International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

SOCIAL MEDIA POPULARITY PREDICTION USING MACHINE LEARNING

[1] Dr. Chaganti B N Lakshmi, [2] G Pranavi Reddy, [3] K Nandini, [4] K Vara Laxmi, [5] G Maneesh Rohan

[1] Professor in Department Of CSE, TKR College Of Engineering and Technology

[2 3 4 5] UG. Scholar Department Of CSE, TKR College Of Engineering and Technology

Abstract

Predicting the popularity of social media posts has become increasingly important due to its practical applications in areas such as advertising, recommendation systems, and trend analysis. However, this task is inherently complex, as social media engagement is influenced by numerous unpredictable factors—including content quality, audience relevance, and real-world events—that are difficult to model accurately. Many existing methods attempt to improve prediction accuracy by incorporating a wide range of features and modalities. However, they often treat all features equally, without considering their varying degrees of importance. To address this limitation, we propose a novel approach that leverages the self-attention mechanism to automatically and effectively fuse different types of features, allowing the model to focus on the most relevant information for popularity prediction. Our model primarily utilizes two categories of features: semantic features (derived from text content) and numeric features (such as metadata and engagement metrics). We validate our approach through extensive experiments and ablation studies conducted on the training and testing sets of the ACM Multimedia SMPD 2020 Challenge dataset. The results demonstrate that our method outperforms existing techniques, highlighting its effectiveness in handling multi-modal data for social media popularity prediction. This work is grounded in machine learning, a key area of data science that uses statistical methods to train algorithms capable of making predictions and uncovering patterns from data. By learning directly from training data, our model can make accurate predictions without the need for hard-coded rules.

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

Such machine learning approaches are especially valuable in scenarios where traditional algorithmic solutions are difficult or infeasible to implement.

I INTRODUCTION

As social media platforms continue to grow rapidly, the ability to predict the popularity of online content has become increasingly important. Whether it's for optimizing digital marketing strategies, powering recommendation engines, or analyzing social trends, accurate popularity prediction can offer significant value across various real-world applications.

However, predicting how a post will perform on social media is far from straightforward. Popularity is influenced by a mix of factors—like how relevant the content is to current events, the interests of the audience, the quality of the content itself, and even when it is posted. These elements are not only diverse but also interdependent, making them difficult to model using traditional approaches.

Many existing models attempt to tackle this problem by incorporating as many features as possible from different data types or "modalities." While this multi-modal approach is a step in the right direction, it often falls short by treating all features equally, failing to capture their varying degrees of importance.

To address this limitation, we introduce a machine learning approach that uses a **self-attention mechanism** to intelligently combine multi-modal features for better prediction. Unlike conventional models, our method learns to prioritize the most relevant features and downplay those that contribute less to the prediction task. In this work, we focus on two primary types of data: **semantic features**, which are extracted from the text content of the posts, and **numeric features**, which include metadata and user interaction metrics.

We evaluate our model using the **ACM Multimedia SMPD 2020 Challenge dataset**, a well-known benchmark that reflects the real-world complexity of social media data. Through a series of experiments

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

and ablation studies, we demonstrate that our model delivers higher accuracy and stronger performance than existing methods, especially in cases involving complex feature interactions.

Our approach is grounded in machine learning, which plays a crucial role in modern data science. By learning directly from data, machine learning models can identify patterns and make predictions without being explicitly programmed. This adaptability makes them particularly effective for tasks like popularity prediction, where traditional rule-based systems often fall short.

In summary, this work presents a more nuanced and adaptive way to predict social media popularity, using self-attention to make better use of multi-modal data and improve prediction accuracy.

II LITERATURE SURVEY

Social Media Popularity Prediction (SMPP) has gained significant attention in recent years, particularly due to its wide applications in marketing, recommendation systems, and content optimization. Researchers have explored various approaches using multimodal data such as visual content, textual features, and user metadata.

In their 2014 study, Khosla, Das Sarma, and Hamid explore the factors that contribute to the popularity of images on social media platforms. They analyze a large dataset of images from Flickr and investigate how visual content and social context influence an image's popularity. The authors propose a predictive model that combines both low-level image features (like color, composition, and object presence) and high-level social features (like user statistics and upload time). Their findings show that while social context is important, visual features alone can also significantly predict image popularity. This research was one of the first to quantitatively analyze and predict the popularity of online images using machine learning techniques.

