Trends in Information Management (TRIM)
ISSN: 0973-4163
13(1), pp. 45-60

Sentiment Analysis of Users' Comments on Budget Speeches by the Indian Finance Minister (2019–2023): A Data Mining Study based on YouTube Videos

Nitesh Kumar Verma
Swagota Saikia
Manoj Kumar Verma

Abstract

This study examines the sentiment analysis of public responses to the Indian Finance Minister's budget speeches from 2019 to 2023, leveraging Python programming and NLP techniques. Objectives include evaluating video characteristics such as duration, resolution, view count, and likes; ranking videos based on engagement metrics; analyzing sentiments in speeches and viewer comments; and exploring trends in comment growth and gender-wise variation. Findings reveal that the Union Budget 2023-24 video garnered significant attention and positive reception, demonstrating the importance of engaging content despite longer durations. Viewer comments for this budget highlight substantial engagement, reflecting its relevance. Sentiment analysis uncovers frequent mentions of government policies, taxation, and economic figures, indicating public interest in financial and policy-related aspects. Overall, the sentiment expressed in both videos and comments is predominantly positive, with moderate subjectivity, showcasing diverse perspectives.

Keywords: *YouTube speeches, Finance Minister, sentiment analysis, data mining, comment analysis, budget speeches, communication strategies.*

Introduction

Budget speeches delivered by a nation's finance minister hold significant importance as they outline expenditure plans and estimates for the country's economic growth and development in the fiscal year. These speeches not only reflect government priorities but also serve as critical tools for driving economic growth and fostering social cohesion (Ferry et al., 2014; Miller & Rose, 1990). Traditionally, budget speeches are presented in parliament. However, with the rapid proliferation of digital platforms, they are now also accessible in real-time on platforms like YouTube. YouTube, ranked as the second most popular social networking site globally, boasts 2.514 billion users (Statista, 2023). The platform has emerged as a prominent medium for sharing videos, including speeches by influential figures on a diverse range of topics, transforming how individuals access and engage with information. Through features such as likes, shares, and comments, YouTube facilitates active user engagement, allowing viewers to express opinions and interact with content. Metrics such as likes, views, shares, and comments help evaluate the audience's reception and engagement with creators' content (Saikia et al., 2023). Budget speech videos available on YouTube present an opportunity to analyse public perceptions and opinions through viewer comments. Sentiment analysis of these comments provides valuable insights into the public's perception, sentiments, and engagement regarding the finance minister's budgetary decisions and policies. This study aims to analyse YouTube videos of Finance Minister Nirmala Sitharaman's budget speeches from 2019 to 2023 to uncover public sentiment surrounding these speeches.

Sentiment analysis, or opinion mining, is a subfield of Natural Language Processing (NLP) that extracts subjective information and sentiments—positive, negative, or neutral—from textual data (Pang & Lee, 2008). Originating in the early 2000s, the field initially employed rule-based approaches to analyse sentiment in online reviews, social media posts, and customer feedback (Liu, 2012; Turney, 2002). Its applications span various domains, including market research, brand reputation management, social media monitoring, customer feedback analysis, and political analysis (Go et al., 2009; Liu, 2012). Sentiment analysis enables organisations to understand public opinion, track customer satisfaction, and make data-driven decisions. One prominent example of government communication on YouTube is the "PM Mann Ki Baat" series, where India's Prime Minister addresses the nation on a range of national issues. These videos bridge the gap between government and citizens, reflecting public sentiments, expectations, and responses to policies and initiatives.

Several studies have explored sentiment analysis on YouTube to understand public opinion. Smith et al. (2022) examined the impact of sentiment on viewer engagement and political debate by analysing political speeches on YouTube. Similarly, Johnson and Gupta (2021) studied public reactions to government policies, offering insights into the dynamics of online discourse. Bringula et al. (2023) analysed comments on political vlogs and news sources related to two Philippine presidential candidates, finding variations in public sentiment and biases based on content type. Alhujaili and Yafooz (2021) evaluated Arabic educational videos using machine learning and deep learning models, achieving high accuracy in sentiment classification. Deori et al. (2021) analysed viewer comments on Hindi news channel videos, uncovering trends in gender-based viewership and sentiment. Porreca et al. (2020) studied sentiments toward vaccination-related videos in Italy, highlighting shifts in public opinion following vaccination campaigns and the role of social media in shaping narratives.

