

ECHO INSIGHTS

YouTube Sentiment Analysis App

A PROJECT REPORT

Submitted by

SENTHILNAATHAN K (2116210701238)
SHARAN ADHITYA C D (2116210701241)
SANJEEV U (2116210701232)

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



RAJALAKSHMI ENGINEERING COLLEGE
ANNA UNIVERSITY, CHENNAI
MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this project titled “**ECHO INSIGHTS, YOUTUBE SENTIMENT ANALYSIS APP**” is the bonafide work of “**SENTHILNAATHAN K (2116210701238)**”, “**SHARAN ADHITYA C D (2116210701241)**”,

“**SANJEEV U (2116210701241)**” who carried out the work under my supervision. Certified further that to the best of my knowledge, the work reported herein does not form part of any other thesis or dissertation based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Mr. Vijay K

PROJECT COORDINATOR

Assistant Professor,

Department of Computer Science and Engineering

Rajalakshmi Engineering College, Chennai - 602 105

Submitted to Project Viva-Voce Examination held on _____

Internal Examiner

External Examiner

ABSTRACT

In an era where social media proves greater, our innovative tool offers a streamlined solution for dissecting the sentiment behind YouTube videos. The proposed project aims to develop a simple social media analytical tool names ECHO INSIGHTS focused on YouTube video sentiment analysis.

The tool will leverage the YouTube API to extract comments from a given video link, and create a data frame with these and translate these comments into English for accessibility of wide audience, and subsequently perform sentiment analysis using pre-trained Vader bots.

The results will be presented in an interactive dashboard created with Streamlit, allowing users to easily visualize and interpret sentiment patterns, presenting the worst, best comments with overall sentiment of the video. Also, word cloud is presented that portrays the maximum used words in comments providing a quick glance over mostly used words.

Overall, ECHO INSIGHTS aims to empower users with a comprehensive understanding of audience sentiment and engagement on YouTube videos, enabling more informed decision-making and content strategy development.

ACKNOWLEDGEMENT

First, we thank the almighty god for the successful completion of the project. Our sincere thanks to our chairman **Mr. S. Meganathan B.E., F.I.E.**, for his sincere endeavor in educating us in his premier institution. We would like to express our deep gratitude to our beloved Chairperson **Dr. Thangam Meganathan Ph.D.**, for her enthusiastic motivation which inspired us a lot in completing this project and Vice Chairman

Mr. Abhay Shankar Meganathan B.E., M.S., for providing us with the requisite infrastructure.

We also express our sincere gratitude to our college Principal,

Dr. S. N. Murugesan M.E., PhD., and **Dr. P. KUMAR M.E., PhD**, **Director computing and information science** , and **Head Of Department of Computer Science and Engineering** and our project coordinator **Mr. Vijay K** , for his encouragement and guiding us throughout the project towards successful completion of this project and to our parents, friends, all faculty members and supporting staffs for their direct and indirect involvement in successful completion of the project for their encouragement and support.

SENTHILNAATHAN K

SHARAN ADHITYA C D

SANJEEV U

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	iii
1.	INTRODUCTION	1
	1.1 Problem statement 1.2 Scope of work 1.3 Aim and Objectives of the project 1.4 Resources	
2.	LITERATURE SURVEY	4
3.	SYSTEM DESIGN	5
	3.1 General 3.2 System Architecture diagram 3.3 Development Environment 3.3.1 Hardware Requirements 3.3.2 Software Requirements	
4.	PROJECT DESCRIPTION	7
	4.1 Methodology 4.2 Module Description	
5.	RESULTS AND DISCUSSIONS	9
	5.1 Output Images 5.2 Result	
6.	CONCLUSION AND FUTURE ENHANCEMENTS	14
	6.1 Conclusion 6.2 Future Enhancement	
	APPENDIX	16
	REFERENCES	21

CHAPTER 1

INTRODUCTION

In today's fast-paced world, Social media sites like YouTube are effective means for content producers to interact with their viewers in the current digital era. To properly customize their content strategies, marketers and content creators must have a thorough understanding of audience sentiment and engagement dynamics. The proposed project, ECHO INSIGHTS, provides a thorough social media analytical tool targeted at YouTube videos in order to meet this demand.

To give consumers practical insights into audience interactions, ECHO INSIGHTS makes use of cutting-edge technology including the YouTube API, sentiment analysis algorithms, and interactive data visualization tools. Through the process of extracting, translating, and analyzing comments from YouTube videos, ECHO INSIGHTS allows users to identify important topics and sentiment patterns that emerge from audience interactions.

Streamlit was used to create an intuitive dashboard interface that allows users to view word clouds representing popular themes, analyze sentiment trends, and access comprehensive sentiment analysis results. To help with additional analysis and documentation, ECHO INSIGHTS also provides easy ways to print or save analysis findings as PDFs.

