Analysis of Sequences of Emotions in Movies

‘Harnessing the Emotional Energy of Films’

Contents

Abstract

Movies are a big part of our daily entertainment nowadays. We enjoy movies because they make us feel different emotions that we do not usually experience daily. These emotions can be anything from happiness to sadness. When we watch a movie, we go through a series of emotions, just like the characters in the film. In this paper, we wanted to understand how filmmakers use emotions to tell their stories. We used different tools and techniques to analyze the emotions shown by the characters. We wanted to know how filmmakers mix these emotions to create engaging and captivating narratives. We also wanted to know how the audience feels when they watch movies. To do this, we asked a group of people to tag the emotions they felt while watching short films manually. Using the background music and video, we also predicted the emotions evoked by the same films. By comparing the actual emotions of the audience with our predictions, we can learn more about the connection between how people feel and the movie's storytelling. Our research helps filmmakers and researchers better understand the power of emotions in movies. By knowing how emotions impact storytelling, filmmakers can create more engaging and meaningful movies for people to enjoy.

1. Introduction

The first motion picture was released in the late 19th century, and from there, we have yet to see anything but progress in that industry. We also feel connected to the movies and their characters and feel like we are part of that fictional world. This is what is known as the *Diegetic effect* [1]. Creators aim to maximize this effect by ensuring the audience feels various emotions in the movies. One can define the word 'emotion' in many ways. One of the definitions is that '*Emotion is a construct that describes the feelings of an individual or group*'[2]. There are various models which depict the various emotions in different manners. One of the most famous models is *Plutchik's wheel of emotion* [3], where he classified all emotions into eight complex emotions (joy, sadness, anger, fear, trust, disgust, surprise, and anticipation). These emotions are further divided into sub-emotions, as shown in Figure-1. Plutchik's model is very popular in sentiment analysis in the field of NLP (Natural Language processing).

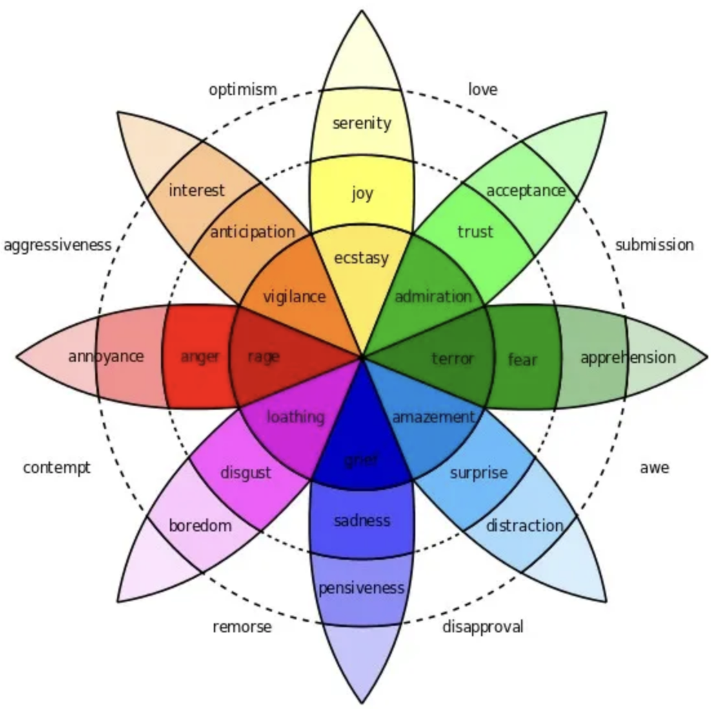


Fig:1 Plutchik's Wheel of Emotion

In a paper by Philippe Aurier and Guergana Guintcheva [4], They manually interviewed 25 individuals, each lasting 40 to 60 minutes. They concluded that emotions play a significant role in the overall experience of a movie. They observed a particular pattern in these emotions, motivating us to observe and analyze them.

We observed that emotion could be analyzed in 3 ways in feature films, the first is through Subtitles (or scripts), the second is the background music (BGM), and lastly, through the facial and physical expressions of the characters in movies.

In this paper, we covered the analysis of emotions using subtitle files in section-3. This will cover the emotion expressed by characters in a movie but not correctly cover the viewer's emotions. We utilized the deep learning method and modification of Bidirectional Encoder Representations from Transformers (BERT) to analyze emotion and observe some patterns in a movie. In section-4, we manually hand-tagged the viewers' emotions for 13 short films; for that, we gathered 5-7 people to watch the short film and discuss the frame emotion and the emotion they felt during the scene. The selection of short films for our study was motivated by their potential to evoke a profound emotional impact on viewers while requiring less time for consumption. The shorter duration of these films facilitated the examination of two crucial dimensions of emotion analysis: background music and facial expressions. We subsequently compared the results of our analysis with manually assigned emotions by us.

2. Related Works

Text-based classification of emotions is one of the mainstream applications of Natural Language Processing (NLP) fields. [8] works as a guide for doing an emotional analysis of the text. They gave thorough explanations of different emotion models and NLP methods which are used based on the use cases.

One of the most common methods utilized for emotion detection is the Lexicon method. These lexicons map individual words to the emotions that could be expressed from that word. NRC lexicon [5] is one of the most commonly used lexicons since it has over 14,000 words with eight emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) and two sentiments(negative and positive).

In [6], they applied a lexicon method to analyze emotion using subtitle files. They used a total of 3 lexicons, one of which was the NRC lexicon. They computed the overall emotional proportion of movies, created a secondary data set of subtitle files, and computed the correlation between emotions and the IMDb score of users. They did not touch on the emotional analysis of each scene in this work.

In [8], they addressed some of the issues with lexicon-based methods; one of them was that the currently existing emotion lexicons are not domain-specific and specify each term's generally comprehensible emotion label. Other than that lexicon-based method does not account for negations and gives the wrong classification in that case.

Somewhat similar work was also done in [7]; instead of subtitles, they used an audio description provided for the visually impaired, which is better than subtitles because audio description files have more details. They used the reverse of the lexicon method. They mapped from emotions to the words which could excite that emotion. They marked emotions at the time stamp at which the word appears. They were able to observe some patterns of emotions in their work.

Some machine learning and deep learning-inspired methods for emotion analysis are generally proven more potent than lexicon-based methods. One such work was done in [9]. They applied deep learning, evaluated Bidirectional Encoder Representations from the Transformers (BERT) model on Twitter data, and reported 89% F1 for emotion detection on four emotions.

The study on short films [16] shows an analysis of some of the successful Indian short films, how the art of storytelling is so powerful with short films, and Youtube as a platform for this kind of short film. This paper helped us choose the short films for our hand-tagging experiment and shed light on the changing impact of short films over time.

3. Analysis of Subtitles of Movies

In this section, we dive into the emotions found in movies and how they affect our experience. We studied the subtitles from 90 different movies to see how emotions were portrayed and if there were any patterns across different types of movies. To do this, we collected the data, prepared it for analysis, divided it into smaller parts, and used advanced models to predict emotions.

By looking at the emotions and how they were distributed, the frequency of different emotion pairs, and how they changed over time, we learned a lot about the emotional journey of movies. We discovered how different emotions relate to each other and how they evolve throughout a film.

Understanding these emotional aspects gives us a deeper understanding of how movies affect us emotionally. This information is useful for filmmakers, researchers, and anyone interested in improving storytelling techniques and exploring the emotional impact of movies.

3.1 Data Collection and Preprocessing

We collected subtitle files from the Opensubtitles website [22], which is freely available for download. We collected subtitle files for three different categories of movies. In the 1st category, we used the top 30 movies with IMDb ratings [23]; in the 2nd and 3rd categories, we chose 30 movies with IMDb ratings between 6 to 8 and 3 to 5, each with a minimum of 10,000 votes. So classify them as good-performing, mediocre-performing, and poorly-performing movies based on audience reviews. We ensured that there was no bias in terms of the rating or genre of the film in the second and third categories. So we have subtitle files of movies of 3 categories, each consisting of 30 movies, in total 90 movies.

In preprocessing, we first converted .srt subtitle files to .xml files to use them for further analysis by making data frames of them. In the next phase, we did some data cleaning on these subtitle files before feeding them into the model. First, we removed some unnecessary parts, which are adpositions(like ‘in,’ ‘to’ etc.), determiners(‘a,’ ‘an’ and ‘the’), numbers, punctuations, and symbols (like ‘@,’ ‘$’ etc.), since they do not contribute to emotion generation at all. We also observe that movies have frequent use of contractions and slang words. To tackle that, we used an open-source Python library called *contractions* [11]; we also added some of our own slang words, which we observe in movies frequently with their actual English word (like ‘Ma’ will be replaced by the word ‘mother’ etc.).

3.2 Method/Our Approach

3.2.1 Segmentation

One of the most crucial tasks for analyzing emotions is deciding the number of dialogues chosen as scenes from a subtitle file. It is essential because we require the proper context of the scene, and then we can implement the proper emotion detection. Without that, we might take fewer dialogues, then we will not have the proper context of the scene, and our emotion prediction model might give the wrong output since the model is missing out on some information. Alternatively, we may have more dialogues than required. In that case, we have some unnecessary context to the scene, which might cause wrong predictions of emotion. So, we needed a generalized approach to implement in all our 90 movies.