Bo Wu et al. (2016) emphasize that time plays a crucial role in predicting social media

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

popularity. Earlier works mainly focused on visual content or user interactions but failed to capture how popularity evolves over different time scales. The authors argue that existing models either use short-term trends or long-term accumulations, missing the multi-scale temporal patterns. They review previous approaches like regression-based models, point process models, and temporal dynamics, and highlight their limitations in adaptability and accuracy. Their study introduces the concept of multi-scale temporalization to fill this gap, showing that combining various time resolutions helps in better understanding and predicting content popularity.

In his 2017 paper, R. van Zwol investigates user behavior on the photo-sharing platform Flickr, focusing on who views the content and how users interact with it. The study analyzes patterns of photo access, user profiles, and metadata, aiming to understand how content gains visibility. Van Zwol highlights the importance of tags, user connections (contacts), and groups in influencing visibility. The research also explores the social dynamics of photo sharing, showing that users with more social connections or active participation in groups tend to receive more views. The findings provide insight into social search and recommendation systems by examining real-world behavior in online communities.

In their 2017 study, Hessel, Lee, and Mimno explore how different factors influence the popularity of social media content. The authors examine two major aspects: multi-modal content (like images and captions) and contextual factors (such as the timing of the post and the identity of the content creator). They reference previous research that either focused on content quality (e.g., visual and textual appeal) or contextual signals (e.g., time of posting, user influence) as separate predictors of online popularity. However, they argue that understanding popularity requires analyzing both content and context together, especially in competitive, time-sensitive environments like Reddit. Their work builds on and integrates findings from

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

studies in computer vision, natural language processing, and social computing.

Mazloom et al. (2016) review prior research on popularity prediction in social media, highlighting two main research streams: content-based approaches (analyzing visual and textual elements) and social/temporal factors (such as user influence and time of posting). They observe that while earlier studies have examined these aspects individually, there has been limited focus on integrating multimodal data (both images and text) specifically in brand-related posts. Their work is motivated by the gap in understanding how visual aesthetics, textual cues, and brand-related features together influence the popularity of social media content. They emphasize the importance of multimodal learning to improve prediction accuracy in a marketing context.

Hsu and colleagues (2019) review prior work that attempts to predict social media popularity using either visual content, textual content, or contextual metadata. They note that earlier studies often focus on single-modal features, such as image aesthetics or post timing, which limits prediction accuracy. Some recent efforts began combining visual and textual data, but these approaches still lacked deep integration of multi-modal information. The authors emphasize the need for multi-modal feature mining, where images, text, and metadata are analyzed together using deep learning techniques. Their study builds on this gap by proposing a comprehensive model that leverages convolutional and recurrent neural networks to improve popularity prediction performance.

Gelli and colleagues (2015) focus on predicting image popularity in social media by combining visual content, sentiment, and contextual features. The authors note that earlier work primarily emphasized visual aesthetics or low-level image features for popularity prediction. However, they point out that contextual factors—such as upload time, user information, and social signals—also play a crucial role. Additionally, they highlight the importance of sentiment,

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

showing that emotional impact can significantly affect engagement. Building on prior research in multimedia analysis and affective computing, the paper introduces a more holistic approach that blends sentiment and context with visual data for more accurate predictions.

Vaswani et al. (2017) introduce the Transformer model, which revolutionizes natural language processing (NLP) tasks by proposing a novel architecture that relies solely on self-attention mechanisms, eliminating the need for recurrent neural networks (RNNs) and convolutional neural networks (CNNs). The authors argue that traditional sequence-to-sequence models, which use RNNs for encoding and decoding, suffer from issues like slow training and limited parallelization. Previous work in NLP, such as LSTM and GRU-based models, focused on improving RNN-based architectures. However, these models still face challenges in capturing long-range dependencies efficiently. The Transformer model overcomes these limitations by using attention layers to process sequences in parallel, improving computational efficiency and performance on tasks like machine translation. The authors build on earlier research in neural networks and sequence modeling, but they show that a model built entirely on attention mechanisms can outperform previous state-of-the-art models, both in speed and accuracy.

He et al. (2016) address the challenge of training deep neural networks for image recognition. Prior work had shown that increasing the depth of neural networks could improve their performance, but also made the training more difficult due to the vanishing/exploding gradient problem. The authors reference prior methods such as unsupervised pre-training and skip connections, which had tried to mitigate this issue. However, these approaches still suffered from performance degradation as the depth of networks increased. He et al. propose the concept of residual learning, where instead of learning the desired output directly, the network learns the residual (or

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

difference) between the input and output. This helps mitigate the degradation problem and allows for the effective training of very deep networks (hundreds or even thousands of layers). Their work builds on previous advancements in deep learning and introduces a novel architecture, ResNet, which significantly improves image recognition performance in challenging benchmarks like ImageNet.

