BENONY GABRIEL

[Email address]

Abstract

[Draw your reader in with an engaging abstract. It is typically a short summary of the document.   
When you’re ready to add your content, just click here and start typing.]

Manual book

Sentiment Analysis on Universities under APERTI BUMN via Google Review

1. **PENDAHULUAN**
2. **SPESIFIKASI PROYEK**
3. **PERSIAPAN LINGKUNGAN**
4. **STRUKTUR DATASET**
5. **LABELLING DATASET**
6. **Tujuan**

Tujuan dari proses pelabelan dataset adalah untuk mengkategorikan ulasan (review) yang diperoleh dari Google Maps ke dalam tiga jenis sentimen utama, yaitu:

* Positif
* Negatif
* Netral

Label sentimen ini diperlukan sebagai target dalam pemodelan machine learning yang bertujuan untuk memahami persepsi pengguna terhadap institusi pendidikan (dalam hal ini, kampus di bawah naungan APERTI BUMN).

1. **Alat Labelling**

Proses labelling dilakukan menggunakan model GPT-3.5 Turbo dari OpenAI melalui API sebagai mesin klasifikasi sentimen karena beberapa alasan utama berikut:

* **Kemampuan Bahasa Alami**: GPT-3.5 Turbo sangat andal dalam memahami konteks bahasa Indonesia, termasuk ekspresi-emosi dalam teks informal seperti review pengguna.
* **Tidak Memerlukan Data Latih**: Berbeda dengan metode supervised learning konvensional yang memerlukan data berlabel terlebih dahulu, GPT-3.5 dapat digunakan langsung (*zero-shot classification*) untuk memberi label pada teks.
* **Konsistensi & Efisiensi Waktu**: Pelabelan manual oleh manusia rawan subjektivitas dan sangat memakan waktu. Dengan API GPT, proses dapat diotomatisasi secara akurat dan konsisten.
* **Fleksibel**: Bisa digunakan untuk berbagai domain dan jenis teks dengan sedikit atau tanpa penyesuaian.

1. **Cara Kerja Proses Labeling**

Proses labeling dilakukan secara otomatis dan terprogram dengan alur sebagai berikut:

1. Membaca Dataset

Dataset review hasil scraping Google Maps dibaca menggunakan library pandas. Fokus utama adalah kolom snippet yang berisi isi ulasan pengguna.

A screenshot of a computer code

AI-generated content may be incorrect.

1. Membangun fungsi Klasifikasi Sentimen

Untuk setiap review, dibuat sebuah prompt atau instruksi dalam bahasa Indonesia yang dikirimkan ke API GPT-3.5 Turbo.

A screenshot of a computer program

AI-generated content may be incorrect.

Prompt tersebut dikirim ke API menggunakan endpoint chat.completions.create() dari library openai versi terbaru (>= 1.0). Response dari model akan berupa satu kata hasil klasifikasi.

1. Memvalidasi dan Menyimpan Label

Hasil respon dari API divalidasi agar sesuai dengan tiga label yang ditentukan. Setelah itu, label disimpan kembali ke dataset pada kolom baru label\_sentimen.

A screen shot of a computer code

AI-generated content may be incorrect.

1. **Hasil**

Setelah semua snippet diberi label, dataset akan memiliki kolom baru **label\_sentimen** yang siap digunakan untuk tahap selanjutnya:

A screenshot of a computer

AI-generated content may be incorrect.

1. **Ss**
2. **Ss**
3. **Ss**
4. **Ss**
5. **Ss**
6. **Ss**
7. **ETC**
8. **CC**
9. **CC**
10. **CC**
11. **CC**
12. **CC**
13. **CC**
14. **CC**