Movie Emotion Analysis: An Interactive Visualization Platform

Team008: Xinyuan Wang, Xinjie Zhu, Yipu Chen, Shiyao Shen, Zihao Xiao

1 INTRODUCTION

Movies evoke diverse emotions that vary across individuals and cultures. Our project develops an interactive visualization platform allowing users to classify globally popular films into five emotion categories based on personal perceptions while concurrently comparing these subjective ratings with multiple algorithmic analyses: extracting sentiment from audience reviews, deriving emotional tones from movie scripts, and perceiving emotion from movie clips. This dual approach not only deepens insights into whether a film's intended emotional impact aligns with audience reactions but also enhances content recommendations, marketing strategies, and our overall understanding of cultural differences in emotional perception. By bridging the gap between a film's emotional intent and audience interpretation, the platform serves filmmakers, critics, streaming services, and moviegoers alike. Impact will be evaluated through user studies, benchmarked algorithmic performance, and interaction metrics, while acknowledging risks such as data bias and the complexity of modeling nuanced or culturally-specific expressions. If successful, the framework could be extended to broader media contexts, offering a valuable tool for enriched media analytics.

2 PROBLEM DEFINITION

Technical Version. Let $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$ denote a collection of n globally popular films. For each film f_i, we aim to assign a vector of emotional intensities $\mathbf{u}_i \in [0,1]^5$ based on user perception, where each dimension corresponds to a predefined emotion category $\mathcal{E} = \{e_1, e_2, e_3, e_4, e_5\}$. These user-generated emotion vectors are collected via an interactive platform, allowing personalized and subjective classification.

In parallel, we compute two additional emotion representations for each film:

(1) **Text-based sentiment vector** $\mathbf{t}_i \in \mathbb{R}^5$, derived from natural language processing (NLP) techniques applied to user reviews and movie scripts. This algorithm takes an input form of text strings.

(2) **Visual-audio emotion vector** $\mathbf{v}_i \in \mathbb{R}^5$, inferred from video analysis models applied to corresponding movie clips or trailers. This algorithm takes an input video of shape $\mathbb{R}^{T \times H \times W \times C}$, where T is the number of frames, H is the height of the video, W is the width of the video, and C is the number of channels of the video, and a prompt string.

The goal is to develop a system that jointly presents $\mathbf{u}_i$, $\mathbf{t}_i$, and $\mathbf{v}_i$ for each film f_i, enabling users to explore the alignment or divergence between:

- the *emotional tone from the script* of the film,
- the *audience's general sentiment* of the film,
- the *perceived emotion from the movie trailer*, and
- the *individual user's perception.*

Jargon-free Version. We aim to build an interactive platform that lets users classify globally popular films into five emotion categories based on personal perception, while comparing these subjective ratings with two computational analyses: one that extracts sentiment from audience reviews and movie scripts, and another that derives perceived emotion tones from movie clips. Our goal is to reveal whether a film's intended emotion aligns with audience reaction, as well as displaying the user's perception side-by-side for comparison.

3 LITERATURE SURVEY

Extracting Sentiments and Emotions from Movies

Understanding emotional content in films has been approached from multiple angles, including metadata analysis, sentiment extraction from text, and narrative modeling. Olesen et al. [9] highlight the potential of computational tools to uncover historical patterns in film archives, primarily through structured metadata. While their focus is on film historiography, it establishes a precedent for using computational analysis to understand broader thematic and emotional trends.

A more direct approach to extracting emotions is to analyze movie scripts and user reviews. Topal and Ozsoyoglu [12] develop a recommendation system that interprets IMDb reviews through the lens of Cambria et al.'s Hourglass Model, categorizing sentiment into dimensions such as pleasantness and attention. Although their system captures nuanced emotional profiles, it is limited to user-generated reviews and does not account for the film's intended emotional tone. Kurzhals et al. [6] take a different approach by aligning scripts with subtitles to map character interactions and scene structure, offering a framework for semantic film analysis.

