The Influence of User Profile and Post Metadata on the Popularity of Image-based Social Media: A Data Perspective

Shintami Chusnul Hidayati
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
shintami@its.ac.id

Muhammad Rizqi Fiqih Thalib
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
05111940000219@student.its.ac.id

Abdul Munif
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
munif@if.its.ac.id

***Abstract*—In modern internet communication, social media platforms have emerged as indispensable tools for self-expression, social engagement, and content distribution. Nevertheless, existing research on predicting social media popularity tends to overlook the complex interactions between different characteristics of user profiles, such as geolocation and account type, and post metadata, such as postdate and hashtags. This paper addresses this gap by investigating the factors derived from user profiles and post metadata contributing to content popularity within social media environments, with a specific emphasis on image-sharing. Utilizing a comprehensive dataset, we employ artificial intelligence-based data analytics techniques to uncover patterns, correlations, and predictive models, thereby providing a detailed insight into the dynamics of popularity within these online spaces.**

***Index Terms*—post popularity prediction, social media, regression task, community**

I. Introduction

In the digital information-sharing landscape, social media platforms serve as pervasive channels for self-expression, interpersonal connection, and global content dissemination [1]. The exponential growth in content on these platforms transforms them into rich data sources [2], [3], creating an opportunity to investigate the determinants of post popularity. Understanding the dynamics governing the success of such content is crucial for content creators, social media platforms, and marketers. Engagement metrics like likes, comments, shares, and view count are vital indicators of content resonance and impact, making predicting and enhancing engagement a central concern in the digital age.

The study of predicting post popularity has received significant attention in social media analytics, focusing on various content types such as text [4]–[7], videos [8]–[10], and images [11]–[13]. While progress has been made in understanding the dynamics of post popularity, the nuances associated with image-based content remain a challenging area that requires further investigation, especially with the pervasive growth of platforms like Flickr, Instagram, and Snapchat, where visual communication is paramount.

Pioneering work by Khosla *et al.* [14] explored visual content factors, such as color, gradients, deep learning features, and objects present, as determinants of image popularity. Works like [11] and [15] demonstrated the efficacy of incorporating visual aesthetics to enhance prediction accuracy. Massip *et al.* [12] extended the investigation by incorporating category-specific information, offering insights into how visual content across specific categories influences user engagement. Simultaneously, studies such as [16] and [17] delved into caption information.

Understanding the role of user profile information, defining user-related features (e.g., geolocation and account type), in image popularity prediction has gained prominence. Users with a substantial following are intuitively more likely to have their posts seen by a larger audience, potentially increasing visibility and popularity. Metadata, including post date and hashtag information, plays a crucial role in image understanding, providing context and details about the content. Previous studies have incorporated metadata with user profile information into predictive models [18], [19], yet a gap exists in systematically understanding each component.

This paper addresses the identified gap through a data-driven approach, investigating the relationship between user profile attributes and metadata associated with image-based posts. By exploring different strategies, ranging from feature engineering to model learning, our objective is to uncover the factors influencing the success of visual content in capturing and resonating with an online audience. Specifically, we explore various methods to extract and transform variables related to users and metadata from raw data. Subsequently, we utilize these transformed variables as input for different learning algorithms to not only identify the best predictive models but also comprehend the impact of each component and their interrelationships. Our findings provide valuable insights to guidance social media platforms in refining algorithms and interface design, thereby enhancing user experiences and fostering meaningful engagement.

The subsequent sections are organized as follows: Section II outlines the dataset used, detailing the collection process and key features. Section III presents the methodology, describing selected machine learning algorithms and their motivation.

TABLE I: A summary of feature attributes (sorted by the name ascending).

