

**Analisis Kinerja Algoritma Naive Bayes, SVM dan
Random Forest dalam Deteksi Hoaks Berita (Studi
Kasus Bahasa Indonesia)**

Proposal Tugas Akhir

Oleh

**Muhammad Nurul Hakim
18222097**



**PROGRAM STUDI SISTEM DAN TEKNOLOGI INFORMASI
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG**

Desember 2025

LEMBAR PENGESAHAN

Analisis Kinerja Algoritma Naive Bayes, SVM dan Random Forest dalam Deteksi Hoaks Berita (Studi Kasus Bahasa Indonesia)

Proposal Tugas Akhir

Oleh

Muhammad Nurul Hakim
18222097

Program Studi Sistem dan Teknologi Informasi
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Proposal Tugas Akhir ini telah disetujui dan disahkan
di Bandung, pada tanggal 20 November 2025

Pembimbing

Dr. Riza Satria Perdana, S.T, M.T.

NIP.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR KODE

BAB I

PENDAHULUAN

I.1 Latar Belakang

Informasi pada era digital dewasa ini telah menjadi salah satu aset paling berharga bagi individu maupun organisasi. Nilainya tidak hanya terletak pada kelimpahan data yang dapat diakses, tetapi juga pada kualitas dan kebenarannya. Informasi yang akurat, relevan, dan tepat waktu menjadi fondasi bagi pengambilan keputusan yang efektif baik dalam konteks pribadi, bisnis, maupun kebijakan publik. Dengan demikian, kualitas informasi berbanding lurus dengan kualitas keputusan yang dihasilkan.

Namun, kemudahan akses terhadap informasi ini juga membawa konsekuensi yang tidak bisa diabaikan. Ekosistem digital yang semula diharapkan menjadi sumber pencerahan kini juga menjadi wadah bagi penyebaran berita palsu atau hoaks. Fenomena ini memunculkan apa yang disebut sebagai “polusi informasi”, yaitu situasi ketika berita palsu, misinformasi, dan disinformasi menyebar dengan kecepatan yang sama, bahkan lebih cepat, daripada informasi yang faktual. Kondisi ini mengancam integritas ekosistem informasi dan mengikis kepercayaan masyarakat terhadap institusi, media, maupun pakar yang seharusnya menjadi sumber kebenaran.

Dampak dari penyebaran hoaks tidak dapat dianggap ringan. Informasi yang menyesatkan dapat memicu kekacauan sosial, kesalahan pengambilan keputusan, hingga menimbulkan kerugian ekonomi dan politik. Masyarakat akhirnya dipaksa untuk terus-menerus memverifikasi setiap informasi yang mereka terima, menciptakan beban kognitif yang besar dan ketidakpastian yang berkepanjangan. Dalam konteks ini, permasalahan hoaks tidak lagi sekadar isu literasi digital, melainkan persoalan sistemik yang membutuhkan solusi komprehensif, termasuk melalui pendekatan teknologi.

Salah satu pendekatan yang dapat dieksplorasi adalah pemanfaatan Machine Learning (ML) untuk mendeteksi pola-pola tertentu yang membedakan berita hoaks dari berita yang valid. Algoritma Machine Learning klasik seperti Naïve Bayes, Support Vector Machine (SVM), dan Random Forest telah banyak digunakan dalam berbagai bidang klasifikasi teks, seperti deteksi spam, analisis sentimen, dan pengelompokan dokumen. Pendekatan ini memungkinkan sistem untuk belajar dari data teks berita, mengenali ciri linguistik atau pola tertentu yang sering muncul dalam berita hoaks, dan menggunakannya untuk mengklasifikasikan berita baru secara otomatis.

Meski penelitian di bidang deteksi hoaks telah berkembang pesat, sebagian besar masih berfokus pada analisis isi teks berita saja (content-based approach). Padahal, dimensi lain seperti kredibilitas sumber berita atau domain penyebar juga berpotensi memengaruhi tingkat keakuratan deteksi. Kondisi ini membuka ruang eksplorasi lebih lanjut untuk memahami sejauh mana algoritma Machine Learning klasik dapat digunakan secara efektif, baik dalam konteks analisis konten maupun kemungkinan pengayaan dengan informasi sumber berita (source-based approach), khususnya dalam bahasa Indonesia yang memiliki karakteristik linguistik berbeda dari bahasa global lainnya.