Content producers, marketers, and researchers can obtain more profound understanding of the attitudes and behaviors of their YouTube audience using ECHO INSIGHTS. This will enable them to make well-informed decisions and maximize their content strategies for maximum effect and success

1.1 PROBLEM STATEMENT

To Create an AI-powered sentiment analysis solution specialized in interpreting emotions within social media content, empowering individuals and organizations to manage their online reputation and perception effectively.

How might we create an AI-powered sentiment analysis solution specialized in interpreting emotions within social media content? Empowering individuals and organizations to manage online reputation and perception effectively.

1.2 SCOPE OF WORK

The scope of this project is to develop a comprehensive social media analytical tool centered on YouTube videos, ECHO INSIGHTS entails integrating multiple technologies and functionalities." The tool will translate comments into English and retrieve them from videos using a smooth connection with the YouTube API. This will assure accessibility.

The emotional tone and sentiment of comments will be analyzed by sophisticated sentiment analysis algorithms, including Vader bots that have been taught, to provide insightful information about audience reactions. The main component of the service will be an interactive dashboard created with Streamlit, which enables users to examine word clouds that represent hot subjects in comments and see sentiment trends over time.

1.3 AIM AND OBJECTIVES OF THE PROJECT

The ECHO INSIGHTS project aims to provide an approachable social media analysis tool with a YouTube video focus. Through the use of interactive data visualization tools, sentiment analysis algorithms, and the YouTube API, the project seeks to offer practical insights into the dynamics of audience interaction and sentiment.

The goal of this project is to develop an extensive tool for YouTube comment analysis. In order to effectively extract comments from videos while respecting API rate limitations, it will make use of the YouTube API. The program will translate comments into English to help with accessibility and analysis in the event of a language barrier. Sentiment analysis will be used to classify comments according to emotions, providing insights into the responses of the audience. Users will be able to observe sentiment patterns and actively explore the data using an interactive dashboard that was created with Streamlit.

Finally, a well-defined project management plan with clear milestones and timelines will ensure efficient collaboration and timely completion.

1.4 RESOURCES

This project has been developed through widespread secondary research of accredited manuscripts, standard papers, business journals, white papers, analysts' information, and conference reviews. Significant resources are required to achieve an efficacious completion of this project.

The following prospectus details a list of resources that will play a primary role in the successful execution of our project:

A properly functioning workstation (PC, laptop, net-books etc.) to carry out desired research and collect relevant content.

Stable internet connection throughout the process.

CHAPTER 2

LITERATURE SURVEY

The realm of social media analytics and audience engagement presents a dynamic landscape where the interaction between content producers and viewers plays a pivotal role in shaping digital content strategies. In response to the growing importance of understanding audience sentiment and engagement dynamics on platforms like YouTube, recent literature has explored innovative approaches leveraging technology to provide valuable insights to marketers and content creators.

Research by Lee et al. (2020) delves into the significance of sentiment analysis in understanding audience reactions on social media platforms, highlighting the role of machine learning algorithms in extracting and analysing textual data to uncover sentiment trends and patterns. Similarly, the work of Wang and Zhang (2019) explores the application of sentiment analysis techniques in social media analytics, demonstrating its effectiveness in gauging audience sentiment and identifying key themes from user-generated content.

Furthermore, studies by Chen et al. (2018) and Liu et al. (2021) investigate the utilization of natural language processing (NLP) techniques in social media data analysis, emphasizing the importance of advanced algorithms in processing and interpreting textual data to extract valuable insights. These studies underscore the potential of NLP-driven approaches in uncovering audience sentiments and providing actionable insights for content creators and marketers.

Moreover, research by Liang and Chen (2017) examines the role of interactive data visualization tools in social media analytics, emphasizing the importance of user-friendly interfaces in facilitating data exploration and interpretation. This highlights the significance of tools like Streamlit in providing intuitive interfaces for users to interact with social media analytics data effectively.

In addition, ethical considerations surrounding data privacy and transparency in social media analytics are addressed in studies by Zhang et al. (2019) and Kim et al. (2020), highlighting the importance of ethical guidelines and regulatory frameworks in ensuring responsible use of user-generated data for analytics purposes.

Overall, the literature survey underscores the multifaceted landscape of social media analytics and audience engagement, showcasing the importance of advanced technologies such as sentiment analysis, NLP, and interactive data visualization tools in providing actionable insights to content producers, marketers, and researchers.