Feature attribute	Description
alltags	Number of tags from users
canbuypro	Users can buy pro members or not
category	First category of posts with 11 different classes
concept	Post concept with 502 different classes
geoaccuracy	The level of accuracy of location information
ispro	Pro member users or not
latitude	Latitude that ranges from -90 to 90
longitude	Longitude that ranges from -180 to 180
photo-count	Number of photos uploaded by the user
postdate	Time when the post was published
subcategory	Subcategory of posts with 75 different classes
timezone-id	User's time zone ID
timezone-offset	User's time zone

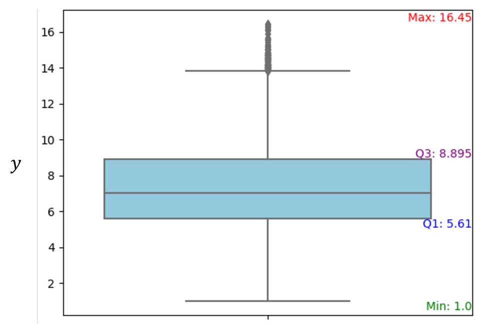


Fig. 1: Distribution of normalized engagements in the collected data.

Section IV discusses experimental results and implications, providing a comprehensive analysis. Finally, Section V concludes the paper by summarizing key contributions, discussing limitations, and suggesting directions for future research in this evolving domain.

II. Dataset

We collected 11,771 posts from 2,926 user accounts on the image-sharing social media platform Flickr[1], where user engagement, representing popularity, is measured by the number of views. Because we focus on user profiles and metadata information, 13 feature attributes were selected as predictor variables, as summarized in Table I. Since the number of views on each post has varying values, log normalization is utilized, which is defined as:

$$y = log_2 \frac{\hat{y}}{d} + 1, \tag{1}$$

where y is the normalized engagement value, $\hat{y}$ is the view count of a photo considered as engagement, and d is the number of days since the image was posted. The distribution of normalized engagement data is presented in Fig. 1.

III. Methodology

A. Feature Engineering

Feature engineering is essential for deriving meaningful insights from raw data and optimizing machine learning model performance by selecting optimal features. Mainly, the diversity in data types and model assumptions necessitates tailored approaches.

To handle categorical data, we employ two methods: One-Hot Encoding, transforming categorical variables into binary vectors (0s and 1s), and Label Encoding, assigning a unique integer to each category. Encoded variables are day, month, and time extracted from the *postdate* attribute and attributes such as *category*, *subcategory*, *concept*, *ispro*, *canbuypro*, and *timezone-id*. Meanwhile, *longitude* and *latitude* in photo location data are processed using their values directly as features or by clustering. K-means clustering is utilized in this study and the optimal number of clusters (k) is determined through the elbow method.

To address the high dimensionality resulting from One-Hot Encoding, we explore the application of PCA (Principal Component Analysis) [20]. PCA is a popular method that effectively reduces data dimensionality while preserving most of its variability [21]. Specifically, the number of feature dimensions obtained from the results of one-hot encoding is 614, whereas, after reduction by PCA, it becomes 95.

B. Training Process

Four regression-based methods are utilized for this task: Support Vector Regression [22], Random Forest Regressor [23], Category Boosting [24], and Artificial Neural Network (ANN) [25]. Each method is optimized to ensure a fair comparison.

1) Support Vector Regression: Support Vector Regression (SVR) is a supervised learning algorithm used for regression tasks, aiming to find a hyperplane that best represents the relationship between the input features and the target variable while allowing for a certain tolerance in prediction errors. Unlike traditional regression models that minimize errors across all data points, SVR focuses on fitting the data within a specified margin, as determined by support vectors.

