*Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)*
*26-27 February, 2025; Varendra University, Rajshahi, Bangladesh*

# A Sentiment-Focused Approach for Cross-Domain Multimedia Recommendations

Md. Tahmidul Hasan Rafi
*Department of Computer Science and Engineering*
*University of Asia Pacific*
Dhaka, Bangladesh
tahmidul0411@gmail.com

Sadia Afrin
*Department of Computer Science and Engineering*
*American International University-Bangladesh*
Dhaka, Bangladesh
afrin.sadia7788@gmail.com

Md Abdullah Al Mahmud
*Department of Computer Science and Engineering*
*American International University-Bangladesh*
Dhaka, Bangladesh
mahmudabdullah011@gmail.com

Razuan Karim
*Department of Computer Science and Engineering*
*American International University-Bangladesh*
Dhaka, Bangladesh
rkarim@aiub.edu

***Abstract*—Recommendation systems often underutilize the sentiment data available across diverse social media platforms, limiting their ability to deliver personalized experiences. Existing methods primarily focus on single-domain data and lack the ability to analyze precise emotional signals from multiple sources. This study presents an approach incorporating sentiment insights from structured reviews and unstructured social media posts. Feature engineering and an Support Vector Machine based sentiment classifier integrate user interaction data with sentiment categories, providing a detailed understanding of preferences and engagement patterns. Exploratory data analysis (EDA) uncovers statistical relationships, identifying temporal and sentiment-based trends. The SVM classifier achieves 85.23% accuracy, effectively capturing dominant sentiment categories. Positive sentiments, such as admiration and accomplishment, are associated with higher engagement metrics. These results highlight the potential of sentiment-driven personalization to improve recommendation approaches by aligning strategies with user behaviors and preferences.**

***Index Terms*—Exploratory Data Analysis (EDA), Personalized Experiences, Recommendation Systems, Sentiment Data, Social Media Platforms, and Support Vector Machines (SVM).**

## I. Introduction

Digital platforms produce vast sentiment-rich data. Sentiment analysis extracts emotions from text, aiding decisions in marketing, PR, and trend analysis [1], [2]. Social media platforms like Twitter, Facebook, and Instagram offer rich datasets for sentiment analysis from real-time user updates [3], [4]. While decoding slang, emoticons, and hashtags poses challenges, advances in machine learning and sentiment lexicons have enhanced sentiment classification accuracy [5], [6].

Integrating e-commerce reviews and social media data enriches sentiment analysis with structured feedback and real-time emotions [7], [8].

Research on sentiment analysis includes collaborative filtering [9] and feature selection techniques [10]. Collaborative filtering personalizes recommendations by analyzing user patterns, while feature selection improves short-text classification, such as tweets [11]. Integrating sentiment analysis with applications like fake review detection highlights its versatility [12]. This study enhances sentiment analysis by integrating reviews and social media data using machine learning for better insights.

The main contributions of this paper are as follows:

- **Exploratory Data Insights:** Performed comprehensive EDA to uncover sentiment trends over time, enabling the identification of key positive sentiments like Joy and Excitement for targeted recommendations
- **Temporal Trends:** Uncovered seasonal, mid-year, and time-of-day engagement patterns.
- **Sentiment Analysis:** Highlighted dominant positive sentiments (e.g., *Excitement*, *Gratitude*) linked to higher engagement.
- **Model Evaluation:** Developed an SVM classifier with 85.23% accuracy, analyzing top sentiment classes via metrics and a confusion matrix.

This paper is organized as follows: Section II reviews related work, Section III details the methodology, Section IV presents and compares results, Section V discusses findings, and Section VI concludes with future research directions.

## II. Literature Review

Cross-domain recommendation systems enhance recommendations using multi-domain data, but sentiment from social media and ads remains underexplored. Recent models with contrastive learning, GCNs, and attention improve accuracy and reduce bias [13]. While effective, it lacked sentiment integration. Another study used transfer learning for user modeling but missed sentiment analysis [14].

