A

Mini Project Report

on

Youtube Sentiment Analyser

Submitted in partial fulfillment of the requirements for the

degree

Third Year Engineering – Computer Science Engineering (Data Science)

by

Sakshi Kadam	**22107032**
Priyanka Barman	**22107004**
Tejas Deshmukh	**22107015**
KisanKumar Jena	**22107049**

Under the guidance of

Ms. Richa Singh

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING (DATA SCIENCE)

A.P. SHAH INSTITUTE OF TECHNOLOGY

G.B. Road, Kasarvadavali, Thane (W)-400615

UNIVERSITY OF MUMBAI

Academic year: 2024-25

CERTIFICATE

This to certify that the Mini Project report on **Youtube Sentiment Analyzer** has been submitted by Sakshi Kadam (22107032), Priyanka Barman (22107004), Tejas Deshmukh (22107015) and KisanKumar Jena (22107049) who are bonafide students of A. P. Shah Institute of Technology, Thane as a partial fulfillment of the requirement for the degree in **Computer Science Engineering (Data Science)**, during the academic year **2024-2025** in the satisfactory manner as per the curriculum laid down by University of Mumbai.

Ms. Richa Singh
Guide

Ms. Anagha Aher
HOD, CSE(Data Science)

Dr. Uttam D. Kolekar
Principal

External Examiner:
1.

Internal Examiner:
1.

Place: A. P. Shah Institute of Technology, Thane
Date:

ACKNOWLEDGEMENT

This project would not have come to fruition without the invaluable help of our guide **Ms. Richa Singh**. Expressing gratitude towards our HOD, **Ms. Anagha Aher**, and the Department of Computer Science Engineering (Data Science) for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our project coordinator **Ms. Sarala Mary** who gave us her valuable suggestions and ideas when we were in need of them. We would also like to thank our peers for their helpful suggestions.

TABLE OF CONTENTS

ABSTRACT

In the era of digital content, YouTube has become a leading platform for video sharing and consumption, where content creators constantly seek to improve audience engagement. This project, titled "YouTube Sentiment Analyzer," is designed to assist YouTubers by providing them with meaningful insights into their channel's performance and viewer interaction. The project focuses on analyzing various metrics such as views, likes, comments, and subscriber growth to help creators better understand their audience and improve their content strategy. Additionally, it offers recommendations for optimizing video titles, enhancing discoverability and viewer interest. Sentiment analysis of user comments is conducted to gauge the overall mood and reactions of the audience, which can be classified as positive, negative, or neutral. Furthermore, the project includes a video summarization feature, which generates concise summaries of videos, making content more accessible and increasing viewer retention. Through these functionalities, the YouTube Sentiment Analyzer aims to provide YouTubers with a powerful toolset to optimize their content creation and audience engagement through data-driven.

Chapter 1

Introduction

The rapid expansion of digital media, YouTube has become one of the largest platforms for content sharing, where millions of creators, commonly known as YouTubers, continuously produce videos to engage with their audiences. The competition for attention on YouTube is fierce, and simply creating content is no longer enough for success. Content creators need to understand how their audience reacts to their videos, what drives engagement, and how to optimize their content strategies to ensure visibility and growth.

While YouTube provides basic analytics like views, likes, and comments, these metrics offer only limited insight into the true sentiments of viewers. To thrive in today's fast-paced digital landscape, YouTubers need more advanced tools that allow them to go beyond surface-level data and truly understand their audience's preferences, emotions, and reactions. The YouTube Sentiment Analyzer is designed to address this need by providing creators with a comprehensive solution that combines data analysis, sentiment detection, and video optimization. This project focuses on several core functionalities: analyzing channel performance, recommending optimized video titles, assessing viewer sentiment through comment analysis, and generating concise video summaries. The Title Recommendation feature, powered by the Naive Bayes classifier algorithm, assists YouTubers in crafting titles that maximize discoverability and engagement. The Sentiment Analysis tool leverages natural language processing (NLP) to assess viewer comments and identify the emotional tone—whether positive, negative, or neutral—allowing creators to gain a deeper understanding of their audience's reactions.

By integrating these features, the YouTube Sentiment Analyzer empowers YouTubers to make data-driven decisions that improve content quality, engage viewers more effectively, and foster channel growth. As the competition for attention on YouTube intensifies, tools like this will prove invaluable in helping creators succeed in the ever-evolving digital content ecosystem.

1.1 Purpose :

The purpose of the YouTube Sentiment Analyzer project is to offer YouTubers a powerful tool that enables them to better understand their audience and optimize their content strategy. In a competitive environment like YouTube, where success relies on effective engagement and consistent growth, basic metrics like views and likes are no longer enough. This project aims to provide a deeper level of analysis by examining viewer behavior, engagement patterns, and emotional responses through advanced data-driven techniques.

A key aspect of the project is helping creators enhance their understanding of how their audience feels about their content. By analyzing comments and engagement data, the tool uncovers the sentiments behind user feedback, offering a clearer picture of whether viewers have a positive, negative, or neutral reaction to the content. Additionally, by addressing audience feedback and optimizing video presentation, creators can improve engagement, foster stronger viewer connections, and ultimately increase audience retention.

In essence, the purpose of the YouTube Sentiment Analyzer is to empower creators with the data and insights they need to make informed, strategic decisions, leading to better content performance and sustainable channel growth.

1.2 Problem Statement :

With the rapid growth of YouTube as a dominant platform for content creation, YouTubers face increasing challenges in maintaining audience engagement and standing out in a highly competitive environment. While YouTube provides basic analytics such as views, likes, and comments, these metrics alone do not offer a comprehensive understanding of how viewers truly feel about the content. Creators lack the tools to assess the emotional reactions of their audience or to efficiently optimize elements such as video titles to improve discoverability.

The lack of a streamlined solution that integrates sentiment analysis, title optimization, and video summarization creates a gap between raw data and actionable insights. This limits the ability of YouTubers to adapt their content based on viewer feedback and make informed decisions to foster long-term growth and engagement. The YouTube Sentiment Analyzer addresses this gap by providing an automated tool that enables YouTubers to gain deeper insights

into audience sentiment, optimize video titles using machine learning algorithms, and generate video summaries, allowing creators to improve their content and grow their channels effectively.

1.3 Objectives :

The YouTube Sentiment Analyzer project is designed to assist content creators in improving their YouTube channels by providing valuable insights and optimizing their content strategy. The project's objectives include:

- **Analyze Channel Performance:** The tool will provide comprehensive analytics on various aspects of a YouTube channel, such as views, likes, comments, and subscriber growth. This analysis helps content creators track their channel's performance, identify successful content trends, and pinpoint areas that need improvement.

- **Sentiment Analysis of Viewer Comments:** Utilizing natural language processing (NLP), the tool will analyze and classify viewer comments as positive, negative, or neutral. This feature allows creators to gauge the emotional tone of their audience's feedback, providing insights into how viewers feel about their videos.

- **Title Recommendation :** The project will incorporate a title recommendation system that to suggest optimized video titles. By analyzing patterns in successful video titles, this system aims to improve video discoverability and increase click-through rates.

- **Video Summarization:** The tool will automatically generate concise summaries of YouTube videos, making it easier for viewers to quickly grasp the main points of the content. This feature enhances content accessibility and can attract new viewers by providing a clear overview of what the video offers.

1.4 Scope:

The YouTube Sentiment Analyzer project aims to empower content creators by offering a range of tools designed to improve channel performance, optimize content, and deepen audience engagement. By leveraging machine learning and data analysis, the project focuses on understanding audience behavior, improving content discoverability, and offering creators insights

into their channel's overall health. The project scope includes several essential features aimed at making content creation and management more efficient and data-driven:

- Channel Performance Analysis: Provide comprehensive analytics on key performance metrics like views, likes, comments, and subscriber growth.
- Sentiment Analysis of Viewer Comments: This enables creators to better understand their audience's reactions and tailor future content based on feedback.
- Title Recommendation: title patterns and suggests improvements that can boost video discoverability and click-through rates.
- Video Summarization: This feature helps potential viewers quickly understand the video's main points without having to watch the full content.
- Engagement Metrics Tracking: Provide data on audience retention, average watch time, and interaction rates to help creators understand how engaging their content is.
- Content Optimization: Offer suggestions on improving video content, including timing for uploads, video length, and content types that resonate well with audiences based on historical data from successful videos.
- Competitor Analysis: Include a feature to analyze competing YouTube channels within the same niche. This will help creators understand what other channels are doing, giving them an edge in staying competitive.
- User Interface and Experience: Develop a highly intuitive and user-friendly interface that enables users to access insights, reports, and recommendations effortlessly, catering to both beginner and advanced content creators.
- Integration with YouTube API: Ensure seamless integration with YouTube's API for real-time data fetching and analysis. This guarantees that creators receive up-to-date insights based on live channel performance.