Let $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), ..., (\mathbf{x}_N, y_N)\}$ denote the training data, where $\mathbf{x}_i$ is the input vector and y_i is the corresponding output. The goal is to find a function $f(\mathbf{x})$ satisfying $|f(\mathbf{x}_i) - y_i| \leq \epsilon$ for all training examples while maximizing the margin, where ϵ represents the acceptable error. The SVR formulation involves optimizing a cost function defined as:

$$\min_{\mathbf{w},b,\xi,\xi^*}(\frac{1}{2}\|\mathbf{w}\|^2 + C\Sigma_{i=1}^{N}(\xi_i + \xi_i^*)), \tag{2}$$

subject to the constraints:

$$y_i - \mathbf{w} \cdot \phi(\mathbf{x}_i) - b \leq \epsilon + \xi_i, \tag{3}$$

$$\mathbf{w} \cdot \phi(\mathbf{x}_i) + b - y_i \leq \epsilon + \xi_i^*, \text{and} \tag{4}$$

$$\xi_i, \xi_i^* \geq 0 \quad \text{for } i = 1, \ldots, N. \tag{5}$$

Here, w represents the weight vector, b is the bias term, $\phi(\cdot)$ denotes the mapping function to a higher-dimensional space, and ξ_i and ξ_i^* are slack variables.

[1]https://www.flickr.com/

TABLE II: Hyperparameter settings used for model training.

Method		Scenario				
		Enc	Enc-Clust	OHE	OHE-PCA	Cat
SVR	kernel	rbf	rbf	rbf	rbf	-
	C	0.1	1	10	10	-
	ϵ	0.1	0.1	1	0.2	-
	γ	0.1	1	0.01	0.1	-
RFR	n estimators	100	300	150	50	-
	max depth	20	30	50	40	-
	min samples leaf	1	1	1	1	-
	min samples split	2	2	2	7	-
	max features	1	None	1	1	-
CatBoost	learning rate	0.05	0.05	0.1	0.05	0.1
	depth	10	10	8	10	10
	iterations	500	500	700	500	600
	L2 leaf reg	5	1	1	0.5	5
MLP	solver	adam	sgd	adam	adam	-
	hidden layer sizes	(30, 20, 10, 5)	(30, 20, 10)	(30, 20, 10)	(20, 10)	-
	activation function	tanh	tanh	tanh	tanh	-
	learning rate	constant	constant	constant	constant	-
	α	0.1	0.01	1	5	-

2) Random Forest Regressor: Random Forest (RF) is an ensemble learning method within the family of decision tree-based models. The key idea behind RF is to build decision trees during training and aggregate the average prediction (for regression tasks) or the mode prediction (for classification tasks) from the individual trees. Let T represent the number of trees in the forest, and $f_t(\mathbf{x})$ denote the prediction made by the *t*-th tree given input features **x**. In our task, which is a regression problem, the overall prediction $f(\mathbf{x})$ is the average of predictions from all individual trees:

$$f(\mathbf{x}) = \frac{1}{2}\Sigma_t^T f_t(\mathbf{x}). \tag{6}$$

3) Category Boosting: Category Boosting (CatBoost) is a machine learning algorithm designed explicitly for gradient boosting on decision trees. The general idea is to build an ensemble of decision trees that minimizes the mean squared error while efficiently handling categorical features. The overall formulation involves finding a set of weak learners f_t such that the following objective function is minimized:

$$f(\mathbf{x}) = \frac{1}{N}\Sigma_{i=1}^N (y_i - \Sigma_{t=1}^T f_t(\mathbf{x}_i))^2 + \Sigma_{t=1}^T \Omega(f_t), \tag{7}$$

where $\Omega(f_t)$ is a regularization term that penalizes complex trees to avoid overfitting. The training process involves iteratively adding trees to the model, each time fitting the negative gradient of the loss function concerning the current model predictions. The process continues until a predefined number of trees are added or until convergence is achieved.

4) Multi-layer Perceptron: A Multi-layer Perceptron (MLP) is an artificial neural network that can be used for classification and regression tasks. The architecture of an MLP consists of an input layer, one or more hidden layers, and an output layer. Each layer contains multiple nodes or neurons, and connections between these nodes have associated weights.

