

# LAPORAN PROYEK NATURAL LANGUAGE PROCESSING (NLP)

**Analisis Sentimen dan Deteksi Spam pada Komentar  
TikTok Institusi Pendidikan**



## **Anggota Kelompok 4:**

1. ATHA DHAIFFATIN (1003230026)
2. AZRA MEUTHIA RINALDY ((1003250051))
3. ZIKRI FIRMANSYAH (1003230043)

**INSTITUT TEKNOLOGI TANGERANG SELATAN**

**2026**

## **DAFTAR ISI**

1. Bab 1: Pendahuluan
  - 1.1 Latar Belakang
  - 1.2 Tujuan
2. Bab 2: Deskripsi Data
  - 2.1 Sumber Data Prediksi (dataNew.csv)
  - 2.2 Data Latih (Training Data)
3. Bab 3: Pipeline NLP & Metodologi
  - 3.1 Preprocessing (Pra-pemrosesan Teks)
  - 3.2 Feature Extraction (Ekstraksi Fitur)
  - 3.3 Model Machine Learning
4. Bab 4: Implementasi Sistem
  - 4.1 Teknologi yang Digunakan
  - 4.2 Cara Menjalankan Aplikasi
  - 4.3 Fitur Aplikasi
5. Bab 5: Evaluasi dan Hasil
  - 5.1 Matrik Evaluasi
  - 5.2 Hasil Visualisasi
6. Bab 6: Kesimpulan
  - 6.1 Analisis Hasil
  - 6.2 Kesimpulan Akhir

## BAB 1: PENDAHULUAN

### 1.1 Latar Belakang

Media sosial, khususnya TikTok, telah menjadi sarana utama bagi institusi pendidikan untuk mempromosikan kegiatan dan berinteraksi dengan calon mahasiswa. Namun, tingginya volume interaksi seringkali membawa dua tantangan utama: sulitnya memetakan sentimen audiens secara manual dan banyaknya komentar *spam* (promosi tidak relevan) yang mengganggu informasi. Oleh karena itu, diperlukan sistem otomatis berbasis *Machine Learning* untuk menganalisis sentimen dan menyaring *spam*.

### 1.2 Tujuan

Tujuan dari proyek ini adalah:

1. Membangun aplikasi berbasis web (Flask) untuk analisis teks otomatis.
2. Mengklasifikasikan komentar menjadi sentimen Positif, Negatif, atau Netral.
3. Mendeteksi komentar yang bersifat Spam atau Bukan Spam.
4. Memvisualisasikan hasil analisis untuk pengambilan keputusan eksekutif.

## BAB 2: DESKRIPSI DATA

### 2.1 Sumber Data Prediksi (dataNew.csv)

Data yang digunakan sebagai objek analisis (data uji/prediksi) berasal dari ekspor aktivitas media sosial TikTok yang disimpan dalam file dataNew.csv.

- **Topik:** Konten seputar institusi pendidikan (ITTS, ITB), event kampus, dan promosi beasiswa.
- **Atribut Data:** video\_url, nama\_akun, caption, comments, likes, shares.
- **Fokus Analisis:** Kolom comments yang berisi opini publik. Karena satu baris data bisa memiliki banyak komentar yang dipisahkan koma, dilakukan proses *explode* data untuk menganalisis per komentar.

### 2.2 Data Latih (Training Data)

Karena dataNew.csv adalah data mentah tanpa label, model dilatih menggunakan dataset terlabel yang didefinisikan secara internal dalam sistem (main.py).

- **Sampel Sentimen:** Berisi kalimat dengan label positif, negatif, dan netral (Contoh: "Kampus ini keren banget" [Positif], "Biaya mahal fasilitas rusak" [Negatif]).
- **Sampel Spam:** Berisi kalimat dengan label spam dan bukan\_spam (Contoh: "Cek IG kita kak" [Spam], "Info pendaftaran kapan?" [Bukan Spam]).

## BAB 3: PIPELINE NLP & METODOLOGI

Proses analisis dilakukan mengikuti tahapan standar NLP sebagai berikut:

### 3.1 Preprocessing (Pra-pemrosesan Teks)

Sebelum masuk ke model, data teks mentah dibersihkan melalui fungsi preprocess\_text di main.py:

1. **Case Folding:** Mengubah semua huruf menjadi huruf kecil (*lowercase*).
2. **Cleaning:**\* Menghapus URL (<http>/<https>).
  - o Menghapus *mention* (@username).
  - o Menghapus karakter non-alfanumerik (tanda baca).
  - o Menghapus angka.
3. **Tokenization:** Memecah kalimat menjadi kata-kata individual.
4. **Stopword Removal:** Menghapus kata hubung umum bahasa Indonesia (seperti "dan", "yang", "di", "ke") menggunakan pustaka NLTK dan daftar kustom.