Chapter 2

Literature Review

The *"Sentiment Analysis for YouTube Videos with User Comment"* by Rawan Fahad Alhujaili and Wael M.S. Yafooz (2021) focuses on analyzing user comments on YouTube to classify them into positive, negative, or neutral sentiments. By doing so, it provides insights into user opinions, helping content creators improve their content strategies and engage better with their audience. The paper employs natural language processing (NLP) techniques and machine learning models to classify comments, comparing the performance of different models to determine which is most accurate for sentiment analysis. The methodology used in this study involves the application of NLP for processing comments and various machine learning models to classify them. The paper evaluates these models based on their accuracy in sentiment classification, which helps in identifying the best approach for sentiment analysis. The merits of this study include providing a deeper understanding of viewer sentiment, allowing content creators to make informed decisions for enhancing their content strategies. Additionally, the paper addresses the challenges associated with sentiment analysis, offering recommendations on the most effective models for classification.

The *"A Classification Scheme for Content Analyses of YouTube Video Comment"* by Amy Madden, Ian Ruthven, and David McMenemy (2017) focuses on developing a detailed classification framework to categorize YouTube comments. The study identifies ten broad categories along with 58 subcategories, aiming to understand the diverse communicative purposes of these comments. This classification helps reveal the various ways users engage with content, whether through criticism, praise, requests, or general conversation. The methodology involved analyzing 66,637 YouTube comments to design a comprehensive schema. This large dataset allowed the authors to ensure that the classification system was thorough and covered the wide variety of comment types found on the platform. The study's merits lie in providing a structured and systematic approach to analyze user-generated content, which benefits researchers across various fields, such as social media analysis, digital communication, and content strategy. The framework makes it easier to extract insights from comments, supporting future research efforts related to YouTube interactions

The*"Engagement and Popularity Dynamics of YouTube Videos and Sensitivity to Meta-Data"* by William Hoiles, Anup Aprem, and Vikram Krishnamurthy (2013) investigates the

impact of meta-level features—such as titles, tags, thumbnails, and descriptions—on the popularity of YouTube videos. It also explores how social dynamics, such as interactions and engagement trends, influence view counts and suggests optimization strategies to enhance video performance. The outcomes of this study help identify the key elements that drive video visibility and popularity.The methodology involved analyzing real-world data from 6 million videos and 25,000 channels to determine how different meta-data and social factors influence engagement. The study focuses on examining the relationship between these factors and metrics like view counts, subscriber growth, and viewer engagement over time.The merits of the paper lie in providing actionable insights for optimizing video meta-data to increase visibility and engagement. It highlights the importance of effective scheduling, maintaining viewer interest through playthrough rates, and building relationships between views and subscriber growth. The findings offer practical recommendations for content creators to strategically manage their content for better performance on YouTube.

Chapter 3

Proposed System

The proposed system is a sophisticated Python-based application tailored to enhance the effectiveness of YouTube content creators. It integrates multiple functionalities into a single, cohesive platform designed to tackle various challenges faced by content creators in managing and optimizing their YouTube presence. The system is built to provide actionable insights and support strategic decision-making.

3.1 Features

- Comprehensive Metrics Tracking: Monitor essential channel metrics such as views, likes, comments, and subscriber growth.
- Trend Visualization: Graphical representations of performance trends to identify patterns and insights.
- Real-Time Suggestions: Provides real-time title suggestions to enhance video discoverability and engagement.
- Comment Sentiment Classification: Categorizes user comments as positive, negative, or neutral.
- Actionable Feedback: Offers insights based on sentiment analysis to refine content strategy.
- Automatic Summaries: Generates brief summaries of video content.
- Summary Customization: Allows customization to focus on key points or specific content segments.

3.2 Functionalities

The proposed system offers a robust suite of features designed to meet the needs of YouTube content creators. It integrates functionalities for analyzing channel performance, optimizing video titles, performing sentiment analysis using the Naive Bayes classifier, and summarizing video content.

Chapter 4

Requirement Analysis

The YouTube Sentiment Analyzer project is a sophisticated application designed to enhance the effectiveness of content creators by providing deep insights into video performance and audience engagement. The system integrates a variety of technologies and libraries to deliver a comprehensive suite of features that address key aspects of content analysis and optimization.

The front-end of the application is constructed using HTML, CSS, and JavaScript. HTML provides the essential structure for the web pages, laying out the foundational elements such as headings, forms, and buttons. This foundational structure ensures that the user interface is both functional and organized. CSS is employed to style these HTML elements, enhancing the visual appeal of the application. It controls the layout, color schemes, fonts, and spacing, ensuring a polished and user-friendly design. JavaScript adds a layer of interactivity, enabling dynamic updates, form validations, and asynchronous content loading. By utilizing JavaScript frameworks or libraries, the application can deliver a more responsive and engaging user experience.

On the back-end, MySQL serves as the relational database management system. It handles the storage and organization of data such as user profiles, video details, and sentiment analysis results. MySQL's reliability and efficiency in data management support the application's performance, ensuring that data is accurately stored and retrieved as needed. Flask, a lightweight Python web framework, manages the server-side operations of the application. It facilitates the creation of RESTful APIs, which enable seamless communication between the front-end and back-end. Flask's simplicity and flexibility make it well-suited for handling HTTP requests, processing data, and integrating with various libraries and services. The core functionalities of the system are built upon advanced technologies and algorithms. For sentiment analysis, the application utilizes the Naive Bayes classifier algorithm. This machine learning technique is well-suited for text classification tasks, allowing the system to categorize YouTube comments into positive, negative, or neutral sentiments. By analyzing these sentiments, content creators can gain valuable insights into viewer reactions and tailor their content strategies to better engage their audience. Video summarization is achieved using transformer models, which are accessed via the Google Transcripts API. This technology processes video transcripts to generate concise

and informative summaries. This feature enables users to quickly understand the main points of a video without needing to watch the entire content, enhancing content management efficiency. Additionally, the application includes a title recommendation feature. This functionality leverages historical data and machine learning techniques to suggest optimized video titles that can improve visibility and viewer engagement. By analyzing past performance and trends, the system provides recommendations that help content creators craft more effective titles.

In essence, the YouTube Sentiment Analyzer integrates these diverse technologies into a cohesive platform that supports content creators in optimizing their YouTube channels. By combining front-end and back-end technologies with advanced machine learning algorithms, the application provides a robust tool for analyzing and improving video content and audience interaction.

Chapter 5

Project Design

Design is the first step in the development phase for any engineering product (or) system. It may be defined as “the process of applying various techniques and principles for the purpose of defining a device, a process, or a system insufficient detail to permit its physical realization.” Software design is an iterative process through which requirements are translated into a Blueprint’ for constructing the software.

