IMPACT OF SOCIAL MEDIA ON EDUCATION: ANALYSIS OF FEEDBACK ON YOUTUBE

Sonal Lakhotia
Institute of Computer Science, University of Goettingen
e-mail: sonal.lakhotia@stud.uni-goettingen.de

Abstract – **Social media is an undeniably large platform with a plethora of information about all topics of interest. Weather, news, sports, politics and education, and entertainment form the most searched domains in social media.Facebook, YouTube, and WhatsApp have aced user acquisition and information sharing. While these platforms are known for catering to entertainment, they prove fruitful from an educative perspective. YouTube provides a large on-demand and self-learning platform for users to learn and engage with other users for a broader perspective and learning experience. Many reputed universities and tech schools have their YouTube channel and provide courses that would benefit all. This report constitutes that feedback from the content of YouTube videos can enable us to determine if it is a viable resource for learning by analyzing the sentiments.**

Keywords – **YouTube learning, Self-directed learning, YouTube Comments, Sentiment Analysis**

I. INTRODUCTION

Social media such as YouTube, WhatsApp,and Facebook promote knowledge sharing. [1] Knowledge sharing refers to actions in which people make their own internally accumulated learning or external proficiency references that they hold at their disposal available to others. Knowledge sharing requires time and effort and has no direct reward for sharing the resources. These portals enable users to modify and examine each other's ideas and discuss the content that leads to knowledge construction. As per the reviews of an experimental setting, discussions on YouTube have a higher productive potential than on Facebook and WhatsApp.

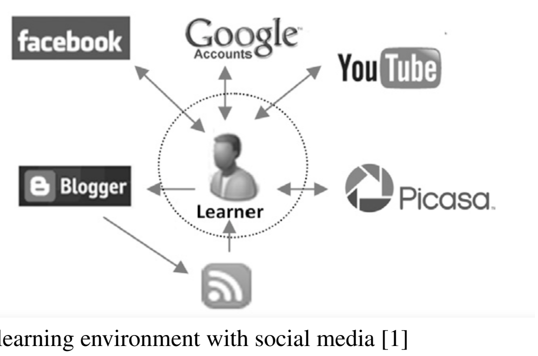

Fig. 1. A learning environment with social media [1]

YouTube provides a platform for upgrading or honing new skills in less time with almost no costs, yet there is no evidence that YouTube learning is as good as MOOC courses or classroom learning. Classroom learning was preferred by many for studying the conceptually rich domains as it might trigger many questions and doubts, or it would include a practical session. Online learning is profoundly popular due to its flexible learning pattern. Learners can take any time of the day, finish a particular lecture and post a comment if a question or an area of doubt arises. Enveloping online learning with technologies like VR and AR has made it possible to experience learning as if it were at the source. YouTube offers features for the VR experience. Several lectures which require intensive observative use cases have been recorded and posted on YouTube.We focus on the role of social media, particularly YouTube, in online learning and education.

YouTube is an excellent platform to understand if informal learning allows knowledge creation and co-construction. It enables the user to upload, create and share content from anywhere and on any topic. Users can consume and analyze the content and provide comments or feedback about a particular video in the comments section. Comments and replies from several users on the quality and the usability of the video content give credibility to the content creators. The quality of the video and its trustworthiness can be determined by the ranking, popularity, and number of views available on that particular video. Some channels like MIT and Harvard do not need a reliability barometer, but most other users aim for maximum subscribers, viewers, and comments on their content.

These statistical factors measure the content's likeability and popularity. Consumer engagement and content consumption can be analyzed through feedback or sentiments from the comments that user posts in the comment section. Users encourage the creator by praising or providing valuable opinions on the methods used in the video. Using YouTube for solving mathematical problems or learning algorithms can enable the user to adopt a similar approach and post whether or not it produced the desired result or if something improves the solution. The discussions allow collective examination and modification of each other's ideas, leading to knowledge co-construction.

YouTube learning also prompts an informal self-learning directive. It requires motivation, intent, and control to curate, create and present content that reaches millions of users and could get both positive and negative feedback. Users might post negative or irrelevant comments in the section to mislead others or defame the person or the content. These sentiments have to be analyzed and filtered from being spam or from a user account. Sentiment analysis of a particular video demonstrates the user's opinions and engagement with the content. The ratio between the number of comments and the views on a video is almost 3:7. Hence it's not efficiently established if YouTube is at par or better than

MOOC resources. Analyzing user comments to determine user sentiments proves that user engages with the content and provide feedback.