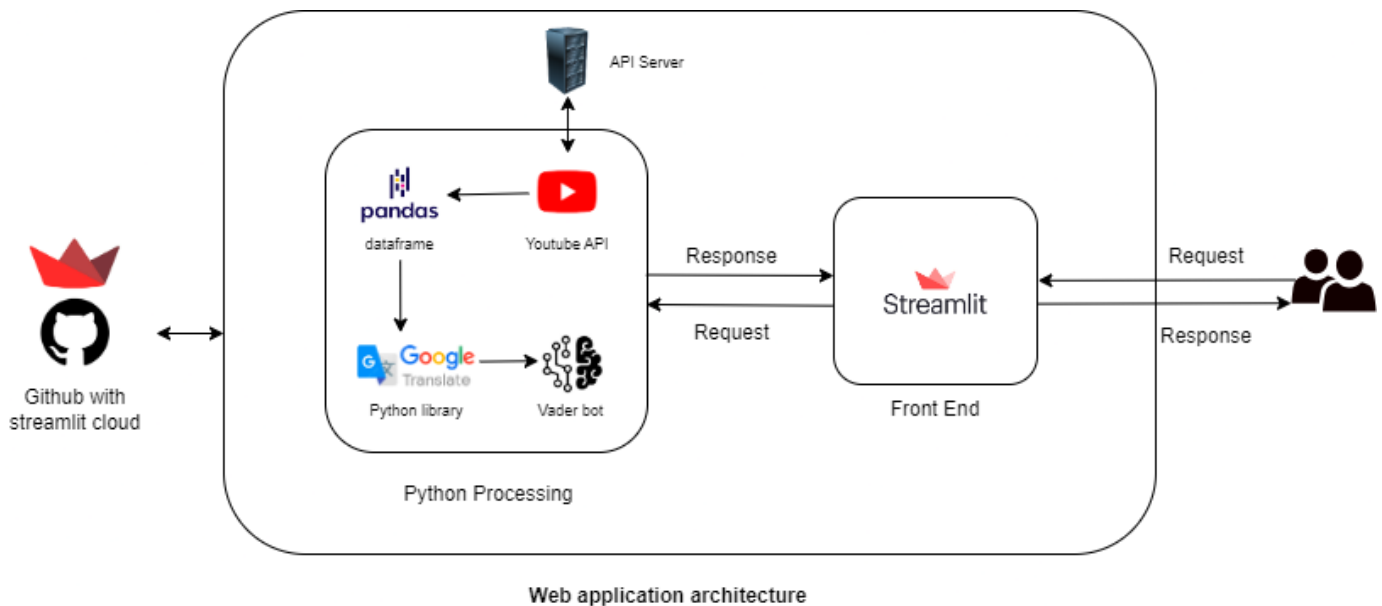
CHAPTER 3

SYSTEM DESIGN

3.1 GENERAL

In this section, we would like to show how the general outline of how all the components end up working when organized and arranged together. It is further represented in the form of a flow chart below.

3.2 SYSTEM ARCHITECTURE DIAGRAM



This is the architecture diagram of ECHO INSIGHTS, with user request and response models connected to streamlit that acts as front-end, the comments are fetched from API V3 then translated using google trans and processed using Vader bots for sentiment. Pushed to GitHub and hosted using streamlit cloud.

Fig 3.1: System Architecture

3.3 DEVELOPMENT ENVIRONMENT

3.3.1 HARDWARE REQUIREMENTS

The hardware requirements may serve as the basis for a contract for the system's implementation. It should therefore be a complete and consistent specification of the entire system. It is generally used by software engineers as the starting point for the system design.

Processor	Core i3 or more
RAM	4GB or more
Browser	Chrome 64+ or Edge 79+
OS	Any OS with supported browsers

TABLE 3.1 HARDWARE REQUIREMENTS

3.3.2 SOFTWARE REQUIREMENTS

The software requirements for "ECHO INSIGHTS" are as follows:

- 1) Any mobile or computing devices with a web browser
- 2) All the libraries installed or via req.txt (Streamlit, Vader, Google trans)

CHAPTER 4

PROJECT DESCRIPTION

4.1 METHODOLOGY

With the help of your YouTube sentiment analysis app, you may gain a deeper understanding of your audience. Through a user-friendly Streamlit interface, you may examine comments related to videos you provide by utilizing the YouTube Data API. The sentiment analysis program VADER then analyzes each comment's underlying emotions and assigns a score to classify it as positive, negative, or neutral. Even comments written in foreign languages can be easily translated for thorough examination with Google Translate.

The software does more than just provide emotion scores. It offers a basic sentiment that encapsulates the atmosphere of the comments area. Additionally, it highlights the most encouraging and unfavorable remarks, which can help you figure out what appeals to your audience and where you might make improvements. An eye-catching word cloud illustrates the most frequently used words and insights.

With an interactive dashboard that lets you easily study trends, filter comments, and dig deeper into the data, Streamlit excels once more. This user-centered strategy guarantees a seamless and enlightening encounter. After deployment and testing, the app is prepared to provide insightful information about the mood of your audience with a few clicks.

4.2 MODULE DESCRIPTION

The development process for "ECHO INSIGHTS – Youtube Sentiment analysis tool encompasses four key modules.

