Analyzing the Relationship Between YouTube Video Popularity and Comment Sentiment using Transformer-Based Models

Abhiraj Ezhil
Department of Statistics
University of Michigan
Ann Arbor, USA
email: abhiraj@umich.edu

***Abstract*—This paper explores the relationship between the popularity of YouTube videos and the sentiment of their most relevant comments using a dataset containing video statistics and top comments. We leverage traditional machine learning techniques as well as transformer-based models from Hugging Face to perform sentiment analysis and evaluate how comment sentiment correlates with engagement metrics such as likes, views, and comment counts. Our findings show that transformer models provide robust sentiment predictions, enabling deeper insights into video-audience interactions.**

***Index Terms*—Sentiment analysis, YouTube, NLP, Hugging Face, Transformer, Logistic Regression, SVM, Naive Bayes, Random Forest**

I. Introduction

Social media platforms such as YouTube play a significant role in shaping public opinion and entertainment consumption. Understanding how the sentiment of user comments relates to the popularity of videos provides key insights for content creators and platforms alike. With advancements in Natural Language Processing (NLP), especially transformer-based models, we can now perform sentiment analysis with much higher accuracy and interpretability.

Previous research has focused on predicting sentiment in comments using traditional approaches like Naive Bayes, SVM, or LSTMs. However, recent models such as BERT, RoBERTa, and DistilBERT have achieved state-of-the-art performance across many NLP tasks. This project uses both traditional TF-IDF based classifiers (Logistic Regression, SVM, Naive Bayes, and Random Forest) and Hugging Face's transformer models to classify comment sentiment and analyze how it correlates with video popularity.

II. Method

A. Problem Formulation

The objective is to predict the sentiment of comments associated with trending YouTube videos and analyze their correlation with popularity metrics such as views, likes and comment count.

Input: Text of a YouTube comment.
Output: Sentiment label – negative (0), neutral (1), or positive (2).

B. Dataset Description and Preprocessing

The dataset consists of two files: `videos-stats.csv` and `comments.csv`. The former contains metadata for each video such as views, likes, and comment count, while the latter contains the top 10 most relevant comments per video along with pre-assigned sentiment scores.

These two files were merged using the `Video ID` column as the key. The final data set included columns such as comment text, comment likes, video views, video likes, and sentiment. For traditional machine learning models, comments were cleaned using techniques such as lower case insertion, punctuation, and URL removal, stop-word filtering, and lemmatization.

For transformer-based predictions, I used a RoBERTa Hugging Face model fine-tuned with Twitter sentiment data. A new column, `Transformer Prediction`, was created by applying the sentiment prediction function on the `Comment` column.

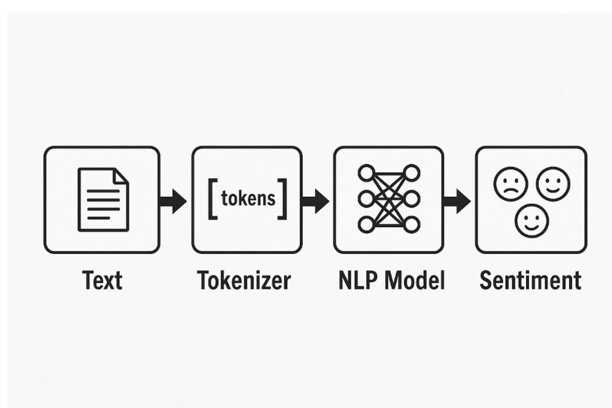

Fig. 1. Sentiment Analysis Pipeline Overview.

C. Model Formulation and Evaluation Metrics

We implemented multiple models:

- **TF-IDF + Logistic Regression**
- **TF-IDF + Linear SVM**
- **TF-IDF + Naive Bayes**
- **TF-IDF + Random Forest**
- **Zero-shot Classification (BART)**
- **Fine-tuned RoBERTa Transformer**

The comments were passed into the zero shot model as a mini batch dataset rather than a csv column to improve evaluation speed. Performance was measured using `accuracy_score`, `f1_score`, `precision_score`, `recall_score`, and a full `classification_report` from scikit-learn.