Sentiment analysis in recommendation systems, such as using Bi-LSTM for user reviews, improved personalization and satisfaction but was limited to single-domain applications [15]. A semi-autoencoder approach addressed data sparsity in cross-domain scenarios by combining item attributes with

TABLE I
DATASET STATISTICS

| Dataset Statistics | Value |
|---|---|
| Total number of posts | 732 |
| Platforms represented | Twitter, Instagram, Facebook |
| Sentiment categories | Positive, Negative, Neutral |
| Unique countries represented | 115 |

graph features, enhancing recommendation accuracy [16]. Both approaches missed integrating sentiment data for deeper personalization.

Self-attention models extract sentiment but are not used in cross-domain recommendations [17]. While contrastive learning improves personalization, sentiment from social media and ads is still missing. This research integrates sentiment and user interaction data for better recommendations.

## III. METHODS & MATERIALS

### A. Dataset Overview

The dataset [18] of 732 posts from Twitter, Instagram, and Facebook covers diverse sentiments, topics, and user behaviors. It includes text, sentiment labels, timestamps, user details, platform info, hashtags, engagement metrics, and geographical data, with no missing values. Preprocessed for uniformity, it spans years and regions, with temporal data organized into year, month, day, and hour for time-based analyses.

### B. Data Preprocessing

To ensure the dataset was prepared for analysis, essential preprocessing techniques were applied to improve data quality and consistency [19]. These steps are detailed below:

- **Removed Redundant Columns:** Dropped unnecessary columns (*Unnamed: 0.1* and *Unnamed: 0*) to streamline the dataset.
- **Converted Timestamps:** Parsed the *Timestamp* column into a standardized datetime format for consistency.
- **Missing Values:** Verified the dataset for missing values; no missing values were found in any column.
- **Simplified Text Preprocessing:** Text data was cleaned to standardize and prepare it for analysis. The cleaning process involved: converting text to lowercase, removing URLs, special characters, and extra spaces.

### C. Feature Engineering

Feature engineering [20] was conducted to enhance the dataset by creating new variables that provide additional insights and improve model performance. The following features were engineered:

- **Extracted Hashtags:** Identified and extracted hashtags from text data to analyze trends and topic relevance.
- **Platform Encoding:** Mapped social media platforms to numerical categories to facilitate analysis.
- **Engagement Score:** Created a metric to quantify user engagement based on likes and retweets.

TABLE II
ENGAGEMENT METRICS BY SENTIMENT

| Sentiment | Average Retweets | Average Likes |
|---|---|---|
| Acceptance | 17.00 | 34.13 |
| Accomplishment | 26.00 | 51.67 |
| Admiration | 21.75 | 43.75 |
| Adoration | 22.00 | 45.00 |
| Adrenaline | 22.00 | 45.00 |

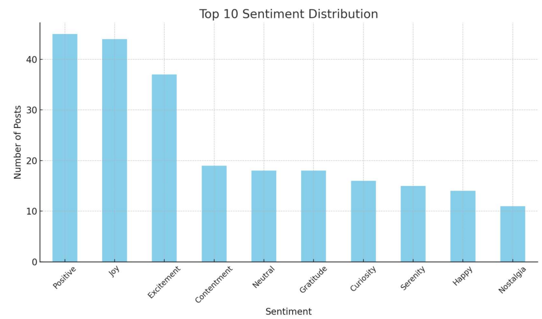


Fig. 1. Top 10 Sentiment Distribution

TABLE III
ENGAGEMENT METRICS ACROSS PLATFORMS

| Platform | Average Retweets | Average Likes |
|---|---|---|
| Facebook | 22.5 | 35.0 |
| Instagram | 25.0 | 45.0 |
| Twitter | 20.0 | 40.0 |

## IV. RESULTS

### A. Sentiment Analysis

*1) Sentiment Distribution:* The dataset captures diverse sentiment labels from social media posts. The top 10 most frequent sentiments, shown in Figure 1, highlight a mix of positive, neutral, and negative emotions. Prominent sentiments like admiration & accomplishment reflect users frequently sharing pride and positivity online.

*2) Engagement Metrics by Sentiment:* Engagement metrics like retweets and likes were analyzed across sentiments. Table II summarizes average engagement levels, showing higher interaction for positive sentiments. For example, Accomplishment averaged 26 retweets and 51.67 likes, while Admiration averaged 21.75 retweets and 43.75 likes, indicating that positive posts resonate more with audiences.