More recent research has combined natural language processing with narrative structure analysis. Haris et al. [5] use sentiment analysis to investigate emotional expression across gendered dialogue in movie scripts. Similarly, Chun et al. [2] introduce *SentimentArcs*, a method for modeling the temporal progression of sentiment in narrative texts. Though originally designed for literature, this technique offers a valuable tool for capturing emotional dynamics in film scripts and user responses, which our project seeks to adapt and optimize.

Cluster-based and network approaches further enhance emotion classification. Ha et al. [4] analyze high-frequency sentiment words in movie reviews, categorizing them into emotional clusters such as "Happy," "Sad," and "Fear." By mapping these categories onto a similarity network of films, they provide an intuitive visualization of emotional proximity between movies. These methods serve as a foundation for our emotion-driven classification system, which integrates both user perception and script-level sentiment.

Visualization Techniques for Film Data

Parallel to advances in sentiment analysis, researchers have explored how visual representations can enhance the interpretability of film data. Dang [3] demonstrates how interactive visualizations can surface hidden gaps in feminist film historiography, illustrating how visualization acts not just as a presentation tool, but also as a mode of discovery. Their structured interface, while historically focused, informs the design of exploratory systems that enable user-driven investigation of film content and emotional patterns.

Building on this, Neware, Paul, and Poulose [8] develop a Tableau-based dashboard to visualize global film production trends, genre distributions, and financial metrics. Their integration of multiple chart types—maps, scatter plots, treemaps—offers a flexible and engaging way to present complex data. Although their work centers on industry metrics, the visualization principles translate well to emotional and cultural analyses.

Qian et al. [10] similarly emphasize interactivity through a multivariate visualization system that maps movie attributes for personalized recommendations. While their framework does not incorporate emotional variables, its design showcases how network-based interaction can help users uncover relationships within large media datasets.

Across these studies, a recurring theme is the importance of aligning visualization design with narrative and analytical goals. Bradbury and Guadagno's [1] emphasis on documentary-style storytelling supports this notion, demonstrating that visualization can do more than inform—it can evoke and guide emotional experience. Our work builds on these insights by integrating sentiment classification directly into the visual interface, enabling users to explore emotion as both a structural and experiential dimension of film.

4 METHOD

In this section we describe our end-to-end pipeline for extracting and visualizing emotional trajectories from movie scripts, subtitles, and user reviews. We first detail the script- and subtitle-based sentiment analysis, then outline the review-based analysis, and finally present the ensemble smoothing and peak-detection steps.

Script- and Subtitle-Based Emotion Analysis

We adapt the SentimentArcs framework to process time-aligned subtitle files, producing a smoothed emotional arc with key highlights:

(1) **Data Acquisition.** We gain movie subtitles from the Open-Subtitle[?] Database, and the subtitles are exported as CSV (see Movie_Script_Sentiments_Line.ipynb[1]).

[1]Code: Movie_Script_Sentiments_Line.ipynb.

(2) **Preprocessing.** Each line is lowercased, stripped of punctuation and special characters, and indexed by its position in the dialogue.
(3) **Sentiment Scoring.** We compute sentiment scores $s_{m,i}$ for each subtitle line i using seven transformer models from Hugging Face:
   - sentiment-roberta-large-english
   - distilbert-base-uncased-finetuned-sst-2-english
   - bert-base-uncased-imdb
   - roberta-base-imdb
   - robertuito-sentiment-analysis
   - distilroberta-finetuned-financial-news-sentiment-analysis
   - a second bertweet-base-sentiment-analysis run for cross-validation
(4) **Ensemble Smoothing.** We average across models to obtain

$$\bar{s}_i = \frac{1}{7}\sum_{m=1}^{7} s_{m,i},$$

then apply LOWESS smoothing (span $\alpha = 0.1$) to yield a continuous arc $\tilde{s}_i$.
(5) **Peak and Valley Detection.** Using a prominence threshold of 0.05, we detect local maxima (peaks) and minima (valleys) on $\tilde{s}_i$ to mark key emotional turning points.
(6) **Visualization.** See the smoothed subtitle-based sentiment arc for 200Meters in Appendix 6, which overlays the raw model trajectories, the LOWESS-smoothed ensemble mean (red), and detected peaks (▲) and valleys (▼).