Consider an MLP with L layers, denoting the weights and biases of layer l as W^l and b^l, respectively. The output of layer l after applying the activation function σ^l is represented as a^l. The forward pass to predict the output for an input vector **x**, denoted as $f(\mathbf{x})$, can be expressed as:

$$f(\mathbf{x}) = \sigma^l(W^l a^{l-1} + b^l). \tag{8}$$

During training, the weights and biases are updated through backpropagation and gradient descent to minimize the loss.

C. Predicting Popularity Score

The primary objective of this step is to predict the popularity score of the input data. Utilizing the optimized trained model described in Section III-B, the prediction is performed on the input data using features extracted through the approach detailed in Section III-A.

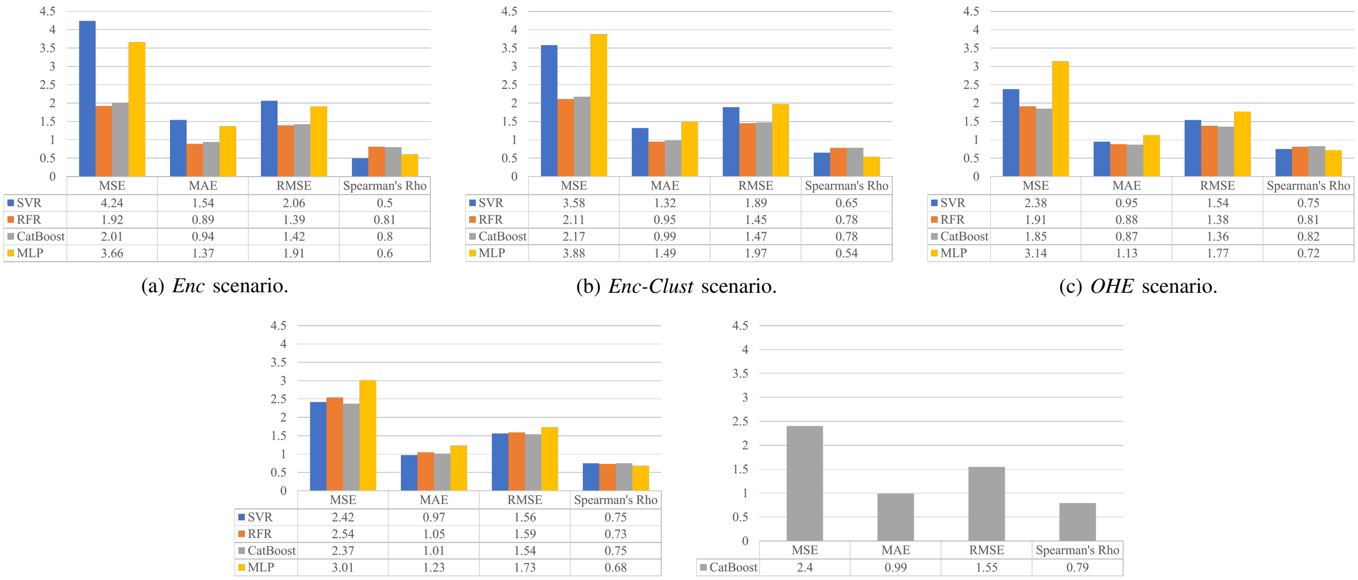
IV. Experimental Results and Discussion

In this section, we begin by outlining the experimental settings. Subsequently, we introduce the chosen evaluation metrics used to assess the performance of the experiments. Following this, we present a comprehensive explanation of the scenarios employed in the study and detailed corresponding results obtained from each scenario.

A. Settings

We conducted our experiments on a personal computer with an Intel i7-9750 CPU, 16GB RAM, and an NVIDIA GT420 GPU. The dataset was split into two sets: 70% for training and the remaining portion for testing. Exploration of the hyperparameter space was achieved using grid search. The hyperparameters for each method are presented in Table II.