III EXISTING SYSTEM

Existing systems for social media popularity prediction typically combine multiple types of features, such as textual content and user engagement metrics, using multi-modal fusion techniques. However, these approaches often treat all features equally, which can reduce the model's effectiveness by failing to prioritize the most relevant information. Many rely on handcrafted features or shallow fusion methods that cannot capture complex interactions between different data types. While some deep learning models have been applied, they generally lack mechanisms to dynamically weigh features based on their importance. Consequently, existing methods often require extensive feature engineering and may not generalize well across datasets, highlighting the need for more adaptive models that can intelligently fuse and prioritize multi-modal features for improved prediction accuracy.

IV PROBLEM STATEMENT

The task is to predict the popularity score of social media posts using the available information such as content, user details, and other related data. Popularity score typically refers to metrics like the number of views or likes a post receives. Accurately estimating this score is difficult because it depends on multiple complex and diverse factors such as content quality, timing, user influence, and social context—that are often hard to measure or represent.

The goal of this research is to build a predictive model that can effectively estimate the popularity of a post by combining **semantic features** (like captions and tags) and **numerical features** (like timestamps, geolocation, and user

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

statistics), using an **attention-based network**. The model aims to overcome the challenges of multimodal data handling and to improve accuracy while maintaining manageable complexity

V PROPOSED SYSTEM

Predicting the popularity of social media posts is both a challenging and valuable task, especially given the growing role of social media in everyday communication and influence. As these platforms become key channels for users, businesses, and content creators, understanding what drives a post's popularity can significantly enhance engagement strategies. However, this prediction is complex due to the many factors involved—such as content quality, user influence, timing, and real-world events—that interact in unpredictable ways. Traditional methods often struggle to effectively capture these complexities. To address this, our project proposes a multi-modal approach that combines semantic (text-based) and numerical (data-based) features, enhanced with attention mechanisms. By transforming diverse data types like images and social graphs into a unified representation, the model can more effectively process and learn from varied information sources, leading to improved accuracy in popularity prediction.

VI IMPLEMENTATION

Data Collection & Preprocessing

- **Image Caption Extraction:** Extract and convert image content into descriptive captions (textual data) to capture visual semantics.
- **Semantic Data Collection:** Gather additional semantic information such as user-generated tags, post text content, and hashtags.
- **Numerical Data Collection:** Collect metadata including timestamps, geolocation information, follower counts, engagement metrics, and other relevant numerical attributes.
- **Data Formatting:** Convert all collected data into standardized,

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

machine-readable formats suitable for model training and evaluation.

Model Design

- **Dual-Branch Network Architecture:** Develop a neural network with two distinct branches to independently process semantic features (text, captions, tags) and numerical features (timestamps, follower counts).
- **Self-Attention Mechanism:** Incorporate self-attention layers in the semantic branch to model contextual relationships and dependencies within the text without relying on convolutional or recurrent networks.
- **Feature Attention Mechanism:** Implement specialized attention modules that focus on relevant textual features, enhancing the semantic representation's quality.

Feature Fusion and Prediction

Feature Integration: Merge semantic and numerical feature representations into a unified embedding space for comprehensive analysis.

Ensemble Learning: Combine predictions from both semantic and numerical branches using ensemble techniques to improve robustness and accuracy of the popularity score prediction.

Performance Evaluation

Benchmarking: Evaluate the proposed model against existing state-of-the-art methods using standard metrics such as accuracy, precision, recall, and F1-score.

Feature Impact Analysis: Conduct ablation studies and sensitivity analysis to understand the contribution of individual semantic and numerical features to overall model performance.

Scalability and Extendability

Modular Design: Ensure the architecture is modular and scalable to easily incorporate additional data modalities such as audio

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

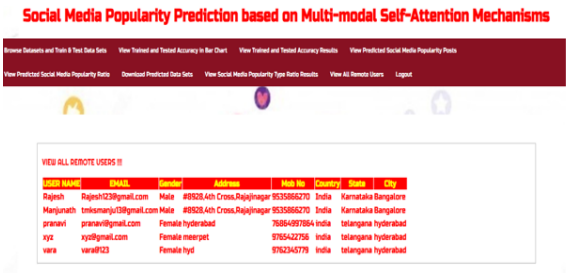
Cosmos Impact Factor-5.86

signals, video content, or user interaction behavior metrics in future iterations.

