TIERS Information Technology Journal
Vol.6, No.2, December 2025, pp. 217~240
P-ISSN: 2723-4533 / E-ISSN: 2723-4541
DOI: https://doi.org/10.38043/tiers.v6i2.7183

Genre-Based Sentiment and Emotion System for Audience Insight

Suwarno[1], Wesly[2], Bayu Syahputra[3]

suwarno.liang@uib.ac.id[1], 2231075.wesly@uib.edu[2], bayu@uib.ac.id[3]

Faculty of Computer Science, University International Batam, Indonesia

ABSTRACT

Movies can influence people's moods in different ways depending on film genre. Fear is commonly induced by horror films, whereas joy is typically associated with comedy. Understanding how genre-based expectations shape audience emotions offers valuable insights for producers and digital platforms. However, previous studies have only briefly examined this relationship, with most focusing on general sentiment analysis. This study develops a genre-based sentiment and emotion model to analyze how film genres influence audience reactions. The Cross-Industry Standard Process for Data Mining (CRISP-DM) framework was applied to 46,173 IMDb reviews using Term Frequency–Inverse Document Frequency (TF-IDF) features and three machine learning models: Logistic Regression, Linear Support Vector Classification, and One-vs-Rest Logistic Regression. The results show that Fear (0.704) and Anger (0.684) are the most dominant emotions, indicating that audiences tend to be more emotionally engaged with intense genres. The model was also implemented in a Flask–React web-based system that allows users to analyze and visualize reviews in real time. These findings provide practical implications for filmmakers, producers, and streaming platforms in adjusting genre design, content recommendation, and promotional strategies to align with audience emotional responses.

Keywords: genre; sentiment; emotion; machine learning; CRISP-DM

Article Info

Received : 01-07-2025
Revised : 11-09-2025
Accepted : 29-12-2025

Correspondence Author:

Suwarno
Department of Information Systems
University International Batam,
Baloi-Sei Ladi, Gajah Mada Street, Tiban Indah, District Sekupang, Batam City, Riau Islands 29426.
Email: suwarno.liang@uib.ac.id

1. INTRODUCTION

The film industry continues to grow rapidly in the digital era, supported by online platforms such as IMDb that allow audiences to share opinions and emotions about the films they watch [1]. The large volume of reviews makes manual reading and interpretation inefficient; therefore, automated sentiment analysis is required to classify opinions into positive and negative categories [2]. With the increasing amount of textual data, recent studies have emphasized the need for automated systems capable of recognizing emotions in text accurately and in real time [3].

Film genres strongly influence audience emotions. Comedy is often associated with happiness and relaxation [4], whereas drama tends to evoke empathy and emotional involvement with the story and characters [5]. In contrast, horror films can trigger fear, anxiety, and even trauma in children [6]. Previous studies have also reported that exposure to horror films may lead to depression [7] and aggressive behavior among teenagers [8]. When the story or emotional tone of a film does not meet audience expectations, viewers are more likely to leave negative reviews [9], which can negatively affect a film's reputation and commercial success [10].

Previous studies have explored the relationship between films and audience emotions using various approaches. For example, the BiProjection Multimodal Transformer model combines textual, visual, and audio features for automatic emotion and genre recognition, achieving strong performance [11]. Other studies employ real-time facial expression analysis to detect audience emotions while watching films [12]. In the textual domain, Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Bidirectional LSTM models have been applied to predict genres and emotions from subtitles or written reviews [13], [14]. However, visual-based models such as Convolutional Neural Networks (CNNs), including ResNet, VGG, and Xception, often perform poorly under varying lighting conditions, camera angles, and complex scenes, indicating that text-based analysis remains more reliable for large-scale emotion detection [15].

Traditional machine-learning methods using Term Frequency–Inverse Document Frequency (TF-IDF) features are also widely adopted for movie review analysis. Support Vector Machine (SVM) models have achieved 79% accuracy on IMDb reviews [16], while SVM combined with TF-IDF reached 82.51% accuracy on Indonesian horror film reviews [17]. Logistic Regression achieved 80.61% accuracy [18], and further optimization using Grid Search improved performance to 90.90% [19]. TF-IDF combined with emotion detection also reported accuracy rates of up to 90% [20]. In addition, text-mining approaches incorporating lexicon-based methods and natural language processing tools have been used to identify sentiment in written reviews, highlighting the importance of effective data preparation and well-defined preprocessing steps in sentiment analysis [21]. Linear SVC optimized using Particle Swarm Optimization achieved an accuracy of 92.03% [22]. Other studies employing hybrid features, such as TF-IDF, n-grams, and Information Gain, have also reported improved classification accuracy [23]. Classical algorithms, including Random Forest and Naïve Bayes, have demonstrated effectiveness across diverse datasets [24], while clustering approaches such as K-Means combined with SEMMA have also shown promising results in text classification [25].

Numerous previous studies have investigated the development of sentiment and emotion analysis models; however, most remain primarily focused on technical performance, particularly classification accuracy. Limited attention has been given to examining differences in audience emotions resulting from exposure to different film genres, despite the fact that genre plays a significant role in shaping audience expectations and emotional responses. When genre context is not considered, the practical value of sentiment analysis results becomes difficult to interpret in real-world applications. Therefore, genre-specific sentiment and emotion analysis is required to better understand how different film categories evoke distinct emotional responses. Such insights are essential for filmmakers and digital platforms when developing marketing strategies, content positioning, and production decisions that align more closely with audience emotional preferences [26].

Although many methods have been proposed, most existing studies continue to emphasize model performance and accuracy rather than exploring how film genres influence audience emotions. This study addresses this gap by applying machine-learning-based sentiment and emotion analysis to IMDb movie reviews. Logistic Regression and Linear SVC are employed with TF-IDF features due to their proven efficiency and interpretability for text-based data [27], [28], [29]. The novelty of this research lies in demonstrating how audience emotions vary across genres, comparing model performance within a genre-based sentiment analysis framework, and providing actionable insights for producers and digital platforms to better align content and marketing strategies with audience emotional preferences.

2. RESEARCH METHOD

This study adopts an applied research approach using a case study based on IMDb movie reviews collected from the Kaggle platform. The objective is to analyze how film genres influence audience emotions through machine-learning-based sentiment analysis. The model developed in this study classifies positive and negative sentiment and identifies emotional patterns across different genres. The findings are expected to assist producers and digital platforms in understanding audience responses more effectively.

The research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, which consists of five main stages: data understanding, data preprocessing, modeling, evaluation, and implementation. This framework ensures a structured and consistent workflow from data collection to final model assessment [30].

The overall research process is illustrated in Figure 1, beginning with data understanding and text cleaning, followed by feature extraction using TF-IDF, model training with machine-learning algorithms, result evaluation, and implementation of the developed model into a Flask–React web application for interactive visualization. The workflow design was adapted and refined based on previous studies to emphasize genre-based sentiment and emotion analysis, providing a clear foundation for examining how genres shape audience emotional responses [11]. After system implementation, a system evaluation is conducted to assess whether the web-based application operates properly and is easy to use. This evaluation ensures that users can perform analyses smoothly and clearly interpret the sentiment and emotion results generated by the system [31].

.

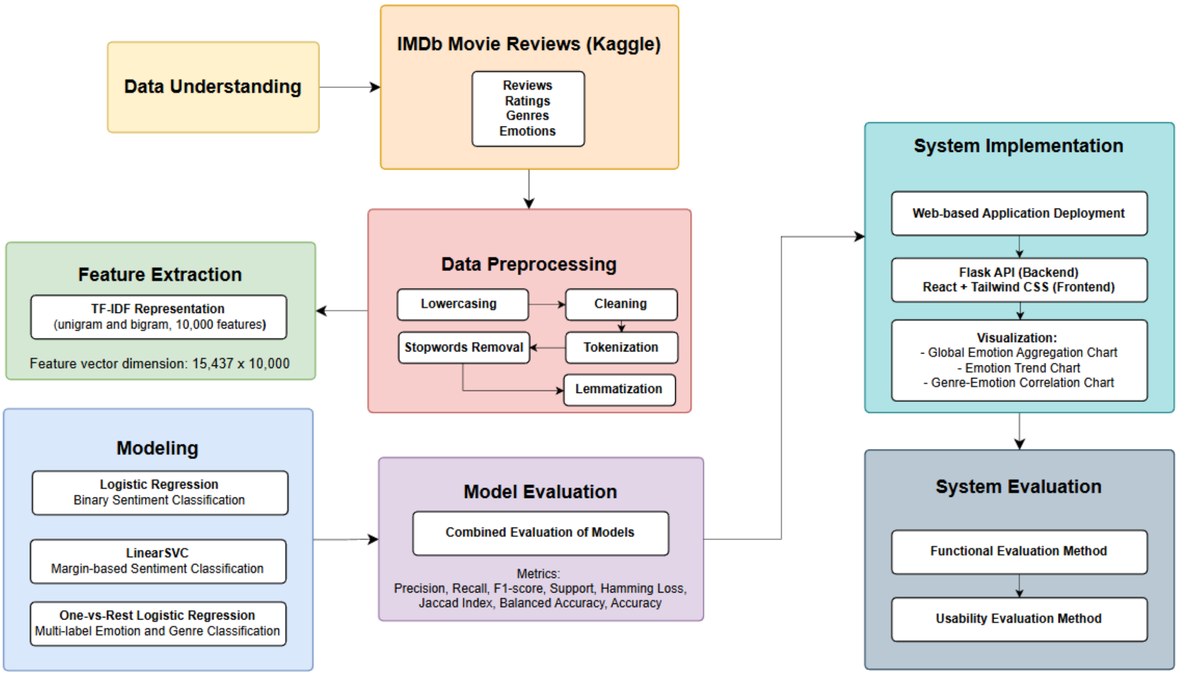


Figure 1. Proposed CRISP-DM Workflow for Genre-Based Sentiment and Emotion Analysis

Figure 1 illustrates the research workflow adopted in this study based on the CRISP-DM framework. The process begins with data understanding of IMDb movie reviews obtained from the Kaggle platform, including review text, ratings, genres, and emotion labels. Data preprocessing is subsequently performed through text cleaning, lowercasing, tokenization, stop word removal, and lemmatization to enhance data quality. TF-IDF feature extraction is then applied to transform textual data into numerical representations for model training. Logistic Regression and Linear SVC are employed for sentiment classification, while emotion–genre relationships are analyzed using One-vs-Rest Logistic Regression. The workflow concludes with model evaluation and deployment of the system into a web-based application. Following deployment, a system evaluation is conducted to assess functional performance and usability, ensuring that the application operates correctly and that users can easily access and interpret the sentiment and emotion analysis results [31].

2.1. Data Understanding

The dataset used in this study consists of 46,173 IMDb movie reviews obtained from the Kaggle platform. Each entry contains four main attributes: review (English-language review text), rating (a numerical score ranging from 1 to 10), genres (film categories, where a single film may belong to more than one genre), and emotions (eight emotion categories: anger, anticipation, disgust, fear, joy, optimism, and sadness) [32].

Rating values are used to generate binary sentiment labels for classification purposes. Reviews with ratings ≥ 6 are categorized as positive, while those with ratings < 6 are classified as negative. This labeling scheme has been adopted in previous IMDb-based sentiment analysis studies [33], [34].

Emotion labels are already provided in the dataset; however, their distribution is highly imbalanced. Sadness constitutes the dominant emotion class, whereas surprise occurs only in a very small number of instances. This imbalance may cause the model to achieve higher accuracy for dominant emotion classes [35]. Therefore, examining the initial distribution of emotional labels is an important step prior to the modeling stage.

Initially, the dataset was cleaned by removing duplicate and blank entries to obtain a consistent set of final reviews. The distribution of reviews across genres and emotional categories is presented in Table 1, providing an overview of the predominance of certain genres and the diversity of emotional responses expressed by viewers. This distribution represents the state of the dataset prior to preprocessing and serves as a basis for characterizing the data used in this study. The importance of examining this initial distribution has also been highlighted in previous research, which indicates that online review datasets are often imbalanced, as the majority of reviews belong to dominant classes; without appropriate handling, such imbalance may bias model performance toward those classes [36].

Table 1. Distribution of Audience Emotion across Movie Genres in IMDb Dataset

Genre	Anger	Anticipation	Disgust	Fear	Joy	Optimism	Sadness	Total
Drama	1,758	3,452	654	1,293	3,691	2,373	9,877	23,098
Romance	1,380	2,666	601	1,371	3,854	1,689	7,568	19,129
Comedy	1,330	2,411	676	805	3,231	1,633	5,804	15,890
Thriller	753	1,452	406	913	894	891	3,270	8,579
Action	1,091	1,117	394	726	721	848	1,906	6,803
Fantasy	264	777	66	667	1,214	604	1,966	5,558
Horror	351	761	217	609	688	432	1829	4887
Family	88	780	73	355	1,354	496	1,734	4,880
Adventure	664	670	184	491	785	882	1,074	4,750
Crime	542	865	133	405	433	210	1,370	3,958

Based on Table 1, reader responses across the ten genres receiving the highest number of reviews are presented. Drama recorded 23,098 responses, while Romance accounted for 19,129 responses, with these genres primarily associated with the emotions of sadness and joy, respectively. Thriller contributed 8,579 responses, which were commonly associated with fear and anger, whereas Horror recorded 4,887 responses and was predominantly linked to fear. These results indicate that different genres evoke distinct emotional tendencies among viewers, highlighting the relevance of this dataset for research on genre-based emotion analysis. This observation is consistent with previous studies demonstrating that emotional traits in film reviews can be used to predict genre, with certain genres exhibiting dominant emotional patterns, such as Drama being closely associated with sadness and Horror emphasizing fear [37].

