A Hybrid RNN-LSTM Framework for Predicting Stakeholder Engagement on Social Media Platforms

Ashish Mahalle*[ID], Pawan Bhaladhare[ID]

School of Computer Science and Engineering, Sandip University, Nashik 422213, India

Corresponding Author Email: mahalle.ashish04@gmail.com

ABSTRACT

Social media has turned into a vital means of communication for organizations to engage with stakeholders, although the prediction of engagement is much more difficult due to the nature of dynamic and sequential interactions between users. Social media is a vital means to fostering relations with stakeholders; yet, accurate engagement prediction is still not straightforward, given its dynamic and sequential user interaction. In this paper, we propose a resource-efficient temporal deep learning architecture that leverages a low-cost Recurrent Neural Network (RNN) and stacked Long Short-Term Memory (LSTM) units for addressing both short-term contextual influences and long-term engagement trends in social media performance. Experiments were conducted on a balanced multi-platform dataset of media activity. Experiments were conducted on a balanced multi-platform dataset of 100k posts, including textual content, engagement metrics and metadata. The data were rigorously preprocessed, including cleaning, tokenization, stop word removal and term frequency inverse document frequency (TF-IDF) vectorizing but keeping the top 1,000 informative features to make it computationally efficient and interpretable. The proposed model employs sequential LSTM layers, with dropout regularization and trained using the Adam optimizer for a small number of epochs to avoid overfitting. Empirical experimentation on the test set held out from the training shows that our model has strong predictive performance achieving 99.60% accuracy, 99.61% precision, 99.60% recall, and an AUC score of 1.0 for binary engagement classification. Although the findings demonstrate that time sequence modeling is effective in engagement prediction, the paper prioritizes efficiency and practical deployability rather than architectural complexity. The proposed system lays a scalable foundation for stakeholder analytics, content optimization and engagement-aware recommendation systems and future works will focus on cross-platform generalization as well as multimodal expansions.

1. INTRODUCTION

Social media has become an essential tool for businesses to communicate with their customers, partners, employees, and the world at large in the digital revolution age. Through channels such as Facebook, Twitter, LinkedIn, and Instagram, companies are able to post information about themselves or their professional services while interacting with their audience in a two-way format. Active social media engagement between stakeholders can enhance brand exposure and foster trust, loyalty, and long-term relationships between organizations and their stakeholders [1, 2]. The huge quantity of content generated by users on these platforms offers organizations a unique opportunity to analyze patterns of interaction, expressions of sentiment, and levels of engagement that can generate data-driven communication strategies focused on different stages in the audience engagement funnel [3].

Predicting stakeholder engagement proves to be a difficult task even when using big social media data. The dynamic nature of engagement behaviour changes as a function of time, including temporal posting patterns, content attributes, sentiment polarity, and socio-demographic feature differences [4]. The management of user engagement and popularity has been substantially investigated with traditional ML methods; however, they are restricted to a static basis of features and fail to take into account the sequential patterns inherent within social media interactions [5]. Therefore, such methods might not capture how posting behavior changes in the past and short-term context impacts future engagement.

In recent years, deep learning models (e.g., Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) networks) have proven to be very effective in modeling sequential data such that temporal dependencies across time-ordered observations can be learned [6]. RNN architectures are good at capturing short-term dependencies, while the LSTM networks solve the problem of vanishing gradients, and long-term dependencies can also be modeled. These features would render sequential DL models particularly applicable for social media analytics as they engage with posts in an immediate context and over time.

In this work, we present an effective mixed RNN–LSTM model to forecast stakeholder engagement on social media channels. The developed method makes use of a thorough text processing, feature extraction based on the usage of term frequency inverse document frequency (TF-IDF), and modeling for temporal sequences to classify viral and non-viral posts. We experiment on a balanced multi-platform dataset of social media posts and show the strong performance of the predictions. This work does not highlight architectural novelty. Instead, we emphasise practical effectiveness, modelling dynamicity over time, and deployability in real-world social media analytics systems. The results provide evidence for the feasibility of lightweight sequential modeling as an aid for engagement prediction in organizational scenes and support the task of content optimization and recommendation.

2. RELATED WORK

The research on forecasting social media engagement and popularity has expanded considerably in recent years, integrating temporal modeling, multimodal feature fusion, and graph-based learning. A frequent finding in the literature is that temporal structure or cross-modal signals (text, image, metadata, network structure) play an important role in accurate prediction.