Oleh karena itu, penelitian ini diharapkan dapat menjadi langkah eksploratif dalam menganalisis kinerja algoritma Machine Learning klasik Naïve Bayes, SVM, dan Random Forest dalam mendeteksi hoaks berita berbahasa Indonesia. Hasilnya diharapkan memberikan pemahaman empiris mengenai sejauh mana pendekatan-pendekatan tersebut dapat dimanfaatkan dalam meningkatkan kualitas informasi di ekosistem digital nasional.

I.2 Rumusan Masalah

Rumusan Masalah berisi masalah utama yang dibahas dalam tugas akhir. Rumusan masalah yang baik memiliki struktur sebagai berikut:

1. Pokok persoalan dari kondisi atau situasi yang ada saat ini. Dengan kata lain, jelaskan kelemahan atau kekurangan dari kondisi, situasi, atau solusi yang dijelaskan pada latar belakang. Ini merupakan pokok rumusan masalah.
2. Elaborasi lebih lanjut urgensi penyelesaian masalah tersebut (mengapa penting untuk diselesaikan dan akibat yang dapat terjadi jika tidak diselesaikan).
3. Usulan singkat terkait dengan solusi yang ditawarkan untuk menyelesaikan

persoalan. Penting untuk diperhatikan bahwa persoalan yang dideskripsikan pada subbab ini akan dipertanggungjawabkan di bab Evaluasi (apakah terselesaikan atau tidak).

I.3 Tujuan

Tuliskan tujuan utama dan/atau tujuan detail yang akan dicapai dalam pelaksanaan tugas akhir. Fokuskan pada hasil akhir yang ingin diperoleh setelah tugas akhir diselesaikan, terkait dengan penyelesaian persoalan pada rumusan masalah. Penting untuk diperhatikan bahwa tujuan yang dideskripsikan pada subbab ini akan dipertanggungjawabkan di akhir pelaksanaan tugas akhir apakah tercapai atau tidak. Tuliskan kriteria keberhasilan tugas akhir ini.

I.4 Batasan Masalah

Tuliskan batasan-batasan yang diambil dalam pelaksanaan tugas akhir. Batasan ini dapat dihindari (bersifat opsional, tidak perlu ada) jika topik atau judul tugas akhir dibuat cukup spesifik.

I.5 Metodologi

Tuliskan semua tahapan yang akan dilalui selama pelaksanaan tugas akhir. Tahapan ini spesifik untuk menyelesaikan persoalan tugas akhir. Khusus untuk penyusunan proposal ini, jelaskan secara detail:

1. Tahapan investigasi pengumpulan fakta di latar belakang untuk merumuskan masalah.
2. Langkah-langkah pencarian, pengelompokan, dan penapisan literatur atau sumber informasi untuk mengumpulkan informasi yang relevan tentang topik yang diangkat, termasuk teori (konsep atau teori apa saja yang perlu dicari), hal-hal yang telah dicapai oleh orang lain (cara mencari dan kata kuncinya), dan berbagai informasi pendukung, untuk mencari solusi terhadap masalah yang dibahas. Gunakan metodologi yang tepat dalam menggali informasi dan dokumentasikan prosesnya (termasuk rekaman wawancara atau survei) di dalam Lampiran, termasuk tautan ke video atau foto. Hasil penggalian informasi ini akan dijelaskan secara sistematis di Bab II Studi Literatur.

BAB II

STUDI LITERATUR

II.1 Penulisan Gambar, Tabel, Rumus, dan Kode

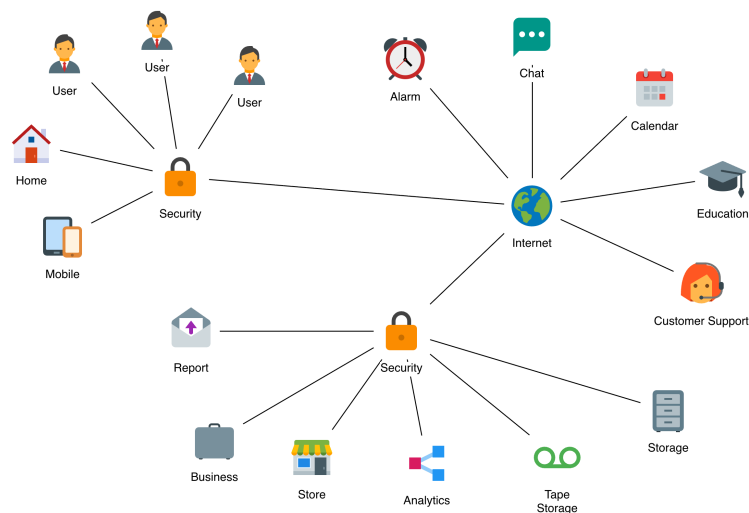
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

