel tersebut adalah sebuah klarifikasi atau pengecekan fakta, sehingga berpotensi menurunkan akurasi klasifikasi terhadap berita asli dalam proses pelatihan maupun pengujian model.

### 3.3 Perbandingan Kinerja Model dengan dan Tanpa Oversampling

Untuk mengukur dampak dari penerapan teknik *oversampling* secara kuantitatif, dilakukan evaluasi kinerja model *Logistic Regression* pada data uji. Kinerja model dievaluasi dalam dua skenario:

- sebelum penerapan *oversampling* (model dilatih pada data tidak seimbang)
- sesudah penerapan *oversampling* (model dilatih pada data seimbang). Metrik yang digunakan meliputi akurasi, presisi, *recall*, F1-Score, dan AUC-ROC.



Gambar 7 Proses Penyeimbangan Kelas Menggunakan Oversampling.

Hasil perbandingan kinerja model dirangkum dalam Tabel 1 berikut:

Tabel 1 Hasil Evaluasi Kinerja Model

Tanpa Oversampling:	Dengan Oversampling:
AKURASI: 0.8323 (83.23%)	AKURASI: 0.8170 (81.70%)
PRESISI: 0.8149 (81.49%)	PRESISI: 0.8129 (81.29%)
RECALL: 0.8323 (83.23%)	RECALL: 0.8170 (81.70%)
F1-SCORE: 0.7791 (77.91%)	F1-SCORE: 0.8148 (81.48%)
AUC-ROC: 0.7662 (76.62%)	AUC-ROC: 0.7660 (76.60%)

Dari tabel di atas, dapat ditarik beberapa pengamatan penting:

- Penurunan Akurasi yang Wajar: Terjadi sedikit penurunan akurasi dari 83.23% menjadi 81.70% setelah oversampling. Hal ini adalah fenomena yang wajar dan dapat diperkirakan. Akurasi pada model tanpa oversampling cenderung terlihat lebih tinggi karena bias model terhadap kelas mayoritas (berita valid, 81.9% dari data). Model tersebut bisa mencapai akurasi tinggi hanya dengan lebih sering menebak kelas mayoritas, namun gagal mengenali kelas minoritas (berita hoax) secara efektif. Fenomena ini menggarisbawahi mengapa metrik akurasi saja tidak cukup untuk mengevaluasi kinerja model pada data tidak seimbang; metrik seperti presisi dan recall menjadi lebih penting untuk memberikan gambaran yang lengkap [13].

- Peningkatan Signifikan pada F1-Score: Metrik yang paling krusial dalam kasus data tidak seimbang ini adalah F1-Score. Terjadi peningkatan yang signifikan pada F1-Score dari 77.91% menjadi 81.48%. F1-Score merupakan rata-rata harmonis dari presisi dan recall, yang digunakan untuk menilai keseimbangan kinerja model. Peningkatan nilai ini mengindikasikan bahwa model yang dilatih menggunakan data yang seimbang memiliki kemampuan yang lebih baik dalam mengidentifikasi berita valid maupun hoaks secara proporsional. Dengan demikian, model menjadi lebih stabil dan minim bias.

- Stabilitas Metrik Lainnya: Nilai presisi, recall, dan AUC-ROC relatif stabil dan tidak menunjukkan perubahan drastis. Ini mengindikasikan bahwa kemampuan dasar model untuk membedakan kelas tidak banyak berubah, namun oversampling berhasil menyesuaikan ambang batas klasifikasi sehingga menghasilkan performa yang lebih seimbang secara keseluruhan.

Pendekatan *oversampling* dipilih karena terbukti lebih baik dibandingkan *undersampling* sebab tidak ada kehilangan informasi dari data mayoritas yang mungkin relevan [13]. Metode *random oversampling* terbukti cukup efektif dalam mengatasi ketidakseimbangan data karena mampu menambah representasi kelas minoritas tanpa memerlukan modifikasi

struktur data yang kompleks, sehingga sering kali sudah cukup untuk meningkatkan performa model klasifikasi [14].

