Journal of Cognitive Human-Computer Interaction

♦ ISSN: 2771-1463 Vol: 08, No. 02, 2024 ♦ PP. 25–29

Article DOI: https://doi.org/10.54216/JCHCI.080206

FULL LENGTH ARTICLE

Visual Harmony Tailoring Video Recommendations through Text

Jayakaran P.[1] **Litheeswaran S.**[1] **Janakiraman S.**[1] **Manikandan**[1] **S. Malathi**[2,*]

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

[2] Department 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

Received: October 22, 2023 Revised: January 12, 2024 Accepted: May 19, 2024 ⋆ Corresponding author

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 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 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. 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.

2. RELATED WORK

The goal of content-based music recommendation is to create systems that make use of musical compositions' acoustic similarity. Two different approaches to improve efficiency are investigated. First, a conventional method of acoustic feature analysis is used, in which 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 and convolutional neural networks to extract complex patterns straight from audio spectrograms [1].

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. Prior work introduces a method for YouTube recommendation that focuses on dynamically modifying recommendations according to user mood. The suggested solution considers recent watching history together with weather, time of day, and month, realizing that user preferences can vary with time and other factors [2].

YouTube must figure out how to best suggest films to its enormous user base among an ever-expanding collection of content. Traditional techniques frequently fail to consider the wide range of tastes and changing interests of users. In response, some systems use information from the YouTube Data API to provide customized recommendation through a cloud-based program and user-friendly online interface [3].

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, affective computing and non-invasive sensing techniques have been used. Research in which participants watched movies while skin-estimated pulse and facial expressions were recorded showed that such methods can estimate dominating emotions with promising accuracy [4].

Online video consumption has become a major component of contemporary entertainment and education. 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. Hybrid approaches use collaborative filtering and content-based filtering with components such as neural networks, web crawlers, and MapReduce algorithms [5].

Recommender systems also encounter the cold-start issue when new movies are added to a platform with little user interaction data. Content-based video recommendation strategies using deep convolutional neural networks have been proposed to handle this problem. Other approaches learn deep video embeddings from visual and audio information alone, avoiding reliance on metadata and enabling recommendations even when interaction data is restricted [6, 7].

Additional work has addressed search and recommendation limitations by analyzing video footage with object identification and feature extraction techniques, or by gathering real-time data about subscribers and live shows to improve recommendations for live broadcast situations. In educational contexts, recommender systems can connect open instructional videos with skills demanded by job markets [8, 9, 10].

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 NLP methods to extract users' emotions from textual inputs. MoodMatch uncovers underlying emotional states and emotional cues in user-generated material by deeply analysing 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 large 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 chosen to match the user's emotional

state. The proposed system is classified into two approaches: detecting the mood of the user and recommending video by matching.

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 comments, descriptions, or other textual data, is first obtained. After that, the text data is processed to remove noise and prepare it 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 feelings. Sentiment analysis algorithms are then applied to subsets $X[i,\ldots,i+1]$ 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. Emotion scale used for mood interpretation.

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 useful information about the user's emotional condition. Overall, this algorithm makes it possible to automatically identify mood from text input by utilizing sentiment analysis methods and the FrameNet framework.

3.1.1 Algorithm for Detecting Mood of the User

Step 1: Start.

Step 2: Get X from the user, where X is the text of the user.

Step 3: Process X to reduce noise by using aggregate functions.

Step 4: Get X_1 using the FrameNet framework, where X_1 stores all textual data related to emotions as a list.

Step 5: Understand $X[i,\ldots,i+1]$ by applying sentiment analysis using NLP.

Step 6: Detect and retrieve M by understanding the list $X[i,\ldots,i+1]$, where M is the detected mood of the user.

Step 7: Stop.

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 system 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 several mood groups. Videos that are appropriate and pertinent for the matching mood category are included in each playlist.

Figure 4. YouTube API key and video metadata retrieval.

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 a selection of films catered to the current mood. The algorithm makes it possible to create emotionally relevant and personalised video recommendations on YouTube by utilising playlist curation and mood analysis.

3.2.1 Algorithm for Recommending Video

Step 1: Start.

Step 2: Understand M and match it with MX, where M is the detected mood of the user and MX is the list of moods.

Step 3: Match M with $MX[i,\ldots,i+1]$ and retrieve the matched mood.

Step 4: Prepare YouTube playlists for different moods.

Step 5: Match the YouTube playlist and play the playlist using the YouTube API key.

Step 6: Stop.

3.2.2 Equation

Given the Mood-Based YouTube Video Recommendation System, the system's essence can be formulated as:

$$R = f(M,V,F,U) \tag{1}$$

where 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, and U represents user engagement metrics and satisfaction levels.

