EMOTIONALLY INTELLIGENT STORYTELLING THROUGH AI AND ML USING EMOTIONAL ARCS OF MOTION PICTURE THAT RESONATE WITH AUDIENCES

Research Paper

Dr. Prashanth Ramappa, Swiss School of Business Management, Geneva, Prashant@ssbm.ch

Abstract

Film and moviemaking production are an intricate process, mainly decision-making among its producers, Motion picture directors, and writers, especially for script selection. Movie or Filmmaking is a huge investment, it takes more of a resource of Capital, Manpower & Creativity. Traditional script evaluation methods: almost entirely subjective measurements; reliant on the vast and varied experiences of industry professionals (style) — thus inevitably creates bias disparities in decision making, often discounting new or unique narratives. Film producer should be clear whether the narrative is appropriate for the audience's typical emotional journey. Understanding emotions in scripts is a difficult and complicated task. Our Approach involved, combined the concept of Shapes of stories, Emotional Analysis using Natural Language Processing (wheel of emotion) and ML Clustering technique to solve challenge. Our Methodology involves the use of subtitle data to demystify the emotions in the story. We also apply NRCLex, a lexicon-based method, to associate with eight basic emotions and cluster similar movies within the six main story shapes. Our findings indicate that stories generally follow the six fundamental forms. We also found that movies in the "Man in the Hole" category where the hero faces an unexpected challenge, overcomes his fear, and succeeds in becoming a leader have the highest box office and popularity success.

Keywords: Emotional AI, Emotional Arc, Emotions, NLP, Machine learning, Shape of Stories, Movie Script.

1. Introduction

Technology and fresh storytelling have become an integral part of the trends in this constantly evolving entertainment landscape. It constitutes a precursor to the evaluation of a cutting-edge dataset with potential for ML and AI in affecting product-and-service innovation through movie emotional arc embedding. The confluence of these fields signals an exciting opportunity to address the critical challenges of film screenplay development and push the industry forward into new frontiers of creative innovation. Leaders in the domain of technology integration are to be seen at the peak with nascent performing arts industry-backed-in-the-throes of creative and technology driven sound insights ensuring leaps in success. Which leads me to suspect that while people have an ever-growing thirst for the strong and personal. The goal, setting, and applicability of emotions with technology are all disclosed in this part, which provides a thorough overview of the study (Del Vecchio et al., 2021)

One of the foundation papers, "Improving productivity in Hollywood with data science-Using emotional arcs of movies to drive product and service innovation in entertainment industries" by(Del Vecchio et al., 2021), serves as a motive for this study. Previous research (Reagan et al., 2016), made the preparation to understand emotional arcs and their effect. Through this research, in the future, we aim to learn more

about the methods to be able to make use of machine learning and artificial intelligence in the entertainment industry.

Storytelling is an aspect of the human experience and runs connection between so much of our social interactions, cultural traditions, and day-to-day living(Behrooz et al., 2015). According to. (Behrooz et al., 2015) research, People feel more connected to these intelligent agents, much as it does when share memories helps strengthen friendships between people. They tested their system by using a game and showing that it was able to pick out interesting stories to tell. They hope this will help make more sociable and relatable robots and virtual assistants in the future.

1.1. Challenges in Traditional Story Evaluation Methods.

- Film and moviemaking production are an intricate process, mainly decision-making among its producers, directors, and writers, especially for script selection.
- It's almost entirely subjective measurements: reliant on the vast and varied experiences of industry professionals (style) — thus inevitably creates bias disparities in decision making, often discounting new or unique narratives.
- Movie or Filmmaking is a huge investment, it takes more of a resource of Capital, Manpower & Creativity. Because scripts are a critical component of filmmaking, they are usually judged on a subjective level, meaning the selection process can often omit out-of-the-box storytelling.
- According to (Reagan et al., 2016),Emotional arcs which are important for engaging storytelling, are often not used enough. This is because traditional methods don't have a systematic and data-driven approach.

1.2. Concepts Integrated to Address Challenges

- According to Kurt Vonnegut (Comberg, 2010) by early 1950, he had this idea about the shape of stories. Here's how it would work: if you could map a story with the horizontal axis representing time and the vertical axis representing our characters' fortunes, you could see how it's structured over time.
- According to Plutchik (Plutchik, 1982) created a wheel of emotions according to Samurai, who explained different emotions. Plutchik first proposed his cone-shaped model 3D or the wheel model 2D back in 1980 to describe how the emotions were related.
- Connecting both the concepts with Artificial intelligence like Natural language processing by Using NRCLex(Mohammad, 2019) which will measure emotional affect from a body of text. Affect dictionary contains approximately 27,000 words and based on National Research Council Canada NRC and Machine learning Clustering Techniques Gaussian Mixture model, Kmeans and mini batch Kmeans used to Classify based on the shapes of Stories.

2. Research Objectives

The main purpose of this study is to use Natural Language Processing (NLP) and Machine Learning techniques to simplify emotional arc analysis in filmmaking.