Despite the growing body of research, no prior studies have specifically analysed YouTube videos of budget speeches delivered by the Indian Finance Minister from 2019 to 2023. This study addresses this gap by examining the sentiment landscape of viewer comments, identifying prominent keywords and bigrams, and interpreting the findings. By employing Python programming and NLP techniques to process and analyse extensive textual data, this research provides valuable insights into public reactions and sentiments toward government financial plans and policies.

Objectives

- To evaluate the characteristics such as duration, resolution, view count, and likes of the selected videos on YouTube.
- To rank the selected videos based on their view count, like count, and comment count.
- To conduct sentiment analysis on the Finance Minister Budget speeches from the years 2019-2023 YouTube videos and gain insights into the overall sentiment expressed by the speakers.
- To conduct sentiment and term analysis on the Finance Minister Budget speeches from the years 2019-2023 YouTube video comments and gain insights into the sentiment expressed by viewers.
- To evaluate the Gender-wise distribution of comments on the videos.

Data and Methodology

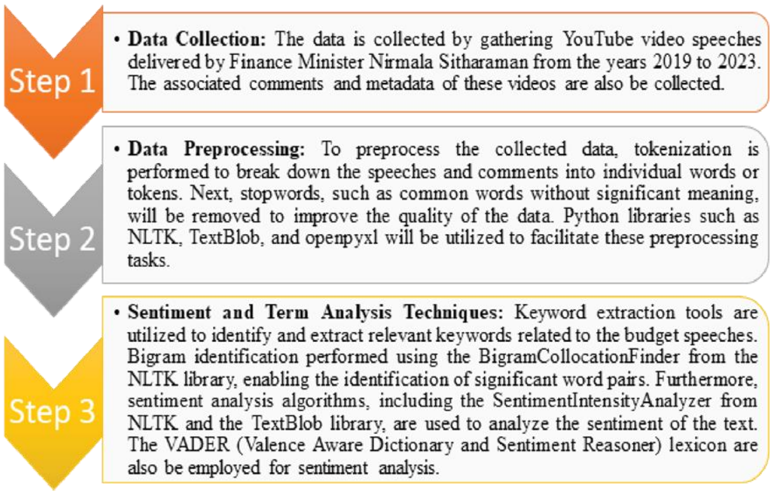
The present study employs a quantitative research design with a cross-sectional approach. This design entails collecting data at a specific time, focusing on Finance Minister Nirmala Sitharaman's budget speeches from 2019-2023 and the associated comments on YouTube. The study employs Python-based text preprocessing techniques such as tokenisation, stopword removal, and lemmatisation, using Python libraries like NLTK, TextBlob, and openpyxl. The dataset used consists of YouTube video speech data, comments, and related metadata. The study aims to provide meaningful insights into the themes, sentiment, and audience engagement of YouTube video speeches and comments. However, it is important to note that the study is limited to the specific dataset and the employed NLP techniques and does not include advanced machine learning algorithms or predictive modeling for future trends.

The video details and associated comments were extracted using the software Webometric Analyst (http://lexiurl.wlv.ac.uk/) on October 2023. Subsequently, English language comments (n=676) were selected for sentiment analysis of YouTube comments. The analysed dataset was then presented through various charts, tables, and graphs using tools such as Tableau, Google Spreadsheet, and MS Excel, facilitating meaningful interpretation and interaction.

The mode of data collection is through the retrieval of YouTube video speeches via YouTube API. The associated comments and metadata are also collected by data mining techniques, including sentiment analysis algorithms, keyword extraction, and bigram identification, to the collected data from these videos.