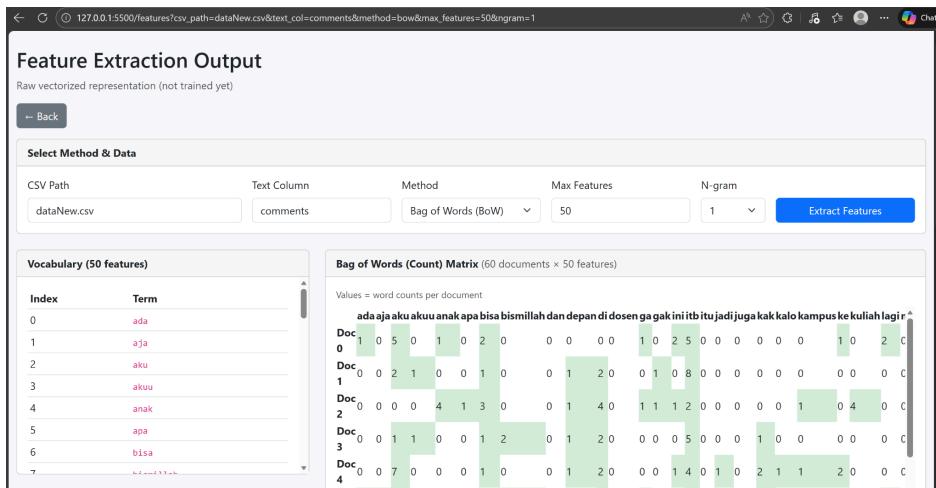
Processed (first 20 rows)			
	original	cleaned_segments	tokens_per_segment
	Tunggu aku ya ITB   tunggu aku 1731 hari lagiITB   boleh ujan ujanan ga mas   min doain besok gua tka semoga nilai bagus suapaya bisa masuk itb sekali lagi IITEEBEE   tahun ini anak ku daftar disini mudahan ada rejekinya bisa lolosFSRD menyal   Pertamina udah setoran kah ke ITB   Doain aku biar masuk sini   TUNGGU AKU YAA ITB   tunggu bentar lagi   ITB seindah ini masa gaada aku nya sih	tunggu aku ya itb   tunggu aku hari lagiitb   boleh ujan ujanan ga mas   min doain besok gua tka semoga nilai bagus suapaya bisa masuk itb sekali lagi iiteebbee   tahun ini anak ku daftar disini mudahan ada rejekinya bisa lolosfsrd menyal   pertamina udah setoran kah ke itb   doain aku biar masuk sini   tunggu aku yaa itb   tunggu bentar lagi   itb seindah ini masa gaada aku nya sih	tunggu ya itb   tunggu lagiitb   ujan ujanan ga mas   min doain besok gua tka semoga nilai bagus suapaya masuk itb iiteebbee   anak ku daftar mudahan rejekinya lolosfsrd menyal   pertamina udah setoran kah itb   doain biar masuk   tunggu yaa itb   tunggu bentar   itb seindah gaada nya sih
	ITB nerima SMK gak sih   bissmillah ITB   mau tanya di itb bisa ambil kelas karyawan tida   universitas impian   maba 20262027 siapa nichh yang di ITB   itb tunggu akuu yaaa   ITB tungguin aku ya   pertamaa   ITB Terima aku plsss tahun depan   ITB plis terima akuu	itb nerima smk gak sih   bissmillah itb   mau tanya di itb bisa ambil kelas karyawan tida   universitas impian   maba siapa nichh yang di itb   itb tunggu akuu yaaa   itb tungguin aku ya   pertamaa   itb terima aku plsss tahun depan   itb plis terima akuu	itb nerima smk gak sih   bissmillah itb   itb ambil kelas karyawan tida   universitas impian   maba nichh itb   itb tunggu akuu yaaa   itb tungguin ya   pertamaa   itb terima plsss   itb plis terima akuu
	ITB Adain jalur kasihan plis   pesona calon orang orang mapan   Jalur itb biasanya lwt apa yaa kaka soalnya tahun depan anak saya kuliah siapa tau bsa dpt rezeki kuliah di kampus ini   Anak saya alumni FSRD ITBTerima kasih ITBSekarang anak saya bekerja di LN   orang2 ga tidur tapi mukanya ttp ok jirr   YANG PERTAMA CAKEPNYA KELEWATAN   punya ruang auditorium sendiri   hanya orang kaya yg bisa masuk kuliah di sinimungkin orang seperti saya mah gak akan bisa masukin anak kuliah di sini   hahaha arek smk bisa nggak ya   Kereen bgt baju wisudanyasmoga tetep keren sampe anakku lulus nanti jgn aneh2 modelnya	itb adain jalur kasihan plis   pesona calon orang orang mapan   jalur itb biasanya lwt apa yaa kaka soalnya tahun depan anak saya kuliah siapa tau bsa dpt rezeki kuliah di kampus ini   anak saya alumni fsrd itbterima kasih itbsekarang anak saya bekerja di ln   orang ga tidur tapi mukanya ttp ok jirr   yang pertama cakepnya kelewan   punya ruang auditorium sendiri   hanya orang kaya yg bisa masuk kuliah di sinimungkin orang seperti saya mah gak akan bisa masukin anak kuliah di sini   hahaha arek smk bisa nggak ya   kereen bgt baju wisudanyasmoga tetep keren sampe anakku lulus nanti jgn aneh modelnya	itb adain jalur kasihan plis   pesona calon orang mapan   jalur itb lwt yaa kaka anak kuliah tau bsa dpt rezeki kuliah kampus   anak alumni fsrd itbterima kasih itbsekarang anak ln   orang ga tidur mukanya ttp ok jirr   cakepnya kelewan   ruang auditorium   orang kaya yg masuk kuliah sinimungkin orang mah gak masukin anak kuliah   hahaha arek smk nggak ya   kereen bgt baju wisudanyasmoga tetep keren sampe anakku lulus jgn aneh modelnya
	BISMILLAH ITB YA ALLAHHHH   bissmilah itbuiugm   di itb bisa ambil kelas karyawan   minn doain akuu yaa biar bisa kesituuu   ITB tungguin aku yah di situ   one day	bismillah itb ya allahhh   bissmilah itbuiugm   di itb bisa ambil kelas karyawan   minn doain akuu yaa biar bisa kesituuu   itb tungguin aku yah di situ   one day	bismillah itb ya allahhh   bissmilah itbuiugm   itb ambil kelas karyawan   minn doain akuu yaa biar bisa kesituuu   itb tungguin yah situ   one

