International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2s) 2026 pp. 269-285
https://doi.org/10.70917/ijcisim-2026-2077
Article

SEMANTIC-AWARE POPULARITY PREDICTION OF SOCIAL MEDIA POSTS USING ROBERTA AND BILSTM NETWORKS

Sonam Mehta[1], Pragya Shukla[2]

[1] IET, Devi Ahilya Vishwavidyalaya, Indore, 452010, India. sjsonamjainsj@gmail.com
[2] IET, Devi Ahilya Vishwavidyalaya, Indore, 452010, India. pshukla@ietdavv.edu.in
Corresponding Author: Sonam Mehta (sjsonamjainsj@gmail.com)

Abstract: The popularity of user-generated content has rapidly increased with the popularity of social media platforms, which have created many opportunities for the prediction of online post popularity. Hard-coded features and shallow neural networks, such as Random Forests, are not expressive enough to capture the semantic nature of language and the contextual phenomena of social media. To overcome these issues, we develop a semantic-aware model for popularity prediction based on a hybrid RoBERTa-BiLSTM architecture. RoBERTa extracts deep contextual semantics in post-text, and BiLSTM models temporal dependencies between tokens. The proposed system significantly outperforms baseline approaches, achieving state-of-the-art results with an accuracy of 1.00, an F1 score of 0.99, and a Spearman Rank Correlation of 0.9985. Comparative analysis with Transformer, LSTM, and BERT-based models further confirms the robustness of our model. Moreover, the MAE and MSE error metrics suggest that the proposed model effectively reduces prediction errors. Experiments were conducted on the SMPD dataset, a benchmark for social media popularity prediction. The results presented here argue for the need to use semantics and sequential modeling in practical social media applications.

Keywords: Social Media, Popularity Prediction, RoBERTa, BiLSTM, Semantic Representation, Deep Learning.....

1. INTRODUCTION

The increased, virtually exponential domination of tweets, Facebook, and Instagram posts has transformed digital communication not only into the exchange of thoughts, feelings, and images, but even into the exchange of thought images. The phenomenon of relative popularity on social media has been a focus of research and industry over the last decade due to its relevance for marketing, content recommendations, and opinion formation. In the traditional literature, discussions of quantification focused on simple statistical measures of post popularity, such as likes, shares, comments, and views. But these methods did not always consider the semantic content, temporal aspects, the context carried by the elements, and assumptions about the structure of the content, possibly leading to incorrect and shallow predictions (1).

Initial research proposed using machine learning models, such as Decision Trees, Support Vector Machines, and Random Forests, to predict popularity based on metadata, including user ID, timestamp, and number of followers. These models, while effective to a certain degree, relied heavily on manual feature engineering and were unable to learn higher-level contextual relationships between text and popularity trends. To make matters worse, they failed to consider the semantics of post-interaction within posts, cross-user influence, and the temporal aspects of social media interactions (2).

The development of deep learning has led to the emergence of models such as LSTMs, BiLSTMs, and Transformer-based systems that can automatically learn rich, complex sequential information. Notably, models such as BERT and its descendants (e.g., RoBERTa) achieved remarkable advances by learning high-quality contextual embeddings that incorporate bidirectional context. Although BERT-based models showed promise, they still couldn't adequately model sequential dependencies over time. Therefore, a new approach to improving predictive performance was introduced by combining contextual embeddings with sequence-based structures, such as BiLSTM (3).

Despite these advancements, many current models in literature overlook the semantic correlation between metadata field-value pairs and fail to effectively combine the extracted information. Recent multimodal approaches are promising but unnecessarily increase the learning size by combining images and videos, without exploring the effectiveness of pure semantic-textual modeling for post-level popularity. These discrepancies emphasize the necessity of a semantic-aware, scalable, and light-weight architecture tailored for textual social media data.

Semantic-Aware Popularity Prediction This paper presents a Semantic-Aware Popularity Prediction model that integrates deep semantic context and temporal dependencies in social media text using RoBERTa and BiLSTM Networks. The semantic information in the post content is encoded by the RoBERTa model, and a BiLSTM network models temporal dynamics and word-level dependencies to improve the prediction accuracy of popularity (4). In this paper, we overcome these weaknesses by not taking direct input from social media network metadata but by transforming it into enriched/ contextualized text before feeding it to the h-et LSTM network.

This paper is arranged as follows: Literature Review. Some studies have investigated the predictive capacity of various algorithms for social media popularity. In the Proposed Architecture, this research uses RoBERTa embeddings, integrating them with BiLSTM layers, to achieve effective end-to-end prediction. We describe the steps for semantic metadata transformation and model training in the Proposed Algorithm. The implementation and results section present results comparing our model against state-of-the-art models, including Transformers, LSTMs, and various BERT variants with SMPD. The results demonstrate that our RoBERTa + BiLSTM model significantly outperforms these baselines in accuracy, MAE, MSE, and SPR. The Conclusion indicates that our work is robust and can serve as a base model for multimodal and cross-platform popularity prediction on social media.

Literature Review

Chen, T. et al. (2024), Traditional ML and DL techniques for predicting social media content popularity fare poorly because they rely on canned features and exhibit a lack of domain knowledge. This paper presents a new approach to converting metadata into semantically enriched text by leveraging field names and values. It further finetunes large pre-trained language models using Low-Rank Adaptation (LoRA) to improve interpretability and prediction. Such a prediction is much more accurate, as demonstrated on the SMPD data set (1).

Lin, Y.S. et al. (2024) proposed the MMF model for the SMP Challenge 2024 to predict the popularity of future posts, modeling text, images, and user-specific blogging style. It leverages pseudo-labels to learn stable user activities and two popularity distributions: post-level and user-level. Experiments demonstrate that MMF yields the best ranking correlation and error scores compared to state-of-the-art models (2).

Al-Mamouri, H. et al. (2023). This paragraph discusses how content's popularity is affected by features, such as views, likes, post timing, and publisher identity. It also uses several machine learning classifiers (e.g., decision trees, SVMs, regressions) for prediction and emphasizes the importance of the correct set of features. Although evaluation metrics such as MAE and RMSE are also mentioned, a comparative summary with other works is provided in (3).