The data are accessed and extracted from the YouTube platform by using the following steps:

Fig 1: Data Extraction process from YouTube



The following techniques are also used:

a) *Converting YouTube video speech into Texts:* Python code is prepared to convert the YouTube video speech into texts. The code uses Python and several libraries like Natural Language Toolkit (NLTK), moviepy, Google's speech recognition API, etc., to extract audio from a YouTube video and write the transcriptions to a .TXT file. The "Webometric Analyst" tool is used to obtain the YouTube Videos comments.

b) *Analyse the Speech text in Python:* The Python code is prepared to assess Natural Language Processing (NLP) tasks on a text document created from the YouTube Video speech and the respective comments. Here is a breakdown of the code and each step:

   i. *Libraries imported:* The first few lines of code import necessary libraries for NLP tasks, such as NLTK, TextBlob, and openpyxl. These libraries provide functionality for tasks such as tokenising text, calculating word frequencies, performing sentiment analysis, and creating Excel files.

   ii. *Resource downloading:* Downloaded additional resources required by NLTK, such as stopwords, a part-of-speech tagger, and the VADER lexicon for sentiment analysis.

   iii. *Text loading:* The code loads the text file to be analysed from the text file.

   iv. *Text preprocessing*: The code tokenises the text, removes stopwords and punctuation, and lemmatises the words. Lemmatisation converts words to their base form (e.g., "running" becomes "run").

c) *Keyword extraction:* The code performs part-of-speech tagging on the tokens and extracts only the nouns and adjectives. It then calculates the frequency distribution of the words to identify the most common keywords in the text.

d) *Bigram extraction:* The code uses the BigramCollocationFinder from NLTK to extract the 30 most common bigrams (pairs of words that appear together frequently) in the text.

e) *Sentiment analysis:* The code uses the SentimentIntensityAnalyzer from NLTK and the TextBlob library to perform sentiment analysis on the text. The VADER (Valence Aware Dictionary and Sentiment Reasoner) tool is also used. It calculates sentiment polarity and subjectivity.

Data Analysis and Interpretation

Duration of the videos

Based on Table 1, the video of Finance Minister Nirmala Sitharaman's Budget Speech for the Union Budget 2023-24 has the highest video length, with a duration of 01:29:10. The video with the shortest duration is the Budget Speech by FM Nirmala Sitharaman for the Union Budget 2020-21, with a duration of 02:41:21. The video for the Union Budget 2023-24 has the highest number of likes, with 4.4k likes, followed by the Union Budget 2019-20 with 1.5k likes. The video with the lowest number of likes is the Budget Speech for the Union Budget 2020-21, with 763 likes. The video for the Union Budget 2023-24 has the highest number of likes, whereas the video for the Union Budget 2020-21, which has the longest duration, has the lowest number of likes. Max Resolution of Finance Minister Nirmala Sitharaman's Budget Speech Videos by Year.

Table 1: Video Length and total likes of on the Videos

S.N.	Title of Video	Video length (HH:MM: SS)	Total Likes
1	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2023-24	01:29:10	4.4k
2	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2022-23	01:34:53	2k
3	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2021-22	01:51:30	2.6k
4	Budget Speech by FM Nirmala Sitharaman | Union Budget 2020 – 21	02:41:21	763
5	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2019-20	02:10:12	1.5k

Resolution of the videos

Based on Table 2, all the videos of Finance Minister Nirmala Sitharaman's Budget Speech have a maximum resolution of either 1080p or 720p. The recent video of the Union Budget 2023-2024 has a maximum resolution of 1080p, while the first video of the Union Budget 2019-2020 has a maximum resolution of 720p. The quality and resolution of the videos have improved over time, with the newest video having the highest resolution. This shows that the makers of these videos are bestowing more resources for creating better videos having good resolution resulting in more viewers.