\*untuk setiap postingan, data komentar di gabung menjadi satu teks dipisahkan dengan penanda |

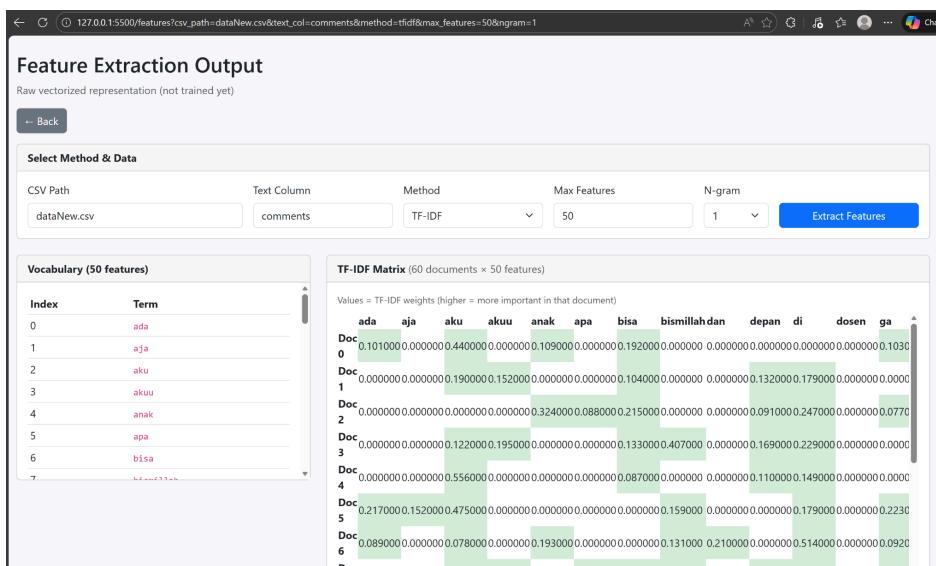
## 3.2 Feature Extraction (Ekstraksi Fitur)

Mengubah teks menjadi format numerik agar bisa diproses mesin.