- 1) The main dashboard, uses streamlit, a python based web framework that ensures easy integration with py models and other components much easier. Inputs the URL from user and manipulates it with the processes and returns the result for user in intuitive dashboard.
- 2) Then initial module is “fetch_comments”. Uses googleclient’s Youtube API V3 to fetch the comments the given youtube video link. Set the max results to 100 to fetch comments. Returns the author’s name , posted time , like count and the text display and return df as dataframe.
- 3) The next module “translate_cmt”.Uses google trans library of python to translate and return the result as a dataframe. The processing speed varies based on devices though.
- 4) Next comes “vader_sentiment”. Uses vader bot , a pre trained model to alter the sentiments for the given texts and pushes a new column to dataframe. Yields a better results when compared to other models.
- 5) “Word cloud” , a simple python library that uses matplotlib lib to plot most frequently used words based on sizes for a better visual glance and return to key page.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 OUTPUT

The following contains the images of the working application

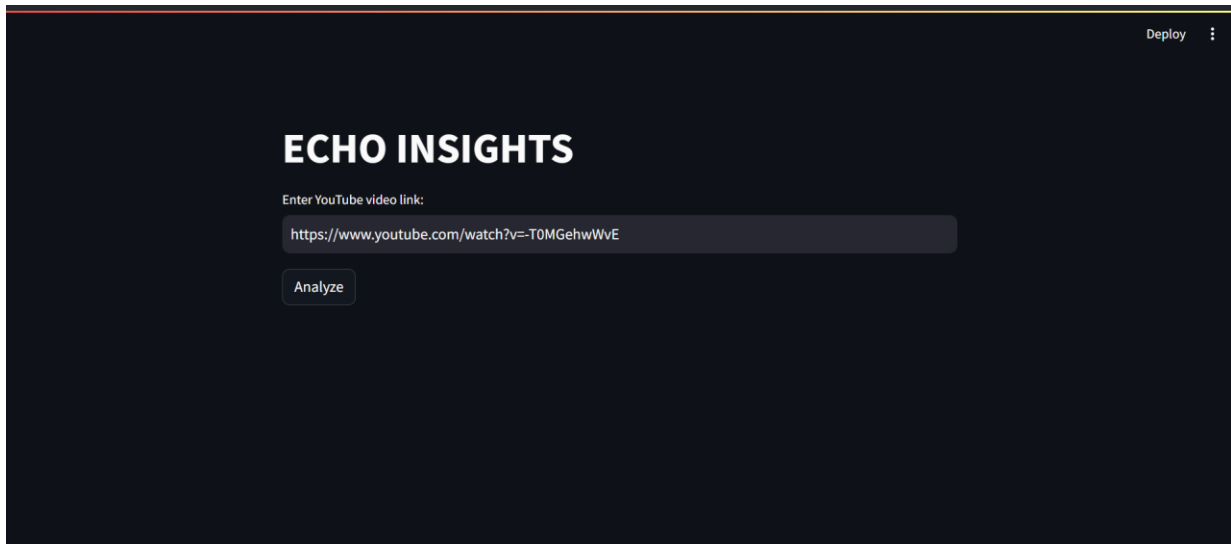


Fig 2 – Welcome page

prompts the user to input the Youtube URL , on pressing analyse button fetches youtube comments and displays them

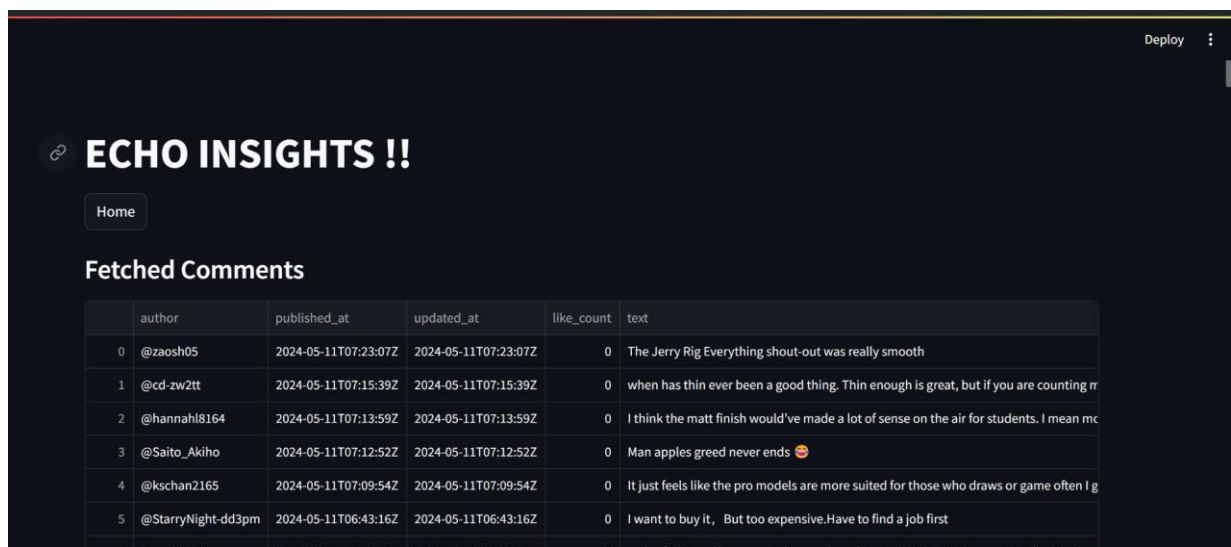


Fig 3 – Displays fetched comments

Secondary page that shows the comments with all required labels like , author , time and like counts along with the text fetched using API