Table 2: Maximum Resolution of Finance Minister Nirmala Sitharaman's Budget Speech Videos by Year

S.N.	Title of Video	Max Resolution
1	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2023-24	1080p
2	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2022-23	1080p
3	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2021-22	720p
4	Budget Speech by FM Nirmala Sitharaman | Union Budget 2020 – 21	720p
5	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2019-20	720p

Number of views received

Based on Table 3, the video of Finance Minister Nirmala Sitharaman's Budget Speech for the Union Budget 2023-2024 has the uppermost number of 208,786 views, followed by the Union Budget 2019-2020 with 120,647 views. The video with the lowest number of views is for the Union Budget 2020-21, with 33,946 views. It is also noteworthy that all the videos were uploaded on YouTube on the same day of the month (i.e. 1st) and the same time of the year (between February and July). The videos for the Union Budget 2021-2022 and 2022-2023 have a similar number of views, with a difference of only around 37,000 views, whereas the video for the Union Budget 2023-2024 has significantly more views than the other videos of the Union Budget.

Number of comments received on the videos

The total number of comments on the videos is shown in Table 4. The video titled "Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2023-24" has received the highest number of comments i.e. 209. This specifies a substantial level of engagement and attention from the viewers concerning the budget speech for the specified year. The video titled "Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2021-22" has received 184 comments, portentous an important level of viewer interest in the budget speech for a particular year. The mean value of 136 comments advocates an average level of viewer interest across these videos, while the median value of 141 comments specifies that half of the videos have obtained mostly this number of comments or less. The range of 154 comments demonstrates the difference in the total number of comments, with the highest and lowest values differing by 154. This range highlights the variation in viewer engagement across the videos.

Table 3: Total Number of Views for Finance Minister Nirmala Sitharaman's Budget Speech Videos from 2019-20 to 2023-24

S.N.	Titles of the Videos	Published on	Total Views
1	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2023-24	01-02-2023	208,786
2	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2022-23	01-02-2022	81,648
3	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2021-22	01-02-2021	118,909
4	Budget Speech by FM Nirmala Sitharaman | Union Budget 2020 – 21	01-02-2020	33,946
5	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2019-20	05-07-2019	120,647

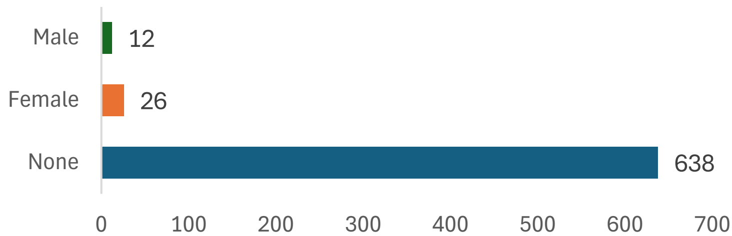
Table 4: Analysis of Total number of comments on the Videos

S.N.	Titles of the Videos	Total number of Comments
1	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2023-24	209
2	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2022-23	93
3	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2021-22	184
4	Budget Speech by FM Nirmala Sitharaman | Union Budget 2020 – 21	55
5	Finance Minister Nirmala Sitharaman's Budget Speech | Union Budget 2019-20	141

From Fig. 2, it can be observed that most comments on YouTube related to the topic are from the "None" category, accounting for approximately 94.37% of the total comments. This indicates that the gender of the commenters is unidentified or maybe the commenters are not showing it for whatsoever reasons. On the other hand, the proportion of comments from male viewers is relatively low, representing only 1.78% of the total comments. Similarly, the proportion of comments from female viewers is also relatively low, accounting for approximately 3.85% of the total comments. These proportions suggest that most comments on selected

YouTube videos are from unidentified or unspecified gender commenters, while comments specifically from male and female viewers are comparatively fewer.

Fig. 2: Distribution of comments by Gender



Top Terms and Sentiment Analysis of YouTube Video Speeches

To find out the top terms and conduct sentiment analysis on YouTube videos, a Python code has been prepared to convert the YouTube video speech into textual format. The analysis of top terms in the YouTube video speech dataset offers a quantitative overview of the most frequently mentioned terms. By scrutinising term frequencies, an understanding resonates with the viewers shown in Table 5.

