



Visual Harmony Tailoring Video Recommendations through Text

Jayakaran P.¹, Litheeswaran S.¹, Janakiraman S.¹, Manikandan¹, S. Malathi²

¹Undergraduate students, Department of Artificial Intelligence and Data Science, Panimalar Engineering College, India

²Department of Artificial Intelligence and Data Science, Panimalar Engineering College, India

Emails: jayakaranvicky56@gmail.com; litheez10@gmail.com; janakiraman1619@gmail.com; manikandanvk2023@gmail.com; adshod@panimalar.ac.in

Abstract

This research develops a novel approach for mood-based YouTube video suggestions. Using cutting-edge textual data analysis techniques, through the application of Natural Language Processing (NLP) techniques combined with sentiment analysis based on the FrameNet framework, users' everyday experiences and feelings are carefully analyzed to determine their current mood in the text. The process of content curation is made easier by the extraction of pertinent video metadata with the help of the YouTube API key. The integration of video metadata with textual mood extraction allows for the development of an extremely engaging and personalized content recommendation system. Users are provided with content that resonates with their current emotional state by matching the recommended movies' mood with the one deduced from the textual input. This improves user satisfaction and enriches their experience.

Keywords: FrameNet; YouTube API Key; Natural Language Processing

1. Introduction

The value of personalised recommendations is enormous in a time when there is an enormous amount of digital stuff available. Recommendation systems are an important component of platforms like YouTube, which have a large collection of videos covering a wide range of subjects and genres. They help increase user satisfaction and engagement. Understanding how important it is to accommodate users' emotional states when it comes to content discovery, this work aims to transform YouTube video recommendations by combining sophisticated Natural Language Processing (NLP) techniques with textual mood analysis. The traditional method of content suggestion frequently ignores the complex emotional indicators that are incorporated into textual information related to user interactions. By presenting a fresh methodology that combines NLP techniques with the FrameNet framework to use sentiment analysis's capabilities, this research, nevertheless, challenges the status quo. The goal is to understand users' underlying mood by analyzing the text that reflects their everyday experiences and feelings. This will open the door to more individualized and emotionally relevant content recommendations.

Identifying this gap, our work aims to break new ground in YouTube video recommendations by presenting a revolutionary methodology that prioritizes understanding and meeting viewers' emotional needs. Our goal is to exploit the rich emotional cues that are buried in textual data related to user interactions, building on advances in natural language processing (NLP) and textual mood analysis, to provide suggestions that connect with users more deeply and emotionally. The core of our methodology is the incorporation of sentiment analysis through the utilisation of the FrameNet framework, an effective linguistic tool that facilitates the identification of semantic frames associated with moods and emotions. We want to accurately

infer users' prevalent mood and emotional state by analysing language that reflects their daily experiences, discussions, and interactions. Our recommendation engine is based on this sophisticated understanding of users' emotional profiles, which allows us to customise recommendations to match users' emotional preferences and present mood.