- **Metode:** BoW (*Bag Of Words*)
- **Alasan:** Bag of Words (BoW) sederhana dan cepat karena hanya menghitung frekuensi kata tanpa memperhatikan urutan. Metode ini efektif sebagai baseline analisis teks, mampu menangkap kata dominan, dan menghasilkan vektor fitur yang siap dipakai untuk model klasifikasi seperti Naive Bayes atau Logistic Regression.



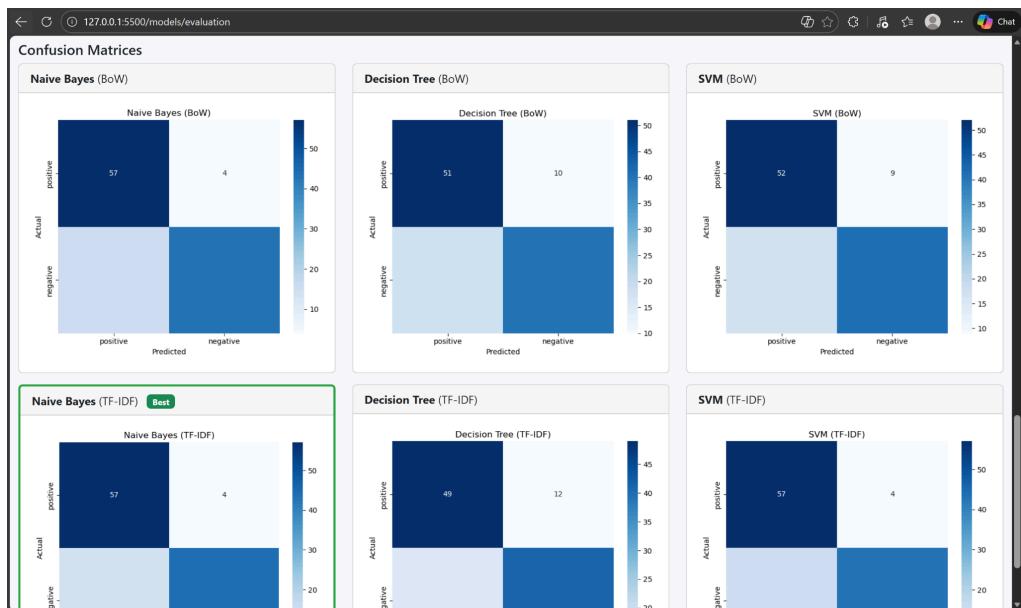
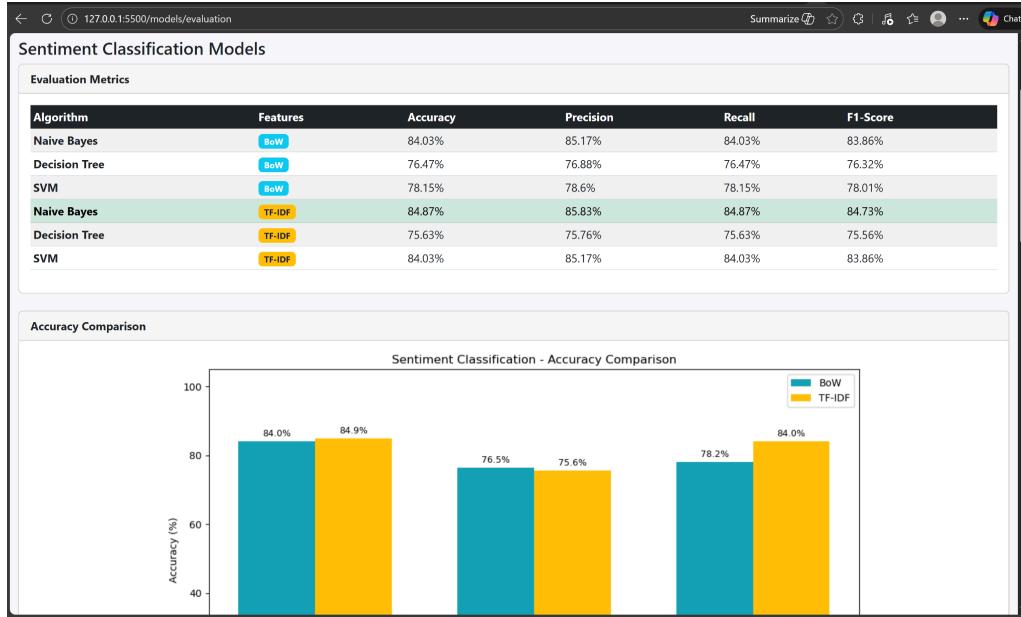
- **Metode:** TF-IDF (*Term Frequency-Inverse Document Frequency*).
- **Alasan:** Metode ini memberikan bobot lebih tinggi pada kata-kata unik yang penting untuk klasifikasi, bukan hanya menghitung frekuensi kemunculan kata semata.