Specifically, in Support Vector Regression (SVR), the hyperparameters include the kernel function, C, ϵ, and γ. Here, C is the regularization parameter controlling the trade-off between a smooth decision boundary and fitting the training data, ϵ represents the margin of tolerance for errors, and γ defines how far the influence of a single training example reaches. For the Random Forest Regression (RFR), the optimized hyperparameters include the number of trees in the forest (n estimators), the maximum depth of the tree, the minimum



(a) *Enc* scenario. (b) *Enc-Clust* scenario. (c) *OHE* scenario.

(d) *OHE-PCA* scenario. (e) *Cat* scenario.

Fig. 2: Model performance on different experimental scenarios.

number of samples required to be at a leaf node, the minimum number of samples required to split an internal node, and the number of features to consider when looking for the best split. Meanwhile, hyperparameters for the CatBoost regressor encompass the depth of the trees, the number of boosting iterations (trees) to be run, and the L2 regularization term on weights that helps prevent overfitting. For the Multilayer Perceptron (MLP) regressor, the hyperparameters consist of the optimization algorithm (solver), the number of neurons in each hidden layer (hidden layer sizes), the activation function for the hidden layers, the step size in updating the weights during training (learning rate), and the L2 penalty (regularization term) to prevent overfitting (α).

B. Evaluation Metrics

To measure the accuracy and generalization capabilities of the models, key evaluation metrics for regression tasks are employed: (1) Mean Absolute Error (MAE), which represents the average absolute difference between the predicted and actual values; (2) Mean Squared Error (MSE) that measures the average squared difference between predicted and actual values; (3) Root Mean Squared Error (RMSE) to provide a measure of the average magnitude of errors in the original units of the target variable; and (4) Spearman's rank correlation coefficient, often denoted as ρ (rho) used to evaluate how well the ranking of predictions aligns with the ranking of actual values.

C. Experimental Scenarios

Five scenarios have been designed to evaluate the predictive efficacy of the various feature processing techniques outlined in Section III-A, utilizing the regression-based methods explained in Section III-B.

1) Enc: This experiment converts categorical data into numeric values using the Label Encoding strategy, while non-categorical data retains its original values. The encoded attributes are then concatenated with non-categorical attributes and used as input for the regression model.

2) Enc-Clust: In this experiment, the *latitude* and *longitude* attributes are clustered so that posts with the same area due to the proximity of *latitude* and *longitude* are in the same cluster category. Categorical data, including the geolocation clusters, are converted to numeric via Label Encoding, while other data uses the original values. The encoded attributes are then combined with non-categorical attributes using concatenation and used as input for the regression model.

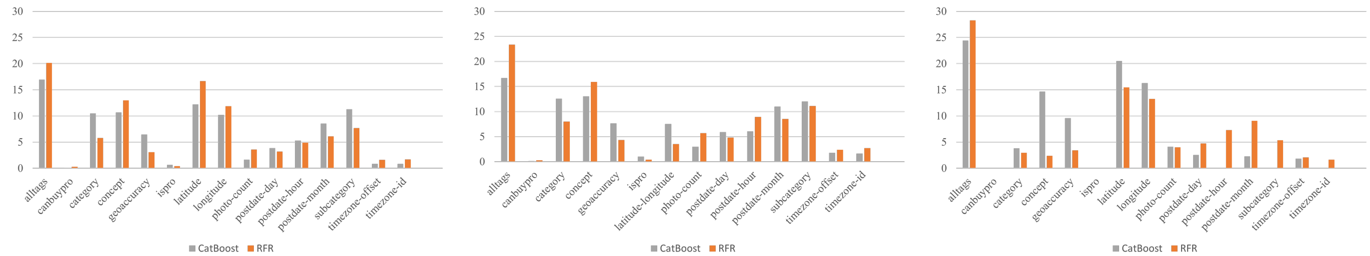
3) OHE: In this experimental scenario, one-Hot Encoding is applied to convert categorical attributes into binary vectors. Subsequently, all attributes, including the original numeric values of *latitude* and *longitude*, are concatenated and used as input for the regression model.