5.1 Use Case diagram:

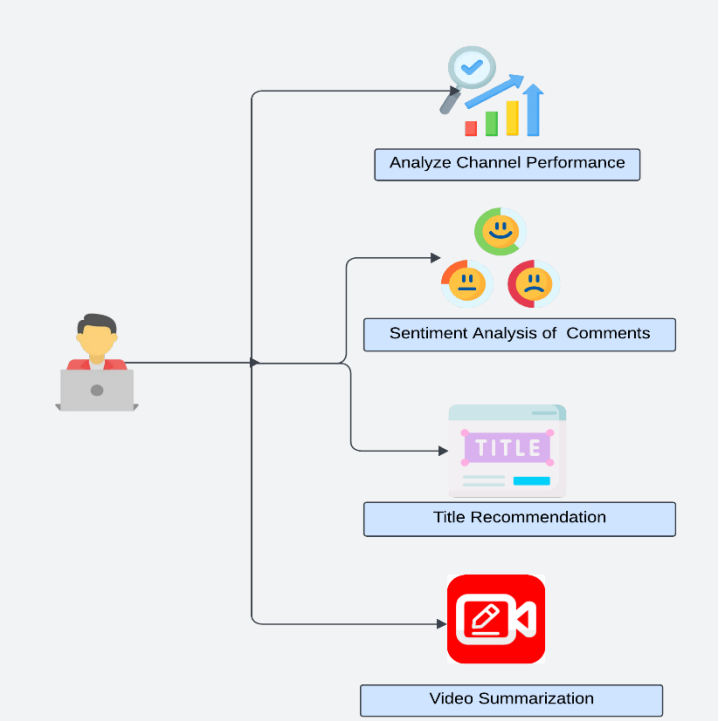


Figure 5.1: Use Case Diagram for our system

In figure 5.1, we can see how The diagram illustrates the process of sentiment analysis on YouTube comments. It starts with collecting YouTube comments and pre-processing them to prepare them for analysis. Then, a subset of the comments is manually labeled with positive, negative, or neutral sentiments. This labeled data is used to train a

machine learning model, which learns to identify patterns and relationships between the words and their corresponding sentiments. The trained model is then tested on a separate set of comments to evaluate its performance. Finally, the model is applied to the entire dataset of YouTube comments to classify each comment as positive, negative, or neutral. This information can be used for various purposes, such as understanding public opinion, market research, and customer feedback analysis

5.2 Data Flow Diagram

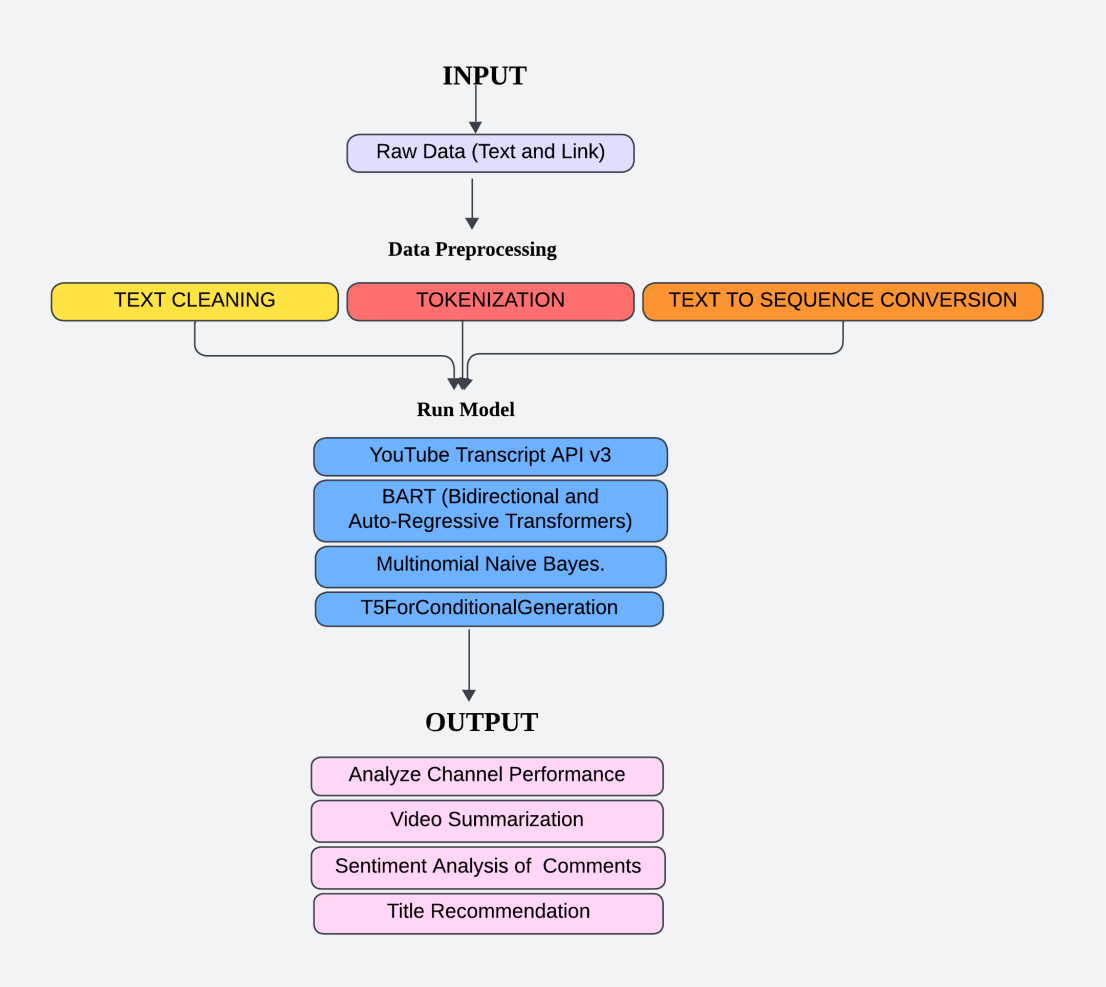


Figure 5.2: DFD (Data Flow Diagram) for our system

In figure 5.2, we can see how the diagram outlines a system for analysing YouTube videos and their comments. It starts by retrieving information about the video, such as its title, description, and upload date, as well as collecting viewer comments. The collected comments are then filtered to remove irrelevant or spam content. Next, the system performs sentiment analysis

on the filtered comments, determining whether they are positive, negative, or neutral. Each comment is classified based on its sentiment, and insights are generated based on the overall sentiment of the video, common themes, and viewer opinions. The results of the analysis, including the video metadata, comments, and insights, are then displayed to the user. Additionally, the system can generate suggestions for alternative titles based on the video content and viewer comments, and it can summarize the main points of the video. Finally, the insights generated can be visualized in various ways, such as charts, graphs, or word clouds, to provide a clear and concise overview of the analysis.

Overall, this system offers a comprehensive analysis of YouTube videos and their comments, enabling users to gain valuable insights into viewer opinions and feedback.

