

**ANALISIS PESAN HOAX PADA FACEBOOK  
MENGUNAKAN METODE NAÏVE BAYES CLASIFIER**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**ANANG WAHYU PRADANA**

**17.11.1132**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2024**

**ANALISIS PESAN HOAX PADA FACEBOOK  
MENGUNAKAN METODE NAÏVE BAYES CLASIFIER**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Program Studi Informatika



disusun oleh

**ANANG WAHYU PRADANA**

**17.11.1132**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS PESAN HOAX PADA FACEBOOK  
MENGUNAKAN METODE NAÏVE BAYES CLASIFIER**

yang disusun dan diajukan oleh

**Anang Wahyu Pradana**

**17.11.1132**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 23 Desember 2022

**Dosen Pembimbing,**

a.n



**Sumarni Adi, S.Kom., M.Cs.**

**NIK. 190302256**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**ANALISIS PESAN HOAX PADA FACEBOOK**  
**MENGGUNAKAN METODE NAÏVE BAYES CLASIFIER**

yang disusun dan diajukan oleh

**Anang Wahyu Pradana**

**17.11.1132**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 23 Desember 2022

**Susunan Dewan Penguji**

**Nama Penguji**

**Anna Baita, M.Kom**  
**NIK. 190302290**

**Arif Akbarul Huda, S.Si., M.Eng**  
**NIK. 190302287**

**Norhikmah, M.Kom**  
**NIK. 190302245**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 27 April 2024

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, S.Kom., M.Kom., Ph.D.**  
**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Anang Wahyu Pradana  
NIM : 17.11.1125

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Pesan Hoax Pada Facebook Menggunakan Metode Naïve Bayes  
Clasifier**

Dosen Pembimbing : Sumarni Adi, S.Kom., M.Cs.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Desember 2022

Yang Menyatakan,



Anang Wahyu Pradana

## HALAMAN PERSEMBAHAN

Segala puji syukur dipanjatkan ke hadirat Tuhan Yang Maha Esa, atas rahmat dan berkat-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul " Analisis Pesan Hoax Pada Facebook Menggunakan Metode Naïve Bayes Clasifier". Terima kasih atas berkah, nikmat, dan karunia yang diberikan-Nya. Penghargaan juga disampaikan atas segala bantuan, kekuatan, kesabaran, ilmu, serta dukungan dari orang-orang baik yang memberikan semangat dan doa, sehingga peneliti berhasil menyelesaikan skripsi ini. Untuk itu, rasa terima kasih peneliti diucapkan kepada:

1. Orang Tua dan seluruh anggota keluarga yang telah memberikan semangat yang besar dan kekuatan yang berarti selama ini dengan doa dan fasilitas yang telah diberikan dalam penyelesaian skripsi.
2. Ibu Sumarni Adi, S.Kom., M.Cs yang telah senantiasa membimbing dan menyemangati peneliti dalam menyelesaikan skripsi.
3. Teman-teman 17 Informatika 04 yang senantiasa memberikan dukungan dan membawa kenangan-kenangan berharga sepanjang perjalanan perkuliahan. Terima kasih atas pengalaman yang tak terlupakan.
4. Perusahaan Applemaxid (Applemax Official) yang telah memberikan fasilitas yang sangat berharga sehingga peneliti bisa menyelesaikan dengan lancar tanpa perangkat yang terhambat.

Akhir kata, semoga skripsi ini dapat bermanfaat bagi pembaca dan dapat memberikan kontribusi positif bagi pengembangan ilmu pengetahuan dan teknologi.

## KATA PENGANTAR

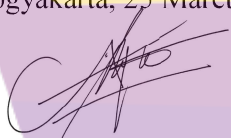
Alhamdulillah, peneliti mengucapkan puji dan syukur kepada Allah SWT yang telah memberikan limpahan karunia-Nya sehingga proses penyelesaian skripsi ini dapat berjalan dengan lancar. Skripsi ini disusun sebagai bagian dari syarat untuk menyelesaikan Program Studi Strata-1 Informatika di Universitas Amikom Yogyakarta.

Penulis ingin menyampaikan rasa terima kasih kepada semua pihak yang telah memberikan bantuan, bimbingan, dan dukungan, terutama kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Sumarni Adi, S.Kom., M.Cs selaku dosen pembimbing.
3. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta
4. Orang tua, wali, dan seluruh anggota keluarga peneliti.
5. Semua rekan mahasiswa Universitas Amikom Yogyakarta.
6. Seluruh teman-teman kelas 17 Informatika 04.