Translated Comments			Deploy
	text	translated_text	
0	The Jerry Rig Everything shout-out was really smooth	The Jerry Rig Everything shout-out was really smooth	
1	when has thin ever been a good thing. Thin enough is great, but if you are counting n	when has thin ever been a good thing. Thin enough is great, but if you are counting n	
2	I think the matt finish would've made a lot of sense on the air for students. I mean mc	I think the matt finish would've made a lot of sense on the air for students. I mean mc	
3	Man apples greed never ends 🤔	Man apples greed never ends 🤔	
4	It just feels like the pro models are more suited for those who draws or game often I g	It just feels like the pro models are more suited for those who draws or game often I g	
5	I want to buy it, But too expensive.Have to find a job first	I want to buy it, But too expensive.Have to find a job first	
6	2:48	2:48	
7	Anyway, did we get a calculator?	Anyway, did we get a calculator?	
8	@MkHBD, it may not be relevant for you and it's ridiculous in terms of power. But, Yo	@MkHBD, it may not be relevant for you and it's ridiculous in terms of power. But, Yo	
9	Until Mac OS comes to iPad they can't do anything to make this thing worth it to me	Until Mac OS comes to iPad they can't do anything to make this thing worth it to me	

Fig 4 – Display translated comments

Uses google trans library to convert all the comments into English(to an extend) and add it as a column in data frame and displays it to user in table form as shown

Top 10 Comments based on Sentiment				Deploy
	author	text	vader_sentiment	
27	@fitdude	With your sarcasm and flippancy style of video you have lost my brother as a subscribe	0.9844	
35	@BrockBarr	Haha! This is MKB at peak sarcasm. Hilarious! 🤔🤔🤔	0.9199	
18	@nickranly6522	It's impressive. I love my Pixel Tablet. No complaints.	0.8676	
23	@TheWildFatKid	Keep it up Marques. I love the videos, and the content. The MKBHD Ridge Wallet look	0.8622	
66	@rcfamily4life74	I've replaced my Windows laptop with a 4th Gen iPad Pro and love it!!! Well I rec	0.8565	
14	@mr-jq2ny	the spec bumped ipads are hella useful for artists and animators. they're not made fo	0.8425	
45	@brandonlee7382	Wow, it's so thin!! I honestly didn't think that was possible	0.8247	
36	@The3pleSix	I think the ultimate question for thin devices like that is "how sturdy it is? I mean isn't	0.8156	
47	@DustinDAmour	I will be upgrading to the new iPad Pro 13 from a 2nd generation iPad Pro 12.9. The 1	0.8074	
55	@njinco	Great hardware that you will never be able to take full advantage of. Yep, that's	0.8074	

Fig 5a – Sorted comments based on sentiment – increasing

After sentimenting the results , the top 10 comments are displayed to the user in a data frame table format with text and author name based on sentiment values

Deploy

Worst 10 Comments based on Sentiment

	author	text	vader_sentiment
93	@titania7768	Apple's new plan with the ipad is to instill its users will a constant fear of literally anyt	-0.9109
39	@truestech4158	Yay, lets make yet another stupid wobbly otherwise flat sitting device that has that ca	-0.886
77	@V1rus0ne	Zack can make that iPad's bend test a live paid event and YouTube would need to ass	-0.7351
40	@bobd429	damn \$999 for the 11 inch? it cost \$1,260 here in Malaysia 🙄	-0.7003
85	@MrRapmaster19	Apple desperately needs to simplify its iPad lineup. It's a product that is far less requi	-0.6652
83	@vedantdesai1	Matte OLED = No OLED. There is a reason why the OLED gaming community is shitting	-0.6124
15	@StanleyKC	MKBHD in his villain stage. Here for it!	-0.5983
67	@user-ti7hp8ld4y	Apple has run out of ideas since Steve Jobs died	-0.5574
48	@24kiryusuf	I'm guessing Apple's gonna add some crazy stuff to iPadOS next month, sin	-0.5574
80	@liban4695	Speaking of overpriced, why are ridge wallets so damn expensive.	-0.541

Fig 5b – Sorted comments based on sentiment

After sentimenting the results , the worst 10 comments are displayed to the user in a data frame table format with text and author name based on sentiment values