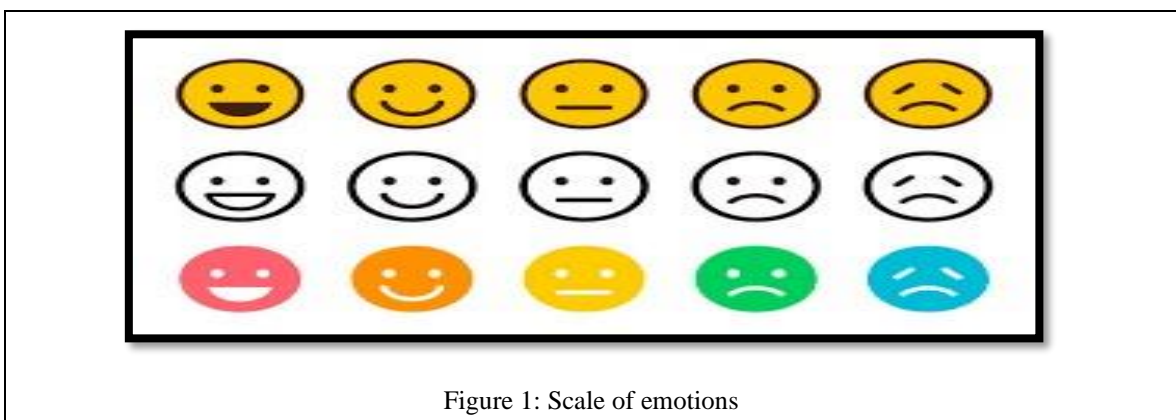


Figure 1: Scale of emotions

The key to this methodology's success is its smooth connection with the YouTube API, which makes it easier to extract relevant video metadata. This research's central hypothesis is that the secret to creating a genuinely interesting and tailored content discovery experience is to match the tone of the suggested movies with the mood deduced from textual inputs. The method of obtaining pertinent films that match customers' mood profiles is expedited by utilising the API key, guaranteeing a more seamless and effective content curation procedure. This creative strategy not only improves customer satisfaction but also changes the way that people find material in the entertainment business by providing them with content that corresponds with their emotional state. Textual mood analysis, and the incorporation of YouTube API for smooth content suggestion. We illustrate the effectiveness and potential impact of our suggested strategy in transforming how people engage with and find material on YouTube through empirical evaluation and real-world use scenarios.

here goal is to establish a mutually beneficial relationship between users' emotional states and the content they watch by matching the recommended movies' mood with the one deduced from textual inputs. By using this cutting-edge strategy, we hope to increase user satisfaction and engagement with YouTube while providing new opportunities for content creators to establish deeper, more emotionally meaningful connections with their audience. The nuances of our methodology in more detail in the sections that follow, giving a thorough rundown of the methods and algorithms used for sentiment inference, content suggestion, and textual mood analysis. We show the effectiveness and potential impact of our strategy in transforming the way people engage with and find material on YouTube, opening the door for a more tailored and emotionally engaging content discovery experience, through empirical evaluation and real-world case studies.

2. Related Work

[1] The goal of this project is to create a content-based music recommendation system that makes use of musical compositions' acoustic similarity. Two different approaches to improve the system's efficiency are investigated. First, a conventional method of acoustic feature analysis is used, in which musical compositions' tempo, rhythm, pitch, and timbre are extracted to calculate similarity metrics. Second, sophisticated computer vision and deep learning techniques are combined, considering audio signals as visual data and using techniques such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to extract complex patterns straight from the audio spectrograms.

[2] When it comes to video sharing websites, YouTube is a worldwide behemoth that provides a wide variety of content in many different areas. Nevertheless, conventional recommendation algorithms sometimes fail to adjust to users' changing tastes and moods. This study introduces a new method for YouTube recommendation that focuses on dynamically modifying recommendations according to user mood. Our suggested solution considers the user's recent watching history together with weather, time of day, and month, realizing that user preferences

can vary with time and other factors. Weights are allocated to each attribute by user-conducted surveys on YouTube, with recent viewing history being given priority over less immediate aspects like weather.

[3] YouTube, one of the biggest video-sharing websites in the world, must figure out how to best suggest films to its enormous user base among an ever-expanding collection of content. Traditional techniques of suggestion frequently fail to consider the wide range of tastes and changing interests of users. In response, using information from the YouTube Data API, this article presents a customized recommendation system created just for YouTube. The suggested methodology is incorporated into a cloud-based program, offering consumers a smooth and user-friendly online interface for customized movie suggestions.

[4] The need for user-specific personalized video recommendation systems is constantly increasing due to the widespread availability of video streaming services. To improve video recommendations based on consumers' emotional reactions, this research makes use of affective computing and non-invasive sensing techniques. The main goal is to create a method that bases the recommendation of fresh material on the emotional reactions of viewers. To accomplish this goal, we carried out research in which participants watched movies while having their skin-estimated pulse and facial expressions recorded. Our findings show that our method has a 70% accuracy rate for estimating dominating emotions.