In our first approach, we calculate the average time taken for one dialogue within different categories of movies (1st category: *2.5s,* 2nd category: *2.4s,* 3rd category: *2.3s*). Considering that the average movie scene lasts between 2 and 3 minutes[13], we set a time threshold of 90 seconds. Based on this threshold, we merge dialogues within each category until the cumulative duration of the merged dialogues reaches or slightly exceeds 90 seconds. This approach allows us to create coherent dialogue segments that capture the emotional dynamics within the predefined time frame.

The second approach of this segmentation was based on the total number of scenes in movies. In [12], they analyzed over 12,000 movie scripts, and one of their important inferences was that a movie has around 110 scenes. So, we took it as a reference, divided the total number of dialogues by 110, took the dividend as the number of dialogues in a scene, and merged those dialogues to make a single scene.

3.2.2 Emotion Prediction

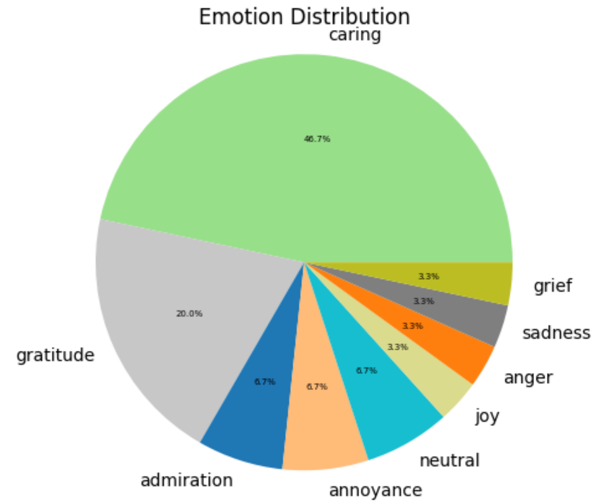
After doing the segmentation, we implemented emotion detection on the segmented scenes. We used the EmoRoBERTa [14] model, accessible on the Hugging Face website. The EmoRoBERTa model modifies the BERT (Bidirectional Encoder Representations from Transformers) model. This model is trained on the GoEmotion dataset [15], which comprises around 58,000 Reddit comments with 28 emotions. The EmoRoBERTa model is made up by changing the key hyper-parameters of the BERT model, like batch size and learning rate.

We used the EmoRoBERTa model on both the segmented methods to see how each performs in emotion detection. The preprocessed movie files have four columns(Start time, end time, text, and cleaned text). We added two more columns: the column of 'time taken' using start and end time, and the detected emotion. To fast-pace this process, we made a dictionary for each category (movie index as key and data frame of subtitle file as value). We made a function that can take these dictionaries as input and give dictionaries with each data frame added with an emotion column.

3.3 Observations

3.3.1 Analysis of Movie Emotion Proportions

We created pie charts to visualize the overall distribution of emotions in movies. Figure-2a represents the pie chart of the movie 'Spirited Away' using approach-1, where scenes were segmented every 90 seconds. Similarly, Figure-2b shows the pie chart of the same movie using approach-2, where the movie was divided into around 110 scenes. We observed that approach-2 had a higher portion of 'neutral' emotions. One possible explanation is that dividing the movie into smaller segments may have provided less context for emotion detection by the model.



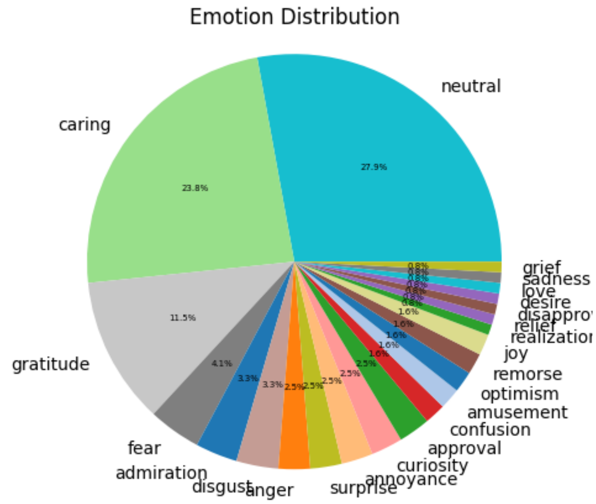


Fig:2a Total emotion proportion of movie 'Spirited Away' (approach-1)

Fig:2b Total emotion proportion of movie 'Spirited Away' (approach-2)

3.3.2 Analysis of Emotion Pairs in Movies

Frequency of Emotion Pairs

Another approach we analyzed was looking at the frequency of emotion pairs in movies. This could give insight into the interconnection between emotions and the aspect of which pair is highly exploited for the movie storytelling process.

We calculated the frequency of emotion pairs within each movie category. Figure-3a and Figure-3b represent the bar charts of the top 30 emotion pairs in the first category, segmented by approach-1 and approach-2, respectively. We observed the dominance of certain emotion pairs in specific categories, highlighting their significance in movie storytelling.

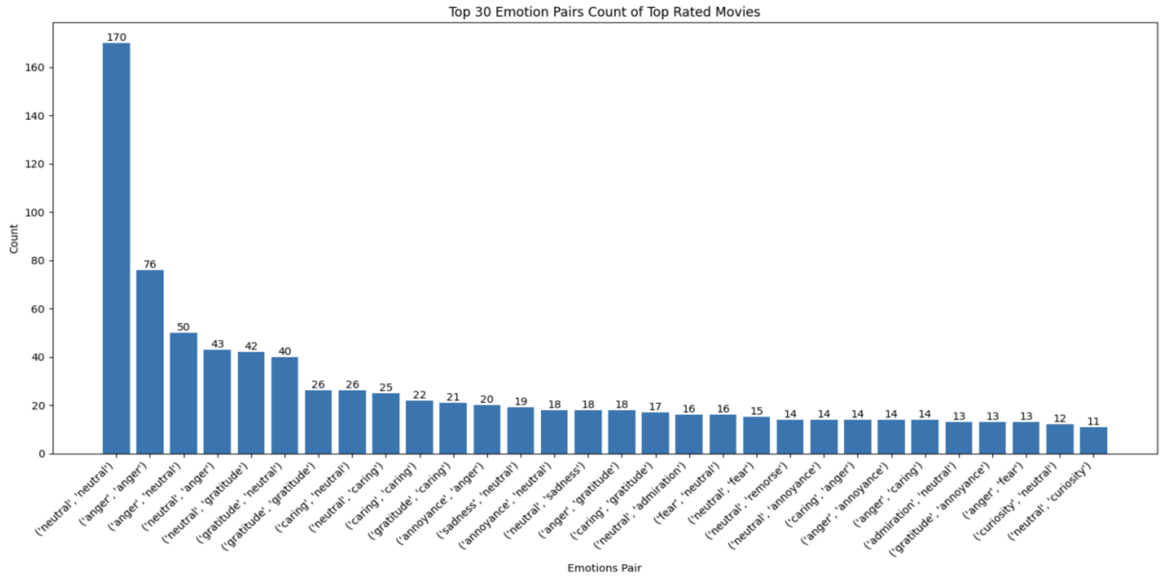


Fig:3a Top 30 emotion pairs of the first category (approach-1)

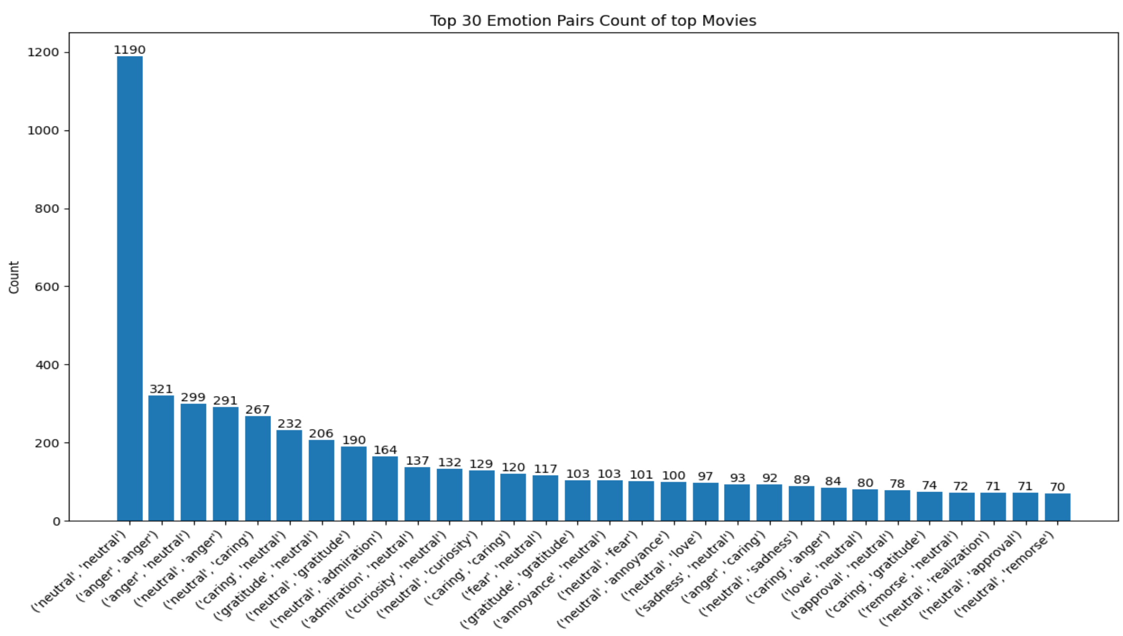


Fig:3b Top 30 emotion pairs of the first category (approach-2)

Dominant Emotion Pairs in Specific Categories

Figure-4a displays the bar chart of the top 30 emotion pairs in the third category, segmented by approach-1. Notably, the pair 'gratitude-gratitude' dominated in this category. It is important to note that when we refer to emotions like 'anger,' we are discussing the emotions portrayed by the characters in the scene, which may or may not be experienced by the audience.