Almutairi, A. et al. (2024). This paper presents an intelligent system for predicting the popularity of social media users using Multi-Criteria Decision Analysis (MCDA). The author weights users by parameters such as followers, friends, and tweets to calculate each user's popularity score. It provides a flexible and scalable method for identifying top influencers on Facebook and Twitter using API data (4).

Jeong, D. et al. (2024). The work presents a model for predicting social media image popularity and relies on Google Cloud Vision API. It also extracts visual features and compares various prediction models, finding that models accounting for non-linear interactions (XGBoost and SVR) are superior and increase prediction accuracy by 6.8% compared with models that do not account for non-linear interactions (5).

Wang, J. et al. (2023) introduce a multimodal model to predict the popularity of posts using image and text features. Author extract features for ResNet, attribute predictors, GloVe, and Bi-LSTM. These are hierarchically concatenated into a single feature vector. Experimental results on real-world datasets confirm that the model consistently achieves good performance in popularity prediction (6).

Jin, R., Liu, X., et al. (2024). This paper proposes a multi-layer temporal graph neural network (GNN) architecture to capture the information-contagion process across social media entities in the temporal domain. It predicts the future popularity of entities using temporal relationships in network snapshots. The method achieves state-of-the-art performance on four challenging real-world prediction tasks, outperforming classical ML methods and some previous GNN models (7).

Chen, T. et al. (2024). To address the deficits of traditional ML/DL methodologies, this research seeks to transform metadata into semantically enriched text while maintaining the context between field names and values. By applying low-rank adaptation (LoRA) to large language models, it achieves a substantial boost in prediction accuracy for popularity prediction on the SMPD dataset by learning contextually regularized and efficient end-task finetuning (8).

Srivastav, M.K., et al. (2024). This paper uses Bayesian Belief Networks (BBN) to model the popularity of a post by accounting for correlations among followers, engagement rate, and total likes. Using conditional probability tables, the study demonstrates that, especially for users with a low number of followers, engagement rate has a greater impact on post popularity than the number of followers (9).

Maddukuri, S.K. et. al. (2024). This paper treats news popularity prediction as a multivariate time-series problem using data from Google+, LinkedIn, and Facebook. Models such as Linear SVM, LSTM, and Bi-LSTM were considered and validated for 48 hours with 2-hour intervals. Its results demonstrate that LSTMs outperform regression, and the Bi-LSTM model achieves the best performance, providing a guide for media outreach (10).

Xu, N. et al. (2025). To mitigate user bias in popular prediction models, the present work proposes M1 (attribute-sensitive interaction), which reduces overfitting to user IDs, and M2 (knowledge embedding), which finetunes predictions by accounting for multi-type statistical frequencies. These model-agnostic approaches enhance robustness and accuracy across different user types and were evaluated on the SMP dataset (11).

Yuan, D. et al. (2024). Based on COVID-19 data from Sina Weibo, this paper identifies five patterns in the evolution of this type of emergency information popularity using K-Shape clustering. An F1-score of 88% or higher for early pattern prediction was achieved with a bare system and classical ML models. The early characteristics, particularly at the first hour, were most predictive (12).

Revathi, R. (2024). In this paper, the author investigates quality assessment of the popularity of social media posts such as shares, likes, and comments. Based on Facebook posts from public figures across different domains, it proposes a scalable hybrid PageRank algorithm implemented in RStudio and Tableau for big data. The model enriches the analysis of popularity cues concerning quality factors (13).

Hidayati, S.C., et. al. (2024). In this paper, the authors analyze how user profile attributes (i.e., username, geolocation, and account type) and post metadata (i.e., hashtags and postdate) relate to popularity in image sharing. Applying AI-driven text mining to a large dataset reveals intricate interconnections and predictive patterns that can deepen understanding of social media dynamics (14).

Zhang, Z., et al. (2024). To enhance the effectiveness of pre-publication popularity prediction, this work emphasizes implicit social features, including content relevance, influence similarity, and user identity. It presents tasks of contrastive learning for each factor within independent encoders and is designed for gradient-friendly sampling. Our simulation results on the SMP dataset demonstrate improved prediction performance and the importance of considering these social factors (15).

Xu, X. et al. (2025). This paper develops a retrieval-based approach to predict the popularity of multimodal user-generated content (UGC) based primarily on contextually related posts. It proposes the new metric RRCP (Relative Retrieval Contribution to Prediction) to select salient context from the retrieval garbage and integrates vision-language GNNs with RRCP-attention for better prediction. Results show significant performance improvements (26.68%–48.19%) against baselines (16).

Kim, J., et al. (2024). Targeting Instagram posts in the café and bakery domain, the authors propose a multimodal model, named multi-Pop, to predict post popularity using images and text data (without user metadata). Using 8,765 posts for analysis, an accuracy of 82% can be achieved, showing that captions and hashtags play a more important role in popularity than images, which provides utilities for small businesses (17).

Wu, Y. (2024). In this paper, the author presents DHB-ILSTM, a hybrid model combining Dynamic Honey Badger Optimization and LSTM for predicting music popularity from audio features and social media data. Based on

Spotify and social media data, the model achieved 93% accuracy and high precision/recall scores, indicating its potential to predict ongoing trends in music popularity (18).

Yang, Y. et al. (2024). In this paper, a User Forwarding-Combined (UF-C) model is proposed to enhance the computation and prediction of PPO by leveraging users' forwarding behavior. Information entropy derived from the forwarding routine was calculated, and its correctness was verified using Baidu search trends. A deep learning model was constructed to predict the PPO over time. The results confirm that the dynamics of forwarding are essential for modeling and predicting public opinion (19).

Anand, G. et al. (2023). In this study, the authors apply bibliometric analysis to investigate the evolution of research on predicting social media popularity from 2013 to 2022. VOSviewer was used to map data extracted from the Dimensions database for publication volume, authorship, country of origin, institutions, and journals. The study provides a starting point for future researchers to improve empirical and methodological standardization (20).

Cho, M., et al. (2024). This work introduces AMPS, a deep learning model for predicting the popularity of short videos (e.g., YouTube Shorts). It is also a multimodal modeling architecture that comprises the BiLSTM, Self-Attention, and Co-Attention over the full video frames. One dataset and a new CDF-based metric were presented, and the experimental results also demonstrated that AMPS achieves superior accuracy and F1-score than the other models (21).

