

**Deteksi Hoax Di Media Sosial Menggunakan *Natural Language Processing* Dan *Naïve Bayes***

**LAPORAN TUGAS AKHIR**

Laporan ini Disusun untuk Memenuhi Salah Satu Syarat Memperoleh Gelar Sarjana Strata 1 (S1) pada Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang



**DISUSUN OLEH:**

**Muhammad Nur Chanif**

**32602100084**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INDUSTRI**

**UNIVERSITAS ISLAM SULTAN AGUNG**

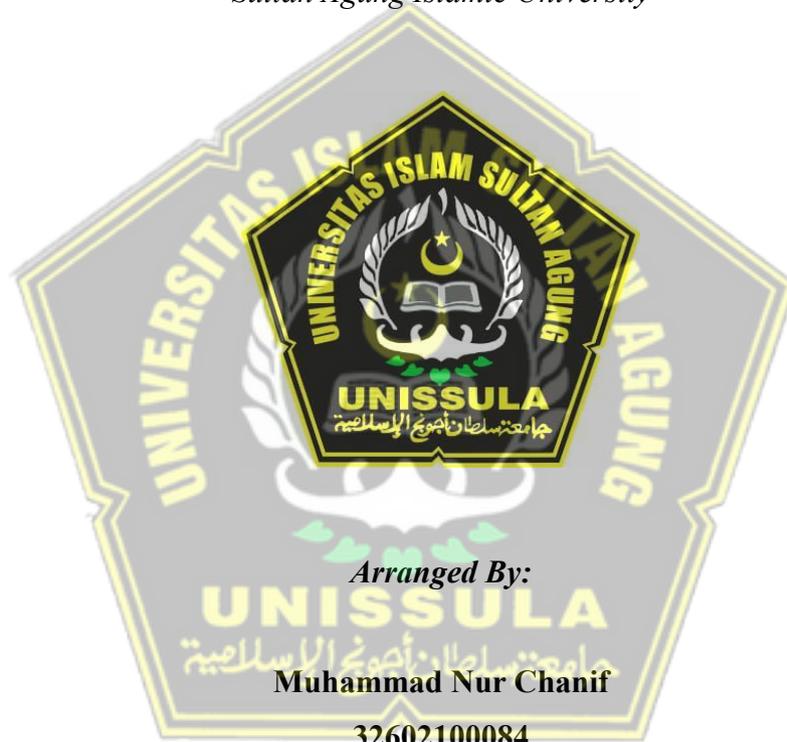
**SEMARANG**

**2025**

***FINAL PROJECT***

**Detecting Hoaxes on Social Media Using Natural Language  
Processing and Naïve Bayes**

*Proposed to complete the requirement to obtain a bachelor's degree (S1)  
at Informatics Engineering Departement of Industrial Technology Faculty  
Sultan Agung Islamic University*



*Arranged By:*

**Muhammad Nur Chanif**

**32602100084**

***MAJORING OF INFORMATICS ENGINEERING  
FACULTY OF INDUSTRIAL TECHNOLOGY  
SULTAN AGUNG ISLAMIC UNIVERSITY***

**SEMARANG**

**2025**

LEMBAR PENGESAHAN  
TUGAS AKHIR

Deteksi Hoax Di Media Sosial Menggunakan *Natural Language Processing*  
Dan *Naïve Bayes*

MUHAMMAD NUR CHANIF  
NIM 32602100081

Telah dipertahankan di depan tim penguji ujian sarjana tugas akhir  
Program Studi Teknik Informatika  
Universitas Islam Sultan Agung  
Pada tanggal : 23.06.2025

TIM PENGUJI UJIAN SARJANA :

Much Taufik, ST, MIT.  
NIDN. 0622037502  
(Ketua Penguji)

10-06-2025

Bagus Satrio Waluyo Poetro,  
S.Kom., M.Cs.  
NIDN. 1027118801  
(Anggota Penguji)

10-06-2025

Imam Much Ibnu Subroto, ST,  
M.Sc., Ph.D  
NIDN. 0613037301  
(Pembimbing)

10-06-2025

Semarang, 10-06-2025

Mengetahui,

Kapredi Teknik Informatika  
Universitas Islam Sultan Agung



Much Taufik, S.T., M.IT  
NIDN. 0622037502

## SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Muhammad Nur Chanif

NIM : 32602100084

Judul Tugas Akhir : Deteksi Hoax Di Media Sosial Menggunakan *Natural Language Processing* Dan *Naïve Bayes*

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, 15 Maret 2025

Yang Menyatakan



Muhammad Nur Chanif

## PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Muhammad Nur Chanif

NIM : 3260210004

Program Studi : Teknik Informatika

Fakultas : Teknologi industri

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul :  
Deteksi Hoax Di Media Sosial Menggunakan *Natural Language Processing* Dan  
*Naive Bayes*

Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan diinternet dan media lain untuk kepentingan akademis selama tetap menyantumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiatisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan Agung.

Semarang, 15 Maret 2025

Yang Menyatakan



Muhammad Nur Chanif

## KATA PENGANTAR

Dengan mengucapkan rasa syukur alhamdulillah atas kehadiran Allah SWT yang telah memberikan Rahmat dan karunianya kepada penulis, sehingga dapat menyelesaikan Laporan Proposal Tugas Akhir dengan judul “Deteksi Hoax Di Media Sosial Menggunakan *Natural Language Processing* Dan *Naïve Bayes*” ini untuk memenuhi salah satu syarat menyelesaikan studi serta dalam rangka memperoleh gelar sarjana (S-1) pada program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Islam Sultan Agung.

Laporan Proposal Tugas Akhir ini disusun dan dibuat dengan adanya bantuan dari berbagai pihak, materi maupun teknis, oleh karena itu saya selaku penulis mengucapkan terima kasih kepada:

1. Rektor UNISSULA Bapak Prof. Dr. H. Gunarto, S.H., M.H yang mengizinkan penulis menimba ilmu dikampus ini.
2. Dekan Fakultas Teknologi Industri Ibu Dr. Ir. Novi Marlyana, S.T., M.T., IPU., ASEAN. Eng.
3. Dosen Pembimbing penulis Bapak Imam Much Ibnu Subroto, ST, M.Sc, Ph.D yang memberikan banyak nasehat dan saran.
4. Orang tua dan keluarga penulis yang menjadi *support system* dan mengizinkan untuk menyelesaikan laporan ini.
5. Para sahabat, teman-teman yang telah memberikan begitu banyak bantuan, semangat, inspirasi, dan diskusi progres penyusunan Tugas Akhir.

Semarang, 15 Maret 2025

Muhammad Nur Chanif

## DAFTAR ISI

LEMBAR PENGESAHAN TUGAS AKHIR.....	iii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH <b>Error! Bookmark not defined.</b>	
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL .....	ix
DAFTAR GAMBAR .....	x
ABSTRAK.....	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah .....	2
1.3 Pembatasan Masalah.....	2
1.4 Tujuan .....	3
1.5 Manfaat .....	3
1.6 Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI .....	5
2.1 Tinjauan Pustaka.....	5
2.2 Dasar Teori.....	7
2.2.1 Sistem Pendeteksi dan Hoax di Media Sosial .....	7
2.2.2 Natural Language Processing (NLP).....	7
2.2.3 Term Frequency – Inverse Document Frequency (TF-IDF) .....	11
2.2.4 Naive Bayes .....	12
2.2.5 Python .....	13
2.2.6 Evaluasi Model klasifikasi.....	14
BAB III METODE PENELITIAN.....	17
3.1 Metode Penelitian .....	17
3.2 Studi Literatur .....	17
3.3 Pengumpulan Data .....	18
3.4 Perancangan Implementasi Sistem.....	18

3.5	Pembuatan Sistem .....	20
BAB IV HASIL DAN ANALISIS PENELITIAN .....		22
4.1	Persiapan Data .....	22
4.2	<i>Preprocessing</i> Data .....	22
4.2.1	<i>Cleaning</i> .....	22
4.2.2	<i>Case Folding</i> .....	23
4.2.3	<i>Tokenizing</i> .....	24
4.2.4	Stopword Removal .....	26
4.2.5	Lematization .....	27
4.3	Fiture Extraction .....	28
4.3.1	Term Frequency-Inverse Document Frequency (TF-IDF) .....	28
4.4	Optimasi model .....	30
4.4.1	Split Data .....	30
4.4.2	Training Model Naive Bayes .....	31
4.5	Hasil .....	32
4.6	Tampilan <i>Streamlit</i> .....	34
4.6.1	Tampilan Analisis Teks .....	34
4.7	Pembahasan .....	35
BAB V KESIMPULAN DAN SARAN .....		37
5.1	Kesimpulan .....	37
5.2	Saran .....	37
DAFTAR PUSTAKA .....		39

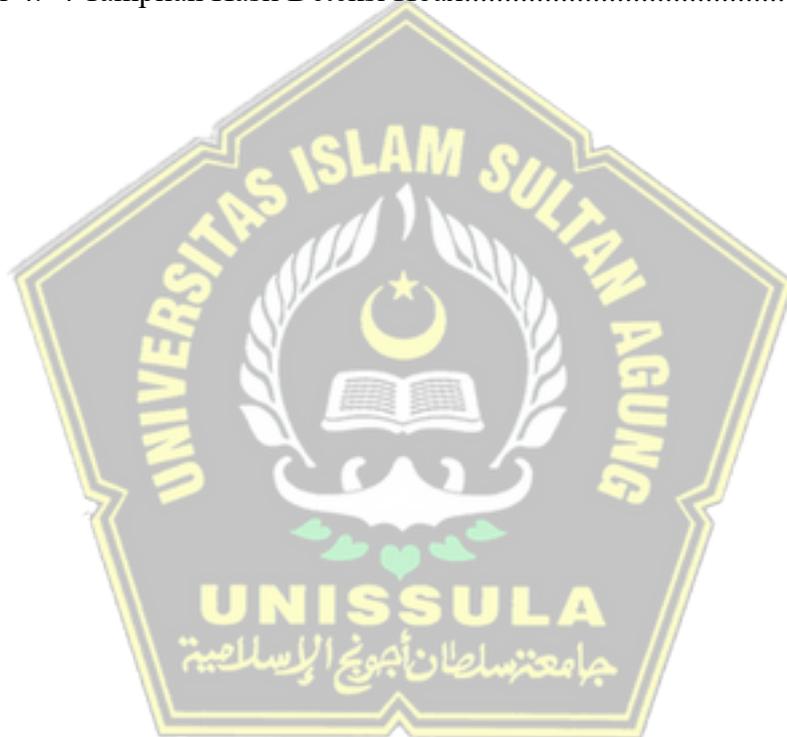
## DAFTAR TABEL

Tabel 4. 1 hasil data sesudah dan sebelum cleaning .....	22
Tabel 4. 2 hasil data sesudah dan sebelum Case Folding.....	23
Tabel 4. 3 hasil data sesudah dan sebelum Case Folding.....	25
Tabel 4. 4 hasil data sesudah dan sebelum Stopword removal .....	26
Tabel 4. 5 hasil data sesudah dan sebelum Lemmatization.....	28
Tabel 4. 6 hasil data sesudah dan sebelum TF IDF .....	29