1. To what extent will technology like AI, computational narratology, and NLP help reshape and interpret the emotional arc to allow in order help Business to get a most profitable motional Arc category in motion pictures.
2. How well a Lexicon based approach NRCLex will help to demystify emotional insights and analyze emotional arcs in movies to categorize to improve efficiency and productivity in script.

3. Methodology and Scope of Study

3.1. Literature Review

- The Wheel of Emotions, devised by Robert Plutchik(Plutchik, 1982), provides a range of human emotions. Emotions were majorly classified into eight categories and by word and emotions associated with it. Plutchik has acknowledged the multi-dimensionality of emotions and their potential influence on behavior beyond mere rationality.
- AI and machine learning-based emotion analytics could be useful in personalizing their emotional arcs while connecting to audience tastes, hence enhancing user engagement, and fueling service innovation in filmmakers (Campbell, n.d.; Khanna & Malviya, 2020; Nader et al., 2024).
- Deep learning algorithms can recognize facial expressions, speech tones, and physiological responses; this helps filmmakers make emotionally enriching situations(Dashtipour et al., 2021; Zhou et al., 2019)
- To improve resource allocation and operational efficiency (Green & Brock, 2000), recommendation algorithms(Covington et al., 2016) that are driven by emotional computing and cognitive systems can change user emotions and predict consumer behavior (Taylor et al., 2015; Tzavella, 2020)
- Use of Algorithmic bias, Adaptation of Script writers and AI Ethical data privacy issue to help protect customer privacy data and implications.(Glymour et al., 2018)

3.2. Research Methodology and Workflow

3.2.1. Qualitative Approach- Highlights

- Scripts and subtitle texts from IMDB are refined and analyzed Based on context analysis findings, applied normalization, special character removal etc.

3.2.2. Quantitative Approach- Highlights

- The IMDB Kaggle dataset, contains around 4500 movies' subtitles and reviews (from 1915 to 2017) and after filter for original language to English remains 3500 movies for Analysis, serves as the primary source for the experimental analysis.
- De-duplicates and redundant features, followed by a joining of movie subtitles and review data, based on IMDB ID.
- Qualitative Approach Includes Subscripts cleaning, removing special characters, Analyzing review text and subscript texts.
- Kmeans, Gaussian Mixture model and Minibatch Kmeans Clustering is applied for all the movies for Six Cluster. Created a Custom methodology to combine all 3-cluster combination to assign emotional Arc based on combination.
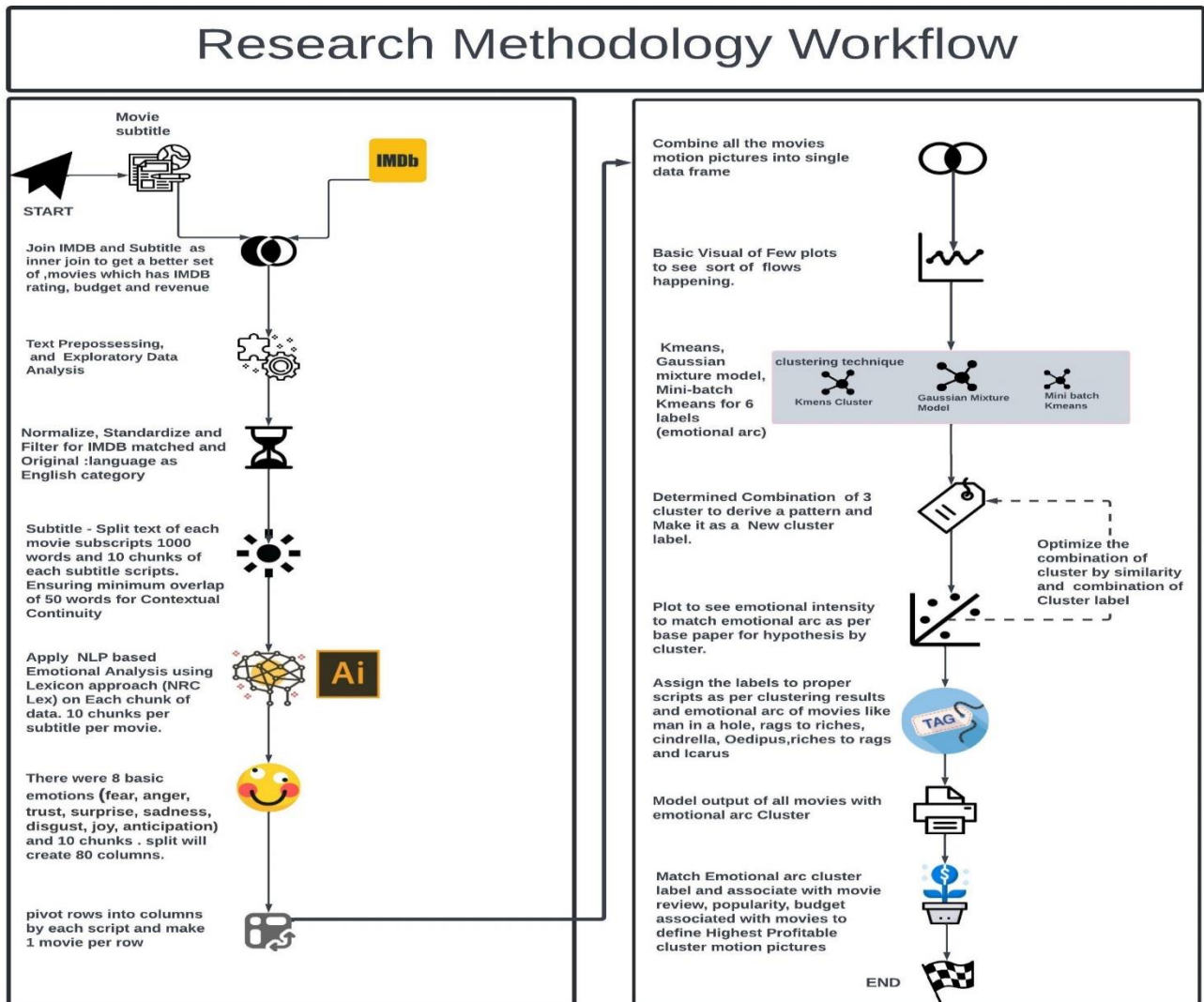
3.3. Proposed Workflow