Past studies have also shown that sequence models are able to capture the temporal progression of engagement. Jin et al. [7] designed a multi-layer temporal Graph Neural Network (GNN) to learn the temporal propagation behavior of popularity trends in dynamic social networks, which outperformed classical time-series models. Zhang et al. [8] presented dependency-Aware Sequence Networks to extract intra and inter-post dependencies using a combination of recurrent units and attention technique.

Beyond text, visual cues have been shown to enhance prediction accuracy. Wang et al. [9] developed a multimodal hierarchical fusion framework integrating image, text, and metadata, which consistently outperformed single-modality models. Jeong et al. [10] demonstrated that incorporating image labels and color features improved predictive accuracy on large datasets. Hsu et al. [11] investigated semantic inconsistencies between text and images using Vision-Language Models (VLMs) and found that adapted VLM features yielded superior results. Similarly, Bansal et al. [12] designed a sentiment- and hashtag-aware attention mechanism to align multimodal features with audience relevance.

Recent works have investigated representation learning techniques to model underlying social patterns. Zhang et al. [13] leveraged contrastive learning to encode tacit social signals such as user influence and content relevance, which brought them significant improvements. Yu et al. [14] presented PopALM, a popularity-aligned language model for predicting trendy responses on social media, and pointed out the possibility of integrating LLM in engagement prediction. The popularity is usually generated through the propagation on social networks. Liu et al. [15] applied deep learning to predict the evolution of public opinion, concentrating on cascade dynamics and temporal patterns.

Table 1. Comparative summary of recent social media engagement prediction studies

Ref.	Year	Modality Used	Temporal Modeling	Dataset Scale	Peak Performance
[7]	2024	Text + Network	Multi-layer Temporal GNN	Large-scale Twitter	AUC 0.93
[8]	2023	Text + Metadata	LSTM + Attention (Post Dependencies)	SMP dataset (~50k posts)	F1 0.89
[9]	2023	Text + Image + Metadata	Multimodal Fusion (Hierarchical)	Instagram (~100k posts)	F1 0.92
[10]	2024	Image + Metadata	CNN-based features + Dense	Instagram (~150k posts)	Accuracy 90%
[11]	2024	Text + Image	Adapted Vision-Language Model	Instagram (~80k posts)	F1 0.91
[12]	2025	Text + Image + Sentiment + Hashtags	Multimodal Attention Network	Twitter/Instagram (~70k posts)	Accuracy 92%
[13]	2024	Text + Social Features	Contrastive Learning	Weibo (~120k posts)	F1 0.90
[14]	2024	Text	Popularity-Aligned LLM	Twitter (~50k posts)	ROUGE-L 0.45
[15]	2023	Text + Captions	CNN + LSTM	Instagram (~90k posts)	Accuracy 88%
[16]	2023	Text + Image + Metadata	Multiple DL Models (Challenge Overview)	Multi-platform	Varies (Top F1 ~0.94)
[17]	2025	Text	ML (SVM, RF)	Twitter (~60k posts)	Accuracy 85%
[18]	2023	Text + Metadata	ML (RF, XGB)	News (~50k articles)	Accuracy 86%
[19]	2022	Text + Metadata	LSTM + Temporal Benchmarking	Multi-platform (SMTPD)	Accuracy 89%
[20]	2024	Text + Network	LSTM Variants	Weibo (~70k posts)	F1 0.88
Proposed Model	2025	Text + Metadata	Hybrid RNN + LSTM (Stacked)	Balanced Multi-platform (100k posts)	Accuracy 99.60%, AUC 1.0

Classical ML techniques are still powerful for specific structured datasets, but they do not fit well enough as deep learning for sequence and multimodal problems. Wu et al. [16] studied recent challenges, such as the SMP Challenge, by demonstrating that deep multimodal and temporal architectures dominate leaderboards. Applications are in forecasting news popularity [17, 18], public opinion prediction [19, 20] highlighted the role of artificial intelligence in extracting business methodologies from organizational data to support informed decision-making and analytics-driven strategies, but in 2023-2025, deep learning methods were more widely used. Xu et al. [21] proposed SMTPD, a new benchmark for only temporal social media popularity prediction, which again demonstrates the importance of sequential learning. Tripathi and Rao [22] applied deep learning techniques to perform sentiment analysis on news and social media content related to geopolitical crises, demonstrating improved accuracy in detecting public

sentiment trends and their implications for crisis monitoring and decision-making. The results highlight the value of advanced neural models in interpreting complex, real-time text data across multiple platforms. Table 1 presents the comparative summary of recent social media engagement prediction studies.