2.2. Data Preprocessing

Data preprocessing is the stage in which the dataset is prepared for machine-learning modeling. Text reviews were processed through several steps, including lowercasing, removal of non-informative characters and stopwords, and lemmatization to obtain base word forms. These procedures help normalize the textual data and reduce noise, which can improve model performance when working with unstructured online text [38].

Subsequently, the dataset underwent consistency checks across the main attributes, namely review, sentiment, emotion, and genre. The resulting dataset comprised 19,297 cleaned reviews, divided into 15,437 training samples and 3,860 testing samples, stratified by sentiment labels. Reviews with ratings ≥ 6 were classified as positive, whereas those with ratings < 6 were classified as negative, resulting in 11,131 positive and 8,166 negative reviews.

Following data cleaning and validation, the dataset was ready for analysis, featuring relatively balanced sentiment classes alongside an imbalanced distribution of emotion labels. A summary of these distributions is presented in Table 2 to provide an overview of the dataset after preprocessing.

Table 2. Distribution of Sentiment and Emotion in IMDb Dataset

Category	Class	Total
Sentiment	Positive	11,131
	Negative	8,166
Emotion	Sadness	7,810
	Joy	3,318
	Anticipation	2,568
	Optimism	2,100
	Fear	1,575
	Anger	1,257
	Disgust	669

Based on Table 2, the dataset contains 11,131 positive reviews and 8,166 negative reviews. In contrast, the distribution of emotion labels is notably imbalanced. Sadness dominates the dataset with 7,810 reviews, followed by Joy (3,318 reviews), Anticipation (2,568 reviews), Optimism (2,100 reviews), Fear (1,575 reviews), Anger (1,257 reviews), and Disgust (669 reviews). These figures indicate that, while the sentiment data are relatively balanced, the emotion distribution is clearly skewed, with a pronounced predominance of certain emotional categories.

.

2.3. Feature Extraction (TF-IDF)

After the preprocessing stage was completed, the cleaned review text was transformed into numerical representations using the Term Frequency–Inverse Document Frequency (TF-IDF) approach. This representation was selected because it emphasizes informative terms by weighting word frequency within individual documents relative to their occurrence across the entire corpus. In addition, TF-IDF is effective in reducing the influence of common but less informative words in sentiment analysis [39].

In this study, TF-IDF was applied using unigram and bigram representations, limited to the top 10,000 features. The vectorizer was fitted exclusively on the training data to prevent data leakage, while the test data were transformed separately. This procedure ensures that the model better reflects real-world performance when applied to unseen data [40].

The resulting TF-IDF representation produced a training matrix of 15,437 × 10,000 and a test matrix of 3,860 × 10,000, indicating that each review was represented as a 10,000-dimensional feature vector. Feature extraction was conducted independently for sentiment and emotion classification tasks to ensure consistent and unbiased evaluation [41].

2.4. Modeling

After the review text was represented as TF-IDF vectors, the next step involved building machine-learning models to classify audience sentiment and emotion within a genre-based context. This study employed two algorithms: Logistic Regression and Support Vector Machine (SVM). These models were selected because they are widely used in text analysis and have demonstrated strong performance in high-dimensional feature spaces such as TF-IDF representations [42].

Logistic Regression learns linear relationships between textual features and sentiment labels. It was selected due to its simplicity, computational efficiency, and frequent use as a robust baseline model for text classification. To ensure stable and balanced learning outcomes, class weighting was applied to adjust the contribution of each sentiment category. This strategy helps mitigate the effects of class imbalance and improves model performance on underrepresented classes. High TF-IDF values indicate stronger feature influence on predictions, enabling the model to capture variations in emotional tone across genres effectively [43].

Although transformer-based models such as BERT often achieve higher accuracy, this study deliberately employs TF-IDF features in combination with Logistic Regression and Linear SVC to enhance model interpretability [44]. Both models allow feature contributions to be examined through their coefficients, enabling transparent and traceable decision-making processes [45]. Such transparency is particularly important for genre-based audience analysis, as filmmakers and content teams require an understanding of why specific emotions are predicted, rather than relying solely on final classification scores. Consequently, the use of TF-IDF represents a methodological choice aligned with Explainable Artificial Intelligence (XAI) principles, allowing linear models to reveal interpretable emotional patterns across film genres [46].

Beyond interpretability, model selection was also guided by computational efficiency and deployment considerations. As the proposed approach is implemented within a web-based system supporting near real-time review analysis, traditional machine-learning models using TF-IDF offer faster inference and lower computational costs compared to deep learning models such as BERT or Bi-LSTM. This characteristic makes the proposed approach well suited for real-time web applications, where responsiveness and scalability are critical, and positions the lightweight model design as a deliberate and practical choice rather than a limitation [47].

In addition to binary sentiment classification, this study also developed a multi-label classification model that jointly analyzes emotion and genre labels. This model was implemented using a One-vs-Rest Logistic Regression approach, in which each label is learned independently. Such an approach is effective for simultaneously examining genre–emotion relationships while maintaining training efficiency and consistency of results [48].

The One-vs-Rest Logistic Regression (OVR-LR) model adopted in this study employs a multi-label learning approach in which each emotion label is learned independently. This design enables the model to simultaneously analyze emotion–genre relationships and capture variations in emotional expression across different film genres. Each genre and emotion label is modeled independently, without direct interference from other labels, resulting in a flexible and computationally efficient framework. This approach facilitates a more accurate and consistent understanding of the complex relationship between emotional expression and audience genre expectations across varying data conditions [49].

The modeling framework developed in this study comprises three main pipelines: Logistic Regression for sentiment classification, Linear SVC for sentiment classification, and One-vs-Rest Logistic Regression for genre–emotion classification. Each pipeline is designed to produce consistent and objective results that reflect model performance in capturing audience emotions and preferences within a genre-based context [50].

The modeling design is presented in Figure 2, which illustrates the overall workflow of the proposed approach. The workflow consists of three primary pipelines Logistic Regression, Linear SVC, and One-vs-Rest Logistic Regression each responsible for specific tasks ranging from binary sentiment classification to multi-label emotion and genre prediction. The modeling design was adapted and refined based on previous studies to align with the focus of this research on TF-IDF feature representation, ensuring that the entire process from feature extraction and model training to final prediction remains consistent and free from data leakage [11].

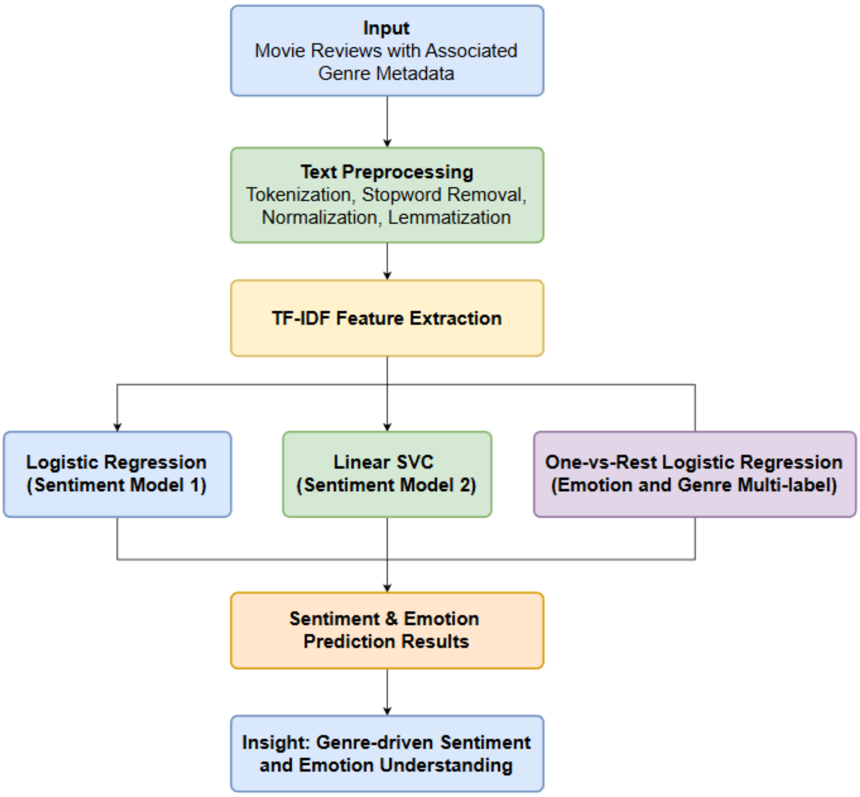


Figure 2. Proposed Genre-Sentiment-Emotion Classification

As illustrated in Figure 2, the workflow begins with text preprocessing, including tokenization, stopword removal, normalization, and lemmatization. Three parallel pipelines Logistic Regression, Linear SVC, and One-vs-Rest Logistic Regression are then executed to perform sentiment classification, identify emotional tendencies, and associate them with film genres, resulting in integrated genre-based sentiment and emotion predictions.

The outcomes of the modeling phase are summarized in Table 3, which presents the primary configurations and hyperparameters of the three models: Logistic Regression, Linear SVC, and One-vs-Rest Logistic Regression. All models were constructed using TF-IDF feature representations to ensure consistency and stability throughout the classification process.

Table 3. Configuration and Hyperparameters of the Proposed Models

Model	Algorithm	Key Hyperparameters	Task
Model 1	Logistic Regression	C=1.0, class_weight='balanced'	Sentiment
Model 2	Linear SVC	C= 1.0, dual=False, class_weight='balanced'	Sentiment
Model 3	One-vs-Rest Logistic Regression	C=1.0, class_weight='balanced'	Multi-label Emotion and Genre

Based on Table 3, Logistic Regression and Linear SVC were applied for sentiment classification, while One-vs-Rest Logistic Regression was employed for multi-label emotion and genre prediction. All models used balanced class weights and consistent regularization (C = 1.0) to ensure stable and reliable performance in genre-based sentiment and emotion analysis.

.

2.5. Model Evaluation

The evaluation phase is conducted to quantify the classification performance of the machine-learning models. This evaluation is based on several widely used metrics, including precision, recall, F1-score, and accuracy, which together enable assessment of the models' ability to consistently and correctly identify both positive and negative sentiment classes [51].

In addition to the primary evaluation metrics, this study also employs balanced accuracy, Hamming loss, and the Jaccard Index to assess model performance in analyzing audience emotions across genres. Balanced accuracy provides a more consistent evaluation in the presence of uneven class distributions across genres [52]. Hamming loss quantifies the average prediction error for each emotion label in a multi-label classification setting [53]. The Jaccard Index is used to measure the similarity between predicted and actual emotion labels, offering additional insight into the model's ability to capture audience emotion patterns based on genre expectations [54].

Using this combination of metrics enables a more robust and objective evaluation. Accuracy reflects overall prediction correctness, while precision and recall indicate the model's ability to minimize classification errors within specific classes. The F1-score balances precision and recall, helping to mitigate bias toward dominant classes. All metrics are applied to both sentiment and emotion classification tasks, resulting in a comprehensive evaluation that objectively and representatively reflects model performance [55].

2.6. System Implementation

At this stage, the developed sentiment and emotion analysis model is deployed within a web-based application. This application was designed to ensure that the research outcomes extend beyond the modeling phase and can be directly utilized to observe audience emotional patterns across different film genres. Previous studies have shown that such systems can enhance the effectiveness of sentiment analysis and provide clearer insights into how viewers emotionally respond to films [56].

The backend of the application is implemented using a Flask API to process data and generate model predictions, while the frontend is developed with React and Tailwind CSS to enhance interactivity and usability. Through this system, users can input movie reviews and immediately view visualizations presented in three main charts: the Global Emotion Aggregation Chart, the Emotion Trend Chart, and the Genre–Emotion Correlation Chart.

These visualizations are interconnected and automatically display analysis results based on the submitted reviews. Consequently, the application serves not only as a tool for examining the relationship between film genres and audience emotional states, but also as a means of demonstrating the practical applicability of the research. This implementation stage provides assurance that the developed model is suitable for real-world use and capable of delivering timely insights into audience emotional responses to films [57].

2.7. System Evaluation

Following the system implementation stage, a system evaluation is conducted to assess the performance of the developed web-based application when applied to genre-based sentiment and emotion analysis. This evaluation is essential, as the study extends beyond model development to ensure that analysis results can be effectively accessed and utilized through the system [58]. The system evaluation focuses on determining whether the application operates as intended, supports the analysis process smoothly, and presents sentiment and emotion information in a clear and interpretable manner. In this study, the system is evaluated from both functional and usability perspectives to ensure that it supports the practical exploration of audience emotional responses across different film genres [31].

2.7.1 Functional Evaluation Method

The functional evaluation method is employed to ensure that the web-based system operates correctly when performing genre-based sentiment and emotion analysis. This evaluation focuses on the system's core functionalities, including inputting movie reviews, selecting genres, executing sentiment and emotion predictions, and displaying analysis results. The objective is to verify that each feature functions as intended and supports the analysis process without technical issues [59].

Functional testing is conducted by running the system using movie reviews from various genres and observing system behavior at each stage of the analysis process [60]. This procedure assesses whether user inputs are processed correctly, prediction models execute smoothly, and sentiment and emotion results are presented consistently. Through this evaluation, the system's readiness to support genre-based sentiment and emotion analysis in practical applications can be reliably assessed [26].

2.7.2 Usability Evaluation Method

The usability evaluation method is applied to assess how easily users can operate the web-based system when performing genre-based sentiment and emotion analysis [61]. This evaluation focuses on interface clarity, ease of access to core features, and users' ability to understand the analysis results presented by the system. The objective of this evaluation is to ensure that the system can be used comfortably and effectively without requiring technical expertise [62].

Usability evaluation is conducted by inviting users to interact directly with the system, including entering movie reviews, selecting genres, and examining the sentiment and emotion visualizations generated by the application. During this process, user interactions are observed to determine whether the workflow is intuitive, navigation is easy to follow, and results can be interpreted without additional explanation [63]. Through this evaluation, the usability of the system is assessed to confirm that it supports practical exploration of audience emotional patterns across different film genres [64].