### 3.3 Model Machine Learning

Kami membandingkan tiga algoritma untuk menentukan model terbaik:

1. **Support Vector Machine (LinearSVC)**: Efektif untuk ruang dimensi tinggi (teks).
2. **Naive Bayes (MultinomialNB)**: Cepat dan efisien untuk klasifikasi teks standar.
3. **Decision Tree**: Model berbasis pohon keputusan yang mudah diinterpretasi.



Berdasarkan evaluasi dalam kode, **Naive Bayes** dipilih sebagai model utama untuk melakukan prediksi pada data baru karena performanya yang stabil pada dataset teks kecil hingga menengah.

## BAB 4: IMPLEMENTASI SISTEM

### 4.1 Teknologi yang Digunakan

- **Bahasa:** Python 3.x
- **Framework Web:** Flask.
- **Library Data:** Pandas, NumPy.
- **Library NLP/ML:** NLTK, Scikit-learn.
- **Visualisasi:** Matplotlib, Seaborn.

### 4.2 Cara Menjalankan Aplikasi

1. Pastikan Python dan library terinstall (pip install flask pandas nltk scikit-learn matplotlib seaborn).
2. Letakkan file main.py dan dataNew.csv dalam satu folder.
3. Buka terminal/command prompt di folder tersebut.
4. Jalankan perintah:  
python main.py
5. Akses aplikasi melalui browser di alamat yang muncul (biasanya <http://127.0.0.1:5000>).

### 4.3 Fitur Aplikasi

- **Dashboard:** Menampilkan preview data dan menu untuk preprocess atau prediksi data

The screenshot shows a web-based user interface titled "NLP Preprocessing - Simple Web UI". At the top, there is a form with fields for "CSV Path" (set to "dataNew.csv"), "Text Column" (set to "comments"), and "Rows to show" (set to "60"). Below the form are several buttons: "Preprocess" (blue), "Show All Data" (disabled), "Feature Extraction (BoW/TF-IDF)" (disabled), "Model Comparison (All 12 Models)" (disabled), "Model Evaluation & Metrics" (disabled), and "Predict (Best Model)" (disabled). Underneath these buttons is a section titled "Data Preview (original)". This section contains a table with the following columns: "video\_url", "nama\_akun", "uploadDate", "likes", "shares", "caption", and "comments". The table displays three rows of data. Row 1: URL [itbofficial](#), upload date 10-27, likes 8055, shares 737, caption "Meromantisasi kampus tempat kami berteduh dan bercerita tentang hari ini 🌟", comments "Tunggu aku ya ITB | tunggu aku 1731 hari lagiITB | boleh ujan ujanan ga mas | min doain besok gua tka semoga nilai bagus suapaya bisa masuk itb sekali lagi IITEEBEE | tahun ini anak ku daftar disini mudahan ada rejekinya bisa lolosFSRD menyalah | Pertamina udah setoran kah ke ITB | Doain aku biar masuk sini | TUNGGU AKU YAA ITB | tunggu bentar lagi | ITB seindah ini masa gaada aku nya sih". Row 2: URL [itbofficial](#), upload date 10-25, likes 2778, shares 46, caption "wisudawan ITB punya playlist lagu tema pas skripsi, kalian apa? coba komen di bawah yaa 🙏", comments "ITB nerima SMK gak sih | bissmillah ITB | mau tanya di itb bisa ambil kelas karyawan tida | universitas impian | maba 20262027 siapa nich yang di ITB | itb tunggu aku yaaa | ITB tungguin aku ya | pertamaa | ITB Terima aku plsss tahun depan | ITB plis terima akuuu". Row 3: URL [itbofficial](#), upload date 10-23, likes 27.3K, shares 1699, caption "Lihat deh, para lulusan ITB hari ini. Langkah mereka penuh makna, matanya penuh harapan. Setiap toga punya kisah, setiap senyum menyimpan perjuangan. Mereka bukan cuma lulus, tapi tumbuh jadi sosok yang siap menebar manfaat untuk Tuhan, Bangsa, dan Alamater ❤️", comments "ITB Adain jalur kasihan plis | pesona calon orang orang mapan | Jalur itb biasanya lwt apa yaa kaka soalnya tahun depan anak saya kuliah siapa tau bsa dpt rezeki kuliah di kampus ini | Anak saya alumni FSRD ITBterima kasih ITBSekarang anak saya bekerja di LN | orang2 ga tidur tapi mukanya ttp ok jirr | YANG PERTAMA CAKEPNYA KELEWATAN | punya ruang auditorium sendiri | hanya orang kaya yg bisa masuk kuliah di sinimungkin orang seperti".