Review-Based Emotion Analysis

We similarly process user reviews to capture overall audience sentiment:

(1) **Data Acquisition.** We load TMDB reviews to all_movies_reviews_emotion_scores.csv (see Movie_Review_Scirpts_Classification_Sentiments.ipynb[2]).
(2) **Preprocessing.** Reviews are segmented into sentences, HTML tags and URLs removed, then lowercased and tokenized.
(3) **Sentiment Scoring.** Each sentence is scored by using transformer sentiment classification model:
   - nateraw/bert-base-uncased-emotion

   Output the probability of six emotions in total: Fear, Sad, Surprise, angry, happy and love.
(4) **Aggregation.** We average sentence scores to obtain per-review sentiment, then compute the sentiment classification score for each film.

[2]Code: Movie_Review_Scirpts_Classification_Sentiments.ipynb.

Video-Based Emotion Analysis

Besides movie descriptions and user comments, we also analyze the invoked sentiment from the movie video itself using Vision Language Models (VLMs). Recent developments in VLMs have shown their strong potential in various image tasks due to their large-scale pertaining. Building from VLMs, people have developed Video Language Models, to solve multiple video-related tasks, including video question answering, video captioning, long video understanding, etc [7]. In this project, we will fine-tune a VLM for movie emotion analysis.

For our base VLM, we choose the NVILA-Lite-15B-Video from the NVILA family [7]. This set of models is the state-of-the-art of many video-related tasks. It is fully open-source with training code, datasets, and model checkpoints. Although it can be directly used for our task by prompting, the accuracy is not high. We decided to fine-tune the model for our task. To fine-tune this model, we use the EmoStim [11] dataset. It presents a collection of 139 videos annotated by around 1.6k user scores of the 16 common emotions such fear, anxiety, etc., they felt while watching the video. Each score is an integer value from one to five (both inclusive). The clips are excerpts from movies with a mean duration of around two minutes. We augment the data to with 20 prompt templates to ask for the emotion ratings of the video clip. For the target scores, we use the average scores of the user ratings for the corresponding video clip. For inference, we ask the model to rate a video using the 20 prompt templates and take the average of the score as the final score for the video.

Interactive Movie Emotion Visualization

We built a three-tiered, interactive visualization platform that guides users from a global overview of movie sentiment down to detailed, per-film emotion analyses. At the top level (Figure7), a responsive choropleth world map encodes each country by its average TMDB rating. Hovering over a nation reveals a custom tooltip with the total number of movies and the mean rating;

clicking any country smoothly scrolls and zooms into that region's movie list.

In the region view (Figure8), users see a fully sortable table of that locale's top films—complete with poster thumbnails, titles, release dates, popularity scores, average ratings, and vote counts. Rows highlight on hover and are clickable, linking to the film's detail page for deeper exploration.

Finally, the film detail interface (Figure9) presents a rich "movie header" card (poster, overview, popularity, rating, votes) alongside an emotion-selection widget. After the viewer tags the feelings a film evoked, three dynamic pie charts appear—showing subtitle-, review-, and video-based emotion scores —while the user's own selected emotions are displayed as colored tags (Figure10). Beneath the charts, the full subtitle script and individual review excerpts can be expanded or collapsed on demand, allowing side-by-side comparison of creator intent versus audience reaction.

This three-level platform takes you from a color-coded world map through sortable regional tables down to a film detail page that layers automated subtitle-, review-, and video-based emotion charts with your own selected emotion tags. By seeing creator intent, general audience reaction, and your personal response side by side, you can instantly spot mismatches—whether a movie's dialogue-driven sadness didn't resonate with you, or two equally rated films hide very different emotional tones. This blend of broad cultural trends and in-place, user-anchored insight invites richer reflection on how emotion shapes—and sometimes surprises—our experience of cinema.