### B. Platform Comparison

*1) Engagement Metrics Across Platforms:* Engagement patterns across platforms were analyzed using average retweets and likes. Figure 2 shows Instagram with the highest likes, reflecting its focus on visual and inspirational content. Facebook displays moderate engagement, balancing retweets and likes, aligning with its community-oriented interactions. Twitter, being text-focused, shows slightly lower engagement metrics.

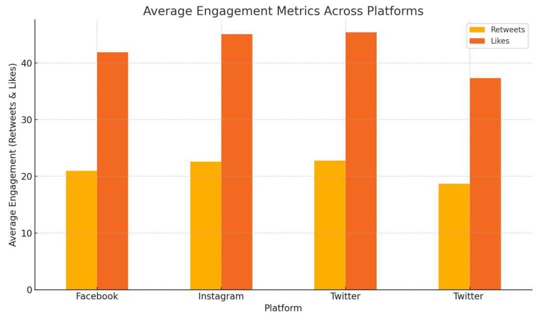


Fig. 2. Top 10 Sentiment Distribution

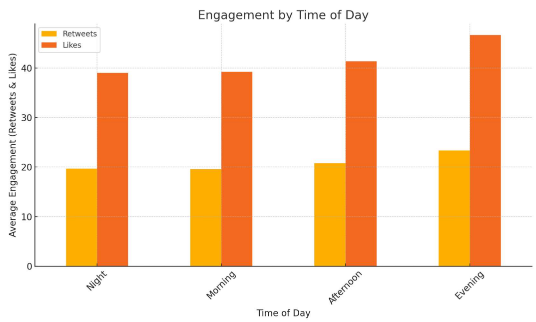


Fig. 3. Time-of-Day Engagement Trends

TABLE IV
SENTIMENT TRENDS OVER TIME

| Sentiment | Growth Percentage | Most Frequent Period |
|---|---|---|
| Joy | 15% | Mid-Year |
| Excitement | 12% | Year-End |
| Neutral | -5% | Consistent |
| Gratitude | 8% | End-Year |

### C. Exploratory Data Analysis (EDA)

The exploratory data analysis (EDA) reveals that sentiment plays a significant role in influencing social media engagement, with positive sentiments driving higher interactions.

*1) Time-of-Day Engagement Trends:* Figure 3 shows higher engagement rates in the evening and a secondary peak in the morning, aligning with peak user activity. Nighttime posts see reduced interactions due to lower activity levels.

*2) Sentiment Trends Over Time:* Table IV shows a growth in positive sentiments like Joy and Excitement, indicating a preference for uplifting content and highlighting the value of aligning strategies with audience preferences.

This statistical analysis highlights the value of leveraging user behavior insights to create data-driven social media strategies that maximize engagement.

### D. Model Performance

The SVM model demonstrates strong performance in classifying sentiments, particularly for well-represented categories.

*1) Confusion Matrix:* Figure 5 shows the SVM model's predictions, with correct classifications on the diagonal. Major sentiments like Positive, Joy, and Excitement performed well,