## DAFTAR GAMBAR

Gambar 3. 1 Flowchart Alur Penelitian.....	17
Gambar 3. 2 Flowchart Alur Sistem.....	18
Gambar 3. 3 Flowchart Alur Kerja Sistem.....	21
Gambar 4. 1 Garafik Confusion Matrix .....	32
Gambar 4. 2 Tampilan Deteksi Hoax .....	34
Gambar 4. 3 Tampilan Hasil Deteksi Bukan Hoax .....	35
Gambar 4. 4 Tampilan Hasil Deteksi Hoax.....	35



## ABSTRAK

Media sosial berfungsi sebagai sarana penyebaran informasi yang sangat cepat, tidak hanya memperkuat interaksi antarindividu, tetapi juga membuka peluang bagi beredarnya informasi yang belum terverifikasi, bahkan mengandung unsur kebohongan. Keberadaan hoaks dapat memicu disinformasi, keresahan publik, hingga menimbulkan dampak negatif dalam aspek sosial maupun ekonomi. Oleh karena itu, upaya untuk mendeteksi hoaks merupakan suatu langkah yang krusial. Penelitian ini bertujuan untuk merancang sebuah sistem yang mampu mendeteksi berita palsu di media sosial melalui pendekatan *Natural Language Processing* (NLP) serta penerapan algoritma *Naïve Bayes*. NLP digunakan dalam proses analisis teks untuk mengenali pola-pola linguistik khas yang sering ditemukan dalam hoaks. Sementara itu, algoritma *Naïve Bayes* dipilih karena efektivitasnya dalam melakukan klasifikasi teks berbasis probabilitas. Untuk menilai kinerja sistem, digunakan metrik evaluasi berupa akurasi, presisi, dan *recall*. Berdasarkan hasil pengujian, sistem yang di bangun mampu mencapai akurasi sebesar 95,87%, presisi sebesar 95,23%, dan *recall* sebesar 96,80%. Temuan ini mengindikasikan bahwa model yang dibangun memiliki kinerja yang tinggi dan andal dalam membedakan antara konten hoaks dan non-hoaks di platform media sosial. Penelitian ini diharapkan dapat menjadi kontribusi awal dalam pengembangan solusi berbasis kecerdasan buatan guna menekan penyebaran misinformasi serta meningkatkan kepercayaan terhadap informasi di ruang digital.

Kata Kunci : Deteksi Hoax, Natural Language Processing, Naïve Bayes, Sosial Media , Berita Palsu

### ABSTRACT

Social media serves as a platform for the rapid dissemination of information, not only strengthening interactions between individuals but also opening opportunities for the spread of unverified information, including those containing elements of falsehood. The presence of hoaxes can trigger disinformation, cause public unrest, and lead to negative impacts on both social and economic aspects. Therefore, efforts to detect hoaxes are a crucial step. This research aims to design a system capable of detecting fake news on social media using a Natural Language Processing (NLP) approach and the application of the Naïve Bayes algorithm. NLP is used in the text analysis process to recognize distinctive linguistic patterns commonly found in hoaxes. Meanwhile, the Naïve Bayes algorithm is chosen for its effectiveness in performing probability-based text classification. To evaluate the system's performance, accuracy, precision, and recall are used as evaluation metrics. Based on testing results, the developed system achieved an accuracy of 95.87%, a precision of 95.23%, and a recall of 96.80%. These findings indicate that the developed model has high performance and reliability in distinguishing between hoax and non-hoax content on social media platforms. This research is expected to serve as an initial contribution to the development of artificial intelligence-based solutions to curb the spread of misinformation and enhance trust in information within digital spaces.

*Keywords:* Hoax Detection, Natural Language Processing, Naïve Bayes, Social Media, Fake News

# BAB I PENDAHULUAN

## 1.1 Latar Belakang.

Saat ini, media sosial telah menjadi bagian integral dari arus penyebaran informasi yang berlangsung sangat cepat dan luas, memungkinkan masyarakat untuk saling berbagi konten secara instan. Meskipun hal ini memberikan kemudahan dalam komunikasi, kemajuan tersebut juga membawa dampak negatif berupa munculnya berita palsu atau hoaks. Informasi yang menyesatkan ini dapat tersebar secara masif dan tanpa pengawasan yang memadai, sehingga menyesatkan publik dan memengaruhi opini serta keputusan sosial dan politik (Pennycook & Rand, 2019). Oleh karena itu, deteksi dan hoaks menjadi salah satu tantangan utama dalam era informasi ini.

Media sosial, sebagai sarana penyebaran informasi yang sangat cepat, tidak hanya meningkatkan komunikasi antar individu tetapi juga memberi ruang bagi penyebaran informasi yang tidak terverifikasi dan sering kali salah. Dalam hal ini, deteksi hoaks menjadi krusial, karena hoaks dapat menyebabkan kebingungan, kecemasan, dan bahkan kerugian ekonomi atau sosial. Oleh karena itu, pembuatan sistem yang mampu mendeteksi dan mencegah penyebaran hoaks di media sosial menggunakan teknologi canggih seperti *Natural Language Processing* (NLP) dan *Machine Learning* (ML) menjadi sangat penting.

*Natural Language Processing* (NLP) telah banyak digunakan dalam analisis teks untuk mendeteksi pola dalam bahasa alami. Dengan NLP, teks dapat dianalisis secara mendalam, membantu dalam mengenali pola-pola bahasa yang sering digunakan dalam penyebaran hoaks. Teknik ini memungkinkan sistem untuk memahami konteks kalimat, mendeteksi kontradiksi, dan menilai kredibilitas informasi. Salah satu metode yang banyak digunakan dalam kombinasi dengan NLP adalah *Naive Bayes*, sebuah algoritma klasifikasi yang efektif untuk memproses data teks dalam jumlah besar dengan hasil yang akurat (Indrayuni et al., 2021). Meskipun sudah ada penelitian yang membahas deteksi hoaks dengan metode NLP dan *Naive Bayes*, tantangan tetap ada. Misalnya, keterbatasan model

dalam menangani variasi bahasa informal yang umum di media sosial. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknologi deteksi hoaks dengan memanfaatkan kombinasi teknik NLP dan *Naive Bayes*. Dengan pendekatan berbasis konteks dan memperhatikan dinamika konten di media sosial, diharapkan sistem ini dapat mendeteksi hoaks dengan akurasi yang tinggi di platform sosial media dan diharapkan mampu membantu meminimalkan dampak negatif hoaks di media sosial dan menciptakan lingkungan informasi yang lebih terpercaya.

## 1.2 Perumusan Masalah

Berdasarkan uraian latar belakang yang telah dijelaskan, permasalahan utama yang menjadi fokus dalam penelitian ini adalah bagaimana cara membangun sistem deteksi hoaks di media sosial menggunakan *Natural Language Processing* (NLP) dan algoritma *Machine Learning Naïve Bayes*?

## 1.3 Pembatasan Masalah

1. Penelitian ini hanya fokus pada deteksi hoaks di media sosial menggunakan teknik *Natural Language Processing* (NLP) dan *Machine Learning*, dengan teks yang diambil dari platform media sosial tertentu seperti Twitter. Tanpa mencakup analisis media lain seperti berita cetak, siaran televisi, atau platform lainnya.
2. Penelitian ini tidak akan membahas aspek psikologis atau sosial dari penyebaran hoaks, melainkan hanya akan menyoroti teknik-teknik dalam pengolahan dan analisis data untuk deteksi hoaks.
3. Keterbatasan dataset yang digunakan dapat mempengaruhi hasil deteksi, sehingga penelitian ini tidak dapat mengklaim generalisasi untuk semua jenis hoaks yang mungkin ada di media sosial.

#### 1.4 Tujuan

Tujuan tugas akhir ini adalah membangun sistem deteksi hoaks di media sosial menggunakan *Natural Language Processing* (NLP) dan algoritma *Naïve Bayes* untuk identifikasi berita palsu.

#### 1.5 Manfaat

Penelitian ini memiliki manfaat sebagai sistem untuk mendeteksi hoaks dan mencegah penyebarannya di media sosial, sehingga dapat membantu masyarakat dalam memperoleh informasi yang lebih akurat dan terpercaya, serta menjadi serta bisa dijadikan contoh untuk mengembangkan teknologi agar bisa melawan penyebaran berita palsu di internet.

#### 1.6 Sistematika Penulisan

Penyusunan laporan tugas akhir menerapkan tata cara penulisan yaitu:

##### BAB 1 : PENDAHULUAN

Bab ini berisi penjelasan mengenai latar belakang, perumusan masalah, pembatasan masalah, tujuan penelitian, dan metodologi penelitian, serta sistematika penulisan.

##### BAB 2 : TINJAUAN PUSTAKA DAN DASAR TEORI

Bab ini mencakup penelitian-penelitian terdahulu yang relevan serta dasar teori yang mendukung penelitian.

##### BAB 3 : METODE PENELITIAN

Bab ini menjelaskan tahapan penelitian secara sistematis, mulai dari pengumpulan data hingga proses klasifikasi data.

##### BAB 4 : HASIL DAN ANALISIS PENELITIAN

Bab ini menyajikan hasil penelitian, khususnya hasil klasifikasi teks dari media sosial ke dalam kategori hoaks dan non-hoaks menggunakan teknik *Natural Language Processing* (NLP) dan algoritma *Naïve Bayes*

## BAB 5 : KESIMPULAN DAN SARAN

Bab terakhir berisi ringkasan keseluruhan penelitian dari tahap awal hingga akhir, serta memberikan saran bersarkan temuan yang diperoleh.



## BAB II

### TINJAUAN PUSTAKA DAN DASAR TEORI

#### 2.1 Tinjauan Pustaka

Tinjauan terhadap penelitian-penelitian sebelumnya memiliki peran penting dalam mendukung penelitian ini. Di satu sisi, tinjauan tersebut memberikan dasar perbandingan yang berguna untuk menilai kelebihan dan kekurangan dari studi yang telah dilakukan sebelumnya. Selain itu, hasil-hasil penelitian terdahulu juga digunakan untuk memperkuat landasan argumen dalam penelitian ini, sehingga peneliti dapat memperoleh pemahaman yang lebih luas dan mendalam terkait topik yang sedang dikaji. Beberapa penelitian serupa dapat ditemukan pada referensi berikut.

Sebuah penelitian lain mengusulkan penggunaan teknik *ensemble* yang mengintegrasikan *transformers*, NLP, dan *machine learning* untuk deteksi berita palsu. Pendekatan multimodal ini menggabungkan analisis teks dengan citra dan video, dan menggunakan BERT (*Bidirectional Encoder Representations from Transformers*) untuk analisis teks serta CNN (*Convolutional Neural Networks*) untuk data visual. Penelitian ini menunjukkan bahwa model *multimodal* memiliki akurasi yang lebih baik dibandingkan dengan model-model sebelumnya, memberikan kontribusi signifikan dalam deteksi *hoax* yang lebih efektif di berbagai media (Transformers & Processing, 2024).