II. Related Work

YouTube comment classification and context analysis enable us to understand the relevance of the discussions made in the comments by the users. Fake accounts or bots tend to spam the comments feed with irrelevant comments that challenge the user to analyze the contents of the feedback. Most negative comments are not related to the content of the video description but express angst towards any other ongoing social, political or religious issue in the community. YouTube provides an opportunity and a platform for thoughts and ideas to collaborate. Most people express their opinions, and the rest negate them if they find them unsuited. Therefore it's viable to understand the relevance and sentiments of the users.

A. YouTube comment classification

1) Approach and Workflow: [2] illustrates YouTube comment classification wherein comments are classified into positive, negative, and relevant. [2] demonstrates comment classification using the bag of words and association word list-based features. Feature extraction techniques for the bag of words model are employed using the most frequent words in the video description. The association word list method uses a bag of words for the video description, which follows the collection of association lists for comments and video descriptions [2]. Positive and negative sentiment word lists are created from a dataset, and the polarity is scored based on the presence or absence of these words in the comment. An association list is created using Synset, and the words are made relevant or irrelevant based on their presence or absence in the feature set.

2) Results and Evaluation: [2] describes the extraction of comments and manually categorizing them into four classes, and the experiment evaluated the automatic categorization of these comments based on precision, recall, and accuracy. A video on MS Word with 46 comments and several others with other domains gave results with both approaches. According to the bag of words model, 24 of the 22 relevant, 12 positive, 1 negative, and 11 irrelevant comments were classified as relevant, 10 as positive, 1 negative, and 11 as irrelevant. 27 were classified as relevant, 8 as positive, 0 as negative, and 11 as irrelevant using the association list approach, [2].

Fig. 2. Comment Classification Comparison using Bag of Words Approach and Association List based Approaches [2]

After classification, the comments were evaluated employing Precision (P), Recall (R), and Accuracy (A) metrics [2]. The calculation of precision is carried out with respect to true positive's ratio to the total number of true comments. The recall [2] represents the ratio of the number of true positives [2] to the total number of true comments [2], and the accuracy represents the total number of true comments to the total number of classified comments [2]. TP denotes the number of relevant comments that are correctly classified as relevant, FP denotes the number of non-relevant comments that are classified as relevant incorrectly , TN denotes the number of non-relevant YouTube comments that are segregated correctly as non-relevant comments, and FN describes number of comments relevant to the video description and are classified as non-relevant comments incorrectly [2]. The results were evaluated for both the bag of words and the association list based approaches.

	Association Word List							Bag of Words						
Category	TP	FP	TN	FN	P (%)	R (%)	A (%)	TP	FP	TN	FN	P (%)	R (%)	A (%)
Relevant	22	5	19	0	21.48	100	89.13	20	4	20	2	83.33	90.90	86.95
Positive	8	0	34	4	100	66.66	91.3	10	0	34	2	100	83.33	95.65
Negative	0	0	45	1	0	0	97.8	0	1	44	1	0	0	95.65
Irrelevant	11	0	35	0	100	100	100	10	1	34	1	90.90	90.90	95.65

Fig. 3. Comment Classification using the Bag of Words and Association List Based Comparisons [2]

3) Limitations: An extensive lexicon accounts for and classifies multilingual phrases and words. The above approach curates the positive and negative word lists manually. It might be extremely time-consuming to create such a word set and generate association lists based on the feature sets. The comments are classified based on the video description. There are times when the video description does not fully communicate what the content is. It might be in the form of a question or an idiomatic phrase. Emotion analysis results in low accuracy due to neglecting the effects of grammar in language words and ignoring the semantics of words.Generating words in a relevant or irrelevant bag of words might be tricky. Although we can classify the user comments as positive or negative polarity or relevant and irrelevant as per the video description, a complete and thorough analysis of the sentiment or the user engagement might not be able to be done using this approach.

B. YouTube as a Self-Directed Learning Platform

1) Approach and Workflow: [3] inspects comments from pupils who watched educational videos on YouTube. Sentiment analysis sheds light on the feedback the learners posted in the comments section after watching educational YouTube videos [3]. Self-directed learning involves four critical dimensions motivation, control, self-efficacy, and initiative [3]. YouTube engages the learners by providing ease of access and independence to approach the person who created the video and other users who benefitted from it. Autonomous learning enhances the social learning experience for the users by engaging in effective communications and collaborations.Sentiment analysis is a branch of computational linguistics that aims to automate the study of people's opinions and assessments concerning various objects mentioned in the text, such as products, services, organizations, people, and events [4].The sentiment is represented as a numerical value on a scale: binary (positive/negative attitude), ternary (adding neutral or contradictory), n-ary(for example, [–5, 5]) .