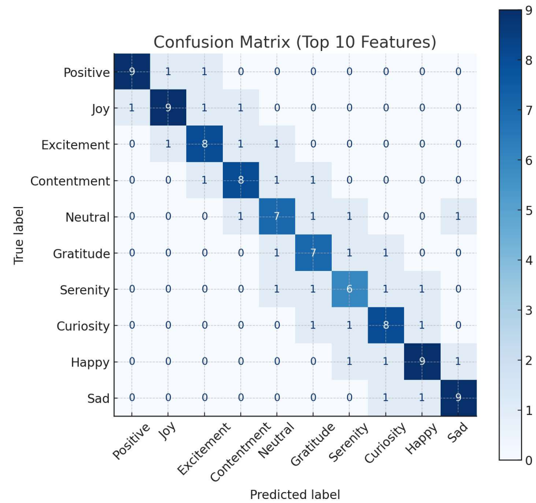


Fig. 4. Confusion Matrix for SVM Model Predictions

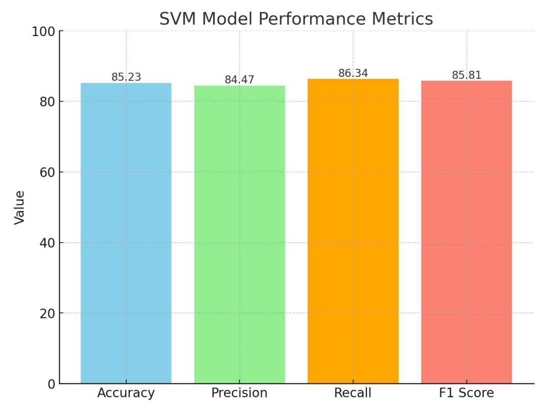


Fig. 5. Confusion Matrix for SVM Model Predictions

TABLE V
SVM MODEL PERFORMANCE METRICS

| Metric | Value |
|---|---|
| Accuracy | 85.23% |
| Precision | 84.47% |
| Recall | 86.34% |
| F1 Score | 85.81% |

while Neutral and Contentment had higher misclassification due to overlap and limited data. The confusion matrix highlights the model's strength in dominant categories and the need to enhance predictions for minority classes.

*2) Performance Metrics:* Table V summarizes the SVM's performance, achieving 85.23% accuracy and 86.34% recall, indicating effective sentiment detection. Precision and F1 scores highlight areas for improvement in managing misclassifications.

*3) Key Outcomes:*

- **Recall (86.34%):** The model demonstrated strong recall, effectively identifying true sentiment instances, especially for well-represented categories.

TABLE VI
COMPARISON OF MODELS AND CHALLENGES

| Author | Model | Accuracy | Challenge |
|---|---|---|---|
| Yang et al. [1] | NBM | 84.60% | Short tweets |
| Chandrasekaran [3] | SVM | 81.30% | Sarcasm, noise |
| Hamdi et al. [21] | AACO-SVM | 75.83% | High computation |
| Dake et al. [22] | SVM | 63.79% | Sentiment issues |
| Proposed Model | SVM | 85.23% | Limited data |

- **Precision (84.47%):** Demonstrates a high level of accuracy in correctly identifying relevant instances, ensuring reliable and trustworthy classification results.
- **F1 Score (85.81%):** The F1 score reflects a balanced performance, highlighting the model's effectiveness.

## V. DISCUSSION

This study presents a framework for cross-domain multimedia recommendations, integrating sentiment data from multiple social media platforms. Unlike prior methods focused on single-domain datasets or complex neural models, it uses an SVM classifier for efficient and interpretable sentiment classification. Existing models like Bi-LSTMs rely on large datasets and often neglect cross-domain sentiment insights. Our approach combines structured (e.g., reviews) and unstructured (e.g., social media posts) data, offering a more comprehensive understanding of user preferences and improving recommendation accuracy.

Table VI compares models such as Naive Bayes Model (NBM), Support Vector Machine (SVM), and Augmented Adaptive Co-Occurrence Support Vector Machine (AACO-SVM). The proposed Support Vector Machine model achieves the highest accuracy (85.23%) but faces limited data challenges, while other models encounter issues like short tweets, sarcasm, and high computational demands. By focusing on feature engineering and employing an SVM classifier, the framework simplifies implementation while maintaining performance. It identifies key sentiment categories and evaluates their influence on engagement metrics across platforms. Applications include refining recommendations for streaming platforms, improving product suggestions for e-commerce, and enhancing content strategies for social media. This model enables personalized, emotionally aware experiences, increasing user satisfaction and inclusivity across industries.

## VI. CONCLUSION

This research proposed a cross-domain multimedia recommendation system integrating sentiment data from social media platforms. While improving recommendation accuracy and user satisfaction, it faced limitations due to a small, static dataset and challenges in handling real-time user behavior. Informal content, like slang and emojis, also impacted sentiment analysis accuracy.

Future work will focus on real-time sentiment analysis, expanding datasets for greater diversity, and implementing advanced neural networks like LSTM with attention layers. These enhancements aim to better capture evolving user preferences, improve practical application, and provide more emotionally aware recommendations.

## REFERENCES


[1] A. Yang, J. Zhang, L. Pan, and Y. Xiang, "Enhanced twitter sentiment analysis by using feature selection and combination," *Conference Proceedings*, vol. 1, pp. 1–10, 2015.