Figure 1. Research Workflow model.

3.3.1. Proposed Workflow details

- Only English language movies is filtered by original language IMDB column which consists approximately 3,500 films and have strong data subset.
- Subtitles are chunked for minimum 10,000-word with overlap of 25 words from each chunk and split it to 10 chunks. If more than 10,000 words (after removing stop words) Will be looped to get always exact 10 chunks are one of key methodology approach used in the next steps
- The outcome is very structured for each film subtitles (10 distinct chunks) time-stamped at every 10 minutes to create an analysis more emotional fine-grained.
- Each of these chunks were ran through Lexicon based NLP model provided by NRCLex which is a very useful advanced library that can extract and quantify emotional Probabilities by 8 basic

emotions (anger, fear trustier surprise sadness joy disgust, anticipation) and respective probabilities.

- Emotional Probabilities of all the 10 chunks will be pivoted and made it as columns to rows like 10 chunks with 8 probabilities will make 80 column per movie or IMDB_ID. Get all the movies processed and create one DataFrame.
- Run the Visualization samples to check the flow of each chunk whether it shows the variations by chunk.
- Apply 3 different clustering Algorithms like K Kmeans, Gaussian Mixture model and Minibatch Kmeans and did Experimental Design to Combine as a Label like Kmean-1, GMM-3 and mini batch means - 6 (136- Probably falls into Cinderella cluster) and verified based on the hypothesis of movies wrt to base paper to assign emotional Arc like Rags to Riches, Riches to Rags, Cinderella, Icarus, Man in a hole and Oedipus.

3.4. Exploratory Data Analysis

- Figure 2: Average vote of all movies and figure 3: Most profitable movies by Imdb profit percentage show that voting average distribution as 6 to 8 has 68% movies fall under umbrella of good rating.
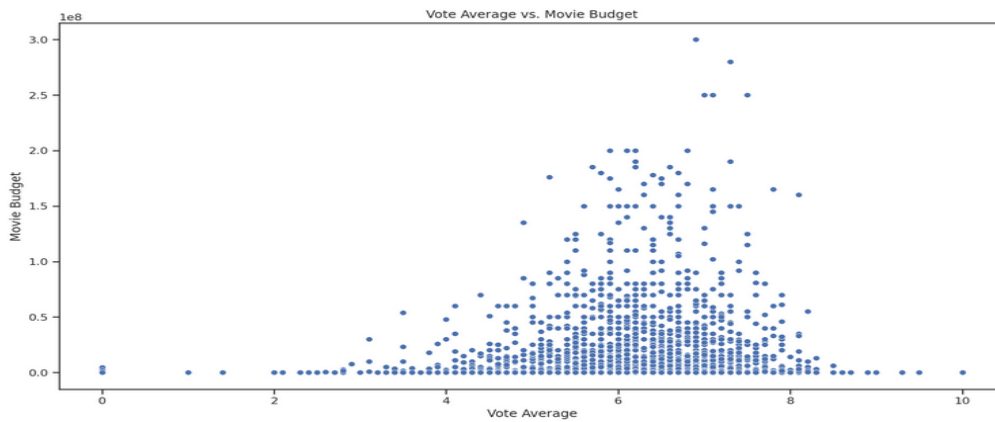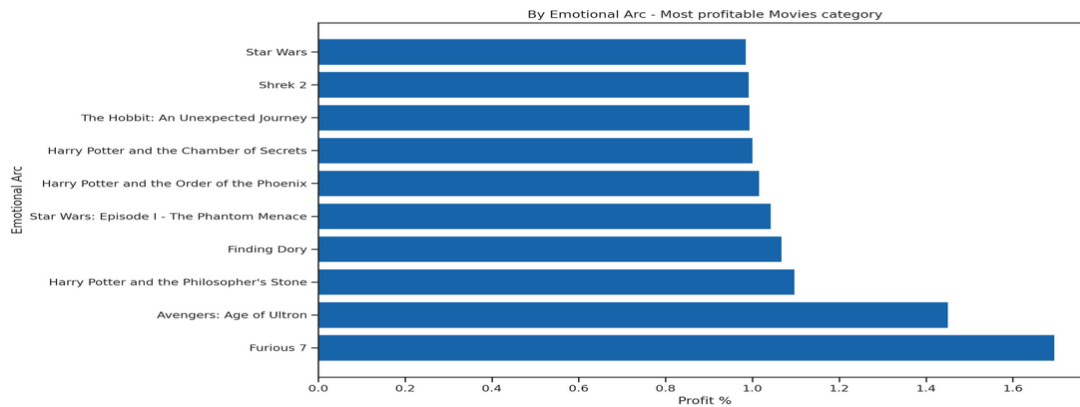