Data collection procedures involve extracting comments from programming, computing, and computer science domains. A customized software extracts about 30,000 video comments from 150 videos. Sentiment analysis is carried out using SentiStrength [3].SentiStrength is a lexicon-based method for sentiment analysis [4]. It assigns two numerical values: the first ranges from -1 for non-negative texts to -5 for extremely negative texts, and the second ranges from 1 for non-positive texts to 5 for highly positive texts [4]. It involves taking a text as input and producing a numerical value that averages the sentiment of the words in the text found from the sentiment lexicon.It is widely used for identifying options and judgments for a variety of web contexts including emoticons and multiple spelling words.

2) Sentiment Analysis Results: YouTube offers a feature to like or dislike a video. The number of views and the number of likes a particular video receives establishes content consumption and appreciation by learners. In a data sample where the videos are 150 in number and the comments extracted are about 30,000, we can analyze the likeability of the video.

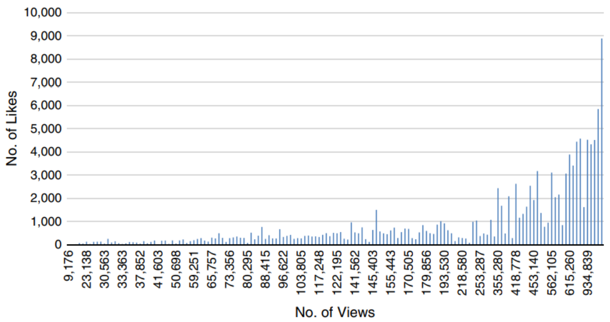

Fig. 4. Like count versus view count for the 150 videos [3]

The number of likes or dislikes alone does not help to analyze user sentiments. A percentage of positive and negative emotions is generated based on the user comments and the number of views of the videos.

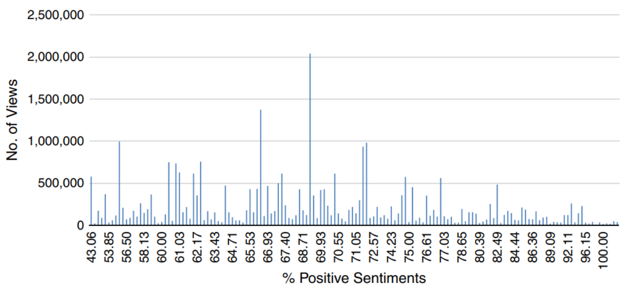

Fig. 5. Viewership and positive sentiments for 150 YouTube videos [3]

The above analysis shows that the number of viewership counts is lesser for videos with positive sentiments. It is an interesting observation as we would have assumed otherwise.

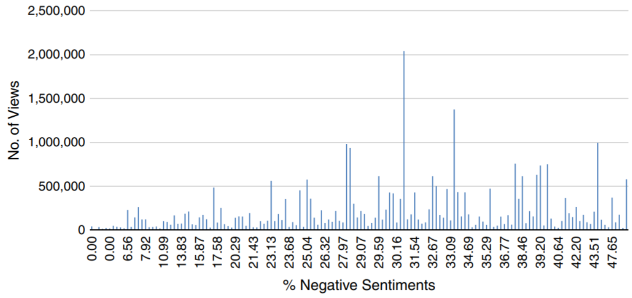

Fig. 6. Viewership and negative sentiments for 150 YouTube videos [3]

The above analysis shows that the number of viewership counts is higher for videos with negative sentiments. A comparison between the above figures illustrates that the percentage of positive sentiments is higher than the negative sentiments for the 150 videos under observation.

3) Limitations: About 150 videos and their comments were analyzed using the lexicon-based technique using SentiStrength.It limits the search to the English dictionary only. Multi-lingual comments and words do not account for this approach. This approach highlights that a learner watches a video and leaves a like, dislike, or comment in the section. It does not explain the learner's collaboration and engagement through an informal learning platform. A Dataset of about 30,000 video comments is analyzed per the SentiStrength lexicon to map if it directs to the self-learning paradigm. Lexicon-based approaches are the most common and oldest techniques for Natural Language Processing tasks such as Sentiment Analysis. A few limitations are that it consumes a lot of time and provides a limited thesaurus. Efficient sentiment analysis using machine learning or deep learning techniques gives us appropriate polarity checks.

III. Using RoBERTa for YouTube Comments Sentiment Analysis

Machine learning algorithms like SVM, Naive Bayes, and Transformer-based deep learning models are widely used for sentiment analysis and determining the polarity of texts. Deep learning-based models give efficient results in sentiment analysis tasks. They account for coherence, semantics, and context while determining the positive, negative, or neutral sentiments. One such deep learning model is BERT(Bidirectional Encoder Representations from Transformers) which uses transfer learning and applies knowledge gained from one task to another similar work [5].It uses a bidirectional learning technique that ensures that the context is maintained. BERT uses encoders which are neural network architectures taken from transformers [5]. Input is passed through an attention layer where all language features are extracted before being sent to the encoder. The encoder creates an encoded representation of the text. The output of the first layer forms the input for the second encoder, thus ensuring the context is preserved. [5] illustrates that BERT is the best word embedding model .It accounts for semantics, syntax, context, and vocabulary of the words in the texts.