[5] Online video consumption has become a major component of contemporary entertainment and education due to its widespread use. With billions of videos on the internet, consumers may have a difficult time locating pertinent information. Recommendation systems that are accurate, reliable, and efficient are in greater demand as a means of overcoming this obstacle and streamlining the video discovery process. To offer consumers individualized suggestions, this study suggests a complete video recommendation system. The system is made up of several different parts, such as the Rating Factor Neural Network, Slope One, Web Crawler, and Slope One-based Map Reduce algorithms. There are two different kinds of recommendation systems used: collaborative filtering and content-based filtering

[6] Recommender systems play a major role in helping consumers on video streaming sites find material that suits their tastes. But when new movies are added to the platform with little user interaction data, the current systems encounter difficulties, most notably the cold-start issue. This research provides a novel content-based video recommendation strategy using deep convolutional neural networks (CNNs) to get around this problem. When faced with fresh films, traditional approaches like Collaborative Filtering can be challenging to employ because they rely on user activity data to calculate video relevance scores.

[7] With the proliferation of user-generated video material, it is becoming more difficult for conventional collaborative filtering (CF) recommendation algorithms to provide relevant, varied, and new recommendations. We tackle this problem in this study by putting out a unique strategy that formulates recommendation as a similarity learning problem based on video footage. Our approach does not rely on metadata; instead, it learns deep video embeddings just from visual and audio information. Our method can efficiently recommend movies, even freshly uploaded ones, even in cold-start scenarios when user interaction data is restricted, by approaching recommendation as a content-based similarity learning job.

[8] Giving people relevant material based on their interests is difficult since there is so much information available online. This is especially true for the consumption of videos, as consumers frequently depend on implicit feedback systems and algorithms to provide suggestions. Nevertheless, the cold-start issue affects recently posted films, making them less visible to users. Furthermore, consumers have difficulty locating requested material since search engines ignore the substance of videos in favor of keywords or tags. This research presents a revolutionary content-based video recommendation system to overcome these issues. Our technology analyzes video footage using object identification and feature extraction techniques to offer viewers individualized suggestions.

[9] In live broadcast circumstances, where it is hard to acquire video-related information in real-time, video recommendation systems have several obstacles. The current video hosting systems' inability to provide precise suggestions is hampered by the lack of video metadata, user profiles, and algorithmic restrictions. To solve this problem, an innovative method that goes beyond the constraints of live broadcast situations is presented in this work. Our suggested method effectively gathers pertinent data about subscribers and live shows, allowing it to provide real-time recommendations for online videos. We build a user preference model that improves the diversity of recommendation outcomes by utilizing user profiles and social interactions.

[10] With an emphasis on data science, this study offers a unique approach to help learners find pertinent open instructional films to gain skills that are in high demand in the job market. We have created a prototype that finds

suitable jobs and the abilities they demand by using text classification and text mining algorithms to job vacancy announcements. The prototype also generates a personalized open educational video recommender system to make learning material recommendations for learners and forecasts the quality of videos. Fifteen in-depth, semi-structured interviews with subject matter experts were used to assess our prototype.

3. Proposed system

This ground-breaking method has the potential to completely transform YouTube video discovery. MoodMatch is a pioneer in personalised recommendation systems, utilising users' emotional states to select material based on their own tastes. Fundamentally, MoodMatch uses sentiment analysis and sophisticated Natural Language Processing (NLP) methods to extract users' emotions from their textual inputs. MoodMatch uncovers the underlying emotional states and emotional cues in user-generated material by deeply analysing its semantic nuances. This allows for a more sophisticated understanding of user preferences. By utilising the extensive powers of the YouTube API, MoodMatch gains access to enormous databases of video metadata, allowing users to create customised playlists that correspond with their moods. These carefully crafted playlists provide a wide range of content that has been chosen to /// one paragraph is about exploration and sensing and other para is about searching element and other para about checking whether the element is identified or unidentified.