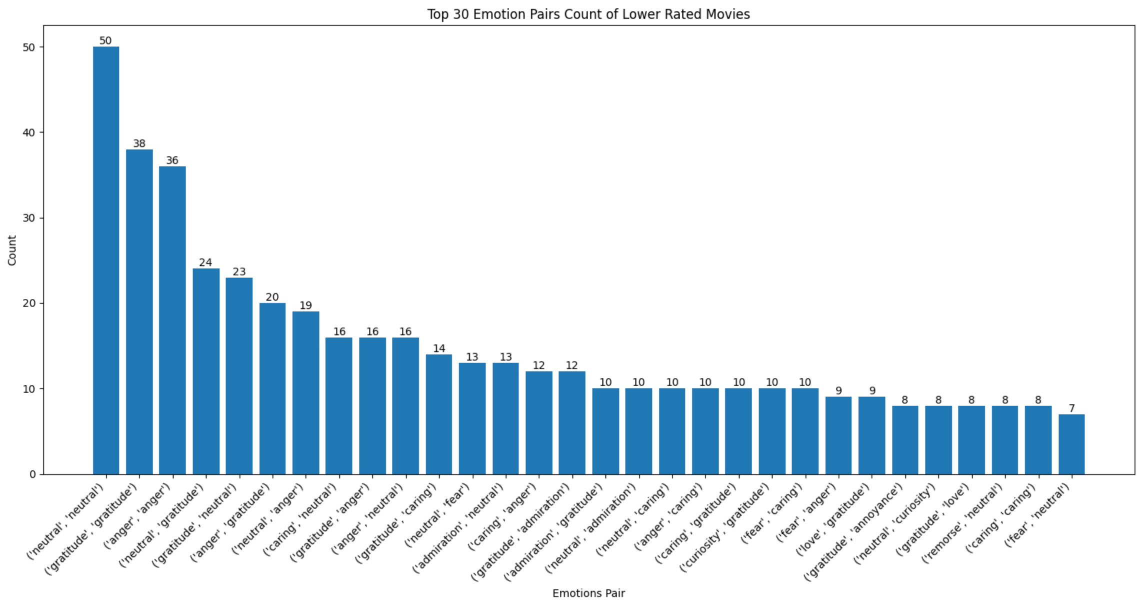


Fig:4a- Top 30 emotion pairs of the third category (approach-1)

Patterns of Distinct Emotion Pairs

As we discussed the pairs of emotions, we realized it would be ideal to ignore the ‘neutral’ emotion and emotion pairs with the same emotion (for example, ‘fear-fear’). we constructed a dictionary excluding 'neutral' emotions and pairs with the same emotion. This allowed us to identify the most frequent pairs of distinct emotions. Figure-4b presents the bar chart of the top 30 distinct emotion pairs using approach-1. Surprisingly, we found that 'gratitude' appeared frequently in the top pairs, despite 'anger' dominating in previous observations. This suggests that while 'gratitude' occurs often, it tends to be present for shorter durations than 'anger,' resulting in a higher frequency of the 'anger-anger' pair.

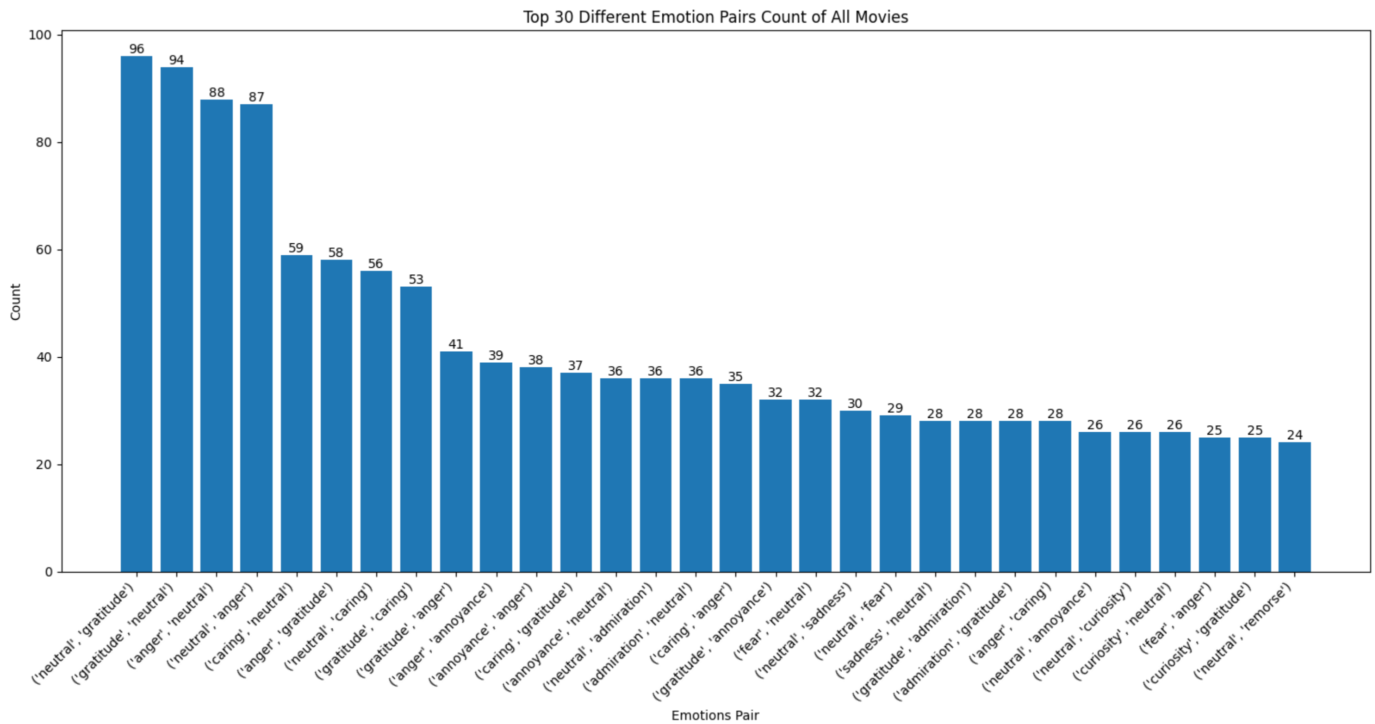


Fig:4b- Top 30 distinct emotion pairs of all 90 movies combined (approach-1)

3.3.3 Temporal Analysis of Movie Emotions

This subsection focuses on understanding how emotions change over time in movies and how they vary across different genres and categories.

Time-Series Scatter Plot of Emotions

The primary thing we are exploring in this paper is how emotions change with time in a movie, which emotion is followed by which, and so many other questions related to the sequences of emotions. For that, we made a time-series scatter plot of emotions for all 90 movies using both segmentation approaches. The x-axis consists of the time, and the y-axis consists of different emotions, so any point has coordinates as (time, emotion). Figure 5, Figure 6, and Figure 7 consist of a graph of 3 films (one from each category) with both approaches.

As you can see in figure-5 and figure-6, the ‘anger’ emotion is dominant overall, but figure-7, which consists of graphs from the movie ‘Scary Movie,’ a horror movie, clearly dominates the emotion ‘fear.’ Using these graphs (Figure-5,6,7), we can visually see how emotions vary with time for different genres and different categories of movies.