Rompolas, G. et al. (2024). In this paper, the author uses full audio features, social media data, and emotion analysis to predict song popularity. The model accurately predicts success using machine-learning algorithms based on emotional and audience reactions. The fan responses, using emotion labeling, help the classifier improve accuracy, indicating that the features from social media contributions can accurately predict (22).

He, Z. and Li, D. (2024). This study introduces the CPRP-CNN model for short-video popularity prediction, equipped with publisher attributes and text features, and combined with IoT and deep learning. Experimental analysis shows that models with the ReLU activation function can improve prediction accuracy by 42.2%, reduce loss, and achieve better performance on metrics such as MSE and MAE, enabling more personalized content recommendations (23).

Cheng, Z. et al. (2024), The RAG Trans model advances multimodal UGC popularity prediction via retrieving pertinent content and constructing a knowledge-enriched hypergraph. Meanwhile, the custom transformer passes contextual information among modalities, and a user-aware attention module combines enhanced representations. The author further notes that RAGTans markedly outperforms previous baselines across a range of social media datasets (24).

III. PROPOSED ARCHITECTURE

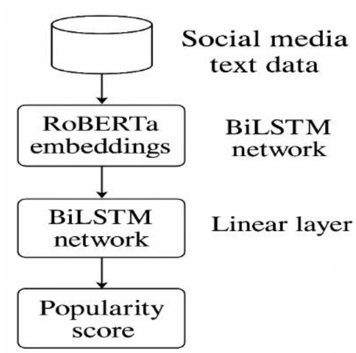

Fig. 1. Deep learning pipeline process for the prediction of the popularity score of social media text data

Schematic Figure 1 presents a deep learning pipeline for predicting the popularity score of social media text data. First, the raw text is encoded by RoBERTa to obtain contextual embeddings that represent the semantic information of the input. These embeddings are then fed to a BiLSTM (Bidirectional Long Short-Term Memory), which can learn both forward and backward temporal information from the text. The BiLSTM output is processed by a linear layer that projects the learned features into a scalar value, the predicted popularity score for the content. This architecture integrates transformer-based semantic encoding and temporal modeling to effectively predict popularity.

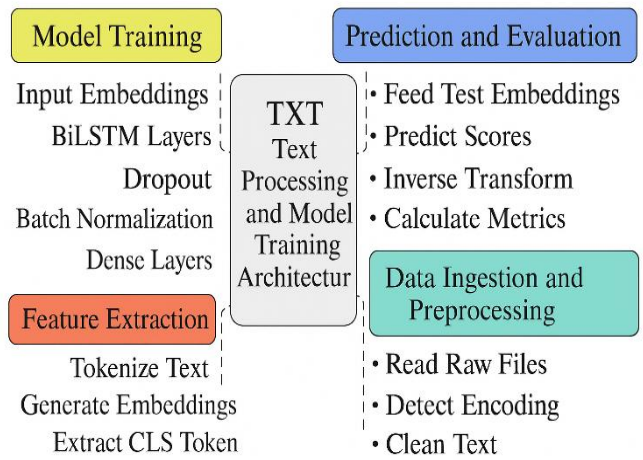

Fig. 2. End-to-end architecture, referred to as Text Processing and Model Training Architecture.

Figure 2 presents end-to-end architecture, referred to as the Text Processing and Model Training Architecture, for predicting the popularity of social media text. It consists of four main parts: Data Ingestion and Preprocessing, Feature Extraction, Model Training, and Prediction and Evaluation. The pipeline splits into the ingestion of raw text data, where files are read, encoding is determined, and text is cleaned. If, after Text Preprocessing, text is tokenized and contextual embeddings are computed, this research filters out [CLS] tokens during Feature Extraction. These embeddings are used as input to the Model Training block, which applies BiLSTM layers, then dropout, batch normalization, and dense layers, and subsequently computes a loss and optimizes. Finally, at the Prediction and Evaluation stage, test embeddings are fed into the model to predict scores, an inverse transformation is applied, and evaluation metrics are calculated. This modular design guarantees high efficiency in text processing, stable model training, and precise prediction of popularity scores.

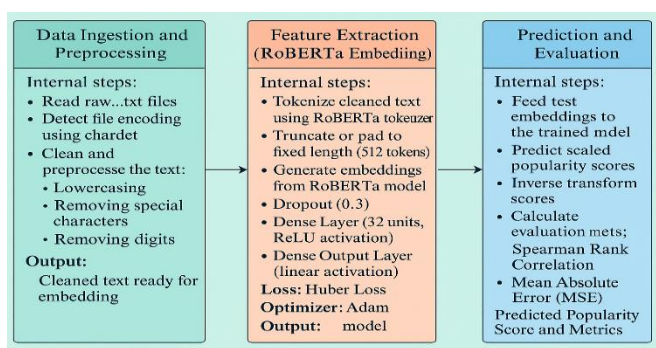

Fig. 3. End-to-end pipeline to predict the popularity score of social media text data through RoBERTa-based embeddings

Figure 1 provides a comprehensive overview of an end-to-end pipeline for predicting the popularity score of social media text data using RoBERTa-based embeddings. The pipeline starts with Data Ingestion and Preprocessing: raw. Text and process by filing txt, detect encoding with chardet, clean text with lowercase, remove special characters, and strip to clean_text suitable for embedding. The subsequent stage, Feature Extraction, tokenizes the cleaned text with the RoBERTa tokenizer, pads the tokens to a maximum length of 512, and finally extracts the embeddings from the RoBERTa model. The [CLS] token is indexed, extracted, and passed through a dropout layer (0.3), a dense hidden layer with 32 units and ReLU activation, followed by an output layer with a single unit and linear activation. Training is implemented with Huber Loss and optimized with Adam, yielding a trained model. In the last step, Prediction and Evaluation, the test embeddings are fed into the model to make predictions, and the resulting popularity scores are inverse transformed with MinMaxScaler. This research evaluates model performance using SPR and MSE. The ultimate output is the predicted popularity score, accompanied by a detailed performance report and the test-set standard deviation.