5.3 System Architecture

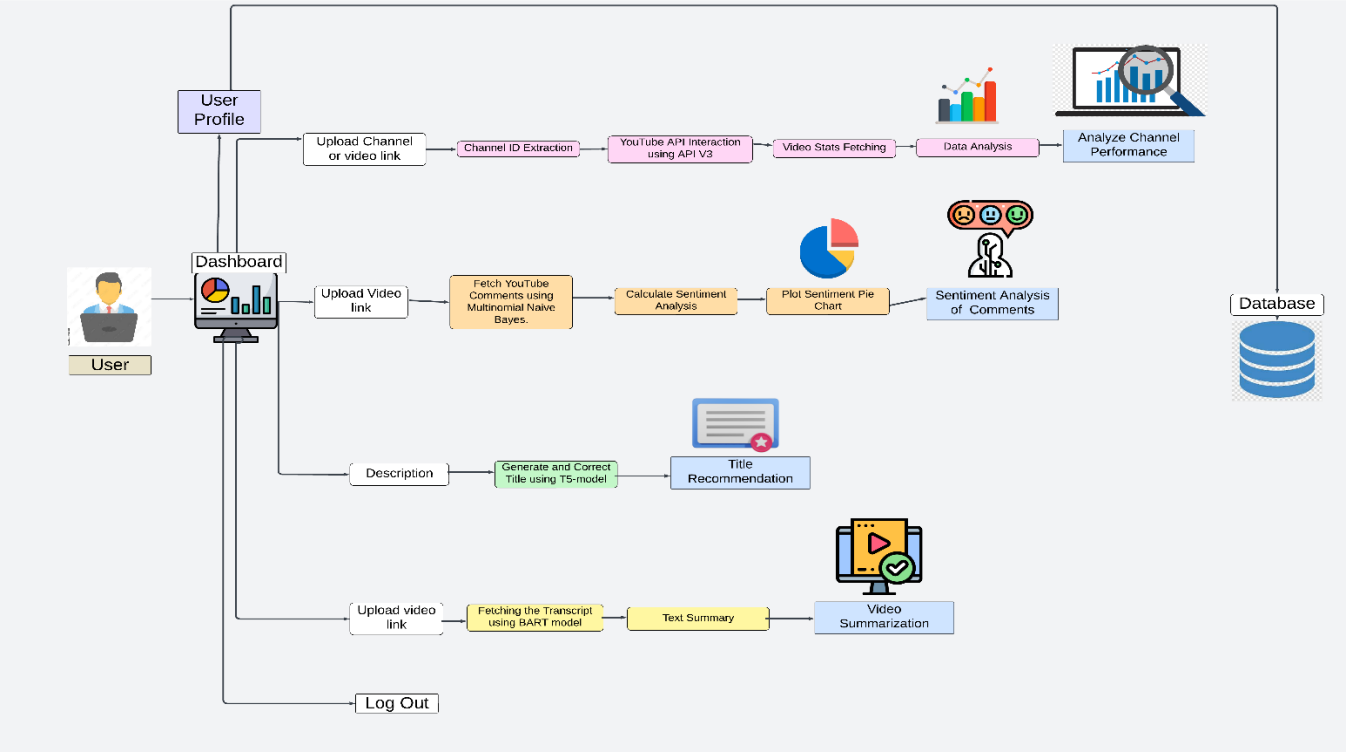


Figure 5.3 System Architecture of our system

The diagram illustrates a system for analysing YouTube videos and providing recommendations. It starts with a user who has a profile and can access a dashboard. The user can either upload a video or enter a channel or video link. Once a video is uploaded or linked, the system analyzes its performance, including metrics like views, likes, and comments. Additionally, it conducts sentiment analysis on the comments to understand viewer opinions. Based on the analysis, the system provides recommendations for titles that might be more engaging. It also summarizes the video content to provide a concise overview. All of this data is stored in a database, which can be used for further analysis or to generate more recommendations. The system also includes a logout feature for user authentication and security.

Overall, this system offers a comprehensive analysis of YouTube videos, providing valuable insights and recommendations for creators to improve their content and reach a wider audience.

5.4 Implementation

Implementing our YouTube sentiment analyzer, fetch comments using the YouTube Data API, preprocess the text by removing stop words and punctuation, analyze sentiment using, and visualize the results. After retrieving the comments, the next step is to preprocess the text by removing stop words and punctuation while also normalizing the text through techniques like lemmatization or stemming, which helps improve the accuracy of your sentiment analysis.Once the text is preprocessed, you can analyze the sentiment using machine learning models such as BERT or RoBERTa, which capture contextual nuances more effectively than traditional methods. After conducting sentiment analysis, you can visualize the results using libraries like Matplotlib or Seaborn, creating informative charts that represent sentiment distribution or trends over time.

Additionally, consider implementing a feedback mechanism that allows users to provide insights on the analysis, helping you fine-tune your sentiment model for better accuracy. As your project grows, think about scalability, potentially using cloud services for hosting to ensure efficient performance with high user traffic.

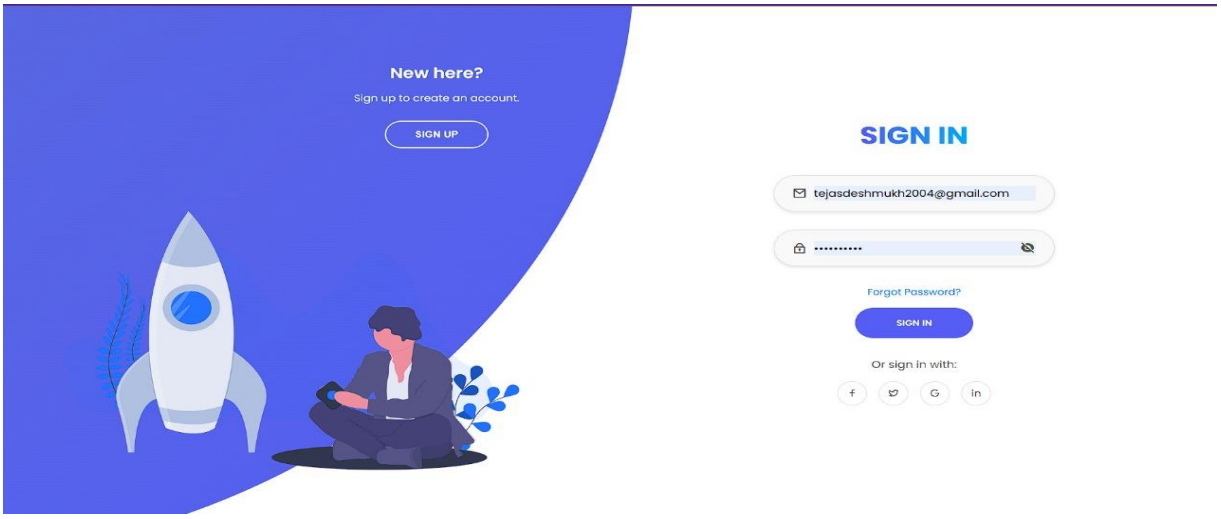


Figure 5.4.1: Account Access Page

This implementation shows a login page for a website or application. The page has two main sections:1. Sign Up: This section is for new users who want to create an account. Clicking on "SIGN UP" will take the user to a registration form where they can provide their personal information.2. Sign In: This section is for existing users who want to access their account. Users can enter their email address and password in the provided fields and click on "SIGN IN" to log in. If they forget their password, they can click on "Forgot Password?" to reset it.The page also provides the option to sign in using social media accounts like Google and Facebook. This allows users to quickly log in without having to create a new account on the website. Overall, the login page provides a simple and convenient way for users to access the website or application.

Figure 5.4.2: User Profile

This implementation shows a user profile page. The page displays the user's profile picture, name, and email address. There is a "Save Changes" button to update the profile information. The page also has options to "Forget Password," "Set Password," "History," and "Logout." The message "Profile updated successfully!" indicates that the user's profile information has been successfully updated.

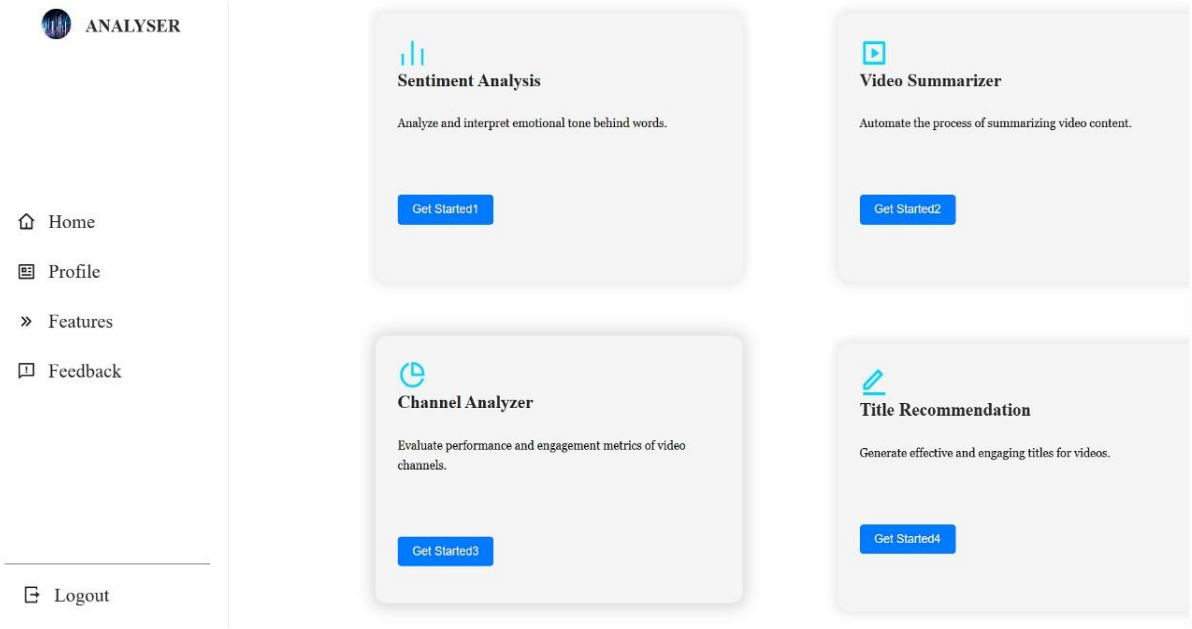


Figure 5.4.3: Dashboard

This implementation depicts the homepage of an online video analysis tool called "ANALYSER." This tool offers a suite of features designed to help users analyse and optimize

their video content.Here's a breakdown of the available features: 1. Sentiment Analysis: This feature allows users to analyze the emotional tone behind words in text data. By understanding the sentiment of viewers' comments or descriptions, creators can gauge the overall reception of their videos and identify areas for improvement.2. Video Summarizer: This feature automates the process of summarizing video content. By generating concise summaries, users can quickly grasp the key points of long videos without having to watch them in their entirety.3. Channel Analyzer: This feature evaluates the performance and engagement metrics of video channels. By analyzing data such as viewership, likes, comments, and shares, creators can gain insights into their audience's preferences and identify successful strategies.4. Title Recommendation: This feature generates effective and engaging titles for videos.

