ACM Multimedia 2026
SMP Challenge (Social Media Prediction Challenge)

Overview - http://smp-challenge.com (will open if accepted)

In recent years, social multimedia have become a dominant content format across various sectors, witnessing exponential growth. Platforms such as Flickr, Instagram, Youtube and Tiktok have introduced their new styles of multimedia content sharing—while they have emerged as global leaders in the Internet. The rapid rise of social multimedia is fueled by their convenience, accessibility, and ease of creation, making them integral to entertainment, marketing, advertising, news production, and education.

As social media continue to shape digital engagement, the ability to predict their trending posts is increasingly important. Accurate prediction models can enable content creators to optimize content production, improve audience engagement, and refine marketing strategies. However, large-scale social media popularity prediction remains a complex challenge due to the change of rapidly dynamic interests and the evolving nature of research methodologies in this field.

Social Media Prediction Challenge (SMP Challenge) is seeking excellent research teams or solutions for social multimodal prediction. Our goal focused on social multimodal popularity prediction. We have curated a dataset of online images with meta-information and engagement metrics. This challenge invites participants to develop innovative predictive models that can assess popularity based on user engagement metrics such as views, likes, comments, and shares. With the ever-changing public attention and individual interests, predicting the exact values accurately of online popularity is even more important than before. The challenges for researchers looking at their models on social media data are large-scale, multimodal, and multivariate.

We released the **Social Media Prediction Dataset (SMPD)** for sociological understanding and forecasting, consisting of over 680K images and 80K users. Beyond popularity scores, the dataset includes diverse social media images, post texts, and attributes such as user profiles, timestamps, locations, and categories, requiring participating teams to develop innovative algorithms to improve prediction accuracy and correlation evaluations over streaming data. Research in social media prediction spans multimedia, artificial intelligence, computer vision, and natural language processing.

The challenge is an annual event that has been held for 7 years of efforts with support from multiple research organizations and companies. It has attracted over 280 global research teams. The goal is to create a comprehensive platform for and next 3-5 years, enabling robust research and advancements in social multimedia prediction.

Task Description

Social Media Popularity Prediction. The overconsumption of online information has its limitations, so online word-of-mouth helps us to efficiently discover emerging topics, interesting news, or new products from the information ocean. Therefore, predicting online popularity became a crucial task for online media, brand owners, social influencers, and individuals.

The task focuses on predicting the impact of sharing different posts for a publisher on social media. Given a multimodal post from a publisher, the goal is to automatically predict the future popularity after the post is publicly shared.

The contestants are asked to develop their dynamic prediction models based on the dataset provided by the SMP Challenge (as training data), plus possibly additional public/private data, to address the given tasks. For evaluation, a contesting system is asked to produce prediction results of popularity. The accuracy will be evaluated by pre-defined quantitative evaluation. The contestants need to introduce their systems and datasets at the conference.

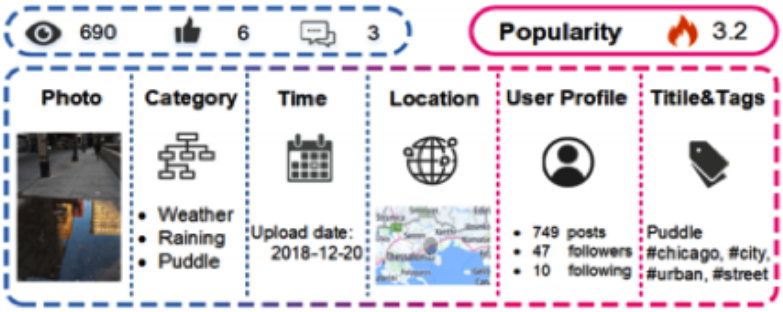

Figure 1: The illustration of social media popularity prediction

Recent Methods

We provide social images for social media prediction, incorporating the dynamics of social multimedia data. A core challenge is learning from data sequences and temporal sequences, which are essential for accurate predictions. While classic multimedia prediction models rely on regression, stochastic processes, convolutional networks, or transformers for multimodal modeling, temporal social media data presents additional challenges, including multi-modal content, discrete sequences, variable-length patterns, and multi-scale variance. We split training/testing data into chronological sets (commonly, by date and time). Addressing these complexities, temporal learning for social multimedia data remains an open problem, requiring further research and innovation.

Dataset

SMPD. SMPD-Image contains 486K social image posts from 70K users, which have anonymized photo-sharing records, user profiles, web images, text, time, location, category, etc. SMPD is a multi-faced, large-scale, temporal web data collection, collected from Flickr.

Dataset	#Post	#User	#Categories	Duration (M)	# Tags
SMPD-Image	486k	70k	756	16	250k

Evaluation

The evaluation will be based on the online evaluation server and platform supported by the Huggingface community. We released the training dataset and withhold the testing data set for evaluation.

SMP-Image: We will measure the received solutions by the prediction correlation metric Spearman's Rho (SR) and the error metric Mean Absolute Error (MAE).

The ranking for the competition is based on quantitative evaluation. The final rank of a team is measured by combining objective evaluation metrics:

$$R(team) = R(team)@SR + R(team)@MAE \text{ (smaller is better)}$$

where R(team) is the rank position of the team on a particular metric, e.g., if the team achieves the best performance in terms of Spearman's Rho, then R(team)@SR is "1".

Timeline

- March 10, 2026: Website Open (Dataset: no later March 15)
- April 20 - May 28, 2025: Result Submission
- May 28 - June 5, 2025: Evaluation and Review
- June 5, 2025: Leaderboard Notification
- July 5, 2025: Solution Paper Submission
- July 16, 2025: Author Notification
- Aug 13, 2025: Camera Ready

Commitment

We will run and maintain the website for the challenge containing the information, tasks, datasets, and the leaderboard for at least 3 years. The team has enough experience in continuously supporting global challenges for the next 3-5 years. We also planned how to promote the challenge to researchers for participation through multiple social media platforms.

Contacts

Bo Wu MIT-IBM Watson AI Lab, bobbywu.cs@gmail.com

Peiye Liu ByteDance

Qiushi Huang University of Surrey

Jia Wang National Yang Ming Chiao Tung University

Bei Liu Microsoft Research Asia

Zhaoyang Zeng International Digital Economy Academy

Jiebo Luo University of Rochester

Wen-Huang Cheng National Taiwan University