Models	Syntactical	Semantics	Contextual	Out of vocabulary
1-Hot encoding	[×]	[×]	[×]	[×]
BOW	[×]	[×]	[×]	[×]
TF-IDF	[×]	[×]	[×]	[×]
Word2vec	[✓]	[✓]	[×]	[×]
GloVe	[✓]	[✓]	[×]	[×]
FastText	[✓]	[×]	[×]	[✓]
BERT	[✓]	[✓]	[✓]	[✓]

Fig. 7. Analysis of the word embedding models [5]

Although BERT is a highly pre-trained bi-directional model, it uses static data masking techniques. An improvement to the BERT model RoBERTa enables dynamic data masking. It trains longer text sequences of data and gives efficient state-of-the-art results. A sentiment analysis pipeline by hugging face that uses the RoBERTa pretrained model to determine the polarity of texts is implemented on some YouTube comments to analyze the sentiment efficiently.An implementation using PyTorch is carried out for sentiment analysis using RoBERTa.

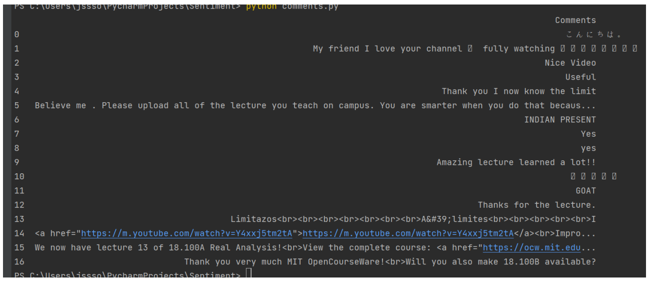

Fig. 8. YouTube comments extracted to be analyzed

A Twitter-RoBERTa-base for Sentiment Analysis pretrained model tokenizes the text sequences and encodes them for analysis.Encoded texts obtained from the tokenizer return tensors when applied to the pre-trained model.

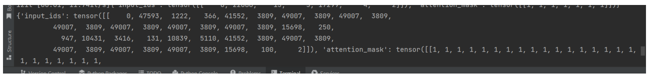

Fig. 9. Text encoding into 1's and 0's embedding for the model to understand

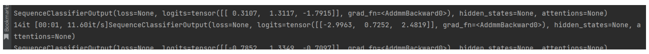

Fig. 10. Tensors returned after the model is run on encoded text

The PyTorch tensors are changed over to NumPy clusters and softmaxed. The negative, neutral, and positive polarities and the sentiment scores are determined from the tensors after processing the texts. Although a very small subset of comments are analyzed in the example, it signifies the efficient sentiment analysis by RoBERTa deep learning transformer model.A sentiment score 1 was assigned for a negative sentiment, 2 for a neutral sentiment and 3 for a positive sentiment.

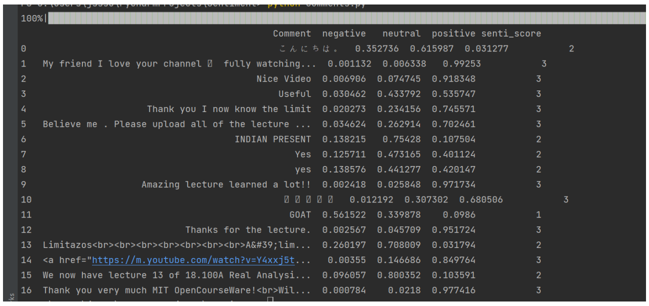

Fig. 11. An overview of the comment, polarities and sentiment score

Although deep learning transformer-based models give very appropriate polarity to the texts to be analyzed, they have certain limitations. The processing of sequences takes an order of n squared and is costly to apply to long text sequences. Transformers are highly demanding on memory, and a larger model would reach a run-time limit. These models work efficiently with a fixed-length string. Whenever there is textual fragmentation, the coherence and context are lost, as a result of which the meaning is manipulated.Sentiment Analysis on YouTube comments using the RoBERTa deep learning transformer-based model enables us to understand the user opinions about the content and quality of the video. The feedback and replies from users on those construct a negative, neutral or positive sentiment model. The likes on the video description show if the content is accepted, but comments ensure that the video aligns with the user's needs and if they benefit from it.