Figure 2. Average vote all movies

Figure 3. Top10 Imdb movies by profit

4. Results – Key Observations for Q2 Research Objectives

4.1. Data Analysis and Results – Emotional Arc- Rags to Riches (RISE)
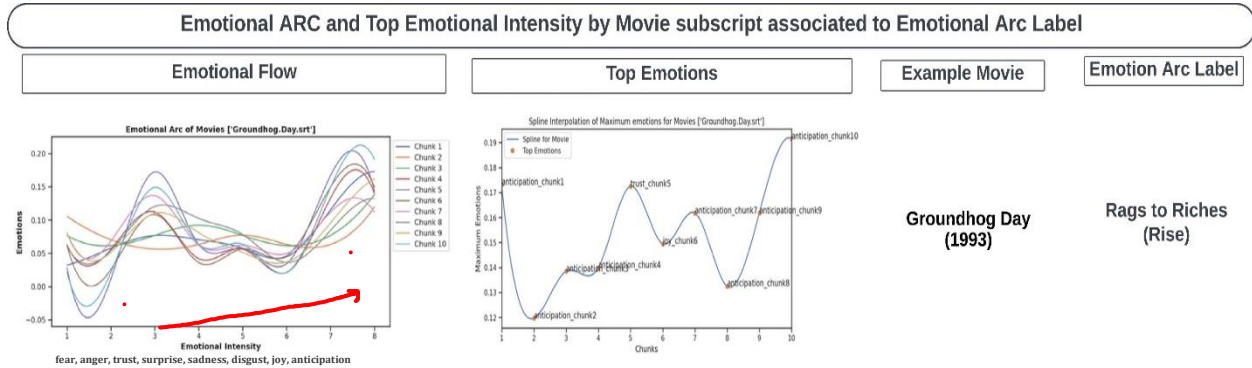

Figure 4. Rags to Riches (RISE)

❖ **Movie Summary:**

Groundhog Day" is a comedy about a self-centered weatherman, Phil Connors, stuck in a time warp that makes him relive the same day repeatedly. However, while learning to use this time loop to his full advantage by developing new skills and helping other people, improving himself as a person, he finally works his way through to the happiness of love after all, truly exemplifying the rise from 'rags to riches'.

❖ **Results Summary:**

- In the movie "Groundhog Day," shares a "Rags to Riches" storyline, positive emotions feelings characterized by trust, anticipation in most of segments. Also patter of flow show the Rise in mist of the chunks where it starts at lower emotions and flow ends in higher positive emotion.
- Emotional "Rags to Riches" arcs typically conclude with the spike of strong positive emotions, in this case high closer on anticipation.

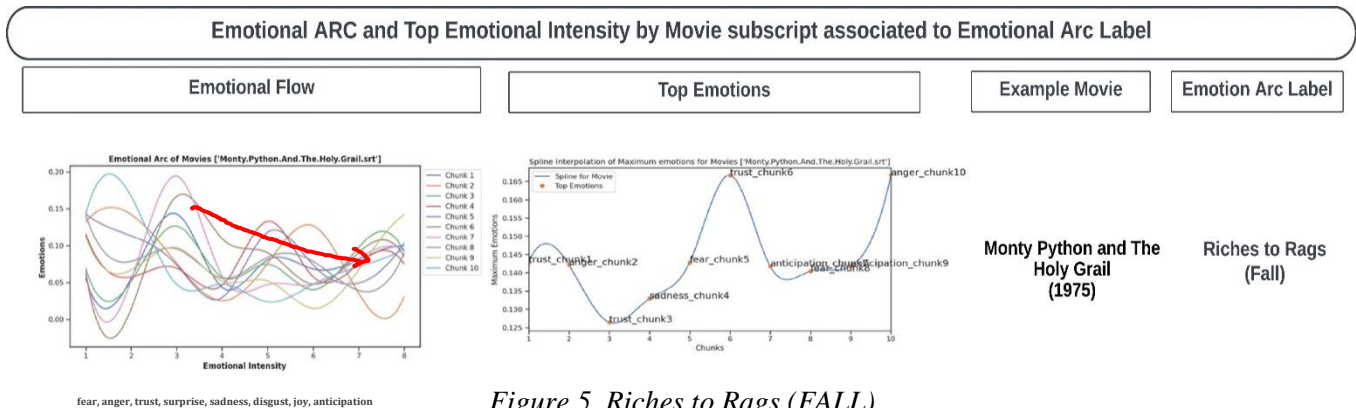## 4.2. Data Analysis and Results – Emotional Arc- Riches to Rags (FALL)