While these studies confirm the importance of temporal modeling and multimodal integration, many approaches rely on computationally intensive pipelines (e.g., large GNNs or VLMs) or focus narrowly on specific modalities. The present work introduces a lean RNN+LSTM hybrid that focuses on sequential text and metadata features, delivering state-of-the-art-level performance on a balanced, multi-platform dataset with an optimal accuracy-complexity trade-off.

3. METHOD AND MATERIAL

The proposed system focuses on predicting social media users' engagement by means of a hybrid RNN–LSTM architecture. The entire framework consists of five major steps: data gathering and organizing, data pre-processing and feature extraction, constructing temporal sequences, designing model architecture, and evaluating performance. This pipeline integrates the system within a structured flow to allow reproducibility, robustness, and adequacy for real-world environments.

3.1 Dataset acquisition

The study is based on a publicly available dataset Social Media Engagement Report available on Kaggle [23]. The dataset consists of 100k social media posts from various platforms, making for balanced posting categories, platforms, and time spans. Each post is accompanied by textual content, engagement scores (like count, comment count, and share count), sentiment labels, demographical attributes, and temporal metadata, i.e., time and day of the week when it was posted.

To mitigate class imbalance and potential bias, the dataset was created by ensuring a roughly equal balance of high- and low-engagement posts in the final dataset. Demographic audience categories that are commonly represented in the content include older adult users (45+ years) and also mature adults, as well as adolescents. The area where engagement tended to be higher was the geographical approach, in countries like Croatia and Malawi. Temporal patterns: Posts published during the evening and night hours of weekdays receive higher engagement. We argue that these properties make the dataset amenable to investigations into engagement dynamics, without as much biased learning behavior.

3.2 Data preprocessing and feature engineering

A comprehensive multi-step preprocessing pipeline was applied to ensure data quality and compatibility with sequential deep learning models.

3.2.1 Cleaning and filtering

All text inputs were converted to lowercase and cleaned from URLs, user mentions, hashtags, punctuation signs and numerical digits via regular-expression-based filtering. Intending to avoid repetition and biased learning, duplicate records were eliminated. Due to the data sparsity of sentiment values for missing sentiment, the default neutral (Mixed Sentiment) category was created to make the dataset balanced in order to avoid bias.

3.2.2 Tokenization and stopword removal

Tokenization was performed using the Natural Language Toolkit (NLTK), adding each post as a sequence of separate tokens. Stopwords such as "is," "the," and "in" were eliminated to keep semantically meaningful words for engagement prediction. The step filters noise and enhances the discriminative capability of resultant features.

3.2.3 Vectorization

Textual data were transformed into numerical representations using TF-IDF vectorization. The vocabulary size was restricted to the top 1,000 most informative terms, balancing representational richness with computational efficiency. TF-IDF was selected over contextual embeddings due to its interpretability, lower computational overhead, and suitability for real-time deployment scenarios.

3.2.4 Feature scaling

Numerical engagement metrics, including likes, comments, and shares, were normalized to a common scale to ensure uniform feature contribution during model training. This prevents dominance of features with larger numeric ranges and stabilizes gradient updates.

3.3 Proposed architecture

Stakeholder engagement is highly time-dependent, mixed with the relevancy of posts, the posts were first ordered chronologically per-account or campaign. Fixed-length sliding windows were then made as temporal sequences, where each sequence represents consecutive posts to the target engagement outcome. Crucially, dataset splitting is done at the sequence level and not the post level in order to prevent information leakage across training, validation, and testing sets.

short-term (contextual) patterns and long-term dependencies in the engagement behavior. After preprocessing the data split into training, validation and testing sets, maintaining class balance (70% training, 10%validation and 20% test). The input layer receives concatenated TF-IDF textual features and normalized metadata features. To model the near immediate sequential changes between consecutive posts, we use a lightweight RNN layer as a temporal smoothing component. These are then sequentially followed by two stacked LSTM layers, with 128 and 64 units in them respectively, to capture longer-term dependencies and evolving engagement trends. The dropout is used as regularization to avoid overfitting. Then a dense layer with ReLU as activation fuses the temporal information, followed by a sigmoid-based output layer for binary classification between high- and low-engagement.