3. RESULTS AND DISCUSSION

This section presents the main findings along with the corresponding discussion. The results are presented through tables and figures to illustrate the model's performance in analyzing audience sentiment and emotions across film genres. The analysis is structured into three stages: sentiment classification, emotion classification, and exploration of the relationship between genre and emotional intensity. Each set of results is discussed within its respective context to explain how film genres influence audience emotional patterns. In addition, several interactive visual outputs generated by the web-based application including the frontend interface and three emotion–genre charts are presented to demonstrate the implementation outcomes.

3.1. Sentiment Classification Results

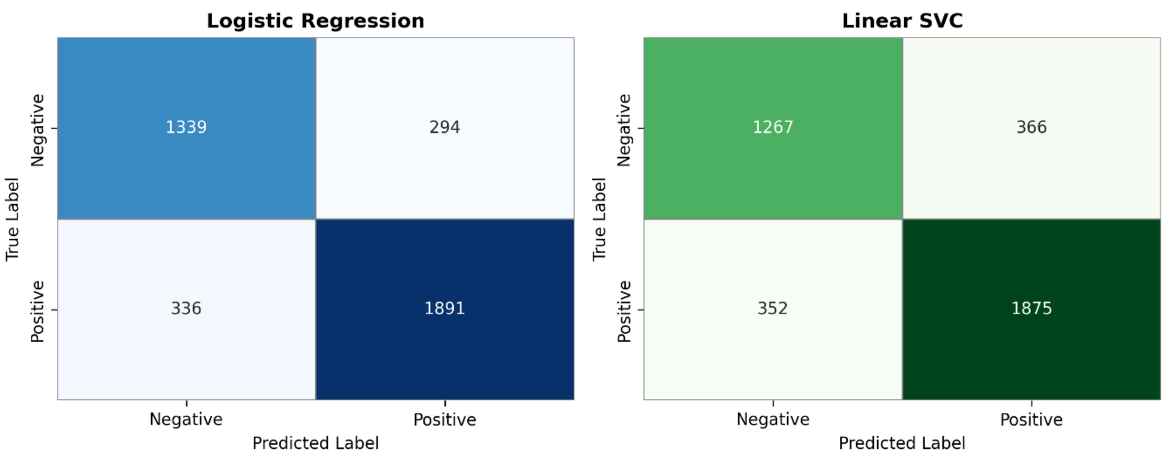
Two machine-learning models, Logistic Regression and Linear Support Vector Classifier (Linear SVC), were employed for sentiment classification. Table 4 presents the results of the sentiment model evaluation, including precision, recall, F1-score, weighted average, macro average, accuracy, and balanced accuracy values. These metrics provide a comprehensive overview of the models' ability to distinguish between positive and negative reviews in the IMDb dataset.

Table 4. Evaluation of sentiment classification models

Model	Class	Precision	Recall	F1-Score	Support
Logistic Regression	Negative	0.80	0.82	0.81	1,633
	Positive	0.87	0.85	0.86	2,227
	Weighted Avg	0.84	0.84	0.84	3,860
	Macro Avg	0.83	0.83	0.83	3,860
	Balanced Accuracy	0.83			3,860
	Accuracy	0.84			3,860
Linear SVC	Negative	0.78	0.78	0.78	1,633
	Positive	0.84	0.84	0.84	2,227
	Weighted Avg	0.81	0.81	0.81	3,860
	Macro Avg	0.81	0.81	0.81	3,860
	Balanced Accuracy	0.81			3,860
	Accuracy	0.81			3,860

Based on Table 4, both models demonstrate consistent performance, with accuracy values exceeding 80% in both cases. Logistic Regression achieved an accuracy of 84% and an F1-score of 0.86 for positive reviews, while Linear SVC attained an accuracy of 81% and an F1-score of 0.84. These results indicate that both models are capable of accurately distinguishing sentiment polarity, with Logistic Regression showing a slight advantage in terms of the balance between precision and recall.

Figure 3 presents the confusion matrices for both models, illustrating their effectiveness in identifying positive and negative sentiment classes. This visualization complements the quantitative evaluation by highlighting the balance of classification performance across both sentiment categories.



.

Figure 3. Confusion Matrices of Logistic Regression and Linear SVC for Sentiment Classification

The confusion matrices for both models are presented in Figure 3. Logistic Regression correctly identified 1,339 negative and 1,891 positive reviews, while Linear SVC correctly classified 1,267 negative and 1,875 positive reviews. Both models exhibited a low rate of misclassification (less than 10% of total predictions), indicating balanced and stable performance in sentiment polarity recognition prior to subsequent genre-based analysis.

Figure 3 not only reports the number of correct predictions but also illustrates how classification errors are distributed between positive and negative sentiment classes. Both Logistic Regression and Linear SVC show a relatively even distribution of errors, suggesting that neither model is biased toward over-predicting one sentiment class at the expense of the other. Such balance is critical for subsequent genre-based analyses, as skewed sentiment predictions could lead to inaccurate inferences regarding audience genre preferences. Accordingly, the confusion matrices reinforce the evaluation results presented in Table 4 by confirming that both models provide stable and reliable sentiment predictions.

Figure 4 compares the F1-scores across the ten most frequent genres, highlighting variations in model performance among different film categories. This comparison demonstrates the consistency of the models in capturing sentiment changes across genres and supports the interpretation of genre-based sentiment tendencies.

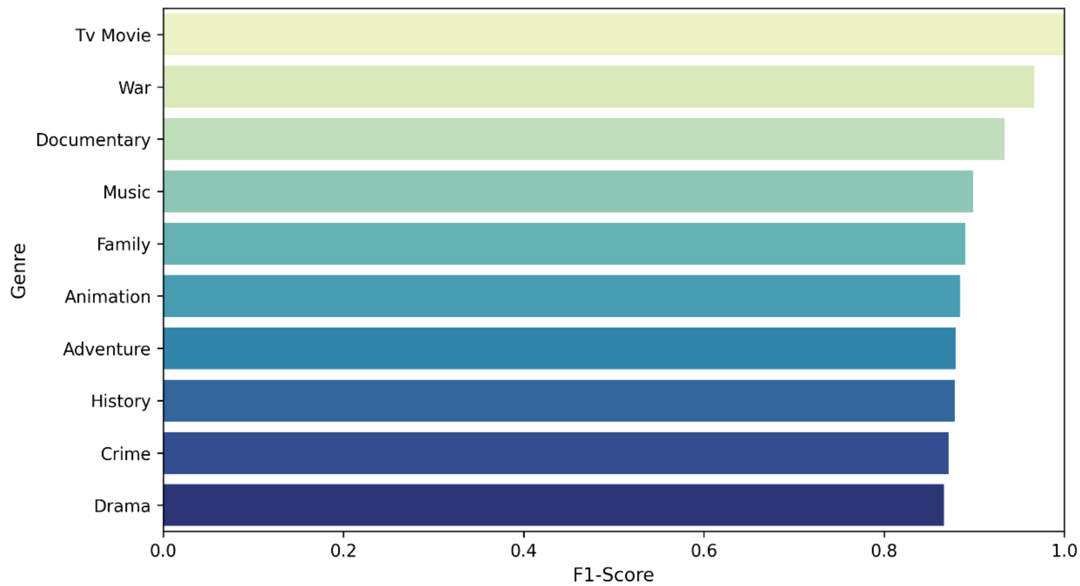


Figure 4. Comparison of F1-Score Across Genres for Sentiment Classification

Figure 4 presents a comparison of F1-scores across the ten most frequent genres. TV Movie (1.00), War (0.96), and Documentary (0.93) achieved the highest scores, indicating strong model performance in identifying sentiment within genres that tend to exhibit well-defined and structured narrative patterns. Slightly lower F1-scores were observed for genres such as Drama (0.86) and Romance (0.87), suggesting that sentiment prediction accuracy may be influenced by greater narrative complexity and variability in emotional expression across these genres.

The results of the sentiment classification stage are further illustrated in Figure 5, which depicts the distribution of positive and negative reviews across different film genres. This visualization provides an initial overview of how audience sentiment varies by genre, facilitating the identification of genres that are more frequently associated with positive or negative evaluations. The figure also highlights overall sentiment patterns that are examined in greater depth in the subsequent discussion.

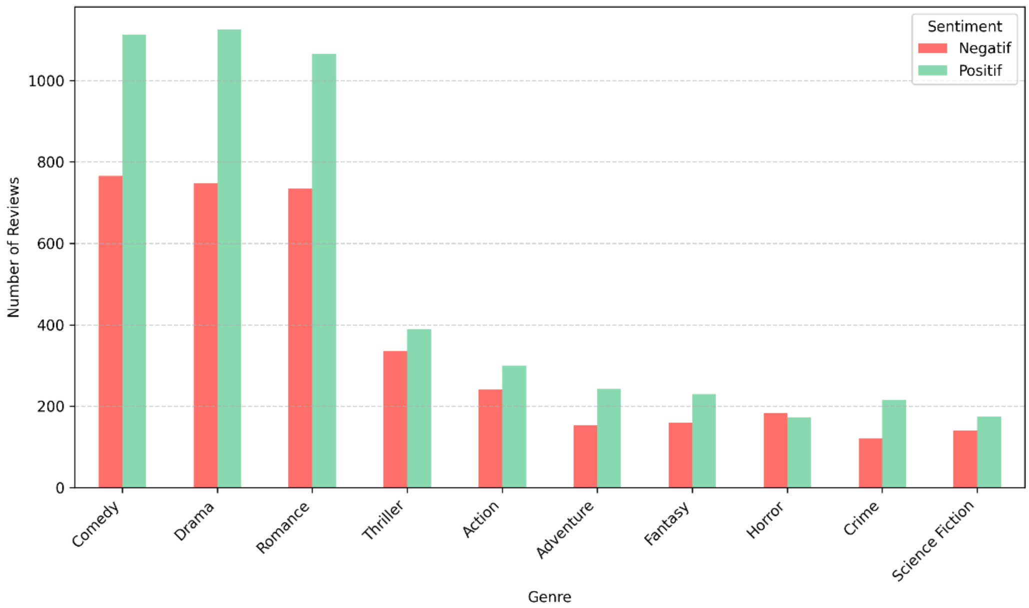


Figure 5. Sentiment Distribution across Genres

Based on Figure 5, the genres of Comedy (1,879 reviews), Drama (1,873), and Romance (1,800) recorded the highest numbers of positive reviews. Together, these three genres accounted for more than 5,000 positive reviews, reflecting a generally favorable audience response to lighthearted and emotionally driven storytelling. The number of positive reviews for Thriller (725) and Action (541) also showed an upward trend; however, the gap between positive and negative reviews in these genres was less pronounced than in the three dominant genres. A similar positive tendency was observed for Adventure, Fantasy, Crime, and Science Fiction, whereas Horror was the only genre with a slightly higher number of negative reviews, suggesting a more divided audience response.

The performance of the proposed models is broadly consistent with findings from previous studies, in which Logistic Regression achieved accuracy levels of approximately 80.61% and SVM reached 82.42%, while TF-IDF combined with SVM attained 82.51%. Other studies employing optimization techniques or hybrid feature representations have reported accuracy levels exceeding 90%. Although the accuracy achieved in this study is comparatively lower, the proposed approach emphasizes interpretability and practical applicability by leveraging basic TF-IDF features to examine how film genres shape audience emotional responses.

3.2. Emotion Classification Results

In addition to sentiment classification, this study also evaluates audience emotion classification. The model employed for this task is Logistic Regression using a multi-label One-vs-Rest classification approach. This method allows each review to be associated with more than one relevant emotion label, thereby providing a more comprehensive representation of the audience's emotional spectrum. The results of the emotion classification evaluation are presented in Table 8, which reports the precision, F1-score, and support values for each emotion class.

Table 5. Evaluation of emotion classification models

Model	Class	Precision	Recall	F1-Score	Support
One-vs-Rest Logistic Regression	Anger	0.84	0.19	0.31	251
	Anticipation	0.65	0.21	0.31	514
	Disgust	0.88	0.05	0.10	134
	Fear	0.58	0.33	0.42	315
	Joy	0.70	0.31	0.43	663
	Optimism	0.73	0.18	0.29	420
	Sadness	0.70	0.52	0.60	1,562
	Micro Avg	0.69	0.35	0.47	3,860
	Macro Avg	0.72	0.25	0.35	3,860
	Weighted Avg	0.70	0.35	0.45	3,860
	Samples Avg	0.34	0.35	0.35	3,860
	Hamming Loss	0.12			
	Jaccard Index	0.34			3,860
	Accuracy	0.34			3,860

.

Table 5 presents the evaluation results for each emotion category using the One-vs-Rest Logistic Regression model. Among the emotion classes, Sadness achieved the highest F1-score (0.60), indicating that the model performed relatively well in identifying sadness-related emotional expressions in audience reviews. Joy and Fear also exhibited more stable F1-scores compared to emotion categories with fewer instances, such as Disgust and Optimism.

Overall multi-label performance is reflected in the Micro and Macro F1-scores of 0.47 and 0.35, respectively, providing a more balanced assessment across emotion categories. The Hamming loss value of 0.12 indicates a relatively low rate of mislabeling across the seven emotion classes. The reported accuracy of 0.34 is included for completeness; however, accuracy is less informative in multi-label classification, as a prediction is considered correct only when all emotion labels for a given review are correctly identified. Consequently, the Hamming loss and F1-scores offer a more appropriate representation of the model's capability to predict multiple emotions within a single review text.

The model's emotion recognition performance is further examined using the confusion matrix shown in Figure 9. This figure illustrates the predictions generated by the One-vs-Rest Logistic Regression model for four primary emotion categories Sadness, Joy, Anticipation, and Optimism which were selected due to their higher frequencies in the dataset and their representativeness of dominant emotional patterns.