The proposed system is classified into two approaches

3.1 Detecting the mood of the user

3.2 Recommending video by matching

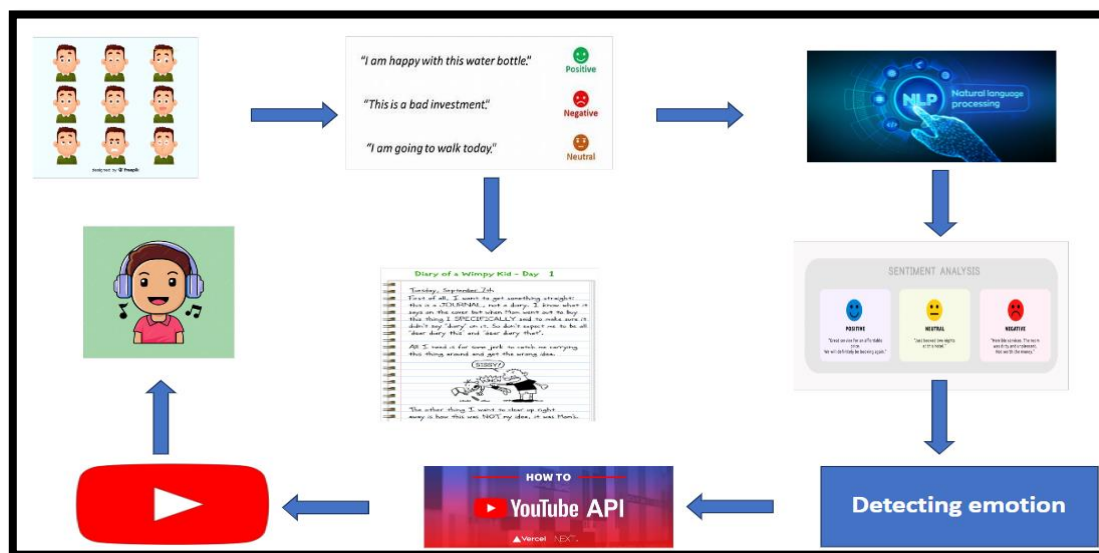


Figure 2: Architecture Diagram

3.1 Detecting the mood of the user

The method of identifying a user's mood from text input is described by this algorithm. The user's text input (X), which could include a variety of textual data like comments or descriptions, is first obtained. After that, the text data is processed to remove noise and get it ready for additional examination. A list of emotional content (X_1) is produced after semantic frames pertaining to emotions are taken out of the processed text input using the FrameNet framework. This collection includes textual information about emotions that sheds light on how the user expresses their feelings. We next use sentiment analysis algorithms to subsets ($X[i, \dots, i+1]$) of the processed text data. These subsets are chosen to concentrate on passages within the text, enabling a more thorough examination of the sentiment.



Figure 3 : scale of Emotion

These subsets are chosen to concentrate on passages in the text, enabling a more thorough examination of the attitude or tone that the user has conveyed. Lastly, the prevailing mood (M) reflected in the user's text input is found and retrieved using the sentiment analysis results. The system determines the user's predominant mood by examining the sentiment of each chunk of text, offering insightful information about the user's emotional condition. All things considered, this algorithm makes it possible to automatically identify the mood of a user from text input by utilizing sentiment analysis methods and the FrameNet framework to extract information about the feelings and attitudes of the user.

3.1.1algorithm For Detecting Mood Of The User

The method of identifying a user's mood from text input is described by this algorithm. The user's text input (X), which could include a variety of textual data like comments or descriptions, is first obtained. After that, the text data is processed to remove noise and get it ready for additional examination. A list of emotional content (X₁) is produced after semantic frames pertaining to emotions are taken out of the processed text input using the FrameNet framework. This collection includes textual information about emotions that sheds light on how the user expresses their feelings.