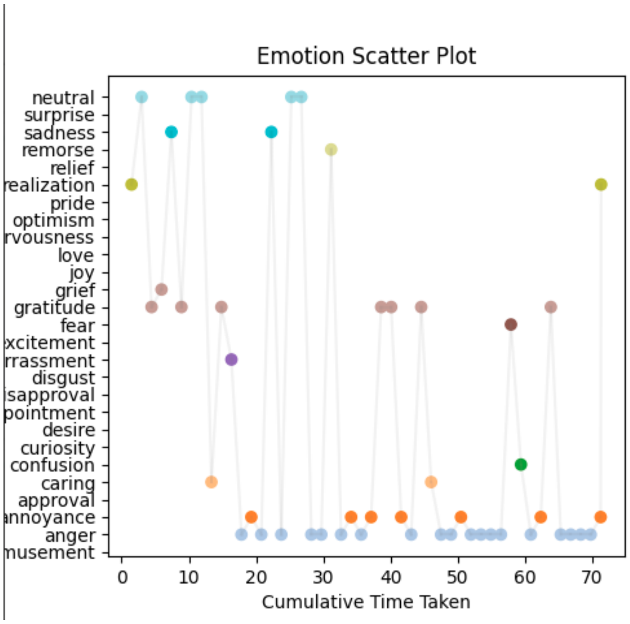


Fig:5a Emotion Vs. Time scatter plot of the movie 'Fight Club' (1999) by approach-1

Fig:5b Emotion Vs. Time scatter plot of the movie 'Fight Club' (1999) by approach-2

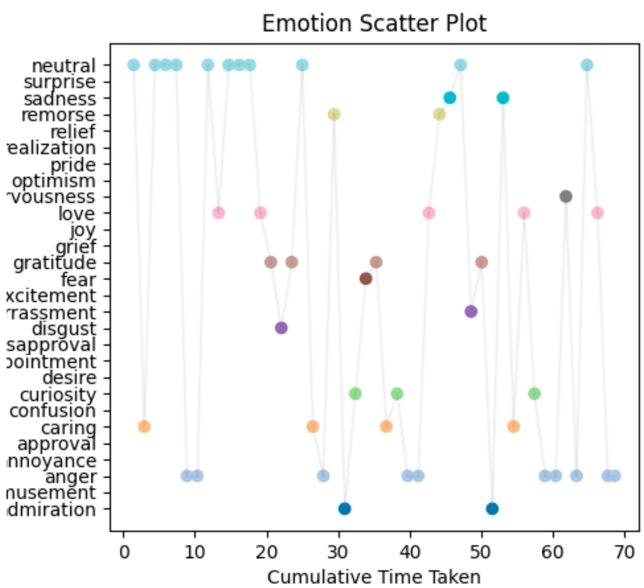


Fig:6a Emotion Vs. Time scatter plot of movie 'August Osage County ' (2013) by approach-1

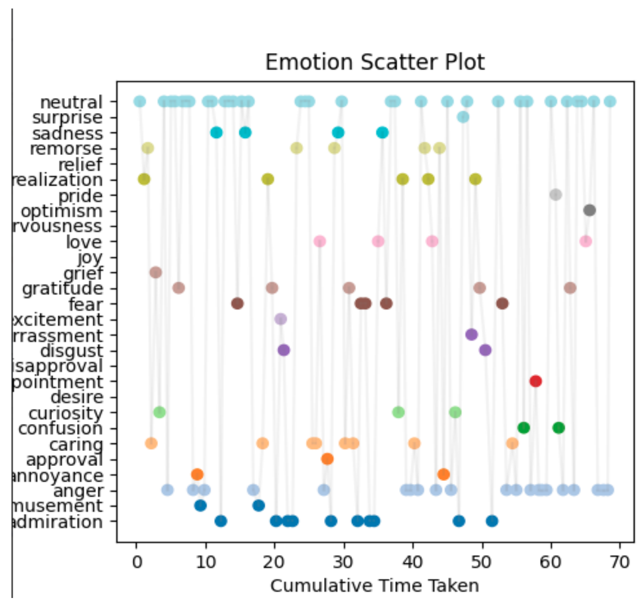


Fig:6b Emotion Vs. Time scatter plot of movie 'August Osage County ' (2013) by approach-2

Fig:7a Emotion Vs. Time scatter plot of the *movie 'Scary Movie' (2013) by approach-1*

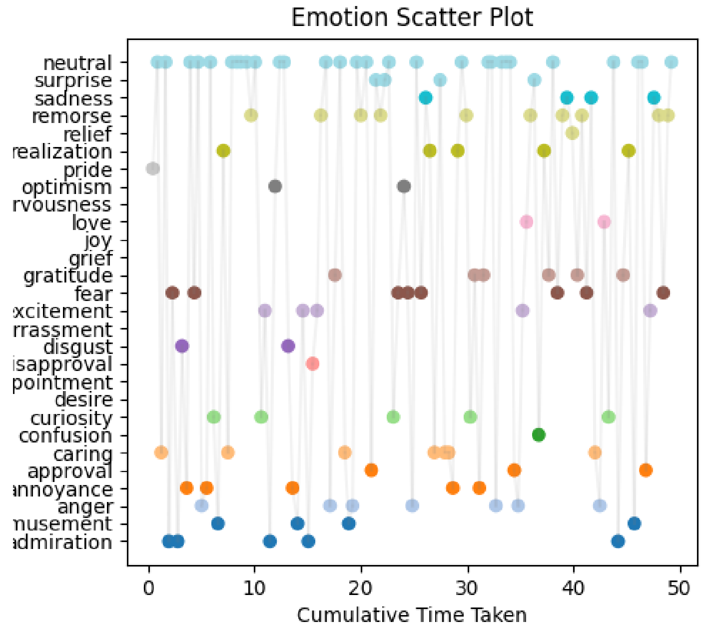


Fig:7b Emotion Vs. Time scatter plot of the movie 'Scary Movie' (2013) by approach-2

3.3.4 Interconnection and Distribution of Emotion Pairs

In this subsection, we explore the interconnection and distribution of emotion pairs across different categories and genres using matrix analysis.

Distribution of Distinct Emotion Pairs:

Figure-8 illustrates the scatter plot distribution of distinct emotion pairs across the three movie categories using approach-1. The graph reveals that certain emotion pairs are more prevalent in specific categories, such as the pair ('gratitude,' 'admiration'), which occurs significantly more in the lower category compared to the other two categories. We also observed that several emotion pairs are unique to the lower and middle categories and not present in the top category.

The Emotion Matrix:

The emotion pair also opens up one more analytics tool for us: ***the matrix***. To analyze the probabilities of emotion occurrence, we constructed a 26x26 emotional matrix (excluding 'neutral' emotions) where each entry represents the probability of one emotion following another.). For e.g., the ('x,' 'y') position will be the probability of occurrence of 'y' emotion after 'x' emotion. We only included the distinct emotion pairs (diagonal elements will be zero). Figure-9a shows the heatmap of the emotional matrix for upper-category movies using approach-1. Lighter colors indicate a higher occurrence of pairs, with 'anger' and 'gratitude' having more lighter-colored cells.

We also generated similar matrices for different genres, as seen in Figure-9b, where 'Romance' showed minimal involvement of the 'anger-anger' pair. Additionally, the genre 'Fantasy' exhibited a more uniform emotion distribution throughout the matrix than other genres.

In addition, we have created a percentage matrix where each row represents the relationship of a specific emotion to the other 25 emotions. The sum of each row adds up to 100%, ensuring that the column represents a complete distribution. An illustrative example of this type of matrix can be seen in Figure-9c.

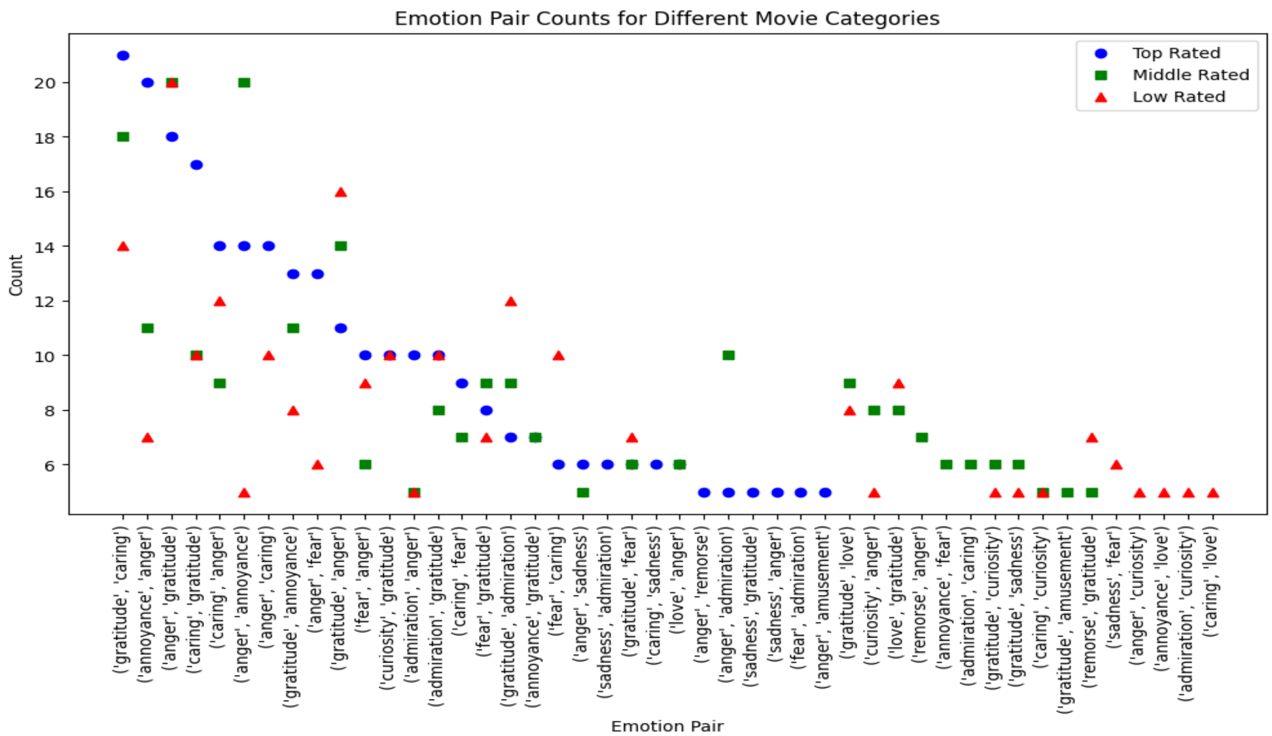


Fig:8- Distinct Emotion pair counts across the categories (approach-1)