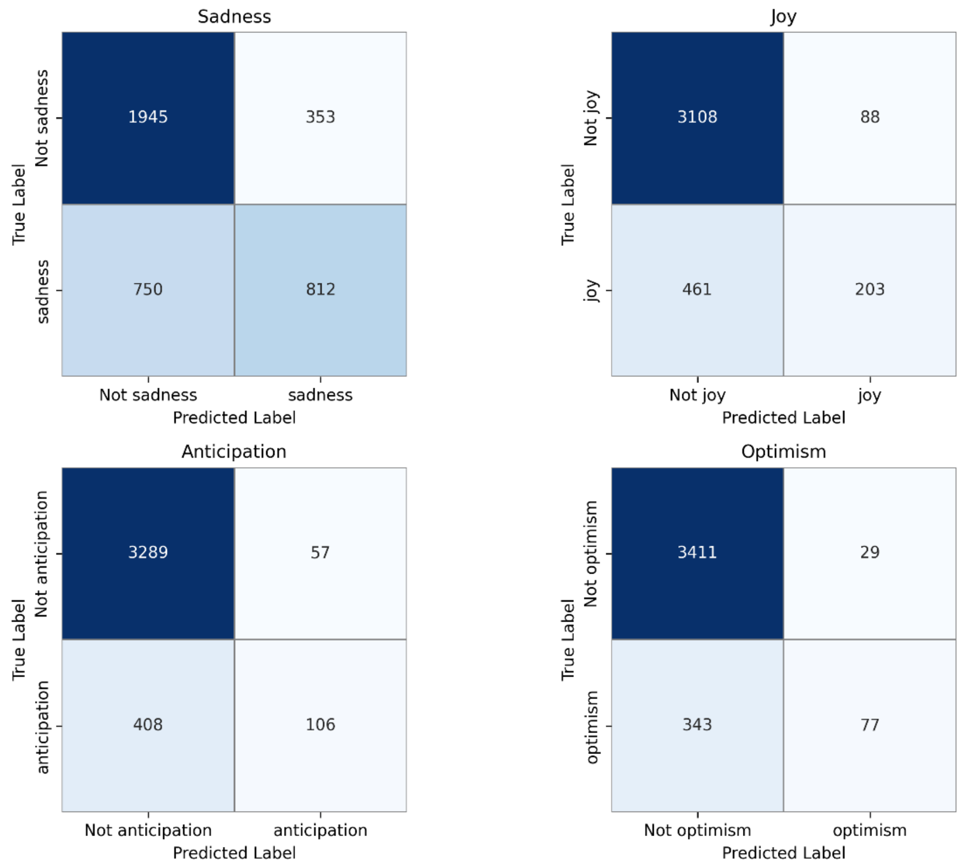


Figure 6. Confusion Matrix of One-vs-Rest Logistic Regression for Emotion Classification

As shown in Figure 6, the model correctly identified Sadness in 812 out of 1,562 reviews and Joy in 203 out of 664 reviews. In addition, it detected Anticipation in 106 out of 514 reviews and Optimism in 77 out of 420 reviews. These results indicate that the model performs more effectively for emotions that are explicitly expressed in the text, such as Sadness and Joy, compared to emotions that are more context-dependent, such as Anticipation and Optimism. This observation is consistent with the findings reported in Table 5, which show stronger performance for emotion classes characterized by clearer and more balanced textual cues.

Despite these strengths, several limitations should be considered when interpreting the overall performance of the emotion classification task. The distribution of emotion labels in the dataset is highly imbalanced; for example, Sadness and Joy occur far more frequently than emotions such as Disgust or Optimism. This imbalance may limit the model's ability to generalize across all emotion categories. Moreover, movie reviews often contain colloquial language, mixed emotions, and subtle expressions, which introduce noise and can reduce the accuracy of predicting context-dependent emotions. Some reviews are also minimally expressive in emotional terms, further increasing the difficulty of multi-label classification. While the model demonstrates satisfactory performance for dominant emotion classes, these factors suggest that its generalizability could be improved in future work through the use of more balanced emotion datasets and clearer emotion annotation schemes.

3.2.1. Emotion Distribution by Genre

The emotion classification results are further examined by analyzing how different emotional responses are distributed across the ten most frequent film genres. Figure 10 illustrates the distribution of emotions such as Joy, Sadness, Fear, Anticipation, and other categories across these genres. This visualization provides an initial overview of audience emotional expressions in relation to film genres and indicates that certain genres tend to evoke more intense emotional responses than others.

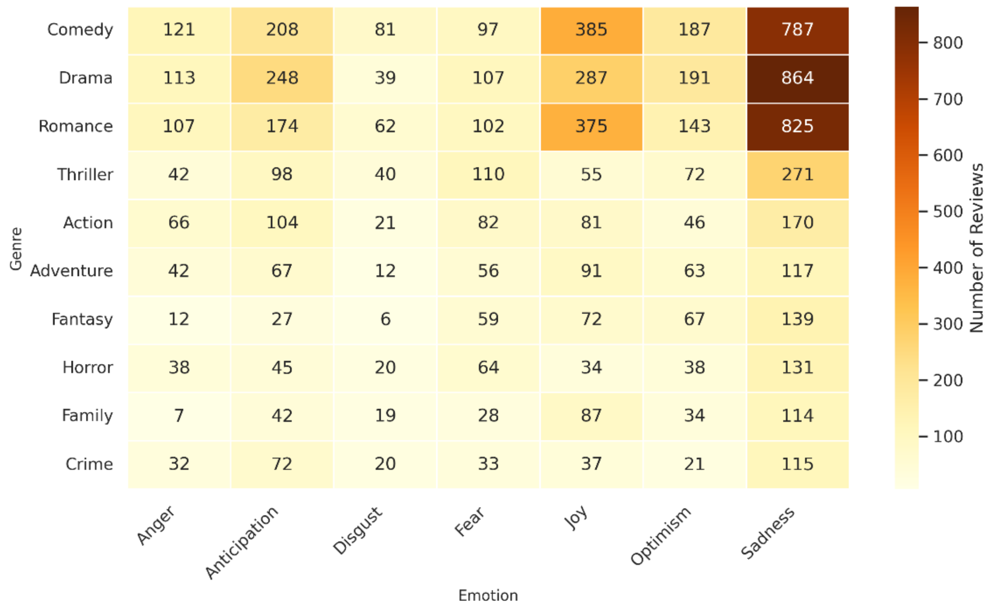


Figure 7. Heatmap of Emotion Distribution across Genres

Figure 7 illustrates the distribution of emotional responses across different film genres. Drama (1,866), Comedy (1,849), and Romance (1,788) recorded the highest numbers of emotional responses, whereas Thriller (688) and Action (570) exhibited moderate levels of emotional diversity. In contrast, genres such as Adventure, Fantasy, Horror, Family, and Crime showed fewer emotional responses (306–448), suggesting relatively more neutral audience reactions. These findings indicate that the intensity of emotional responses varies according to film genre and narrative tone.

Figure 7 further demonstrates that emotional expressions are not evenly distributed across genres. Drama, Comedy, and Romance display a wider range and higher frequency of emotional reactions, suggesting stronger audience emotional engagement. In contrast, genres with fewer emotional responses tend to elicit more neutral reactions, which may be associated with simpler narratives or a lower emphasis on emotional storytelling. These results support the research objective by confirming that film genre is a key factor influencing audience emotional responses.

Despite the presence of imbalanced emotion labels, the One-vs-Rest Logistic Regression model successfully identified multiple emotional patterns across genres. While previous studies have reported higher accuracy by combining textual, visual, and audio features or by applying optimized TF-IDF–based SVM approaches, the proposed model prioritizes simplicity and interpretability. This design enables effective exploration of genre-based emotional variations in movie reviews without relying on complex multimodal architectures.

3.2.2. Model Performance Comparison on Emotion Detection

To further examine the model’s capability to identify emotional states across different film genres, Figure 8 presents a comparison of F1-scores by genre. This visualization highlights the proficiency of the One-vs-Rest Logistic Regression model in capturing emotion-related language within user reviews. Moreover, the chart makes apparent how variations across genres influence the model’s ability to recognize and categorize emotional tendencies in textual data.

.

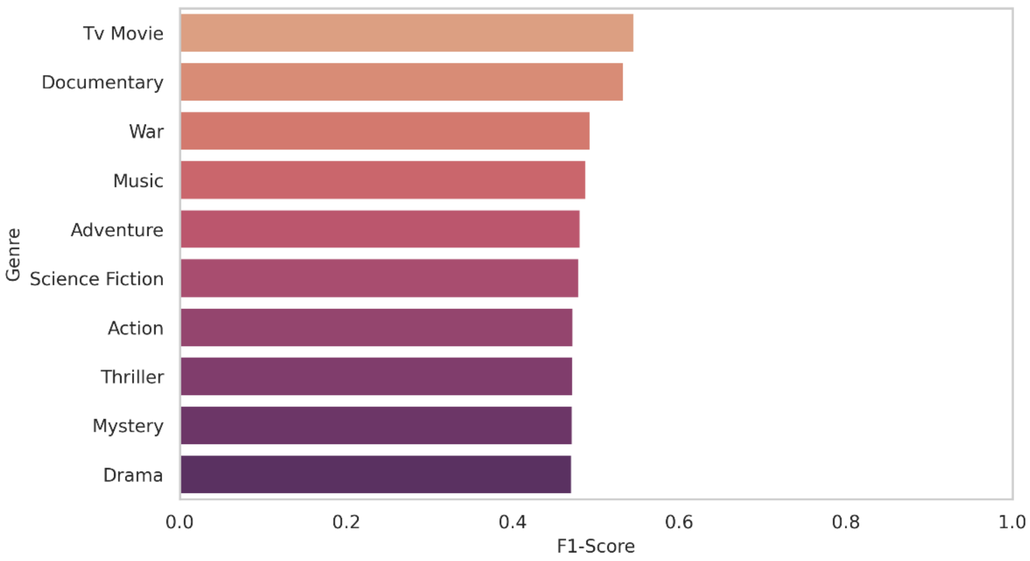


Figure 8. Comparison of F1-Score Across Genres for Emotion Classification

Figure 8 presents a comparison of F1-scores for emotion classification, illustrating model performance across different film genres. The TV Movie genre achieved the highest F1-score (0.55), followed by Documentary (0.53) and War (0.49). Genres such as Music, Adventure, and Science Fiction also demonstrated relatively strong performance, with F1-scores ranging from 0.48 to 0.49. In contrast, Action, Thriller, Mystery, and Drama exhibited slightly lower F1-scores (approximately 0.47), indicating that genres characterized by more complex and diverse emotional expressions may pose greater challenges for accurate emotion detection.

3.3. System Evaluation Results

In addition to evaluating the performance of the sentiment and emotion classification models, this study also assesses the developed web-based system as a whole. As the research extends beyond model development to include practical implementation, system evaluation is essential to ensure that analysis results can be effectively accessed and utilized by users. This evaluation examines whether the system operates as intended, supports the analysis process efficiently, and presents sentiment and emotion information in a clear and interpretable manner. The system evaluation encompasses both functional aspects and usability, providing insight into the system's effectiveness in delivering genre-based sentiment and emotion analysis to end users.

3.3.1. Functional Evaluation of the Web-Based System

The functional evaluation was conducted to verify that the web-based application operates correctly and supports the intended sentiment and emotion analysis process. This evaluation examines whether the system's core features function as designed, including review input, sentiment and emotion prediction, genre selection, and visualization of analysis results. The system was tested using movie reviews from various genres to ensure that the prediction models run smoothly and produce consistent outputs. The evaluation results indicate that the system can process user input without errors, generate sentiment and emotion classifications accurately, and display genre-based visualizations in a timely manner. Overall, the system demonstrates reliable functionality and effectively supports the analysis of audience emotional responses across film genres.

3.3.2. Usability Evaluation

The usability evaluation was conducted to assess how easily users can interact with the web-based system when performing sentiment and emotion analysis. This evaluation focuses on interface clarity, ease of use of the core features, and users' ability to read and interpret the analysis results without additional explanation. Users were asked to input movie reviews, select genres, and examine the visual outputs generated by the system in order to observe how smoothly the analysis process operates. The results indicate that users were able to follow the analysis steps and understand the sentiment and emotion outputs with minimal difficulty. Overall, the system provides a clear and user-friendly interface that supports practical usage and facilitates exploration of audience emotional patterns across different film genres.

3.3.3. Discussion of System Evaluation Results

Based on the system evaluation results, the developed web-based application demonstrates reliable functionality and satisfactory usability in supporting genre-based sentiment and emotion analysis. The functional evaluation confirms that the system's core features operate as intended, from processing user input to generating sentiment and emotion predictions and presenting results through visualizations. These findings indicate that the proposed models can be effectively implemented within a practical system and used without technical interruptions.

The usability evaluation further indicates that users can interact with the system comfortably and understand the analysis process without requiring additional explanation. The interface design and workflow enable users to input movie reviews, select genres, and interpret sentiment and emotion outputs with ease. This aspect is particularly important, as the system is intended not only for technical users but also for practitioners such as filmmakers, content creators, and decision-makers who require clear and accessible insights.

Overall, the system evaluation supports the conclusion that the developed application is suitable for real-world use and effectively complements the model evaluation results. Although the evaluation was conducted at a general level, it provides sufficient evidence that the system is functional, user-friendly, and capable of delivering meaningful insights into audience emotions based on film genres. Future work may involve more detailed usability testing with a larger user group; however, the current findings already demonstrate that the system fulfills the objectives of this study.