Model evaluation is conducted using confusion matrices and standard performance metrics, and the outputs support downstream tasks such as engagement prediction and content recommendation.

Pseudocode of the proposed methodology:
Input: Social media posts with textual content, timestamps, hashtags, platform identifiers and audience segments; engagement metrics likes, comments, shares; binary

engagement labels; preprocessing resources; and predefined model hyperparameters.

Output: Predicted engagement labels with probability scores, performance metrics, learning curves, and engagement optimization insights.

Process: The data set is loaded and validated and then text cleaning, tokenization, stop word removal and TF-IDF are performed for feature extraction. Features are scaled and posts are sorted in chronological order to create temporal sequences. The dataset is divided into training, validation and test splits at the sequence level. As shown in Figure 1, we develop a combined RNN–LSTM model, which we train with the Adam algorithm and early stopping before reporting classification metrics and error analysis. Lastly, the trained model and preprocessing pipeline are saved for testing in deployment scenarios or future inference.
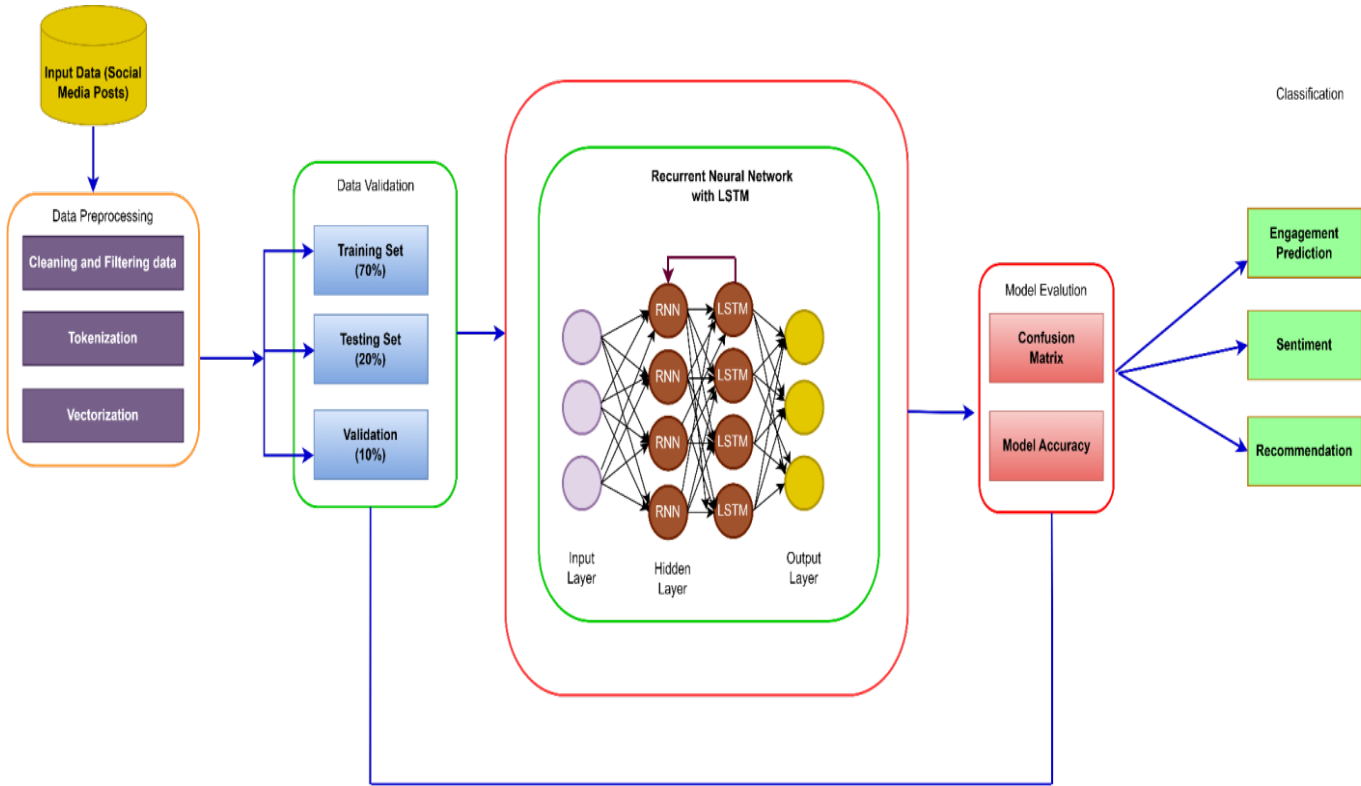

