



Hans Santoso - 535220129

MANUAL BOOK

Dashboard Klasifikasi Topik dan
Intensi pada Twitter Gojek dan Grab

JUDUL SKRIPSI

Klasifikasi Topik dan Intensi Ulasan Konsumen Gojek dan Grab pada Twitter
Menggunakan Metode Stacking Ensemble Xgboost dan Lightgbm



FAKULTAS TEKNOLOGI INFORMASI
PROGRAM STUDI TEKNIK INFORMATIKA

Dosen Pembimbing Tri Sutrisno, S.Si., M.Sc.

Versi 2025 Oktober

Daftar Isi

Daftar Isi

i

01

Pendahuluan - Mengenai Aplikasi Dashboard

01

| | |
|---------------------------|----|
| 1.1 Tujuan Aplikasi | 02 |
| 1.2 Prasyarat Sistem | 02 |
| 1.3 Tata Cara Menjalankan | 02 |

02

| | |
|--------------------------------|----|
| Menjalankan Dashboard Local | 03 |
| 2.1 Persiapan File | 04 |
| 2.2 Jalankan Script | 04 |
| 2.3 Buka Aplikasi pada Browser | 04 |

03

| | |
|---------------------------|----|
| Interface User | 05 |
| 3.1 Tampilan Utama | 06 |
| 3.2 Sidebar Navigasi | 07 |
| 3.2.1 Judul Aplikasi | 07 |
| 3.2.2 Unggah Dataset | 07 |
| 3.2.3 Navigasi Menu Utama | 07 |
| 3.2.4 Status Model | 07 |

04

| | |
|-----------------------------------|----|
| Menu Dataset | 08 |
| 4.1 Tampilan Utama | 09 |
| 4.2 Menampilkan Data | 09 |
| 4.3 Filter Merek, Topic Intention | 10 |
| 4.4 Statistik Ringkas | 10 |
| 4.5 Menampilkan Seluruh Data | 10 |
| 4.6 Informasi Dataset Asli | 11 |

05

| | |
|-----------------------------|----|
| Menu Dashboard | 12 |
| 5.1 Tampilan Awal | 13 |
| 5.2 Filter Merk dan Dimensi | 13 |
| 5.3 Membaca Visualisasi | 14 |

06

| | |
|--------------------------------|----|
| Menu Prediction | 15 |
| 6.1 Tampilan Awal | 16 |
| 6.2 Prediksi Tunggal | 16 |
| 6.3 Prediksi Batch | 17 |
| 6.3.1 Unggah Dataset | 17 |
| 6.3.2 Prediksi Batch dan Hasil | 17 |

07

| | |
|--|----|
| Menu Evaluation | 18 |
| 7.1 Tampilan Awal | 19 |
| 7.2 Memilih Skenario Evaluasi (Filter) | 19 |
| 7.3 Hasil Evaluasi | 19 |
| 7.4 Mengunduh Classification Report | 20 |
| 7.5 Confusion Matrix Ternormalisasi | 20 |
| 7.6 Hasil Evaluasi | 20 |

08

| | |
|-------------------|----|
| Menu About | 21 |
| Tampilan About | 22 |



01

Pendahuluan

Mengenai Aplikasi Dashboard



Tujuan

Aplikasi ini adalah prototipe dashboard interaktif yang berfungsi untuk menganalisis dan mengklasifikasikan ulasan pengguna Gojek dan Grab dari Twitter berdasarkan Topik layanan dan Intensi pengguna. Aplikasi ini memfasilitasi eksplorasi data melalui visualisasi dan memungkinkan prediksi otomatis pada data baru menggunakan model machine learning.

Prasyarat Sistem

- ❖ **Python** Versi 3.12.xx direkomendasikan
- ❖ **Library Python** Streamlit, Pandas, Plotly, Matplotlib, WordCloud, Scikit-learn, XGBoost, LightGBM, Streamlit-Antd-Components, Pickle, Numpy.
- ❖ **File model (.pkl)** tersimpan dalam struktur folder model/topic/ dan model/intention/

Tata Cara Menjalankan

Online :

- ❖ Buka <https://dashboard-klasifikasi-tweet-gojek-grab.streamlit.app/> pada browser web, aplikasi akan terbuka

localhost :

- ❖ Buka terminal atau command prompt.
- ❖ Navigasi ke direktori tempat file app.py disimpan.
- ❖ Jalankan perintah: streamlit run app.py
- ❖ Aplikasi akan terbuka di browser web

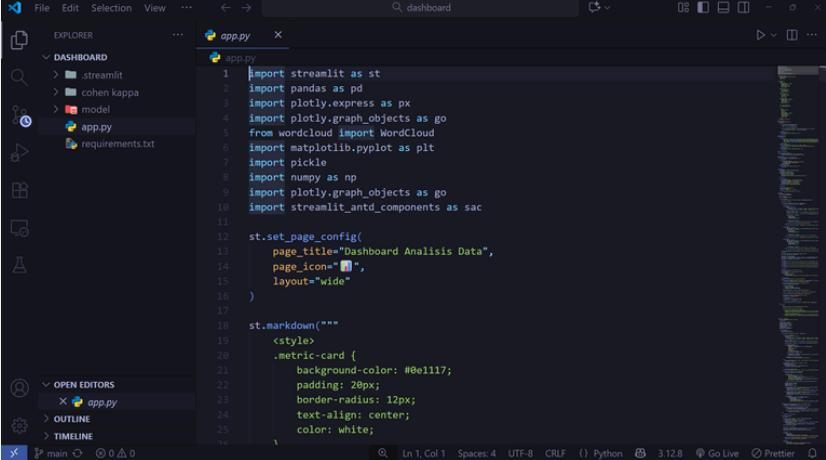
02

Menjalankan Dashboard Local



2.1 Persiapan File

Keseluruhan File Harus sudah disediakan dan didownload (**app.py** beserta dengan model **.pkl**) pada IDE yang dipilih.