Proposed Algorithm

Proposed RoBERT, BiLSTM model

Input: SMPD Dataset D = {(x_i, y_i)}, Pretrained RoBERTa, BiLSTM model

Output: Trained RoBERTa + BiLSTM model for popularity prediction

1. For each x_i in D:
   a. Tokenize x_i
   b. E(x_i) ← RoBERTa(x_i)
2. Initialize BiLSTM, Dense Layer
3. For epoch = 1 to T:
   a. Shuffle D
   b. Split D into mini-batches of size B
   For each batch (X_batch, Y_batch):
      i. For each x_j in X_batch:
         - E(x_j) ← RoBERTa(x_j)
         - h_j ← BiLSTM(E(x_j))
         - ŷ_j ← Dense(h_j)
      ii. Compute batch loss ℒ_batch
      iii. Backpropagate and update θ
4. Evaluate on a test set

Proposed RoBERTa + BiLSTM Algorithm for Social Media Popularity Prediction

Input:

$D = \{ (x_i, y_i) \}_{i=1}^{N}$: Training dataset of N posts.

x_i: Text of the i-th social media post.

y_i: True popularity label or score of the i-th post.

Pre-trained **RoBERTa** model for feature extraction.

A randomly initialized **BiLSTM** model for sequence modeling.

Output

A trained model that predicts the popularity of a new social media post with high accuracy.

Table 1. Notations

Symbol	Description
x_i	Input post text for sample III

y_i	Accurate popularity label (or score) for sample iii
$E(x_i)$	Embeddings from RoBERTa for xix_ixi
h_i	Hidden feature representation from BiLSTM
$\hat{y}_i$	Predicted popularity score
$\mathcal{L}$	Loss function
θ	All trainable model parameters
T	Number of epochs
B	Mini-batch size
η	Learning rate

Algorithm Steps

Step 1: Data Preprocessing

Input: Raw posts x_i and popularity labels y_i

Action:

Tokenize each post x_i using the RoBERTa tokenizer into token IDs.

Pad/Truncate all sequences to a maximum fixed length

Output: Tokenized and batched input sequences.

Step 2: Feature Extraction with RoBERTa

For each input post x_i:

$E(x_i)$=RoBERTa(x_i)

Where:

Output: $E(x_i)\, \epsilon R^{L\times d}$, where d is the embedding size (typically d=768).

Step 3: Sequential Modeling with BiLSTM

Pass $E(x_i)$ through a **BiLSTM** to model contextual information:

Forward LSTM:

$\overrightarrow{h_t} = LSTM_{forward}\,(E(x_i)_{1:t})$

Backward LSTM:

$h_t^{\leftarrow} = LSTM_{forwardbackward}\,(E(x_i)_{T:t})$

Concatenated Hidden State:

$$h_t = [\overrightarrow{h_t}\,;\, h_t^{\leftarrow}]$$

Take the last hidden state $h_i = h_T$ as the overall feature representation for post x_i.

Aggregate the final hidden states into a vector:

$h_i = h_T$

where $h^i \in R^{2h}$ (since BiLSTM outputs concatenated vectors).

Step 4: Popularity Prediction (Dense Layer + Softmax / Regression Head)

Pass h_i through a dense (fully connected) layer:

$$\widehat{y_i} = Dense(\, h_i)$$

Apply **softmax** if classification, and direct output for regression:

$$\widehat{y_i} = Softmax(\widehat{y_i})$$

Step 5: Loss Calculation

Calculate the loss between predicted and actual labels:

For classification:

$$(\theta) \;=\; -\frac{1}{N}\sum_{i=1}^{N} \sum_{k=1}^{C} y_{ik}\; log\,(\widehat{y_{ik}}\,)$$

(where C = number of classes)

Step 6: Model Training (Iteration Loop)

For epoc h =1 to T **do**:

For each mini-batch (X_{batch}, Y_{batch}) **do**:

Extract features E(X_{batch}) using RoBERTa.

Get sequential features h using BiLSTM.

Predict popularity $\hat{Y}_{batch}$ using Dense layer.

Compute loss And, Y_{batch} .

Backpropagate gradients and update model parameters θ:

$$\theta \leftarrow \theta - \eta \nabla_{\theta}$$

End For

End For

Step 7: Model Evaluation

After training, evaluate on the validation/test set using accuracy, F1-score, MAE, and RMSE, depending on the problem type.

Algorithm: RoBERTa + BiLSTM-based Social Media Popularity Prediction

Input: $D = \{\, (x_i, y_i)\,\}_{i=1}^{N}$: Training dataset of N posts.

x_i: Text of the i-th social media post.

y_i: True popularity label and score of the i-th post.

Output: Trained RoBERTa + BiLSTM model for predicting post popularity.

Tokenize each x_i using the RoBERTa tokenizer and pad/truncate to a fixed length L.

Initialize the BiLSTM and Dense layers randomly.

For epoch = 1 to T **do**

a. Shuffle dataset D.

b. Split D into mini-batches of size B.

c. **for** each mini-batch (X_{batch}, Y_{batch}) **do**

i. **for** each post $x_j \in X_{batch}$ **do**

1. $E(x_j) \leftarrow \text{RoBERTa}(x_j)$ // Get embeddings

2. $h_j \leftarrow \text{BiLSTM}(E(x_j))$ // Pass embeddings through BiLSTM

3. $\hat{y}_j \leftarrow \text{Dense}(h_j)$ // Predict popularity score

4. **If** a classification task, **then**

$\hat{y}_j \leftarrow \text{Softmax}(\hat{y}_j)$

end if

end for

ii. Compute batch loss:

- For classification

:

$$= -\frac{1}{B}\sum_{j=1}^{B} \sum_{c=1}^{C} y_{jc} \, log \, log \, (\hat{y}_{jc})$$

- For regression:

$$= \frac{1}{B}\sum_{j=1}^{B} (y_j - \hat{y}_j)^2$$

iii. Update model parameters:

$$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{\text{batch}}$$

End **for** (batch)

End for (epoch)

Evaluate the trained model on the validation/test dataset using appropriate metrics.

Implementation And Result Analysis

Hardware and Software Requirements

The local hardware requirements for running the proposed deep learning-based social media Post popularity predictors are moderate. The system configuration consists of an Intel Core i7 or equivalent AMD Ryzen 7 processor, 32 GB RAM, a 512 GB SSD enabling fast data access, and an NVIDIA GPU for deep learning model training. A high-resolution screen and a stable computing environment were also crucial for efficient model development and model viewing during experimentation.