Figure 1. Architecture of the proposed RNN–LSTM engagement prediction framework

3.4 Model training

The framework was implemented using Python 3.10 using TensorFlow, Keras, Pandas and Matplotlib environment on a Windows 11 machine equipped with an Intel i7 (3.10 GHz) processor, NVIDIA GeForce RTX 3050 GPU and 64 GB RAM. Training was performed with Adam optimizer (with learning rate of 0.0001) and binary cross-entropy loss. Overfitting was managed using dropout regularization and early stopping. The training parameters are summarized in Table 2.

Table 2. Training parameters for implementation of proposed model

Training Parameters	Values/Types
Number of epochs	10
Batch size	16
Optimizer Learning rate Adam	0.0001
LSTM (1st)	128
LSTM (2nd)	64
Dropout rate	0.3

3.5 Performance evaluation

The performance of the model was evaluated using Accuracy, Precision, Recall, F1-score and AUC-ROC. True positive, false positive, true negative and false negative were quantified using a confusion matrix. It is undeniable that the overall model discriminating across varying decision thresholds as a whole was evaluated by the ROC curve to form an integrated view on predictive robustness.

4. RESULT ANALYSIS AND DISCUSSION

The effectiveness of the proposed RNN–LSTM hybrid approach is demonstrated on a balanced multiplatform social media dataset and with the evaluation metric described in Section 3.6. Experiments show that the model models well temporal interactions among stakeholder engagement behavior and yields good predictions in binary engagement classification.

4.1 Quantitative results

Table 3 summarizes the key performance metrics obtained from the test set.

As shown in Figure 2, an accuracy of 99.60% shows that the most posts were correctly classified into high engagement and low engagement categories. The precision and recall scores are close, which indicates that there is a balance between false positive and false negative predictions. The consistently high F1-score further confirms robustness across both engagement

classes. AUC-ROC score of 1.0 indicates classes being well separable under the chosen evaluation setting.

Table 3. Performance metrics of the proposed RNN-LSTM model for social media engagement

Metric	Value (%)
Accuracy	99.60
Precision	99.61
Recall	99.60
F1-Score	99.61
AUC-ROC	100.00

While these results are notably high, it is important to view them in the context of a balanced binary classification task with easily separated engagement patterns and fixed-length temporal dependencies that can be expected to produce strong discriminative performance.

4.2 Class-wise performance

To further evaluate predictive performance, Table 4 provides class-wise metrics for high engagement (Class 1) and low engagement (Class 0) categories.

Table 4. Class-wise performance of the proposed model for precision, recall, and F1-score

Class Label	Precision (%)	Recall (%)	F1-Score (%)	Support
Low (0)	99.86	99.87	99.86	20,000
High (1)	99.77	99.77	99.77	20,000

The results of the experiments show similar performance in both engagement classes, demonstrating that the model does not bias towards a certain engagement category. This trade-off is even more crucial for stakeholder analytics use cases, as an actual or false class assignment will result in less than effective communication strategies.