- **Analisis Data:** Tabel interaktif yang menampilkan komentar asli, hasil prediksi sentimen, hasil deteksi spam.

Naive Bayes Per Comment Classification

Page 1 / 15

post_index	comment	sentiment	spam
0	Tunggu aku ya ITB	neutral (50%)	not spam (58%)
0	tunggu aku 1731 hari lagiITB	neutral (50%)	spam (58%)
0	boleh ujian ujanan ga mas	neutral (50%)	not spam (51%)
0	min doain besok gua tka semoga nilai bagus suapaya bisa masuk itb sekali lagi IITEESEE	positive (100%)	not spam (100%)
0	tahun ini anak ku daftar disini mudahan ada rejekinya bisa lolosFSRD menala	positive (100%)	spam (96%)
0	Pertamina udah setoran kah ke ITB	neutral (50%)	not spam (51%)
0	Doain aku biar masuk sini	neutral (50%)	not spam (93%)
0	TUNGGU AKU YAA ITB	neutral (50%)	not spam (93%)
0	tunggu bentar lagi	neutral (50%)	not spam (93%)
0	ITB seindah ini masa gaada aku nya sih	neutral (50%)	not spam (100%)
1	ITB nerima SMK gak sih	negative (100%)	not spam (93%)
1	bissmillah ITB	neutral (50%)	not spam (51%)
1	mau tanya di itb bisa ambil kelas karyawan tida	neutral (50%)	not spam (96%)
1	universitas impian	positive (65%)	not spam (51%)
1	maba 20262027 siapa nichih yang di ITB	neutral (50%)	not spam (58%)
1	itb tunggu aku yaaa	neutral (50%)	not spam (51%)
1	ITB tungguin aku ya	neutral (50%)	not spam (58%)
1	pertamaaa	neutral (50%)	not spam (51%)
1	ITB Terima aku plsss tahun depan	positive (91%)	not spam (99%)

SVM Per-Comment Classification

Page 1 / 15

Summarize

post_index	comment	sentiment	spam
0	Tunggu aku ya ITB	positive (56%)	spam (56%)
0	tunggu aku 1731 hari lagiITB	positive (56%)	spam (74%)
0	boleh ujian ujanan ga mas	positive (56%)	not spam (51%)
0	min doain besok gua tka semoga nilai bagus suapaya bisa masuk itb sekali lagi IITEESEE	positive (99%)	not spam (90%)
0	tahun ini anak ku daftar disini mudahan ada rejekinya bisa lolosFSRD menala	positive (100%)	spam (81%)
0	Pertamina udah setoran kah ke ITB	positive (56%)	not spam (51%)
0	Doain aku biar masuk sini	positive (56%)	not spam (82%)
0	TUNGGU AKU YAA ITB	positive (56%)	not spam (82%)
0	tunggu bentar lagi	positive (56%)	not spam (83%)
0	ITB seindah ini masa gaada aku nya sih	negative (86%)	not spam (93%)
1	ITB nerima SMK gak sih	negative (100%)	not spam (79%)
1	bissmillah ITB	positive (56%)	not spam (51%)
1	mau tanya di itb bisa ambil kelas karyawan tida	positive (56%)	not spam (58%)
1	universitas impian	positive (56%)	not spam (51%)
1	maba 20262027 siapa nichih yang di ITB	positive (56%)	not spam (55%)
1	itb tunggu aku yaaa	positive (56%)	not spam (51%)
1	ITB tungguin aku ya	positive (56%)	not spam (56%)
1	pertamaaa	positive (56%)	not spam (51%)
1	ITB Terima aku plsss tahun depan	positive (89%)	not spam (91%)

- **Evaluasi Model:** Halaman khusus yang membandingkan akurasi antar algoritma.