### 3.4 Kriteria Evaluasi Berita Valid dan Hoax

Berdasarkan dokumentasi resmi dari dataset *Indonesia False News Dataset for Text Classification in Bahasa Indonesia* yang disusun oleh Muhammad Ghazi Muharam dan tersedia di platform Kaggle, tidak dijelaskan secara rinci mengenai kriteria atau metode evaluasi yang digunakan dalam menentukan label “hoaks” (False) maupun “valid” (Valid) pada setiap data. Akibatnya, tidak ada penjelasan eksplisit yang merinci dasar kriteria untuk mengklasifikasikan sebuah berita sebagai hoaks. Absennya justifikasi pelabelan ini menjadi tantangan tersendiri, karena idealnya sebuah sistem deteksi tidak hanya memberikan label “**hoaks**” atau “**valid**”, tetapi juga penjelasan yang dapat dipahami manusia (explainability) untuk membangun kepercayaan [15]. Tidak disebutkan apakah pelabelan dilakukan berdasarkan:

- Verifikasi manual oleh tim ahli
- Acuan dari lembaga pemeriksa fakta seperti **TurnBackHoax.id** atau **Mafindo**
- Penggunaan algoritma otomatis
- Referensi terhadap sumber berita kredibel

Akibat dari tidak adanya dokumentasi tersebut, validitas label (khususnya kategori hoaks) tidak dapat diverifikasi secara independen, dan berpotensi menimbulkan ambiguitas dalam interpretasi hasil klasifikasi model. Meski begitu, dataset ini tetap banyak digunakan sebagai benchmark awal dalam penelitian klasifikasi berita palsu dalam Bahasa Indonesia, karena sudah menyediakan struktur label biner yang siap dipakai untuk eksperimen klasifikasi teks.

## 4. KESIMPULAN

Studi ini berhasil merancang dan membangun sebuah sistem deteksi berita palsu dalam bahasa Indonesia berbasis web, dengan memanfaatkan kombinasi pendekatan *TF-IDF* sebagai ekstraksi fitur dan algoritma *Logistic Regression* untuk proses klasifikasinya. Sistem ini mampu memproses input baik dalam bentuk teks langsung maupun URL artikel, dan memberikan hasil klasifikasi secara real-time dengan visualisasi yang membantu pengguna dalam memahami keaslian informasi.

Evaluasi menunjukkan bahwa penanganan data tidak seimbang menggunakan teknik oversampling secara signifikan meningkatkan kinerja model, yang dibuktikan dengan peningkatan F1-Score dari 77.91% menjadi 81.48%. Hasil ini menegaskan bahwa model yang dilatih pada data seimbang lebih andal dalam membedakan berita valid dan hoax. Dengan demikian, sistem ini memiliki potensi besar untuk digunakan sebagai alat bantu masyarakat dalam mengidentifikasi berita hoax secara mandiri dan efisien.

## 5. SARAN

Salah satu kelemahan dalam penelitian ini berkaitan dengan sumber data yang digunakan. Dataset tersebut tidak mencantumkan secara jelas keterangan mengenai kriteria penilaian atau tahapan verifikasi yang digunakan dalam menetapkan label “Valid” maupun

“False” pada data. Tidak diketahui apakah pelabelan dilakukan melalui referensi lembaga pemeriksa fakta seperti TurnBackHoax.id atau melalui metode manual lainnya. Hal ini dapat memengaruhi validitas label yang dijadikan acuan dalam pelatihan dan pengujian model. Dengan demikian, pada penelitian berikutnya disarankan untuk menggunakan dataset yang menyediakan dokumentasi pelabelan secara lebih terbuka dan jelas, atau melakukan proses validasi ulang terhadap data yang digunakan guna memastikan keakuratan label.

Pengembangan lanjutan dari penelitian ini dapat dilakukan dengan menghubungkan sistem ke database pengecekan fakta nasional, menambahkan kemampuan deteksi judul clickbait, dan menyertakan analisis sentimen terhadap konten berita. Selain itu, eksplorasi terhadap metode klasifikasi yang lebih kompleks seperti *ensemble learning* dan *deep learning* dapat dilakukan untuk meningkatkan akurasi sistem dalam menangani kasus-kasus yang lebih beragam dan kompleks. Rekomendasi ini didukung oleh berbagai tinjauan literatur yang menyarankan penggunaan model hybrid, ensemble, dan deep learning sebagai arah penelitian masa depan untuk mengatasi kompleksitas berita palsu yang terus berkembang [6].

## 6. UCAPAN TERIMA KASIH

Dengan segala kerendahan hati, penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, baik secara moral maupun teknis, selama proses penyusunan jurnal ini. Penulis menyampaikan apresiasi yang tulus kepada **Muhammad Ghazi Muharam**, selaku penyedia *Indonesia False News (Hoax) Dataset* melalui platform Kaggle, yang sangat membantu dalam mendukung penelitian ini.

Penulis menyampaikan apresiasi yang setinggi-tingginya kepada Bapak Assoc. Prof. Dr. Abdul Hamid Arribathi, S.Ag., M.M. atas segala bentuk bimbingan, dukungan, serta dorongan semangat yang telah diberikan selama proses penyusunan jurnal ini. Kontribusi beliau, baik dari segi ilmu, waktu, maupun perhatian, memiliki peran yang sangat penting dalam mendukung kelancaran dan penyelesaian karya ini secara optimal.