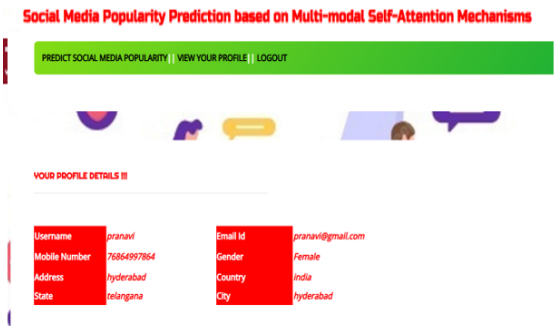
VII RESULTS

Home Page

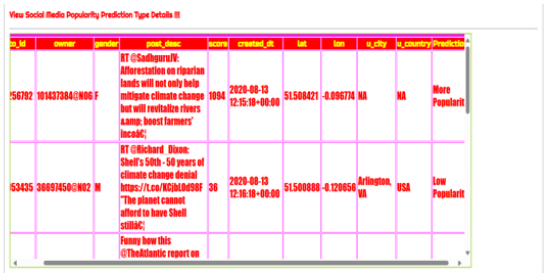
Admin Login



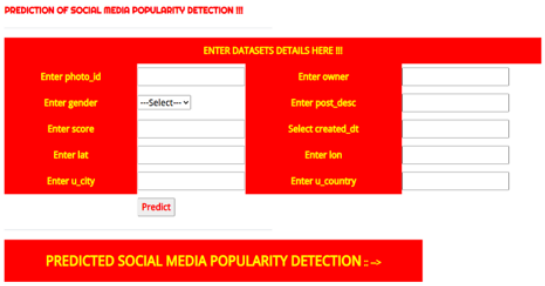
Remote Users



User Details



Datasets



Prediction Table

The view function processes a social media dataset to predict the popularity of posts

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

based on their descriptions. It reads data from Datasets1.csv, where each post is labelled as either "Low Popularity" (score ≤ 100) or "More Popularity" (score > 100). Using machine learning models like Naive Bayes, SVM, Logistic Regression, Decision Tree, and MLP Classifier, it trains and evaluates performance using accuracy, confusion matrix, and classification report. These models are then combined into a Voting Classifier for better prediction. When a user submits a post description, the system predicts its popularity and stores the result along with the input details in the database. The process aims to help users understand how likely their posts are to become popular based on content.

VIII CONCLUSION

we addressed the complex problem of predicting social media post popularity by leveraging both semantic (textual) and numerical features. Our method employs a self-attention mechanism to effectively fuse these heterogeneous data types, outperforming traditional models that treat features uniformly. Extensive experiments on the ACM Multimedia SMPD 2020 Challenge dataset validated the superior accuracy and efficiency of our approach. The integration of machine learning techniques provided deeper insights and enabled automated decision-making, highlighting the value of intelligent systems in practical applications such as advertising and trend analysis.

IX FUTURE WORK

To further improve the model, future work will focus on incorporating additional data modalities, such as visual content including images and videos, to enhance prediction performance. We also plan to explore advanced deep learning architectures, especially transformer-based models, for richer and more contextual feature representation. Extending the framework to support real-time popularity prediction and adapting it for different social media platforms will increase its versatility and applicability. Additionally, efforts will be made to improve the interpretability of the model's predictions and to reduce

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

computational costs, ensuring scalability and transparency.

REFERENCES

[1] Aditya Khosla, Atish Das Sarma, and Raffay Hamid, "What makes an image popular?," International Conference on World Wide Web., p.p.867–876. 2014.

[2] Bo Wu, Wen-Huang Cheng, Yongdong Zhang, and Tao Mei, "Time matters: Multi-scale temporalization of social media popularity," ACM International Conference on Multimedia., p.p. 1336–1344. 2016.

[3] R. van Zwol, "Flickr: Who is Looking?," IEEE/WIC/ACM International Conference on Web Intelligence., p.p. 184-190. 2017.

[4] Jack Hessel, Lillian Lee, and David Mimno, "Cats and captions vs. creators and the clock: Comparing multi-modal content to context in predicting relative popularity," International Conference on World Wide Web., p.p. 927–936. 2017.

[5] Masoud Mazloom, Robert Rietveld, Stevan Rudinac, MarcelWorring, and Willemijn Van Dolen, "Multimodal Popularity Prediction of Brand-related Social Media Posts," ACM International Conference on Multimedia., p.p. 179-201. 2016.