3.4. Model Comparison Results

To provide a broader evaluation of the proposed approach, this study compares the performance of TF-IDF–based models with transformer-based and deep learning models under different experimental settings. The comparison is presented in Table 6, which reports F1-macro, F1-micro, and Hamming loss values. These metrics were selected to more accurately reflect multi-label classification performance, rather than relying on accuracy alone. Based on the results in Table 6, the TF-IDF models combined with Logistic Regression and Linear SVC demonstrate stronger performance than the zero-shot BERT model, particularly in terms of overall F1-scores and label-level error rates.

Table 6. Model Performance Comparison Between Conventional Machine Learning and Deep Learning

Model	F1 Macro	F1 Micro	Hamming Loss
Logistic Regression	0.83		
Linear SVC	0.81		
One-vs-Rest Logistic Regression	0.35	0.47	0.12
BERT (Zero-Shot)	0.21	0.29	0.57
CNN (Supervised)	0.05	0.17	0.15
LSTM (Supervised)	0.04	0.09	0.15

The results presented in Table 6 indicate that TF-IDF combined with linear models consistently outperforms the zero-shot BERT model. The One-vs-Rest Logistic Regression approach achieves an F1-micro score of 0.47 with a Hamming loss of 0.12, demonstrating its ability to effectively identify multiple emotion labels within a single review. In contrast, the zero-shot BERT model records a substantially lower F1-micro score of 0.29 and a higher Hamming loss of 0.57, indicating a high rate of mislabeling and limited stability across emotion categories. Although transformer-based models generally achieve superior performance when fine-tuned, these findings demonstrate that the proposed approach remains effective without incurring high computational costs. Moreover, TF-IDF-based linear models provide clear and interpretable decision boundaries, facilitating the analysis of emotional patterns within genre-specific contexts. This interpretability aligns with the study's objective of delivering actionable insights for filmmakers and content creators.

In addition, supervised deep learning models based on CNN and LSTM architectures were evaluated in response to reviewer suggestions regarding the inclusion of more recent classification methods. As shown in Table 6, the CNN model achieved an F1-macro score of 0.05 and an F1-micro score of 0.17, with a Hamming loss of 0.15. Similarly, the LSTM model obtained an F1-macro score of 0.04 and an F1-micro score of 0.09, with the same Hamming loss value. These results suggest that, under the current experimental conditions, both CNN and LSTM models struggle to effectively capture complex and overlapping emotional expressions in multi-label movie reviews. This limitation is likely attributable to the relatively limited amount of labeled emotion data and the diverse linguistic expressions of emotions in written reviews. In comparison, TF-IDF–based linear models yield more consistent performance with lower error rates, while remaining more interpretable and computationally efficient for real-time, genre-based emotion analysis.

.

3.5. Genre-Emotion Relationship Exploration

To examine how audience emotions vary across different film genres, Figure 9 visualizes the relationship between genre categories and emotional intensity using a heatmap. This visualization highlights genres that elicit stronger emotional responses such as Joy, Fear, Sadness, and Anticipation while indicating that other genres exhibit a more even distribution of emotions. The heatmap provides a clear representation of the impact of genre on audience emotional engagement and offers an overview of prevailing emotional patterns associated with different types of films.

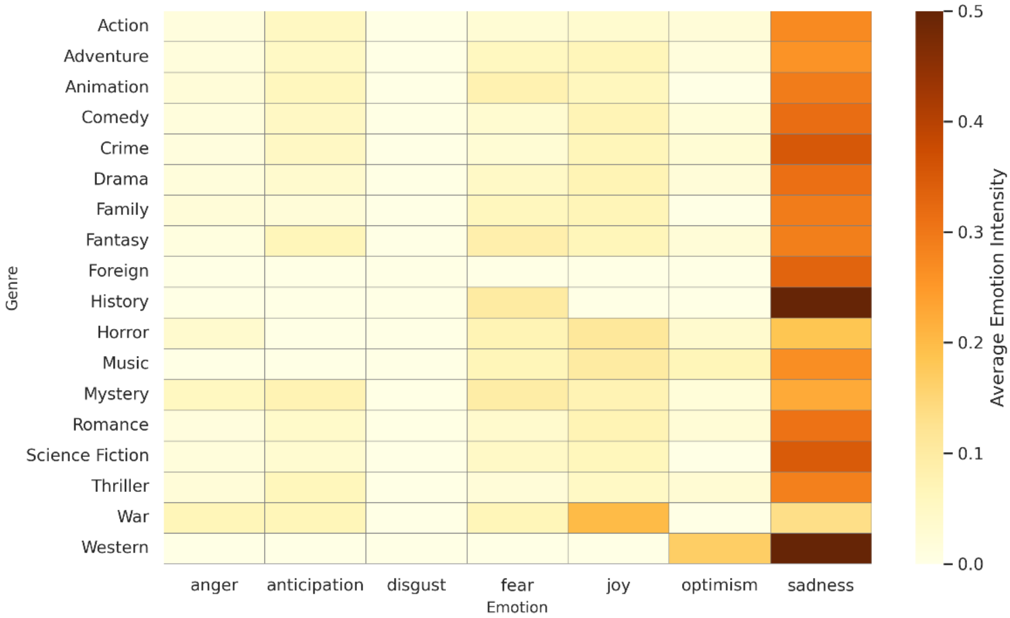


Figure 9. Heatmap of Genre and Emotion Relationships

Figure 9 illustrates the relationship between film genres and seven emotion categories using a heatmap, where darker colors represent higher emotional intensity. The visual patterns, together with the average intensity values, indicate that the Western genre exhibits the highest overall emotional intensity (0.095), driven primarily by sadness and anticipation. History and Mystery follow with average intensity values of 0.085 and 0.078, respectively, reflecting similarly deep emotional response patterns. Conflict-oriented genres such as Crime, Fantasy, and War also fall within the higher-intensity range, with average values of approximately 0.076. In contrast, more lighthearted genres, including Comedy, Animation, and Science Fiction, display lower emotional intensity across most categories, with average values of around 0.072. This distribution suggests that films characterized by conflict-driven narratives or historical themes tend to elicit stronger emotional reactions, whereas lighter genres are associated with more moderate emotional responses, as reflected by the differing color patterns in the heatmap.

Overall, these findings indicate that emotionally intense genres generally foster higher levels of audience engagement, while lighter genres tend to evoke more neutral emotional responses. Compared with multimodal and hybrid TF-IDF–SVM approaches that report higher accuracy, this study proposes a straightforward and interpretable text-based framework that effectively maps genre-driven emotional variations, offering practical insights for producers and digital platforms.

3.6. Web-Based Application Interface Results

To demonstrate the practical implementation of the proposed model, Figure 10 presents the main interface of the developed web-based system. This interface illustrates how the genre-based sentiment and emotion analysis framework is deployed in an interactive application that users can access directly. It represents the final stage of the CRISP-DM process, in which the analytical model is implemented within a functional system that links audience reviews to genre-based sentiment and emotion analysis.

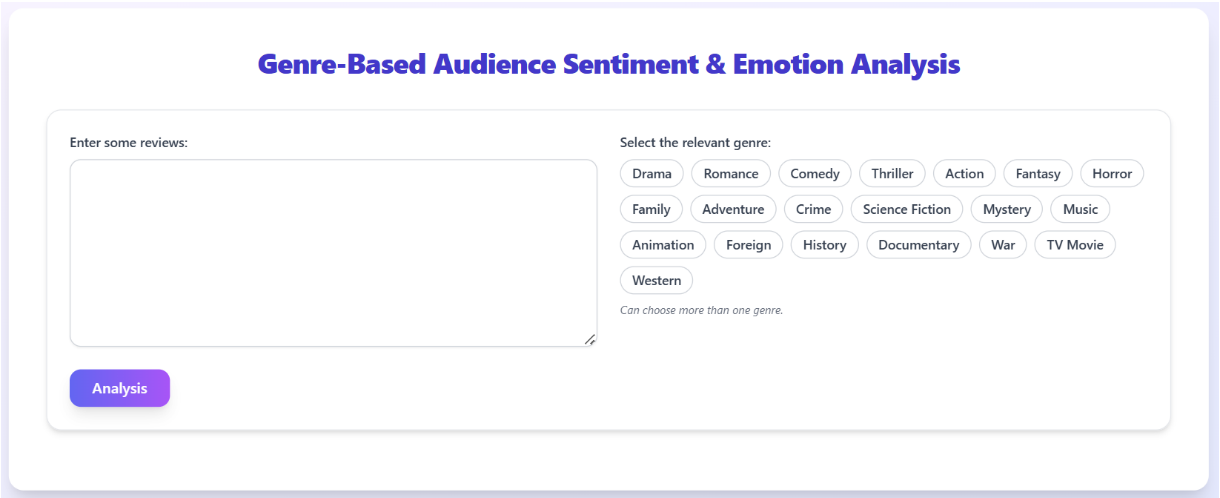


Figure 10. Main Input Interface of the Web-Based System

Figure 10 presents the main interface of the developed web-based system. The backend is implemented using Flask, while the frontend is built with React and Tailwind CSS. On the left side of the interface, a large text input area allows users to enter multiple movie reviews simultaneously. On the right side, users can select one or more movie genres, such as Drama, Romance, Comedy, and Action. After entering the reviews and selecting the relevant genres, users can initiate the analysis by clicking the *Analysis* button. At this stage, the input data are transmitted to the backend model, which performs sentiment and emotion classification for the selected genres. This interface serves as the connection between users and the analytical model, facilitating a clear and user-friendly exploration of how genre expectations influence audience emotions.

To illustrate how the analysis results are presented, Figure 11 displays the results interface of the web-based system. This stage follows the submission of movie reviews and genre selection by the user. The figure demonstrates how the frontend presents the analysis output generated by the backend model in a structured and intuitive manner, reflecting the final step of system implementation in which audience reviews are processed and analyzed in near real time.

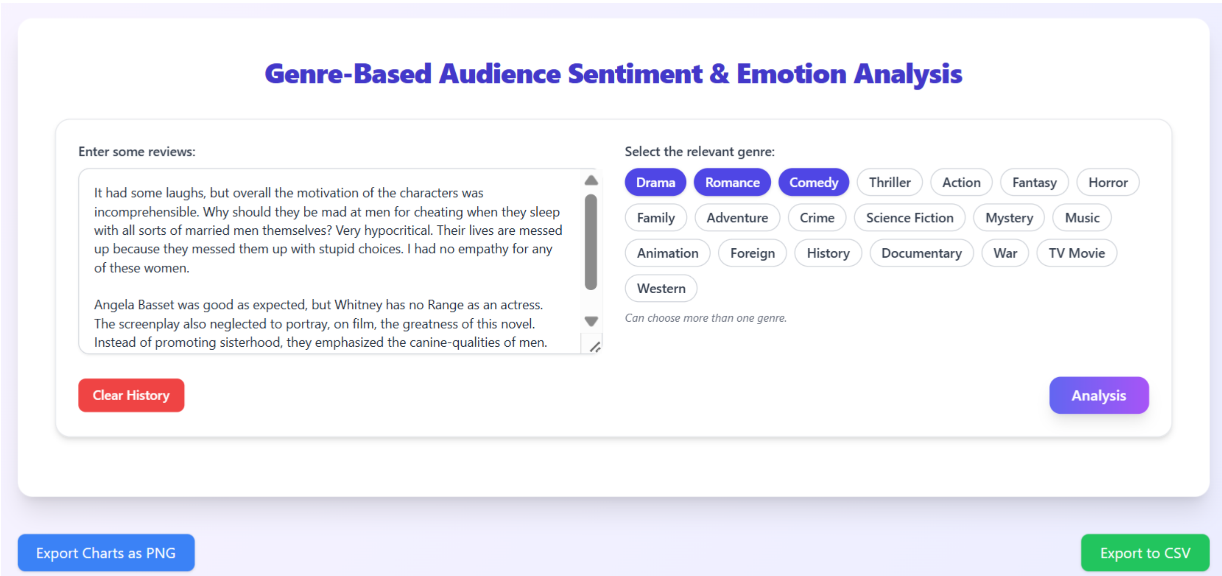


Figure 11. Result Interface of the Web-Based System

Figure 11 presents the main results interface displayed after the sentiment and emotion analysis has been completed. Once processing is finalized on the backend, the frontend dynamically updates to present the analyzed data without requiring a page refresh. The interface includes functions such as *Clear History* to remove previous inputs and *Export Charts as PNG* or *Export to CSV* options, which allow users to save and review the generated outputs at a later time. This results interface demonstrates how the integration of Flask and React effectively transforms the

.

analytical model into a functional and interactive web application. It also highlights the practical contribution of this research by making genre-based sentiment and emotion analysis accessible and interpretable through a clean, responsive, and user-friendly web system.

3.7. Global Emotion Aggregation Result

To illustrate the overall emotional patterns identified through the web-based system, Figure 12 presents the aggregated emotion results derived from user-submitted movie reviews. This visualization depicts the average intensity of each detected emotion, summarizing how the system interprets audience expressions and providing an overall view of the dominant emotions that emerge after all reviews are processed by the model.

