

Klasifikasi Berita Palsu Pada Pemilihan Umum Presiden 2024 Menggunakan Metode K-Nearest Neighbor

Arief Fathul Hikam¹, Triana Harmini², Aziz Mustafa³, Faisal Reza Pradhana⁴

^{1,2,3,4}Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Darussalam Gontor

E-mail: ^{*1}arieffathulhikam@mhs.unida.gontor.ac.id, ²triana@unida.gontor.ac.id,

³aziz@unida.gontor.ac.id, ⁴faisalrezapradhana@unida.gontor.ac.id

Abstrak – Pemilihan Umum Presiden 2024 menjadi salah satu momen krusial dalam demokrasi Indonesia. Namun, penyebaran berita palsu (hoaks) semakin marak dan berdampak negatif terhadap opini publik serta proses pemilu yang sehat. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi berita palsu menggunakan metode K-Nearest Neighbor (KNN). Data yang digunakan berasal dari situs Turn Back Hoax oleh MAFINDO (Masyarakat Anti Fitnah Indonesia) yang menyediakan sumber berita terverifikasi benar dan telah melabeli berita salah yang beredar di masyarakat. Proses klasifikasi melibatkan tahap prapemrosesan data, seperti tokenisasi, penghapusan kata tidak bermakna (stopwords), dan stemming. Selanjutnya, fitur data diekstraksi menggunakan metode TF-IDF untuk meningkatkan kualitas klasifikasi. Berdasarkan pengujian yang dilakukan, model KNN berhasil mencapai akurasi sebesar 97%, yang menunjukkan efektivitas metode ini dalam mendeteksi berita palsu. Hasil penelitian ini diharapkan dapat berkontribusi dalam memberikan solusi praktis untuk memitigasi penyebaran hoaks, khususnya dalam konteks pemilu, serta mendorong partisipasi pemilih yang lebih terinformasi.

Kata Kunci — Klasifikasi, Berita palsu, pemilu, K-Nearest Neighbor

1. PENDAHULUAN

Pemilihan Umum Presiden 2024 merupakan salah satu momen penting dalam kehidupan demokrasi Indonesia. Dalam proses pemilu, informasi yang akurat dan terpercaya memegang peranan kunci untuk membangun opini publik yang sehat serta mendukung pengambilan keputusan pemilih secara rasional. Namun, fenomena penyebaran berita palsu (hoaks) semakin marak dan menjadi tantangan serius, terutama dengan meningkatnya penggunaan media sosial sebagai sumber utama informasi masyarakat. Hoaks adalah usaha untuk memengaruhi audiens dengan cara memanipulasi mereka agar menerima opini tertentu. Salah satu contohnya adalah menyebarkan informasi yang menggambarkan suatu peristiwa dengan cara yang tidak sesuai dengan fakta di lapangan [1]. Berita palsu dapat menyesatkan masyarakat, memengaruhi persepsi publik, dan menciptakan polarisasi di tengah masyarakat [2]. Berita palsu dapat merusak citra dan kredibilitas serta mempengaruhi banyak orang [3]. penyebaran hoaks (berita bohong) yang dapat menyesatkan dan mempengaruhi opini publik [4].

Dalam konteks pemilu, penyebaran berita palsu tidak hanya merugikan pihak tertentu, tetapi juga dapat mengancam integritas proses demokrasi. Berita palsu sering kali dirancang untuk memanipulasi opini publik, mengaburkan fakta, dan memicu konflik di masyarakat [5]. Oleh karena itu, diperlukan solusi yang efektif untuk mengidentifikasi dan menangani penyebaran berita palsu, terutama melalui pendekatan teknologi yang berbasis kecerdasan buatan.

Salah satu metode yang dapat digunakan untuk klasifikasi berita palsu adalah K-Nearest Neighbor (KNN). Metode KNN dikenal sebagai algoritma pembelajaran mesin yang sederhana namun efektif dalam melakukan klasifikasi berdasarkan kedekatan data dalam ruang multidimensi [6]. K-Nearest Neighbor (KNN) adalah algoritma supervised learning yang menentukan hasil klasifikasi data uji baru berdasarkan kategori mayoritas dari K tetangga terdekat [7]. KNN memiliki kelebihan dalam mengeksekusi data dengan waktu yang relatif singkat, sehingga sangat efektif untuk menganalisis penyebaran informasi yang tidak jelas isinya, termasuk berita yang mengandung unsur hoaks, secara cepat [8]. Penggunaan KNN dalam klasifikasi berita palsu telah menunjukkan hasil yang menjanjikan karena algoritma ini mampu mengklasifikasikan data berdasarkan pola distribusi fitur dengan tingkat akurasi yang tinggi.