Pertumbuhan pesat media sosial telah mendorong peningkatan signifikan dalam akses terhadap informasi, namun di saat yang sama mempercepat peredaran informasi yang tidak akurat. Kondisi ini menjadikan proses identifikasi terhadap informasi yang tidak valid sebagai tantangan utama dalam bidang *Natural Language Processing* (NLP). Untuk menghadapi persoalan tersebut, diperlukan penerapan metode klasifikasi yang lebih canggih, salah satunya melalui pemanfaatan teknologi *Bidirectional Encoder Representations from Transformers* (BERT), yang terbukti akurat dalam mendeteksi keberadaan berita palsu (Sukmawati et al., 2024)

Perkembangan pesat dari berbagai platform media sosial telah menyebabkan peningkatan dalam kemudahan akses terhadap informasi, sekaligus mempercepat persebaran konten yang tidak akurat. Hal ini menimbulkan tantangan krusial dalam bidang *Natural Language Processing* (NLP), khususnya dalam hal identifikasi informasi yang tidak valid. Untuk menghadapi persoalan ini secara lebih efektif, diperlukan penerapan metode klasifikasi yang canggih, salah satunya melalui pemanfaatan teknologi *Bidirectional Encoder Representations from Transformers* (BERT) yang dinilai efektif dan akurat dalam proses deteksi berita palsu (Manzoor et al., 2019).

Pendekatan berbasis analisis media dapat memberikan performa yang lebih baik dengan memanfaatkan integrasi data teks dan sinyal sosial. Hal ini dilakukan dengan mencakup kombinasi antara konten teks dan sinyal media sosial dan penggunaan model ensemble berbasis NLP untuk meningkatkan akurasi deteksi berita palsu di berbagai *platform* media sosial (Alim Al Ayub et al., 2021)

Peningkatan signifikan dalam pertumbuhan platform media sosial tidak hanya memperluas akses masyarakat terhadap informasi, tetapi juga mempercepat laju penyebaran informasi yang keliru. Dalam konteks ini, tantangan utama yang dihadapi dalam bidang *Natural Language Processing* (NLP) adalah kemampuan untuk mengidentifikasi informasi yang tidak valid secara akurat dan efisien. Pendekatan manual dalam mendeteksi berita palsu dinilai tidak lagi memadai mengingat volume informasi yang terus bertambah secara cepat, khususnya di media sosial. Oleh karena itu, dibutuhkan sistem klasifikasi otomatis berbasis kecerdasan buatan, di mana teknologi *Bidirectional Encoder Representations from Transformers* (BERT) menjadi salah satu solusi yang ditawarkan untuk meningkatkan efektivitas dalam proses pendeteksian berita palsu. Selain itu, strategi deteksi dapat dilakukan melalui empat pendekatan utama, yaitu berbasis pengetahuan (*knowledge-based*), gaya penulisan (*style-based*), pola penyebaran (*propagation-based*), dan kredibilitas sumber (*source-based*) (Yunanto et al., 2021)

## 2.2 Dasar Teori

### 2.2.1 Sistem Pendeteksi Hoax di Media Sosial

Sistem pendeteksi hoax merupakan salah satu pendekatan modern yang bertujuan untuk mencegah penyebaran informasi yang tidak benar dan menyesatkan yang beredar luas di media sosial. Teknologi ini menggantikan metode manual dalam metode memeriksa keabsahan yang tidak hanya memakan waktu, tetapi juga kurang efektif dalam menghadapi lalu lintas informasi yang tinggi. Sistem deteksi *hoax* berbasis NLP dan algoritma *Naive Bayes* merupakan solusi efektif untuk menghadapi penyebaran informasi palsu di media sosial. Sistem ini menggantikan metode manual dalam memeriksa validitas informasi, yang sering kali memakan waktu dan kurang efisien. Dengan menggabungkan kemampuan *Natural Language Processing* (NLP) dan algoritma *Naive Bayes*, sistem mampu mengidentifikasi pola linguistik yang menunjukkan adanya *hoax*. NLP membantu menganalisis teks dengan memahami struktur kalimat, konteks bahasa, serta sentimen yang terkandung di dalamnya. Sementara itu, algoritma *Naive Bayes* digunakan untuk klasifikasi berita *hoax* dan *non-hoax* berdasarkan probabilitas dari fitur-fitur yang telah diekstraksi.

*Naive Bayes* memiliki kinerja yang baik dalam mendeteksi berita *hoax* di media sosial dengan memanfaatkan model CRISP-DM untuk analisis data. Data pelatihan diperoleh dari situs berita dengan jumlah data yang signifikan, dan sistem ini diuji menggunakan *confusion matrix* untuk evaluasi hasil klasifikasi. Implementasi sistem deteksi *hoax* ini dapat membantu meminimalisir penyebaran informasi menyesatkan dan meningkatkan literasi digital masyarakat (Agustina et al., 2022). Kombinasi algoritma *Naive Bayes* dan NLP dalam klasifikasi berita *hoax* memberikan tingkat akurasi yang tinggi, terutama ketika data yang digunakan telah melalui proses tokenisasi dan analisis sintaksis untuk mengekstraksi fitur-fitur penting dari teks (’Ulhaq et al., 2022).

### 2.2.2 Natural Language Processing (NLP)

*Natural Language Processing* (NLP) merupakan salah satu bidang dalam kecerdasan buatan yang berfokus pada pengembangan kemampuan sistem

komputer untuk berinteraksi dengan bahasa manusia. Teknologi ini dirancang agar komputer mampu menganalisis, memahami, dan menginterpretasikan bahasa alami dengan pendekatan yang mendekati cara manusia memproses informasi linguistik. Teknologi ini mengkombinasikan teknik linguistik dan *machine learning* untuk menerjemahkan, menganalisis, dan menghasilkan teks atau ucapan, membuat komputer mampu menangani teks atau suara dari berbagai bahasa dengan lebih efektif. NLP sangat penting untuk memungkinkan komputer memahami bahasa manusia, dan menggunakan berbagai tahapan, termasuk tokenisasi, penguraian sintaksis, dan analisis semantik.

*Natural Language Processing* (NLP) memiliki beberapa kelebihan yang menjadikannya alat penting dalam berbagai aplikasi. Salah satu kelebihan utama NLP adalah kemampuannya dalam memahami dan menghasilkan bahasa alami, memungkinkan interaksi antara manusia dan komputer lebih mudah melalui teks atau suara. NLP bermanfaat untuk analisis data skala besar, yaitu text mining dan analisis sentimen. NLP telah digunakan secara efektif dalam banyak aplikasi, termasuk deteksi hoax, di mana algoritma seperti Transformer, BERT, dan GPT digunakan untuk menganalisis dan mengklasifikasikan teks dengan tingkat akurasi yang tinggi (Sawicki et al., 2023).

Dalam praktiknya, NLP banyak digunakan dalam *chatbot*, asisten virtual, dan aplikasi personalisasi rekomendasi yang bergantung pada pemahaman konteks dan bahasa pengguna (Strubell & Mccallum, n.d.). Secara keseluruhan, NLP berperan penting dalam mendorong kemampuan teknologi untuk memproses dan memahami bahasamanusia secara lebih mendalam, memungkinkan aplikasi yang efisien dan akurat. *Natural Language Processing* (NLP) terdiri dari beberapa komponen inti yang membentuk struktur pemrosesan bahasa, yaitu:

- **Tokenisasi** merupakan tahapan untuk memisahkan teks menjadi bagian-bagian kecil, seperti kata atau kalimat, sehingga lebih mudah diproses.
- **Stemming dan Lematisasi** bertujuan untuk menyederhanakan kata ke bentuk dasarnya agar makna inti dari kata tersebut tetap dapat dikenali.

- **Part-of-Speech (POS) Tagging** digunakan untuk menentukan jenis kata untuk mengidentifikasi kategori gramatikal dari setiap kata dalam suatu kalimat, seperti kata benda, kata kerja, maupun kata sifat.
- **Named Entity Recognition (NER)** berfungsi untuk mendeteksi informasi penting dalam teks, seperti nama orang, tempat, waktu, atau nama organisasi.
- **Parsing** adalah proses menganalisis struktur kalimat guna memahami hubungan gramatikal antar kata dalam sebuah teks.

Proses *Natural Language Processing* (NLP) melibatkan serangkaian tahap yang didesain untuk memahami dan memproses bahasa manusia agar dapat diinterpretasikan oleh komputer. Menurut literatur ilmiah terkini, proses ini melibatkan beberapa langkah penting, diantaranya :

a. *Cleaning*

*Cleaning* dalam *Natural Language Processing* (NLP) adalah proses awal dalam *preprocessing* data teks yang bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan atau mengganggu analisis. Secara umum, proses ini mencakup:

- Menghapus simbol-simbol atau karakter khusus (misalnya: @, #, %, !)
- Menghilangkan angka yang tidak diperlukan
- Menghapus tag HTML, URL, dan *mention* (@username)
- Menghapus spasi berlebih dan tanda baca yang tidak penting

*Cleaning* penting karena data teks mentah dari media sosial sering kali mengandung banyak "noise" atau gangguan yang dapat memengaruhi akurasi model NLP. Dengan membersihkan teks, sistem dapat fokus menganalisis informasi yang benar-benar bermakna.

b. *Case Folding*

*Case folding* adalah proses mengubah seluruh teks menjadi huruf kecil (*lowercase*). Tujuannya adalah untuk menyamakan bentuk kata yang secara makna sama tetapi secara penulisan berbeda huruf besar-kecilnya. Contohnya Kata "*Hoax*", "*hoax*", dan "*HOAX*" akan diubah menjadi "*hoax*" semua.

Ini penting karena model NLP memperlakukan kata dengan huruf besar dan kecil sebagai entitas yang berbeda. Tanpa *case folding*, kata yang sama bisa

dihitung sebagai kata yang berbeda, sehingga mengganggu analisis dan klasifikasi teks.

c. *Tokenizing*

Tokenisasi merupakan tahapan dalam pemrosesan teks yang bertujuan untuk membagi suatu kalimat atau dokumen menjadi bagian-bagian yang lebih kecil yang disebut token. Token tersebut dapat berupa kata, frasa, atau bahkan karakter, tergantung pada konteks dan tujuan analisis yang dilakukan..

Contohnya:

Kalimat: "Hoax tersebar luas di media sosial."

Setelah tokenizing menjadi:

["Hoax", "tersebar", "luas", "di", "media", "sosial", "."]

Proses ini sangat penting dalam *Natural Language Processing* karena model hanya dapat menganalisis teks setelah dipecah menjadi bagian-bagian yang bermakna. *Tokenizing* menjadi dasar untuk proses NLP selanjutnya, seperti *stopword removal* dan *lemmatization*.

d. *Stopword Removal*

*Stopword Removal* adalah proses menghapus kata-kata umum (disebut *stopword*) yang tidak memiliki makna signifikan dalam analisis teks. Kata-kata ini biasanya berupa kata **hubung**, kata depan, atau kata bantu seperti: "yang", "di", "ke", "dan", "adalah", dll.

Tujuan dari proses ini adalah Mengurangi kebisingan dalam data teks, Memfokuskan analisis pada kata-kata yang lebih bermakna dan relevan secara konteks. Contoh:

- **Kalimat asli:** "Informasi yang tersebar di media sosial sangat cepat."
- **Setelah *stopword removal*:** "Informasi tersebar media sosial sangat cepat"

Dengan menghapus *stopword*, teks menjadi lebih sederhana dan efisien untuk dianalisis oleh model NLP.

e. *Lematization*

*Lematization* merupakan suatu teknik dalam pemrosesan bahasa alami yang bertujuan untuk mengonversi suatu kata ke bentuk dasarnya atau *lemma*, dengan memperhatikan konteks serta makna kata tersebut dalam kalimat. Tidak seperti *stemming* yang hanya memotong kata ke bentuk dasarnya tanpa memperhatikan arti, *lemmatization* mempertahankan struktur bahasa yang benar. Proses ini bertujuan untuk Menyatukan berbagai bentuk kata menjadi satu bentuk dasar yang umum dan Meningkatkan akurasi dalam analisis teks karena model akan mengenali kata-kata berbeda sebagai satu konsep.