Table 5: Top Terms and Frequencies in YouTube Video Speech

Rank	Keywords	Frequency	Rank	Keywords	Frequency
1	government	285	5	India	220
2	year	269	6	sector	152
3	tax	260	7	income	147
4	crore	225	8	scheme	139

The term that occurs most frequently in the dataset is "government," with a frequency of 285. This indicates that discussions surrounding various aspects of the government are prevalent in the YouTube videos. Following closely is the term "year" with a frequency of 269, suggesting that the videos often address events, developments, or trends that occur within specific timeframes. The keyword "tax" ranks third with a frequency of 260 times, emphasising the connotation of discussions about taxation matters. Moreover, the other keywords such as "crore" (225 times), "India" (220 times), and "sector" (152 times) feature highly, signifying that the issues such as monetary terms, the country of India, and the industry segments are recurrently chatted.

Other keywords that appear prominently in the top ten include "income" (147 times), "scheme" (139 times), "state" (131 times), and "infrastructure" (130 times). These keywords deliver valuable perceptions into the various aspects roofed in the selected YouTube videos, encircling discussions on income-related matters, various government schemes, state-level matters, and infrastructure growth. The frequencies associated with each term reflect the relative prominence and importance of these topics within the dataset.

Fig. 3: Keyword Analysis of the speech

Analysis of Top Ten Bigrams on Budget Speech

The analysis of these bigrams provides valuable insights into the key phrases and combinations of words frequently occurring in the videos. The most frequently occurring bigram, "income tax," with a frequency of 64, indicates a strong emphasis on taxation matters. This suggests that the videos extensively cover topics such as tax policies, implications, and potential reforms. Additionally, the recurring bigrams "prime minister," "custom duty," and "lakh crore," each with a frequency of 46, point towards discussions centred around the country's leadership, government policies, and financial aspects, particularly related to customs duties and significant financial amounts.

Furthermore, the presence of the bigram "speaker sir" with a frequency of 37 suggests that the videos incorporate parliamentary or official proceedings. Additionally, the bigram "fiscal deficit," occurring 23 times, indicates discussions surrounding fiscal policies, budgetary concerns, and financial management. Moreover, the analysis highlights the bigrams "affordable housing," "pradhan mantri," and "cooperative society," each with frequencies of 16, 14, and 14, respectively, indicating that the videos touch upon topics related to housing affordability, government initiatives led by the Prime Minister, and cooperative societies. It is likely that these discussions revolve around policies, schemes, and developments in these specific areas. Lastly, the presence of the bigram "electric vehicle" with a frequency of 13 suggests that the videos engage in discussions about electric vehicles, including their adoption, infrastructure requirements, and associated policies.

Table 6: Top Ten Bigrams in YouTube Video Speech

Rank	Bigram terms	Frequency
1	income tax	64
2	prime minister	46
3	custom duty	46
4	lakh crore	46
5	speaker sir	37
6	fiscal deficit	23
7	affordable housing	16
8	pradhan mantri	14
9	cooperative society	14
10	electric vehicle	13

Sentiment Analysis of the Videos

Based on the sentiment analysis conducted on YouTube video speech dataset, the emotional and subjective aspects of the videos emerge. Fig. 4 displays the multiple sentiments of the viewers, ranging from negative, neutral, positive, subjectivity, and polarity in the videos. The analysis reveals that the majority of sentiment falls into the neutral category, accounting for approximately 82.7% of the sentiments. This suggests that the videos primarily aim to present information objectively and without significant bias. Moreover, the analysis indicates the presence of positive sentiment, accounting for around 14.7% of the sentiments expressed. Inversely, the negative sentiments appear relatively less (only 2.60%) of the total sentiments.

The subjective analysis reveals a modest level with a subjectivity of 0.376. The subjectivity value reveals that the videos contain individual comments and opinions. The polarity value of 0.097 resembles a marginally positive polarity. The overall sentiment analysis of the videos reveals a commonly positive sentiment scenery with a trivial slanting near to positivity.