[6] SMP Challenge Organization. 2020. Social Media Prediction Challenge.

Available: http: //smp-challenge.com

[7] Chih-Chung Hsu, Li-Wei Kang, Chia-Yen Lee, Jun-Yi Lee, Zhong-Xuan Zhang, and Shao-Min Wu, "Popularity Prediction of Social Media based on Multi-Modal Feature Mining," ACM International Conference on Multimedia., p.p. 2687–2691. 2019.

[8] Francesco Gelli, Tiberio Uricchio, Marco Bertini, Alberto Del Bimbo, and Shih-Fu Chang, "Image popularity prediction in social media using sentiment and context features," ACM International Conference on Multimedia., p.p. 907–910. 2015.

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

Polosukhin, "Attention is all you need," International Conference on Neural Information Processing Systems., p.p. 6000–6010. 2017.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," IEEE Conference on Computer Vision and Pattern Recognition., p.p. 770–778. 2016.

[11] Y. Liu and Myle Ott and Naman Goyal and Jingfei Du and Mandar Joshi and Danqi Chen and Omer Levy and M. Lewis and Luke Zettlemoyer and Veselin Stoyanov, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arxiv.org/abs/1907.11692., 2019.

[12] Douwe Kiela, Suvrat Bhooshan, Hamed Firooz and Davide Testuggine, "Supervised multi-modal Bitransformers for Classifying Images and Text," arXiv preprint arxiv.org/abs/1909.02950., 2019.

[13] Bo Wu, Wen-Huang Cheng, Yongdong Zhang, Qiushi Huang, Jintao Li, and Tao Mei, "Sequential Prediction of Social Media Popularity with Deep Temporal Context Networks," International Joint Conference on Artificial Intelligence., p.p. 3062–3068. 2017.

[14] Zehang Lin and Feitao Huang and Yukun Li and Zhenguo Yang and W. Liu, "A layer-wise deep stacking model for social image popularity prediction," International Conference on World Wide Web., p.p. 1639-1655. 2018.

[15] Keyan Ding, Ronggang Wang, and Shiqi Wang, "Social Media Popularity Prediction: A Multiple Feature Fusion Approach with Deep Neural Networks," ACM International Conference on Multimedia., p.p. 2682–2686. 2019.

[16] Ziliang He, Zijian He, Jiahong Wu, and Zhenguo Yang, "Feature Construction for Posts and Users Combined with LightGBM for Social Media Popularity Prediction," ACM International Conference on Multimedia., p.p. 2672–2676. 2019.

[17] Junhong Chen, Dayong Liang, Zhanmo Zhu, Xiaojing Zhou, Zihan Ye, and Xiuyun

Mo, "Social Media Popularity Prediction Based on Visual-Textual Features with XGBoost," ACM International Conference on Multimedia., p.p. 2692–2696. 2019.

[18] Peipei Kang, Zehang Lin, Shaohua Teng, Guipeng Zhang, Lingni Guo, and Wei Zhang, "Catboost-based Framework with Additional User Information for Social Media Popularity Prediction," ACM International Conference on Multimedia., p.p. 2677–2681. 2019.

[19] Liuwu Li, Sihong Huang, Ziliang He, and Wenyin Liu, "An Effective Textbased Characterization Combined with Numerical Features for Social Media Headline Prediction," ACM International Conference on Multimedia., p.p. 2003–2007. 2018.

[20] E Ethem F. Can, Hüseyin Oktay, and R. Manmatha, "Predicting retweet count using visual cues," ACM international conference on Information & Knowledge Management., p.p. 1481–1484. 2013.

[21] Chih-Chung Hsu, Chia-Yen Lee, Ting-Xuan Liao, Jun-Yi Lee, Tsai-Yne Hou, Ying-Chu Kuo, Jing-Wen Lin, Ching-Yi Hsueh, Zhong-Xuan Zhang, and Hsiang-Chin Chien, "An iterative refinement approach for social media headline prediction," ACM International Conference on Multimedia., p.p. 2008–2012. 2018.

[22] Chih-Chung Hsu, Ying-Chin Lee, Ping-En Lu, Shian-Shin Lu, Hsiao-Ting Lai, Chihg-Chu Huang, Chun Wang, Yang-Jiun Lin, and Weng-Tai Su, "Social media prediction based on residual learning and random forest," ACM International Conference on Multimedia., p.p. 1865–1870. 2017.