Peneliti menyadari bahwa masih terdapat kekurangan dalam penyusunan skripsi ini. Oleh karena itu, segala kritik dan saran akan diterima dengan baik. Semoga skripsi ini dapat memberikan manfaat dan pengetahuan kepada para pembaca.

Yogyakarta, 25 Maret 2024

  
Anang Wahyu Pradana

## DAFTAR ISI

<b>COVER</b> .....	<b>I</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>III</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>IV</b>
<b>HALAMAN PERNYATAAN KEASLIAN SKRIPSI</b> .....	<b>V</b>
<b>HALAMAN PERSEMBAHAN</b> .....	<b>VI</b>
<b>KATA PENGANTAR</b> .....	<b>VII</b>
<b>DAFTAR ISI</b> .....	<b>VIII</b>
<b>DAFTAR GAMBAR</b> .....	<b>X</b>
<b>DAFTAR TABEL</b> .....	<b>XI</b>
<b>INTISARI</b> .....	<b>XII</b>
<b>ABSTRACT</b> .....	<b>I</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
<b>1.1 LATAR BELAKANG</b> .....	<b>1</b>
<b>1.2 RUMUSAN MASALAH</b> .....	<b>4</b>
<b>1.3 BATASAN MASALAH</b> .....	<b>4</b>
<b>1.4 MAKSUD DAN TUJUAN PENELITIAN</b> .....	<b>5</b>
<b>1.5 MANFAAT PENELITIAN</b> .....	<b>5</b>
<b>1.6 METODE PENELITIAN</b> .....	<b>5</b>
<b>1.6.1 STUDI PUSTAKA</b> .....	<b>6</b>
<b>1.6.2 METODE CRAWLING DATA</b> .....	<b>6</b>
<b>1.6.3 METODE TESTING</b> .....	<b>7</b>
<b>1.7 SISTEMATIKA PENULISAN</b> .....	<b>7</b>
<b>BAB II DASAR TEORI</b> .....	<b>9</b>
<b>2.1 TINJAUAN UMUM</b> .....	<b>9</b>
<b>2.2 FACEBOOK</b> .....	<b>16</b>
<b>2.3 HOAKS</b> .....	<b>17</b>
<b>2.4 DATA MINING</b> .....	<b>17</b>