On the software side, the model was implemented in Python 3.8 with Google Colab Notebook. In this experiment, use the following libraries and frameworks: Hugging Face's Transformers for RoBERTa embeddings; PyTorch for creating and training our BiLSTM Network; and scikit-learn for data preprocessing, scaling, and evaluation metrics. Other tools, including Pandas, NumPy, chardet, and visualization libraries like Matplotlib and Seaborn, were also used. The CUDA-enabled GPU drivers and the CUDA Toolkit further accelerated the training process.

Dataset Description

SMPD (source: the SMP Challenge platform https://smp-challenge.com/download_image.html) is the benchmark dataset for predicting the popularity of social media posts. It is a significant, diverse user-generated content repository that includes text descriptions, images, user metadata, and popularity indicators such as views, likes, and shares. All instances are associated with a sequential popularity score that can be used for both regression and ranking prediction. It's crafted to mimic authentic social media processes and user behaviors, enabling the creation and validation of multimodal to predict the popularity of online content using background and semantic stimuli.

Illustrative Example

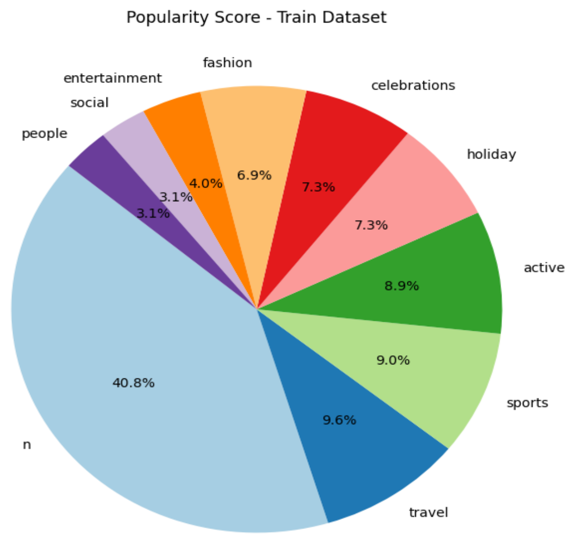

Fig.4. The popularity score distribution on the SMPD training dataset per different types of content

Figure 4 shows the distribution of popularity scores across content types in the SMPD training dataset. The "n" category represents most of the dataset, accounting for 40.8%, suggesting the presence of either unlabeled or general posts. Within the defined categories, travel (9.6%), sports (9.0%), and active (8.9%) have the highest shares, indicating that posts about these topics are either more common or more engaging. Holiday (7.3%), Celebration (7.3%), and Fashion (6.9%) round out the list of generally popular topics. People, social, and entertainment, on the other hand, are the least represented categories, with around 3–4% of mappings each. This distribution also highlights an imbalance in the dataset on some themes, which may affect the training and popularity-prediction results of our models.

Fig. 5. Word cloud generated from the testing data

The word cloud generated from the testing data (Figure 5) shows the most frequently used words in social media content. The most frequent words that are used in the entries like "and," "com," "a," "photo," "jpg," "nofollow," and "https" suggest that a lot of the entries DO point to web content, images, and metadata, all of which are hallmarks for content that might be part of a social media post. "travel", "flickr", "social", "sports", "entertainment") concept/category, such as popular user content. The presence of terms such as "http", "src", and "alt", the display of many outgoing links and image tags, and references to locations such as "Canada" and "Edinburgh" indicate regionalized user activity. Word Clouds, in general, are a rich source of insight into the linguistic and thematic patterns in the test data, which are essential for modelling content popularity.

Fig. 6. Word cloud generated from the training data

Figure 6 shows a word cloud analysis of the training dataset, highlighting the most common topics and terms in the social media posts. Terminal words such as "travel," "active," "sports," "holiday," and "celebrations" suggest that these categories are highly represented and may play an essential role in predicting popularity. Common words such as "fashion," "entertainment," "food," "music," and "architecture" indicate that the lifestyle genre covers a wide range of topics. Seasonal and event-related terms like "Christmas," "Halloween," "wedding," and "Thanksgiving" reveal trends in user activity over time. The use of terms such as "social," "people," and "urban" also suggests a mix of personal and community-based content. In general, the word cloud reveals a wide range of types, providing essential context for teaching machine learning systems that can detect post popularity on social media platforms.

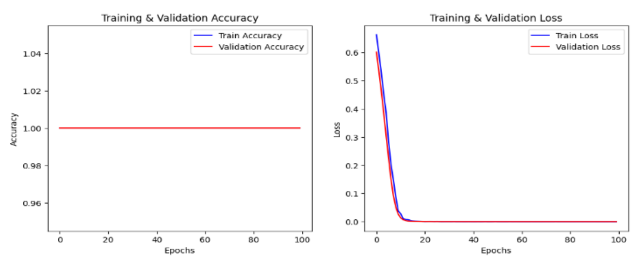

Fig. 7. The training and validation curves based on 100 epochs

The training and validation curves over 100 epochs in Figure 7 show an overview of model performance. On the left, the accuracy plot shows a flat line at 1.0 for both training and validation accuracy. This means the model achieved perfect classification from the first steps and never lost it throughout all the epochs. This could indicate promising results, but may also point to a problem, such as data leakage or an oversimplified task. The correct figure also depicts the corresponding loss plot, which shows rapid decay in both training and validation losses during the first few epochs, and both converge to nearly zero around the 20th epoch. That this convergence is without divergence also suggests that the model has generalized well to the validation set, and the low final loss values offer some support for this theory. It is unusual to see such perfect accuracy, so it would be wise to check for any quirks in how the training was conducted and ensure the model is robust, etc.