III. Results

A. Data Pipeline and Evaluation

The TF-IDF based models performed almost similarly, with logistic regression performing the best. The Zero-shot BART had lower performance in comparison as expected. The transformer model outperformed the traditional classifiers by a large margin as shown in Table 1.

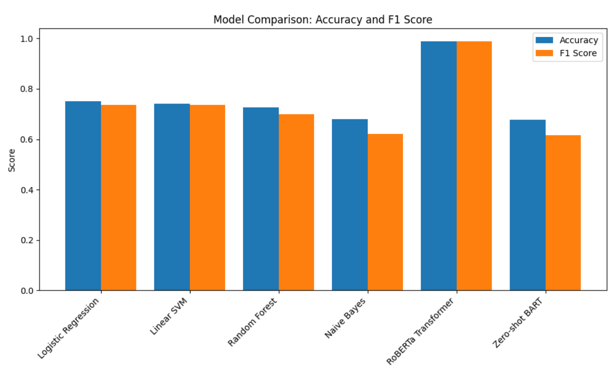

Fig. 2. Model Comparisons with Accuracy and F1 score

TABLE I
PERFORMANCE COMPARISON OF SENTIMENT MODELS

Model	Accuracy	F1-score	Notes
TF-IDF + LR	75%	0.63	Baseline model
TF-IDF + SVM	74%	0.65	Generalized
TF-IDF + NB	73%	0.57	Fast, low precision
TF-IDF + RF	68%	0.47	Ensemble slower
Zero-shot BART	68%	0.62	No fine-tuning
RoBERTa Fine-Tuned	99%	0.99	Best performance

B. Insights

To analyze whether positive sentiment correlates with video engagement, we merged `comments.csv` with `videos-stats.csv` using the `Video ID` column and calculated the average sentiment score (0=negative, 1=neutral, 2=positive) for each video. This was then correlated with metrics such as likes and views.

We computed correlations, finding weak linear correlation but slightly stronger monotonic trends. Visualizations confirmed that videos with more positive comments tended to have higher engagement.

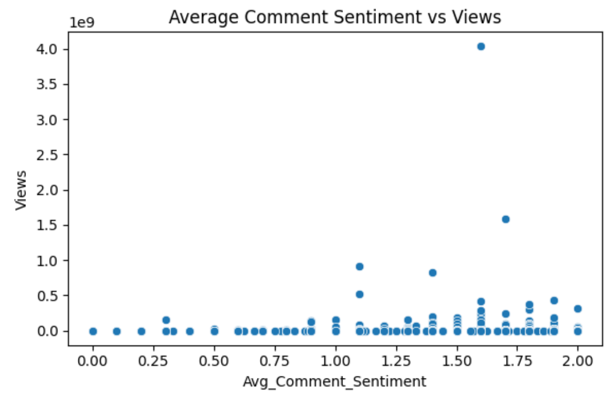

Fig. 3. Average Comment Sentiment vs Video Views

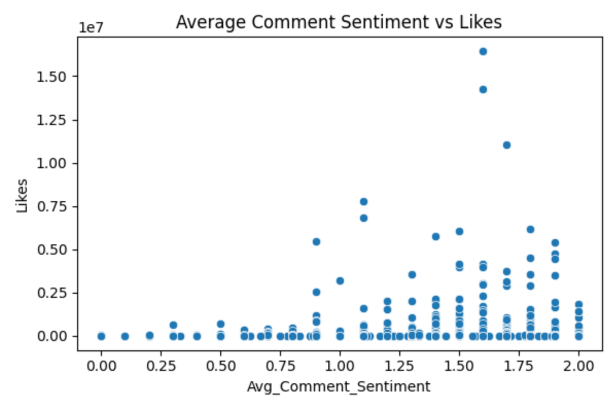

Fig. 4. Average Comment Sentiment vs Video Likes

IV. Conclusion

This project highlights the strength of transformer-based models for sentiment classification of YouTube comments. With superior accuracy and precision, they offer more insightful analytics compared to traditional models. The analysis also suggests that positive sentiment among the top comments is modestly associated with increased engagement. Future work could explore multilingual support and temporal sentiment shifts to understand how public perception evolves over time.

References

[1] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," NAACL, 2019.
[2] Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv:1907.11692, 2019.
[3] Hugging Face, Available: https://huggingface.co/models