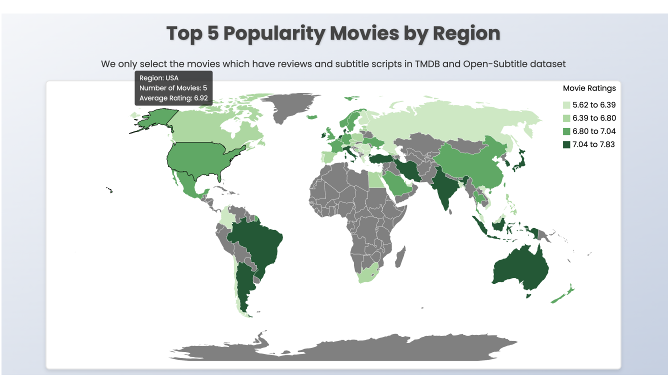

Figure 1: Example visualization of global movie rating distribution

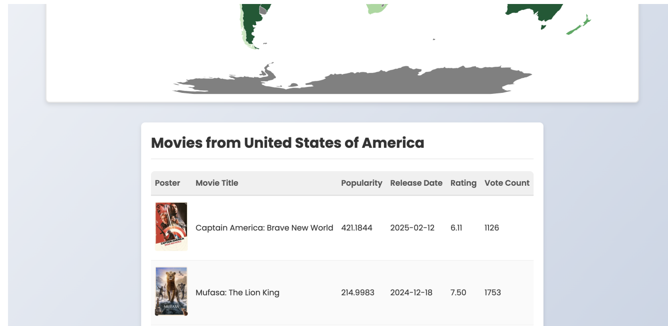

Figure 2: Example table of movies from a selected region

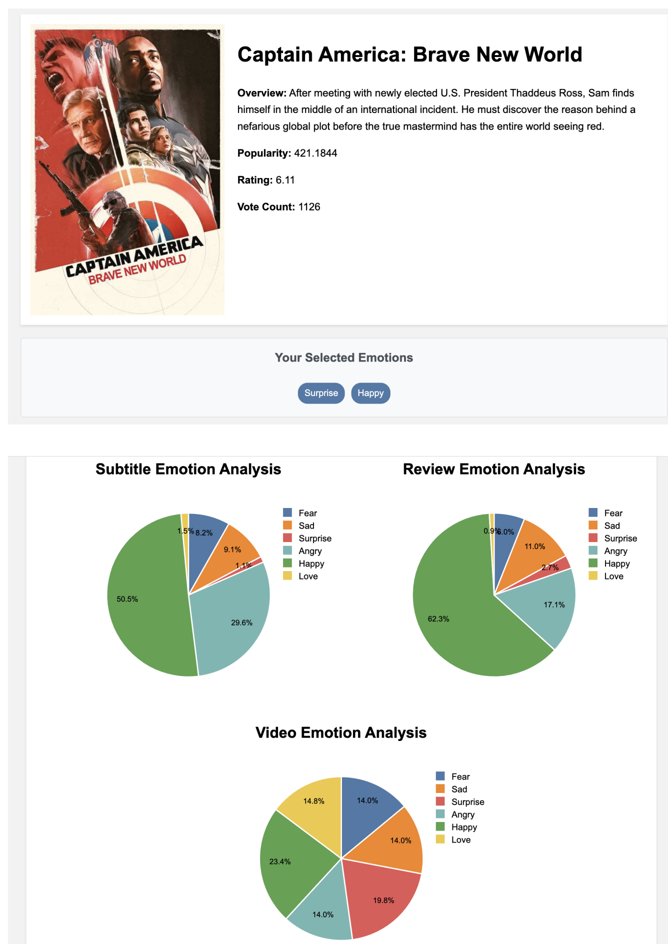

Figure 3: Example sentiment analysis plots for a selected movie

5 EVALUATIONS

Evaluation of Scripts and Review-Based Emotion Analysis

For script-based analysis, in the absence of direct ground truth labels, we utilized an ensemble approach, aggregating the sentiment trajectories generated by multiple transformer-based models to derive a consensus emotional trajectory. Each individual model's performance was then quantitatively assessed against this consensus trajectory using RMSE to measure deviations and determine each model's effectiveness in capturing the underlying emotional dynamics. Figure 12 illustrates the RMSE across different transformer models. The `roberta-base-imdb` model achieved the lowest RMSE, indicating superior performance, while the `pysentimiento/robertuito` model exhibited the highest RMSE, highlighting potential limitations in accurately capturing script-based emotional dynamics.