IV. YouTube Actor and Activity Networks

User comments and replies networks give a vivid description of the users and their interaction through feedback. An actor and activity network of user and user comments would enable us to understand learner engagement while watching an educational video. It allows us to determine if the watchers collaborate and augment their learning experience through informal learning platforms.

A. Data Collection

Videos from Edureka's YouTube channel on Data Science are used to analyze the actor and activity network of YouTube. Data Science is the most desired topic of interest among all kinds of learners. Pupils from schools, universities, and working professionals want to learn and improve their skills in Data Science. Edureka offers compact and detailed courses with a focus on necessary topics. Four videos from Edureka related to Data Science are selected, and their YouTube Data is collected.

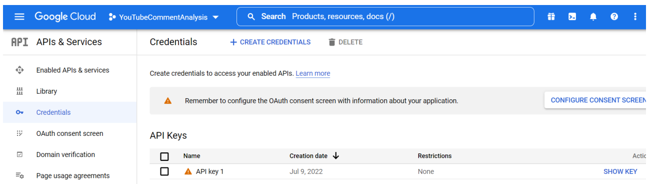

Fig. 12. An API key is required to access public data from YouTube

YouTube data collection requires an API key from Google.

A new project is created in the Google Developer Console wherein it should be specified that YouTube Data API is intended to be used.The latest videos from the channel are not selected due to the lack of comments, views, or likes. Lecture videos chosen for analysis on the topic of Data Science include the following.

- Data Science Full Course for Beginners in 10 hours with 2,537,036 views and 54k likes, and 556 comments
- Data Science Full Course in 3 hours with 307,387 views, 5.3K likes, and 54 comments
- R for Data Science with 18,929 views, 512 likes, and 21 comments
- Python for Data Science Full course with 277,147 views, 7k likes and 69 comments.

Comments are extracted using the YouTube Data API and collected in a CSV file.

Comment	AuthorDis	AuthorPro	AuthorCha	AuthorCha	ReplyCour	LikeCount	Published	UpdatedA	Comment	ParentID	VideoID
Got a	edureka!	https://yt3	http://ww	UCkw4JCw	15	151	2019-08-1	2021-09-2	UgxV4zu9'	NA	#NAME?
Excellent c	Ashwen V	https://yt3	http://ww	UCjnResu3	1	0	2022-07-2	2022-07-2	UgyCCRhn	NA	#NAME?
Hi Team, a	Hasmath I	https://yt3	http://ww	UCcvLS91C	1	0	2022-07-2	2022-07-2	UgyF6Bna	NA	#NAME?
Thanks so	okeke vict	https://yt3	http://ww	UCt9iqdLh	1	0	2022-07-1	2022-07-1	Ugy1BYYN	NA	#NAME?
Good cour	Akudo Agl	https://yt3	http://ww	UCcspkVe	1	0	2022-07-1	2022-07-1	UgxAEwvN	NA	#NAME?
Can I have	Reenu Sha	https://yt3	http://ww	UC1S70Gju	1	0	2022-07-0	2022-07-0	UgyIfI6L14	NA	#NAME?
Could you	ANIL MEG	https://yt3	http://ww	UClSFGW9	1	0	2022-07-0	2022-07-0	UgwSQi5v	NA	#NAME?
Can you m	you tube	https://yt3	http://ww	UCir-zT6D	1	0	2022-07-0	2022-07-0	UgxeVgz18	NA	#NAME?
Pls can we	Ayush Figs	https://yt3	http://ww	UCxjWgUv	1	1	2022-07-0	2022-07-0	UgzWnDd	NA	#NAME?
Kindly Sha	Harsh Gup	https://yt3	http://ww	UC3t-_54r	1	0	2022-07-0	2022-07-0	UgxNxXHH	NA	#NAME?
Can you p	shibna pc	https://yt3	http://ww	UCURnYGr	1	0	2022-07-0	2022-07-0	UgxFL2ZeF	NA	#NAME?

Fig. 13. Data collected from YouTube for further analysis

B. YouTube User Network Analysis

1) Learner Interaction in a YouTube lecture course: A user network for YouTube interaction consists of a unique node, the video under view and other nodes are learners who commented on the video. The edges are the interactions between the users through comments and replies. The YouTube interface allows the users to have top-level comments and reply threads. A top-level comment is on the video, and several other users can reply or comment on it. If learners make a top-level comment or reply to a top-level comment by another user, an edge is created, and the user interaction can be viewed.