The screenshot shows a code editor window with the file 'app.py' open. The code imports various libraries including streamlit, pandas, plotly, numpy, and wordcloud. It sets up a Streamlit configuration with a title, icon, and layout. The code then defines a markdown block containing CSS for a metric card, which has a background color of #0e1117, padding of 28px, a border radius of 12px, centered text, and white text color. The code editor interface includes a sidebar with 'EXPLORER', 'OPEN EDITORS', 'OUTLINE', and 'TIMELINE' sections, and a bottom status bar showing file details like 'Ln 1, Col 1' and 'Python 3.12.8'.

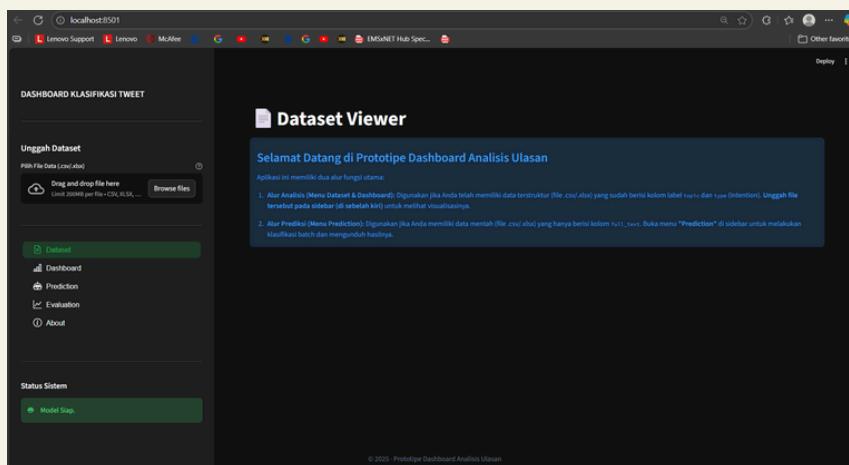
2.2 Jalankan Script

Kemudian buka Terminal atau Command Prompt. Pastikan direktori ada di lokasi dimana app.py streamlit berada. Run streamlit **run app.py / py -3.12 -m streamlit runn app.py**

```
PS C:\Data Science\SKRIPSI FINAL> py -3.12 -m streamlit run app.py
```

2.3 Buka Aplikasi pada Browser

Setelah Aplikasi dijalankan melalui termina, akan muncul pada browser default



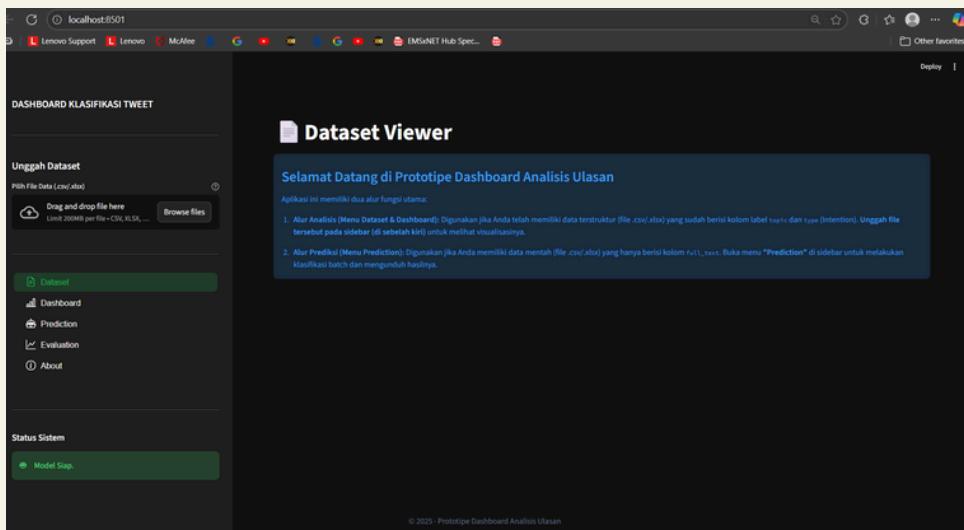
03

Interface User



3.1 Tampilan Utama

Ketika aplikasi terbuka. Interface aplikasi terdiri dari dua area utama yaitu Sidebar Navigasi di sebelah kiri untuk kontrol dan pemilihan menu, serta Area Konten Utama di sebelah kanan yang menampilkan informasi sesuai menu terpilih.