II.1.0.1 Gambar

Contoh gambar dapat dilihat pada Gambar ???. Gambar dan judulnya diposisikan di tengah. Nomor gambar tidak diakhiri tanda titik. Gambar tersebut dibuat menggunakan aplikasi draw.io dan disimpan ke format PNG setelah dengan zoom setting pada angka 300%. Ukuran gambar yang ditampilkan dapat diatur dengan mengubah nilai *width* dalam sintaks *includegraphics*.

Gambar umumnya tidak jelas atau kabur jika gambar tersebut:

- a. diperoleh dari hasil cropping pada suatu halaman buku atau situs web;
- b. hasil pembesaran gambar yang gambar aslinya sebenarnya berukuran kecil;
atau
- c. disimpan dalam resolusi kecil



Gambar II.1 Contoh gambar jaringan

Tabel II.1 Tabel harga bahan pokok

Nama	Satuan	Harga
Buku	Exemplar	25000
Komputer	Unit	2500000
Pensil	Buah	118900

Ketidakjelasan gambar ini dapat dilihat pada garis-garis diagram yang tidak tegas dan tulisan-tulisan dalam gambar yang tampak kabur dan kurang jelas terbaca.

Untuk mendapatkan gambar yang tidak kabur (blur), langkah-langkah berikut dapat digunakan:

- Gambar yang didapat di suatu pustaka atau referensi sebaiknya digambar ulang, misalnya menggunakan PowerPoint, Canva, Figma, draw.io, atau yang lainnya.
- Jika diagram atau ilustrasi digambar menggunakan draw.io, saat gambar disimpan ke format PNG atau JPG (export as), lakukan zoom ke minimal 300
- Jika diagram digambar dengan menggunakan PowerPoint, gambar dapat langsung di-copy-paste ke Word.

II.1.0.2 Tabel

Contoh tabel dapat dilihat pada Tabel ?? dan ??. Tabel dan judulnya dibuat rata kiri dan judul tabel diletakkan di atas tabel. Usahakan tabel dapat ditulis dalam satu halaman, tidak terpotong ke halaman berikutnya.

Tabel II.2 Tabel harga bahan sekunder

Nama	Satuan	Harga
Buku	Exemplar	25000
Komputer	Unit	2500000
Pensil	Buah	118900

II.1.0.3 Rumus

Contoh rumus matematika dapat ditulis seperti pada Persamaan ?? di bawah ini. Penomoran persamaan diletakkan di sebelah kanan, dan rumus ditulis dalam mode *display math*.

$$E = mc^2 \quad (\text{II.1})$$

II.1.1 Algoritma, Pseudocode, atau Kode

Contoh penulisan algoritma atau pseudocode dapat ditulis seperti pada Kode ?? di bawah ini. Gunakan paket *listings* untuk menulis source code dalam bahasa pemrograman tertentu, seperti pada Kode ??.

Kode II.1 Contoh pseudocode

```
ALGORITHM HelloWorld
    PRINT "Hello, World!"
END ALGORITHM
```

Kode II.2 Contoh source code Python

```
def hello_world():
    print("Hello, World!")
hello_world()
```

II.2 Beberapa Kesalahan Penulisan yang Sering Terjadi

II.2.1 Penggunaan kata "di mana" atau "dimana"

Banyak yang menuliskan kata "di mana" atau "dimana" sebagai pengganti kata "which" dalam bahasa Inggris. Padahal, penggunaan kata "di mana" atau "dimana" tidak tepat dalam konteks tersebut. Demikian juga untuk kata serupa, misalnya "yang mana". Kata "di mana" atau "dimana" ini harus diganti dengan kata lain, seperti "dengan", "tempat", "yang", dan sebagainya tergantung kalimatnya. Penjelasan lengkap dapat dilihat pada (BPBI).