Analysis of the Top Ten keywords in the video's comments

The most commonly mentioned keyword in the video's comments is "budget," with a frequency of 132 times, this represents those discussions in the video's comments are connected to monetary and fiscal matters. The keyword "hindi" ranks second with a frequency of 74, showing an interest in the Indian language penchants and it also echoes the availability of content in Hindi and the preference for consuming videos in the Hindi language. "India" keyword positions third with a frequency of 63, signifying a sturdy attention on the matters specific to the country. The keywords "people" and the "government" having frequencies of 51 and 47, respectively, signifying a substantial curiosity in the societal and party-political issues and it also reflects the debates on the influence of government strategies on the common people's lives. The keyword "modi" appears 37 times, representing the discussions related to the Prime Minister Shri Narendra Modi and it advocates that the addressees are vigorously follow and discuss about the Prime Minister. The frequencies of the other keywords like "english", "speaking", "minister" and "language" range from 34 to 31 and these keywords specify the discussions adjacent to language preferences, speeches given by government ministers, and the importance of diverse languages.

Fig. 4: Sentiment Analysis of YouTube videos

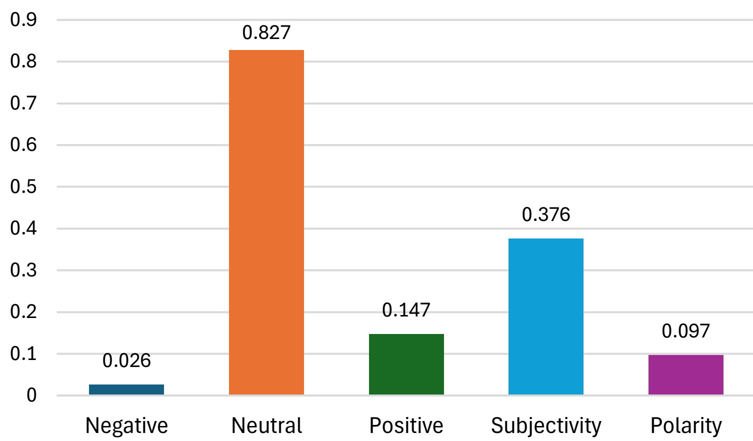


Table 7: Top Ten terms and frequencies in YouTube video comments

Rank	Keywords	Frequency	Rank	Keywords	Frequency
1	budget	132	6	Modi	37
2	hindi	74	7	English	34
3	India	63	8	speech	34
4	people	51	9	minister	32
5	government	47	10	language	31

Top Ten Bigrams in the Comments

The most common bigram term is "finance minister" that shows considerable attention on the topics pertaining to the finance minister and their policies. It also suggests that the audience are vigorously in discussions and communicating their sentiments about the fiscal matters. The bigrams terms are "modi parimenstr" and "pm modi" having a frequency of 21, representative the sturdy presence of discussions associated with the Prime Minister of India. The bigram term "maham dhanyawad" appears 11 times, which resembles the expression of thankfulness and gratefulness. The bigrams "middle class", "nirmala sitaraman" and "jai hind" have frequencies of 9, representative that the expressions associated to the middle-class, Finance Minister Nirmala Sitaraman, and nationalistic sentiments. These bigrams reflect the diverse range of terms and sentiments expressed by commenters in the YouTube video comments. Furthermore, the bigrams "parimenstr rashtriyapati", "best person" and "nath kovind" have frequencies of 9 and 8 respectively. These bigrams indicate discussions related to the President of India, expressions of admiration for certain individuals, and mentions of Hon'ble President of India i.e., Shri Ram Nath Kovind.

Table 8: Top Ten Bigrams in YouTube video comments

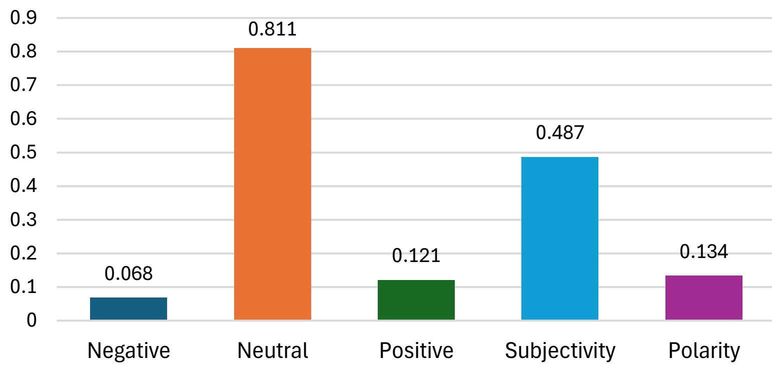
Rank	Bigram term	Frequency
1	finance minister	22
2	modi parimenstr	21
3	pm modi	21
4	maham dhanyawad	11
5	middle class	9
6	nirmala sitaraman	9
7	jai hind	9
8	parimenstr rashtriyapati	9
9	best person	9
10	nath kovind	8