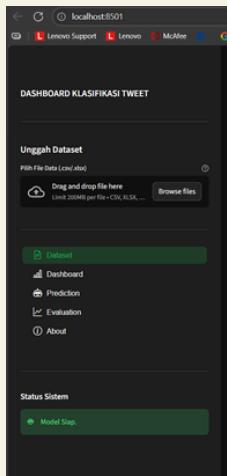


Sidebar

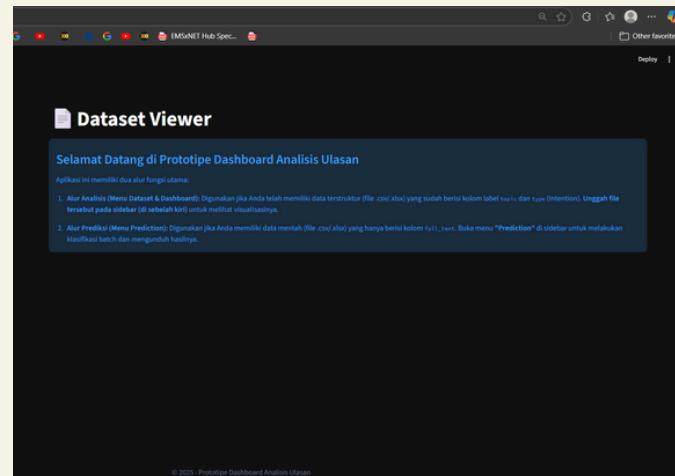
Terdapat Judul, Area unggah file dan navigasi menu ,dan status model

Konten Utama

Konten utama tampilan dari navigasi sidebar, ada 5 menu yang akan muncul yaitu Dataset, Dashboard, Prediction, Evaluation, dan About



Side Bar



Konten Utama

3.2 Sidebar Navigasi

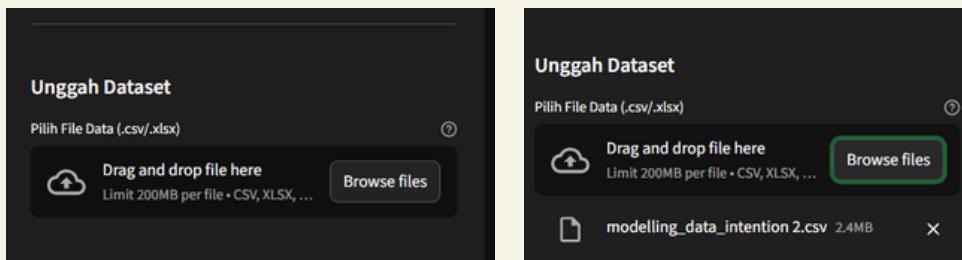
3.2.1 Judul Aplikasi

Menampilkan "DASHBOARD KLASIFIKASI TWEET".



3.2.2 Unggah Dataset

Area untuk mengunggah file data. Klik tombol "Pilih File Data (.csv/.xlsx)" untuk memilih file dari komputer Anda. Data setelah terpilih akan dibaca aplikasi dan ditampilkan untuk menu dataset dan dashboard. Format yang didukung adalah CSV dan Excel (XLSX, XLS).

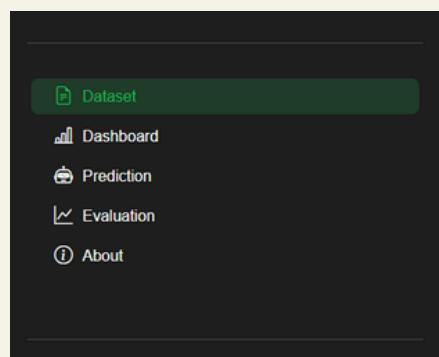


3.2.3 Navigasi Menu Utama

Daftar menu interaktif untuk navigasi antar fitur:

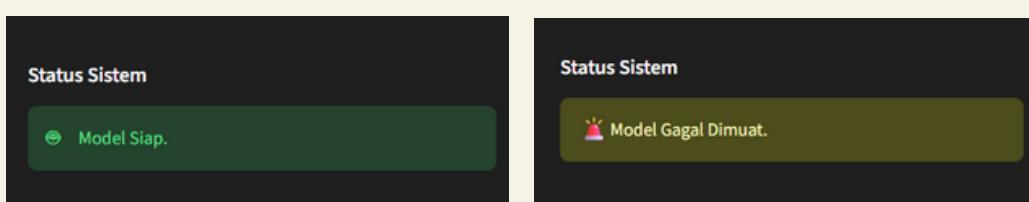
- Dataset: Melihat data mentah/terfilter.
- Dashboard: Visualisasi data interaktif.
- Prediction: Prediksi topik/intensitas.
- Evaluation: Hasil evaluasi performa model.
- About: Informasi aplikasi.

User dapat klik menu tersebut untuk pindah pindah tampilan



3.2.4 Status Model

Menunjukkan apakah model machine learning berhasil dimuat dan siap digunakan. Status bisa berupa "Model Siap." (hijau) atau "Model Gagal Dimuat." (kuning/merah).