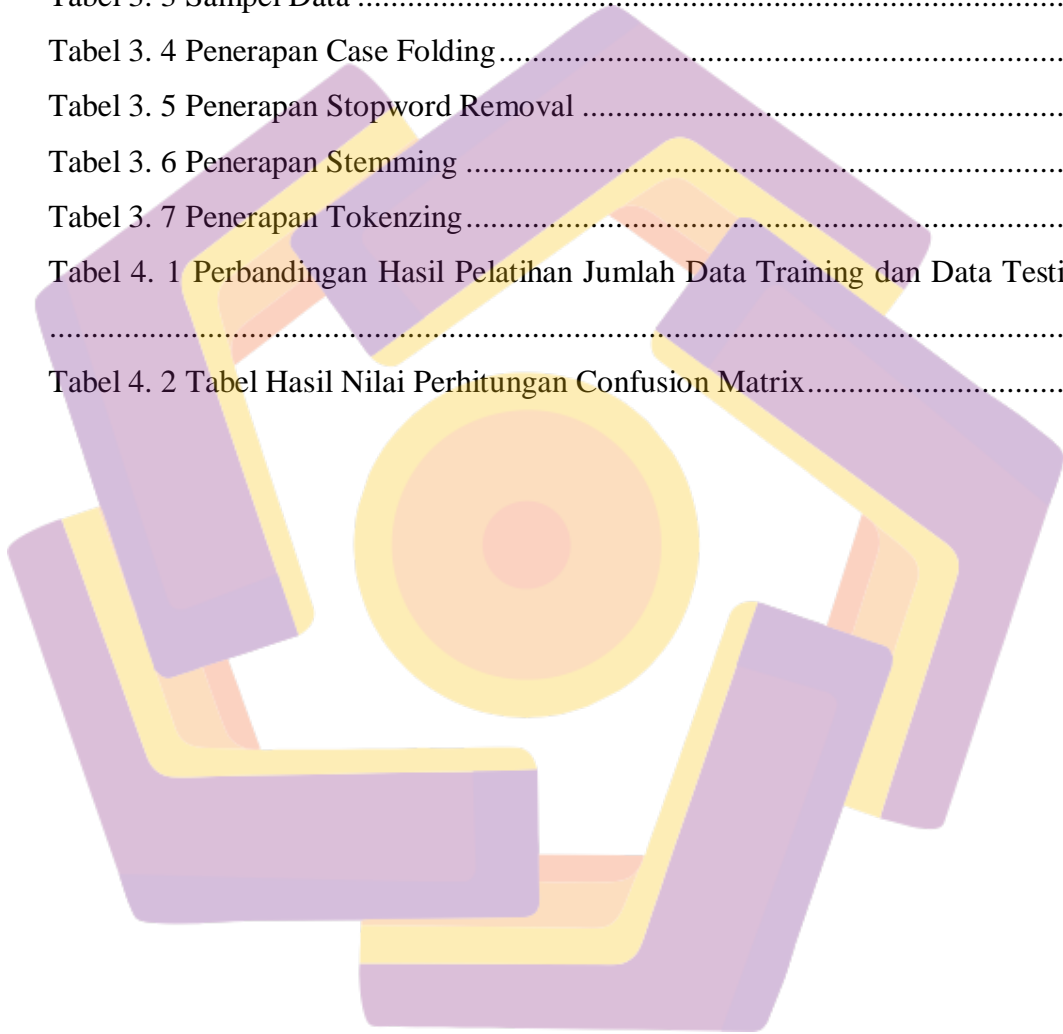
2.5	<i>TEXT MINING</i> .....	17
2.6	<i>TEXT PROCESSING</i> .....	18
2.7	<b>KLASIFIKASI</b> .....	19
2.8	<b>NAÏVE BAYES CLASIFIER</b> .....	19
2.9	<i>PHYTON</i> .....	21
2.10	<i>CONFUSION MATRIX</i> .....	21
<b>BAB III METODE PENELITIAN</b> .....		<b>24</b>
3.1	<b>TINJAUAN UMUM</b> .....	24
3.2	<b>TAHAPAN PENELITIAN</b> .....	25
3.3	<b>ANALISIS KEBUTUHAN</b> .....	26
3.4	<b>PENGUMPULAN DATA</b> .....	28
3.5	<i>PREPROCESSING DATA</i> .....	30
3.6	<b>PERANCANGAN DAN PELATIHAN MODEL NBC</b> .....	33
3.7	<b>TAHAPAN MODEL</b> .....	37
3.7.1	<i>DATA FLOW DIAGRAM</i> .....	37
3.8	<b>PENGUJIAN</b> .....	39
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....		<b>41</b>
4.1	<b>PERSIAPAN DATA</b> .....	41
4.2	<i>PREPROCESSING DATA</i> .....	43
4.3	<i>PEMBAGIAN DATA</i> .....	47
4.4	<b>PROSES PELATIHAN</b> .....	48
4.5	<b>HASIL PELATIHAN</b> .....	51
4.6	<b>PENGUJIAN <i>CONFUSION MATRIX</i></b> .....	53
<b>BAB V PENUTUP</b> .....		<b>57</b>
1.1	<b>KESIMPULAN</b> .....	57
1.2	<b>SARAN</b> .....	57
<b>DAFTAR PUSTAKA</b> .....		<b>58</b>

## DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Tahapan Penelitian .....	25
Gambar 3. 2 Data Flow Diagram Level 0 .....	38
Gambar 3. 3 Data Flow Diagram Level 1 .....	39
Gambar 4. 1 Source Code Memuat Data.....	41
Gambar 4. 2 Sampel Data Menggunakan Library Pandas .....	42
Gambar 4. 3 Sampel Seluruh Data.....	43
Gambar 4. 4 Source Code Case Folding .....	43
Gambar 4. 5 Hasil Case Folding .....	44
Gambar 4. 6 Source Code Stopword Removal.....	44
Gambar 4. 7 Hasil Stopword Removal.....	45
Gambar 4. 8 Source Code Stemming .....	46
Gambar 4. 9 Hasil Stemming.....	46
Gambar 4. 10 Source Code Tokenzing .....	47
Gambar 4. 11 Hasil Tokenzing .....	47
Gambar 4. 12 Source Code Pembagian Data.....	48
Gambar 4. 13 Source Code Bag of Words .....	48
Gambar 4. 14 Hasil Bag of Words.....	49
Gambar 4. 15 Source Code TF-IDF.....	50
Gambar 4. 16 Hasil TF-IDF.....	50
Gambar 4. 17 Source Code Pelatihan NBC.....	51
Gambar 4. 18 Source Code Menampilkan Hasil Prediksi.....	52
Gambar 4. 19 Hasil Prediksi Data Testing .....	53
Gambar 4. 20 Source Code Confusion Matrix .....	54
Gambar 4. 21 Hasil Confusion Matrix .....	54
Gambar 4. 22 Source Code Mencari Nilai dari Confusion Matrix .....	55
Gambar 4. 23 Hasil Perhitungan Confusion Matrix.....	55