Sentiment Analysis of the comments

Fig. 5 displays the multiple sentiments of the viewers ranging from negative, neutral, positive, subjectivity and polarity of the comments. The analysis reveals that the majority of comments have a neutral sentiment, accounting for 81.1% of the total. This suggests that commenters are expressing their thoughts and opinions in a balanced and objective manner, without a strong inclination towards positive or negative sentiments. Negative sentiment is relatively low, with a score of 0.068, indicating that commenters are not predominantly expressing negative emotions or criticism. On the other hand, positive sentiment has a score of 0.121, indicates that commenters are actively engaging in discussions with a constructive and positive approach.

The analysis also highlights a moderate level of subjectivity, with a score of 0.487. This suggests that commenters are expressing their personal opinions, perspectives, and experiences in their comments, contributing to a diverse range of viewpoints within the dataset. Furthermore, the overall polarity score is 0.134, leaning slightly towards positive sentiment. This indicates that the comments, as a whole, exhibit a slightly positive tone or inclination.

Fig. 5: ***Sentiment Analysis of YouTube videos comments***



Results and Discussions

Various studies were conducted on YouTube Comments sentiment analysis. But in the present study, Python programming and NLP techniques were used to convert YouTube videos into text for sentiment analysis along with their respective video comments. The findings reveal that the video for the Union Budget 2023-24, with the longest duration and highest number of likes, garnered substantial attention and positive reception from viewers. This suggests that longer videos can still attract viewer engagement if the content is compelling and relevant. Additionally, the improved video quality over time indicates a conscious effort by the creators to enhance the viewing experience.

The videos of Union Budget 2023–2024 received the most views, but the viewers' count speckled among the budget talks and it demonstrates that how crucial the subject and the Finance Minister's speech were in piquing the public's attention. It's interesting to mention that the majority of comments fall under the "None" gender category that shows either the viewers did not declare their gender or it was not obvious and also highlight the understanding of the YouTube audience's demographics and viewers tastes. The evaluation of the bigrams and keywords that used the most frequently termed focus on the subjects covered in the videos. The frequency of keywords referring to the government, taxes, and certain fiscal data suggests emphasis on these topics and it also indicates that viewers have keen interest in knowing government policies, financial repercussions, and the state of the economy of India. The sentiment analysis shows that the videos and comments have an overall positive sentiment expressed which also express that the videos and comments seeks to convey the facts in an unbiased manner. A modest level of subjectivity within the videos and comments indicates that personal thoughts and perspectives are also conveyed. Furthermore, the selected research approach offers a strong framework for exploring and knowing the sentiments of the budget speeches, offering insightful contributions to the study of political communication and sentiment analysis.

Conclusion

Finally, the sentiment analysis of YouTube videos and comments relating to Finance Minister Nirmala Sitharaman's budget speeches deliver insightful evidence on user engagement, video attributes, common subjects, and sentiment expressed. This study and findings will help future communication policies and the knowledge of how budget speeches are received and perceived by politicians, academics, and the general public. Longitudinal analysis to track the changes in viewer engagement and sentiment over an extended period of time may be conducted for long-term effects of budget speeches. To gain a deeper understanding of viewer motivations and expectations, qualitative research methods such as interviews or focus groups could be employed alongside quantitative analysis and audience segmentation analysis may also be conducted to understand specific needs and preferences of different viewer segments.