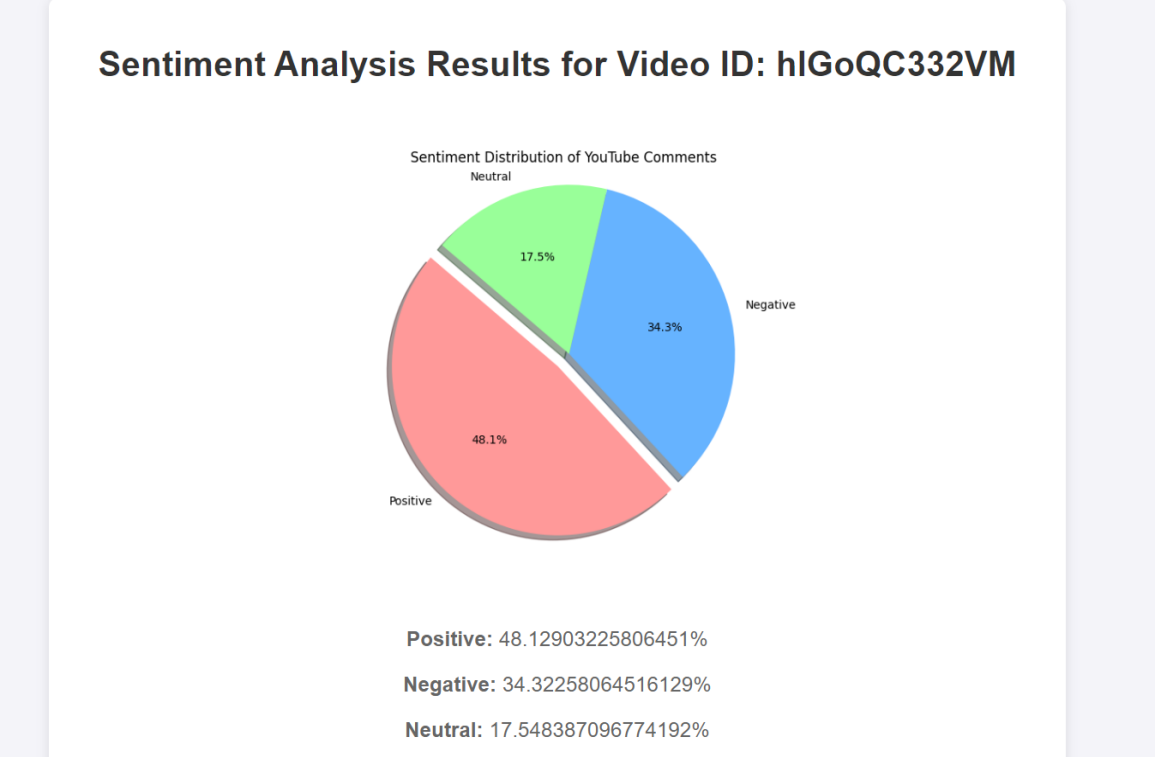


Figure 5.4.4: YouTube Sentiment Analysis

This analysis is used to gauge the overall sentiment expressed in the comments, which can provide valuable insights into the public's perception of the video. Sentiment Distribution The pie chart in the image visually represents the distribution of sentiments among the analyzed comments. It consists of three segments, each corresponding to a different sentiment polarity

1.Positive: This segment represents the percentage of comments that express a positive sentiment towards the video. The color used for this segment is typically a shade of green or blue. 2.Negative**:** This segment represents the percentage of comments that express a negative sentiment towards the video. The color used for this segment is typically a shade of red or orange. 3.Neutral: This segment represents the percentage of comments that do not express a clear positive or negative sentiment towards the video. The color used for this segment is typically a shade of gray or yellow.

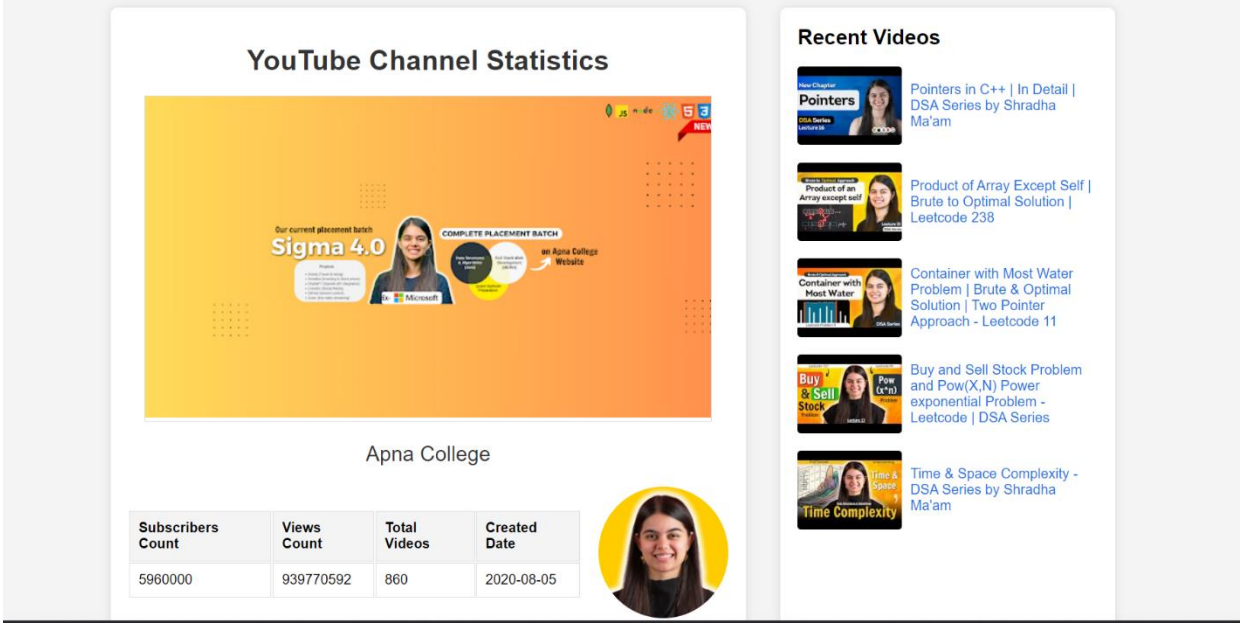


Figure 5.4.5: Channel Statistics

To find detailed information about a YouTube video, go to YouTube and search for it using the search bar at the top of the page. Once you find the video you want, click on it to start playing it. While the video is playing, you can find detailed information about it in the video description. The description typically includes the title of the video, a detailed explanation of its content, the date it was uploaded, the number of times it has been viewed, the number of likes and dislikes it has received, the comments section where viewers can leave their thoughts and feedback, and information about the channel that uploaded the video, such as the channel name, subscriber count, and other relevant details.