4) OHE-PCA: After obtaining the attribute feature vector from Section IV-C3, its dimensions are reduced using PCA before inputting the regressor.

5) Cat: In this scenario, CatBoost is employed as the regression model. All attributes are used as input without applying Label Encoding or One-Hot Encoding to convert categorical data.

D. Results and Discussion

Fig. 2 summarizes the performance of each scenario across different learning methods. It is evident that both the learning method and testing scenario significantly influence performance. In other words, no learning method consistently outperforms others in every experimental scenario. The least favorable result was observed in the *Enc* scenario with SVR as



(a) *Enc* scenario. (b) *Enc-Clust* scenario. (c) *OHE* scenario.

Fig. 3: Feature importance of each feature (in percentage).

the regressor. Conversely, the best performance was achieved by CatBoost in the *OHE* scenario. The differences between them, based on MSE, MAE, RMSE, and Spearman's Rho, are 2.39 (56.37%), 0.67 (43.51%), 0.70 (33.90%), and 0.32 (39.02%), respectively. Notably, CatBoost consistently ranks within the top two performers in every scenario.

While CatBoost exhibits superior overall performance in the *OHE* scenario, its performance in the *Enc* and *Enc-Clust* scenarios remains inferior to RFR across all evaluation matrices. This underscores that CatBoost's effectiveness, in this case, could be better when employing Label Encoding to convert categorical attributes. Notably, using CatBoost with categorical attributes directly as input (without conversion using Label Encoding or One-Hot Encoding) yields the lowest performance compared to employing Label Encoding or One-Hot Encoding before utilizing categorical attributes as input for CatBoost.

Geolocation clustering, encompassing *longitude* and *latitude* attributes (cf. Fig. 2(b)), demonstrates a performance improvement solely in SVR. Specifically, comparing SVR performance with geolocation attribute clustering (*Enc-Clust* scenario) against no geolocation attribute clustering (*Enc* scenario), the performance increases by 0.66 (15.57%), 0.22 (14.29%), 0.17 (8.25%), and 0.15 (23.08%) on the MSE, MAE, RMSE, and Spearman's Rho evaluation matrices, respectively. When comparing SVR performance against existing scenarios, superior results were obtained when applied to features categorized via One-Hot Encoding (*OHE* scenario).

The utilization of PCA has a substantial impact on MLP performance, with this method yielding the best performance compared to other scenarios. The poorest MLP performance was observed in the *Enc-Clust* scenario, exhibiting differences of 0.87 (22.42%) for MSE, 0.26 (17.45%) for MAE, 0.24 (12.18%) for RMSE, and 0.48 (20.59%) for Spearman's Rho compared to the best MLP performance. When compared to employing One-Hot Encoding solely on categorical attributes, integrating PCA into the framework enhances performance by 0.13 (4.14%) for MSE, 0.04 (2.26%) for RMSE, and 0.18 (25.00%) for Spearman's Rho, while MAE decreases by 0.10 (8.13%). These results indicate that transforming original features into a new set of unrelated features (principal components) can be advantageous for MLP, notably reducing the obtained errors. Further exploration into training MLP with autoencoders, such as in [26], could be a promising avenue for future research.

In contrast to MLP, which demonstrated improved performance with the application of PCA, RFR experienced a decline in performance when PCA was applied. In the *OHE* scenario, the MSE, MAE, RMSE, and Spearman's Rho for RFR were 1.91, 0.88, 1.38, and 0.81, respectively. However, these metrics respectively decreased by 0.63 (24.80%), 0.17 (16.19%), 0.21 (13.21%), and 0.08 (9.87%) with the implementation of PCA. This result suggests that the One-Hot Encoding technique is more effective in capturing data patterns for RFR.