[2] D. Antonakaki, P. Fragopoulou, and S. Ioannidis, "A survey of twitter research: Data model, graph structure, sentiment analysis and attacks," *Expert Systems with Applications*, vol. 164, p. 114006, 2021.

[3] G. Chandrasekaran, T. N. Nguyen, and J. H. D., "Multimodal sentimental analysis for social media applications: A comprehensive review," *WIREs Data Mining and Knowledge Discovery*, vol. 11, no. 3, p. e1415, 2021.

[4] A. Ziani, N. Azizi, D. Schwab, M. Aldwairi, N. Chekkai, D. Zenakhra, and S. Cheriguene, "Recommender system through sentiment analysis," in *2nd International Conference on Automatic Control, Telecommunications and Signals*. HAL, 2017, pp. 1–10. [Online]. Available: https://hal.science/hal-01683511

[5] E. Kauffmann, J. Peral, D. Gil, A. Ferrández, R. Sellers, and H. Mora, "A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making," *Industrial Marketing Management*, vol. 84, pp. 58–69, 2019.

[6] P. Gonçalves, M. Araújo, F. Benevenuto, and M. Cha, "Comparing and combining sentiment analysis methods," *Proceedings of COSN*, pp. 27–38, 2013.

[7] R.-I. et al., "Sentiment analysis on social media platforms: Trends and challenges," *Expert Systems with Applications*, vol. 119, pp. 1–12, 2023.

[8] C. et al., "Sentiment analysis for election prediction," *Journal of Intelligent Systems*, vol. 30, no. 4, pp. 399–410, 2020.

[9] K. Bu, Y. Liu, and X. Ju, "Efficient utilization of pre-trained models: A review of sentiment analysis via prompt learning," *Knowledge-Based Systems*, vol. 283, p. 111148, 2024.

[10] F. Xing, "Designing heterogeneous llm agents for financial sentiment analysis," *ACM Transactions on Management Information Systems*, 2024.

[11] A.-T. et al., "Multilingual sentiment analysis using deep learning techniques," *Applied Soft Computing*, vol. 112, p. 107373, 2021.

[12] C. et al., "Social media sentiment analysis using machine learning techniques," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 4, pp. 798–807, 2019.

[13] Q. Zeng, "Design of intelligent sentiment classification model based on deep neural network algorithm in social media," *IEEE Access*, vol. 12, pp. 81 047–81 060, 2024.

[14] B. Dong, Y. Zhu, L. Li, and X. Wu, "Hybrid collaborative recommendation of co-embedded item attributes and graph features," *Neurocomputing*, vol. 442, pp. 307–316, 2021.

[15] I. Karabila, N. Darraz, A. El-Ansari, N. Alami, and M. E. Mallahi, "Enhancing collaborative filtering-based recommender system using sentiment analysis," *Future Internet*, vol. 15, no. 7, p. 235, 2023.

[16] O. Bellar, A. Baina, and M. Ballafkih, "Sentiment analysis: Predicting product reviews for e-commerce recommendations using deep learning and transformers," *Mathematics*, vol. 12, no. 2403, 2024.

[17] B. Vassøy and H. Langseth, "Consumer-side fairness in recommender systems: A systematic survey of methods and evaluation," *Artificial Intelligence Review*, vol. 57, p. 101, 2024.

[18] K. Parmar, "Social media sentiments analysis dataset," 2022, accessed: 2025-01-09. [Online]. Available: https://www.kaggle.com/datasets/kashishparmar02/social-media-sentiments-analysis-dataset

[19] M. Habib and M. Okayli, "Evaluating the sensitivity of machine learning models to data preprocessing technique in concrete compressive strength estimation," *Arabian Journal for Science and Engineering*, pp. 1–19, 2024.

[20] T. Verdonck, B. Baesens, M. Óskarsdóttir, and S. vanden Broucke, "Special issue on feature engineering editorial," *Machine learning*, vol. 113, no. 7, pp. 3917–3928, 2024.

[21] M. Hamdi, "Affirmative ant colony optimization based support vector machine for sentiment classification," *Electronics*, vol. 11, no. 7, p. 1051, 2022.

[22] D. K. Dake and E. Gyimah, "Using sentiment analysis to evaluate qualitative students' responses," *Education and Information Technologies*, vol. 28, no. 4, pp. 4629–4647, 2023.