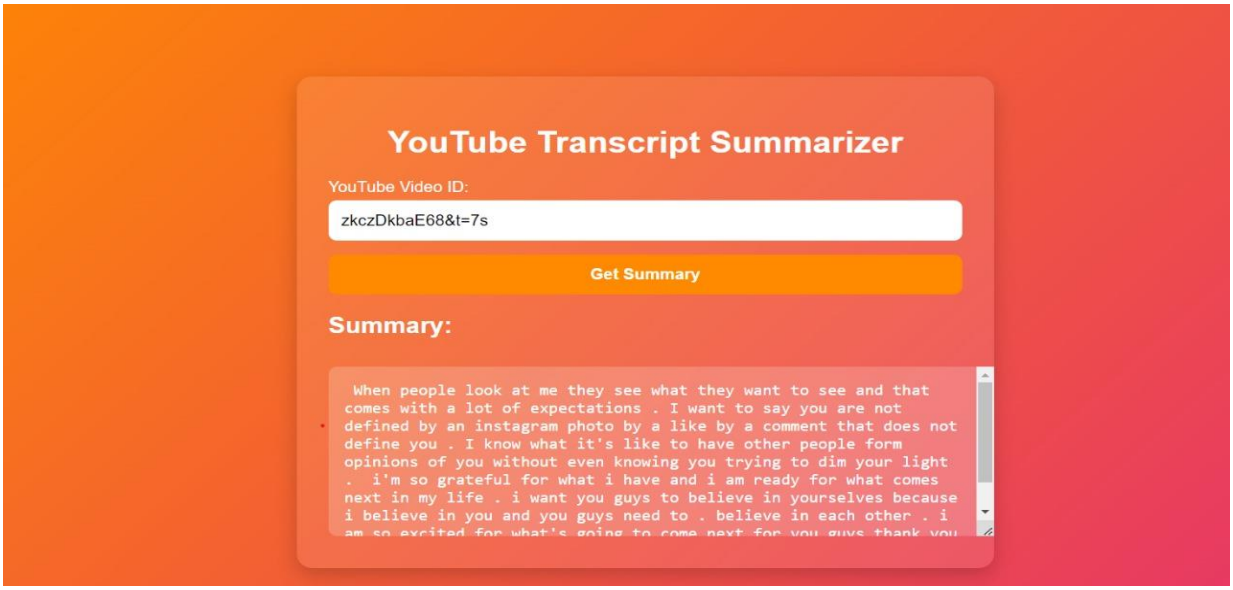


Figure 5.4.6:Youtube Transcript Summarizer

This implementation you provided depicts a web application designed to summarize YouTube transcripts. This tool can be useful for quickly understanding the key points of a video without having to watch the entire content. Here's a breakdown of how the application works: 1.YouTube Video ID Input: The user is required to enter the YouTube video ID into the designated field. The video ID is a unique identifier that can be found in the URL of the video. 2.Get Summary Button: Once the video ID is entered, the user clicks the "Get Summary" button to initiate the summarization process. 3.Summary Generation: The application then processes the transcript of the specified video and generates a concise summary of its content. 4.Summary Display: The generated summary is displayed below the input field. This summary highlights the main points and ideas expressed in the video's transcript.

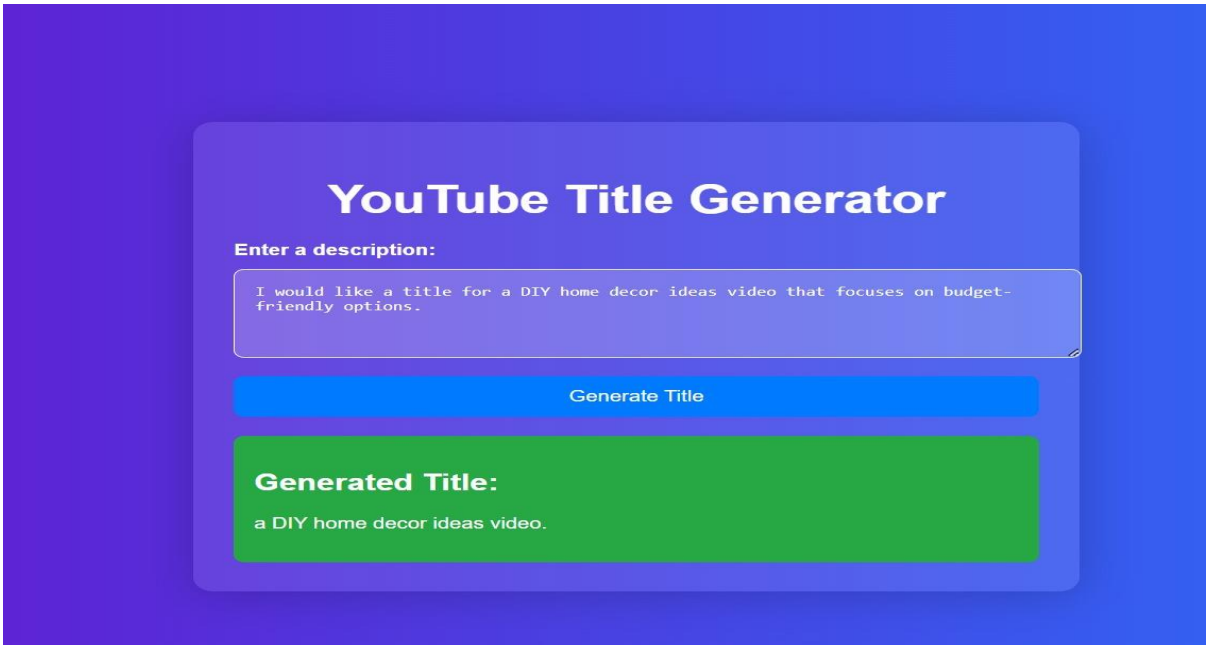


Figure 5.4.7 :Youtube Title Generator

This implementation shows a web application that generates YouTube video titles based on a given description. The user enters a description of their video in the provided field, and then clicks the "Generate Title" button. The application then generates a suggested title for the video, which is displayed in the "Generated Title" field. In the example shown in the image, the user entered a description requesting a title for a DIY home decor ideas video that focuses on budget-friendly options. The generated title is simply "a DIY home decor ideas video," which is not very informative or engaging. The application could be improved by generating more creative and descriptive titles.

Chapter 6

Technical specification

1. **Frontend Technology**

   - HTML: Used to structure the web pages, providing the layout and content display.
   - CSS: Used for styling the web pages, ensuring a visually appealing user interface.
   - JavaScript: Used to implement interactive features, form validation, and dynamic content manipulation on the client side.

2. **Backend Technology**

   - Flask: A lightweight Python web framework used to manage the server-side logic and handle HTTP requests from the frontend. It facilitates the connection between the frontend and backend, serving the results of all features.
   - MySQL: A relational database management system used to store and manage user login details.

3. **Algorithms and Models**

   - Multinomial Naive Bayes (Sentiment Analysis):
     - The sentiment analysis model is implemented using the Multinomial Naive Bayes algorithm. It classifies YouTube comments as positive, negative, or neutral based on the training data.
     - YouTube Data API: The API is used to fetch video comments dynamically for analysis. Comments are retrieved and processed as input to the sentiment analysis model.

- T5 Base Model (Title Recommendation):
    - The T5 (Text-To-Text Transfer Transformer) base model is used to generate video title recommendations based on video descriptions. The model is fine-tuned to enhance performance on the dataset obtained via YouTube's API.
- BART (Bidirectional and Auto-Regressive Transformers:
    - The BART model is a state-of-the-art approach for text summarization due to its combination of bidirectional encoding and auto-regressive decoding. It excels in generating accurate, fluent summaries of long texts like YouTube video transcripts, making it an ideal choice for your project. Its robust pretraining on noisy data also ensures that it handles the variability and complexity of real-world language data effectively.

4. External APIs

- YouTube Data API:
    - Utilized to fetch video metadata (comments, descriptions,channel data) required for both sentiment analysis and title recommendation functionalities and channel statistics.

5. Deployment Environment

- The application is hosted locally using Flask for backend server management.
- The database is deployed using MySQL, with Flask managing the connection and queries.

Chapter 7

Project Scheduling

In project management, a schedule is a listing of a project's milestones, activities, and deliverables. Usually, dependencies and resources are defined for each task, then start and finish dates are estimated from the resource allocation, budget, task duration, and scheduled events. A schedule is commonly used in the project planning and project portfolio management parts of project management. The development and maintenance of the project schedule is the responsibility of a full-time scheduler or team of schedulers, depending on the size and the scope of the project. The project schedule is a calendar that links the tasks to be done with the resources that will do them. It is the core of the project plan used to show the organization how the work will be done, commit people to the project, determine resource needs, and used as a kind of checklist to make sure that every task necessary is performed.