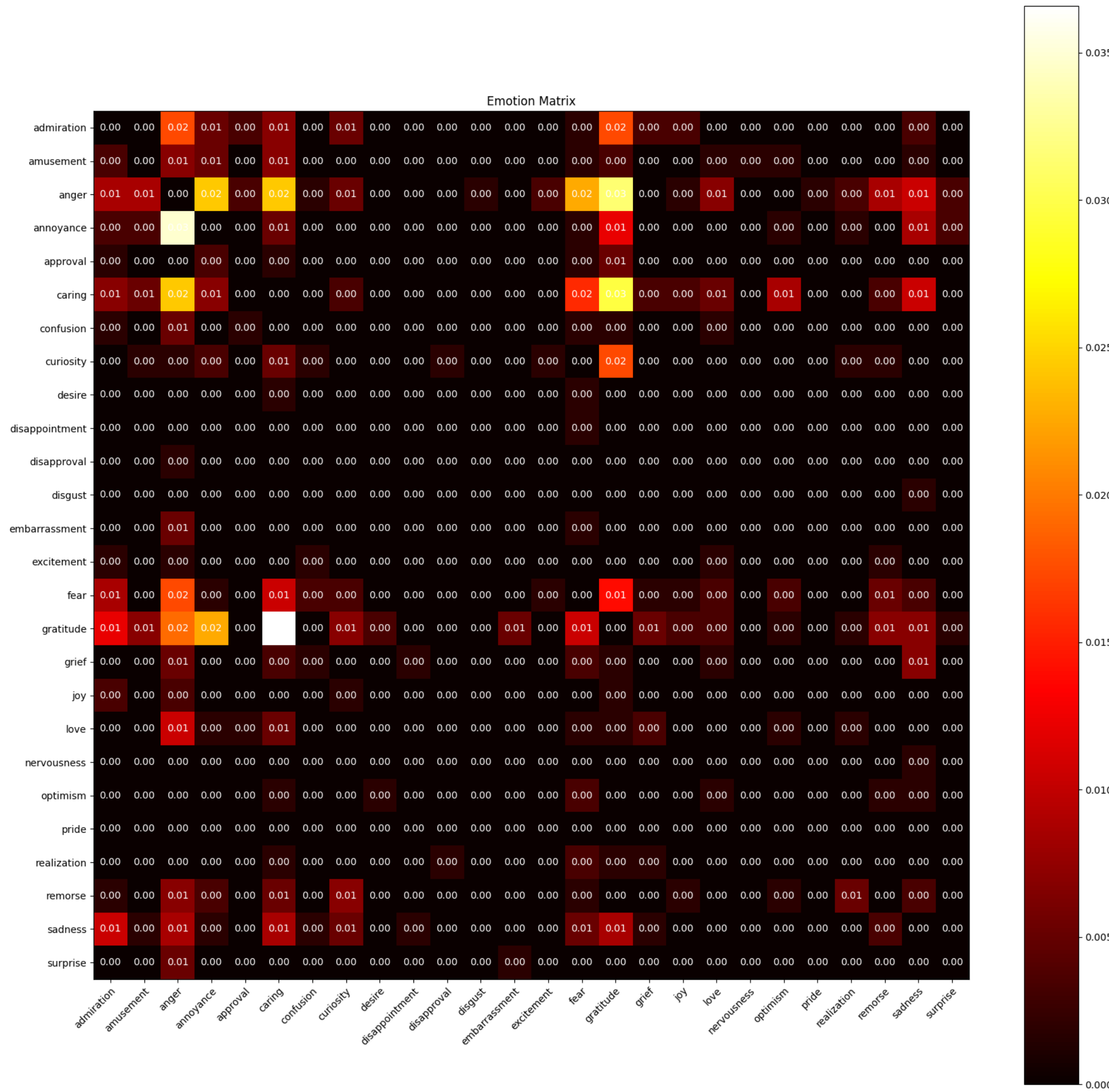


Fig:9a- 26×26 Heat map/Matrix of Upper category(approach-1)

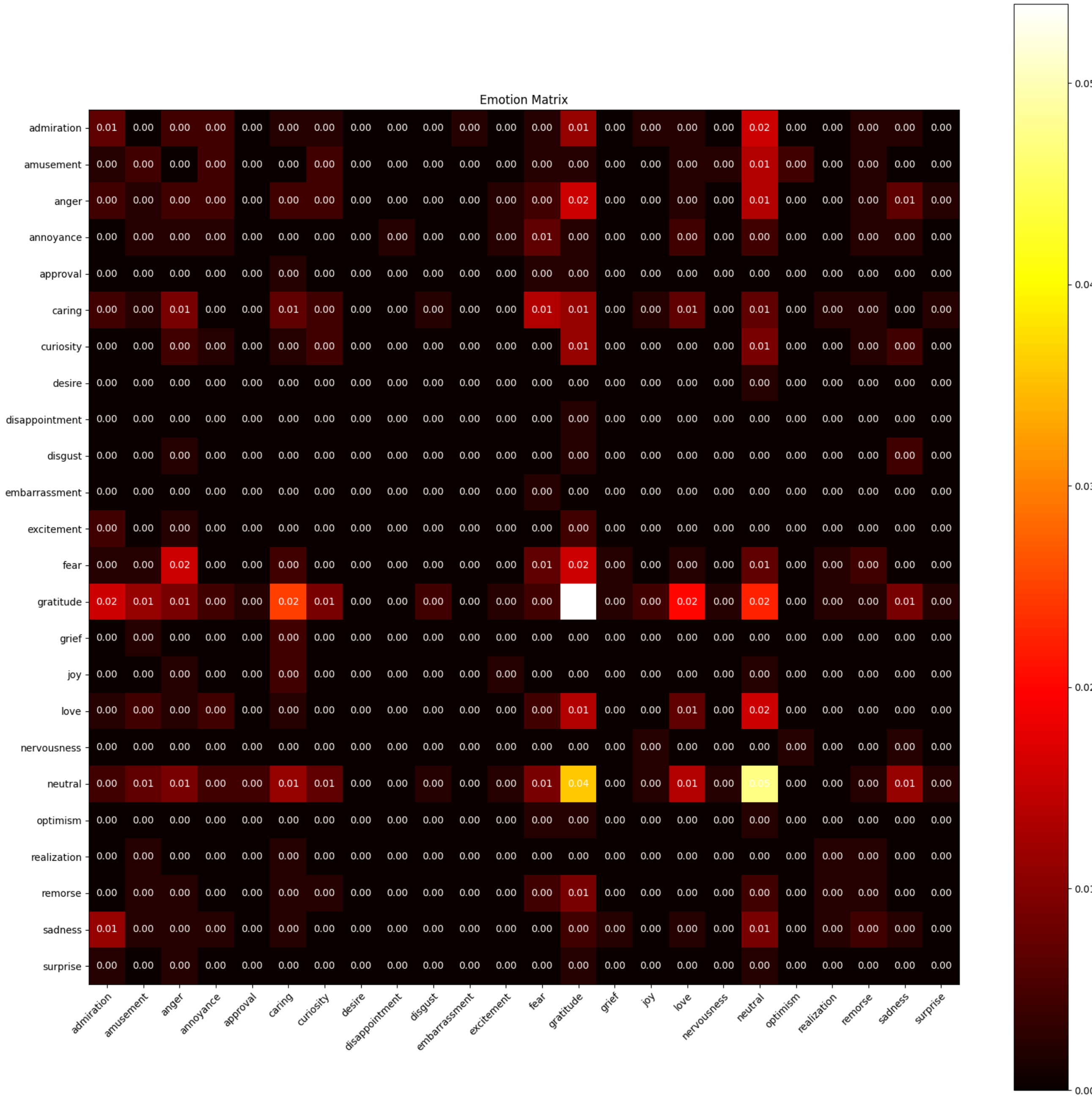


Fig:9b- 26×26 Heat map/Matrix of Romance genre(approach-1)