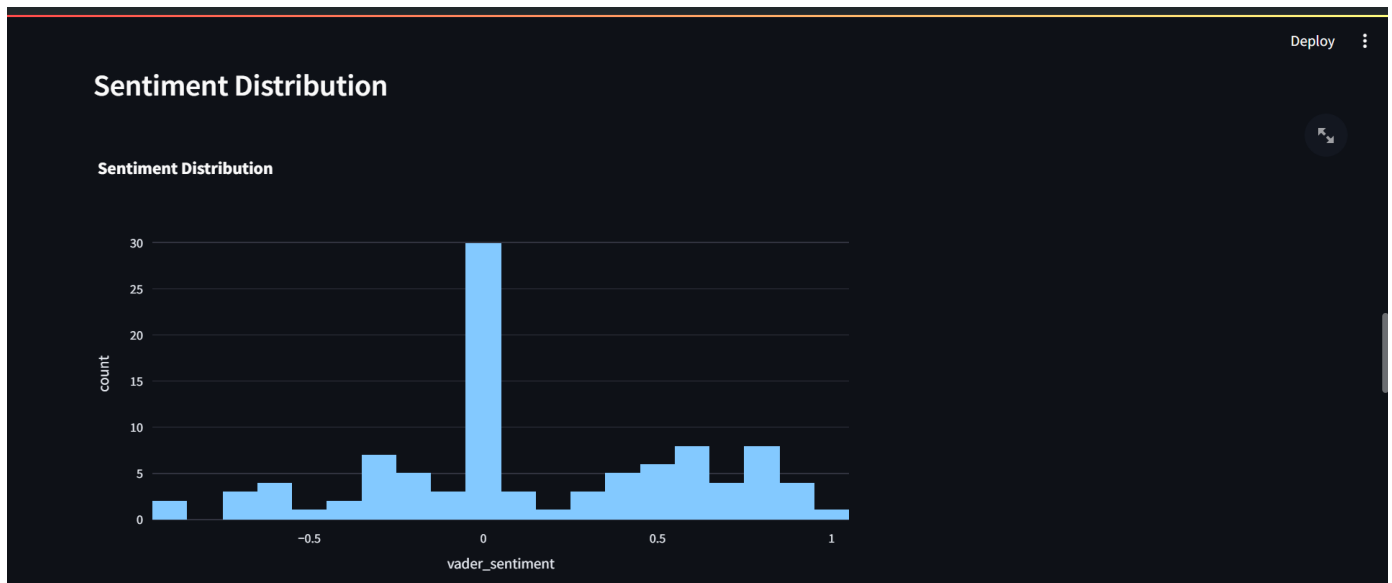


Fig 6 – Sentiment distribution graph

On processing the data, the graphs are plotted using plotly library of python . this shows the sentiment distribution over -1 to 1 with count (as of 100 results fetched)

5.2 RESULT

With our in-depth study of YouTube comments, you can unleash the power of audience insights! This analysis reveals the general sentiment—positive, negative, or neutral—by unlocking the emotional undertones of the comments area of your video. We go above and above by highlighting the most encouraging and critical remarks, serving as valuable nuggets of insight into what appeals to your audience and possible areas for development. However, the examination doesn't end there! The most often used words are graphically represented in a stunning word cloud that takes center stage. This offers an overview of the main ideas and subjects that generated the greatest conversation in the comments.

We care about the user experience more than just statistics. With Streamlit, we created an interactive dashboard that gives you control. Sort comments according to sentiment. For an analysis that is laser-focused, filter comments according to sentiment or keywords. Examine sentiment patterns over time to observe changes in audience perception. Investigate specific comments and their sentiment scores to learn more about your audience's responses.

You may adjust your material to the tastes of your audience by utilizing the power of this study. By cultivating a supportive and active community around your channel, you can eventually soar to new heights with your YouTube presence.

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

6.1 CONCLUSION

To sum up, ECHO INSIGHTS is an essential resource for researchers, marketers, and content creators that want to understand the nuances of audience emotion and interaction on YouTube. Through the utilization of cutting-edge technology such as the YouTube API and sentiment analysis algorithms, customers are enabled to obtain practical insights regarding audience interactions. Comprehensive sentiment analysis results and interactive data visualization tools are easily accessible through the dashboard's user-friendly design, which is powered by Streamlit. Users may maximize effect and refine content strategy by using ECHO INSIGHTS to uncover sentiment trends, emerging topics, and patterns within audience interactions.

6.2 FUTURE ENHANCEMENT

Multi-platform Support: To give consumers a complete picture of their digital footprint, extend ECHO INSIGHTS to evaluate sentiment and interaction patterns across a variety of social media sites, including Facebook, Instagram, and Twitter.

Real-time Data Integration: Combine real-time data streams to give users the most recent information on audience interactions. This way, content strategies may be quickly adjusted in response to conversations and trends that are starting to emerge.

Improved Collaboration Features: Put in place collaborative features that make it easier for academics, marketers, and content creators to work together and share knowledge. This will allow them to analyze data together, exchange ideas, and co-create plans.

Report generation: Introducing configurable reporting options will enable customers to adapt analysis results to their own requirements. These options include the capacity to produce comprehensive reports in a variety of formats, including interactive dashboards, CSV files, and PDFs.