Penelitian ini ingin menerapkan metode KNN untuk mendeteksi berita palsu yang terkait dengan Pemilihan Umum Presiden 2024. Dengan menggunakan pendekatan ini, diharapkan dapat diperoleh model klasifikasi yang akurat dan dapat digunakan sebagai alat bantu dalam mitigasi penyebaran berita palsu, sehingga mendukung pelaksanaan pemilu yang lebih transparan dan terpercaya

2. METODE PENELITIAN

Penelitian ini dilakukan untuk mengembangkan model klasifikasi berita palsu menggunakan metode K-Nearest Neighbor (KNN) dalam konteks Pemilihan Umum Presiden 2024. Metode penelitian yang digunakan terdiri atas beberapa tahapan utama, yaitu pengumpulan data, prapemrosesan data, ekstraksi fitur, pembangunan model, dan evaluasi kinerja model. Berikut adalah uraian lengkap tiap tahapan:

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri atas berita-berita yang relevan dengan Pemilihan Umum Presiden 2024. Berita ini diperoleh dari berbagai sumber, seperti media sosial, portal berita daring, dan situs web resmi. Dataset terdiri dari berita yang telah diberi label sebagai berita palsu (fake news) dan berita asli (real news) berdasarkan verifikasi manual oleh pakar atau data dari sumber yang terpercaya.

2.2 Prapemrosesan Data

Tahap ini bertujuan untuk membersihkan dan mempersiapkan data agar siap digunakan dalam analisis. Langkah-langkah yang dilakukan meliputi:

- Tokenisasi merupakan proses memecah teks menjadi unit-unit kecil yang disebut token, yang biasanya berupa kata, frasa, atau simbol. Tujuan proses ini untuk mengubah teks menjadi bentuk yang dapat diproses oleh algoritma pembelajaran mesin. Contohnya dalam teks asli berbunyi "Pemilu Presiden 2024 sangat penting", maka hasil tokenisasi menjadi ["Pemilu", "Presiden", "2024", "sangat", "penting"]. Tokenisasi biasanya dilakukan berdasarkan spasi atau tanda baca. Dalam bahasa Indonesia, proses ini bisa lebih kompleks jika ada gabungan kata atau istilah tertentu (misalnya, "bukan tidak mungkin" dianggap sebagai satu konsep).
- Penghapusan Stopwords merupakan proses menghapus kata-kata umum yang sering muncul tetapi tidak memiliki kontribusi signifikan terhadap makna atau konteks teks. Dengan tujuan: mengurangi dimensi data dan meningkatkan efisiensi analisis. Stopwords umum: kata-kata seperti "dan", "atau", "adalah", "yang", "di", dan sebagainya. Contohnya dalam teks hasil tokenisasi: ["Pemilu", "Presiden", "2024", "sangat", "penting"]. Setelah penghapusan stopwords: ["Pemilu", "Presiden", "2024", "penting"]. Pendekatan: penghapusan dilakukan dengan menggunakan daftar stopwords yang telah disiapkan sebelumnya untuk bahasa Indonesia.
- Stemming merupakan proses mengubah kata-kata ke bentuk dasarnya (kata dasar) tanpa memperhatikan imbuhan atau bentuk turunan. Tujuan proses ini untuk menyederhanakan analisis teks dengan mengelompokkan kata yang memiliki akar yang sama. Contohnya dalam teks asli berbunyi ["berlari", "lari-lari", "berlari-lari"]. Maka hasil stemming menjadi ["lari", "lari", "lari"]. Stemming dalam bahasa Indonesia, menggunakan algoritma populer seperti Nazief-Adriani Stemming Algorithm sering digunakan untuk menangani berbagai pola imbuhan (me-, ber-, -kan, -i, dll.).
- Normalisasi Teks merupakan proses membersihkan teks dengan menghapus elemen-elemen yang tidak relevan, seperti tanda baca, angka, dan simbol khusus. Tujuan proses ini untuk menghilangkan elemen yang tidak memiliki nilai informasi untuk analisis atau klasifikasi. Contohnya dalam teks asli berbunyi "Pemilu Presiden 2024!!! #Demokrasi". Maka hasil normalisasi menjadi : "Pemilu Presiden". Langkah -langkah Normalisasi: Menghapus tanda baca seperti !, ?, dan .. Menghapus angka yang tidak relevan, misalnya tahun atau nilai tertentu yang tidak penting. Menangani slang atau singkatan dengan mengonversinya ke bentuk standar (misalnya, "gak" menjadi "tidak"). Dalam beberapa kasus, normalisasi juga mencakup konversi teks menjadi huruf kecil agar lebih konsisten.