## Model Comparison

All trained models: 3 Algorithms × 2 Feature Extractors × 2 Tasks = 12 Models

[← Back](#) [Test Models with Custom Text](#) [Model Evaluation & Metrics](#)

### Sentiment Classification Models

Algorithm	Feature Extraction	CV Accuracy (%)	Std Dev	Training Samples
Naive Bayes	BoW	84.17%	±14.48%	119
Decision Tree	BoW	76.67%	±18.41%	119
SVM	BoW	78.33%	±19.61%	119
Naive Bayes	TF-IDF	<b>85.0%</b>	±13.39%	<b>119</b>
Decision Tree	TF-IDF	75.83%	±19.4%	119
SVM	TF-IDF	84.17%	±14.48%	119

### Spam Detection Models

Algorithm	Feature Extraction	CV Accuracy (%)	Std Dev	Training Samples
Naive Bayes	BoW	67.06%	±9.55%	55
Decision Tree	BoW	50.88%	±4.7%	55
SVM	BoW	58.19%	±2.28%	55
Naive Bayes	TF-IDF	<b>67.06%</b>	±9.55%	<b>55</b>
Decision Tree	TF-IDF	50.88%	±4.7%	55
SVM	TF-IDF	58.19%	±2.28%	55

#### Best Sentiment Model

Naive Bayes + TF-IDF

Accuracy: **85.0%** (±13.39%)

#### Best Spam Model

Naive Bayes + BoW

Accuracy: **67.06%** (±9.55%)

#### Model Details

##### Algorithms

- **Naive Bayes:** Probabilistic classifier, fast, works well with text
- **Decision Tree:** Rule-based, interpretable, max\_depth=10
- **SVM:** Linear SVC, good for high-dimensional text data

##### Feature Extraction

- **BoW (Bag of Words):** Raw word counts using CountVectorizer
- **TF-IDF:** Term frequency-inverse document frequency weights

##### Evaluation

- **CV Accuracy:** 3-fold cross-validation mean
- **Std Dev:** Variation across folds
- Green row = best model for that task

- **Uji Prediksi Manual:** Form input untuk menguji kalimat baru secara *real-time*.

127.0.0.1:5500/predict

## Predict with Selected Model

Choose your preferred model or use the best performing model (pre-selected)

[← Back](#) [View All Model Metrics](#)

### Enter Text & Select Models

**Text to Classify**

kampusnya jelek

**SENTIMENT MODEL**

Naive Bayes + TF-IDF (F1: 84.73%) ★ BEST

★ indicates best performing model

**SPAM MODEL**

Naive Bayes + BoW (F1: 64.02%) ★ BEST

★ indicates best performing model

[Predict](#) [Clear Text](#)

### Prediction Results

<b>Sentiment Analysis</b> <span style="background-color: red; color: white; padding: 2px;">👎 NEGATIVE</span> Confidence: 89.0%	<b>Spam Detection</b> <span style="background-color: green; color: white; padding: 2px;">✓ NOT SPAM</span> Confidence: 51.0%
--	--

## BAB 5: EVALUASI DAN HASIL

### 5.1 Matrik Evaluasi

Model dievaluasi menggunakan teknik *Cross Validation* dengan metrik:

- **Accuracy:** Ketepatan keseluruhan prediksi.
- **Precision:** Tingkat ketepatan data positif yang diprediksi.
- **Recall:** Tingkat keberhasilan model menemukan kembali informasi.
- **F1-Score:** Rata-rata harmonik antara Precision dan Recall.

Sentiment Classification	Spam Detection	BoW vs TF-IDF			
Sentiment Classification Models					
Evaluation Metrics					
Algorithm	Features	Accuracy	Precision	Recall	F1-Score
Naive Bayes	BoW	84.03%	85.17%	84.03%	83.86%
Decision Tree	BoW	76.47%	76.88%	76.47%	76.32%
SVM	BoW	78.15%	78.6%	78.15%	78.01%
Naive Bayes	TF-IDF	84.87%	85.83%	84.87%	84.73%
Decision Tree	TF-IDF	75.63%	75.76%	75.63%	75.56%
SVM	TF-IDF	84.03%	85.17%	84.03%	83.86%