Evaluation of Video-Based Emotion Analysis

We split the EmoStim [11] dataset into train and test sets with 80/20 ratios. We fine-tune the base model on the training set for one epoch, two epochs, and three epochs. Thus, we have three fine-tuned models in the end. We choose these small numbers to avoid overfitting on the training set. Training three separate models with different approaches and evaluating them later on the testing set also allows us to select the best model in the end.

After training on the training set, we run the proposed inference pipeline on the testing set. First, we calculate the Mean Absolute Error (MAE) between our predicted scores and the ground truth scores over the 28 videos and 16 emotions. We also provide two baselines. 1. Fixed score of 3 (the middle of the scoring range of 1-5). 2. Uniform random guess between 1 and 5. These baselines completely ignore the video modality and effectively perform a random guess of the emotion scores. They provide a baseline of the worst performance of any algorithm. The result is shown in Figure 4.

As an ablation study, we also report the MAE of the pretrained NVILA-Lite-15B-Video model in Figure 4. The base model performs closely to the fixed score guess baseline (Baseline). Our fine-tuned models again perform significantly better than the pretrained model.

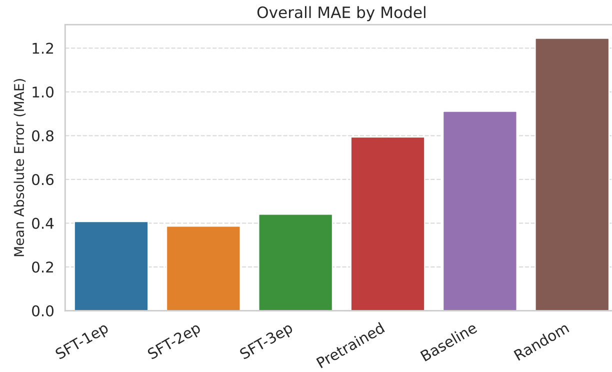

Figure 4: Mean Absolute Error of Models.

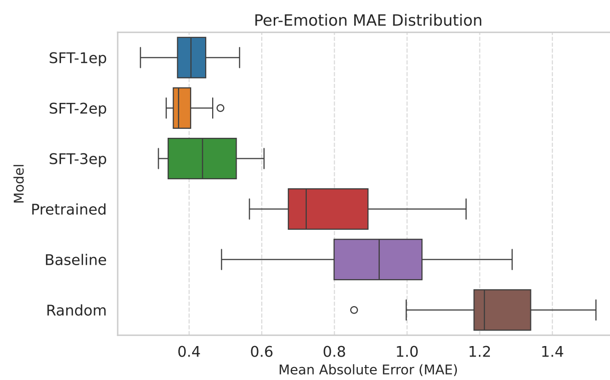

Figure 5: Boxplot of the Per-Emotion Model MAE.

A single MAE value may not be informative enough to tell the performance of our method. We would like to know if the model consistently understands different emotion concepts. To do this, we analyze the MAE error between the model prediction and the ground truth over the 28 videos. This gives the model MAE error per emotion. We report this with a box plot in Figure 5. Note that the model that was trained two epoch performs the best. It understands almost all of the 16 emotion concepts better than the baseline methods and the pretrained model. A more detailed per-emotion MAE figure is included in the appendix (See Figure 11).

6 CONCLUSION AND DISCUSSION

All team members have contributed a similar amount of effort.

In this project, we developed an interactive platform that analyzes emotional content in films through three modalities: movie scripts, audience reviews, and trailers. By combining these computational insights with user input, the system enables nuanced comparisons between creator intent, collective sentiment, and personal emotional response.

Our results demonstrate that ensemble-based script analysis, review sentiment classification, and fine-tuned video-language models can effectively extract emotional signals. The platform's intuitive design allows users to explore emotional alignment across cultural and individual perspectives, offering new tools for media recommendation, audience research, and film analysis.