The recommendation output is determined by inferred mood, pertinent video metadata, feedback systems, and user engagement metrics. This equation summarises the operation of the

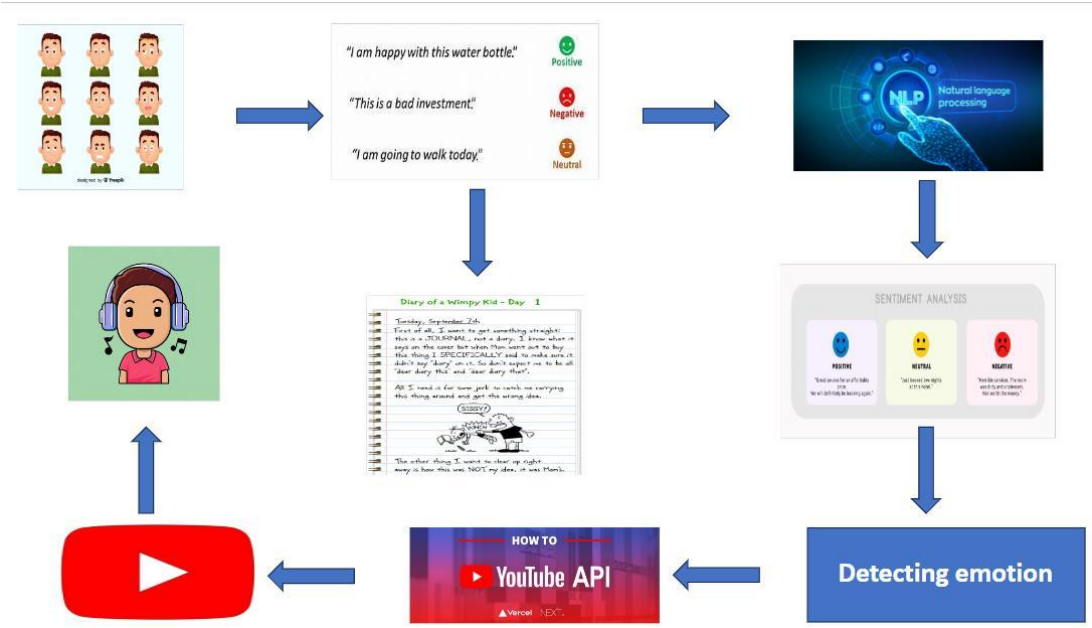

Figure 2. Architecture diagram of the proposed MoodMatch system.

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.

4. 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.

5. SCOPE

The goal of MoodMatch is to improve consumers' content discovery experience by utilising their emotional states to produce personalised recommendations on YouTube. The system covers textual mood analysis, sentiment inference, FrameNet-based semantic interpretation, YouTube metadata extraction, playlist generation, and continuous improvement through user feedback.

By analysing users' textual inputs, MoodMatch identifies emotional tone and retrieves 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 creates customised playlists based on users' moods and offers a mechanism for evaluating recommendation accuracy, engagement, satisfaction, and retention.

6. 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 eliminate noise and unimportant data, guaranteeing relevance and clarity for further analysis.

Subsequently, 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 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.

7. CONCLUSION

MoodMatch, which uses users' emotional states to generate personalised recommendations, is a significant 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 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, 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 indicators for future refinement.

REFERENCES

[1] A. Niyazov, E. Mikhailova, and O. Egorova, "Content-based music recommendation system," in *2021 29th Conference of Open Innovations Association (FRUCT)*. IEEE, 2021, pp. 274–279.

[2] M. S. R. Sajib, M. A. I. Malik, and M. A. Islam, "Video recommendation system for youtube considering users feedback," *Global Journal of Computer Science and Technology*, vol. 18, no. G1, pp. 11–15, 2018.

[3] M. Brbić, E. Rožić, and I. P. Žarko, "Recommendation of youtube videos," in *2012 Proceedings of the 35th International Convention MIPRO*. IEEE, 2012, pp. 1775–1779.

[4] Y. Diaz, C. O. Alm, I. Nwogu, and R. Bailey, "Towards an affective video recommendation system," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2018, pp. 137–142.

[5] S. Jain, T. Pawar, H. Shah, O. Morye, and B. Patil, "Video recommendation system based on human interest," in *2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT)*. IEEE, 2019, pp. 1–4.

[6] Y. Li, H. Wang, H. Liu, and B. Chen, "A study on content-based video recommendation," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 4581–4585.

[7] J. Lee and S. Abu-El-Haija, "Large-scale content-only video recommendation," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 987–995.

[8] A. Mohamed, A. Sherif, F. Osama, Y. Roshdy, M. A. Hassan, and W. H. El Ashmawi, "A new challenge on video recommendation by content," in *2019 14th International Conference on Computer Engineering and Systems (ICCES)*. IEEE, 2019, pp. 336–341.

[9] Z. Dai, G. Sheng, Z. Honggang, C. Guang, Z. Yongsheng, T. Jifeng, and G. Jun, "A real-time video recommendation system for live programs," in *2014 4th IEEE International Conference on Network Infrastructure and Digital Content*. IEEE, 2014, pp. 498–502.

[10] M. Tavakoli, S. Hakimov, R. Ewerth, and G. Kismihok, "A recommender system for open educational videos based on skill requirements," in *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2020, pp. 1–5.