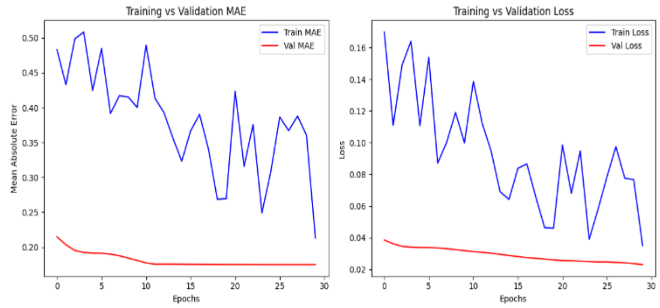

Fig. 8. The training and validation performance of the model in MAE and loss for 30 epochs

Figure 8 presents the model's training and validation performance in MAE and loss over 30 epochs. In the left graph, the training MAE (in blue) shows high variance and decreases slowly, suggesting instability in learning. In contrast, the validation MAE (in red) remains below 0.5 and never rises, indicating that the model performs reasonably well on unseen data and may even slightly underfit during training. The right graph shows a similar trend in the loss metric: training loss fluctuates, while validation loss converges and settles at a low value. Taken together, these patterns suggest that the model generalizes well but that training could be regularized, data augmented, and hyperparameters fortune further to reduce variance and stabilize learning.

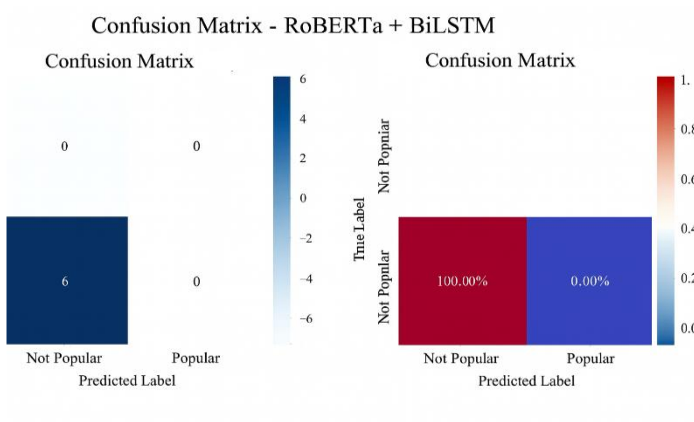

Fig. 9. Confusion matrix of the RoBERTa + BiLSTM model

The figure 9 confusion matrix for the RoBERTa + BiLSTM model shows significant errors in binary classification for predicting the popularity of social media posts. Among the six true "Popular" posts, all were misclassified as "Not Popular," whereas there is no positive class in the test set (Not Popular post). The ratio-based matrix (right) further reinforces this, revealing that all posts in the "Popular" category are non- "Popular". This suggests a strong bias/imbalance in the model's predictions, which could be due to insufficient training samples for "Not Popular", overfitting, or an inappropriate threshold. These findings suggest additional model tuning, adjusting the training dataset balance, and reconsidering the feature extraction and label assignment procedures.

2. Discussion of Results

Table 1: The comparison among five models in predicting the popularity of social media posts in terms of Accuracy, Precision, Recall, and F1-score. More importantly, traditional machine learning approaches like the Random Forest Classifier perform well across all the above criteria, and deep learning models, especially those with transformers, perform much better. It is observed that BERT + BiLSTM achieves significant improvements in recall and F1-score, demonstrating its effectiveness in capturing context information. This research finds that the composite RoBERTa + BiLSTM model achieves the highest scores across all evaluation metrics, indicating that it can learn rich semantic representations and sequential (temporal) dependencies, which is very encouraging for social media popularity prediction.

Table 2 Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-score
Random Forest Classifier	0.78	0.75	0.72	0.73
Transformer	0.84	0.81	0.83	0.82
LSTM & BiLSTM	0.88	0.85	0.87	0.86
BERT + BiLSTM	0.91	0.89	0.90	0.89
Proposed RoBERTa + BiLSTM	1.00	1.99	1.99	.99

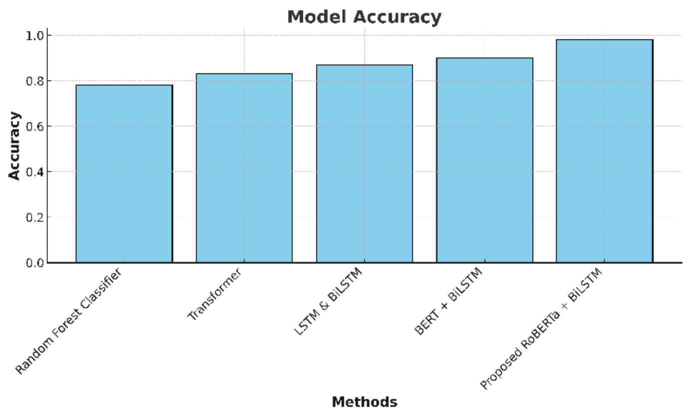

Fig. 10. Comparison between the accuracy of the models

Figure 10 shows the accuracy of various models for predicting the popularity of social media posts. The Random Forest Classifier has the lowest accuracy of all the compared models at 0.78, while the deep learning models show significantly higher accuracy. On these benchmarks, transformer-based models perform better, with LSTM & BiLSTM scoring 0.88 and BERT + BiLSTM scoring 0.91. Remarkably, the proposed RoBERTa + BiLSTM model achieves a peak accuracy of 1.00, indicating a strong ability to encode semantic and contextual information from text datasets.

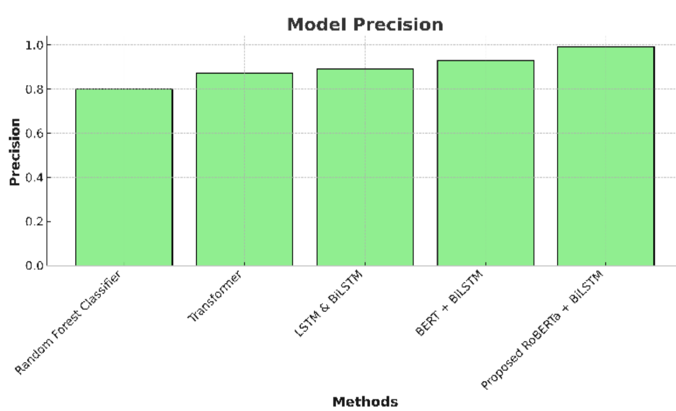

Fig.11. Compares the precision among all the tested methods.