Figure 5. Riches to Rags (FALL)

About Movie Summary:

❖ **Movie Summary:**

"Monty Python and the Holy Grail" in this movie King Arthur and his knights quenching into quest for the Holy Grail but who get frequently gets ridiculed due to mishaps that separated them from their Status of King's men. They Tried all their adventures and solves riddles, obstacles, and disappointments that they forgot why they have come and lost their Rich life to survival mode, and this convey a story line of "Riches to Rags" story.

❖ **Results Summary:**

- This Motion picture depicts "Riches to Rags" narrative, as negative emotions such as fear and anger are predominantly present in most sections.
- Typically, a "Riches to Rags" emotional arc begins with low-intensity positive or negative feelings and concludes with a high-intensity negative emotion, which, in this case, is anger as end emotion. The emotional Flow show the downward pattern from where story started, will help us to classify the Riches to Rags. Riches to Rags and Rags to Riches are opposite to each other in emotional flow.

4.3. Data Analysis and Results – Emotional Arc- Man in a hole (RISE-FALL)
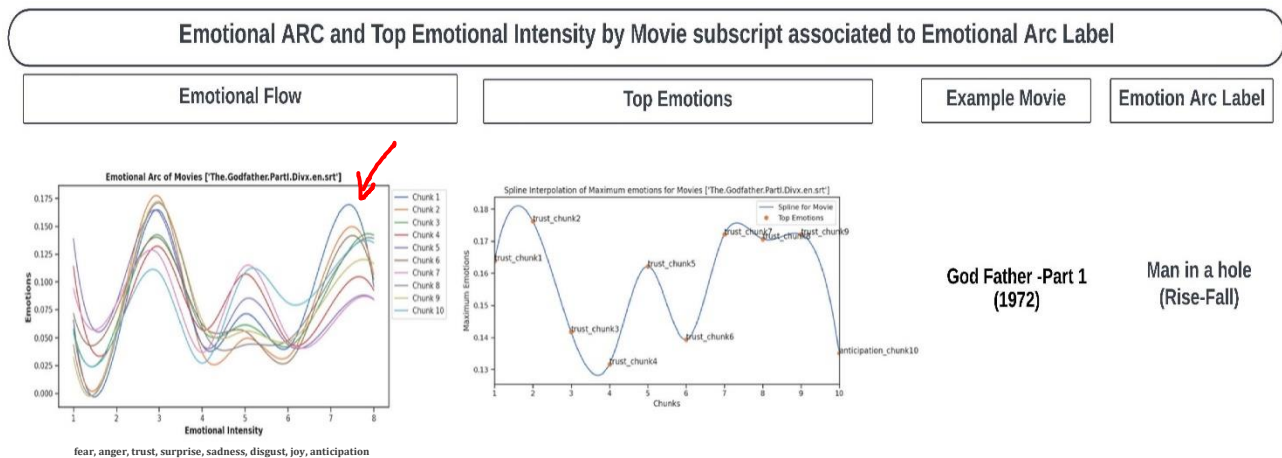

Figure 6. Man in a hole (RISE-FALL)

❖ **Movie Summary:**

"The Godfather Part I" is a crime drama on the Corleone family, an influential Italian American Mafia dynasty that shows how Michael Corleone transforms from being an outsider to a ruthless leader in the family business after his father, Don Vito Corleone, has been shot.

❖ **Results Summary:**

- Script utilizes the "Man in a hole" storyline in the movie "God Father - Part 1," beginning with a strong positive emotional arc, suddenly going downhill mid-story, and ends with another impact blow of low positive emotions.
- Man in a hole emotional trajectory: It seems that audiences are particularly engaged by stories where a protagonist's otherwise happy journey suddenly takes a turn into negative emotions but then ends in a positive note. Man in hole and Icarus are on opposite ends of the Emotional flow. Impact of Story line shows Protagonist Grows up and at the end of movie there is big blow where attack happens on him, and he loses his dear family members and takes a revenge (where emotions will dip).