Contoh :

- Kata “berlari”, “berlari-lari”, “lari” → menjadi “lari”
- Kata “mengetik”, “diketik”, “pengetikan” → menjadi “ketik”

*Lemmatization* penting dalam proses NLP karena membantu menyederhanakan dan menstandarkan kata sebelum digunakan untuk pelatihan model atau klasifikasi.

### 2.2.3 *Term Frequency – Inverse Document Frequency (TF-IDF)*

TF-IDF (*Term Frequency – Inverse Document Frequency*) adalah metode pemberian bobot pada setiap kata dalam sebuah dokumen yang digunakan dalam pengolahan teks dan *Natural Language Processing* (NLP), khususnya untuk ekstraksi fitur. TF-IDF digunakan untuk mengukur seberapa penting suatu kata dalam suatu dokumen relatif terhadap kumpulan dokumen (corpus).

#### 1. *Term Frequency (TF)*

Frekuensi kemunculan suatu kata dalam sebuah dokumen dapat diukur dengan menentukan nilai *Term Frequency* (TF). Semakin tinggi frekuensi kemunculan kata tersebut, maka semakin besar pula nilai TF yang dihasilkan.

$$TF(t) = \frac{\text{Jumlah kemunculan kata } t \text{ dalam dokumen}}{\text{Jumlah total kata dalam dokumen}} \quad (1)$$

## 2. *Inverse Document Frequency (IDF)*

Pengukuran ini bertujuan untuk mengetahui tingkat keberadaan suatu kata secara keseluruhan dalam kumpulan dokumen. Kata yang muncul pada banyak dokumen akan memperoleh bobot yang lebih rendah, karena dianggap memiliki nilai informasi yang kurang spesifik.

$$IDF(t) = \log \left( \frac{N}{df(t)} \right) \quad (2)$$

- $N$  merepresentasikan total keseluruhan dokumen
- $df(t)$  menunjukkan jumlah dokumen yang mengandung kata  $t$  didalamnya

## 3. *TF - IDF Score*

Skor TF-IDF dihitung dengan mengalikan TF dengan IDF, dengan rumus :

$$TF-IDF(t) = TF(t) \times IDF(t) \quad (3)$$

Penggunaan metode TF-IDF dalam penelitian ini bertujuan untuk mentransformasikan data teks yang telah melalui tahap praproses menjadi bentuk representasi numerik, sehingga memungkinkan data tersebut untuk diproses lebih lanjut oleh algoritma klasifikasi, seperti *Naïve Bayes*. Contoh :

Misalnya, kata “*hoax*” muncul 3 kali dalam dokumen yang berisi 100 kata, dan hanya muncul di 2 dari 50 dokumen. Maka bobot TF-IDF dari kata tersebut akan lebih tinggi karena kata tersebut sering muncul di dokumen namun jarang di dokumen lain, menunjukkan kata tersebut relevan.

### 2.2.4 *Naïve Bayes*

*Naïve Bayes* merupakan salah satu algoritma klasifikasi yang berbasis pada prinsip probabilistik dan banyak digunakan dalam tugas-tugas klasifikasi teks. Algoritma ini bekerja dengan asumsi bahwa setiap fitur dalam data bersifat independen secara kondisional terhadap fitur lainnya, sehingga memungkinkan penerapannya secara efisien meskipun dengan pendekatan yang sederhana. Keunggulan tersebut menjadikan *Naïve Bayes* efektif dalam berbagai konteks, termasuk dalam upaya mendeteksi hoaks di media sosial. Dalam implementasinya

pada klasifikasi teks, algoritma ini menghitung probabilitas suatu dokumen tergolong ke dalam kategori tertentu dengan mempertimbangkan distribusi kemunculan kata dalam dokumen tersebut. ( Shengfeng et al., 2021). Algoritma *Naive Bayes* menentukan probabilitas suatu kelas target dengan menganalisis frekuensi kemunculan masing-masing fitur dalam kelas tersebut, kemudian menetapkan kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi akhir.

Algoritma ini efisien dalam perhitungan probabilitas kondisional, sehingga sangat cocok untuk data berukuran besar dan berstruktur sederhana. Contoh penggunaan *Naive Bayes* meliputi klasifikasi teks, deteksi spam, dan pengenalan pola berbasis data kategori. Keunggulan *Naive Bayes* terletak pada efisiensinya, baik dalam perhitungan maupun penggunaan memori, sehingga sangat cocok untuk dataset berukuran besar algoritma ini tetap menunjukkan kinerja yang baik dalam berbagai aplikasi, seperti : Klasifikasi teks (misalnya, deteksi spam), Analisis sentimen.

### 2.2.5 Python

*Python* adalah bahasa pemrograman umum yang termasuk tingkat tinggi, mudah dipelajari, dan dinamis. Pengembangan *Python* dimulai dari hobi bagi penciptanya, karena ia ingin membuat bahasa yang indah dan mudah dibaca oleh semua orang. Bahasa ini menggunakan indentasi untuk mendefinisikan blok kode, bukan tanda kurung kurawal, sehingga lebih aman dan jelas. Meskipun awalnya tidak mendapatkan banyak perhatian dibandingkan dengan bahasa lain, *Python* kini menjadi sorotan terutama dalam bidang *machine learning* dan kecerdasan buatan. *Python* membuat pekerjaan lebih produktif dan lebih mudah dilakukan karena daya komputasi yang luar biasa yang tersedia saat ini, yang memungkinkan kita menyelesaikan pekerjaan dengan lebih efisien meskipun memerlukan banyak waktu (Ajay Rawat, 2020). *Python* mendukung berbagai paradigma pemrograman, mencakup pemrograman berorientasi objek, fungsional, serta terstruktur, sehingga sangat fleksibel untuk beraneka ragam proyek dan aplikasi. *Python* mempunyai beragam pustaka yang luas, seperti *NumPy* dan *Pandas* untuk analisis data, *Matplotlib* dan *Seaborn* untuk visualisasi, *TensorFlow* dan *PyTorch*

untuk machine learning, serta *Flask* dan *Django* untuk pengembangan web.

Keunggulan utama *Python* terletak pada kemampuannya dalam menangani tugas-tugas kompleks dengan kode yang relatif lebih sedikit dibandingkan bahasa lain. Selain itu, *Python* memiliki komunitas pengguna yang besar, yang berkontribusi pada banyaknya sumber daya dan dukungan *online* yang tersedia. Selain itu, Dalam bidang ilmiah dan teknik, *Python* sering digunakan karena kemampuannya mengelola data skala besar dan mendukung analisis statistik yang mendalam. Misalnya, *Python* memungkinkan peneliti untuk membangun dan mengotomatisasi *pipeline* data yang kompleks, serta menyediakan fleksibilitas dalam pengembangan alat khusus sesuai kebutuhan.

### 2.2.6 Evaluasi Model klasifikasi

Memahami cara mengevaluasi kinerja model merupakan bagian penting dalam proses analisis data, baik data terstruktur maupun tidak terstruktur. Proses evaluasi ini dilakukan untuk mengukur tingkat akurasi model dalam menghasilkan prediksi yang sesuai dengan nilai atau kategori yang sebenarnya. Terdapat berbagai metrik yang bisa digunakan untuk menilai performa model secara objektif, seperti akurasi, presisi, recall, dan F1-score. Dengan melakukan evaluasi, kita dapat mengetahui kekuatan dan kelemahan model yang dibangun, sehingga dapat dilakukan perbaikan agar hasil prediksinya menjadi lebih baik.

#### 1. Evaluasi *Confusion Matrix*

*Confusion matrix* adalah tabel yang digunakan untuk melihat seberapa baik model klasifikasi bekerja. Tabel ini biasanya dipakai untuk mengevaluasi model yang membagi data ke dalam satu kelas tertentu, di mana setiap data hanya masuk ke satu kelas pada satu waktu. (Krstinić et al., 2020). Matriks ini menyajikan hasil klasifikasi model dalam empat kategori: *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Keempat metrik ini membantu mengukur berbagai aspek akurasi model, seperti presisi, *recall*, dan *F1-score*. Dalam konteks deteksi *hoax*, *confusion matrix* sangat penting untuk mengetahui sejauh mana model dapat membedakan antara informasi *hoax* dan non-*hoax*.

Perhitungan Metode Evaluasi:

a. Akurasi

Pengukuran ini dilakukan dengan membandingkan jumlah prediksi yang akurat terhadap keseluruhan data yang tersedia.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

b. Presisi

Pengukuran ini menilai tingkat akurasi dari prediksi positif dengan menghitung proporsi prediksi positif yang terbukti benar.

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (5)$$

c. Recall

Pengukuran ini bertujuan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data yang termasuk kategori positif.

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (6)$$

Penjelasan :

a. TP (*True Positive*)

Merupakan banyaknya data yang sebenarnya positif dan berhasil diprediksi dengan tepat sebagai positif.

b. TN (*True Negative*)

Mengacu pada jumlah data yang sebenarnya negatif dan berhasil diklasifikasikan secara akurat sebagai negatif.

c. FP (*False Positive*)

Menunjukkan banyaknya data yang sebenarnya negatif namun salah diprediksi sebagai positif.

d. FN (*False Negative*)

Jumlah data positif yang salah diprediksi sebagai negatif.

2. *F1 Score*

*F1 Score* adalah salah satu cara untuk menilai kinerja model dengan

menggabungkan nilai presisi dan *recall*. Nilai ini menunjukkan seberapa baik model dalam membuat prediksi yang benar dan menangkap data yang benar. *F1 Score* sangat berguna ketika data yang digunakan tidak seimbang antara kelas satu dengan yang lainnya. Nilai ini dihitung dari rata-rata harmonis antara presisi dan *recall*, dengan rumus sebagai berikut:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

Penjelasan Rumus :

1. *Precision* merupakan ukuran yang menunjukkan proporsi prediksi positif yang benar-benar tepat dari keseluruhan prediksi positif yang dihasilkan oleh model.
2. *Recall* mengindikasikan kemampuan model dalam mendeteksi seluruh data relevan dengan menghitung persentase data positif yang berhasil ditemukan dari total data positif sesungguhnya.
3. Rumus ini menggabungkan nilai *Precision* dan *Recall* menjadi satu skor harmonis, dengan memberikan bobot yang sama pada keduanya.