Despite its strengths, the system is limited by its fixed emotion categories and reliance on potentially noisy or biased data sources. It also lacks adaptive learning from user interaction. Future work could include broader emotional taxonomies, multilingual support, and real-time user feedback integration.

Ultimately, this project highlights the power of multimodal emotion analysis to deepen our understanding of how films resonate emotionally—both individually and across global audiences.

REFERENCES

[1] J. D. Bradbury and R. E. Guadagno. 2020. Documentary Narrative Visualization: Features and Modes of Documentary Film in Narrative Visualization. *Information Visualization* 19, 4 (2020), 339–352. https://doi.org/10.1177/1473871620925071

[2] Jon Chun. 2021. SentimentArcs: A Novel Method for Self-Supervised Sentiment Analysis of Time Series Shows SOTA Transformers Can Struggle Finding Narrative Arcs. *arXiv preprint arXiv:2110.09454* (2021). https://arxiv.org/abs/2110.09454

[3] S.-M. Dang. 2023. The Women Film Pioneers Explorer: What Data Visualizations Can Tell Us About Women in Film History. *Feminist Media Histories* 9, 2 (2023), 76–86. https://doi.org/10.1525/fmh.2023.9.2.76

[4] Hyoji Ha, Hyunwoo Han, Seongmin Mun, Sungyun Bae, Jihye Lee, and Kyungwon Lee. 2019. An Improved Study of Multilevel Semantic Network Visualization for Analyzing Sentiment Word of Movie Review Data. *Applied Sciences* 9, 12 (2019), 2419. https://doi.org/10.3390/app9122419

[5] Muhammad Junaid Haris, Aanchal Upreti, Melih Kurtaran, Filip Ginter, Sebastien Lafond, and Sepinoud Azimi. 2023. Identifying gender bias in blockbuster movies through the lens of machine learning. *Humanities and Social Sciences Communications* 10, 1 (2023), 1–11. https://doi.org/10.1057/s41599-023-01576-3

[6] Kuno Kurzhals, Markus John, Florian Heimerl, Paul Kuznecov, and Daniel Weiskopf. 2016. Visual Movie Analytics. *IEEE Transactions on Multimedia* 18, 11 (2016), 2149–2160. https://doi.org/10.1109/TMM.2016.2614184

[7] Zhijian Liu, Ligeng Zhu, Baifeng Shi, Zhuoyang Zhang, Yuming Lou, Shang Yang, Haocheng Xi, Shiyi Cao, Yuxian Gu, Dacheng Li, Xiuyu Li, Yunhao Fang, Yukang Chen, Cheng-Yu Hsieh, De-An Huang, An-Chieh Cheng, Vishwesh Nath, Jinyi Hu, Sifei Liu, Ranjay Krishna, Daguang Xu, Xiaolong Wang, Pavlo Molchanov, Jan Kautz, Hongxu Yin, Song Han, and Yao Lu. 2024. NVILA: Efficient Frontier Visual Language Models. arXiv:2412.04468 [cs.CV] https://arxiv.org/abs/2412.04468

[8] Govindram Neware, John Paul, and Alwin Poulose. 2024. Exploring the Cinemaverse: Data-Driven Analysis of Movie Production, Revenue, and Trends Using Tableau Visualization Tool. *Proceedings of the 2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)* (2024). https://doi.org/10.1109/ICUFN61752.2024.10625227

[9] C. G. Olesen, E. Masson, J. Van Gorp, G. Fossati, and J. Noordegraaf. 2016. Data-Driven Research for Film History: Exploring the Jean Desmet Collection. *The Moving Image* 16, 1 (2016), 82–105. https://www.jstor.org/stable/10.5749/movingimage.16.1.0082

[10] Dong Qian, Cheng Yang, Chen Li, and Ling Bai. 2016. A visualization and interaction system of multivariate movie network data. In *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*. 379–383. https://doi.org/10.1109/SIPROCESS.2016.7888288

[11] Rukshani Somarathna, Patrik Vuilleumier, and Gelareh Mohammadi. 2024. EmoStim: A Database of Emotional Film Clips With Discrete and Componential Assessment. *IEEE Transactions on Affective Computing* 15, 3 (2024), 1202–1212. https://doi.org/10.1109/TAFFC.2023.3328900