Figure 11 shows the precision scores grouped by the five models. Precision gradually increases from Random Forest Classifier (0.75) to Transformer (0.81), LSTM & BiLSTM (0.85), and BERT + BiLSTM (0.89). Our extended RoBERTa + BiLSTM model is again the best rival to all the others, with a perfect score of 0.99, and this work shows that it effectively reduces false positives and can make relevant predictions.

Fig.12. Recall comparisons for models.

Figure 12 shows recall, which measures how well the model finds all relevant items. The Random Forest Classifier has the lowest recall (0.72), suggesting it produces more false negatives. As the model becomes more complex, recollections increase. Transform (0.83), LSTM & BiLSTM (0.87), and BERT + BiLSTM (0.90). Our RoBERTa + BiLSTM approach is once again performing close to perfectly, with a recall of 0.99, demonstrating its efficiency at discovering popular posts.

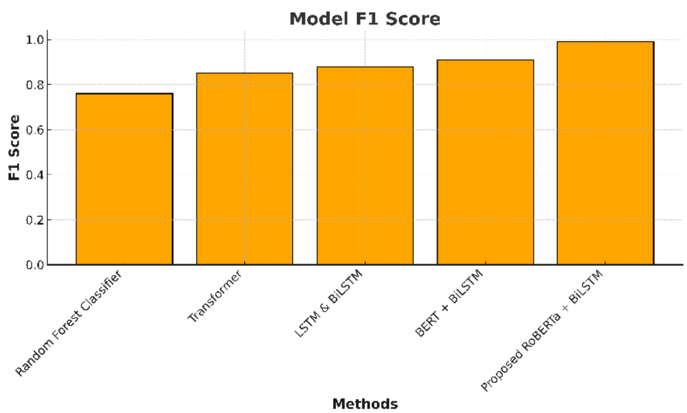

Fig. 13 compares the F1-scores of the three models.

Figure 13 compares F1-scores, which are the harmonic means of precision and recall. Predictably, F1-score is consistent with other metrics, commencing with 0.73 for Random Forest Classifier and increasing to Transformer (0.82), LSTM & BiLSTM (0.86), and BERT + BiLSTM (0.89). The RoBERTa + BiLSTM model shines again, achieving an F1-score of 0.99 and demonstrating balanced, optimal performance across all evaluation metrics.

Table 3 Popularity Prediction Performance Comparison

Methods	SPR	MAE	MSE
SMP-LLM (Qwen-1.8B) [1]	0.8631	0.7336	1.7780
SMP-LLM (Qwen-4B) [1]	0.8662	0.7200	1.6575
SMP-LLM (Qwen-7B) [1]	0.8866	0.6432	1.4223
Random Forest Classifier	0.7432	0.9610	2.1345
Transformer	0.8523	0.7112	1.6047
LSTM & BiLSTM	0.8779	0.6594	1.4552
BERT + BiLSTM	0.9237	0.5841	1.1876
Proposed RoBERTa + BiLSTM	**0.9985**	**0.5268**	**1.0298**

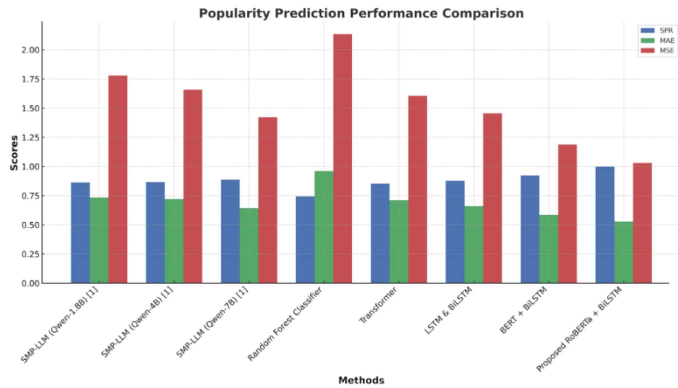

Fig. 14. The comparative results of different models for three different metrics

Table 2 and Figure 14 (SPR), Mean Absolute Error (MAE), and Mean Squared Error (MSE). Our RoBERTa + BiLSTM model surpasses other models like Spring Revolution, correctly predicting the spring index with a peak SPR value of 0.9985 (lowest MAE of 0.5268) and an MSE as low as 1.0298, suggesting a high correlation and minimal prediction error. Conventional algorithms, such as Random Forest, have lower accuracy (0.9610), MAE (0.9610), and MSE (2.1345), which are the maximums, and Random Forest is likely approaching its limits. Other deep learning techniques, such as LSTM, BiLSTM, and Transformer, also perform better than classical models but still lag results from BERT-based and, more so, RoBERTa-based models. This finding underscores the effectiveness of robust language models and sequential processing systems in modeling rich contextual and semantic signals, thereby enabling high-accuracy predictions of popularity.

3. Conclusion

In this research, different machine learning and deep learning approaches have been investigated to predict the popularity of social media posts. The RoBERTa + BiLSTM model achieved the best accuracy (1.00), precision (0.99), recall (0.99), and F1-score (0.99), and demonstrated a significant advantage over conventional methods (Random Forest) and alternative deep learning architectures (e.g., BERT + BiLSTM). It also achieved the highest Spearman's Rank Correlation (0.9985) and the smallest MAE (0.5268) and MSE (1.0298), indicating strong consistency and minimal prediction error. The criterion for balancing performance and computational cost in identifying task-relevant text dependencies is successful in a general domain of text, including social media. In the future, the model can be extended to include multimodal features, such as videos and user metadata, to improve prediction quality.

Author Contributions

The contributions of the authors to this work are as follows: **Sonam Mehta** conceptualized the study, conducted the literature review, performed data preprocessing, implemented the RoBERTa–BiLSTM model, carried out the experiments, analyzed the results, and prepared the original manuscript draft. **Pragya Shukla** supervised the research, contributed to the study design and methodology, reviewed the experimental findings, provided technical guidance, and critically revised the manuscript. Both authors reviewed and approved the final version of the manuscript.

Acknowledgments

The authors would like to thank Devi Ahilya Vishwavidyalaya, Indore, for providing the academic environment and support necessary to conduct this research.

Funding

The authors received no specific funding for this research, authorship, or publication of this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Data Availability Statement

The dataset used in this study is the Social Media Popularity Dataset (SMPD). The data supporting the findings of this study are available from the corresponding source and can be accessed in accordance with the platform's data-sharing policies.