Predictive analytics: Allow users to proactively modify their content strategies for optimal effect and success by integrating predictive analytics tools to foresee future audience behavior and trends based on historical data.

APPENDIX

SOURCE CODE:

Fetch.py

```
import googleapiclient.discovery
from googleapiclient.discovery import build
import pandas as pd
def fetch_comments(video_id):

    api_service_name = "youtube"
    api_version = "v3"
    DEVELOPER_KEY = "AIzaSyB47XVqTefMFpRGzMavn1KMc9W1Jf8BrCg"

    youtube = googleapiclient.discovery.build(api_service_name, api_version,
        developerKey=DEVELOPER_KEY)

    request = youtube.commentThreads().list(
        part="snippet",
        videoId=video_id,
        maxResults=100
    )
    response = request.execute()

    comments = []

    for item in response['items']:
        comment = item['snippet']['topLevelComment']['snippet']
        comments.append([
            comment['authorDisplayName'],
            comment['publishedAt'],
            comment['updatedAt'],
            comment['likeCount'],
            comment['textDisplay']
        ])

    df = pd.DataFrame(comments, columns=['author', 'published_at', 'updated_at',
        'like_count', 'text'])

    return df
```

Pdftp.py

```
# import pandas as pd
# from pandas_profiling import ProfileReport

# def generate_profile(df):
#     # Generate the profile report
#     profile = ProfileReport(dataframe, explorative=True)
#     return profile
```

```

import pandas as pd
from pandas_profiling import ProfileReport
# from pydantic import BaseModel, BaseSettings, Field, PrivateAttr

def generate_profiling_report(df):
    profile = ProfileReport(df, explorative=True)
    return profile.to_html()

```

Index.py

```

import streamlit as st
import pandas as pd
import plotly.express as px
from fetch import fetch_comments
from trans import translate_comments
from vader import analyze_sentiment
from wordcld import generate_wordcloud
from pdpf import generate_profiling_report

def input_page():
    st.title("ECHO INSIGHTS")

    # Get YouTube video link from user
    youtube_link = st.text_input("Enter YouTube video link:")
    if st.button("Analyze") and youtube_link:
        # Extract video ID from the link
        video_id = youtube_link.split("=")[-1]
        # Set query parameters to navigate to the next page
        st.session_state["query_params"] = {"youtube_link": youtube_link, "video_id":
video_id}

def result_page():
    st.title("ECHO INSIGHTS !!")

    if st.button("Home"):
        # Clear session state
        st.session_state.clear()
        # Redirect to input page
        input_page()

    # Get video ID from query parameters
    query_params = st.session_state.get("query_params", {})
    youtube_link = query_params.get("youtube_link", "")
    video_id = query_params.get("video_id", "")

```

```

# Fetch comments
df = fetch_comments(video_id)
st.subheader("Fetched Comments")
# Display selected columns' heads in Streamlit
st.write(df.head(15))

# Translate comments
translated_comments = translate_comments(df)
selected_columns = ['text', 'translated_text']
selected_df = df[selected_columns]
st.subheader("Translated Comments")
st.write(selected_df.head(15))

# Perform sentiment analysis
sentiment_analysis = analyze_sentiment(translated_comments)
selected_columns = ['author', 'text', 'translated_text', 'vader_sentiment']
selected_df = df[selected_columns]

# Combine sentiment analysis with original data
df['vader_sentiment'] = sentiment_analysis['vader_sentiment']

# Sort comments based on sentiment score
top_comments = df.sort_values(by='vader_sentiment', ascending=False).head(10)

# Display top 10 comments with sentiments
st.subheader("Top 10 Comments based on Sentiment")
st.write(top_comments[['author', 'text', 'vader_sentiment']])

worst_comments = df.sort_values(by='vader_sentiment', ascending=True).head(10)

# Display worst 10 comments with sentiments
st.subheader("Worst 10 Comments based on Sentiment")
st.write(worst_comments[['author', 'text', 'vader_sentiment']])

# Display sentiment distribution plot
st.subheader("Sentiment Distribution")
fig = px.histogram(df, x='vader_sentiment', nbins=30, title="Sentiment
Distribution")

st.plotly_chart(fig)

# Display overall sentiment
overall_sentiment = sentiment_analysis['vader_sentiment'].mean()
st.subheader("Overall Sentiment")
st.write(f"Overall sentiment: {overall_sentiment}")

st.subheader("Word Cloud")
generate_wordcloud(translated_comments['text'])

# st.subheader("Pandas Profiling Report")
# profiling_report = generate_profiling_report(df)
# st.write(profiling_report)

```

```

# Comments over time
df['comment_published'] = pd.to_datetime(df['published_at'])
df['comment_date'] = df['comment_published'].dt.date
comments_over_time =
df.groupby('comment_date').size().reset_index(name='comment_count')
st.subheader("Comments Over Time")
fig_comments_over_time = px.line(comments_over_time, x='comment_date',
y='comment_count', title='Comments Over Time')
st.plotly_chart(fig_comments_over_time)

def main():
    # Render the appropriate page based on query parameters
    if st.session_state.get("query_params"):
        result_page()
    else:
        input_page()

if __name__ == "__main__":
    main()

