Nama : Rendi Hendra Syahputra  
Prodi : S1 Rekayasa Perangkat Lunak  
NIM : 23020800005

Laporan Singkat: Analisis Crawling Komentar YouTube

Penelitian ini bertujuan untuk mengumpulkan data komentar dari video YouTube menggunakan YouTube Data API v3. Data komentar diambil dari video dengan ID "0jFfhru66zA" untuk menganalisis interaksi pengguna, seperti tanggal publikasi, nama pengguna, isi komentar, dan jumlah suka. Laporan ini menjelaskan dataset yang dihasilkan, metode pengumpulan data, hasil analisis awal, dan kesimpulan.

1. Dataset

Dataset berupa daftar komentar dari video YouTube dengan ID "0jFfhru66zA ". Setiap entri mencakup:

- Tanggal Publikasi: Waktu komentar dipublikasikan (format ISO 8601).

- Nama Pengguna: Nama akun pengguna yang berkomentar.s

- Isi Komentar: Teks komentar, termasuk emoji dan format HTML.

- Jumlah Suka: Jumlah suka yang diterima komentar.

Data mencakup komentar utama dan balasan, disimpan dalam format list Python, yang dapat dikonversi ke CSV atau DataFrame untuk analisis lebih lanjut.

Dataset berhasil dikumpulkan dengan temuan:

Total number of comments: 2339

Number of positive comments: 143

Number of negative comments: 163

Number of neutral comments: 2033

Positive Percentage: 6.11%

Negative Percentage: 6.97%

Neutral Percentage: 86.92%

1. Metode

Pengumpulan data dilakukan dengan:

* Pengaturan API: Menggunakan YouTube Data API v3 dengan kunci API dan library googleapiclient.discovery.
* Fungsi Crawling: Fungsi video\_comments(video\_id) mengambil komentar utama dan balasan, mengekstrak informasi seperti tanggal, pengguna, komentar, dan suka, serta menangani paginasi dengan nextPageToken.
* Eksekusi: Kode dijalankan untuk video ID "0jFfhru66zA ", menghasilkan daftar komentar.

1. Hasil

Hasil klasifikasi sentimen positif dan negatif menggunakan scikit-learn, TensorFlow, atau Pytorch:

Accuracy: 0.8931623931623932

precision recall f1-score support

Negative 1.00 0.43 0.60 35

Neutral 0.89 1.00 0.94 401

Positive 1.00 0.06 0.12 32

accuracy 0.89 468

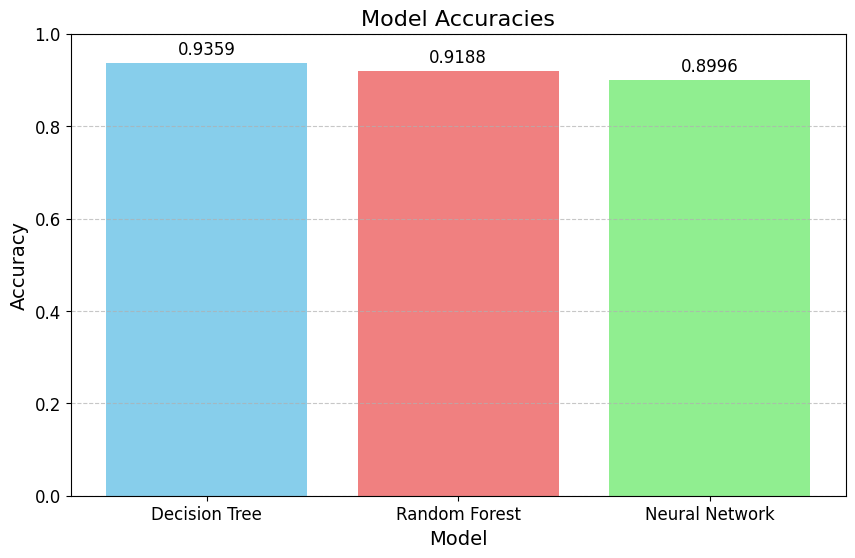
macro avg 0.96 0.50 0.55 468

weighted avg 0.91 0.89 0.86 468

5. Kesimpulan

Pengumpulan data komentar YouTube berhasil, menghasilkan dataset yang kaya informasi. Metode crawling efektif. Hasil menunjukkan interaksi yang aktif dengan tema yang di usung dalam konten.

Grafik hasil accuracy Decision Tree, Random Forest, Neural Network:

:

Model yang paling akurat adalah: Decision Tree dengan akurasi sebesar 0.9359.