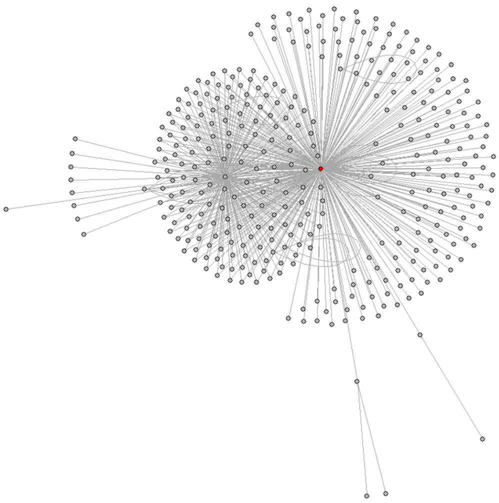

Fig. 14. YouTube Actor Network for a video on Data Science from Edureka

The node attributes include name, screen name, and node type. The name signifies the Channel ID or YouTube's unique user ID. The Screen name implies the user's displayed name, and node type refers to the node being an actor or the video. In the network diagram, the red node specifies the video. The edge attributes include video id, comment id, and edge type. The video id corresponds to the video of which the data is collected, the comment id is the ID of the comment posted by the user, and the edge type signifies if the comment is a top-level comment or a reply to the previous comment made by another user. If a user comment consists of the text from the video description, the edge type is a self-loop. The network consists of 384 nodes and 557 edges. This actor network can be better visualized if we consider a sub-network.A user sub-network consists of only top-level comments by the user and mentioning data science-related terms in their feedback. Few isolate nodes are created while removing edges that consist of top-level comments and self-loop. The texts with mentioned terms like Data, Data Science, Algorithms, Python, and R are collected, and the user interaction is visualized.The nodes in green depict the users who discuss these topics in the comments.

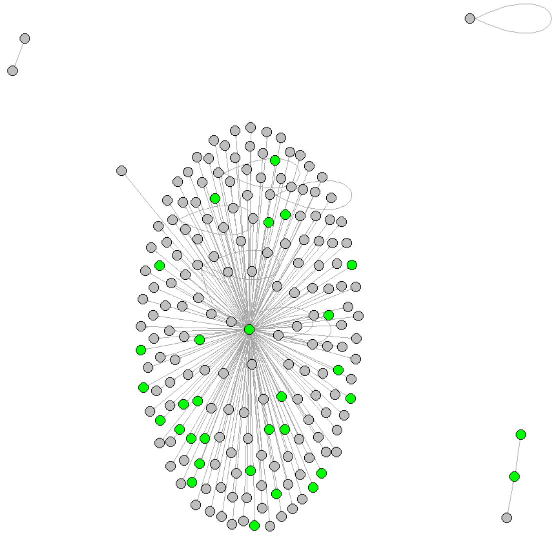

Fig. 15. The sub-network depicting the users sending and receiving replies.

In the above actor network users with comments concerning the Topic of the video and related terms like Data science algorithms, Python programming are viewed particularly. This sub-network highlights that there are numerous comments that are not related to the topic and could be a feedback from the users. It is evident from the network graph that the nodes are connected which shows that users collaborate on the topic and an efficient self-learning and learner engaging platform's is created. A sentiment analysis on these user comments and replies would give us a better understanding of the user views on the topic.

2) User Interaction in the selected YouTube videos on Data Science from Edureka: Videos on Data Science from Edureka's YouTube channel form the data set for studying user and user-activity interaction. The data science lecture videos are a complete lecture or tutorial in data science for beginners and provide a comprehensive description of all the essential concepts. The motive behind choosing the videos from the same channel is to understand if users actively watch videos of the same source on this topic. Video Tutorials on Data Science in R and Python are also analyzed. It is evident from YouTube interface statistics that more users are inclined toward studying data science in Python. YouTube data from these videos are collected, and a YouTube actor network is created. It helps us visualize the relationships between the actors and the comments from these videos. The actor nodes in the network graph correspond to the video and the users who have a unique ID or name. The edges correspond to the comments and replies that the users have exchanged.

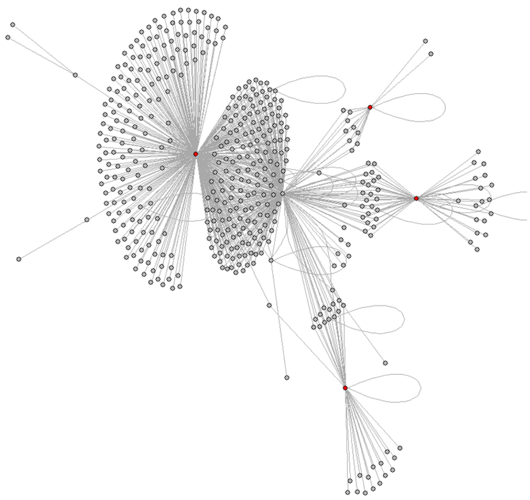

Fig. 16. An Actor Network illustrating user interaction in four YouTube videos.