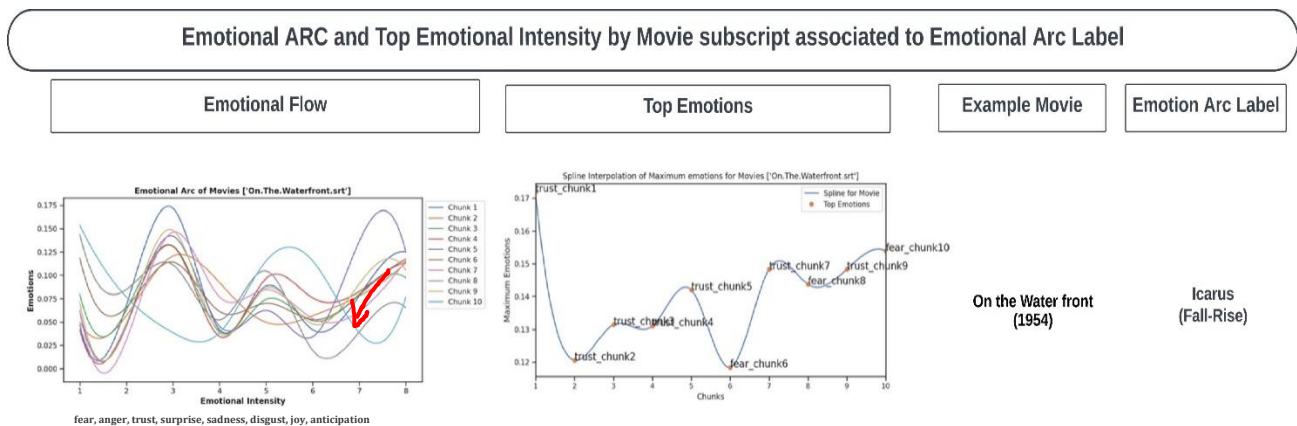## 4.4. Data Analysis and Results – Emotional Arc- Icarus (FALL-RISE)

Figure 7. Icarus (FALL-RISE)

❖ **About Movie Summary:**

"On the Waterfront" is a drama. About The film's protagonist, Terry Malloy, who begins as a promising dock worker, soon gets cusped in to corrupt activities of unions, only to experience an Icarus-like moral growth, followed by a fall from his untouched status but an eventual redemption and fight for justice, symbolized as a hopeful rise—the hallmark of the "Icarus" narrative.

❖ **Results Summary:**

- This film "on the waterfront" that encapsulates the "Icarus" story is a possible emotional roller coaster.
- One inference would be that watching a movie takes less time than reading a book. One possible explanation as to why movies on Icarus don't do as well is that people may not like to have strong negative emotions in such a short period of time unless the positive emotional lift is correspondingly large.
- its showing in our plot above and with Last chunk (chunk-10) shows the Fall-Rise- pattern to justify Icarus Shape of story category. Icarus and Man in a hole are opposite to each other in Emotional flow.

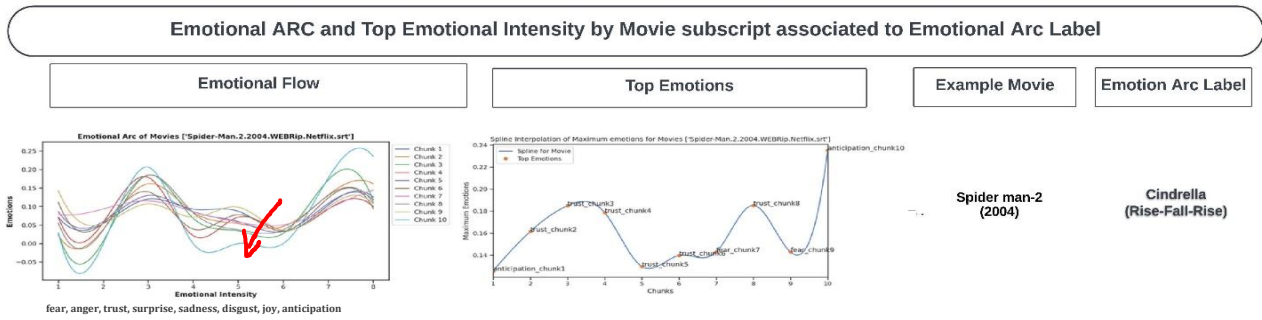## 4.5. Data Analysis and Results – Emotional Arc- Cinderella (RISE-FALL-RISE)

Figure 8. Cinderella (RISE-FALL-RISE)

❖ **Movie Summary:**

"Spider-Man 2" wanted to live a life of Peter Parker, but he struggles with his personal life, love life with Mary Jane, as great power comes with great responsibility as Spider-Man. Genius scientist Dr. Otto Octavius becomes the villain Dr Ock When his experiment fails. Spider-Man must come for rescue of New York City from getting destroyed. During the journey Peter Parker learns life lessons about great power comes with great responsibility and sacrifice.

❖ **Results Summary:**

- The movie "Spiderman 2" which has similarity of "Cinderella" narrative category, starts with high positive emotion (trust), and drops back to low intensity emotion (trust), and concludes with high end positive emotion (anticipation).
- This Flow has emotional rollercoaster of human being struggling inside to make decision and how he understand his powers. This roller coaster ride and Cinderella category will make movie more Interesting to viewers.