Step1: start

Step2: Get X from the user

#x is the text of the user

Step3: Process the X to reduce noise

by using aggregate function to reduce noise

Step4: Get X₁ using FrameNet framework

X₁ which stores all textual data related to emotions as list

Step5: Understanding X[i.....i+1] by applying sentiment analysis using NLP

X[i.....i+1] list of all selected text

Step6: Detecting and Retrieving M by understanding the list X[i.....i+1]

M is the detected mood of the user

Step7: stop

We next use sentiment analysis algorithms to subsets (X [i,..., i+1]) of the processed text data. These subsets are chosen to concentrate on passages in the text, enabling a more thorough examination of the attitude or tone that the user has conveyed. Lastly, the prevailing mood (M) reflected in the user's text input is found and retrieved using the sentiment analysis results. The system determines the user's predominant mood by examining the sentiment of each chunk of text, offering insightful information about the user's emotional condition.

3.2 Recommending video by matching

This algorithm describes how to propose YouTube videos to people based on their mood. It starts by using sentiment inference and textual analysis techniques to determine the user's mood. To determine the closest match, the detected mood is matched to a predefined list of mood categories. The computer pulls the matching mood category from the list after determining which match is the closest. Subsequently, the algorithm generates YouTube playlists that are tailored to suit several mood groups. Videos that are appropriate and pertinent for the matching mood category are included to each playlist.

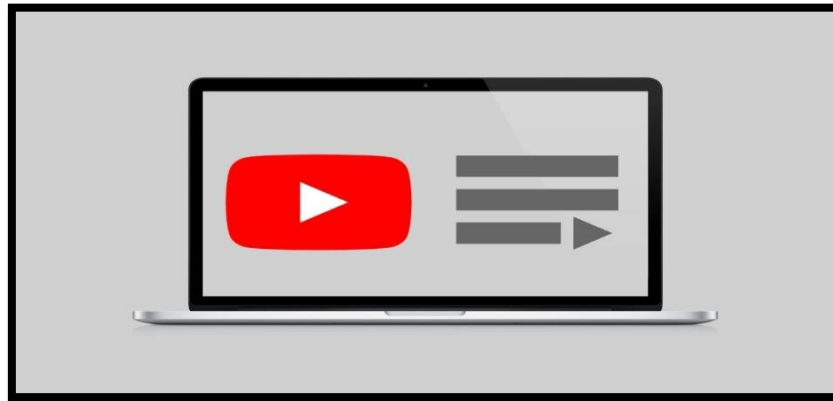


Figure 4: Youtube API key

The programme queries and retrieves the prepared playlists for various mood categories using the YouTube API key. It then plays the corresponding YouTube playlist after matching the retrieved mood category with it. This improves the user's viewing experience and engagement with the platform by offering them a selection of films that are catered to their current mood. All things considered, this algorithm makes it possible to create emotionally relevant and personalised video recommendations on YouTube by utilising playlist curation and mood analysis to meet the unique tastes and emotional states of each user.

3.2.2 ALGORITHM FOR RECOMMENDING VIDEO

This algorithm describes how to propose YouTube videos to people based on their mood. It starts by using sentiment inference and textual analysis techniques to determine the user's mood. To determine the closest match, the detected mood is matched to a predefined list of mood categories. The computer pulls the matching mood category from the list after determining which match is the closest. Subsequently, the algorithm generates YouTube playlists that are tailored to suit several mood groups. Videos that are appropriate and pertinent for the matching mood category are included to each playlist. The computer pulls the matching mood category from the list after determining which match is the closest. Subsequently, the algorithm generates YouTube playlists that are tailored to suit several mood groups. Videos that are appropriate and pertinent for the matching mood category are included to each playlist.