04

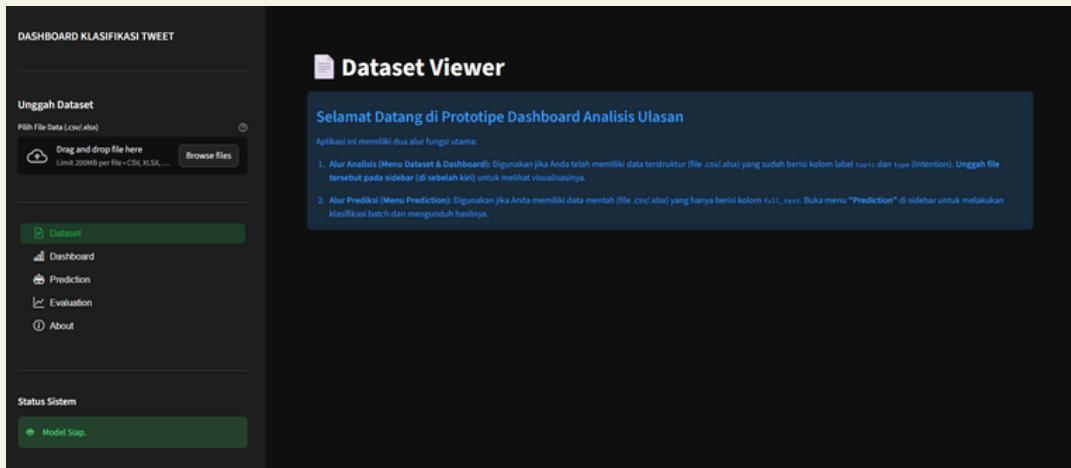
MENU

DATASET



4.1 Tampilan Menu Awal

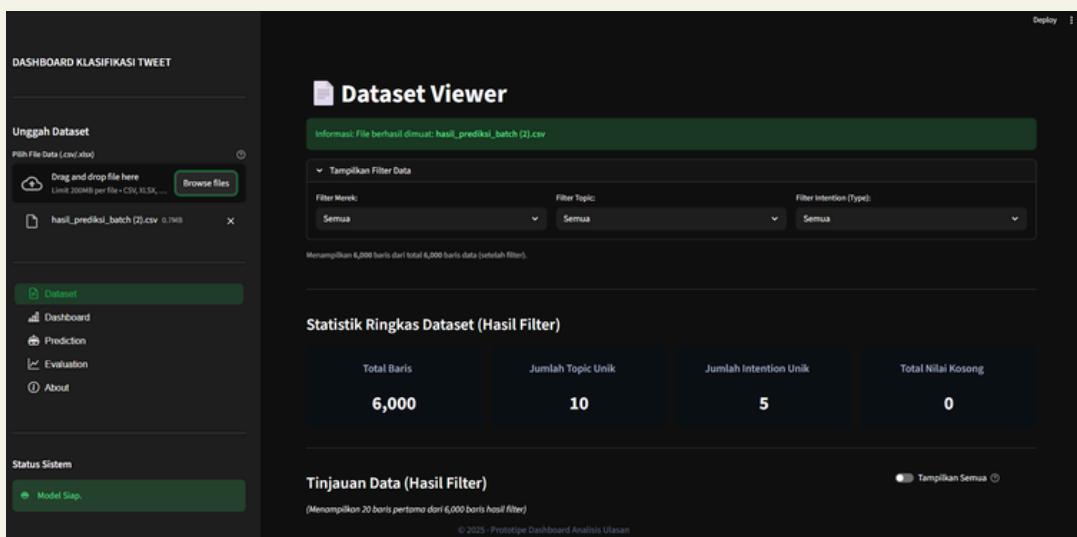
Halaman ini digunakan untuk melihat data tabular dan menerapkan filter. Berikut adalah tampilan sebelum dataset diupload dari sidebar, terdapat penjelasan dan langkah untuk alur analisis dan prediksi.



4.2 Menampilkan Data

Setelah data diunggah pada sidebar bagian unggah dataset, halaman ini akan menampilkan:

- Pesan sukses.
- Panel filter (dalam expander).
- Statistik ringkas.
- Tabel pratinjau 20 baris pertama data hasil filter.



4.3 Filter Merek, Topic Intention

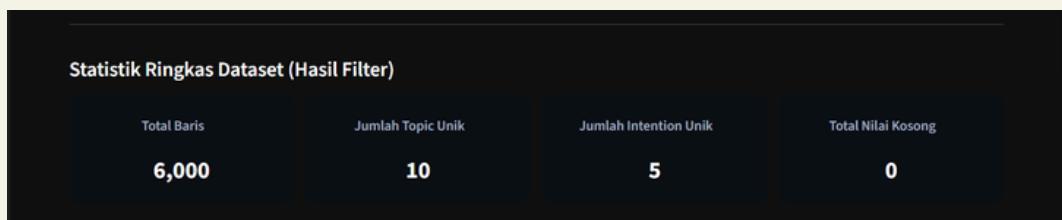
Buka expander "Tampilkan Filter Data". Gunakan dropdown untuk memilih:

- Filter Merek: Pilih "Semua", "Gojek", atau "Grab".
- Filter Topic: Pilih "Semua" atau topik spesifik.
- Filter Intention (Type): Pilih "Semua" atau intensi spesifik.

Data pada tabel dan statistik akan otomatis terfilter dan hanya menampilkan data sesuai pilihan filter tersebut.

4.4 Statistik Ringkas

Empat kartu metrik menampilkan jumlah baris data saat ini (setelah filter), jumlah topik unik, jumlah intensi unik, dan total nilai kosong (null).