To comprehend the impact of each component in user profiles and post metadata on post popularity, we assess their contribution to the model's predictions. This is achieved by calculating the percentage of the mean decrease in impurity (or Gini importance) to the total, gauging each feature's effectiveness in reducing uncertainty. Fig. 3 illustrates the feature importance for the two best models in *Enc*, *Enc-Clust*, and *OHE* scenarios, which are CatBoost and RFR. We do not analyze the *OHE-PCA* scenario because PCA has already transformed the original features into a new set of uncorrelated features.

Based on Fig. 3, it is evident that tags with hashmarks (also known as hashtags) play the most crucial role in determining the popularity of posts. Intuitively, hashtags make posts searchable and facilitate connections with users more likely to engage with and appreciate the content. The strategic use of popular or trending hashtags emerges as a critical strategy for expanding post reach to a broader audience. Additionally, the consistently top-ranking components across various scenarios include the content in the image, represented by concepts, geolocation, and postdate. Understanding and leveraging these insights can empower content creators and strategists to optimize their approach, enhancing the overall visibility and impact of image-based content.

V. Conclusion

This study investigates the complex dynamics influencing the popularity of image-based social media content, emphasizing the interaction between user profiles and post metadata. The findings underscore the crucial role of feature engineering strategies in determining content popularity. Adopting a data-driven methodology contributes to a deeper understanding of the multifaceted factors influencing the success of visual content on social media platforms. Highlighted by experimental results showcasing the crucial role of hashtags in shaping post popularity, it provides valuable insights for content creators and strategists aiming to optimize their approach. Future research could extend the analysis to explore more intricate components of social media posts and incorporate additional ranking criteria.

ACKNOWLEDGMENT

This work is partially supported by Institut Teknologi Sepuluh Nopember under Grant No. 1409/PKS/ITS/2023.

REFERENCES

[1] T.-H. Tsai, W.-C. Jhou, W.-H. Cheng, M.-C. Hu, I.-C. Shen, T. Lim, K.-L. Hua, A. Ghoneim, A. Hossain, and S. C. Hidayati, "Photo sundial: Estimating the time of capture in consumer photos," *Neurocomputing*, vol. 177, pp. 529–542, 2016.

[2] S. C. Hidayati, K.-L. Hua, Y. Tsao, H.-H. Shuai, J. Liu, and W.-H. Cheng, "Garment detectives: Discovering clothes and its genre in consumer photos," in *Proc. IEEE Conf. Multimedia Information Processing and Retrieval*, 2019, pp. 471–474.

[3] S. C. Hidayati, T. W. Goh, J.-S. G. Chan, C.-C. Hsu, J. See, L.-K. Wong, K.-L. Hua, Y. Tsao, and W.-H. Cheng, "Dress with style: Learning style from joint deep embedding of clothing styles and body shapes," *IEEE Trans. Multimedia*, vol. 23, pp. 365–377, 2021.

[4] Z. D. Meilani, I. M. Nizar, M. F. Sunandar, and S. C. Hidayati, "Lawyering social: Navigating legal issues on social media posts with a low-cost data algorithm," in *Proc. Int. Conf. Informatics and Computational Sciences*, 2021, pp. 266–271.

[5] X. Gao, Z. Zheng, Q. Chu, S. Tang, G. Chen, and Q. Deng, "Popularity prediction for single tweet based on heterogeneous bass model," *IEEE Trans. Knowledge and Data Engineering*, vol. 33, no. 05, pp. 2165–2178, 2021.

[6] I. N. Lymperopoulos, "RC-Tweet: Modeling and predicting the popularity of tweets through the dynamics of a capacitor," *Expert Systems with Applications*, vol. 163, p. 113785, 2021.

[7] T. Elmas, S. Stephane, and C. Houssiaux, "Measuring and detecting virality on social media: The case of twitter's viral tweets topic," in *Companion Proc. ACM Web Conf.*, 2023, pp. 314–317.

[8] Y.-H. Hsieh, S. C. Hidayati, W.-H. Cheng, M.-C. Hu, and K.-L. Hua, "Who's the best charades player? mining iconic movement of semantic concepts," in *Proc. Int. Conf. MultiMedia Modeling*, 2014, pp. 231–241.