References

Alhujaili, R. F., & Yafooz, W. M. S. (2021). Sentiment analysis for YouTube videos with user comments: Review. In *Proceedings of IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI)* (pp. 238–244). IEEE.Changchun, China. DOI: 10.1109/icais50930.2021.9396049

Bringula, R. P., Tabo, R. M. R., Alcazar, F. T. S. L., Delica, J. M. I., & Sayson, J. E. A. (2023). YouTube videos on the achievements of presidential candidates: Sentiment and content analysis. *Journal of Political Marketing,* 1-17. DOI: 10.1080/15377857.2023.2202617

Deori, M., Kumar, V., & Verma, M. K. (2021). Analysis of YouTube video contents on Koha and DSpace, and sentiment analysis of viewers' comments. *Library Hi Tech, 41*(3), 711–728. DOI: 10.1108/lht-12-2020-0323

Ferry, L., Zakaria, Z., & Eckersley, P. (2014). The role of budget speech: A Malaysian government study. *International Journal of Public Sector Management, 27*(7), 1–42. DOI: 10.1108/IJPSM-01-2014-0010

Go, A., Huang, L., Bhayani, R., & Cheng, X. (2009). *Twitter sentiment classification using distant supervision* (CS224N Project Report, Vol.1, No.12). Stanford University. https://wwwcs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

Johnson, E., & Gupta, S. (2021). Analysing sentiment in user comments on YouTube videos related to government policies. *Journal of Information Science, 35*(4), 567–582.

Liu, B. (2012). *Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5*(1), 1–167. DOI: 10.2200/S00416ED1V01Y20120 4HLT016

Miller, P., & Rose, N. (1990). Governing economic life. *Economy and Society, 19*(1), 1–31. DOI: 10.1080/03085149000000001

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval, 2*(1–2), 1–135. http://dx.doi.org/10.1561/1500000 011

Porreca, A., Scozzari, F., & Di Nicola, M. (2020). Using text mining and sentiment analysis to analyse YouTube Italian videos concerning vaccination. *BMC Public Health, 20*(259), 1–9. DOI: 10.1186/s12889-020-8342-4

Saikia, S., Kumar, V., & Verma, M. K. (2023). Analyzing user sentiments toward selected content management software: a sentiment analysis of viewer's comments on YouTube. *Information Discovery and Delivery, 52*(4), 378-387. DOI: 10.1108/idd-01-2023-0009

Smith, A., Johnson, B., & Thompson, C. (2022). Sentiment analysis of political speeches on YouTube. *Journal of Communication and Politics, 45*(2), 78–95.

Statista. (2023). Most popular social networks worldwide as of January 2023. *Statistics*. Retrieved October 26, 2023, from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/2023

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)* (pp. 417–424).Association for Computational Linguistics. Philadelphia, PA, USA. DOI: 10.3115/1073083.1073135

Corresponding Author

Manoj Kumar Verma be contacted at: manojdlis@mzu.edu.in

Author Biographies

Dr Nitesh Kumar Verma is working as an Assistant Professor in the Department of Library and Information Science at Babasaheb Bhimrao Ambedkar University, Lucknow. He previously worked as an Assistant Librarian at BBAU and has 10 years of professional and research experience.

Along with academics, he has attended various national and international conferences and seminars and has 30+ publications in various National & International Journals, Conference Proceedings, and Book Chapters. His research interests include Open Access, Sentiment Analysis, Scientometrics, Bibliometrics, and Digital Literacy etc.

Swagota Saikia is a research scholar in the Department of Library and Information Science at Mizoram University, Aizawl, India. She completed her Master of Library and Information Science (MLISc) from Assam. Her research interests include Open Access, Scientometrics, Bibliometrics, Sentiment Analysis, and Data Carpentry. She has actively participated in various national and international seminars and conferences.

Dr. Manoj Kumar Verma is a Professor in the Department of Library and Information Science at Mizoram University, Aizawl, India. He obtained a PhD in Library and Information Science from Guru Ghasidas University (Central University), Bilaspur, India. His areas of interest include Bibliometrics, Scientometrics, Webometrics, Open-access publications, and Sentiment analysis. He has published over 193 research papers in peer-reviewed national and international journals, 9 authored/ edited books, 36 book chapters, and 82 conference papers.