Prapemrosesan ini sangat penting untuk memastikan data yang digunakan dalam analisis atau model pembelajaran mesin bersih, terstruktur, dan relevan untuk menghasilkan hasil yang lebih akurat dan efektif.

2.3 Ekstraksi Fitur

Setelah data diproses, fitur teks diekstraksi menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini menghitung bobot setiap kata berdasarkan frekuensi kemunculannya di dokumen dan seluruh dataset, sehingga dapat menangkap informasi yang signifikan untuk klasifikasi.

2.4 Pembangunan Model KNN

Pembangunan model KNN melalui tiga tahap yaitu pemilihan parameter K, menghitung jarak antara data uji dan data Latihan, dan menentukan kelas data uji.

a. pemilihan parameter K

Parameter K dalam algoritma K-Nearest Neighbor (KNN) adalah nilai yang menentukan jumlah tetangga terdekat yang akan digunakan untuk memutuskan kelas suatu data uji. Pemilihan nilai yang tepat sangat penting untuk memastikan bahwa model dapat memberikan hasil klasifikasi yang paling akurat. Proses pemilihan nilai K yang optimal dilakukan melalui pengujian dengan berbagai nilai K, seperti 1, 3, 5, 7, dan seterusnya. Untuk setiap nilai K, performa model dievaluasi menggunakan metrik tertentu, seperti akurasi atau teknik cross-validation, guna menilai seberapa baik model dapat menggeneralisasi data baru. Setelah evaluasi, nilai K yang menghasilkan performa terbaik akan dipilih sebagai parameter akhir. Dengan memilih

nilai K yang optimal, algoritma KNN dapat mencapai keseimbangan antara sensitivitas terhadap data dan kemampuan untuk mengklasifikasikan data secara akurat.

- b. menghitung jarak antara data uji dan data Latihan

Menghitung jarak antara data uji dan data pelatihan dalam algoritma K-Nearest Neighbor (KNN) adalah langkah penting untuk menentukan tetangga terdekat. Jarak ini dihitung menggunakan metrik seperti jarak Euclidean [9]

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots(1)$$

di mana x_i dan y_i adalah nilai fitur data uji dan pelatihan.

Setelah jarak dihitung, data pelatihan dengan jarak terkecil diidentifikasi sebagai tetangga terdekat. Kelas data uji ditentukan berdasarkan kategori mayoritas dari tetangga ini, memungkinkan KNN memanfaatkan pola lokal untuk klasifikasi.

- c. menentukan kelas data uji

Menentukan kelas data uji dalam algoritma K-Nearest Neighbor (KNN) dilakukan berdasarkan kategori mayoritas dari tetangga terdekat yang telah diidentifikasi. Setelah jarak antara data uji dan semua data pelatihan dihitung, sejumlah KKK tetangga terdekat dipilih berdasarkan jarak terkecil. Setiap tetangga memiliki label atau kelas yang telah diketahui dari data pelatihan. Kelas data uji ditentukan dengan menghitung frekuensi kemunculan setiap kelas di antara tetangga tersebut. Kelas yang paling sering muncul di antara KKK tetangga ini dianggap sebagai kelas data uji. Proses ini memungkinkan KNN untuk memanfaatkan pola lokal dalam dataset untuk membuat keputusan klasifikasi yang sederhana namun efektif.

2.5 Evaluasi Kinerja Model

Evaluasi model yang dibangun dalam algoritma K-Nearest Neighbor (KNN) bertujuan untuk mengukur seberapa baik model mampu melakukan klasifikasi pada data baru. Metrik yang digunakan mencakup:

- Akurasi: Mengukur persentase prediksi yang benar dari total prediksi, menunjukkan kinerja keseluruhan model.
- Presisi: Mengukur proporsi prediksi positif yang benar-benar positif, berguna dalam kasus di mana salah prediksi positif memiliki konsekuensi serius.
- Recall (Sensitivitas): Mengukur kemampuan model untuk mendeteksi semua data positif yang sebenarnya, penting untuk menghindari kelalaian dalam klasifikasi positif.
- F1-Score: Kombinasi presisi dan recall dalam satu metrik harmonis, yang memberikan gambaran seimbang, terutama ketika terdapat ketidakseimbangan data.