[12] Kamil Topal and Gultekin Ozsoyoglu. 2016. Movie review analysis: Emotion analysis of IMDb movie reviews. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 1170–1176. https://doi.org/10.1109/ASONAM.2016.7752387

APPENDICES

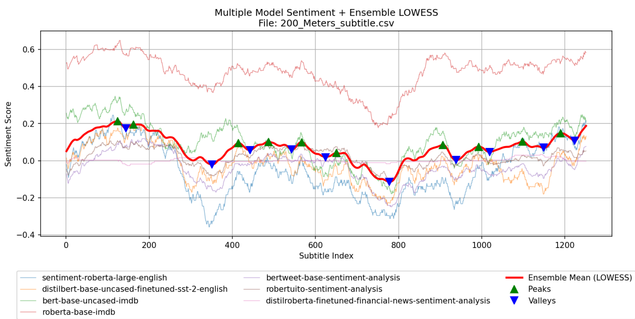

Figure 6: Subtitle-based sentiment trajectory for *200 Meters*. The red curve is the LOWESS-smoothed ensemble mean; green and blue triangles mark peaks and valleys.

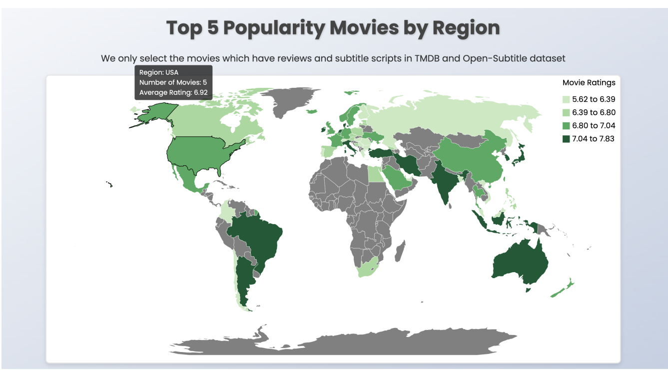

Figure 7: Example visualization of global movie rating distribution