A Gantt chart is a type of bar chart that illustrates a project schedule. Modern Gantt charts also show the dependency relationships between activities and the current schedule status. This chart lists the tasks to be performed on the vertical axis, and time intervals on the horizontal axis. The width of the horizontal bars in the graph shows the duration of each activity. Gantt charts illustrate the start and finish dates of the terminal elements and summary elements of a project. Terminal elements and summary elements constitute the work breakdown structure of the project. Modern Gantt charts also show the dependency (i.e., precedence network) relationships between activities. Gantt charts can be used to show current schedule status using percent-complete shadings.

Sr.No	Group Member	Time duration	Work to be done
1	Priyanka Barman Tejas Deshmukh Sakshi Kadam KisanKumar Jena	1st and 2nd week of July.	Group Formation, Topic finalization and Identify Objectives.
		3rd and 4th week of July.	Identify Functionalities and discuss the project with a paper prototype.
2	KisanKumar Jena Sakshi Kadam Tejas Deshmukh Priyanka Barman	1st and 2nd week of August.	Implementation of Graphical User Interface (GUI).
		3rd and 4th week of August.	Connections of all the GUI pages and Presentation I.
3	Sakshi Kadam Priyanka Barman Tejas Deshmukh KisanKumar Jena	1st, 2nd and 3rd week of September	Implementing the machine learning algorithm required for the project
		4th week of September.	Database Design and Database Connectivity.
4	Tejas Deshmukh Priyanka Barman KisanKumar Jena Sakshi Kadam	1st week of October.	Report Writing. and Presentation II

Table 7.1: Project Task Distribution

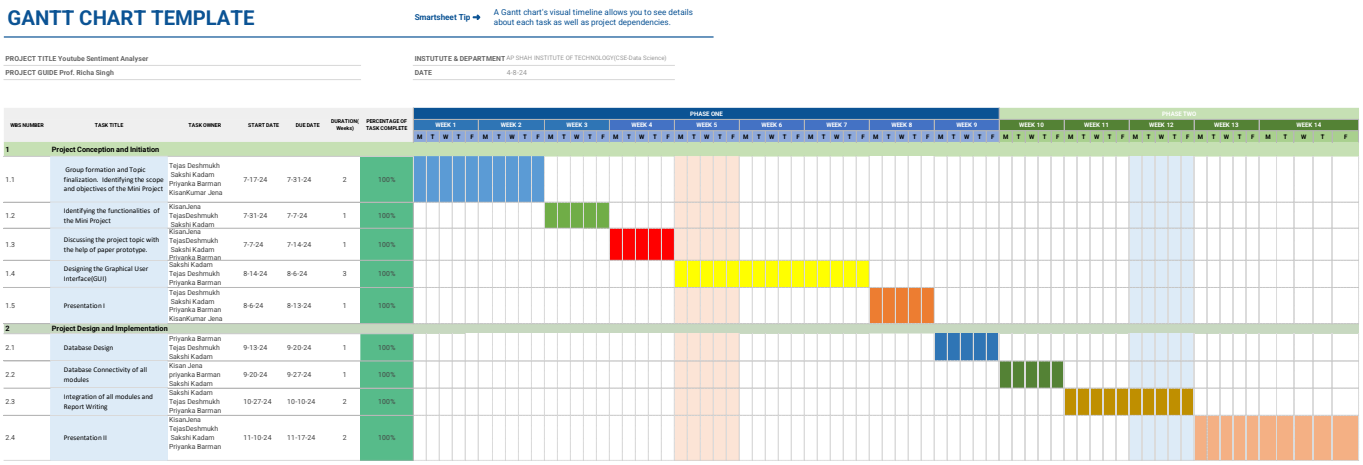


Figure 7.2 : Gantt Chart

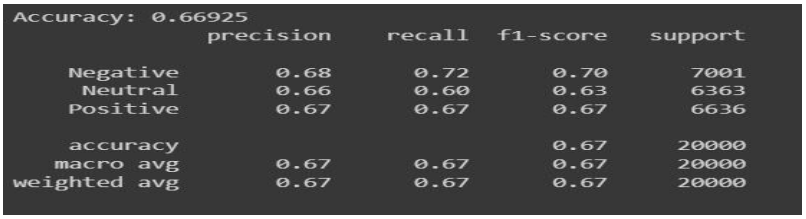
The Gantt chart represents a detailed timeline of the project, breaking it into two main phases: Project Conception and Initiation and Project Design and Implementation. In the first phase, tasks such as group formation, feasibility study, and project proposal are assigned to Sakshi Kadam, KisanKumar Jena, Tejas Deshmukh and Priyanka Barman, with some tasks already completed, like presentations. The second phase involves more technical work, such as database design, which is set to take place in weeks 6-8, and the development of core functionality, spanning weeks 8-12. Each task is marked with different colors to indicate the work duration, and the timeline extends across 14 weeks to ensure smooth progress from planning to final implementation

Chapter 8

Results

The proposed system, designed to enhance the efficiency of YouTube content creators, successfully integrates multiple advanced functionalities. The results demonstrate the effectiveness of the implemented models and features, contributing to better decision-making and content optimization for YouTube creators.

- **Sentiment Analysis**: The sentiment analysis model, powered by the Multinomial Naive Bayes algorithm, efficiently categorized user comments into positive, neutral, and negative sentiments. Insights: The sentiment analysis provided creators with actionable insights to tailor their content based on user feedback, fostering audience engagement and improving channel performance.

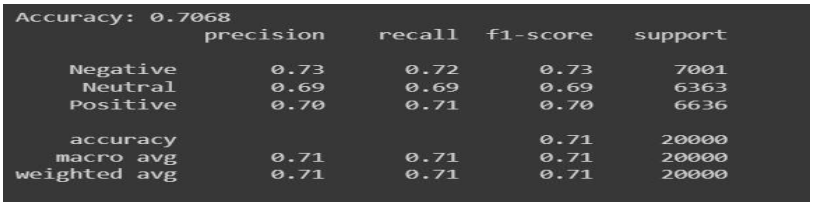


Accuracy: 0.66925

	precision	recall	f1-score	support
Negative	0.68	0.72	0.70	7001
Neutral	0.66	0.60	0.63	6363
Positive	0.67	0.67	0.67	6636
accuracy			0.67	20000
macro avg	0.67	0.67	0.67	20000
weighted avg	0.67	0.67	0.67	20000

8.1 Sentiment analysis model using the Naive Bayes algorithm

The evaluation metrics for a sentiment analysis model using the Naive Bayes algorithm. This overall metric indicates that the model correctly classified 66.925% of the total instances in the dataset.

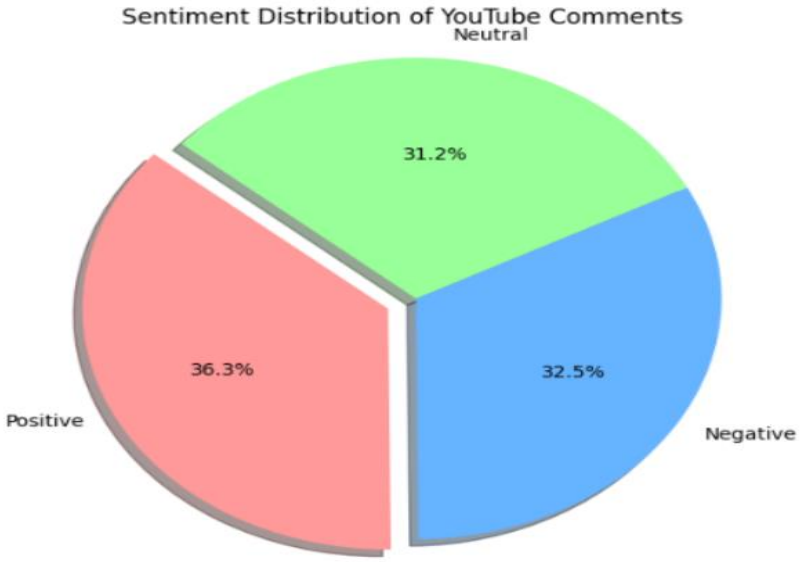


Accuracy: 0.7068

	precision	recall	f1-score	support
Negative	0.73	0.72	0.73	7001
Neutral	0.69	0.69	0.69	6363
Positive	0.70	0.71	0.70	6636
accuracy			0.71	20000
macro avg	0.71	0.71	0.71	20000
weighted avg	0.71	0.71	0.71	20000