Pengujian model dilakukan menggunakan teknik cross-validation, di mana dataset dibagi menjadi beberapa subset (lipatan). Model dilatih pada sebagian subset dan diuji pada subset lainnya secara bergantian hingga semua subset menjadi bagian data uji. Teknik ini memastikan hasil evaluasi lebih andal karena semua data diuji, dan hasil akhir merupakan rata-rata performa dari semua iterasi. Selain itu, cross-validation membantu mengurangi risiko overfitting, yaitu ketika model terlalu cocok dengan data pelatihan tetapi gagal menggeneralisasi data baru. Dengan evaluasi yang menyeluruh ini, model dapat dinilai secara objektif dan dioptimalkan untuk performa terbaik.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada tahapan ini data yang sudah dikumpulkan dipersiapkan dan dibersihkan untuk pengolahan data. Dataset yang didapatkan dari website turnbackhoax.id dengan jumlah total 2076 berita. Data yang berhasil dikumpulkan kemudian dibersihkan dengan menghapus duplikat, dan membuang data yang tidak berguna untuk penelitian secara manual. Kemudian, proses labeling divalidasi oleh dosen pakar dari Ilmu Komunikasi agar tidak terjadi kesalahan dalam proses label. Hasil dari pelabelan disajikan pada Tabel 1.

Tabel 1. Hasil Pelabelan Dataset

Label	Jumlah
Positif	1038
Negatif	1038

3.2 Prapemrosesan Data

Tahap *Case-folding* mengubah semua huruf disetiap kalimat pada data menjadi huruf kecil. Contoh dari *case-folding* seperti “Pemilihan” setelah diproses menjadi “pemilihan”. Hasil dari Prapemrosesan Data disajikan dalam Tabel 2.

Tabel 2 Case Folding

Content	Case Folding
Pemilihan Presiden 2024 hampir pasti akan diikuti oleh tiga pasangan calon.	pemilihan presiden hampir pasti akan diikuti oleh tiga pasangan calon

3.2.a Tokenisasi

Tahap Tokenisasi merubah teks menjadi unit-unit kecil yang disebut token. Hasil dari Tokenisasi disajikan dalam Tabel 3

Tabel 3. Tokenisasi

Content	Tokenisasi
pemilihan presiden 2024 hampir pasti akan diikuti oleh tiga pasangan calon.	[pemilihan, presiden, hampir, pasti, akan, diikuti, oleh, tiga, pasangan, calon]

3.2.b Penghapusan Stopwords

Tahap *stopwords* removal menghilangkan kata-kata yang tidak penting atau tidak memiliki arti. Pada tahapan ini data teks disaring dari kata-kata yang kurang berguna pada proses klasifikasi program menggunakan library stopword bahasa indonesia yang ada pada library NLTK (Natural Language Tool Kit). Hasil dari penghapusan *Stopwords* disajikan dalam table 4

Tabel 4. Stopwors

Content	Stopwords
Pemilihan Presiden 2024 hampir pasti akan diikuti oleh tiga pasangan calon.	pemilihan presiden hampir pasti akan diikuti oleh tiga pasangan calon

3.2.c Stemming

Tahap *stemming* mengubah kata-kata yang telah diproses menjadi kata-kata dasar. Contoh dari *stemming* seperti kata “pemilihan” menjadi “pilih”. Pada tahapan ini setiap kata dalam data teks diubah ke bentuk dasar dari kata tersebut menggunakan library sastrawi. Setelah proses stemming didapatkan hasil dari proses stemming pada tabel 5.

Tabel 5. Stemming

Content	Stemming
Pemilihan Presiden 2024 hampir pasti akan diikuti oleh tiga pasangan calon.	pilih presiden hampir pasti akan ikut oleh tiga pasang calon

3.2.d Normalisasi Teks

Tahap normalisasi mengubah semua kata-kata dalam komentar menjadi sesuai dengan aturan bahasa Indonesia. Pada tahap ini data teks dilakukan proses normalisasi teks dengan mengubah kata singkatan, asing atau yang belum jelas ke kata yang lebih jelas/baku dengan data kata normalisasi yang telah disiapkan didalam dataset normalisasi teks. Hasil dari Normalisasi Teks disajikan dalam Tabel 6.