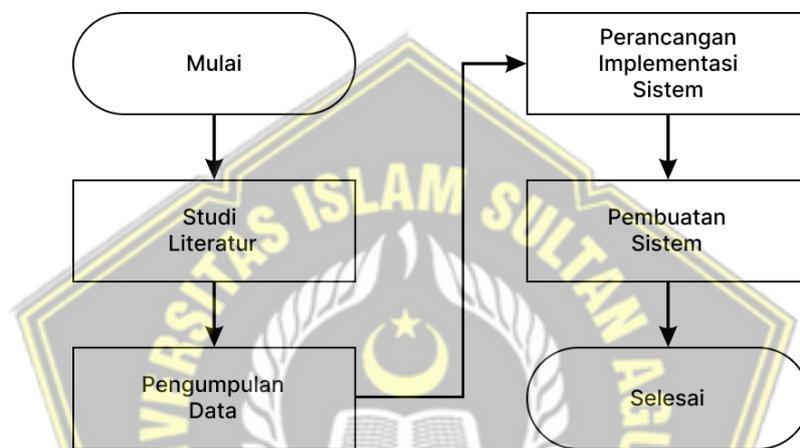
Metrik ini memberikan skor yang lebih akurat tentang kinerja model dengan mempertimbangkan kedua aspek tersebut. Tingginya nilai *F1 Score* mengindikasikan bahwa model mampu menjaga keseimbangan yang optimal antara pengurangan kesalahan klasifikasi positif dan negatif.

Dalam konteks klasifikasi *biner* atau multi-kelas, *F1 Score* sangat efektif karena memperhitungkan baik *false positives* maupun *false negatives*. *F1 Score* membantu mengukur keseimbangan antara presisi dan *recall*, memberikan hasil yang lebih informatif dibandingkan akurasi saja, terutama dalam kondisi data yang tidak seimbang (Powers, 2020). *F1 Score* juga membantu memperjelas hasil evaluasi model karena mempertimbangkan penyebaran hasil klasifikasi secara keseluruhan, sehingga dapat memberikan gambaran yang lebih akurat tentang kinerja sistem. (Yacouby & Axman, 2020).

## BAB III METODE PENELITIAN

### 3.1 Metode Penelitian

Penelitian ini akan membangun sistem deteksi hoax pada media sosial dengan menggunakan pendekatan *Natural Language Processing* (NLP) dan algoritma klasifikasi *Naive Bayes*. Alur penelitian tertera pada Gambar 3.1



Gambar 3.1 Flowchart Alur Penelitian

### 3.2 Studi Literatur

Studi literatur dilakukan dengan mengumpulkan dan mempelajari berbagai artikel, jurnal ilmiah sebelumnya yang berkaitan dengan penulisan tugas akhir dari daftar pustaka yang dikaji. Penelitian ini mengacu pada 14 jurnal akademik dari berbagai institusi dan penerbit seperti *Elsevier*, *IEEE Xplore*, *Springer*, dan *Google Scholar*. Referensi yang digunakan mencakup topik-topik seperti teknik pra-pemrosesan data teks, implementasi algoritma Naive Bayes untuk klasifikasi. Referensi ini menjadi dasar dalam penyusunan tugas akhir dan sangat membantu peneliti dalam merancang sistem deteksi hoax secara lebih terarah dan ilmiah.

### 3.3 Pengumpulan Data

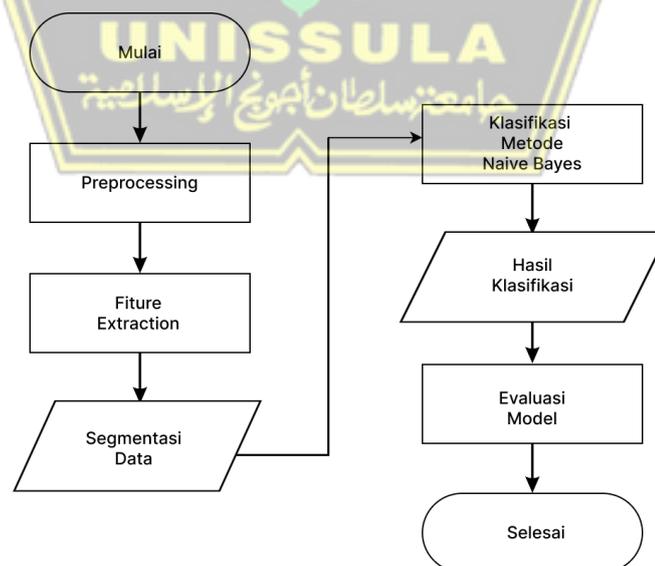
Peneliti menggunakan dataset *Truth Seeker Model Dataset* yang diperoleh dari platform Kaggle sebagai sumber data untuk penelitian ini. Dataset dapat diakses melalui link berikut :

<https://www.kaggle.com/datasets/sudishbasnet/truthseekertwitterdataset2023>.

Dataset ini berisi kumpulan teks dari media sosial tweeter yang telah diklasifikasikan ke dalam dua kategori, yaitu hoax dan non-hoax. Dataset ini terdiri dari 134.198 sampel teks yang telah melalui proses anotasi sebelumnya. Data disimpan dalam format CSV dan telah memiliki label yang menunjukkan apakah suatu teks termasuk hoax atau bukan. kategori pelabelannya yaitu 68.930 data non hoax dan 65.268 data hoax.

### 3.4 Perancangan Implementasi Sistem

Pada tahap ini, sistem deteksi hoax dirancang menggunakan pendekatan Natural Language Processing (NLP) dengan algoritma *Naive Bayes* sebagai metode klasifikasi. Perancangan model ini bertujuan untuk mengklasifikasikan teks dari media sosial *tweeter* ke dalam kategori hoax dan non-hoax. Alur sistem tertera pada Gambar 3.2



Gambar 3.2 Flowchart Alur Sistem

a. Tahap *Preprocessing*

Setelah data terkumpul, tahap pra-pemrosesan bertujuan untuk melakukan pembersihan serta penyiapan data teks agar layak dan optimal digunakan dalam tahapan klasifikasi selanjutnya.. Tahap ini mencakup beberapa langkah, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, *lematization*. Proses ini bertujuan untuk mengurangi *noise* dalam data agar model *Naive Bayes* dapat bekerja lebih optimal dalam mendeteksi *hoax* dan *non-hoax*.

b. *Feature Extraction*

Setelah preprocessing, data teks diproses pada tahap *feature extraction* untuk mengubahnya menjadi bentuk numerik yang dapat dipahami oleh algoritma *Naive Bayes*. Metode yang digunakan adalah TF-IDF (*Term Frequency–Inverse Document Frequency*), yaitu teknik pembobotan kata berdasarkan frekuensi kemunculan dan tingkat keumumannya dalam seluruh data. Hasilnya berupa vektor numerik yang merepresentasikan setiap dokumen, yang kemudian digunakan dalam proses klasifikasi. Tahap ini penting untuk meningkatkan akurasi deteksi hoaks dan non-hoaks.

c. Segmentasi Data

Setelah melalui tahap pra-pemrosesan, dataset dibagi ke dalam dua komponen utama, yakni data pelatihan (*data training*) dan data pengujian (*data testing*). Data pelatihan berfungsi untuk membangun dan melatih model *Naive Bayes*, sementara data pengujian dimanfaatkan untuk mengevaluasi kinerja model yang telah terbentuk. Proses pembagian data dilakukan dengan proporsi 80:20, di mana sebanyak 80% dari keseluruhan data dialokasikan sebagai data pelatihan, sedangkan sisanya sebesar 20% digunakan sebagai data pengujian. Strategi ini diterapkan guna memastikan kemampuan model dalam melakukan generalisasi terhadap data baru, serta menghindari kecenderungan model untuk sekadar mengingat data yang telah digunakan dalam pelatihan.

d. Klasifikasi Metode *Naive Bayes*

Pada tahapan ini, algoritma *Naive Bayes* diterapkan untuk melakukan proses klasifikasi terhadap teks dengan membedakannya ke dalam dua kelas, yakni

hoaks dan non-hoaks. Model ini bekerja dengan menghitung probabilitas suatu teks berdasarkan fitur-fitur yang diperoleh dari dataset. Proses klasifikasi diawali dengan ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), yang mengubah teks menjadi bentuk numerik agar dapat diproses oleh model. Setelah itu, model *Naive Bayes* dilatih menggunakan data pelatihan guna mempelajari dan mengidentifikasi pola-pola yang terdapat dalam teks. Setelah proses pelatihan selesai, model diuji menggunakan data uji untuk mengukur kemampuannya dalam mengklasifikasikan teks baru. Model kemudian memberikan hasil prediksi apakah sebuah teks termasuk dalam kategori *hoax* atau *non-hoax* berdasarkan pola yang telah dipelajari sebelumnya. Tahap ini bertujuan untuk memastikan sistem dapat mengenali dan membedakan berita *hoax* dari berita valid dengan tingkat akurasi yang optimal.

d. Hasil Klasifikasi

Sistem memberikan output berupa prediksi apakah suatu teks termasuk *hoax* atau *non-hoax* berdasarkan hasil klasifikasi menggunakan model *Naive Bayes*. Hasil klasifikasi ini digunakan untuk mengevaluasi sejauh mana model dapat mendeteksi *hoax* dengan benar.

e. Evaluasi Model

Tahapan evaluasi model dimaksudkan untuk menilai tingkat efektivitas kinerja model dalam mengidentifikasi konten *hoaks*. Proses evaluasi ini dilakukan dengan memanfaatkan dua metrik utama, yakni *Confusion Matrix* dan *F1 Score*.

### 3.5 Pembangunan Sistem

Tahap ini merupakan implementasi dari model deteksi *hoax* yang telah dirancang sebelumnya. Model *Naive Bayes* yang telah dilatih dan diuji kemudian disimpan dalam bentuk file untuk digunakan dalam sistem. Sistem ini dikembangkan menggunakan framework *Streamlit*, yang memungkinkan tampilan antarmuka sederhana dan interaktif bagi pengguna. Alur sistemnya bisa dilihat pada Gambar 3.3



Gambar 3.3 *Flowchart* Alur Kerja Sistem

Berikut adalah penjabaran *flowchart* berdasarkan tampilan *Streamlit* yang telah dibuat:

1. Mulai

Proses dimulai ketika pengguna mengakses aplikasi berbasis *Streamlit*.

2. Input Data

Pengguna memasukkan teks atau memilih dataset yang akan dianalisis. Kemudian sistem melakukan *preprocessing*, seperti pembersihan teks, tokenisasi, dan normalisasi agar data siap untuk diproses oleh model *Naïve Bayes*.

3. Klasifikasi Model Menggunakan *Naïve Bayes*

Sistem menggunakan algoritma *Naive Bayes* untuk mengklasifikasikan teks yang di input ke dalam dua kategori, yaitu *hoax* dan *non-hoax*.

4. Hasil Prediksi

Sistem akan menampilkan hasil prediksi secara langsung, yaitu apakah teks tersebut termasuk dalam kategori *Hoax* atau *Non-Hoax*. Selain itu, sistem juga menampilkan probabilitas *hoax* dan *non hoax* dari teks yang dimasukkan.

5. Selesai

Proses dapat diulang dari tahap input jika pengguna ingin menganalisis teks lain.

## BAB IV HASIL DAN ANALISIS PENELITIAN

### 4.1 Persiapan Data

Data yang digunakan dalam penelitian adalah Tweeter Dataset. Dataset ini diambil dari *Kaggle*. Dataset dapat diakses melalui link berikut : <https://www.kaggle.com/datasets/sudishbasnet/truthseekertwitterdataset2023>.

Dataset yang digunakan berjumlah 134.198 data, yang sudah memiliki label kategori *Hoax* dan *Non-Hoax*. Data disimpan dalam *file csv*, dengan jumlah data *non hoax* adalah 68.930 data dan data *hoax* adalah 65.268 data.