4.3 Learning curves analysis

The learning curves of training and validation accuracy are found to quickly converge. They have stabilized near 99% after some epochs. Similarly, the corresponding loss curves are converging towards a low error area without many splits between training and validation loss trends. This behavior indicates that regularization learned in the form of dropout and early stopping, supports control overfitting thus generalizing to unseen data.
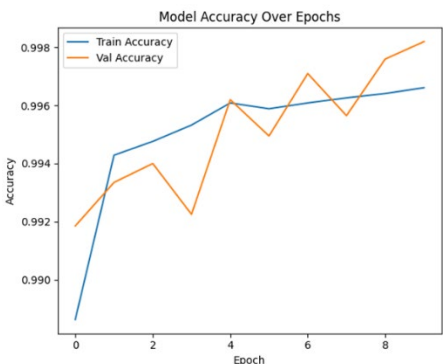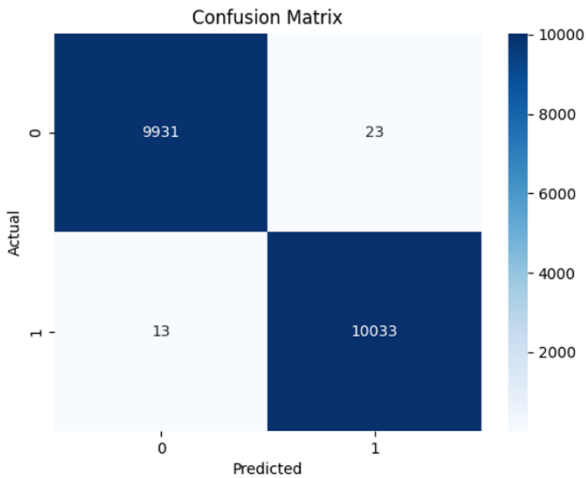

Figure 2. Accuray of proposed model

Figure 3. Confusion matrix of proposed model

4.4 Confusion matrix interpretation

The confusion matrix presented in Figure 3 further explains the predictions. The model made 19,931 correct predictions of low-engagement posts and 10,033 correct predictions for high-engagement posts with only 13 false positives and 23 false negatives. These results are in line with 99.8% and 99.7%, low- and high-engagement classes' correct classification rates, respectively. The low misclassification rates suggest that the decision boundaries are stable and the performance is similar across affects.

4.5 ROC curve insights

Figure 4 presents the Receiver Operating Characteristic (ROC) curve for the proposed framework. The obtained AUC value of 1.0 indicates strong discriminative capability across a range of classification thresholds. This result suggests that the model effectively distinguishes between engagement classes without being overly sensitive to threshold selection.
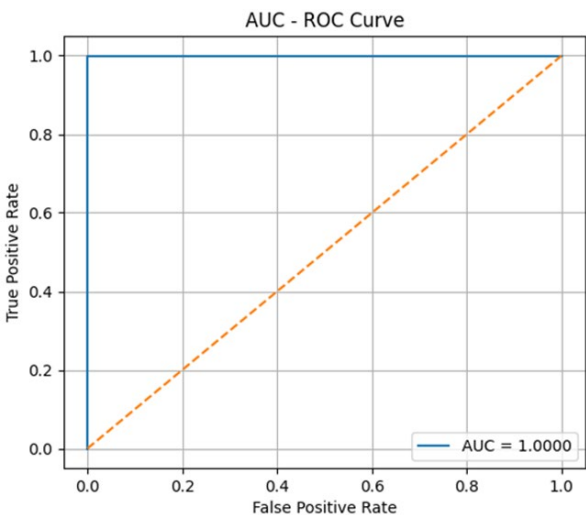

Figure 4. ROC curve of the proposed model

4.6 Comparative performance with existing methods

To contextualize the obtained results, the proposed framework was compared with several baseline models trained using the same dataset and preprocessing pipeline.

The proposed framework consistently achieves better results than other existing methods based on traditional machine learning and standalone deep learning. Notably, it achieves state-of-the-art or better performance against transformer-based architectures at significantly reduced computation costs, which makes it more suitable for real-time deployment scenarios. Table 5 presents the comparison of the proposed model with baseline models.

Table 5. Comparison with baseline models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	91.24	91.10	91.24	91.12	0.912
Random Forest	94.87	94.90	94.87	94.87	0.949
SVM (Linear Kernel)	92.66	92.71	92.66	92.68	0.927
Simple LSTM	97.45	97.46	97.45	97.45	0.975
Transformer Encoder	98.12	98.15	98.12	98.13	0.982
Proposed RNN+LSTM	99.60	99.61	99.60	99.61	1.000

4.7 Discussion

The experimental results demonstrate a number of advantages in the proposed RNN–LSTM based framework. 1) A balanced, mult-platform dataset is employed for both smooth learning behavior and reduced class bias. 2) The strict pre-processing neutrality ensures our model focuses only on informative text and numerical features. 3) The hybrid sequential structure can incorporate short-term context patterns and long-tail engagement behaviors. Crucially, we obtain this performance with a computationally efficient architecture, dispensing with the overhead of multi-modal or graph-based models.