4.5 Menampilkan Seluruh Data

Aktifkan toggle "Tampilkan Semua" di kanan atas tabel pratinjau untuk menampilkan seluruh baris data hasil filter dalam tabel.

| full_text | topic | type |
|---|----------|--------------|
| 0 @Jhooyzx gojek only (ga pasang pin) 3k bisa qriss | gojek | pernyataan |
| 1 sararin makanan enak aku mau gofood | gofood | pernyataan |
| 2 pls bapak gojek saya se oon itu buat dimainin cowo | gojek | pernyataan |
| 3 @neverccryyy @FOOD_FESS Bisaa dong kakkk Mau? Pake gosend sameday sekitar 15rb | gosend | pertanyaan |
| 4 McDONALDS LAGI LAPER BARENGAN? Ini ada Paket 3 ayam + 3 nasi cuma 53rb aja. GILAAA cuma 17 ribuan Cocok banget | grabfood | pujian |
| 5 @tehnit4 @Bluebirdgroup Udah gak pernah pake taksi online Grab/Gocar 100% udah pake Blue Bird sekarang lebih murah | grabcar | komplain |
| 6 @yumilicells Abis berapa kak? Terakhir di UB 30 menit habis 15an pls mahalan beam drg gojek | gojek | pertanyaan |
| 7 app gojek ini sensitif banget g sih kayak kok bisa lagi ngecheck harga rute tapi malah ke booking? kek otomatis?? ah f | gojek | komplain |
| 8 Pilihan PC Set: PC A Set dari Web Plus Chat PC B Set dari Mnet Plus Merch bisa pilih antara dua set eksklusif ini! Suda | gopay | saran-kritik |
| 9 Gasuka bgt klo dpt gojek motornya vespa takut ngejengkang bgt ya allah | gojek | pernyataan |

4.6 Informasi Dataset Asli

Buka expander "Tampilkan Info Dataset Asli (Sebelum Filter)" untuk melihat detail kolom (nama, tipe, jumlah null) dan penggunaan memori dari data awal sebelum filter diterapkan.

| ▼ Tampilkan Info Dataset Asli (Sebelum Filter) | | | | |
|--|-----------|-----------------|-------------|--|
| Informasi Kolom | | | | |
| Kolom | Tipe Data | Jumlah Non-Null | Jumlah Null | |
| full_text | object | 6000 | 0 | |
| topic | object | 6000 | 0 | |
| type | object | 6000 | 0 | |

Total Baris Asli: 6,000 | Penggunaan Memori: 1.62 MB

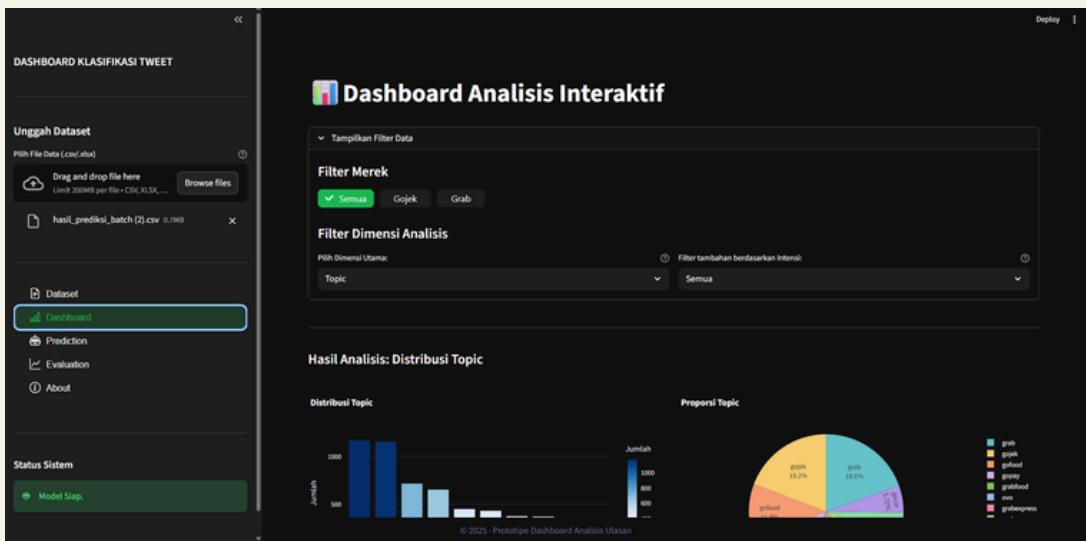
05

Menu Dashboard



5.1 Tampilan Awal

Halaman ini menyajikan visualisasi interaktif dari data yang diunggah.