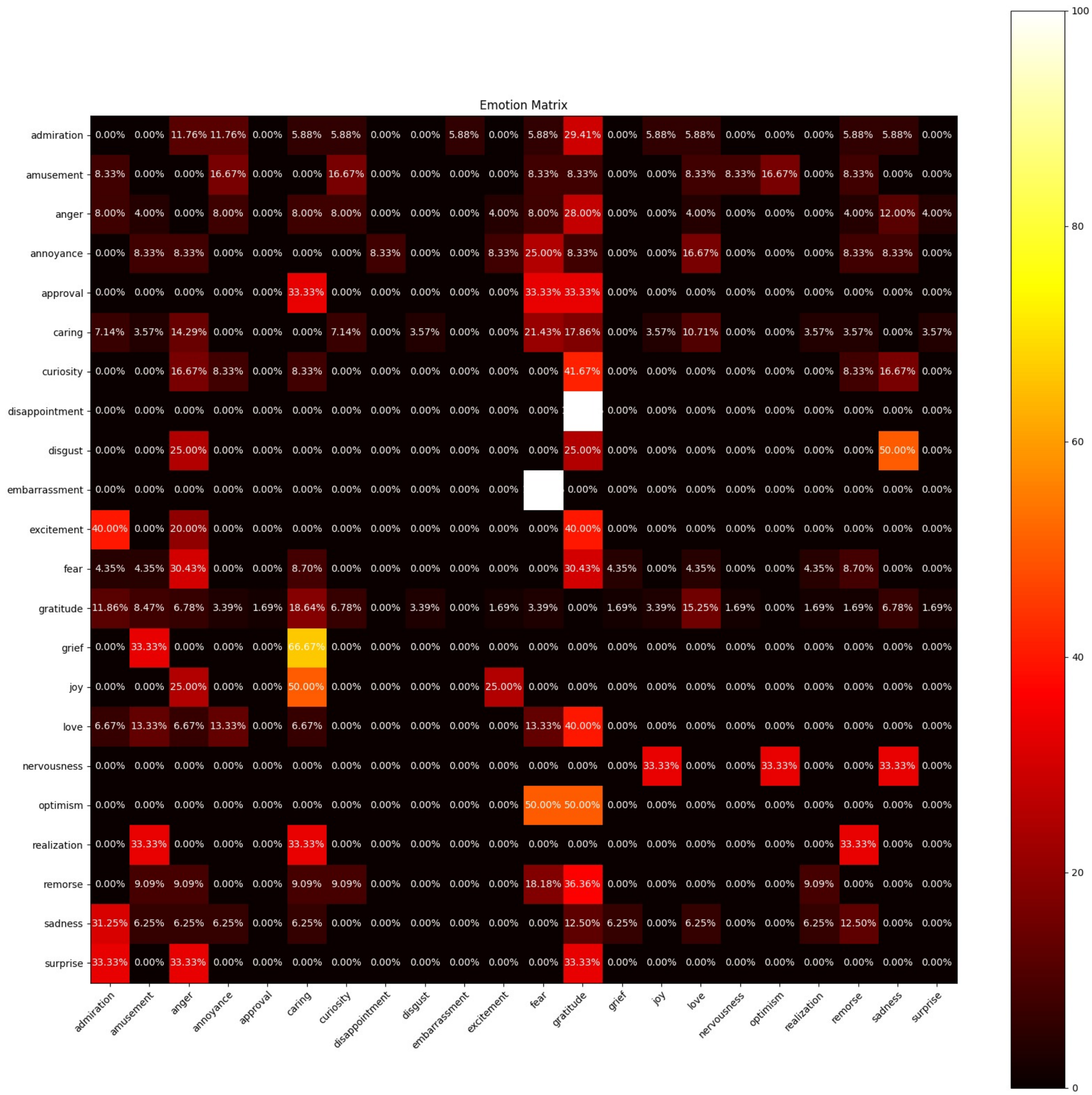


Fig:9c- 26×26 Heat map/ precentageMatrix of Romance genre(approach-1)

4. Analysis of Short Films

In this section, we manually tagged the emotions of 13 Indian short films and predicted emotions based on the characters' background music and facial expressions. We used a group discussion approach with colleagues to determine the emotions in each scene. We also used a model to predict emotions from the background music and another model to detect emotions from facial expressions. Overall, our analysis highlighted the differences between predicted and felt emotions in short films, emphasizing the importance of understanding audience emotions for evaluating a film's quality.

4.1 Data Collection and Preprocessing

Table-1 includes 13 rows, one for each short film we chose (we took [16] as a reference for choosing the short films). Each short film involves considerable emotional involvement of the audience. In Table-1, we gathered additional information about the films, including their genre, release year, runtime, director, IMDb rating, and rating from Shorted India[17]. To extract the background music, we utilized [18] and processed the original audio files.

Short Film	Genre	Release year	Runtime (in Minutes)	Director	SORTED rating	IMDb
Kheer	Romance	2017	7	Surya Balakrishnan	7.6	6.2
Interior Cafe Night	Drama, Romance	2014	13	Adhiraj Bose	7.3	7.6
Taandav	Comedy, Drama	2016	11	Devashish Makhija	7.8	7.3
Delivery Girl	Short	2019	12	Sreejone	7.3	0
Ahalya	Thriller	2015	14	Sujoy Ghosh	7.4	8
Kriti	Mystery, Thriller	2016	19	Shirish Kunder	7.8	7.2
Aai Shapat	Drama	2017	15	Gautam Vaze	6.8	7.1
Chhuri	Drama	2017	12	Mansi Jain	7.6	7
Chutney	Comedy, Drama	2016	17	Jyoti Kapur Das	7.6	7.7
Juice	Drama, Fantasy	2017	15	Neeraj Ghaywan	8	8.5
Pencil Box	Drama	2020	30	Amit Sanouria	6.7	0
Plus Minus	Drama	2018	18	Jyoti Kapur Das	7	8.6
The School Bag	Drama, Family	2017	16	Dheeraj Jindal	7.3	8.3

Table-1: Short films detail

4.2 Method/Our Approach

4.2.1 Manual Hand-Tagging of Emotions

For the manual hand-tagging of emotion task, we used nine Rasas of Natyashastra (Joy/Hasya, Wonder/Adhuta, Courage/Vira, Peace/Shanta, Sadness/Karuna, Anger/Raudra, Fear/Bhayanaka, and Disgust/Vibhasta) and 'Neutral.' We used a different approach than just going with one person survey. Instead, we went with the discussion approach, together with 5 to 7 of our colleagues, watched all 13 short films, and discussed several things. The first task of the group of viewers was to decide on when to segment a scene. The main bases of segmentation were when an emotion of the viewer starts to shift, or a whole together place and camera angles change. After segmenting the scene, the group discusses and decides on the emotion of the frame or scene. Here, everyone in the group gives their opinion on the overall emotion of the scene (not what they felt but what the characters felt). Furthermore, we tag the frame emotion of the scene as what the majority in the group agrees on. Similarly, everyone states what emotion they felt in the scene, and we tag the emotion of the scene with the emotion with which most of the group goes. Table-2 the results from group hand-tagging of the '*Interior Cafe Night (2014)*' short film.

StartTime	EndTime	FrameEmotion	Emotion
0:00:30	0:01:40	Wonder	Joy
0:01:40	0:02:00	Sadness	Wonder
0:02:00	0:02:36	Joy	Joy
0:02:36	0:04:10	Anger, Sadness	Sadness
0:04:10	0:06:53	Joy, Sadness	Joy
0:06:53	0:07:20	Anger	Love
0:07:20	0:08:00	Sadness	Sadness
0:08:00	0:09:20	Sadness, Love	Love, Sadness
0:09:20	0:10:35	Sadness, Anger	Sadness
0:10:35	0:11:56	Joy	Joy

Table-2: Emotion Hand-tagging of 'Interior Cafe Night (2014)' short film

4.2.2 Background Music (BGM) Emotion Prediction

After the hand-tagging, we started with the emotion prediction of background music. We use the same time frames that we segmented during the hand tagging. For emotion prediction, we used a model available on GitHub [19]. This model classifies the emotion of audio files into one emotion out of 'neutral', 'calm', 'happy', 'sad', 'angry', 'fearful', 'disgusted', and 'surprised'. This model is a deep learning classifier trained using two datasets, the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) [20] and the Toronto emotional speech set (TESS) dataset [21], and has an overall F1 score of 80% on eight classes. So, using this model, we predicted the emotion of each scene. In order to get table-3, We added one extra column of predicted emotion by background music to table-2. We created such a table for all 13 short films.

StartTime	EndTime	Frame Emotion	Emotion	Predicted Emotion
0:00:30	0:01:40	Wonder	Joy	disgust
0:01:40	0:02:00	Sadness	Wonder	angry
0:02:00	0:02:36	Joy	Joy	sad
0:02:36	0:04:10	Anger, Sadness	Sadness	calm
0:04:10	0:06:53	Joy, Sadness	Joy	calm
0:06:53	0:07:20	Anger	Love	disgust
0:07:20	0:08:00	Sadness	Sadness	sad
0:08:00	0:09:20	Sadness, Love	Love, Sadness	disgust
0:09:20	0:10:35	Sadness, Anger	Sadness	disgust
0:10:35	0:11:56	Joy	Joy	calm

Table-3: Emotion Prediction of 'Interior Cafe Night (2014)' short film

4.2.3 Facial Emotion Prediction from Video

The essential part of movie-making is the actor's ability to express emotions. It is important to study the facial expressions of the actors. Hence after adding the emotions from the background music, we began to work on predicting the emotions through the characters' expressions in a scene. We achieved the results by using an emotion detection function from DeepFace[24]. The model classifies facial expressions as anger, fear, neutral, sad, disgust, happy, and surprise. The model is trained with four million images uploaded by Facebook users. To arrive at an emotion, we checked frames every one-fourth of a second and collated them every 5 seconds. Then, the top two dominant emotions present during that time are saved. Further, in the following sub-section, we will compare these predicted emotions from facial expressions with predicted emotions from background music and with which emotions the audience felt. This can give good insights into a movie's quality. Then after getting the top two dominant emotions, we got those emotions for the exact time frames of scenes(as shown in table-3). We did this for all short films and added two more columns of top predicted emotions to an already existing table. Table-4 is the table after the facial emotion prediction of the 'Interior Cafe Night (2014)' short film.