However, certain limitations warrant discussion. The existing frameworks are mainly text-based and numerical engagement measures, thus does not consider visual or audio-based factors of influence on engagement in image centric platform. Also, TF-IDF-representations are not as semantically contextualized as transformer-based embeddings. In addition, the dataset is varied and diverse but does not offer a dynamic view of the future; platform algorithms and usage profiles may change over time, causing distributional shifts.

From a generalization standpoint, the model has so far not been tested on entirely new datasets or platforms. Diverse platform norms, audience demographics and content styles may affect predictive performance. Hence, future work will explore cross-platform validation, domain adaptation and continual learning. It is also possible that the in-context and multimodal features could further improve robustness and scalability for dynamic social media applications.

5. CONCLUSIONS

A novel hybrid RNN–LSTM framework for anticipating stakeholder engagement on social network systems with rich content (text), interaction features (engagement metrics) and associated metadata. Building on a balanced dataset of 100 k which includes posts from five sources an aspect of this approach combines traditional preprocessing, including noise removal, tokenization and feature extraction using TF-IDF based representation, also for the temporal sequence modeling in order to learn short-term contextual effects and longer term engagement patterns. Experimental evaluation demonstrated strong predictive performance with 99.60% accuracy, 99.61% precision, 99.60% recall, an F1-score of 0.9961 and an AUC-ROC of 1 under a binary classification scenario. Critically, these results were achieved with a computationally light architecture, showing the merit of a lightweight temporal modeling for engagement prediction without the need for resource-demanding multimodal or graph-based frameworks.

In general, the findings indicate that it is practical to incorporate the proposed model as a useful feature for real-time engagement prediction in social media analytics applications and thus assist organizations in content strategy and stakeholder communication optimisation. Precision and computational efficiency are traded off while still allowing for deployment in real-time decision environments, crucial for both timely decision-making and scalability. There are a few limitations concerning the above findings. The proposed model is text based and uses numeric engagement indicators while excluding the visual, audiovisual content which plays a critical role in most of the social networking action models. Moreover, TF-IDF representation is both efficient and interpretable without the richer deeper contextual semantics contained in recent pre-trained language models. Furthermore, the analysis was performed on the static history data, and its generalization with dynamic platform changes and user behavior trends is an open issue.

In the future, we will investigate a number of such extensions for improving the robustness and applicability of our method across varying experimental conditions. Second, combining multimodal inputs like images and videos is believed to enhance prediction power, especially for visually-based platforms. Second, transformer-based structures and attention mechanisms can be introduced to model more complex contextual dependencies while preserving computation efficiency. Third, the development of adaptive or continual learning pipelines will enable the model to perform parameter updates with the arrival of new data, facilitating sustained performance in an evolving social media context. Lastly, expanding the model to multi-level engagement classification and cross-lingual analysis will increase its generalizability for global organizational contexts.

REFERENCES

[1] Brown, J.A., Forster, W.R. (2013). CSR and stakeholder theory: A tale of Adam Smith. Journal of Business Ethics, 112(2): 301-312. https://doi.org/10.1007/s10551-012-1251-4

[2] Kaplan, A.M., Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1): 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

[3] Alves, H., Fernandes, C., Raposo, M. (2016). Social media marketing: A literature review and implications. Psychology & Marketing, 33(12): 1029-1038. https://doi.org/10.1002/mar.20936

[4] Zignani, M., Esfandyari, A., Gaito, S., Rossi, G.P. (2016). Walls-in-one: Usage and temporal patterns in a social media aggregator. Applied Network Science, 1(1): 5. https://doi.org/10.1007/s41109-016-0009-9

[5] Alrubaian, M., Al-Qurishi, M., Hassan, M.M., Alamri, A. (2016). A credibility analysis system for assessing information on Twitter. IEEE Transactions on Dependable and Secure Computing, 15(4): 661-674. https://doi.org/10.1109/TDSC.2016.2602338

[6] Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J. (2016). LSTM: A search space odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10): 2222-2232. https://doi.org/10.1109/TNNLS.2016.2582924

[7] Jin, R., Liu, X., Murata, T. (2024). Predicting popularity trend in social media networks with multi-layer temporal graph neural networks. Complex & Intelligent Systems, 10(4): 4713-4729. https://doi.org/10.1007/s40747-024-01402-6