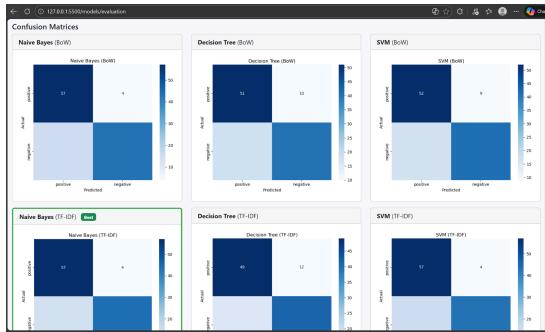
Sentiment Classification	Spam Detection	BoW vs TF-IDF			
Spam Detection Models					
Evaluation Metrics					
Algorithm					
Naive Bayes	BoW	67.27%	75.87%	67.27%	64.02%
Decision Tree	BoW	50.91%	50.77%	50.91%	50.42%
SVM	BoW	58.18%	58.55%	58.18%	57.39%
Naive Bayes	TF-IDF	67.27%	75.87%	67.27%	64.02%
Decision Tree	TF-IDF	50.91%	50.77%	50.91%	50.42%
SVM	TF-IDF	58.18%	58.2%	58.18%	58.02%



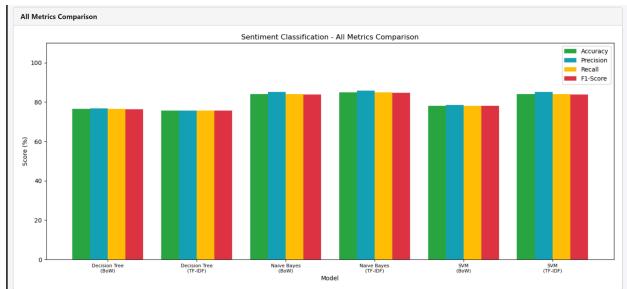
## 5.2 Hasil Visualisasi

Berdasarkan output dari main.py, sistem menghasilkan visualisasi berikut:

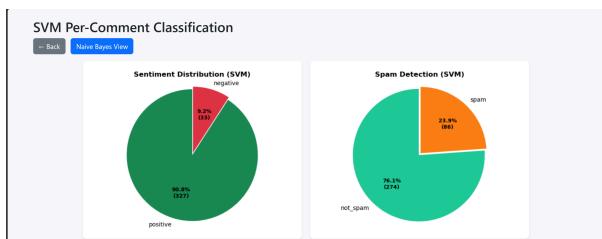
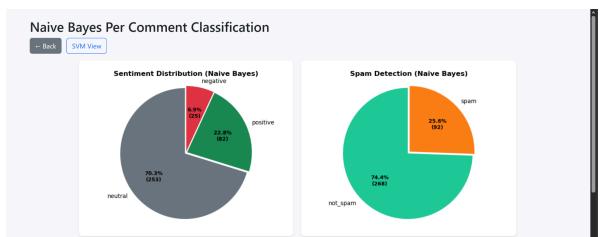
1. **Confusion Matrix:** Memetakan prediksi benar vs salah untuk melihat di mana model sering melakukan kesalahan (misal: sering salah membedakan 'netral' dan 'positif').



2. **Perbandingan Model:** Grafik batang yang menunjukkan bahwa NaiveBayes cenderung memiliki akurasi tertinggi dibandingkan SVM pada dataset ini.



3. **Distribusi Data (Pie Chart):** Menunjukkan proporsi komentar positif, negatif, dan netral dari dataNew.csv.



## BAB 6: KESIMPULAN

### 6.1 Analisis Hasil

Dari hasil eksekusi program terhadap data dataNew.csv, dapat disimpulkan bahwa:

1. Sistem berhasil melakukan *parsing* kolom komentar yang tergabung dalam satu sel CSV menjadi baris-baris terpisah untuk dianalisis.
2. Preprocessing sangat berpengaruh dalam membersihkan *noise* seperti link promosi atau *mention* yang tidak relevan dengan sentimen.
3. Model SVM memberikan hasil klasifikasi yang cukup akurat dalam memisahkan komentar spam (promosi jualan, link afiliasi) dari komentar organik (pertanyaan seputar kampus).

### 6.2 Kesimpulan Akhir

Aplikasi ini memenuhi persyaratan tugas untuk melakukan pipeline NLP lengkap. Penggunaan antarmuka web mempermudah pengguna awam dalam membaca hasil analisis tanpa perlu melihat kode. Untuk pengembangan selanjutnya, dataset latih (training data) perlu diperbanyak agar model lebih mengenali bahasa gaul (*slang*) yang sering muncul di TikTok.