StartTime	EndTime	Frame Emotion	Emotion	Predicted Emotion	Top Dominant Emotion 1	Top Dominant Emotion 2
0:00:30	0:01:40	Wonder	Joy	disgust	angry	sad
0:01:40	0:02:00	Sadness	Wonder	angry	happy	neutral
0:02:00	0:02:36	Joy	Joy	sad	angry	happy
0:02:36	0:04:10	Anger, Sadness	Sadness	calm	fear	sad
0:04:10	0:06:53	Joy, Sadness	Joy	calm	sad	neutral
0:06:53	0:07:20	Anger	Love	disgust	sad	fear
0:07:20	0:08:00	Sadness	Sadness	sad	angry	fear
0:08:00	0:09:20	Sadness, Love	Love, Sadness	disgust	happy	fear
0:09:20	0:10:35	Sadness, Anger	Sadness	disgust	sad	fear
0:10:35	0:11:56	Joy	Joy	calm	angry	fear

Table-4: Emotion Prediction of 'Interior Cafe Night (2014)' short film

4.3 Observations

4.3.1 Visualization of Emotion Occurrence

To analyze the data generated from the hand-tagging and emotion prediction, we began by visualizing the occurrence of each emotion in the three emotion categories: emotions felt by viewers, emotions predicted by background music, and emotions predicted by facial features. This visualization aimed to identify dominant emotions and identify any gaps or connections between the predicted emotions and the emotions experienced by the audience. We created bar charts of occurrence frequency versus emotion for each category in each short film. Figure 10 illustrates the bar charts for the 'Chhuri' short film as an example. The analysis revealed that while the audience felt 'joy,' the dominant emotions predicted by DeepFace were 'neutral' and 'fear,' indicating a gap between the predicted emotions and the audience's emotions. Similar gaps were observed in most of the short films, where the dominant emotion portrayed in the frames differed from the emotions felt by the viewers. Additionally, we observed significant differences between the predicted emotions by background music and the emotions felt by the audience.

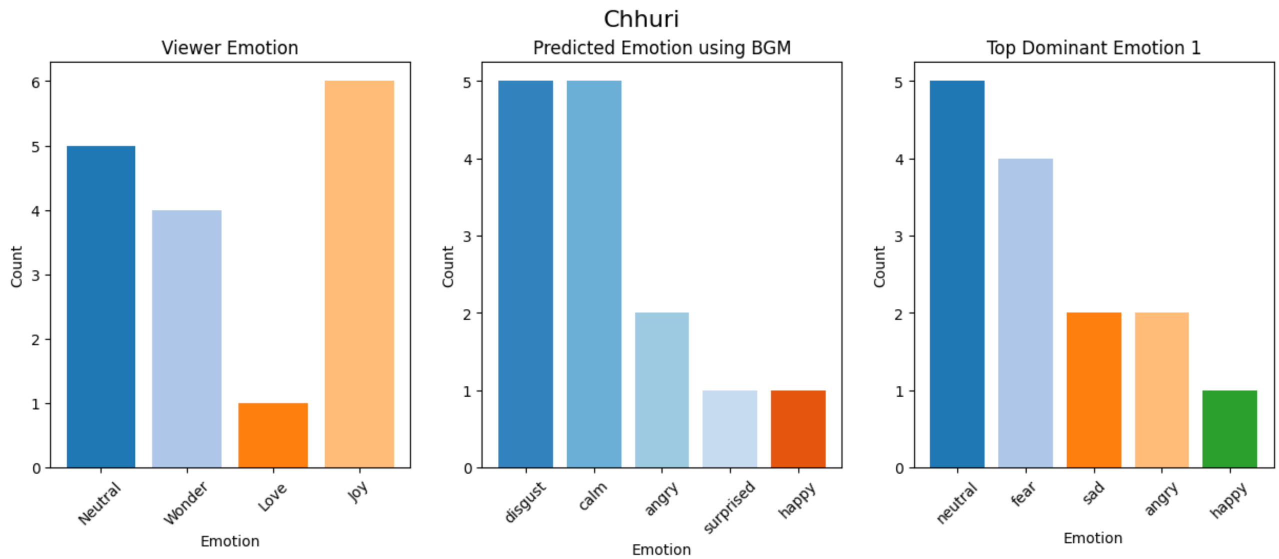


Fig:10 Analysis of emotions of 'Chhuri (2017)' short film

4.3.2 Sequences of Emotion

To further analyze the sequences of emotions, we counted the occurrence of emotion pairs for each short film. Each count pair ('x,' 'y') represented the frequency of emotion 'y' following emotion 'x.' This analysis was performed for all three emotion paradigms: emotions felt by viewers, emotions predicted by background music, and emotions predicted by facial features. These emotion pairs allowed us to calculate the probability of emotion 'y' following emotion 'x' and generate emotion matrices for each paradigm, as described in the previous section. We analyzed the gaps and patterns within the three paradigms by examining these matrices.

4.3.3 Emotion Matrices

We constructed heat-map matrices using the emotion pair counts from all 13 short films. Figure 11 displays the three emotion matrices: the background music emotion matrix (Figure 11a), the DeepFace emotion matrix (Figure 11b), and the viewers' emotion matrix (Figure 11c). The background music matrix (Figure 11a) revealed a significant occurrence of the 'calm-calm' emotion pair, likely due to the selection of short films focused on life and learning themes without violent or intense background music. The DeepFace emotion matrix (Figure 11b) showed a more uniform distribution, indicating that characters displayed various emotions across the short films. 'Sadness' emerged as a dominant emotion, potentially due to the similarity between 'anger' and 'sadness,' making their distinction ambiguous. In the viewers' emotion matrix (Figure 11c), 'joy' and 'wonder' emerged as dominant emotions. The dominance of 'wonder' can be attributed to the endings of the short films, as the endings must be gripping and unique to captivate the audience's appreciation. Therefore, the audience often experienced a sense of 'wonder' at the conclusions of the short films, leading to their overall appreciation.