## DAFTAR TABEL

Tabel 2. 1 Tabel Perbandingan Penelitian Terdahulu .....	11
Tabel 2. 3 Confusion Matrix.....	21
Tabel 3. 1 Spesifikasi Perangkat Keras .....	27
Tabel 3. 2 Spesifikasi Perangkat Lunak .....	27
Tabel 3. 3 Sampel Data .....	28
Tabel 3. 4 Penerapan Case Folding.....	31
Tabel 3. 5 Penerapan Stopword Removal .....	31
Tabel 3. 6 Penerapan Stemming .....	32
Tabel 3. 7 Penerapan Tokenzing.....	32
Tabel 4. 1 Perbandingan Hasil Pelatihan Jumlah Data Training dan Data Testing .....	51
Tabel 4. 2 Tabel Hasil Nilai Perhitungan Confusion Matrix.....	55



## INTISARI

Peningkatan teknologi Era Revolusi 4.0 menyebabkan penyampaian peristiwa yang sangat cepat di berbagai media sosial. Salah satu yang digunakan oleh masyarakat dunia yaitu *facebook*. *Facebook* yang menjadikan media sosial populer saat ini untuk memberikan informasi terkini yang dapat dikonsumsi oleh publik. Setiap pengguna bebas mengunggah, memposting dan berkomentar tanpa ada batasan. Akibatnya, terjadi berita hoax yang merambah secara masif dan terstruktur. Membuat berita dikonsumsi oleh penduduk Indonesia bahkan dunia tanpa mengetahui sumber aslinya.

Penelitian ini bertujuan untuk menganalisis pesan hoax atau kabar bohong yang terjadi di masyarakat pada media sosial *Facebook* supaya memberikan berita positif dan menyuguhkan informasi berharga yang disampaikan dari berbagai akun *Facebook*. Metode yang dilakukan yaitu pengambilan berupa komentar data teks dari postingan yang ditulis oleh berbagai akun *facebook* dengan topik “Perubahan Biaya Layanan BRI” kemudian diproses melalui *text preprocessing* melewati proses *case folding*, *stopword removal*, *stemming*, *tokenizing*. Selanjutnya dilakukan pembagian data antara data latih dan testing. Selanjutnya di klasifikasi oleh algoritme *Naïve Bayes Classifier* melewati bahasa pemrograman *Python*.

Dari Pengujian algoritme menggunakan *confusion matrix* mendapatkan hasil akurasi sebesar 95%, bahwasanya berita tersebut merupakan berita hoax dengan tingkat presisi 85,7%, mendapatkan nilai *recall* 92,3% dan nilai *F1-score* 88,8%.

**Kata Kunci : Text Mining, Pembobotan TF-IDF, Naïve Bayes Classifier**

## ABSTRACT

*Technological improvements in the Revolutionary Era 4.0 led to very fast delivery of events on various social media. One that is used by the world community is Facebook Facebook which makes social media popular today to provide up-to-date information that can be consumed by the public. Every user is free to upload, post and comment without any restrictions. As a result, hoax occurred which spread massively and in a structured manner. Making news consumed by the people of Indonesia and even the world without knowing the original source.*

*This study aims to analyze hoax messages or fake news that occur in society on Facebook social media in order to provide positive news and present valuable information conveyed from various Facebook accounts. The method used is taking the form of text data comments from posts written by various Facebook accounts with the topic "Changes in BRI Service Fees" then processed through text preprocessing through the process of case folding, stopword removal, stemming, tokenizing. Furthermore, the division of data between training and testing data is carried out. Then it is classified by the Naive Bayes Classifier algorithm through the Python programming language.*

*From testing the algorithm using the confusion matrix, an accuracy of 95% is obtained, in fact the news is hoax news with a precision level of 85.7%, a recall value of 92.3% and an F1-score of 88.8%*

**Keywords: Text Mining, TF-IDF Weighting, Naïve Bayes Classifier**