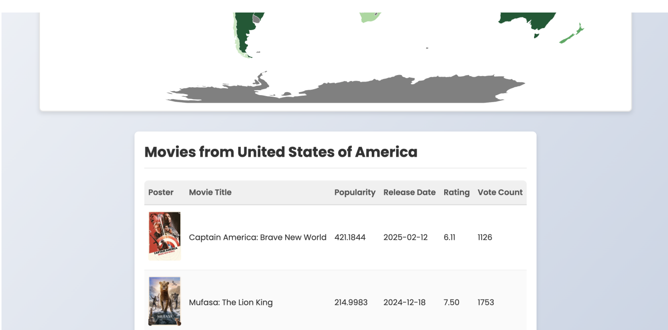

Figure 8: Example table of movies from a selected region

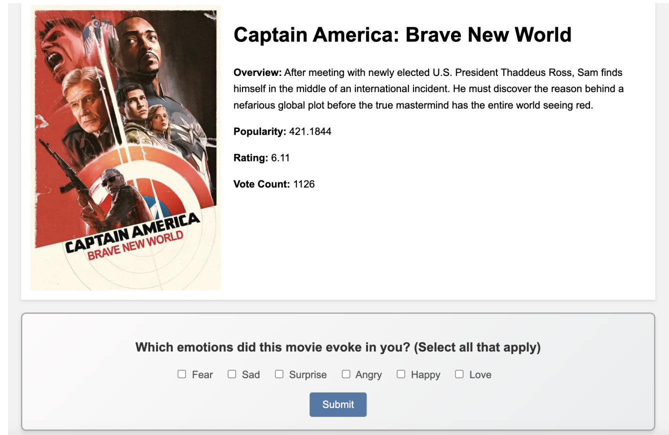

Figure 9: Example of Emotion Selection Interaction Interface

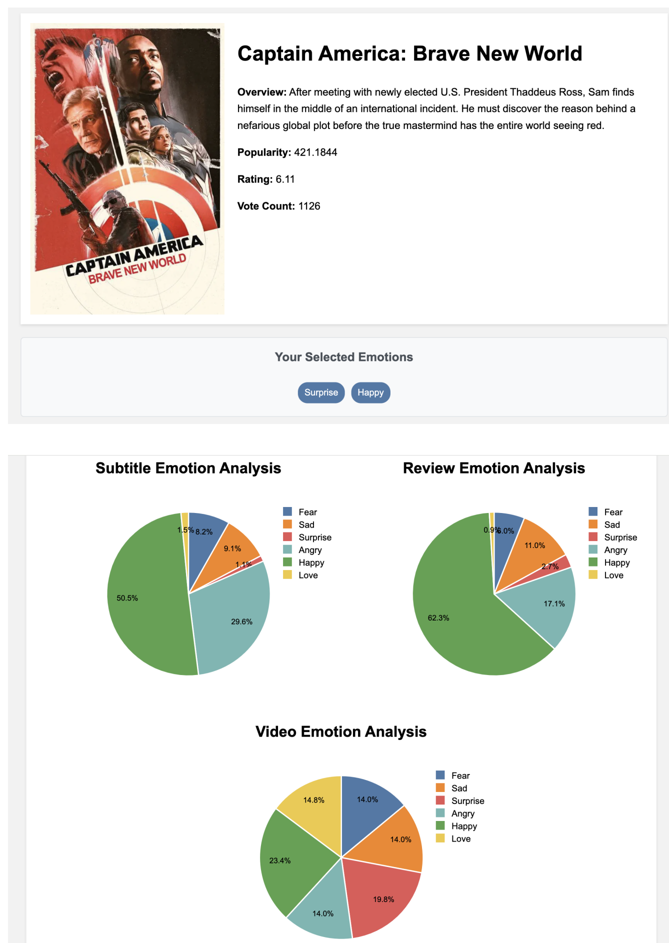

Figure 10: Example sentiment analysis plots for a selected movie

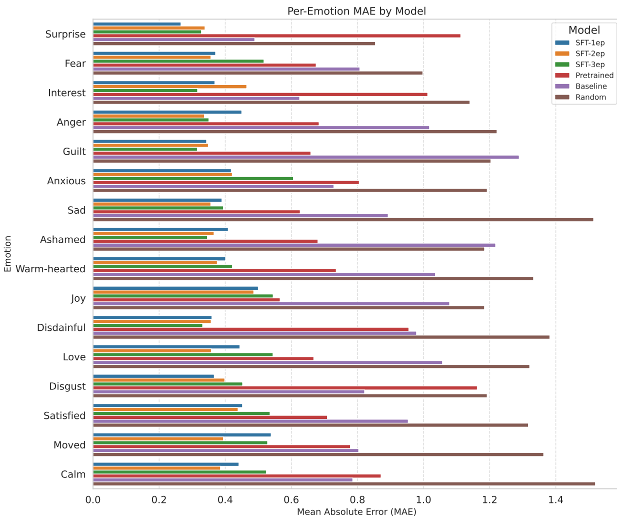

Figure 11: Per-Emotion Model MAE

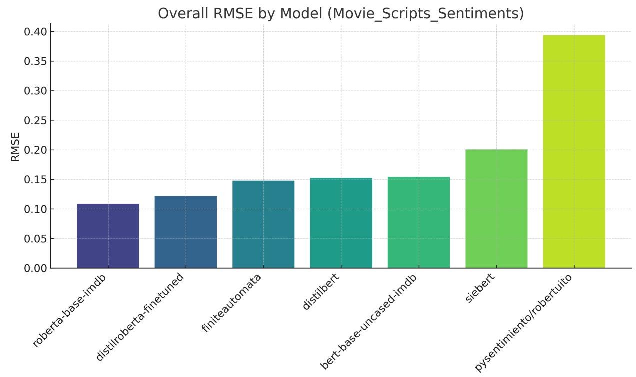

Figure 12: Overall RMSE by Model for Movie Script Sentiments Analysis.