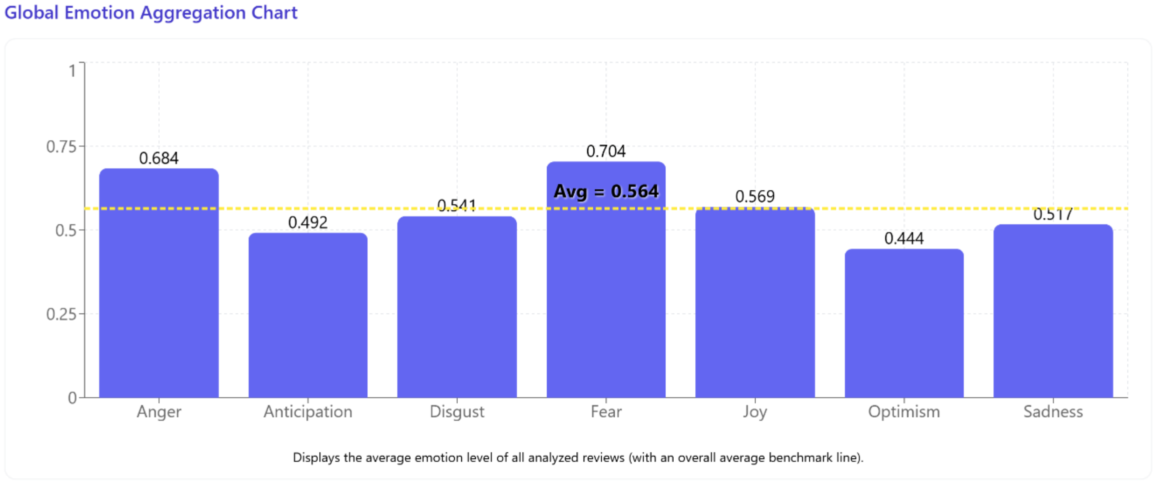


Figure 12. Overall Emotion Intensity Distribution Chart

Figure 12 presents the average emotion intensity values, with a reference line at 0.564 indicating the overall emotional mean across all reviews. Fear (0.704) and Anger (0.684) exhibit the highest intensity levels, suggesting that expressions of tension and conflict dominate audience reactions. These are followed by Joy (0.569) and Disgust (0.541), indicating frequent occurrences of excitement and negative evaluative responses. Meanwhile, Sadness (0.517) and Anticipation (0.492) remain close to the average intensity level. In contrast, Optimism (0.444) falls below the mean, indicating that this emotion occurs less frequently. Overall, the findings suggest that audience engagement is more strongly associated with intense and suspense-driven emotions than with lighter emotional states, which is consistent with the study's emphasis on the influence of genre context on emotional responses.

3.8. Emotion Trend Results

To observe the emotional dynamics across multiple user reviews, Figure 13 presents the emotion trend visualization that tracks how each emotion changes over ten sequential reviews. This figure gives a concise view of how different emotional categories fluctuate over time as users express their opinions about films through the system.

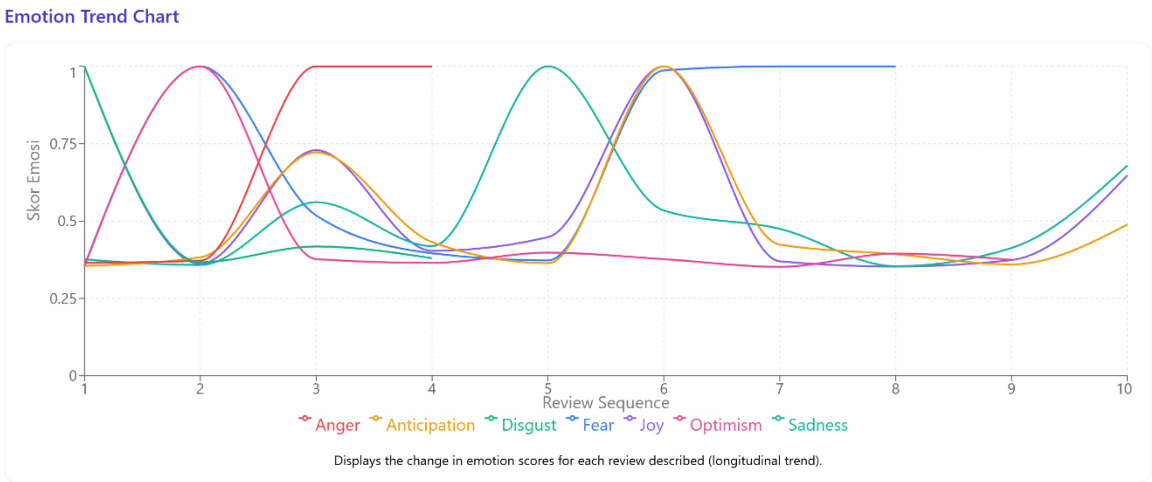


Figure 13. Emotion Dynamics across Reviews Chart

Figure 13 indicates that emotion patterns are not static but change dynamically across reviews. In particular, Joy, Fear, and Anticipation exhibit multiple peaks, representing moments of heightened user engagement and emotional intensity. Anger also shows a sharp increase and maintains a pronounced peak around the middle of the sequence, reflecting another period of intensified emotional expression. In contrast, Sadness and Optimism display more moderate fluctuations, although Optimism demonstrates a notable early peak. Disgust similarly shows a strong early peak, followed by several smaller increases throughout the review sequence.

Overall, this trend illustrates the dynamic nature of audience emotional responses, which fluctuate as users share their viewing experiences. These variations indicate that emotional reactions are situational and strongly influenced by the specific content and context of the films being discussed.

3.9. Genre-Emotion Correlation Results

To illustrate how different emotions manifest across film genres, Figure 14 presents the correlation results between genre categories and emotion intensity levels. This visualization depicts the extent to which specific emotions such as Joy, Fear, and Sadness are associated with particular genres. It provides a comparative overview of emotional tendencies, highlighting the emotions that predominate within certain genres based on user-submitted reviews processed by the system.

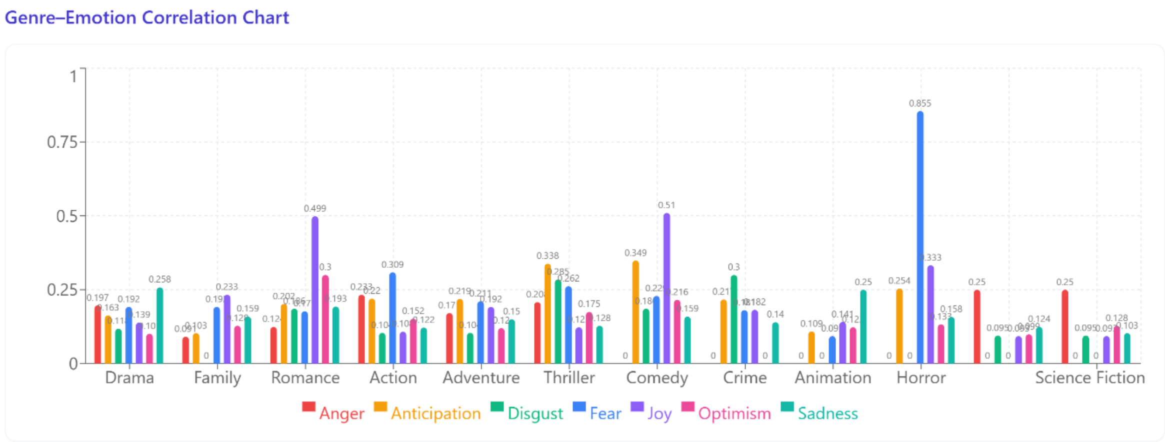


Figure 14. Emotion Distribution across Movie Genres Chart

Figure 14 demonstrates that certain film genres exhibit stronger emotional associations than others. In particular, Horror shows the strongest association with Fear (0.855), indicating that this genre most intensively evokes tension and suspense. Comedy, in contrast, reaches its highest intensity in Joy (0.510), reflecting its primary function of eliciting amusement and laughter. Drama and Romance display more balanced emotional profiles, with Disgust occurring more frequently in Drama (0.258) and Joy being most prominent in Romance (0.499), highlighting their emphasis on emotional depth and interpersonal relationships. Action and Thriller exhibit moderate emotional intensity, with Fear in Action (0.309) and Anticipation in Thriller (0.338), suggesting that emotional responses in these fast-paced genres are more varied rather than dominated by a single emotion. Overall, these findings confirm that different genres are characterized by distinct emotional patterns shaped by their narrative structure and tonal emphasis.

The strong association between Horror and Fear is intuitive and consistent with genre expectations. In contrast, Action films demonstrate a more nuanced emotional pattern in which Fear and Joy co-occur. Fear reflects the tension, conflict, and hazardous situations typical of action-driven narratives, whereas Joy emerges during moments of resolution, such as the protagonist’s victory or the defeat of an antagonist. This combination indicates that Action films do not rely solely on suspense and danger but also generate positive emotional responses through narrative resolution and emotional release. Consequently, emotional reactions in Action films are not dominated by a single emotion; instead, they reflect a more complex emotional structure that adds depth to the genre–emotion relationship.

3.10. Sentiment and User Review Analysis Results

To illustrate how the system generates analysis outputs for individual user reviews, Figure 15 presents examples of the per-review analysis results displayed on the web interface. Each output includes the input review text, the selected film genres, the predicted sentiment polarity, and the most prominent detected emotions visualized using a horizontal bar chart. This feature demonstrates the model's ability to process diverse textual inputs and produce sentiment and emotion predictions that vary according to both the review content and the selected genres.

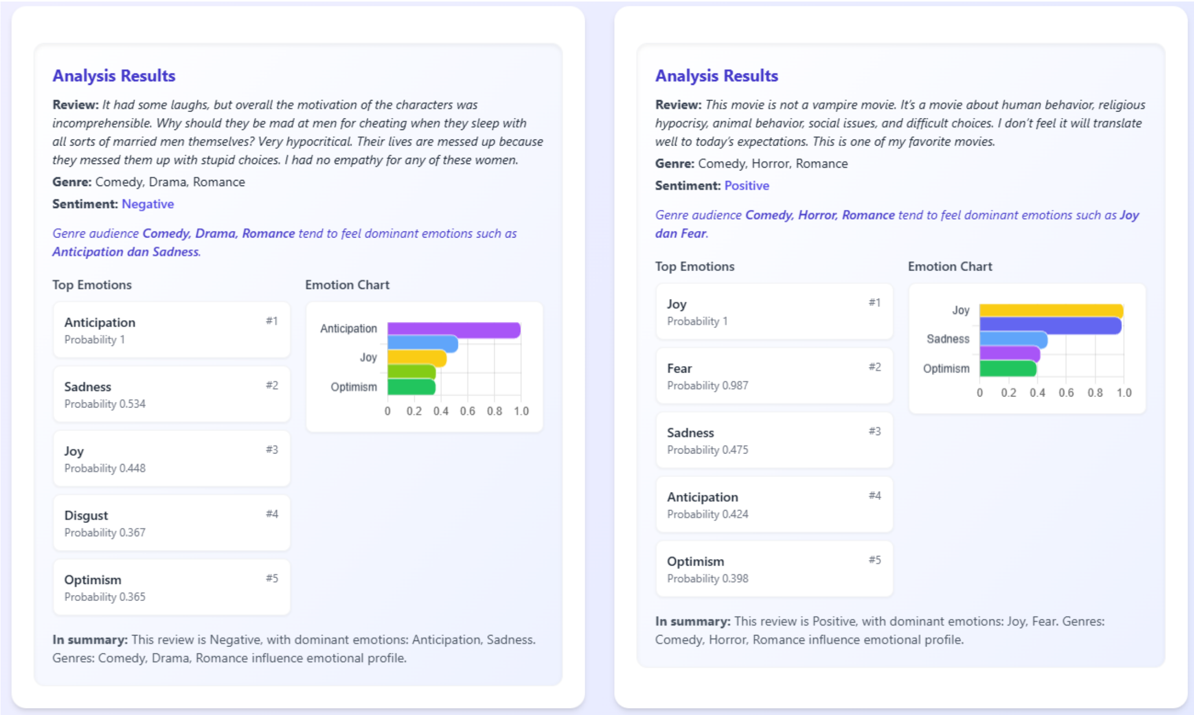


Figure 15. Sentiment and Emotion Analysis Outputs from User Reviews

Figure 15 presents two example analysis results that illustrate how review wording and genre context influence sentiment and emotion predictions. In the first review, the user writes: "It had some laughs, but overall the motivation of the characters was incomprehensible. Why should they be mad at men for cheating when they sleep with all sorts of married men themselves? Very hypocritical. Their lives are messed up because they messed them up with stupid choices. I had no empathy for any of these women." For the genres Comedy, Drama, and Romance, the system predicts a negative sentiment, which is consistent with the strongly evaluative language used in the review, including expressions such as "very hypocritical," "messed up," and "I had no empathy." The dominant emotion, Anticipation (1.000), reflects unmet expectations, as the reviewer initially notes that the movie "had some laughs," suggesting an expectation of entertainment that was ultimately not fulfilled. Sadness (0.534) is supported by the expressed emotional detachment in "I had no empathy," while Joy (0.448) appears at a lower level, corresponding to the brief acknowledgment of humor despite the overall negative tone.

In the second review, the user states: "This movie is not a vampire movie. It's a movie about human behavior, religious hypocrisy, animal behavior, social issues, and difficult choices. I don't feel it will translate well to today's expectations. This is one of my favorite movies." For the genres Comedy, Horror, and Romance, the system predicts a positive sentiment, strongly supported by the explicit statement "one of my favorite movies." The dominant emotion, Joy (1.000), aligns with this positive evaluation, while Fear (0.987) is associated with thematic elements such as "religious hypocrisy" and "difficult choices," which imply tension and moral conflict. Sadness (0.475) appears at a lower intensity and is reflected in the reviewer's concern that the film "will not translate well to today's expectations." Overall, these examples demonstrate that the system accurately identifies sentiment polarity and captures nuanced emotional responses arising from the interaction between review content and genre context.

3.11. Error Analysis Results

To provide a clearer understanding of the limitations of the proposed model, Figure 16 presents several qualitative error analysis examples obtained from the web-based system. This figure highlights review instances in which the predicted emotions do not fully align with the reviewer's intended meaning, particularly in cases involving mixed opinions or subtle emotional expressions. Each example shown in Figure 16 includes the original review text, selected genres, predicted sentiment, and the corresponding emotion distribution. These examples illustrate how specific review characteristics such as nuanced criticism or emotionally layered narratives can challenge the model's interpretive capability, thereby offering insight into where and why misclassifications occur in genre-based emotion analysis.