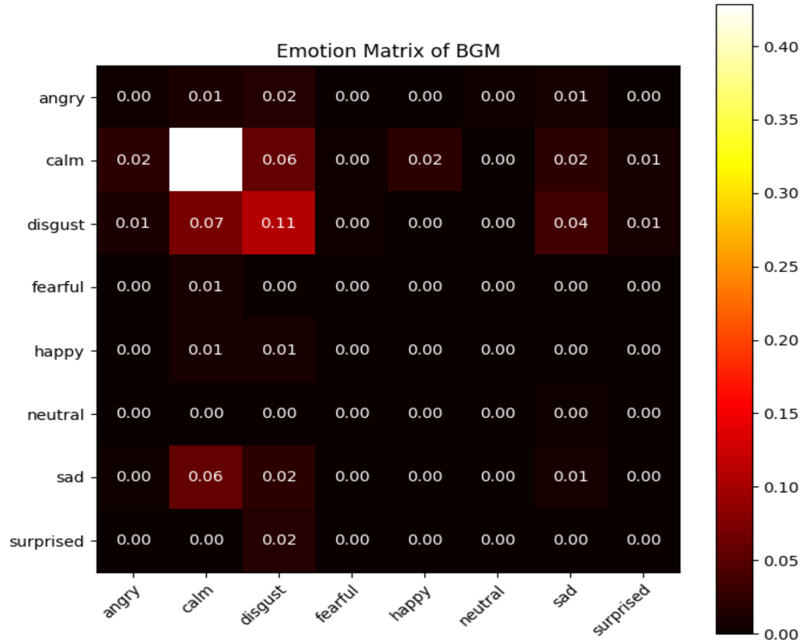


Fig:11a Background music Emotion matrix of short films

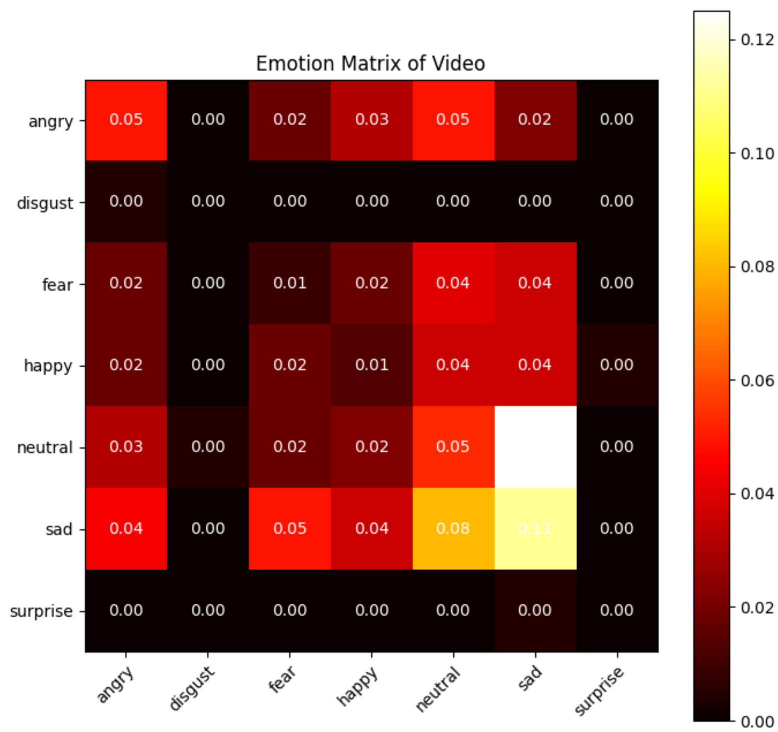


Fig:11b Deepface(video) Emotion matrix of short films

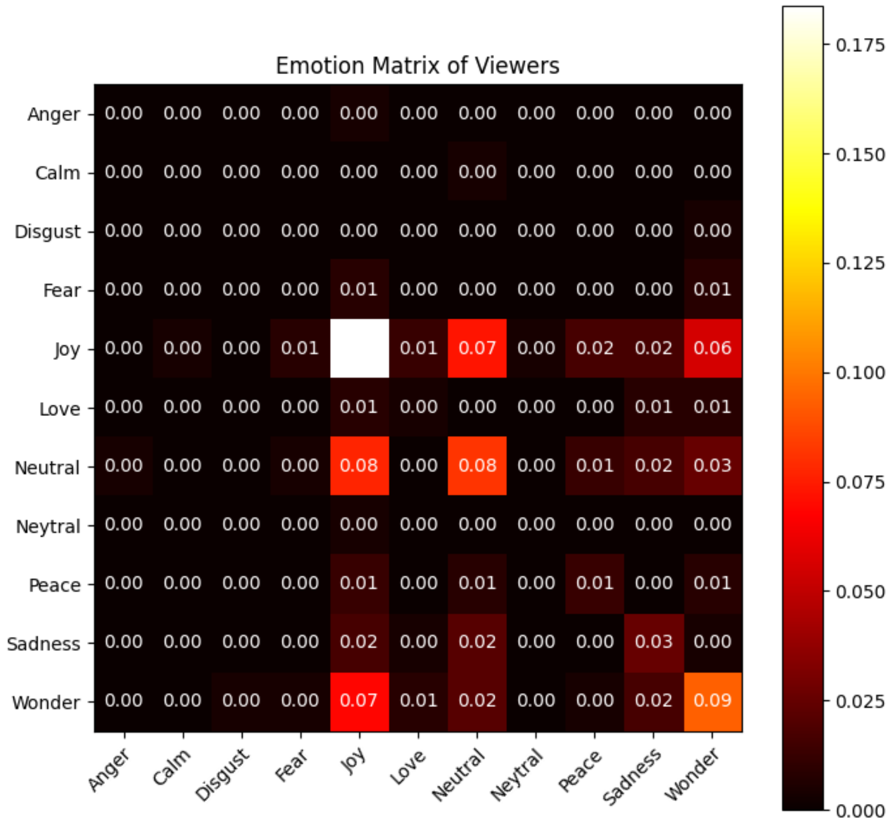


Fig:*11c Viewers Emotion matrix of short films*

Conclusion

In conclusion, our analysis of 90 movie subtitle files revealed that the emotion of "anger" is frequently used by movie creators in their storytelling. Furthermore, we observed that certain emotions are more prevalent in specific movie categories, with the genre of "Romance" standing out as distinct from others by not relying heavily on the emotion of "anger." On the other hand, the genre of "Fantasy" exhibited a more balanced distribution of emotions, adapting to various emotional storytelling approaches.

In our analysis of short films, we found that viewers often experience a sense of wonder and positive arousal at the endings of these films. However, we also identified a discrepancy between the characters' emotions and the emotions the audience felt. Through analyzing background music and facial expressions, we discovered that short films feature a wide range of facial expressions by characters. With the help of matrices, we could conclude that characters go through all sorts of facial expressions in short films.

Overall, our research sheds light on the emotional aspects of movies and their impact on viewers. By understanding the dynamics of emotions in films, filmmakers can refine their storytelling techniques, and researchers can deepen their exploration of the emotional effects of movies.

Future Works

Integration of Other Modalities

There is potential for further analysis by integrating all three dimensions of subtitles, background music, and facial expressions using the same model. Integrating multiple modalities such as subtitles, background music, facial expressions, and even physiological data can offer a more comprehensive analysis of emotions in movies. Combining these modalities can provide a richer understanding of emotional experiences and their impact on viewers.

Viewer emotion analysis

Our study primarily focused on analyzing the emotions expressed by characters in movies. Future research could involve analyzing the emotions experienced by viewers while watching films. This can be achieved through various methods such as surveys, physiological measurements, or sentiment analysis of social media data related to the movies.

Cross-cultural analysis

Emotions and their expression can vary across cultures. Conducting a cross-cultural analysis of movies could shed light on cultural differences in emotional storytelling and audience responses. This could involve comparing movies from different regions or analyzing the emotional portrayals in international co-productions.

Emotional Analysis in streaming platforms

With the rise of streaming platforms, analyzing emotions in movies and TV shows available on these platforms can be an interesting avenue for future research. This could involve examining patterns of emotions in binge-watching behavior, the emotional impact of content recommendation algorithms, or the emotional dynamics of original content produced by streaming platforms.

Emotion-based recommendation systems

The findings from this research can be utilized to develop emotion-based recommendation systems for movies. By analyzing the emotional journey of films and understanding individual viewers' emotional preferences, personalized movie recommendations can be made to enhance the viewer's emotional engagement.

References

[1] Tan, Ed S. Emotion and the structure of narrative film : film as an emotion machine / Ed S. Tan. p. cm.

[2] A. Ortony, G. L. Clore, and A. Collins, The cognitive structure of emotions. Cambridge university press, 1990.

[3] R. Plutchik, “The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice,” American scientist, vol. 89, no. 4, pp. 344–350, 2001.

[4] Philippe, Aurier & Guintcheva, Guergana. (2015). The Dynamics of Emotions in Movie Consumption: A Spectator-Centred Approach. International Journal of Arts management. 17. 5.

[5] S. Mohammad and P. Turney. Crowdsourcing a word-emotion association lexicon.Computational Intelligence,29(3):436–465,2013

[6] Kayhani, Amirkazem & Meziane, Farid & Chiky, Raja. (2020). Movies Emotional Analysis Using Textual Contents. 10.1007/978-3-030-51310-8_19.

[7] Andrew Salway and Mike Graham. 2003. Extracting information about emotions in films. In Proceedings of the eleventh ACM international conference on Multimedia (MULTIMEDIA '03). Association for Computing Machinery, New York, NY, USA, 299–302. https://doi.org/10.1145/957013.957076

[8] S. Zad, M. Heidari, J. H. J. Jones and O. Uzuner, "Emotion Detection of Textual Data: An Interdisciplinary Survey," 2021 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2021, pp. 0255-0261, doi: 10.1109/AIIoT52608.2021.9454192.

[9] A. Chiorrini, C. Diamantini, A. Mircoli, and D. Potena, “Emotion and sentiment analysis of tweets using bert,” 2021.

[10] Sheng-Yeh Chen, Chao-Chun Hsu, Chuan-Chun Kuo, Ting-Hao, Huang, & Lun-Wei Ku. (2018). EmotionLines: An Emotion Corpus of Multi-Party Conversations.

[11] van Kooten, P. (2022). contractions · PyPI. PyPI. https://pypi.org/project/contractions/#description

[12] Follows, S. (2019, February 4). *Defining the average screenplay, via data on 12,000+ scripts*. Stephen Follows. https://stephenfollows.com/what-the-average-screenplay-contains/

[13] Huey, N. (2021, February 22). *How Much Time Does it Take to Shoot a Typical Scene?* Beverly Boy Productions. https://beverlyboy.com/filmmaking/how-much-time-does-it-take-to-shoot-a-typical-scene/

[14] Ghoshal, A. (2023, June 1). *arpanghoshal/EmoRoBERTa · Hugging Face*. Hugging Face. https://huggingface.co/arpanghoshal/EmoRoBERTa

[15] Demszky, Dorottya & Movshovitz-Attias, Dana & Ko, Jeongwoo & Cowen, Alan & Nemade, Gaurav & Ravi, Sujith. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. 4040-4054. 10.18653/v1/2020.acl-main.372.

[16] Global Media Journal. (2020, June). A descriptive study of the Content Orientation of Selected Contemporary Indian Short Films. Global Media Journal, Indian Edition. https://gmj.manipal.edu/issues/june2020/8%20A%20descriptive%20study%20of%20the%20Content%20Orientation%20of%20Selected%20Contemporary%20Indian%20Short%20Films.pdf
[17] *Shorted India*. (2022). Shorted India – Watch handpicked short films from across India. https://shorted.in
[18] Vocali.se. (2022). Vocali.se. https://vocali.se/en
[19] Lal, P. (2021, Octobar 15). *purnima99/EmotionDetection*. GitHub. https://github.com/purnima99/EmotionDetection
[20] Livingstone SR, Russo FA (2018) The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PLoS ONE 13(5): e0196391. https://doi.org/10.1371/journal.pone.0196391
[21] Pichora-Fuller, M. Kathleen; Dupuis, Kate, 2020, "Toronto emotional speech set (TESS)", https://doi.org/10.5683/SP2/E8H2MF , Scholars Portal Dataverse, V1
[22] https://www.opensubtitles.org/en/search/subs
[23] https://www.imdb.com/chart/top/
[24] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 1701-1708, doi: 10.1109/CVPR.2014.220.