



Sentiment & Issue Classification of Airline Tweets Using IBM Granite

CAPSTONE PROJECT

Raw Dataset Link

- ▶ Dataset : Tweets.csv
- ▶ Sumber :
<https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment>

Project Overview

- ▶ **Tujuan Proyek :**

Menganalisis tweet pelanggan maskapai penerbangan untuk mengidentifikasi sentimen (positif, negatif, campuran) serta kategori masalah layanan (keterlambatan, pelayanan kasar, atau bagasi hilang). Hal ini bertujuan untuk memberikan insight strategis terhadap peningkatan layanan pelanggan maskapai.

- ▶ **Latar Belakang :**

Industri penerbangan sangat bergantung pada reputasi dan kepuasan pelanggan. Media sosial seperti Twitter menjadi sarana utama bagi pelanggan untuk menyuarakan pengalaman mereka. Dengan memanfaatkan analisis berbasis AI terhadap data tweet, maskapai dapat lebih cepat memahami dan menanggapi masalah pelanggan.

- ▶ **Permasalahan :**

Tweet pelanggan memiliki volume tinggi, ragam bahasa informal, serta subjektivitas tinggi, sehingga sulit untuk dianalisis secara manual.

- ▶ **Pendekatan :**

Menggunakan model IBM Granite (granite-3.3-8b-instruct) untuk memproses dan mengklasifikasikan tweet berdasarkan sentimen dan isu utama. Model ini diakses melalui API Replicate dan diintegrasikan dalam pipeline Python.

Analysis Process

1.Data Loading : DatasetTweets.csv dimuat menggunakan pandas.

2.Preprocessing :

- Hanya kolom text yang digunakan.
- Data dibagi menjadi batch untuk mematuhi konteks maksimum model LLM.

3.Model Prompting :

- Setiap batch diproses menggunakan prompt yang meminta klasifikasi sentimen dan pengelompokan isu (Delay, Rude Service, Lost Baggage).

4.Refinement Prompts :

- Prompt tambahan digunakan untuk meningkatkan akurasi dan struktur hasil model, seperti :
 - Klasifikasi dua Langkah
 - Format output terstruktur

5.Output Parsing & Aggregation:

- Hasil disimpan dan dicetak untuk setiap batch, lalu dapat diekstrak untuk analisis lebih lanjut atau visualisasi.

Insight, Findings, and Visualization

Temuan Utama :

- Sentimen Negatif Mendominasi :
 - Mayoritas tweet berisi keluhan daripada pujian.
 - Sentimen negatif lebih tinggi pada maskapai tertentu (dapat diverifikasi dengan kolom airline).
- Masalah Utama Pelanggan :
 - Delay (Keterlambatan): Dominan dan paling sering dikeluhkan.
 - Rude Service (Pelayanan Kasar): Sering muncul di tweet bernada negatif.
 - Lost Baggage (Bagasi Hilang): Umumnya berasosiasi dengan sentimen negatif atau campuran.
- Sentimen Campuran :
 - Banyak pengguna menyisipkan keluhan dan pujian dalam satu tweet, menandakan kompleksitas dalam persepsi layanan pelanggan.

Visualisasi :

- Bar chart distribusi sentimen

Conclusion & Recommendations

► Kesimpulan :

- Sentimen pelanggan pada media sosial dapat menjadi indikator penting kualitas layanan maskapai.
- Isu utama seperti keterlambatan dan pelayanan kasar perlu menjadi perhatian utama.

► Rekomendasi :

- Tingkatkan komunikasi proaktif selama keterlambatan.
- Pelatihan ulang untuk frontliner maskapai.
- Integrasi sistem pelacakan bagasi real-time untuk mengurangi keluhan kehilangan barang.

AI Support Explanation

► Peran AI (IBM Granite LLM) :

- **Text Understanding** : Mengurai bahasa informal dan ekspresif dalam tweet.
- **Sentiment Classification** : Menggunakan pemahaman kontekstual untuk menilai sentimen pelanggan.
- **Issue Tagging** : Mengelompokkan keluhan dalam kategori Delay, Rude Service, dan Lost Baggage secara otomatis.

► Alasan Penggunaan :

- Manual tagging tidak efisien untuk ribuan tweet.
- Model IBM Granite terbukti unggul dalam instruksi NLP dan klasifikasi teks kompleks.
- Prompt engineering digunakan untuk mengarahkan model pada hasil yang diinginkan.