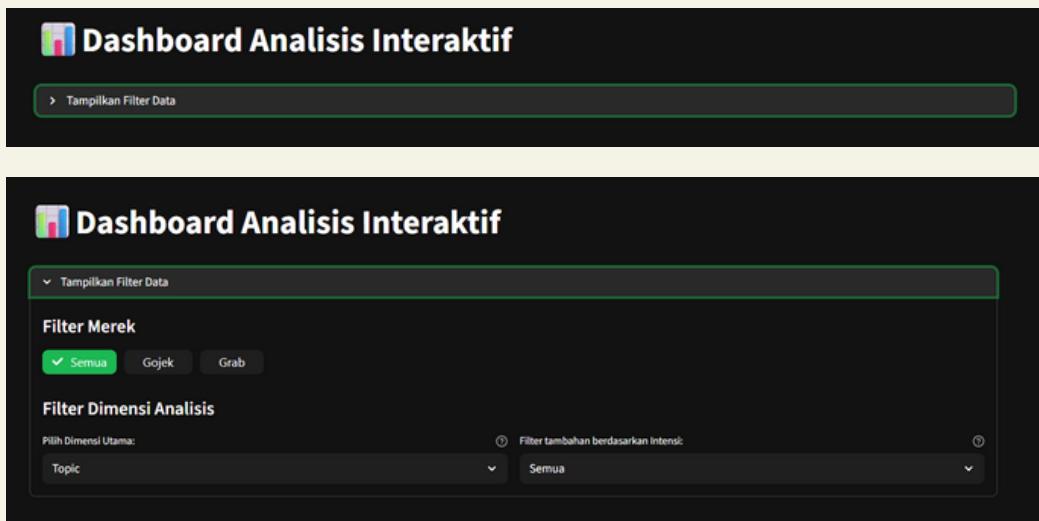


5.2 Filter Merk dan Dimensi

Gunakan tombol chip di bagian atas (di dalam expander filter) untuk memilih antara "Semua", "Gojek", atau "Grab". Semua visualisasi di bawahnya akan otomatis terfilter berdasarkan merek yang dipilih.

Di bawah filter merek, terdapat filter dimensi:

1. Pilih Dimensi Utama: Pilih apakah visualisasi utama (grafik batang dan pie) akan menampilkan distribusi berdasarkan "Topic" atau "Intention".
2. Filter Tambahan: Dropdown ini akan muncul otomatis. Jika Dimensi Utama adalah "Topic", filter tambahan adalah "Intention", begitu pula sebaliknya. Pilih nilai spesifik untuk memfilter visualisasi lebih lanjut.



5.3 Membaca Visualisasi

- Grafik Batang: Menunjukkan jumlah ulasan untuk setiap kategori pada Dimensi Utama yang dipilih, setelah semua filter (Merek, Dimensi Utama, Tambahan) diterapkan.
- Grafik Pie: Menunjukkan proporsi persentase ulasan untuk setiap kategori pada Dimensi Utama, setelah semua filter diterapkan.
- Word Cloud: Menampilkan kata-kata yang paling sering muncul dalam kolom full_text dari data yang sudah terfilter.

Semua visualisasi ini interaktif dan akan diperbarui secara otomatis saat filter diubah.



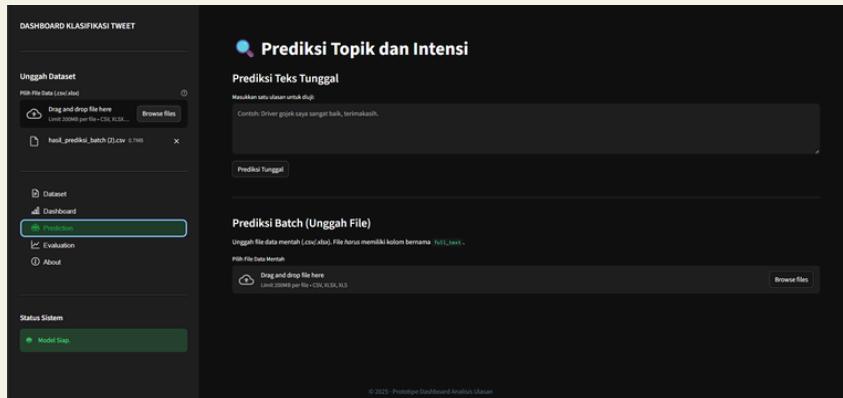
06

Menu Prediction



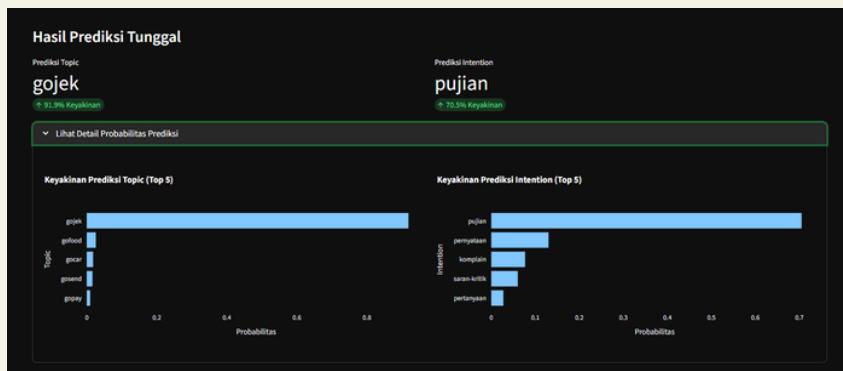
6.1 Tampilan Awal

Halaman ini digunakan untuk mengklasifikasikan teks ulasan baru menggunakan model yang telah dimuat.



6.2 Prediksi Tunggal

- Masukkan satu teks ulasan ke dalam kotak "Masukkan satu ulasan untuk diuji:".
- Klik tombol "Prediksi Tunggal".
- Hasil prediksi untuk Topik dan Intensi akan muncul di bawahnya, lengkap dengan persentase keyakinan.
- Buka expander "Lihat Detail Probabilitas Prediksi" untuk melihat grafik 5 besar kemungkinan prediksi.



