

Nama: M. Rajikh Nashirul Khaqqi

NIM: 231240001356

Panduan Teknis & Penggunaan Cepat: AI Toxic Comment Detector

1. Ringkasan Sistem

Aplikasi web berbasis **client-side** (berjalan di browser) untuk mendeteksi tingkat toksitas teks menggunakan metode **Rule-Based**.

- **Bahasa:** Indonesia & Inggris.
- **Privasi:** 100% data diproses lokal, tidak ada yang dikirim ke server.
- **Dependensi:** Hanya butuh internet saat load pertama (Google Fonts), selebihnya offline.

2. Logika Deteksi (Detail Teknis)

Sistem bekerja dengan alur pipa (pipeline) berikut:

1. Input & Preprocessing:

- **Lowercase:** Semua huruf dikecilkan.
- **Tokenization:** Memecah kalimat menjadi per kata.
- **Cleaning:** Hapus tanda baca.
- **Stemming (Penyederhanaan Kata):**
 - *Indo:* Hapus imbuhan akhir (nya, lah, kah, pun).
 - *Inggris:* Hapus suffix (ing, ed, es, s).

2. Matching & Scoring:

- Mencocokkan token dengan database toxic_words.json.

- **Whitelist Check:** Mencegah *false positive* (contoh: kata "ass" di dalam "class" tidak dihitung).
- **Severity Weight:**
 - Light = 1 poin
 - Medium = 2 poin
 - Heavy = 3 poin

3. Klasifikasi Skor:

-  Aman (< 20%)
-  Cukup Toxic (20% - 50%)
-  Berbahaya (> 50%)

3. Struktur Data

Dictionary (data/toxic_words.json)

Jika ingin menambah kata baru, ikuti format ini:

```
{  
  "word": "kata_toxic",  
  "severity": "light|medium|heavy",  
  "lang": "id|en"  
}
```

Penyimpanan Lokal (LocalStorage)

Riwayat tersimpan di browser user dengan key toxic_history_v1.

- **Kapasitas:** Max 10 entry terakhir.
- **Format:** JSON (bisa di-export).

4. Cara Penggunaan

1. Instalasi/Akses:

- Buka index.html langsung di browser atau gunakan local server (python -m http.server 8000) untuk performa terbaik.

2. Analisis:

- Ketik teks -> Klik "**Analyze Sentiment**".
- Lihat hasil: Badge warna (Hijau/Kuning/Merah) dan kata yang di-highlight.

3. Manajemen Data:

- **Export:** Klik "Download JSON" untuk backup riwayat.
- **Hapus:** Klik "Hapus Riwayat" untuk membersihkan cache browser.

5. Troubleshooting Cepat

Masalah	Penyebab & Solusi
Gagal Fetch JSON	File toxic_words.json hilang atau salah path. Sistem akan otomatis pakai <i>Fallback Dictionary</i> .
Highlight Gagal	Refresh halaman. Pastikan kata tersebut ada di file JSON.
Riwayat Hilang	Cek apakah browser dalam mode <i>Incognito</i> (data tidak tersimpan di mode ini).