Tabel 6. Normalisasi

Content	Normalisasi
Pemilihan Presiden 2024 hampir pasti akan diikuti oleh tiga pasangan calon.	pemilihan presiden hampir pasti akan diikuti oleh tiga pasangan calon

3.3 Pembangunan Model KNN

Pada tahapan ini dilakukan pembangunan model sistem yang telah dirancang untuk mengklasifikasi berita benar dan berita salah pada masa pemilihan umum, proses ini akan dilakukan dengan beberapa tahapan sebagai berikut:

Train & Test Split

Pada proses ini data akan dibagi menjadi data latih dan data uji. Data latih digunakan untuk mempelajari bentuk dari data benar dan data salah yang akan diklasifikasi, proses ini menggunakan 1660 data atau 80% dari total dataset yang telah diproses, data uji digunakan untuk menguji program yang telah melakukan latihan dengan data benar dan data salah sehingga dapat dijadikan untuk bahan evaluasi, proses ini menggunakan 416 data atau 20% dari total dataset yang telah diproses.

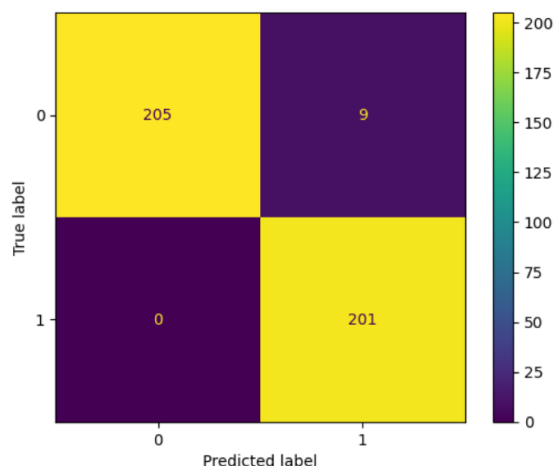
Proses Modeling

Pada skenario uji, setiap berita dalam dataset yang terdiri dari 2076 berita diklasifikasikan menggunakan metode K-Nearest Neighbor (KNN) ke dalam dua kategori: berita benar dan berita salah. Dataset tersebut, setelah melalui proses pembersihan dan prapemrosesan teks (case-folding, tokenisasi, penghapusan stopwords, stemming, dan normalisasi), dibagi menjadi data latih (80%) dan data uji (20%). Data uji diproses menggunakan tahapan prapemrosesan yang sama, kemudian dianalisis dengan model KNN yang telah dilatih menggunakan data latih untuk menentukan kelasnya berdasarkan kedekatan dengan data tetangga terdekat (jumlah K optimal). Hasil klasifikasi berupa label berita benar atau berita salah dievaluasi menggunakan Confusion Matrix untuk mengukur akurasi, presisi, recall, dan F1-score, sehingga validitas model dalam mendeteksi berita palsu dapat dipastikan.

Evaluation

Pada tahapan ini hasil dari program dievaluasi performa metode klasifikasi yang digunakan menggunakan *Confusion Matrix*, proses ini dilakukan untuk mengetahui seberapa baik program bekerja dalam mengklasifikasi berita pemilihan umum. Adapun nilai yang diukur dalam penelitian ini adalah *nilai Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

Pada tahap ini ditunjukkan hasil klasifikasi berita dari metode *K-Nearest Neighborn* menggunakan *Confusion Matrix*. Metode ini menunjukkan berapa banyak data yang diklasifikasi dengan benar dan berapa banyak data yang diklasifikasi dengan salah, lalu hasil evaluasi dari klasifikasi akan dijumlahkan dan dapat diketahui berapa nilai dari Accuracy-nya. Hasil dari Confusion Matrix pada gambar 1 menunjukkan data yang diklasifikasikan dengan benar adalah 205, sedangkan data yang diklasifikasikan dengan salah 201. Adapun hasil dari proses evaluasi *Confusion Matrix* adalah sebagai berikut:



Gambar 1. *Confusion Matrix*

Adapun nilai *accuracy*, *prescission*, *recall*, dan *F-1 score* sebagai berikut

```
Akurasi 0.9783132530120482
Precision 0.9571428571428572
Recall 1.0
F1-skor 0.9781021897810218
```

Gambar 2. Nilai Akurasi

Setelah itu diketahui hasil Accuracy pada gambar 2 menunjukkan bahwa kinerja metode klasifikasi *K-Nearest Neighborn* bekerja dengan baik dengan rata-rata performa *K-Nearest Neighborn* 0.97, dan dapat disimpulkan bahwa metode klasifikasi ini bagus untuk digunakan dalam klasifikasi berita benar dan berita salah.