6.3 Prediksi Batch

6.3.1 Unggah Dataset

- Klik tombol "Pilih File Data Mentah".
- Pilih file CSV atau Excel yang harus memiliki kolom bernama full_text.
- Jika file valid, pratinjau 5 baris pertama akan ditampilkan, beserta pesan konfirmasi.

The screenshot shows a user interface for batch prediction. At the top, there's a header 'Prediksi Batch (Unggah File)'. Below it, a message says 'Unggah file data mentah (.csv/.xlsx). File harus memiliki kolom bernama full_text.' A 'Pilih File Data Mentah' button is present. A file named 'hasil_prediksi_batch [2].csv' (0.71MB) is shown being uploaded. A 'Browse files' button is also visible. Below this is a preview table titled 'Tinjauan Data Mentah' with columns 'full_text', 'topic', and 'type'. The table contains five rows of sample data. At the bottom, a green bar displays the message 'Informasi: Kolom full_text terdeteksi. (6,000 baris siap diproses)'.

6.3.2 Prediksi Batch dan Hasil

Setelah file valid terunggah, klik tombol "Mulai Prediksi Batch". Tunggu proses prediksi selesai (akan muncul spinner). Waktu proses tergantung jumlah data. Setelah prediksi selesai, akan ditampilkan:

- Tinjauan Hasil: Tabel pratinjau data asli ditambah kolom topic, topic_probability, type, dan type_probability.
- Ringkasan Hasil: Grafik pie yang menunjukkan distribusi Topik dan Intensi hasil prediksi.
- Tombol Unduh: Klik "Unduh Hasil Prediksi Lengkap (.csv)" untuk menyimpan file CSV berisi semua data input beserta hasil prediksinya.

The screenshot shows the results of the prediction batch. It includes a preview table 'Tinjauan Hasil Prediksi Batch' with columns 'full_text', 'topic', 'type', 'topic_probability', and 'type_probability'. The table has the same five rows as the preview in the upload interface. Below this are two pie charts: 'Ringkasan Hasil Prediksi' (Topic Distribution) and 'Distribusi Intention yang Diprediksi' (Intention Distribution). The Topic Distribution chart shows proportions for grab, gojek, gofood, gojipay, grabfood, saranan, grabexpress, grabcar, gosend, and pujuan. The Intention Distribution chart shows proportions for permintaan, kumpulan, saran_kritik, pujuan, and pertanyaan.

07

Menu Evaluation



7.1 Tampilan Awal

Halaman ini menampilkan metrik-metrik evaluasi performa model yang telah dihitung sebelumnya (offline).

| Label | Precision | Recall | F1-Score | Support |
|-----------|-----------|--------|----------|---------|
| gocar | 0.91 | 0.89 | 0.9 | 280 |
| gefod | 0.92 | 0.82 | 0.86 | 280 |
| grabtrans | 0.79 | 0.87 | 0.83 | 280 |
| grabcar | 0.98 | 0.92 | 0.95 | 280 |
| groovy | | | | |

7.2 Memilih Skenario Evaluasi (Filter)

Gunakan tiga dropdown di bagian atas halaman untuk memilih kombinasi skenario yang ingin dilihat hasilnya:

- Pilih Model: Topic atau Intention.
- Pilih Tipe Eksperimen: Parameter Default atau Hyperparameter Tuned.
- Pilih Split Data: 70/30 atau 80/20.

7.3 Hasil Evaluasi

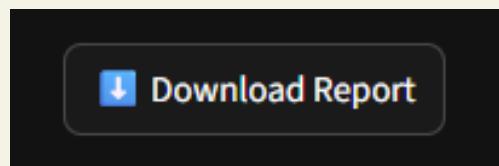
Setelah filter dipilih, halaman akan menampilkan:

- Akurasi: Nilai Train Accuracy dan Test Accuracy.
- Classification Report: Tabel berisi Precision, Recall, F1-Score, Support per kelas, beserta rata-rata Macro dan Weighted.
- Confusion Matrix: Heatmap visual yang menunjukkan jumlah prediksi benar (diagonal) dan salah (non-diagonal) antar kelas.

| Label | Precision | Recall | F1-Score | Support |
|-----------|-----------|--------|----------|---------|
| groovy | 0.98 | 0.93 | 0.95 | 280 |
| gefod | 0.97 | 0.89 | 0.93 | 280 |
| grabtrans | 0.96 | 0.92 | 0.94 | 280 |
| grabcar | 0.92 | 0.81 | 0.86 | 280 |
| gocar | | | | |

