

INTRODUCTION

In the digital age, social media platforms generate vast amounts of user-generated content. YouTube, being one of the largest video-sharing platforms, hosts millions of videos with billions of comments. Analyzing these comments can provide valuable insights into user opinions, feedback, and trends.

This project focuses on automated sentiment analysis of YouTube comments to classify them into Positive, Neutral, or Negative categories. By leveraging Natural Language Processing (NLP) techniques, this system extracts and processes comments to provide a sentiment-based summary of a given video.

EXISTING PROJECT

The existing project for analysing YouTube comments are limited in functionality and efficiency. This project can only fetch comments from a single YouTube video at a time, which restricts their usability when working with multiple videos or larger datasets. Moreover, it is constrained by the following issues:

- 1. Limited Comment Extraction: The current project is only capable of fetching a maximum of 100 comments per video. This limitation makes it ineffective for analysing videos with a large volume of comments, which are often required for meaningful sentiment analysis and trend identification.*
- 2. Lack of Scalability: The inability to fetch comments from multiple videos prevents users from performing a comparative analysis or*

gaining insights across a group of related videos. For creators, businesses, or researchers who require large-scale analysis, this limitation proves to be a significant drawback.

3. *Inability to Save Results:* The existing project does not allow users to save the extracted comment data in a structured file format (e.g., CSV, Excel). This restricts users from storing, sharing, or further analysing the data using other tools or platforms.

4. *User Experience Constraints:* The lack of a user-friendly interface and flexible options for specifying the desired number of comments or video URLs further diminishes the usability of the existing system. Users cannot customize their input or output according to their requirements.

PROPOSED PROJECT

The scope of this project is defined to ensure that the system provides a comprehensive solution for fetching, analyzing, and reporting YouTube video comments. The project aims to deliver a tool that is scalable, efficient, and user-friendly for users seeking insights into audience sentiments from YouTube comments.

In-Scope Features:

1. *YouTube Comment Extraction:*

☐ Fetch comments from any YouTube video using its URL.

☐ Allow users to define the number of comments to be extracted.

2. Flexible Comment Limit:

☐ Users can fetch up to thousands of comments (based on API limitations and availability).

3. Data Storage:

☐ Save the fetched comments into a CSV file.

☐ Ensure the stored file is well-structured and compatible with data

analysis tools like Excel, Pandas, and other data visualization libraries.

4. Sentiment Analysis:

☐ Analyse the extracted comments using sentiment analysis techniques.

☐ Categorize the comments into three sentiment types:

☐ Positive

☐ Negative

☐ Neutral

5. Summary Reports:

☐ Generate overall sentiment distribution reports, including:

☐ Total number of comments analysed.

☐ Count of positive, negative, and neutral comments.

☐ Sentiment analysis accuracy.

☐ Provide insights for understanding user engagement and audience reactions.

6. User-Friendly Interface:

☐ Allow users to interact with the system by providing simple inputs

(like YouTube Video URL, desired comment count and file name).

☐ Automate the process of fetching, analysing, and saving results.

PROJECT ANALYSIS

SYSTEM ANALYSIS

System analysis is a critical phase of project development that involves

understanding the functional and non-functional requirements, limitations

of existing solutions, and the enhancements introduced in the proposed

system. It ensures a comprehensive understanding of the system's

architecture and functionalities.

HARDWARE AND SOFTWARE SELECTION

The proper selection of appropriate hardware and software components

determines the efficient development and smooth operation of the project.

Below is a detailed overview of the hardware and software requirements

for implementing this project.

H/W Req..

S/W Req..

Design

5.1 DATA DESIGN

DESIGN

The data design defines how the system organizes and stores data during

the fetching, analysis, and reporting processes.

1. Input Data:

- ☐ *Video URL, number of comments, file name.*

2. Raw Data:

- ☐ *Comments fetched from the YouTube Data API stored temporarily for processing.*

3. Processed Data:

- ☐ *Sentiment scores and classifications (Positive, Negative, Neutral).*
- ☐ *Count of each sentiment category.*

4. Output Data:

- ☐ *CSV file containing comments and sentiment labels.*
- ☐ *Sentiment distribution summary.*

Project Flow

Flow of the YT Sentiment Explorer Project is as follows:

1. User Input

- ☐ *The user enters a YouTube video URL in the web interface.*
- ☐ *The user specifies the number of comments to fetch.*
- ☐ *The user provides a filename to save the results in CSV format.*

2. Extract Video URL

- ☐ *The system extracts the YouTube Video ID from the provided URL using regex.*
- ☐ *If the URL is invalid, an error message is displayed.*

3. Fetch Comments

- ☐ *The system connects to the YouTube API using the provided API key.*
- ☐ *It fetches the specified number of top-level comments from the video.*

4. Perform Sentiment Analysis

☐ *Each comment is processed using VADER Sentiment Analysis to*

classify it as:

☐ *Positive (score ≥ 0.05)*

☐ *Neutral ($-0.05 < \text{score} < 0.05$)*

☐ *Negative (score ≤ -0.05)*

5. Calculate Metrics

☐ *The system counts the number of positive, negative, and neutral*

comments.

☐ *It calculates the accuracy of sentiment classification.*

6. Save & Display Results

☐ *The system saves the comments along with their sentiment labels in*

a CSV file.

☐ *The results are displayed on the web page, including:*

☐ *Total Positive, Neutral, and Negative comments.*

☐ *Overall Accuracy.*

☐ *A Pie Chart showing sentiment distribution.*

7. Download CSV File

☐ *The user can download the CSV file containing the comments and*

their sentiments.

8. Error Handling

☐ *If an incorrect URL is entered, the system prompts an error message.*

Applications of the Project

◆ Content Creators & Influencers

Helps YouTubers understand audience reactions to their videos.

Assists in improving content based on sentiment trends.

◆ *Digital Marketing & Brand Analysis*

Enables businesses to analyze customer feedback on promotional videos.

Helps brands gauge audience perception of their campaigns.

◆ *Social Media Monitoring*

Identifies trends and public opinion on various topics.

Detects potential controversies or negative feedback early.

◆ *Customer Feedback Analysis*

Companies can analyze sentiment on product review videos.

Assists in making data-driven business decisions.

◆ *Academic & Research Purposes*

Useful for sentiment analysis research and NLP-based studies.

Helps in developing AI models for social media insights.

Pro's

Audience Feedback Analysis : Helps creators and brands understand how their audience feels about specific content, products, or events.

Improved Content Strategy : Allows creators to tailor their content based on the feedback and sentiment patterns, boosting engagement.

Brand Monitoring : Useful for monitoring brand reputation and managing customer relations by identifying negative sentiment quickly.

Data-Driven Decisions : Provides data-backed insights, which can assist in making informed decisions about future content or products.

Time Efficiency : It automates the process of scanning and categorizing thousands of comments, saving time compared to manual analysis.