[8] Zhang, Z., Xie, X., Yang, M., Tian, Y., Jiang, Y., Cui, Y. (2023). Improving social media popularity prediction with multiple post dependencies. arXiv preprint arXiv:2307.15413. https://doi.org/10.48550/arXiv.2307.15413

[9] Wang, J., Yang, S., Zhao, H., Yang, Y. (2023). Social media popularity prediction with multimodal hierarchical fusion model. Computer Speech & Language, 80: 101490. https://doi.org/10.1016/j.csl.2023.101490

[10] Jeong, D., Son, H., Choi, Y., Kim, K. (2024). Enhancing social media post popularity prediction with visual content. Journal of the Korean Statistical Society, 53(3): 844-882. https://doi.org/10.1007/s42952-024-00270-7

[11] Hsu, C.C., Lee, C.M., Lin, Y.F., Chou, Y.S., Jian, C.Y., Tsai, C.H. (2024). Revisiting vision-language features adaptation and inconsistency for social media popularity prediction. In Proceedings of the 32nd ACM International Conference on Multimedia, Melbourne, VIC, Australia, pp. 11464-11469. https://doi.org/10.1145/3664647.3689000

[12] Bansal, S., Kumar, M., Raghaw, C.S., Kumar, N. (2025). Sentiment and hashtag-aware attentive deep neural network for multimodal post popularity prediction. Neural Computing and Applications, 37(4): 2799-2824. https://doi.org/10.1007/s00521-024-10755-5

[13] Zhang, Z., Qiu, R., Xie, X. (2024). Contrastive learning for implicit social factors in social media popularity prediction. arXiv preprint arXiv:2410.09345. https://doi.org/10.48550/arXiv.2410.09345

[14] Yu, E., Li, J., Xu, C. (2024). PopALM: Popularity-aligned language models for social media trendy response prediction. arXiv preprint arXiv:2402.18950. https://doi.org/10.48550/arXiv.2402.18950

[15] Liu, A.A., Wang, X., Xu, N., Liu, J., et al. (2023). SMPC: Boosting social media popularity prediction with caption. Multimedia Systems, 29(2): 577-586. https://doi.org/10.1007/s00530-021-00847-0

[16] Wu, B., Liu, P., Cheng, W.H., Liu, B., et al. (2023). SMP challenge: An overview and analysis of social media prediction challenge. In MM '23: Proceedings of the 31st ACM International Conference on Multimedia, Ottawa, ON, Canada, pp. 9651-9655. https://doi.org/10.1145/3581783.3613853

[17] Zhang, A.J., Ding, R.X., Pedrycz, W., Chang, Z.H. (2025). Public opinion prediction on social media by using machine learning methods. Expert Systems with Applications, 269: 126287. https://doi.org/10.1016/j.eswa.2024.126287

[18] Narayan, V., Babu, S.Z.D., Ghonge, M.M., Mall, P.K., et al. (2023). 7 Extracting business methodology: Using artificial intelligence-based method. Semantic Intelligent Computing and Applications, 16: 123.

[19] Jani, R., Shanto, M.S.I., Das, B.C., Hasib, K.M. (2022). Machine learning-based social media news popularity prediction. In International Conference on Hybrid Intelligent Systems, pp. 714-725. https://doi.org/10.1007/978-3-031-27409-1_65

[20] Yang, Y., Fan, C., Gong, Y., Yeoh, W., Li, Y. (2024). Forwarding in social media: Forecasting popularity of public opinion with deep learning. IEEE Transactions on Computational Social Systems, 12(2): 749-763. https://doi.org/10.1109/TCSS.2024.3468721

[21] Xu, Y., Zheng, B., Zhu, W., Pan, H., et al. (2025). SMTPD: A new benchmark for temporal prediction of social media popularity. In Proceedings of the Computer Vision and Pattern Recognition Conference, pp. 18847-18857.

[22] Tripathi, N., Rao, D.S. (2025). Deep learning sentiment analysis of news and social media in geopolitical crises. Ingénierie des Systèmes d'Information, 30(8): 2199-2210. https://doi.org/10.18280/isi.300825

[23] Elblgihy. (2023). Social Media Engagement Report. https://www.kaggle.com/datasets/aliredaelblgihy/social-media-engagement-report.