The actor nodes in red are the four YouTube videos selected for analyzing the user interaction.The grep nodes are the users or learners who left comments or replies on the top-level comments. It is observed that multiple users watched one or more videos on Data Science from Edureka's YouTube channel and left comments and feedback. A comprehensive collaboration of users is depicted in the network. YouTube users interact, collaborate and provide feedback through comments and augment the self learning prospects of the platform.

3) User Activity Analysis in the selected YouTube videos on Data Science from Edureka: A YouTube activity network represents the nodes that are user comments or the video. The node attributes include name, video ID, published information, updated information, author ID, screen name, and node type. Name attributes to the String ID for the comment or the video. Published and updated describes the timestamp at which the comment is created and updated. Node type demonstrates if the node is a comment and the comment type if it is a top-level, reply comment, reply to reply comment or a reply to the video. The edges of the network describe an edge type. The edge type corresponds to the comment that connects all the top-level comments or the reply comments to the video.

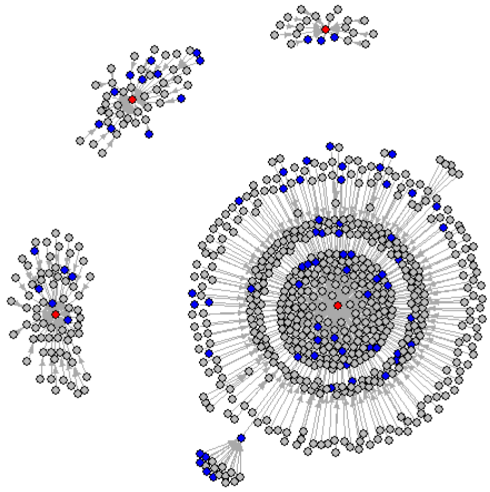

Fig. 17. YouTube Activity Network for the video and the comments by the users

The red nodes depict the YouTube video and the blue nodes describe the comments which have a mention of the terms related to Data Science particularly the comments consisting of the mention of terms Data Science, Algorithms, Python and R programming. The above network is an extensive illustration of the relationship between the video and the comments of the users. Four different videos on the same topic gives us a scattered network wherein the red node describes the video and the rest of the nodes are the comments or the feedback obtained on the video. Most comments on YouTube are an expression of joy, gratefulness or a negative comment about the video therefore it is important to explicitly locate comments related to topic of the video. The Blue nodes indicate the comments made with respect to the video description and the course content. There might be many more comments which relate to the specific topics occurred in the video but the network was created with respect to the broad topics of interest in Data Science namely the Algorithms and the programming languages. YouTube data enables us to understand understand user engagement and active participation through network analysis.YouTube actor network illustrates the relationship between the video and the user interaction through comments. The activity network depicts the relationship between the video and the comments related to the video. Depicting YouTube data as network illustrates user association with the video and other users. We could determine the sentiments associated with these interactions by performing a sentiment analysis on the YouTube comments.

C. Sentiment Analysis of the YouTube comments

In the above section, sentiment analysis using RoBERTa transformer deep learning model was discussed. RoBERTa works efficiently to determine the polarity of the texts. It classifies a text as negative, positive or neutral. An example discussed represented the polarity of the texts and emoticons were efficiently derived using the pretrained model. Sentiment analysis includes checking the positive, negative or neutral polarity of the texts and also the determination of various emotions such as sadness, joy, anger, anticipation, surprise and trust. The comments from four selected videos from Edureka's YouTube channel on Data Science are collected and analyzed for sentiments.

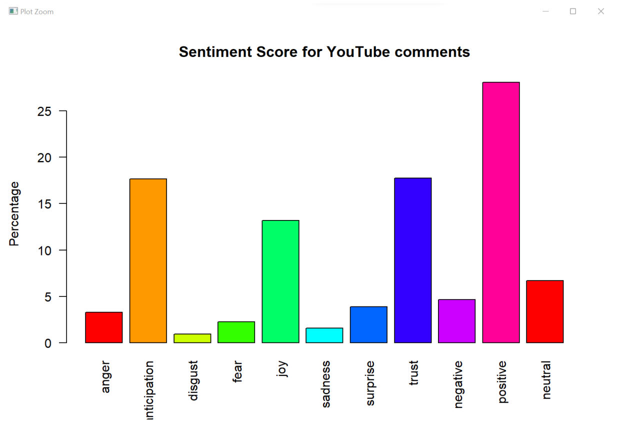

Fig. 18. Sentiment Analysis of the selected YouTube videos on Data Science

The above analysis depicts that most users have positive sentiment about the video and the content. About 690 comments from videos on Data Science have been analyzed. Educational videos on trending topics such as self-driving cars, Network security, game theory, quantum computing and self-learning robots are selected for analysis of sentiments. We analyze the user opinions about these topics and if they anticipate better courses on these topics or have highly satisfied positive feedback. The selected videos include:

- Self-Driving Car with JavaScript Course – Neural Networks and Machine Learning by freeCodeCamp.org with 389,177 views, 14k likes and 806 comments.
- Network Security MIT OpenCourseWare with 112,521 views, 1.2k likes and 34comments.
- What game theory teaches us about war by TED Archive with 3,462,378 views, 78k likes and 3,728 comments.
- Quantum Computing for Computer Scientists by Microsoft Research with 1,754,831 views, 30K likes and 2,077 comments.
- The Power of Self-Learning Systems by DeepMind with 87,096 views, 2.1k likes and 113 comments.

There is no limit to choosing the number of videos and comment, in this experiment variant videos from cutting edge technologies are chosen and their comments are collected using the YouTube API.There is a possibility that users that comments on one video from a particular channel also actively participates in the user feedback in other channel, YouTube actor and activity network would enable its visualization. User satisfaction and content popularity is seen by the likes and counts on the videos. A vast population of learners views the video and leaves a like if they agree with the content or if the content was useful. Comments on the video describe the user's engagement with other users and the video creator or publisher. Feedback from learners motivates the creators to create more content and make available to the learner community for free. YouTube promotes cost-free learning and it directs the students to a self-learning approach. About 6178 comments from the selected videos are collected their sentiment is analyzed.

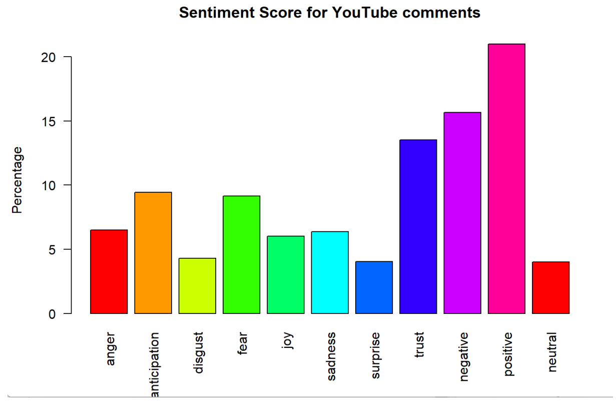

Fig. 19. Sentiment Analysis of the selected YouTube videos

A cumulative analysis of the selected YouTube videos give an elaborate visualization of sentiments expressed by the learner community. Most sentiments from users are positive, a considerable percentage is negative. Other sentiments like anticipation, fear, trust , surprise are also well established.

V. Future Scope of Work

YouTube data from various domains of science and education can be analyzed to provide an elaborate analysis of user interaction on the platform. Several users connect and communicate with many video publishers and reply on comments which provide qualitative feedback for the content or presentation of the video. A large actor network and activity network can be created with 20 videos or more to analyze the user networks and interaction. A keyword dictionary can be created with respect to the videos. Comments related to those specific keywords can be found and a vivid sub network analysis can be performed.

VI. Conclusion

It can be concluded that YouTube provides an user or a learner ample opportunity to learn and develop new skills. It is the most efficient social media for sharpening or honing new expertise. It is established as a self-learning platform wherein users collaborate and exchange their views. User comments enable us to depict the activity of users into a activity network where nodes form comments and the association determines their engagement. It can be concluded that YouTube has a very influential role in education.

REFERENCES

[1] N. Friesen, S. Lowe, "The questionable promise of social media for education: connective learning and the commercial imperative",*Journal of Computer Assisted Learning Volume 28, Issue 3, June 2012*

[2] K.M.Kavitha,Asha Shetty, Bryan Abreo, Adline D'Souza, Akarsha Kondana, "Analysis and Classification

of User Comments on YouTube Videos". *Procedia Computer Science, Volume 177, 2020, Pages 593-598*

[3] Lee, C.S., Osop, H., Goh, D.H.-L. and Kelni, G. (2017), "Making sense of comments on YouTube educational videos: a self-directed learning perspective", *Online Information Review, Vol. 41 No. 5, pp. 611-625.*

[4] Anastasiya Kotelnikova, Danil Paschenko and Elena Razova, "Lexicon-based methods and BERT model for sentiment analysis of Russian text corpora"

[5] Sayyida Tabinda Kokab, Sohail Asghar, ShehneelaNaz, "Transformer-based deep learning models for the sentiment analysis of social media data", *Array Volume 14, July 2022, 100157*