II.2.2 Penggunaan kata "sedangkan" dan "sehingga"

Tabel II.3 Contoh penggunaan kata "sedangkan" dan "sehingga"

Penggunaan kata	Salah	Benar
sedangkan	Sedangkan sistem lama masih digunakan oleh banyak pengguna.	Sistem lama masih digunakan oleh banyak pengguna, sedangkan sistem baru belum siap.
sehingga	Sehingga sistem lama masih digunakan oleh banyak pengguna.	Sistem lama masih digunakan oleh banyak pengguna sehingga sistem baru belum siap.

Kata "sedangkan" dan "sehingga" adalah kata hubung atau konjungsi. Konjungsi adalah kata atau ungkapan yang menghubungkan satuan bahasa (kata, frasa, klausa, dan kalimat). Konjungsi dapat dibagi menjadi konjungsi intrakalimat dan antarkalimat. Kata "sedangkan" menghubungkan dua klausa yang bersifat kontrasif, sedangkan "sehingga" menghubungkan dua klausa yang bersifat kausal. Dalam ragam formal, kata hubung "sedangkan" dan "sehingga" hanya dapat digunakan sebagai konjungsi intrakalimat sehingga kedua konjungsi itu **tidak dapat diletakkan pada awal kalimat**. Selain itu, penggunaan kata "sedangkan" harus didahului oleh koma (,), sedangkan kata "sehingga" tidak perlu didahului oleh koma (,). Contoh penggunaan yang benar dan salah dapat dilihat pada Tabel ??.

II.2.3 Penggunaan Istilah yang Tidak Baku

Ada beberapa istilah yang sering digunakan dalam pembicaraan sehari-hari, tetapi tidak baku dalam penulisan ilmiah. Beberapa istilah tersebut antara lain:

1. analisa → analisis
2. eksisting atau existing → yang ada atau saat ini
3. bisnis proses → proses bisnis
4. user → pengguna
5. system → sistem
6. database → basis data
7. aktifitas → aktivitas
8. efektifitas → efektivitas

9. sosial media → media sosial

BAB III

ANALISIS MASALAH

III.1 Analisis Kondisi Saat Ini

Menurut Laudon **and** Laudon (2020), gambarkan terlebih dahulu model konseptual sistem yang ada saat ini. Model konseptual ini berisi berbagai komponen atau subsistem dan interaksi antarsubsistem tersebut. Setelah itu, berikan penjelasan tentang masalah yang ada pada sistem tersebut. Paragraf berikut berisi contoh penjabaran masalah sistem informasi fasilitas kesehatan untuk pasien (Pressman 2019).

III.2 Analisis Kebutuhan

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

III.2.1 Identifikasi Masalah Pengguna

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat

lacus vel est. Curabitur consectetur.

III.2.2 Kebutuhan Fungsional

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

III.2.3 Kebutuhan Nonfungsional

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

III.3 Analisis Pemilihan Solusi

III.3.1 Alternatif Solusi

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

III.3.2 Analisis Penentuan Solusi

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetur a, feugiat vitae, porttitor eu,

libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consetetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

BAB IV

DESAIN KONSEP SOLUSI

Ilustrasikan desain konsep solusi dalam bentuk model konseptual dan penjelasan secara ringkas, beserta perbedaannya dengan sistem saat ini. Ilustrasi harus dapat dibandingkan (*before and after*). Karena masih berupa proposal, bab ini hanya berisi gambar desain konsep solusi tersebut dan penjelasan perbandingannya dengan gambar sistem yang ada saat ini (yang tergambar di awal Bab III).

BAB V

RENCANA SELANJUTNYA

Jelaskan secara detail langkah-langkah rencana selanjutnya, hal-hal yang diperlukan atau akan disiapkan, dan risiko dan mitigasinya, yang meliputi:

1. Rencana implementasi, termasuk alat dan bahan yang diperlukan, lingkungan, konfigurasi, biaya, dan sebagainya.
2. Desain pengujian dan evaluasi, misalnya metode verifikasi dan validasi.
3. Analisis risiko dan mitigasi, misalnya tindakan selanjutnya jika ada yang tidak berjalan sesuai rencana.

DAFTAR PUSTAKA

Laudon, Kenneth C. **and** Jane P. Laudon. 2020. *Sistem Informasi Manajemen*. Jakarta: Pearson Education.

Pressman, Roger S. 2019. *Rekayasa Perangkat Lunak: Pendekatan Praktisi*. Yogyakarta: McGraw-Hill Education.

LAMPIRAN A: SOURCE CODE

LAMPIRAN B: HASIL SURVEI