[9] W. Hoiles, A. Aprem, and V. Krishnamurthy, "Engagement and popularity dynamics of youtube videos and sensitivity to meta-data," *IEEE Trans. Knowledge and Data Engineering*, vol. 29, no. 7, pp. 1426–1437, 2017.

[10] J. Wang, Y. Wang, N. Weng, T. Chai, A. Li, F. Zhang, and S. Yu, "Will you ever become popular? learning to predict virality of dance clips," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 18, no. 2, 2022.

[11] S. C. Hidayati, Y.-L. Chen, C.-L. Yang, and K.-L. Hua, "Popularity meter: An influence- and aesthetics-aware social media popularity predictor," in *Proc. ACM Int. Conf. Multimedia*, 2017, pp. 1918–1923.

[12] E. Massip, S. C. Hidayati, W.-H. Cheng, and K.-L. Hua, "Exploiting category-specific information for image popularity prediction in social media," in *Proc. IEEE Int. Conf. Multimedia & Expo Workshops*, 2018, pp. 45–46.

[13] J. Wang, S. Yang, H. Zhao, and Y. Yang, "Social media popularity prediction with multimodal hierarchical fusion model," *Computer Speech & Language*, vol. 80, p. 101490, 2023.

[14] A. Khosla, A. Das Sarma, and R. Hamid, "What makes an image popular?" in *Proc. Int. Conf. World Wide Web*, 2014, pp. 867–876.

[15] P. Mudgal and F. Liu, "Are high-quality photos more popular than low-quality ones? a quantitative study," in *Proc. IEEE Int. Wkshp. Multimedia Signal Processing*, 2022, pp. 1–5.

[16] S. C. Hidayati, R. B. R. Prayogo, S. A. V. Karuniawan, M. F. Hasan, and Y. Anistyasari, "What's in a caption?: Leveraging caption pattern for predicting the popularity of social media posts," in *Proc. Int. Conf. Vocational Education and Electrical Engineering*, 2020, pp. 1–5.

[17] A.-A. Liu, X. Wang, N. Xu, J. Liu, Y. Su, Q. Zhang, S. Zhang, Y. Tang, J. Guo, G. Jin, and X. Li, "SMPC: boosting social media popularity prediction with caption," *Multimedia Systems*, vol. 29, pp. 577–586, 2023.

[18] J. Wang, W. Jiang, K. Li, G. Wang, and K. Li, "Incremental group-level popularity prediction in online social networks," *ACM Trans. Internet Technol.*, vol. 22, no. 1, 2021.

[19] S. Ji, X. Lu, M. Liu, L. Sun, C. Liu, B. Du, and H. Xiong, "Community-based dynamic graph learning for popularity prediction," in *Proc. ACM SIGKDD Conf. Knowledge Discovery and Data Mining*, 2023, pp. 930–940.

[20] I. T. Jolliffe, *Principal Component Analysis*. Springer New York, NY, 2002.

[21] A. Widyadhana, S. C. Hidayati, D. A. Navastara, and Y. Anistyasari, "ASF-LLRDA: Locality-regularized linear regression discriminant analysis with approximately symmetrical face preprocessing for face recognition," in *Proc. APSIPA Annual Summit and Conference*, 2023, pp. 2031–2036.

[22] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Int. Conf. Neural Information Processing Systems*, 1996, pp. 155–161.

[23] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[24] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in *Proc. Int. Conf. Neural Information Processing Systems*, 2018, pp. 6639–6649.

[25] A. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: a tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.

[26] Z.-M. Liu, Y.-Y. Chen, S. Hidayati, S.-C. Chien, F.-C. Chang, and K.-L. Hua, "3d model retrieval based on deep autoencoder neural networks," in *Proc. Int. Conf. Signals and Systems*, 2017, pp. 290–296.