8.2 Sentiment analysis model using logistic regression

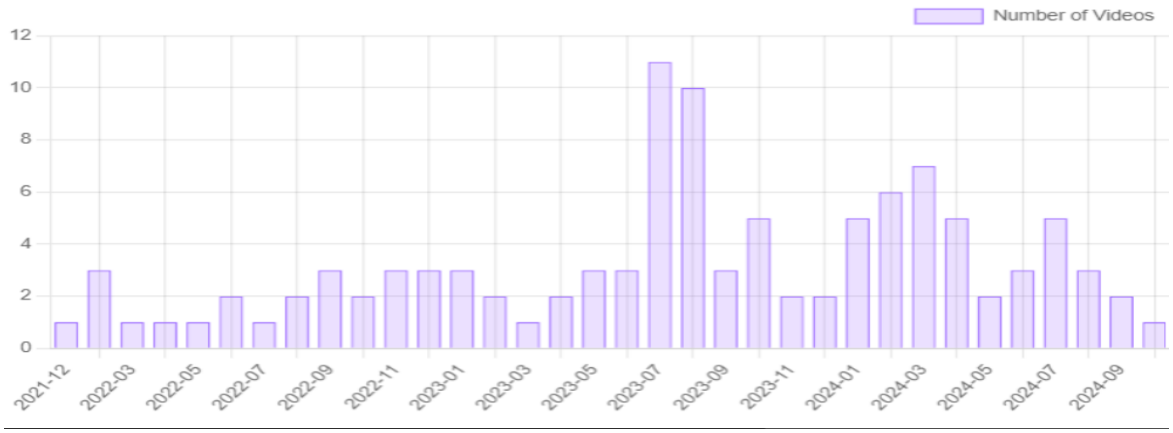
the evaluation metrics for a sentiment analysis model using logistic regression. This overall metric indicates that the model correctly classified 70.68% of the total instances in the dataset



8.3 Sentiment Distribution of Youtube Comments

The pie chart displays the sentiment distribution of YouTube comments, divided into positive, negative, and neutral categories. Positive comments make up the largest portion at 36.3%, indicating a generally favorable response. Negative comments account for 32.5%, showing a significant proportion of criticism or dissatisfaction. Meanwhile, 31.2% of the comments are neutral, suggesting a balance of objective or indifferent feedback. This distribution helps analyze audience perception, identifying both areas of success and those needing improvement. Such insights are valuable for content creators to adjust their strategies and enhance viewer engagement.

- **Channel Performance Analysis**: The system effectively tracked channel metrics such as views, likes, comments, and subscriber growth. Trend Visualization: Graphical representations allowed creators to identify patterns in content performance and make data-driven decisions to enhance content strategy. Growth Insights: The platform enabled creators to monitor trends in audience behavior and content performance over time, facilitating the identification of successful content strategies.



8.4 No. of videos per month

The chart shows the number of videos uploaded per month, highlighting the channel's content strategy and consistency. Peaks between May and July 2023 suggest targeted efforts, such as seasonal campaigns or trending topics, to attract viewers. In contrast, months with fewer uploads may indicate breaks, resource constraints, or shifting priorities. Irregular posting can affect audience retention and visibility in YouTube's algorithms, which favor consistent uploads. The steady activity seen in early 2024 suggests a renewed focus on content creation. Overall, maintaining consistency is crucial for sustaining growth and engagement

- **Title Recommendation**: The T5 Base Model used for title recommendation successfully generated engaging and optimized titles based on video descriptions.
- **Effectiveness**: The system provided real-time suggestions aimed at improving video discoverability and increasing viewer engagement.
- **User Feedback**: Creators found the title recommendations useful for crafting attention-grabbing titles that aligned with YouTube's SEO practices.
- **Video Summarization**: The video summarizer feature generated concise summaries of video content, providing users with an efficient way to understand key points without watching the full video.
- **Time Efficiency**: This feature significantly reduced the time required to grasp the essential content, making it highly valuable for users looking to quickly assess video relevance.
- **Customization**: The ability to customize summaries based on specific content segments allowed creators to highlight critical sections of their videos, enhancing user experience.

Chapter 9

Conclusion

The YouTube Sentiment Analyser is a powerful and innovative tool that empowers content creators by providing critical insights into audience sentiment. Through features such as sentiment analysis, title optimization, video summarization, and metrics tracking, the platform equips creators with actionable data to refine their content strategies. By analyzing user feedback, monitoring channel growth, and optimizing video discoverability, the system fosters a deeper understanding of audience preferences.

The real-time suggestions and summaries enhance creators' productivity by reducing the time spent on content management and enabling strategic decision-making. As a result, creators can focus on producing more engaging content, thereby boosting audience retention and fostering growth. In the highly competitive landscape of digital media, tools like the YouTube Sentiment Analyser are indispensable for creators looking to achieve sustained success. By leveraging advanced machine learning models such as Multinomial Naive Bayes for sentiment analysis and T5 base model for title generation, the platform ensures that creators stay ahead of trends and audience expectations, optimizing their content for maximum impact.

The proposed system not only simplifies the management of YouTube channels but also contributes significantly to the optimization of content strategy, audience engagement, and long-term growth for content creators.

Chapter 10

Future Scope

The future scope of the YouTube Sentiment Analyser project includes significant advancements aimed at enhancing functionality and expanding its usability. One of the key areas of improvement is advanced sentiment analysis, where the system can be expanded to support multiple languages, thereby catering to a broader global audience. In addition to basic sentiment analysis, emotion detection can be integrated, allowing the tool to identify specific emotions such as joy, anger, or sadness. This would provide much deeper feedback on viewer reactions, making the insights far more granular. Real-time sentiment analysis during live streams is another potential addition, which would enable content creators to engage dynamically with their audience as events unfold.

Personalized insights represent another future possibility for the project. By analyzing audience demographics and engagement patterns, the system could provide highly targeted content strategies. This would include offering recommendations on video lengths, ideal posting schedules, and trending topics that align with audience interests, thus helping creators make data-driven decisions. Additionally, the project could evolve into a cross-platform analysis tool, extending its capabilities beyond YouTube to platforms like Instagram and Twitter. API integrations could be added to facilitate competitor analysis and deliver a more comprehensive view of audience engagement across various social media outlets.

To accommodate the increasing volume of data and provide real-time insights, the system could be moved to a cloud-based infrastructure, ensuring greater scalability and improved performance. Additionally, continuous updates to the machine learning models would allow for better accuracy and adaptability, ensuring that the system evolves alongside changing user needs and trends.

References

[1] Madden, A., Ruthven, I., & McMenemy, D. (2013). A classification scheme for content analyses of YouTube video comments. Journal of Documentation, 69(5), 693-714. https://doi.org/10.1108/JD-06-2012-0078

[2] Asghar, M.Z., Ahmad, S., Marwat, A., Kundi, F.M.: "Sentiment Analysis on YouTube: A Brief Survey," *arXiv preprint*, arXiv:1511.09142, 2015.

[3] Sarakit, P., Theeramunkong, T., Haruechaiyasak, C., Okumura, M.: "Classifying Emotion in Thai YouTube Comments," *6th International Conference on Information and Communication Technology for Embedded Systems (IC-ICTES)*, IEEE, 2015.

[4] G. Gürsun, M. Crovella, and I. Matta, “Describing and forecasting video access patterns,” inProc. IEEE INFOCOM, 2011, pp. 16–20. Google Scholar

[5] Prabha, M.I., Srikanth, G.U.: "Survey of Sentiment Analysis Using Deep Learning Techniques," *1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, IEEE, 2019.

[6] Jelodar, H., Wang, Y., Rabbani, M., et al.: "A NLP Framework Based on Latent-Topic Detection and Sentiment Analysis via Fuzzy Lattice Reasoning on YouTube Comments," *Multimedia Tools and Applications*, 2021.

[7] Rawan Fahad Alhujaili, Wael M.S. Yafooz, "Sentiment Analysis for YouTube Videos with User Comments: Review," IEEE Xplore, 2021. DOI: 10.1109/ICAIS50930.2021.9396049.