```

Translate.py

```

def translate_comments(df):
    # Your code to translate comments here
    from googletrans import Translator
    def translate_text(text, src='auto', dest='en'):
        translator = Translator()

        translated_text = translator.translate(text, src=src, dest=dest)
        return translated_text.text

    # Apply translation to the 'text' column
    df['translated_text'] = df['text'].apply(lambda x: translate_text(x))
    return df

```

Vader.py

```

def analyze_sentiment(df):
    from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

    # Function to perform sentiment analysis using VADER
    def get_vader_sentiment(text):
        analyzer = SentimentIntensityAnalyzer()
        sentiment_score = analyzer.polarity_scores(text)

```



```
# We extract the compound score which represents the overall sentiment
return sentiment_score['compound']

# Apply sentiment analysis to the 'translated_text' column
df['vader_sentiment'] = df['translated_text'].apply(get_vader_sentiment)
return df
```

Wordcloud.py

```
import streamlit as st
from wordcloud import WordCloud
import matplotlib.pyplot as plt

def generate_wordcloud(df_column):
    comment_text = ' '.join(df_column)
    wordcloud = WordCloud(width=800, height=400,
background_color='white').generate(comment_text)
    fig, ax = plt.subplots(figsize=(10, 5))

    ax.imshow(wordcloud, interpolation='bilinear')
    ax.axis('off')
    st.pyplot(fig)
```

REFERENCES

- 1) Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel et al. "Scikit-learn: Machine learning in Python." *the Journal of machine Learning research* 12 (2011): 2825-2830.
- 2) Singh, Ritika, and Ayushka Tiwari. "Youtube comments sentiment analysis." *Int. J. Sci. Res. Eng. Manag.(IJSREM, no. May, p. 5, 2021,[Online]. Available: <https://www.researchgate.net/publication/351351202>* (2021).
- 3) Elbagir, Shihab, and Jing Yang. "Sentiment analysis on Twitter with Python's natural language toolkit and VADER sentiment analyzer." In *IAENG Transactions on Engineering Sciences: Special Issue for the International Association of Engineers Conferences 2019*, pp. 63-80. 2020.
- 4) Rana, Manish, and Mohammad Atique. "LANGUAGE TRANSLATION: ENHANCING BI-LINGUAL MACHINE TRANSLATION APPROACH USING PYTHON." *I-Manager's Journal on Computer Science* 7, no. 2 (2019).
- 5) Pradhan, Rahul. "Extracting Sentiments from YouTube Comments." In *2021 Sixth International Conference on Image Information Processing (ICIIP)*, vol. 6, pp. 1-4. IEEE, 2021.
- 6) Sahoo, Kabita, Abhaya Kumar Samal, Jitendra Pramanik, and Subhendu Kumar Pani. "Exploratory data analysis using Python." *International Journal of Innovative Technology and Exploring Engineering* 8, no. 12 (2019): 4727-4735.
- 7) Khorasani, Mohammad, Mohamed Abdou, and J. Hernández Fernández. "Web Application Development with Streamlit." *Software Development* (2022): 498-507.
- 8) McKinney, Wes. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.
- 9) Bernard, Joey, and Joey Bernard. "Python data analysis with pandas." *Python Recipes Handbook: A Problem-Solution Approach* (2016): 37-48.
- 10) Kabir, Ahmed Imran, Ridoan Karim, Shah Newaz, and Muhammad Istiaque Hossain. "The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R." *Informatica Economica* 22, no. 1 (2018).
- 11) Rathee, Nisha, Nikita Joshi, and Jaspreet Kaur. "Sentiment analysis using machine learning techniques on Python." In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 779-785. IEEE, 2018.
- 12) Alhujaili, Rawan Fahad, and Wael MS Yafooz. "Sentiment analysis for youtube videos with user comments." In *2021 International conference on artificial intelligence and smart systems (ICAIS)*, pp. 814-820. IEEE, 2021.
- 13) Kolhar, Manjur, and Abdalla Alameen. "Artificial Intelligence Based Language Translation Platform." *Intelligent Automation & Soft Computing* 28, no. 1 (2021).
- 14) Dashtipour, Kia, Soujanya Poria, Amir Hussain, Erik Cambria, Ahmad YA Hawalah, Alexander Gelbukh, and Qiang Zhou. "Multilingual sentiment analysis: state of the art and independent comparison of techniques." *Cognitive computation* 8 (2016): 757-771.
- 15) Saif, Hassan, Yulan He, and Harith Alani. "Semantic sentiment analysis of twitter." In *The Semantic Web—ISWC 2012: 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I* 11, pp. 508-524. Springer Berlin Heidelberg, 2012.