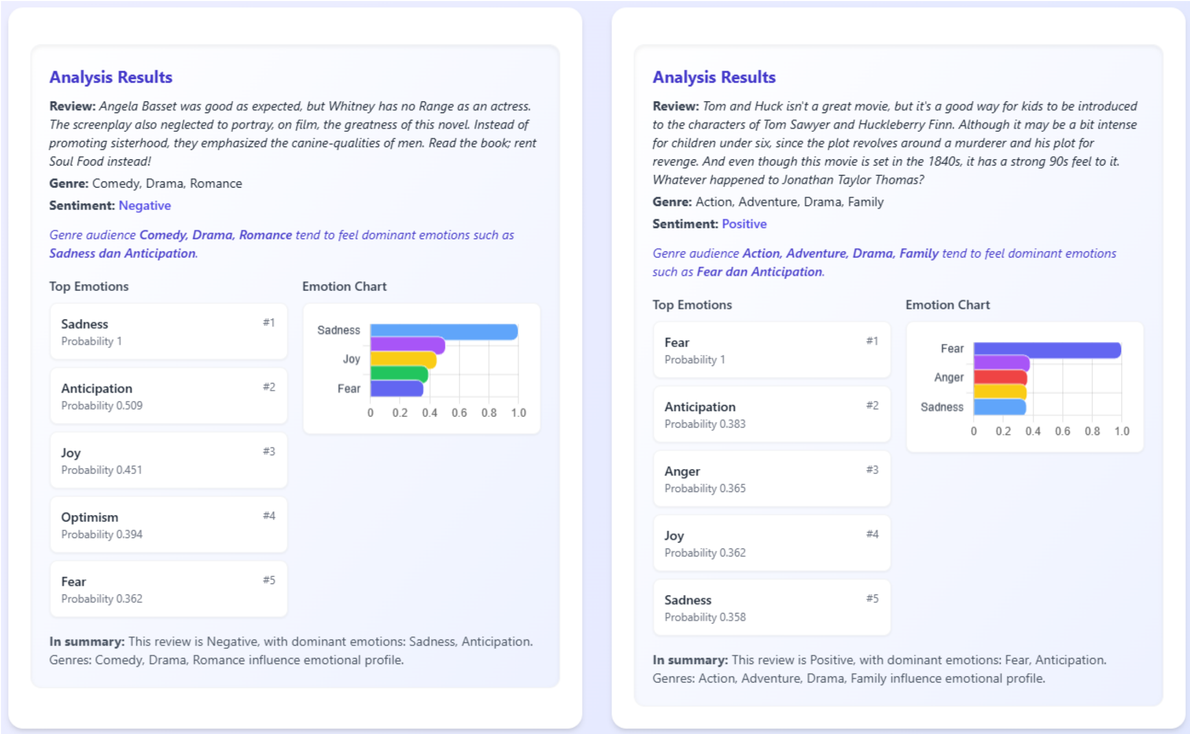


Figure 16. Qualitative Error Analysis Examples from User Reviews

Figure 16 presents two review examples that illustrate cases in which the proposed model struggles to fully capture the reviewer's intended emotional emphasis. In the first review, the user writes: "Angela Basset was good as expected, but Whitney has no range as an actress. The screenplay also neglected to portray, on film, the greatness of this novel. Instead of promoting sisterhood, they emphasized the canine qualities of men. Read the book; rent *Soul Food* instead!" Within the Comedy, Drama, and Romance genres, the system correctly predicts a negative sentiment, as supported by strongly critical expressions such as "has no range as an actress," "neglected to portray," and the dismissive suggestion "rent *Soul Food* instead." The dominant emotion, Sadness (1.000), reflects disappointment with the film's execution, while Anticipation (0.509) corresponds to unmet expectations implied by the phrase "good as expected." However, the relatively high Joy score (0.451) is incongruent with the overall sarcastic and critical tone of the review. This misclassification occurs because the model captures isolated positive lexical cues but fails to account for ironic and sarcastic context.

In the second review, the user states: "*Tom and Huck* isn't a great movie, but it's a good way for kids to be introduced to the characters of Tom Sawyer and Huckleberry Finn. Although it may be a bit intense for children under six, since the plot revolves around a murderer and his plot for revenge. And even though this movie is set in the 1840s, it has a strong 90s feel to it. Whatever happened to Jonathan Taylor Thomas?" Under the Action, Adventure, Drama, and Family genres, the sentiment is predicted as positive, supported by evaluative phrases such as "a good way for kids to be introduced." However, the dominant emotions Fear (1.000), Anticipation (0.383), and Anger (0.365) are largely driven by narrative-related terms such as "murderer," "revenge," and "intense," as well as the mildly critical closing question. In this case, the model places greater emphasis on threatening narrative elements than on the reviewer's overall balanced and informative intent.

Overall, these examples demonstrate that mixed sentiments, sarcasm, and strongly narrative-driven language pose challenges for the model in accurately capturing shifts in emotional emphasis. This limitation highlights the difficulty of modeling contextual and pragmatic meaning in genre-based emotion analysis using text-based approaches.

4. CONCLUSION

This study investigated the relationship between film genres and audience emotions by analyzing 46,173 IMDb movie reviews using machine-learning-based sentiment and emotion analysis. Logistic Regression, Linear Support Vector Classification (Linear SVC), and One-vs-Rest Logistic Regression models were evaluated using TF-IDF features within the CRISP-DM framework.

.

The experimental results indicate that Logistic Regression achieved the strongest overall performance, with an accuracy of 0.84 and a macro F1-score of 0.83, followed closely by Linear SVC with an accuracy of 0.81. In the multi-label emotion classification task, the One-vs-Rest Logistic Regression model successfully identified key emotional patterns, particularly Sadness, which achieved the highest F1-score, while Joy was detected at a moderate level despite the presence of class imbalance. Genre-level evaluation further showed that TV Movie (1.00), War (0.96), and Documentary (0.93) achieved the highest F1-scores, indicating that audience emotional expression varies substantially across different genre categories.

The visualization results and web-based implementation further demonstrate that emotionally intense genres are more strongly associated with audience engagement. Fear (0.704) and Anger (0.684) emerged as the most dominant emotions expressed in the reviews. The integrated Flask–React framework enables near real-time processing and visualization of user input, highlighting the practical applicability of the proposed approach. Future work may extend this study by incorporating deep learning techniques or multimodal data sources, such as visual and audio information, to enhance emotion recognition and provide a more comprehensive representation of audience emotional responses.

CONFLICT OF INTEREST STATEMENT

The Authors state no conflict of interest.

REFERENCES

[1] H. Nuraliza, O. N. Pratiwi, and F. Hamami, "Analisis Sentimen IMBd Film Review Dataset Menggunakan Support Vector Machine (SVM) dan Seleksi Feature Importance," *J. Mirai Manaj.*, vol. 7, no. 1, pp. 1–17, Aug. 2022, doi: https://doi.org/10.37531/mirai.v7i1.2222.

[2] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.

[3] B. Paneru, B. Thapa, and B. Paneru, "Sentiment analysis of movie reviews: A flask application using CNN with RoBERTa embeddings," *Syst. Soft Comput.*, vol. 7, pp. 1–9, Dec. 2025, doi: 10.1016/j.sasc.2025.200192.

[4] H. Shofiyyah, T. Indah, F. M. Jannah, and Natasya Alya Putri, "Pengaruh Jenis Tayangan Film Comedy Terhadap Mood Positif Pada Mahasiswa Akhir," *J. Psychol. Today*, vol. 1, no. 4, pp. 233–241, Dec. 2023, [Online]. Available: https://digamed.net/index.php/psychologytoday/article/view/197/31

[5] Ö. Çalışkan, "Film Genres and Emotions: Genres Watched by Netflix Türkiye Viewers and Emotion Variations During the Pandemic," *Erciyes İletişim Derg.*, vol. 10, no. 1, pp. 1–22, Jan. 2023, doi: 10.17680/erciyesiletisim.1154927.

[6] R. P. A. L. Fajar and T. Lestari, "Pengaruh Film Horor Terhadap Perkembangan Kepribadian Dan Emosional Anak," *J. Kaji. Gend. dan Anak*, vol. 5, no. 1, pp. 17–26, Jun. 2021, doi: 10.24952/gender.v5i1.3731.

[7] R. N. Masaoy and A. F. Ramdhan, "Kultivasi di Era Digital: Studi Kasus Pada Anak-anak yang Gemar Menonton Film Horor di Cijati Majalengka," *J. Mhs. Sos. Humaniora)*, vol. 1, no. 1, pp. 143–156, Sep. 2024, doi: 10.31949/jumash.v1i1.11330.

[8] I. L. Esa, Chairunnisa, S. R. Khoirunnisa, N. Mona, and S. F. N. Zaferina, "The Effect of Watching Horror Film on Health Children and Adolescents in Indonesia," in *The 6th International Conference on Vocational Education Applied Science and Technology (ICVEAST 2023)*, Atlantis Press SARL, 2023, pp. 521–532. doi: 10.2991/978-2-38476-132-6_46.

[9] M. I. Wardah and S. D. Putra, "Implementasi Machine Learning Untuk Rekomendasi Film Di Imdb Menggunakan Collaborative Filtering Berdasarkan Analisa Sentimen IMDB," *J. Manajamen Inform. Jayakarta*, vol. 2, no. 3, pp. 243–249, Jul. 2022, doi: 10.52362/jmijayakarta.v2i3.868.

[10] D. N. M. Tampubolon, V. V. Hulu, R. O. Sipahutar, and O. Sihombing, "Analisis Prediksi Genre Film Pada Internet Movie Database Indonesia Menggunakan Metode Long Short Term Memory," *J. Tekinkom (Teknik Inf. Dan Komputer)*, vol. 6, no. 2, pp. 290–298, Dec. 2023, doi: 10.37600/tekinkom.v6i2.925.

[11] D. A. Moreno-Galván, R. López-Santillán, L. C. González-Gurrola, M. Montes-Y-Gómez, F. Sánchez-Vega, and A. P. López-Monroy, "Automatic movie genre classification & emotion recognition via a BiProjection Multimodal Transformer," *Inf. Fusion*, vol. 113, pp. 1–15, Jan. 2025, doi: 10.1016/j.inffus.2024.102641.

[12] E. Mishra, P. Nikam, S. Vidhyadharan, and R. Cheruvalath, "An affect-based approach to detect collective sentiments of film audience: Analyzing emotions and attentions," *Acta Psychol. (Amst).*, vol. 230, pp. 1–8, Oct. 2022, doi: 10.1016/j.actpsy.2022.103736.

[13] P. S. Poshe, D. V. A. Bharadi, G. A. Yadav, and R. S. Nandiwadekar, "Movie Genre Prediction and Sentiment Analysis Using Natural Language Processing," *IRE Journals*, vol. 4, no. 11, pp. 63–68, May 2021, [Online]. Available: https://www.irejournals.com/paper-details/1702699

[14] Mureed Hussain and M. Naseer, "Comparative Analysis of Logistic Regression, LSTM, and Bi-LSTM Models for Sentiment Analysis on IMDB Movie Reviews," *J. Artif. Intell. Comput.*, vol. 2, no. 1, pp. 1–8, Jun. 2024, doi: 10.57041/6t2gfw23.

[15] J. Almeida, L. Vilaça, I. N. Teixeira, and P. Viana, "Emotion identification in movies through facial expression recognition," *Appl. Sci.*, vol. 11, no. 15, pp. 1–19, Jun. 2021, doi: 10.3390/app11156827.

[16] N. G. Ramadhan and T. I. Ramadhan, "Analysis Sentiment Based on IMDB Aspects from Movie Reviews using SVM," *Sink. J. dan Penelit. Tek. Inform.*, vol. 7, no. 1, pp. 39–45, Jan. 2022, doi: 10.33395/sinkron.v7i1.11204.

[17] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, Apr. 2024, doi: 10.34010/jamika.v14i1.11946.

[18] D. S. Ramdan, R. D. Apnena, and C. A. Sugianto, "Film Review Sentiment Analysis: Comparison of Logistic Regression and Support Vector Classification Performance Based on TF-IDF," *J. Appl. Intell. Syst.*, vol. 8, no. 3, pp. 341–352, Nov. 2023, doi: 10.33633/jais.v8i3.9090.

[19] M. A. Faridi *et al.*, "Sentimen Analisis Review Film Pada IMDb Menggunakan Algoritma Logistic Regression," *JITSI J. Ilm. Teknol. Sist. Inf.*, vol. 6, no. 2, pp. 132–139, Jun. 2025, doi: 10.62527/jitsi.6.2.422.

[20] A. Hermawan, R. Yusuf, B. Daniawan, and Junaedi, "Enhanced sentiment analysis and emotion detection in movie reviews using support vector machine algorithm," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 23, no. 1, pp. 138–146, Feb. 2025, doi: 10.12928/TELKOMNIKA.v23i1.26377.

[21] Y. Christian, "Analysis Study to Detect Student Learning Problems in Online Analysis Study to Detect Student Learning," *J. Theor.*

Appl. Inf. Technol., vol. 100, no. 5, pp. 1875–1884, Mar. 2022, [Online]. Available: http://www.jatit.org/volumes/Vol100No6/23Vol100No6.pdf

[22] M. I. Adrian and E. A. Laksana, “Optimalisasi Parameter Support Vector Machine dengan Algoritma PSO untuk Tugas Klasifikasi Sentimen Ulasan IMDb,” *J. Algoritm.*, vol. 22, no. 1, pp. 288–299, Jun. 2025, doi: 10.33364/algoritma/v.22-1.2306.

[23] Sutriawan, Muljono, Khairunnisa, Z. Alamin, T. A. Lorosae, and S. Ramadhan, “Improving Performance Sentiment Movie Review Classification Using Hybrid Feature TFIDF, N-Gram, Information Gain and Support Vector Machine,” *Math. Model. Eng. Probl.*, vol. 11, no. 2, pp. 375–384, Feb. 2024, doi: 10.18280/mmep.110209.