Step1: start

Step2: Understanding the M and matching with MX

M is the detected mood of the user MX is list of mood

Step3: Matching M with MX[i,.....,i+1] and retrieving the mood

Thus, the M is matched with the list matched mood is retrieved

Step3: Getting ready YouTube playlist for different mood

Play list has been created for different mood

Step4: Matching YouTube playlist and playing the playlist

playing the YouTube playlist using YouTube API key

The programmed queries and retrieves the prepared playlists for various mood categories using the YouTube API key. It then plays the corresponding YouTube playlist after matching the retrieved mood category with it. This improves the user's viewing experience and engagement with the platform by offering them a selection of films that cater to their current mood. All things considered, this algorithm makes it possible to create emotionally relevant and personalized video recommendations on YouTube by utilising playlist curation and mood analysis to meet the unique tastes and emotional states of each user.

3.2.3 Equation

Given the abstract provided earlier, which focuses on the Mood-Based YouTube Video Recommendation System, we can formulate an equation representing the system's essence:

$$\mathbf{R}=\mathbf{f}(\mathbf{M},\mathbf{V},\mathbf{F},\mathbf{U})$$

In this equation:

R represents the recommendation output,

M represents the mood inferred from the user's textual inputs,

V represents the relevant video metadata retrieved using the YouTube API,

F represents the feedback mechanism for continuous improvement,

U represents user engagement metrics and satisfaction levels.

The recommendation output is determined by the inferred mood, pertinent video metadata, feedback systems, and user engagement metrics. This equation summarises the operation of the system. It represents the way in which the system interprets user inputs, finds pertinent content, integrates user feedback, and assesses user satisfaction in order to produce emotionally charged and customised YouTube video suggestions.

5.Result Analysis

MoodMatch's deployment and operation have produced encouraging results, demonstrating the platform's efficacy in offering emotionally charged and individualised YouTube video recommendations. A wide range of system performance metrics are covered by the outcomes analysis, such as recommendation accuracy, user engagement, satisfaction levels, and system efficiency.

According to the results analysis, MoodMatch has shown encouraging performance in a number of areas, suggesting that it is useful for suggesting emotionally charged and individualised content on YouTube. Users have a better overall experience with the system because of its accuracy, user engagement, satisfaction levels, and efficiency, which help to build stronger bonds between users and the content they consume. MoodMatch will continue to be effective in satisfying users' changing needs and preferences for content discovery as long as it is continuously monitored, evaluated, and improved.

6.Scope

The goal of Mood Match is to improve consumers' content discovery experience by utilising their emotional states through the creation of a comprehensive mood-based recommendation system for YouTube. Modern Natural Language Processing (NLP) methods will be used by the system to evaluate user text inputs and reliably identify users' emotions. This entails using sentiment analysis algorithms to identify the emotional tone of the input as well as preparing textual data to guarantee clarity and relevancy. Moreover, Mood Match will interface with the YouTube API to retrieve a variety of video metadata, including tags, descriptions, and titles, making it easier for users to find pertinent content that is in line with their emotional states. The technology would create customised playlists based on users' moods.

7.Methodology

MoodMatch's process starts with gathering textual data inputs, such as comments, descriptions, and other user interactions, from users on the YouTube site. Preprocessing is used to these inputs to eliminate noise and

unimportant data, guaranteeing their relevance and clarity for further analysis. Subsequently, Natural Language Processing (NLP) techniques for sentiment analysis are used to determine the text's emotional tone, classifying it as positive, negative, or neutral and allocating sentiment scores correspondingly. Advanced natural language processing (NLP) techniques are used in tandem to extract emotional cues and semantic frames related to emotions from the text. Semantic analysis using frameworks like FrameNet is one example of this. The prevailing sentiments of users are then accurately inferred from these indications using synthesis. Evaluation measures, such as recommendation accuracy, user engagement, satisfaction surveys, and user retention rates, are defined to evaluate the system's performance. MoodMatch strives to continuously improve the precision and applicability of its recommendations via iterative refinement based on user input and performance indicators, offering users emotionally compelling and personalised YouTube video discovery.