### 4.2 Preprocessing Data

#### 4.2.1 Cleaning

Tahap ini dilakukan untuk membersihkan data dari elemen-elemen yang tidak relevan dalam proses klasifikasi *hoax*, seperti tanda baca, angka, karakter khusus, serta *link* atau URL yang mungkin terdapat dalam teks. Selain itu, proses *cleaning* juga menghilangkan spasi berlebih dan karakter non-alfabet agar data lebih bersih dan siap untuk tahap pemrosesan selanjutnya. Berikut pada Tabel 4.1 merupakan contoh data sebelum dan sesudah dilakukan *cleaning*.

Tabel 4. 1 hasil data sesudah dan sebelum *cleaning*

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
@Burkles1 It is a fact and 100% true that:\n\n1. The vaccine will not prevent #COVID19 \n2. The vaccine is not guaranteed to keep you out of the hospital \n3. The vax will not prevent you from spreading it\n4. The vax is not guaranteed to prevent death. \n\nDo not rely on covid 19 vaccines.	fact true vaccine prevent vaccine guaranteed keep hospital vax prevent spreading vax guaranteed prevent death rely covid vaccine

<p>rmlubash @shawver_richard @TheBrandonMorse The Covid vaccines do nothing to prevent catching or spreading Covid so no benefit there. Also, by not receiving the vaccine, he'll be able to guarantee that he won't get any of the negative side effects. His is the most optimal approach.</p>	<p>covid vaccine nothing prevent catching spreading covid benefit also receiving vaccine hell able guarantee wont get negative side effect optimal approach</p>
<p>@therecount Can someone please help me understand mandating a "vaccine" that not only does NOT prevent infection, but may actually increase the chances of spreading the virus to others?</p>	<p>someone please help understand mandating vaccine prevent infection may actually increase chance spreading virus others</p>

#### 4.2.2 Case Folding

*Case folding* adalah proses mengubah semua huruf dalam teks menjadi huruf kecil (*lowercase*) agar lebih seragam dan memudahkan analisis. Tahap ini penting untuk memastikan bahwa model tidak membedakan kata berdasarkan kapitalisasi. Berikut pada Tabel 4.2 merupakan contoh data sebelum dan sesudah dilakukan *Case Folding*

Tabel 4. 2 hasil data sesudah dan sebelum *Case Folding*

Sebelum <i>Case Folding</i>	Setelah <i>Case Folding</i>
<p>THE SUPREME COURT is siding with super rich property owners and over poor struggling AMERICANS by blocking the eviction moratorium during a pandemic which isn't even over yet. ASSHOLES! get ready for</p>	<p>the supreme court is siding with super rich property owners and over poor struggling americans by blocking the eviction moratorium during a pandemic which isnt even over yet assholes get ready for more homeless</p>

more homeless people	people
I Have said this before, but it really is incredibly the way in which Afghanistan has completely crowded out the things that are actually affecting ordinary Americans in national media coverage, from COVID to the eviction moratorium to climate change	I have said this before but it really is incredibly the way in which afghanistan has completely crowded out the things that are actually affecting ordinary americans in national media coverage from covid to the eviction moratorium to climate change
There is not and never has been stimulus checks in the plan,Joey Biden has already stated Americans dont need stimulus checks, unemployment, eviction moratorium, or rental and utility help, thats why he and the democrats let it all expire. Welcome to a democrat ran America	there is not and never has been stimulus checks in the planjoey biden has already stated americans dont need stimulus checks unemployment eviction moratorium or rental and utility help thats why he and the democrats let it all expire welcome to a democrat ran america

#### 4.2.3 Tokenizing

Tokenisasi merupakan proses pemisahan teks menjadi satuan-satuan kecil yang disebut token. Tahapan ini berfungsi sebagai langkah awal dalam pemrosesan teks, yang memungkinkan model untuk menginterpretasikan serta mengelola informasi secara lebih efektif. Dalam algoritma *Naïve Bayes*, token yang dihasilkan digunakan sebagai fitur utama dalam proses analisis teks, di mana setiap kata diperlakukan sebagai unit informasi independen yang digunakan untuk menghitung probabilitas suatu teks tergolong dalam kategori hoaks atau non-hoaks. Pada tahap ini, setiap kalimat dalam dataset akan dipecah menjadi kata-kata individual. Proses ini penting karena model *Naïve Bayes* bekerja dengan prinsip probabilitas berdasarkan frekuensi kemunculan kata dalam kategori tertentu. Berikut pada Tabel 4.3 merupakan contoh data sebelum dan sesudah dilakukan *tokenizing*

Tabel 4. 3 hasil data sesudah dan sebelum *Tokenizing*

Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
biden blunders month update inflation delta mismanagement covid for kids abandoning americans in afghanistan arming the taliban border crisis breaking job growth abuse of power many exec orders t through reconciliation eviction moratoriumwhat did i miss	['biden', 'blunders', 'month', 'update', 'inflation', 'delta', 'mismanagement', 'covid', 'for', 'kids', 'abandoning', 'americans', 'in', 'afghanistan', 'arming', 'the', 'taliban', 'border', 'crisis', 'breaking', 'job', 'growth', 'abuse', 'of', 'power', 'many', 'exec', 'orders', 't', 'through', 'reconciliation', 'eviction', 'moratoriumwhat', 'did', 'i', 'miss']
the supreme court is siding with super rich property owners and over poor struggling americans by blocking the eviction moratorium during a pandemic which isnt even over yet assholes get ready for more homeless people	['the', 'supreme', 'court', 'is', 'siding', 'with', 'super', 'rich', 'property', 'owners', 'and', 'over', 'poor', 'struggling', 'americans', 'by', 'blocking', 'the', 'eviction', 'moratorium', 'during', 'a', 'pandemic', 'which', 'isnt', 'even', 'over', 'yet', 'assholes', 'get', 'ready', 'for', 'more', 'homeless', 'people']
i agree the confluence of events right now is unprecedented afghan disaster bombing overturning of the eviction moratorium collapse of biden admin credibility if images come out of americans being executed by taliban shtf	['i', 'agree', 'the', 'confluence', 'of', 'events', 'right', 'now', 'is', 'unprecedented', 'afghan', 'disaster', 'bombing', 'overturning', 'of', 'the', 'eviction', 'moratorium', 'collapse', 'of', 'biden', 'admin', 'credibility', 'if', 'images', 'come', 'out', 'of', 'americans', 'being', 'executed', 'by', 'taliban', 'shtf']

#### 4.2.4 *Stopword Removal*

*Stopword removal* merupakan proses eliminasi terhadap kata-kata umum yang tidak memberikan kontribusi bermakna dalam analisis teks. Kata-kata tersebut umumnya memiliki frekuensi kemunculan tinggi dalam bahasa sehari-hari, namun tidak relevan dalam menentukan klasifikasi teks sebagai hoaks maupun non-hoaks. Dalam penelitian ini, *stopword removal* dilakukan untuk mengurangi dimensi data dan meningkatkan efisiensi pemrosesan. Dengan mengeliminasi kata-kata yang tidak memiliki makna signifikan, model *Naïve Bayes* dapat memusatkan perhatian pada kata-kata yang memiliki pengaruh terhadap proses klasifikasi. Oleh karena itu, tahap *stopword removal* berperan dalam menyaring informasi yang relevan sehingga mendukung analisis lanjutan secara lebih efektif. Berikut pada Tabel 4.4 merupakan contoh data sebelum dan sesudah dilakukan *Stopword Removal*

Tabel 4. 4 hasil data sesudah dan sebelum *Stopword removal*

Sebelum <i>Stopword Removal</i>	Setelah <i>Stopword Removal</i>
the supreme court is siding with super rich property owners and over poor struggling americans by blocking the eviction moratorium during a pandemic which isnt even over yet assholes get ready for more homeless people	supreme court siding super rich property owners poor struggling americans blocking eviction moratorium pandemic isnt even yet assholes get ready homeless people
i agree the confluence of events right now is unprecedented afghan disaster bombing overturning of the eviction moratorium collapse of Biden admin credibility if images come out of americans being executed by taliban shtf	agree confluence events right now is unprecedented afghan disaster bombing overturning eviction moratorium collapse Biden admin credibility images come americans executed taliban shtf

ive said this before but it really is incredibly the way in which afghanistan has completely crowded out the things that are actually affecting ordinary americans in national media coverage from covid to the eviction moratorium to climate change	ive said really incredibly way afghanistan completely crowded things actually affecting ordinary americans national media coverage covid eviction moratorium climate change
---	---

#### 4.2.5 Lemmatization

*Lemmatization* adalah proses mengubah kata ke bentuk dasar atau lema-nya berdasarkan makna dan konteks gramatikal. Tidak seperti *stemming* yang hanya memotong akhiran kata secara kasar, *lemmatization* mempertimbangkan struktur bahasa dan memastikan bahwa hasil akhirnya merupakan kata yang valid dalam bahasa Indonesia. Proses ini penting dalam analisis teks karena membantu menyamakan berbagai bentuk kata yang memiliki arti dasar yang sama, seperti "berlari", "berlarian", dan "berlari-lari" menjadi "lari". Dalam penelitian ini, *lemmatization* diterapkan untuk memperbaiki kualitas representasi kata dalam data teks. Dengan mengonversi kata-kata ke bentuk dasarnya, model *Naïve Bayes* mampu mengidentifikasi pola secara lebih konsisten serta mengurangi variasi kata yang tidak perlu. Pendekatan ini secara signifikan berkontribusi terhadap peningkatan akurasi dalam proses klasifikasi hoaks. karena informasi yang diperoleh menjadi lebih terstruktur dan bermakna. *Lemmatization* juga turut mendukung proses reduksi dimensi dan normalisasi teks sebelum tahap pelatihan model dilakukan. Berikut pada Tabel 4.5 merupakan contoh data sebelum dan sesudah dilakukan *Lemmatization*.

Tabel 4. 5 hasil data sesudah dan sebelum *Lematization*

Sebelum <i>Lematization</i>	Setelah <i>Lematization</i>
supreme court siding super rich property owners poor struggling americans blocking eviction moratorium pandemic isnt even yet assholes get ready homeless people	supreme court side super rich property owner poor struggle americans block eviction moratorium pandemic be not even yet asshole get ready homeless people
agree confluence events right unprecedented afghan disaster bombing overturning eviction moratorium collapse biden admin credibility images come americans executed taliban shtf	agree confluence event right unprecedented afghan disaster bombing overturn eviction moratorium collapse biden admin credibility image come americans execute taliban shtf
ive said really incredibly way afghanistan completely crowded things actually affecting ordinary americans national media coverage covid eviction moratorium climate change	I ve say really incredibly way afghanistan completely crowd thing actually affect ordinary americans national medium coverage covid eviction moratorium climate change

### 4.3 *Feature Extraction*

#### 4.3.1 *Term Frequency-Inverse Document Frequency (TF-IDF)*

*Term Frequency-Inverse Document Frequency (TF-IDF)* merupakan suatu metode ekstraksi fitur yang berfungsi untuk merepresentasikan data teks dalam bentuk numerik, sehingga dapat diolah oleh model klasifikasi. Teknik ini menetapkan bobot pada setiap kata dalam suatu dokumen berdasarkan dua komponen utama, yaitu:

- *Term Frequency (TF)*: Mengukur seberapa sering suatu kata muncul dalam sebuah dokumen dibandingkan dengan panjang dokumen tersebut. Kata yang sering muncul akan memiliki nilai TF yang lebih tinggi.