[24] Herman and Y. Christian, “Analisis Performa Akademik Mahasiswa Menggunakan Distributed Random Forest,” *J. Appl. Informatics Comput.*, vol. 6, no. 2, pp. 180–186, Dec. 2022, doi: https://doi.org/10.30871/jaic.v6i2.4854.

[25] Y. Christian and A. Ricardo Hannuella, “Analisa Clustering Untuk Mengklasifikasi Kinerja Akademik di Lembaga Pendidikan,” *Technol. J. Ilm.*, vol. 15, no. 4, pp. 775–786, Oct. 2024, doi: 10.31602/tji.v15i4.16325.

[26] A. Ocal, “Cinematic Narratives as Socio-Technical Systems : Emotion Mining and Script – Audience Emotional Fidelity,” *Systems*, vol. 13, no. 11, pp. 1–37, Nov. 2025, doi: 10.3390/systems13110994.

[27] N. Pavitha *et al.*, “Movie recommendation and sentiment analysis using machine learning,” *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 279–284, Jun. 2022, doi: 10.1016/j.gltp.2022.03.012.

[28] M. T. Razaq, D. Nurjanah, and H. Nurrahmi, “Analisis Sentimen Review Film Menggunakan Naive Bayes Classifier dengan Fitur TF-IDF,” in *e-Proceding of Engineering*, 2023, pp. 1698–1712. [Online]. Available: https://repositori.telkomuniversity.ac.id/home/catalog/id/181205/slug/analisis-sentimen-review-film-menggunakan-naive-bayes-classifier-dengan-fitur-tf-idf.html

[29] A. J. Abimanyu, M. Dwifebri, and W. Astuti, “Sentiment Analysis on Movie Review from Rotten Tomatoes Using Logistic Regression and Information Gain Feature Selection,” *Build. Informatics, Technol. Sci.*, vol. 5, no. 1, pp. 162–170, Jun. 2023, doi: 10.47065/bits.v5i1.3595.

[30] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Comput. Sci.*, vol. 181, pp. 526–534, Jan. 2021, doi: 10.1016/j.procs.2021.01.199.

[31] D. Drungilas, I. Ramašauskas, and M. Kurmis, “Emotion Recognition in Usability Testing : A Framework for Improving Web Application UI Design,” *Appl. Sci.*, vol. 14, no. 11, pp. 1–19, May 2024, doi: 10.3390/app14114773.

[32] F. Rehman, “IMDb Movie Reviews Genres Description and Emotions.” 2023. [Online]. Available: https://www.kaggle.com/datasets/fahadrehman07/movie-reviews-and-emotion-dataset

[33] M. Z. Naeem, F. Rustam, A. Mehmood, Mui-zzud-din, I. Ashraf, and G. S. Choi, “Classification of movie reviews using term frequency-inverse document frequency and optimized machine learning algorithms,” *PeerJ Comput. Sci.*, vol. 8, pp. 1–28, Mar. 2022, doi: 10.7717/PEERJ-CS.914.

[34] R. Ramachandran, S. Sudhir, and A. B. Unnithan, “Exploring the relationship between emotionality and product star ratings in online reviews,” *IIMB Manag. Rev.*, vol. 33, no. 4, pp. 299–308, Dec. 2021, doi: 10.1016/j.iimb.2021.12.002.

[35] V. V. N. Raju *et al.*, “Enhancing emotion prediction using deep learning and distributed federated systems with SMOTE oversampling technique,” *Alexandria Eng. J.*, vol. 108, pp. 498–508, Dec. 2024, doi: 10.1016/j.aej.2024.07.081.

[36] I. Zakarija, F. Škopljanac-Mačina, H. Marušić, and B. Blašković, “A Sentiment Analysis Model Based on User Experiences of Dubrovnik on the Tripadvisor Platform,” *Appl. Sci.*, vol. 14, no. 18, pp. 1-27[1] I. Zakarija, F. Škopljanac-Mačina, H. Maru, Sep. 2024, doi: 10.3390/app14188304.

[37] A. M. Golubovikj, O. Mokryn, and M. Tkalčič, “Intended Movie Experience: Linking Elicited Emotions to Eudaimonic and Hedonic Characteristics,” in *CEUR Workshop Proceedings*, 2024, pp. 103–112. [Online]. Available: https://ceur-ws.org/Vol-3815/paper9.pdf

[38] P. Nandwani and R. Verma, “A review on sentiment analysis and emotion detection from text,” *Soc. Netw. Anal. Min.*, vol. 11, no. 1, pp. 1–19, Aug. 2021, doi: 10.1007/s13278-021-00776-6.

[39] Y. Albalawi, J. Buckley, and N. S. Nikolov, “Investigating the impact of pre-processing techniques and pre-trained word embeddings in detecting Arabic health information on social media,” *J. Big Data*, vol. 8, no. 1, Jul. 2021, doi: 10.1186/s40537-021-00488-w.

[40] M. S. A. Alzaidi *et al.*, “An Efficient Fusion Network for Fake News Classification,” *Mathematics*, vol. 12, no. 20, pp. 1–20, Oct. 2024, doi: 10.3390/math12203294.

[41] Z. Ahanin, M. A. Ismail, N. S. S. Singh, and A. AL-Ashmori, “Hybrid Feature Extraction for Multi-Label Emotion Classification in English Text Messages,” *Sustain.*, vol. 15, no. 16, Aug. 2023, doi: 10.3390/su151612539.

[42] A. Salhi, R. Alshamrani, A. Althbiti, A. Ismail, and M. Abd-elrahman, “Optimizing high dimensional data classification with a hybrid AI driven feature selection framework and machine learning schema,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–27, Oct. 2025, doi: 10.1038/s41598-025-08699-4.

[43] N. A. Azeez, S. Misra, D. O. Ogaraku, and A. P. Abidoye, “A Predictive Model for Benchmarking the Performance of Algorithms for Fake and Counterfeit News Classification in Global Networks,” *Sensors*, vol. 24, no. 17, pp. 1–34, Sep. 2024, doi: 10.3390/s24175817.

[44] A. Saranya and R. Subhashini, “A systematic review of Explainable Artificial Intelligence models and applications : Recent developments and future trends,” *Decis. Anal. J.*, vol. 7, pp. 1–14, Jun. 2023, doi: 10.1016/j.dajour.2023.100230.

[45] G. Huang, Y. Li, S. Jameel, Y. Long, and G. Papanastasiou, “From explainable to interpretable deep learning for natural language processing in healthcare : How far from reality ?,” *Comput. Struct. Biotechnol. J.*, vol. 24, pp. 362–373, Dec. 2024, doi: 10.1016/j.csbj.2024.05.004.

[46] P. Fantozzi and M. Naldi, “The Explainability of Transformers : Current Status and Directions,” *Computers*, vol. 13, no. 4, pp. 1–34, Apr. 2024, doi: 10.3390/computers13040092.

[47] L. Zhang, “Features extraction based on Naive Bayes algorithm and TF-IDF for news,” *PLoS One*, vol. 20, no. 7, pp. 1–17, Jul. 2025, doi: 10.1371/journal.pone.0327347.

[48] W. Hu, Q. Fan, H. Yan, X. Xu, S. Huang, and K. Zhang, “A Survey of Multi-Label Text Classification Under Few-Shot Scenarios,” *Appl. Sci.*, vol. 15, no. 16, pp. 1–42, Aug. 2025, doi: 10.3390/app15168872.

[49] L. Wei, Y. Li, Y. Zhu, B. Li, and L. Zhang, “Prompt Tuning for Multi-Label Text Classification: How to Link Exercises to Knowledge Concepts?,” *Appl. Sci.*, vol. 12, no. 20, pp. 1–13, Oct. 2022, doi: 10.3390/app122010363.

[50] Y. Ni and W. Ni, “A multi-label text sentiment analysis model based on sentiment correlation modeling,” *Front. Psychol.*, vol. 15, pp. 1–9, Dec. 2024, doi: 10.3389/fpsyg.2024.1490796.

[51] C. Miller, T. Portlock, D. M. Nyaga, and J. M. O’Sullivan, “A review of model evaluation metrics for machine learning in genetics and genomics,” *Front. Bioinforma.*, vol. 4, pp. 1–13, Sep. 2024, doi: 10.3389/fbinf.2024.1457619.

[52] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, “Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems,” *PLOS Digit. Heal.*, vol. 2, no. 11, pp. 1–19, Nov. 2023, doi: 10.1371/journal.pdig.0000290.

[53] N. Ashraf, L. Khan, S. Butt, H. T. Chang, G. Sidorov, and A. Gelbukh, “Multi-label emotion classification of Urdu tweets,” *PeerJ Comput. Sci.*, vol. 8, pp. 1–25, Apr. 2022, doi: 10.7717/peerj-cs.896.

[54] M. S. Durkee *et al.*, “Generalizations of the Jaccard index and Sørensen index for assessing agreement across multiple readers in object detection and instance segmentation in biomedical imaging,” *J. Med. Imaging*, vol. 10, no. 6, pp. 1–17, Dec. 2023, doi: 10.1117/1.jmi.10.6.065503.

[55] M. C. Hinojosa Lee, J. Braet, and J. Springael, “Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores,” *Appl. Sci.*, vol. 14, no. 21, pp. 1–21, Oct. 2024, doi: 10.3390/app14219863.

[56] M. Ria Ambrocio Sagum, “An Application of Emotion Detection in Sentiment Analysis on Movie Reviews,” *Turkish J. Comput. Math. Educ.*, vol. 12, no. 3, pp. 5468–5474, Apr. 2021, [Online]. Available: https://turcomat.org/index.php/turkbilmat/article/view/2204

[57] S. Su, “Sentimental Analysis Applied on Movie Reviews,” *J. Educ. Humanit. Soc. Sci.*, vol. 3, pp. 188–195, Sep. 2022, doi: 10.54097/ehss.v3i.1685.

[58] A. D. Mahmoodabadi, H. Dehghan, K. Kimiafar, and M. Karami, “Usability evaluation methods for hospital information systems : a systematic review,” *BMC Health Serv. Res.*, vol. 25, no. 1, pp. 1–14, Nov. 2025, doi: 10.1186/s12913-025-13745-y.

[59] M. O. D. Jr, “A Domain-Specific Evaluation of the Performance of Selected Web-based Sentiment Analysis Platforms,” *Int. J. Softw. Eng. Comput. Syst.*, vol. 9, no. 1, pp. 1–9, Jan. 2023, doi: 10.15282/ijsecs.9.1.2023.1.0105.

[60] M. A. Palomino and F. Aider, “Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis,” *Appl. Sci.*, vol. 12, no. 17, pp. 1–21, Aug. 2022, doi: 10.3390/app12178765.

[61] A. G. Silva *et al.*, “Procedures of User-Centered Usability Assessment for Digital Solutions : Scoping Review of Reviews Reporting on Digital Solutions Relevant for Older Adults,” *JMIR Hum. Factors*, vol. 8, no. 1, pp. 1–14, Jan. 2021, doi: 10.2196/22774.

[62] S. Almasi, K. Bahaadinbeigy, H. Ahmadi, and S. Sohrabei, “Usability Evaluation of Dashboards : A Systematic Literature Review of Tools,” *Biomed Res. Int.*, vol. 2023, no. 1, pp. 1–11, Feb. 2023, doi: 10.1155/2023/9990933.

[63] K. Kucher, S. Bouchama, A. Ebert, and A. Kerren, “Task-based Evaluation of Sentiment Visualization Techniques,” in *13th International Conference on Information Visualization Theory and Applications (IVAPP)*, 2022, pp. 187–194. doi: 10.5220/0010916400003124.

[64] F. Caldeira, J. Lourenço, and T. Chambel, “Multimodal Search and Visualisation of Movies Based on Emotions Along Time in As Movies Go By,” *SN Comput. Sci.*, vol. 6, no. 3, pp. 1–18, Mar. 2025, doi: 10.1007/s42979-025-03690-2.

BIOGRAPHIES OF AUTHORS

Suwarno, S.T., M.M, received his Bachelor's degree in Informatics Engineering from the University of Surabaya in 1999 and his Master's degree in Management from Universitas Internasional Batam in 2015. Since 2001, he has been a Lecturer in the Department of Information Systems at Universitas Internasional Batam, Indonesia. His research interests include computer science, data science, machine learning, artificial intelligence, and information systems. He has authored and co-authored several publications in national and international journals, focusing on software quality, digital innovation, and educational technology. In addition, he is an active member of APTIKOM (the Indonesian Association of Higher Education in Informatics and Computing) and the Ikatan Ahli Informatika Indonesia. He can be contacted via email at suwarno.liang@uib.ac.id

Wesly is an Information Systems student at Universitas Internasional Batam (UIB), Indonesia. He began his studies in 2022 and is currently in his seventh semester. His primary academic interests include information systems development, data management, and educational technology. In 2023, he co-authored a paper that was presented at the National Conference for Community Service Projects (NaCosPro). During his internship program, he served as an extracurricular assistant at SMAK Yos Sudarso Batam from March to May 2025, where he guided students in creative and digital activities. He is committed to leveraging technology to enhance learning and support community development. He can be contacted via email at 2231075.wesly@uib.edu

Bayu Syahputra, S. Kom., M.A. received his Bachelor's degree in Information Systems from Universitas Internasional Batam, Indonesia, and his Master's degree in Visual Communication Design from Ming Chi University of Technology, Taiwan. He is currently a Lecturer in the Department of Information Systems at Universitas Internasional Batam. His areas of expertise include cinematography and photography. His primary professional interest lies in integrating visual communication and digital creativity to support education and the creative industries. He can be contacted via email at bayu@uib.ac.id.

.