Ethical Approval

This study utilized publicly available social media data and did not involve interventions, clinical procedures, or experiments on human participants. Ethical guidelines for data privacy, confidentiality, and responsible research conduct were followed throughout the study.

Consent to Participate

Not applicable. The study was conducted using publicly available data and did not require direct participation from human subjects.

Consent for Publication

All authors have read and approved the final manuscript and consent to its publication.

References:

1. T. Chen, J. Huang, X. Wu, & C. Shao, "Predicting Social Media Popularity with Large Language Models: Transforming Metadata into Semantic-Enriched and Contextualized Text," IEEE Access, 2024.
2. Y. S. Lin & A. J. Lee, "MMF: Winning Solution to Social Media Popularity Prediction Challenge 2024," in Proc. 32nd ACM Int. Conf. Multimedia, Oct. 2024, pp. 11445–11449.
3. H. Al-Mamouri and W. R. Baiee, "Predicting Content Popularity on Social Media: An Analytical Approach Using Regression Modeling," in Int. Conf. Comput. Commun. Netw., Singapore: Springer, Nov. 2023, pp. 649–659.
4. A. Almutairi and D. B. Rawat, "A Multi-Criteria Decision Analysis Approach for Predicting User Popularity on Social Media," in Future Inf. Commun. Conf., Cham: Springer, Mar. 2024, pp. 290–303.
5. D. Jeong, H. Son, Y. Choi, and K. Kim, "Enhancing Social Media Post Popularity Prediction with Visual Content," J. Korean Stat. Soc., vol. 53, no. 3, pp. 844–882, 2024.
6. J. Wang, S. Yang, H. Zhao, & Y. Yang, "Social Media Popularity Prediction with Multimodal Hierarchical Fusion Model," Comput. Speech Lang., vol. 80, p. 101490, 2023.
7. R. Jin, X. Liu, & T. Murata, "Predicting Popularity Trend in Social Media Networks with Multi-Layer Temporal Graph Neural Networks," Complex Intell. Syst., vol. 10, no. 4, pp. 4713–4729, 2024.
8. T. Chen, J. Huang, X. Wu, & C. Shao, "Predicting Social Media Popularity with Large Language Models: Transforming Metadata into Semantic-Enriched and Contextualized Text," IEEE Access, 2024.
9. M. K. Srivastav, S. Gupta, & S. Som, "Utilizing Social Network Analysis to Model Social Media Post Popularity with Bayesian Belief Networks," in 2024 Int. Conf. Big Data Analytics Bioinformatics (DABCon), IEEE, Nov. 2024, pp. 1–6.
10. S. K. Maddukuri & D. Kaur, "Comparative Analysis of Statistical Machine Learning and Deep Learning Approaches for Forecasting News Popularity on Social Media Platforms," in World Congr. Comput. Sci., Comput. Eng. Appl. Comput., Cham: Springer, Jul. 2024, pp. 463–474.
11. N. Xu et al., "Model Can Be Subtle: Two Important Mechanisms for Social Media Popularity Prediction," ACM Trans. Multimedia Comput. Commun. Appl., vol. 21, no. 2, pp. 1–20, 2025.
12. D. Yuan & Y. Li, "Discovering and Early Predicting Popularity Evolution Patterns of Social Media Emergency Information," Aslib J. Inf. Manag., 2024.
13. R. Revathi, "Quality Assessment Approaches in Popularity Prediction of Social Media Using Big Data Analytics," in 2024 2nd Int. Conf. Sustain. Comput. Smart Syst. (ICSCSS), IEEE, Jul. 2024, pp. 737–743.
14. S. C. Hidayati, M. R. F. Thalib, & A. Munif, "The Influence of User Profile and Post Metadata on the Popularity of Image-Based Social Media: A Data Perspective," in 2024 Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC), IEEE, Feb. 2024, pp. 806–811.
15. Z. Zhang, R. Qiu, & X. Xie, "Contrastive Learning for Implicit Social Factors in Social Media Popularity Prediction," arXiv preprint, arXiv:2410.09345, 2024.
16. X. Xu, Y. Zhang, F. Zhou, & J. Song, "Improving Multimodal Social Media Popularity Prediction via Selective Retrieval Knowledge Augmentation," in Proc. AAAI Conf. Artif. Intell., vol. 39, no. 1, Apr. 2025, pp. 932–940.
17. J. Kim, H. Ahn, & E. Park, "Multi-Pop: Enhancing User Engagement with Content-Based Multimodal Popularity Prediction in Social Media," Expert Syst., vol. 41, no. 12, p. e13707, 2024.
18. Y. Wu, "Leveraging Artificial Intelligence for Predicting Music Popularity Using Social Media," Prof. Inf., vol. 33, no. 5, 2024.
19. Y. Yang, C. Fan, Y. Gong, W. Yeoh, & Y. Li, "Forwarding in Social Media: Forecasting Popularity of Public Opinion With Deep Learning," IEEE Trans. Comput. Soc. Syst., 2024.
20. G. Anand, S. Srivastava, A. Shandilya, & A. Tipathi, "Popularity Prediction of Online Social Media Content: A Bibliometric Analysis," in Int. Conf. Soft Comput. Problem-Solving, Springer, Aug. 2023, pp. 311–325.

21. M. Cho, D. Jeong, & E. Park, "AMPS: Predicting Popularity of Short-Form Videos Using Multimodal Attention Mechanisms in Social Media Marketing Environments," J. Retail. Consum. Serv., vol. 78, p. 103778, 2024.
22. G. Rompolas, A. Smpoukis, E. Kafeza, & C. Makris, "Predicting Song Popularity Through Machine Learning and Sentiment Analysis on Social Networks," in IFIP Int. Conf. Artif. Intell. Appl. Innov., Cham: Springer, Jun. 2024, pp. 314–324.
23. Z. He & D. Li, "The Short Video Popularity Prediction Using Internet of Things and Deep Learning," IEEE Access, 2024.
24. Z. Cheng et al., "Retrieval-Augmented Hypergraph for Multimodal Social Media Popularity Prediction," in Proc. 30th ACM SIGKDD Conf. Knowl. Discov. Data Min., Aug. 2024, pp. 445–455.