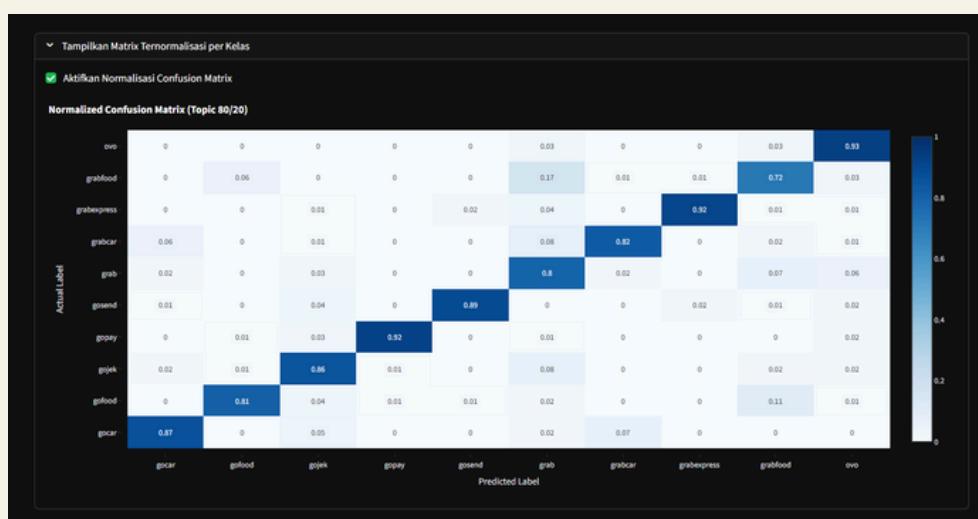
7.4 Mengunduh Classification Report

Klik tombol "Download Report" di bawah tabel Classification Report untuk mengunduh data tabel tersebut dalam format CSV.



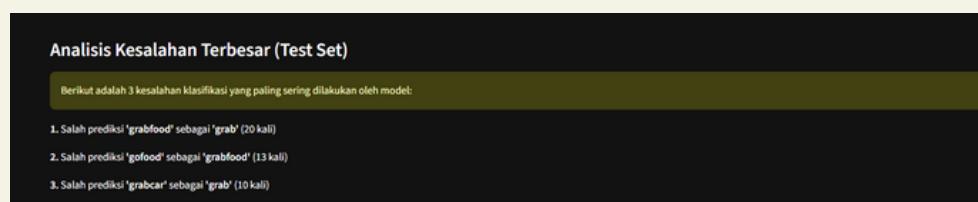
7.5 Confusion Matrix Ternormalisasi

Buka expander "Tampilkan Matrix Ternormalisasi per Kelas". Centang checkbox di dalamnya untuk melihat Confusion Matrix yang nilainya dinormalisasi menjadi proporsi (0-1) berdasarkan baris (kelas aktual).



7.6 Hasil Evaluasi

Bagian ini secara otomatis menampilkan 3 pasangan kelas (aktual vs prediksi) yang paling sering salah diklasifikasikan oleh model pada skenario yang dipilih, berdasarkan data Confusion Matrix.





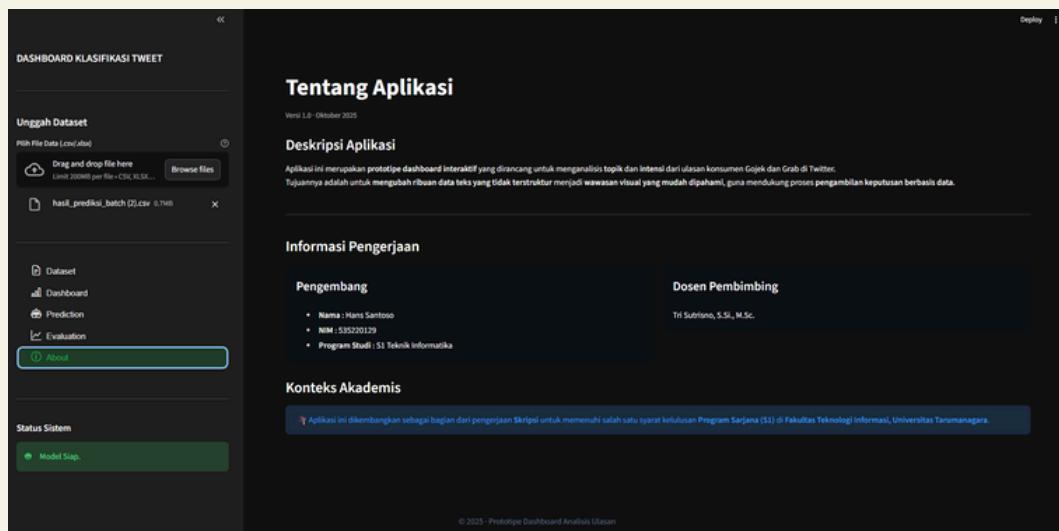
08

Menu

About

8.1 Tampilan About

Halaman ini berisi informasi statis tentang aplikasi. Cukup pilih menu "About" di sidebar untuk melihat deskripsi aplikasi, detail pengembang dan pembimbing, serta konteks penggerjaan skripsi.





Thank You

Terima kasih telah meluangkan waktu untuk membaca dan menggunakan buku panduan ini. Semoga panduan ini dapat membantu Anda dalam memahami dan mengoperasikan Aplikasi Dashboard Klasifikasi Tweet Gojek & Grab dengan efektif.

Panduan ini disusun untuk memberikan gambaran yang jelas mengenai fitur dan fungsionalitas aplikasi. Jika terdapat pertanyaan lebih lanjut atau masukan, jangan ragu untuk menyampaikannya.

Selamat menggunakan aplikasi!

Contact Information

Linkedin <https://www.linkedin.com/in/hans-santoso/>

Instagram hanssantoso

Email hans.5353220129@stu.untar.ac.id