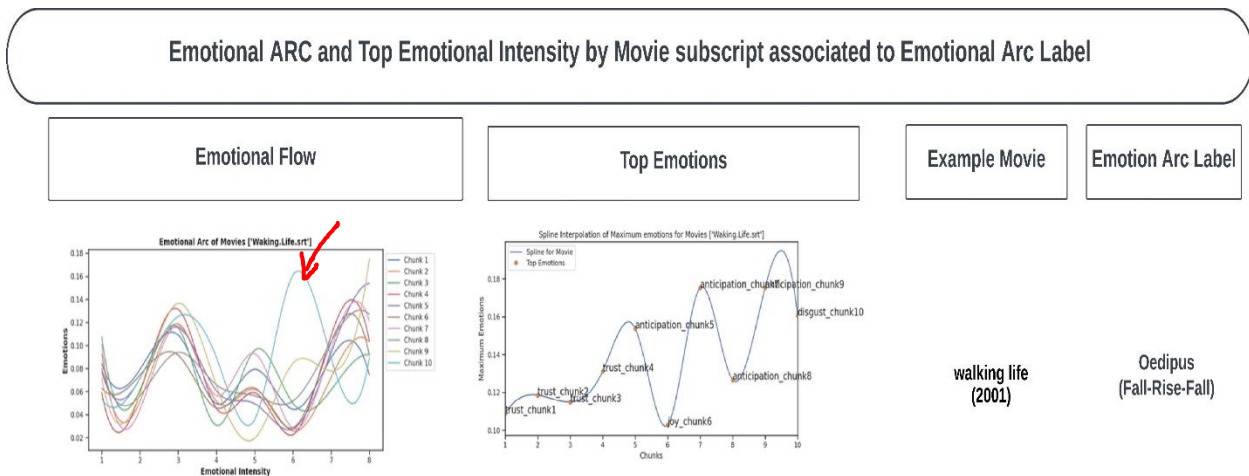## 4.6. Data Analysis and Results – Emotional Arc- Oedipus (FALL-RISE)

Figure 9. Oedipus (FALL-RISE)

❖ **About Movie Summary**:

In "Walking Life", As the protagonist goes deeper into his dreams and, at last, experiences lucid dreaming, and comes back to reality of realization that is somewhat like Oedipus's story of who he is. Both understandings shift one's viewpoint, like the tragic self-realization and ensuing fall of the "Oedipus" storyline.

❖ **Results Summary:**

- The "Walking Life" tier, including the "Oedipus" narrative, begins with trust and progresses to anticipation before ending with disgust. It concludes with a significant negative emotion.
- In Oedipus category movies, it's all about when protagonist understand reality and suggests that the ending might be unexpected or unfavorable, potentially creating a feeling of discomfort or dissatisfaction among the audience.
- its showing in our plot above and with Last chunk (chunk-10) shows the Fall-Rise-Fall pattern to justify Oedipus Shape of story category.

5. Results – Key Observations for Q1 Research Objectives

5.1. Final Output of Emotional Arc of Movies

In the below Figure 10: "Emotion Arc by counts" plot shows that man in hole movie has been the most categorized one and Riches to Rags were the least. And Figure 11: "Top 10 most profitable movies in Emotional Arc" more than 50% of high-income profit movies are in "man in a hole" category and Next best is the with rags to riches which constitute of 20% of overall movie distribution.
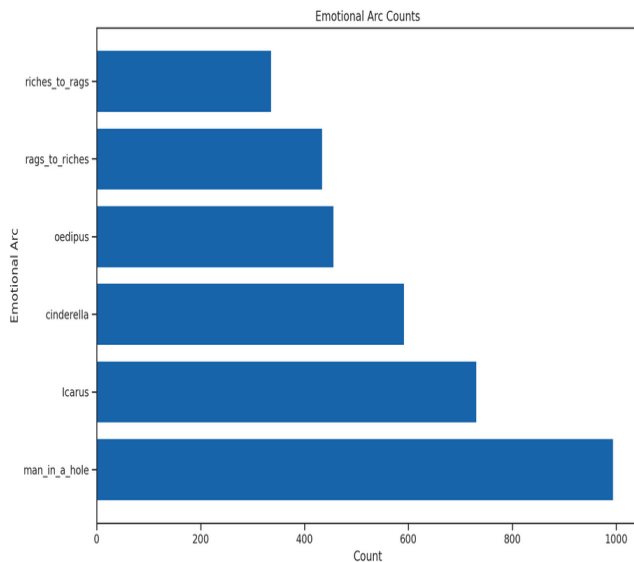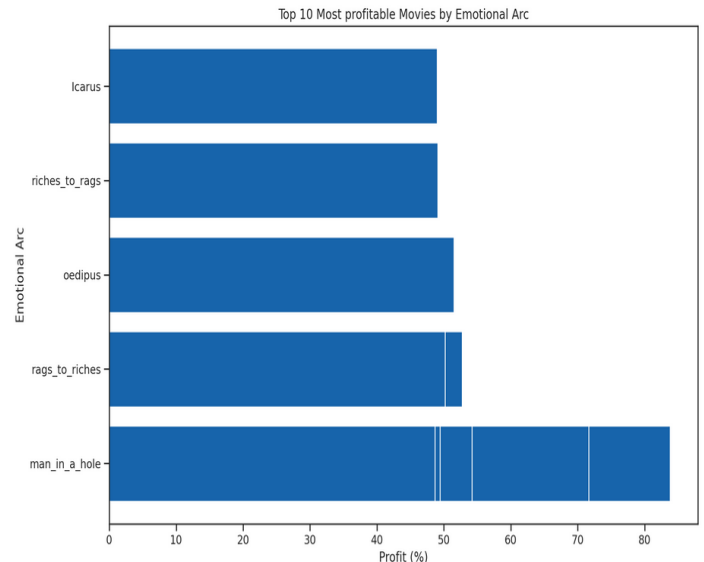