## DAFTAR PUSTAKA

- [1] B. León, M. P. Martínez-Costa, R. Salaverría, and I. López-Goñi, “Health and science-related disinformation on COVID-19: A content analysis of hoaxes identified by fact-checkers in Spain,” *PLoS One*, vol. 17, no. 4 April, Apr. 2022, doi: 10.1371/journal.pone.0265995.
- [2] R. Rafique, R. Gantassi, R. Amin, J. Frnda, A. Mustapha, and A. H. Alshehri, “Deep fake detection and classification using error-level analysis and deep learning,” *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-34629-3.
- [3] S. Lohitha, S. D. Reddy, B. R. Krishna, and N. S. Krishna, “Fake News Detection Using Machine Learning,” *Lect. Notes Networks Syst.*, vol. 645 LNNS, pp. 463–470, 2023, doi: 10.1007/978-981-99-0769-4\_41.
- [4] S. Shahsavari, P. Holur, T. Wang, T. R. Tangherlini, and V. Roychowdhury, “Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news,” *J. Comput. Soc. Sci.*, vol. 3, no. 2, pp. 279–317, Nov. 2020, doi: 10.1007/s42001-020-00086-5.
- [5] J. A. García-Díaz, S. M. Jiménez-Zafra, M. A. García-Cumbreras, and R. Valencia-García, “Evaluating feature combination strategies for hate-speech detection in Spanish using linguistic features and transformers,” *Complex Intell. Syst.*, vol. 9, no. 3,

- pp. 2893–2914, Jun. 2023, doi: 10.1007/s40747-022-00693-x.
- [6] H. F. Villela, F. Corrêa, J. S. de A. N. Ribeiro, A. Rabelo, and D. B. F. Carvalho, “Fake news detection: a systematic literature review of machine learning algorithms and datasets,” *J. Interact. Syst.*, vol. 14, no. 1, pp. 47–58, 2023, doi: 10.5753/jis.2023.3020.
- [7] M. Maalouf, “Logistic regression in data analysis: An overview,” *Int. J. Data Anal. Tech. Strateg.*, vol. 3, no. 3, pp. 281–299, 2011, doi: 10.1504/IJDATS.2011.041335.
- [8] J. Adeleke Adeyiga, P. Gbounmi Toriola, T. Elizabeth Abioye, and A. Esther Oluwatosin, “Fake News Detection Using a Logistic Regression Model and Natural Language Processing Techniques,” pp. 1–18, 2023, [Online]. Available: <https://doi.org/10.21203/rs.3.rs-3156168/v1>
- [9] R. Couronné, P. Probst, and A. L. Boulesteix, “Random forest versus logistic regression: A large-scale benchmark experiment,” *BMC Bioinformatics*, vol. 19, no. 1, Jul. 2018, doi: 10.1186/s12859-018-2264-5.
- [10] A. Shinde, E. Q. Shahra, S. Basurra, F. Saeed, A. A. AlSewari, and W. A. Jabbar, “SMS Scam Detection Application Based on Optical Character Recognition for Image Data Using Unsupervised and Deep Semi-Supervised Learning,” *Sensors*, vol. 24, no. 18, p. 6084, Sep. 2024, doi: 10.3390/s24186084.
- [11] I. Dimitrov, D. R. Flower, and I. Doytchinova, “AllerTOP - a server for in silico prediction of allergens,” *BMC Bioinformatics*, vol. 14, no. SUPPL6, Apr. 2013, doi: 10.1186/1471-2105-14-S6-S4.
- [12] R. Kusumaningrum, I. Z. Nisa, R. P. Nawangsari, and A. Wibowo, “Sentiment analysis of Indonesian hotel reviews: from classical machine learning to deep learning,” *Int. J. Adv. Intell. Informatics*, vol. 7, no. 3, pp. 292–303, 2021, doi: 10.26555/ijain.v7i3.737.
- [13] I. Y. Agarwal and D. P. Rana, “Fake news and imbalanced data perspective,” *Data Preprocessing, Act. Learn. Cost Perceptive Approaches Resolv. Data Imbalance*, no. June, pp. 195–210, 2021, doi: 10.4018/978-1-7998-7371-6.ch011.
- [14] R. Blagus and L. Lusa, “SMOTE for high-dimensional class-imbalanced data,” *BMC Bioinformatics*, vol. 14, 2013, doi: 10.1186/1471-2105-14-106.
- [15] B. Wang *et al.*, “Explainable Fake News Detection with Large Language Model via Defense Among Competing Wisdom,” *WWW 2024 - Proc. ACM Web Conf.*, pp. 2452–2463, 2024, doi: 10.1145/3589334.3645471.