- *Inverse Document Frequency (IDF)*: Mengukur seberapa unik suatu kata dalam keseluruhan kumpulan dokumen. Kata yang muncul di banyak dokumen akan memiliki nilai IDF yang lebih rendah, karena dianggap kurang bermakna dalam membedakan dokumen satu dengan yang lain.

Melalui penggabungan kedua komponen tersebut, TF-IDF memberikan penekanan bobot yang lebih tinggi pada kata-kata yang memiliki informasi penting dalam suatu dokumen namun jarang muncul di keseluruhan dataset. Hal ini membantu model *Naïve Bayes* dalam mengenali pola teks dan meningkatkan akurasi dalam mengklasifikasikan teks sebagai *hoax* atau *non-hoax*. Pada tahap ini, teks yang telah melalui preprocessing (*cleansing, case folding, tokenization, dan stopword removal*) diubah menjadi representasi numerik menggunakan TF-IDF. Hasilnya berupa matriks vektor dengan setiap fitur mewakili bobot kata dalam dokumen, yang kemudian digunakan sebagai input dalam proses klasifikasi. Berikut pada Tabel 4.6 merupakan contoh data sebelum dan sesudah dilakukan TF IDF.

Tabel 4. 6 hasil data sesudah dan sebelum TF IDF

Sebelum TF IDF	Setelah TF IDF
supreme court side super rich property	homeless people 0.316193
owner poor struggle americans block	struggle 0.286299
eviction moratorium pandemic be not	ashole 0.269346
even yet asshole get ready homeless	block 0.255282
people	ready 0.250053
	super 0.248590
	homeless 0.234693
	Evection 0.230972
	moratorium
	owner 0.230900
	eviction 0.229011
	moratorium 0.226706
	poor 0.210368
	property 0.206662
	rich 0.205050

	Supreme court	0.201043
	supreme	0.199171
	pandemic	0.193917
	court	0.177602
	people	0.110396
agree confluence event right	unprecedented	0.280236
unprecedented afghan disaster	credibility	0.278540
bombing overturn eviction moratorium	Bombing	0.272859
collapse biden admin credibility image	Collapse	0.269550
come americans execute taliban shtf	Biden admin	0.266542
	Overturn	0.257320
	Image	0.251730
	Disaster	0.246165
	Afghan	0.244946
	Admin	0.223482
	Event	0.219029
	eviction moratorium	0.219029
	Eviction	0.217169
	Moratorium	0.214983
	Taliban	0.201317
	Agree	0.196010
	Come	0.164770
	Right	0.135422
	biden	0.108831

#### 4.4 Optimasi model

##### 4.4.1 Split Data

Pada tahapan ini, dataset yang telah melewati proses pra-pemrosesan selanjutnya dipisahkan ke dalam dua subset utama, yakni data pelatihan (*training data*) dan data pengujian (*testing data*). Pembagian tersebut dilakukan dengan

proporsi 80:20, di mana sebanyak 80% dari keseluruhan data dimanfaatkan untuk melatih model, sedangkan 20% sisanya digunakan untuk mengevaluasi kinerja model yang telah dibangun. Dengan jumlah total dataset sebanyak 134.198 data, pembagian dilakukan dengan alokasi 107.358 data (80%) sebagai data pelatihan dan 26.840 data (20%) sebagai data pengujian. Pembagian tersebut bertujuan untuk memungkinkan model mempelajari pola-pola dari data pelatihan sekaligus menguji kemampuannya dalam mengklasifikasikan teks hoaks dan non-hoaks pada data yang belum pernah ditemui sebelumnya. Dengan demikian, diharapkan model memiliki kemampuan generalisasi yang memadai.

#### 4.4.2 Training Model *Naive Bayes*

Pada tahap ini, proses pelatihan model dilakukan menggunakan algoritma *Naive Bayes*, yang merupakan salah satu metode klasifikasi berbasis probabilitik. Model ini dipilih karena memiliki kinerja yang baik untuk data teks dan mampu menangani jumlah data yang besar dengan efisien. Sebelum proses pelatihan dilakukan, data teks yang telah melalui tahap *preprocessing* (*cleansing, case folding, tokenization, stopword removal, dan lematization*) diubah menjadi bentuk representasi numerik dengan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Representasi ini memungkinkan model untuk mengenali bobot kata berdasarkan frekuensi dan tingkat keunikannya dalam dokumen. Data yang digunakan untuk pelatihan adalah 80% dari total dataset (107.358 data), sedangkan 20% sisanya (26.840 data) digunakan untuk pengujian. Proses pelatihan dilakukan tanpa pengaturan *hyperparameter* tambahan, karena algoritma *Naive Bayes* yang digunakan sudah cukup optimal dengan pengaturan default dari pustaka *Scikit-learn*. Selama proses pelatihan, model mempelajari pola kemunculan kata dalam kategori *hoax* dan *non-hoax*, sehingga mampu melakukan klasifikasi terhadap data baru berdasarkan probabilitas kemunculan kata-kata dalam masing-masing kategori. Setelah pelatihan selesai, model kemudian digunakan untuk melakukan prediksi pada data uji.

## 4.5 Hasil

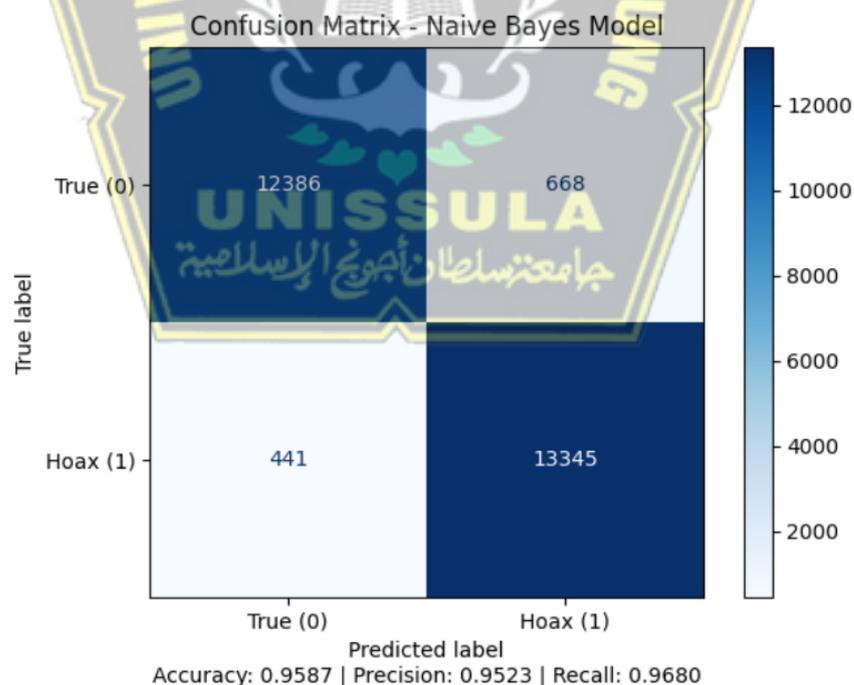
Setelah model *Naive Bayes* dilatih menggunakan data yang telah melalui tahapan *preprocessing* dan *feature extraction*, dilakukan pengujian terhadap data uji untuk melihat performa model dalam mengklasifikasikan teks ke dalam dua kategori, yaitu *hoax* dan *non-hoax*. Evaluasi model dilakukan dengan menggunakan dua metrik utama, yaitu *Confusion Matrix*, dan *F1 Score*.

### 1. *Confusion Matrix*

*Confusion Matrix* digunakan untuk menggambarkan kinerja model dalam membedakan antara data *hoax* dan *non-hoax* berdasarkan hasil prediksi terhadap data uji. Matriks ini menunjukkan jumlah prediksi benar dan salah untuk setiap kelas, yaitu:

- *True Negative* (TN) : 12.386 data *true* terdeteksi sebagai *True*
- *True Positive* (TP) : 13.345 data *hoax* terdeteksi sebagai *hoax*
- *False Positive* (FP) : 668 data *true* salah diklasifikasikan sebagai *hoax*
- *False Negative* (FN) : 441 data *hoax* salah diklasifikasikan sebagai *true*

Visualisasi *confusion matrix* dapat dilihat pada Gambar 4.2 berikut:



Gambar 4. 1 Garafik *Confusion Matrix*

Model *Naive Bayes* yang telah menjalani proses pelatihan memperlihatkan performa yang sangat baik dalam klasifikasi teks hoaks dan non-hoaks. Penilaian dilakukan dengan menggunakan metrik akurasi, *presisi*, dan *recall*. Berdasarkan hasil evaluasi tersebut, diperoleh nilai-nilai sebagai berikut:

- Akurasi sebesar 95,87%, yang menunjukkan bahwa sebagian besar prediksi model sesuai dengan label sebenarnya.
- Presisi sebesar 95,23%, mengindikasikan bahwa sebagian besar data yang diprediksi sebagai hoax memang benar-benar merupakan hoax.
- *Recall* sebesar 96,80%, yang berarti model mampu mengenali hampir semua teks hoax yang ada dalam data uji.

Nilai akurasi yang tinggi menunjukkan bahwa sebagian besar prediksi model benar. Presisi yang tinggi berarti sebagian besar data yang diklasifikasikan sebagai *hoax* memang benar-benar *hoax*. Sementara *recall* yang tinggi menunjukkan bahwa model mampu menemukan hampir semua data *hoax* yang ada dalam data uji. Dengan hasil ini, dapat disimpulkan bahwa *model Naive Bayes* sangat efektif dalam mendeteksi teks *hoax*, meskipun masih terdapat kesalahan kecil dalam klasifikasi berupa *false positive* dan *false negative*. Namun, tingkat kesalahan ini tergolong rendah dan masih dalam batas yang dapat diterima.

## 2. *F1 Score*

Selain *confusion matrix*, metrik lain yang digunakan untuk mengevaluasi performa model adalah *F1-Score*. Metrik ini merupakan rata-rata harmonis dari presisi dan *recall*, sehingga memberikan gambaran yang lebih seimbang antara keduanya, terutama pada data yang tidak seimbang. Hasil evaluasi menunjukkan bahwa model memiliki nilai *F1-Score* sebesar 96,01%, yang mengindikasikan keseimbangan yang baik antara kemampuan model dalam mengenali teks *hoax* (*recall*) dan ketepatannya dalam mengklasifikasikan teks sebagai *hoax* (presisi). Nilai *F1-Score* yang tinggi ini menunjukkan bahwa model *Naive Bayes* tidak hanya mampu mengenali sebagian besar teks *hoax*, tetapi juga jarang salah dalam mengklasifikasikan teks *non-hoax* sebagai *hoax*.

Dengan demikian, model ini dapat diandalkan untuk mendeteksi *hoax* secara konsisten dan akurat.

