NAMA : MOHAMMAD REZA
NIM : A11.2022.14741
MATA KULIAH : DATA MINNING

DRAFT EKSPERIMEN TUGAS AKHIR

A. **TOPIK ATAU JUDUL:** Analisis Sentimen Komentar Youtube menggunakan Algoritma Klasifikasi Naïve Bayes

B. DESKRIPSI SINGKAT (BUSSINES UNDERSTANDING/DATA UNDERSTANDING):

Dalam eksperimen ini, saya akan menerapkan metode klasifikasi seperti naïve bayes untuk menganalisis sentimen dari komentar pengguna di platform YouTube. Tujuan dari eksperimen ini adalah untuk mengidentifikasi sentimen dalam komentar, apakah positif, negatif, atau netral.

C. TUJUAN:

- Masalah: Bagaimana cara mengklasifikasikan sentimen dalam komentar YouTube menjadi kategori positif, negatif, atau netral?
- Tujuan: Menggunakan algoritma klasifikasi naïve bayes untuk mengidentifikasi dan menganalisis sentimen pada komentar YouTube serta mengevaluasi kinerja model dalam mengklasifikasikan sentimen.

D. ALUR / TAHAPAN / KERANGKA EKSPERIMEN

- **BU (BUSSINES UNDERSTANDING):** Mengidentifikasi tujuan analisis sentimen untuk memahami pandangan penonton terhadap konten di YouTube.
- **DU (DATA UNDERSTANDING):** Mengumpulkan dataset komentar YouTube. Dataset ini bisa diperoleh melalui scraping komentar dari video tertentu atau menggunakan dataset komentar YouTube yang tersedia di platform seperti Kaggle.
- **Pre-processing Data:** Membersihkan teks komentar dari simbol, emoji, URL, dan karakter khusus, Memberikan StopWords, Stemming.
- **Modelling:** Menggunakan algoritma klasifikasi seperti **Naïve Bayes**, SVM, atau Logistic Regression untuk menentukan sentimen komentar (positif, negatif, atau netral).
- Evaluasi: Mengukur akurasi model dengan menggunakan metrik seperti confusion matrix.
- Implementasi: Menyajikan hasil dari analisis sentimen dan memberikan kesimpulan mengenai sentimen yang terdeteksi.

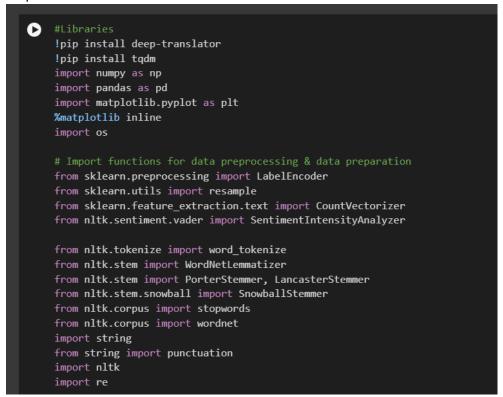
E. PENJELASAN DATASET

 DATASET: Dataset yang digunakan dalam eksperimen ini terdiri dari kumpulan komentar dari platform YouTube. Data ini dapat diperoleh melalui web scraping, API Youtube atau menggunakan dataset yang sudah tersedia di Kaggle.

• ATTRIBUT UTAMA:

- a. Comment: Beirisi komentar oengguna
- b. Sentiment: Label sentiment (jika tidak ada label pada dataset maka bisa dilakukan proses labeling terlebih dahulu).

Eksperimen:





```
from deep_translator import GoogleTranslator from topd import altk from nltk.sentiment.vader import SentimentintensityAnalyzer

# Pastikan untuk mendowiload lextcon VADER nltk.download('vader_lextcon VADER nltk.download('vader_lextcon')

# Terjembkian teks komentar ke bahasa Inggris dengan progress bar datal('coment_B') = [GoogleTranslator(cource'id', target='en').translate(coment) if comment else "" for comment in todm(datal['Comment_B') = datal('Comment_B').fillna("')

# Isi None atau NaN dalam Comment_B' dengan string kosong untuk menghindari error datal('comment_B') = datal('comment_B').fillna("')

# Analisis sentimen pada teks yang sudah diterjemakkan datal('comment_B') = datal('comment_B').datal('comment_B') = datal('comment_B').datal('megative') = [sentiments.polarity.scores(3)["post] if i else # for i in datal('comment_B')] datal('megative') = [sentiments.polarity.scores(3)["compound"] if i else # for i in datal('comment_B')] datal('megative') = [sentiments.polarity.scores(3)["compound"] if i else # for i in datal('comment_B')] datal('compound') = [sentiments.polarity.scores(3)["compound'] if i else # for i in datal('comment_B')] else # for i in datal('comment_B') | datal('comment_B') | datal('compound') = [sentiments.polarity.scores(3)["compound'] if i else # for i in datal('comment_B') | el
```