Figure 10. Emotion Arc by counts

Figure 11. Top 10 profitable movies emotion arc

5.2. Summary:

- In a recent study, we conducted Natural Language Processing methodologies, using the NRCLex package, to understand how emotions modulate consumer preferences for media and entertainment information.
- We observed thousands of subtitles dominated by six distinct emotional trajectories: Rags to Riches, Riches to Rags, Man in a Hole, Icarus, Cinderella, and Oedipus.
- Our methodology shows commercial enterprise associated with the 'Man in a Hole' emotional trajectory. That insight could mean strategic decisions in script selection, development, and production resources towards content that has better shot on success. From that end, this could mean much-improved bottom line performance and market competitiveness in businesses dealing in the entertainment sector.
- The consideration of emotional arcs in the scripts enables screenwriters to come up with narratives that will have a better connection with the audience and achieve greater viewer engagement and satisfaction.

6. Conclusion- Implication and Recommendation

❖ **Implication (discoveries)**

- With a 'Man in the Hole' story, the chances of success are higher, but any story can succeed if the budget is proper.
- Selecting the right script, use of data science, and combining that information will help increase profits and better connect with audiences.
- Sentiment Analysis will help in letting the audience preferences guide the choice of movie selection and can alter the way movies are made today.

❖ **Recommendation**

- With data science and AI, as was the case in Netflix, it boosts profit, outperforming traditional studios.
- Betting on sure hits like the "Man in the Hole" movie type balances risk reduction and probability of reward in content creation.
- Continuously improve our clustering and Emotional Analysis methods to have a better way to measure creativity and productivity in future studies.

References

Behrooz, M., Swanson, R., & Jhala, A. (2015). Remember that time? Telling interesting stories from past interactions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9445*, 93–104. https://doi.org/10.1007/978-3-319-27036-4_9

Campbell, J. (n.d.). *The Hero With A Thousand Faces, Commemorative Edition (Princeton University Press; 2004)*.

Comberg, D. (2010). Kurt Vonnegut on the Shapes of Stories. In *YouTube*.

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for youtube recommendations. *RecSys 2016 - Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198. https://doi.org/10.1145/2959100.2959190

Dashtipour, K., Gogate, M., Adeel, A., Larijani, H., & Hussain, A. (2021). Sentiment analysis of persian movie reviews using deep learning. *Entropy*, *23*(5). https://doi.org/10.3390/e23050596

Del Vecchio, M., Kharlamov, A., Parry, G., & Pogrebna, G. (2021). Improving productivity in Hollywood with data science: Using emotional arcs of movies to drive product and service innovation in entertainment industries. *Journal of the Operational Research Society*, *72*(5), 1110–1137. https://doi.org/10.1080/01605682.2019.1705194

Glymour, B., fairness, J. H.-P. of the conference on, & 2019, undefined. (2018). Measuring the biases that matter: The ethical and casual foundations for measures of fairness in algorithms. *Dl.Acm.Org*. https://dl.acm.org/doi/abs/10.1145/3287560.3287573

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721. https://doi.org/10.1037/0022-3514.79.5.701

Khanna, P., & Malviya, M. (2020). *Identification of Success Factors for Mirzapur Web Series*. *October*. https://doi.org/10.34293/management.v8i2.3351

Mohammad, S. M. (2019). Word affect intensities. *LREC 2018 - 11th International Conference on Language Resources and Evaluation*.

Nader, K., Toprac, P., Scott, S., & Baker, S. (2024). Public understanding of artificial intelligence through entertainment media. *AI and Society*, *39*(2), 713–726. https://doi.org/10.1007/s00146-022-01427-w

Plutchik, R. (1982). A psychoevolutionary theory of emotions. *Social Science Information*, *21*(4–5), 529–553. https://doi.org/10.1177/053901882021004003/ASSET/053901882021004003.FP.PNG_V03

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, *5*(1). https://doi.org/10.1140/epjds/s13688-016-0093-1

Taylor, P., Hobbs, J. N., Burroni, J., & Siegelmann, H. T. (2015). The global landscape of cognition : hierarchical aggregation as an organizational principle of human cortical networks and functions. *Nature Publishing Group*, *June*, 1–18. https://doi.org/10.1038/srep18112

Tzavella, L. (2020). *Behavioural measures and training interventions for food-related cognition , motivation and affect Behavioural measures and training interventions for food-related cognition , motivation and affect PhD Thesis by Loukia Tzavella In Partial Fulfillment of t*. *January*.

Zhou, Y., Zhang, L., & Yi, Z. (2019). Predicting movie box-office revenues using deep neural networks. *Neural Computing and Applications*, *31*(6), 1855–1865. https://doi.org/10.1007/s00521-017-3162-x