## 4.6 Tampilan *Streamlit*

### 4.6.1 Tampilan Analisis Teks

Tab Analisis Teks pada aplikasi *Streamlit* yang dikembangkan memungkinkan pengguna untuk melakukan deteksi *hoax* secara langsung dari teks yang dimasukkan. Pengguna dapat mengetikkan atau menyalin teks ke dalam kotak input yang telah disediakan. Setelah teks dimasukkan, pengguna cukup menekan tombol "Prediksi" untuk memproses teks tersebut menggunakan model *Naive Bayes* yang telah dilatih sebelumnya. Sistem akan menampilkan hasil klasifikasi apakah teks yang dimasukkan termasuk ke dalam kategori *Hoax* atau Bukan *Hoax*. Selain hasil klasifikasi, sistem juga memberikan probabilitas hasil prediksi untuk memberikan gambaran tingkat keyakinan model terhadap hasil yang diberikan. Tampilan ini memudahkan pengguna dalam menganalisis informasi secara instan, serta dapat berperan sebagai alat bantu dalam menangkal penyebaran informasi palsu. Tampilan analisis teks bisa dilihat pada Gambar 4.3



Gambar 4. 2 Tampilan Deteksi *Hoax*

Setelah memasukkan teks tunggu *running* hingga keluar hasil prediksi seperti pada gambar 4.4

Masukkan teks:

vaccine prevent getting covid spreading others ingests lessens effect like brotherinlaw side effect vaccine brotherinlaw got second shot he dealing leg problem

Prediksi

✓ Hasil: BUKAN HOAX

Probabilitas:

- ▶ Hoax : 0.63%
- ▶ Benar : 99.37%

Gambar 4. 3 Tampilan Hasil Deteksi Bukan Hoax

Masukkan teks:

THE SUPREME COURT is siding with super rich property owners and over poor struggling AMERICANS by blocking the eviction moratorium during a pandemic which isn't even over yet. ASSHOLES! get ready for more homeless people!

Prediksi

🚫 Hasil: HOAX

Probabilitas:

- ▶ Hoax : 98.75%
- ▶ Benar : 1.25%

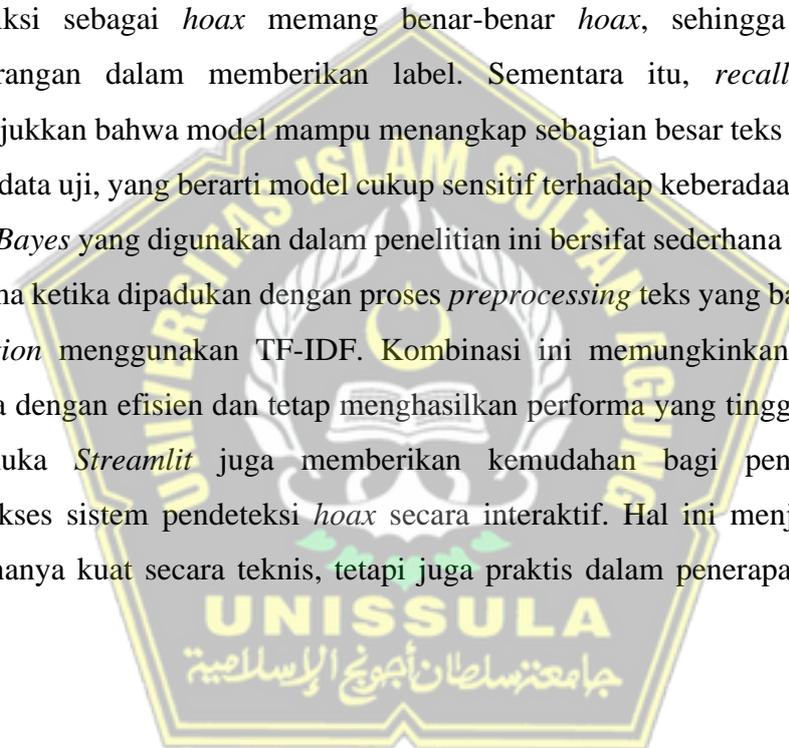
Gambar 4. 4 Tampilan Hasil Deteksi Hoax

#### 4.7 Pembahasan

Model yang dikembangkan untuk deteksi hoaks di media sosial dengan menggunakan metode *Naive Bayes* menunjukkan kinerja yang sangat baik. Hasil evaluasi menunjukkan bahwa model mencapai tingkat akurasi sebesar 95,87%, presisi sebesar 95,23%, *recall* sebesar 96,80%, serta *F1-score* sebesar 96,01%.

Nilai-nilai tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi dengan tingkat kesalahan yang sangat rendah. Selain itu, analisis melalui *confusion matrix* menunjukkan bahwa model berhasil mengklasifikasikan mayoritas teks secara akurat pada kedua kelas, yaitu hoaks dan non-hoaks.. Namun, masih terdapat sedikit kesalahan prediksi, yaitu dalam bentuk *False Positive* (FP) dan *False Negative* (FN). Kesalahan ini tergolong kecil dan masih dalam batas wajar untuk sistem klasifikasi berbasis machine learning.

Presisi yang tinggi mengindikasikan bahwa sebagian besar teks yang diprediksi sebagai *hoax* memang benar-benar *hoax*, sehingga model tidak sembarangan dalam memberikan label. Sementara itu, *recall* yang tinggi menunjukkan bahwa model mampu menangkap sebagian besar teks *hoax* yang ada dalam data uji, yang berarti model cukup sensitif terhadap keberadaan *hoax*. Model *Naive Bayes* yang digunakan dalam penelitian ini bersifat sederhana namun efektif, terutama ketika dipadukan dengan proses *preprocessing* teks yang baik dan *feature extraction* menggunakan TF-IDF. Kombinasi ini memungkinkan model untuk bekerja dengan efisien dan tetap menghasilkan performa yang tinggi. Penggunaan antarmuka *Streamlit* juga memberikan kemudahan bagi pengguna dalam mengakses sistem pendeteksi *hoax* secara interaktif. Hal ini menjadikan sistem tidak hanya kuat secara teknis, tetapi juga praktis dalam penerapannya di dunia nyata.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Penelitian ini bertujuan untuk membangun sistem deteksi hoaks di media sosial menggunakan *Natural Language Processing* (NLP) dan *Naïve Bayes*. Sistem yang dikembangkan dirancang untuk dapat mendeteksi berita apakah termasuk hoaks atau valid, dengan memanfaatkan tahapan *preprocessing* seperti *cleansing*, *case folding*, *tokenization*, *stopword removal*, dan *lemmatization*, serta proses ekstraksi fitur menggunakan TF-IDF. Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa model yang dikembangkan dalam penelitian ini menunjukkan performa yang sangat baik, dengan perolehan nilai evaluasi berupa akurasi sebesar 95,87%, presisi sebesar 95,23%, *recall* sebesar 96,80%, dan *F1-score* sebesar 96,01%. Selain itu, sistem telah diimplementasikan dalam bentuk aplikasi berbasis web menggunakan *Streamlit*, sehingga memudahkan pengguna dalam melakukan analisis teks secara langsung. Secara keseluruhan, metode dan sistem yang dikembangkan dalam penelitian ini dapat menjadi pendekatan awal yang cukup efektif dalam membantu mendeteksi dan meminimalkan penyebaran hoaks di media sosial.

Untuk bisa mendeteksi berita hoaks terbaru, maka harus ada pembanding atau dataset yang dimasukkan secara rutin dan otomatis ke dalam dataset yang akan di training. Setelah dataset di training ulang, maka model dapat mendeteksi berita hoaks terbaru apakah berita tersebut hoaks atau tidak.

#### 5.2 Saran

Berdasarkan hasil penelitian saran untuk penelitian selanjutnya adalah sebagai berikut:

1. Penggunaan dataset yang lebih besar dan beragam untuk meningkatkan generalisasi model, serta berasal dari berbagai platform media sosial.
2. Pengembangan Sistem ke Deteksi *Real-Time*, Sistem yang dikembangkan dalam penelitian ini bersifat *batch* dan belum bekerja secara *real-time*. Untuk

penerapan di dunia nyata, pengembangan sistem yang dapat mendeteksi *hoax* secara langsung saat informasi muncul akan sangat bermanfaat, khususnya dalam konteks pemantauan media sosial.



3. Pengembangan Sistem ke Deteksi *Real-Time*, Sistem yang dikembangkan dalam penelitian ini bersifat batch dan belum bekerja secara *real-time*. Untuk penerapan di dunia nyata, pengembangan sistem yang dapat mendeteksi *hoax* secara langsung saat informasi muncul akan sangat bermanfaat, khususnya dalam konteks pemantauan media sosial.



## DAFTAR PUSTAKA

- 'Ulhaq, D., Suarna, N., & Dwilestari, G. (2022). Klasifikasi Berita Kriminal Menggunakan Algoritma Naïve Bayes Berbasis PSO. *INFORMATICS FOR EDUCATORS AND PROFESSIONAL : Journal of Informatics*, 6(2), 116. <https://doi.org/10.51211/itbi.v6i2.1701>
- Agustina, N., Adrian, A., & Hermawati, M. (2022). Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media. *Faktor Exacta*, 14(4), 206. <https://doi.org/10.30998/faktorexacta.v14i4.11259>
- Ajay Rawat. (2020). A Review on Python Programming. *International Journal of Research in Engineering, Science and Management*, 3(12), 2581–5792.
- Et al., A. A. A. A. (2021). Detecting Fake News using Machine Learning: A Systematic Literature Review. *Psychology and Education Journal*, 58(1), 1932–1939. <https://doi.org/10.17762/pae.v58i1.1046>
- Gan, S., Shao, S., Chen, L., Yu, L., & Jiang, L. (2021). Adapting hidden naive bayes for text classification. *Mathematics*, 9(19), 1–14. <https://doi.org/10.3390/math9192378>
- Indrayuni, E., Nurhadi, A., & Kristiyanti, D. A. (2021). Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc. *Faktor Exacta*, 14(2), 64. <https://doi.org/10.30998/faktorexacta.v14i2.9697>
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). *Multi-label Classifier Performance Evaluation with Confusion Matrix*. 01–14. <https://doi.org/10.5121/csit.2020.100801>
- Manzoor, S. I., Singla, J., & Nikita. (2019). Fake news detection using machine learning approaches: A systematic review. *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, June*, 230–234. <https://doi.org/10.1109/ICOEI.2019.8862770>
- Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2521–2526. <https://doi.org/10.1073/pnas.1806781116>

- Powers, D. M. W. (2020). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. 37–63.  
<http://arxiv.org/abs/2010.16061>
- Sawicki, J., Ganzha, M., & Paprzycki, M. (2023). The State of the Art of Natural Language Processing—A Systematic Automated Review of NLP Literature Using NLP Techniques. In *Data Intelligence* (Vol. 5, Issue 3).  
[https://doi.org/10.1162/dint\\_a\\_00213](https://doi.org/10.1162/dint_a_00213)
- Strubell, E., & McCallum, A. (n.d.). *Energy and Policy Considerations for Modern Deep Learning Research*.
- Sukmawati, E. C., Suryaningrum, L., & Angelica, D. (2024). *SisInfo Klasifikasi Berita Palsu Menggunakan Model Bidirectional Encoder Representations From Transformers (BERT)* *SisInfo*. 6(2), 76–85.
- Transformers, I., & Processing, N. L. (2024). *Ensemble Techniques for Robust Fake News Detection* :
- Yacoub, R., & Axman, D. (2020). *Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models*. 79–91. <https://doi.org/10.18653/v1/2020.eval4nlp-1.9>
- Yunanto, R., Purfini, A. P., & Prabuwisesa, A. (2021). Survei Literatur: Deteksi Berita Palsu Menggunakan Pendekatan Deep Learning. *Jurnal Manajemen Informatika (JAMIKA)*, 11(2), 118–130.  
<https://doi.org/10.34010/jamika.v11i2.5362>