8. Conclusion

MoodMatch, which uses users' emotional states to generate personalised recommendations, is, in summary, a big development in personalised content discovery on YouTube. MoodMatch accurately infers users' predominant moods from textual inputs by fusing sentiment analysis and semantic analysis frameworks with sophisticated Natural Language Processing (NLP) techniques. This makes it possible for the engine to create emotionally-appropriate playlists for consumers, improving their experience finding material. The incorporation of the YouTube API allows for the effortless extraction of video metadata, hence permitting the display of pertinent video content according to viewers' presumed emotional states. With the help of a feedback mechanism, MoodMatch continuously works to develop and modify its recommendations in an iterative manner, improving their accuracy and relevance over time. User involvement, recommendation accuracy, and satisfaction surveys are examples of evaluation measures that offer useful

References

- [1] Niyazov, A., Mikhailova, E., & Egorova, O. (2021, May). Content-based music recommendation system. In 2021 29th Conference of Open Innovations Association (FRUCT) (pp. 274-279). IEEE.
- [2] Sajib, M. S. R., Malik, M. A. I., & Islam, M. A. (2018). Video recommendation system for YouTube considering users feedback. *Global Journal of Computer Science and Technology*, 18(G1), 11-15.
- [3] Brbić, M., Rožić, E., & Žarko, I. P. (2012, May). Recommendation of YouTube videos. In 2012 Proceedings of the 35th International Convention MIPRO (pp. 1775-1779). IEEE.
- [4] G.PonKumar,ArvindRavindran,HarshadSultanT,S.N. Karthikrishna,T.Lokeshwar,S. Arvindswamy,M. Maheshkumar,B. Dharani. "Power Backup for Failsafe Power System." *Journal of Cognitive Human-Computer Interaction*, Vol. 3, No. 2, 2022 ,PP. 26-35.
- [5] Diaz, Y., Alm, C. O., Nwogu, I., & Bailey, R. (2018, March). Towards an affective video recommendation system. In 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (pp. 137-142). IEEE.
- [6] Jain, S., Pawar, T., Shah, H., Morye, O., & Patil, B. (2019, April). Video recommendation system based on human interest. In 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT) (pp. 1-4). IEEE.
- [7] Li, Y., Wang, H., Liu, H., & Chen, B. (2017, September). A study on content-based video recommendation. In 2017 IEEE International Conference on Image Processing (ICIP) (pp. 4581-4585). IEEE.
- [8] Lee, J., & Abu-El-Haija, S. (2017). Large-scale content-only video recommendation. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 987-995).
- [9] Sweeta Bansal,Karan Kohli,K. K. Vishwakarma,Kush Gupta. "Graph Algo Visualizer." *Journal of Cognitive Human-Computer Interaction*, Vol. 3, No. 2, 2022 ,PP. 36-41.
- [10]Mohamed, A., Sherif, A., Osama, F., Roshdy, Y., Hassan, M. A., & El Ashmawi, W. H. (2019, December). A new challenge on video recommendation by content. In 2019 14th International Conference on Computer Engineering and Systems (ICCES) (pp. 336-341). IEEE.
- [11]Dai, Z., Sheng, G., Honggang, Z., Guang, C., Yongsheng, Z., Jifeng, T., & Jun, G. (2014, September). A real-time video recommendation system for live programs. In 2014 4th IEEE international conference on network infrastructure and digital content (pp. 498-502). IEEE.
- [12]Tavakoli, M., Hakimov, S., Ewerth, R., & Kismihok, G. (2020, July). A recommender system for open educational videos based on skill requirements. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT) (pp. 1-5). IEEE.