4. SIMPULAN

Penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan berita benar dan berita salah terkait Pemilihan Umum Presiden 2024. Dataset yang terdiri dari 2076 berita dari turnbackhoax.id diproses melalui tahap pembersihan data, pelabelan dengan validasi pakar, serta prapemrosesan teks seperti *case-folding*, tokenisasi, penghapusan stopwords, stemming, dan normalisasi. Model KNN diuji dengan pembagian data latih (80%) dan data uji (20%), menghasilkan performa tinggi dengan akurasi 97%, presisi 95%, recall 100%, dan F1-score 97%. Hasil ini menunjukkan bahwa metode KNN efektif untuk mendeteksi berita palsu, didukung oleh proses prapemrosesan yang tepat, sehingga dapat menjadi solusi dalam menghadapi disinformasi selama pemilu.

5. SARAN

Berdasarkan hasil penelitian, disarankan untuk menguji metode KNN pada dataset yang lebih besar dan lebih beragam untuk memastikan generalisasi model terhadap berbagai jenis berita palsu. Selain itu, penggabungan metode lain seperti algoritma berbasis pembelajaran mendalam atau pemodelan topik dapat dieksplorasi untuk meningkatkan akurasi dan efisiensi klasifikasi. Penting juga untuk memperbarui dataset secara berkala agar sistem tetap relevan dengan tren berita terbaru. Terakhir, integrasi model ini ke dalam platform pendeteksi berita palsu yang mudah diakses oleh masyarakat dapat menjadi langkah praktis untuk meminimalkan dampak disinformasi selama pemilu.

DAFTAR PUSTAKA

- [1] F. Nur Rozi and D. Harini Sulistyawati, “Klasifikasi Berita Hoax Pilpres Menggunakan Metode Modified K-Nearest Neighbor Dan Pembobotan Menggunakan TF-IDF,” 2019.
- [2] A. Yonathan, H. Sujaini, E. Esyudha Pratama, J. H. Nawawi, and K. Barat, “Perbandingan Algoritma Klasifikasi dalam Pendeteksian Hoax pada Media Sosial,” *Jurnal Aplikasi dan Riset Informatika*, vol. 1, 2022, doi: 10.26418/juara.v1i1.53126.
- [3] M. Diki Hendriyanto and N. Sari, “Penerapan Algoritma K-Nearest Neighbor Penerapan Algoritma K-Nearest Neighbor dalam Klasifikasi Judul Berita Hoax.”
- [4] M. U. Shalih, T. Endra, and E. Tju, “Pembangunan Fitur dalam Identifikasi Cerdas Hoaks dengan Naïve Bayes dan Klasifikasi Decision Tree”.
- [5] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” *Science* (1979), vol. 359, no. 6380, pp. 1146–1151, Mar. 2018, doi: 10.1126/science.aap9559.
- [6] J. Li *et al.*, “KRA: K-Nearest Neighbor Retrieval Augmented Model for Text Classification,” *Electronics (Switzerland)*, vol. 13, no. 16, Aug. 2024, doi: 10.3390/electronics13163237.
- [7] N. Isnaini, Adiwijaya, M. S. Mubarak, and M. Y. A. Bakar, “A multi-label classification on topics of Indonesian news using K-Nearest Neighbor,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, May 2019. doi: 10.1088/1742-6596/1192/1/012027.
- [8] B. Kaida Palma, D. Triantoro Mardiansyah, and W. Astuti, “Klasifikasi Teks Artikel Berita Hoaks Covid-19 dengan Menggunakan Algotrima K-Nearest Neighbor.”
- [9] Siti Nur Asiyah and Kartika Fithriasari, “Klasifikasi Berita Online Menggunakan Metode Support Vector Machine dan K- Nearest Neighbor,” *JURNAL SAINS DAN SENI ITS*, vol. 5, no. 2, 2016.