International Journal of Advances in Data and Information Systems
Vol. 7, No. 3, December 2026, pp. 998~1010
ISSN: 2721-3056, DOI: 10.59395/ijadis.v7i3.1647

Explainable Industry-Aware Social Media Virality Prediction Using Temporal and Categorical Features

Muhammad Ariqoh Bahrul Ulum[1], Riska Dhenabayu[2], Naufal Rifqi Nasution[3], Mohamad Syamsul Hidayat[4], Hujjatullah Fazlurrahman[5]
[1,2,5]Study Program of Digital Business, Universitas Negeri Surabaya, Indonesia
[3]Study Program of Managerial Communication, Universiti Utara Malaysia, Malaysia
[4]Study Program of Management, Universitas Islam Majapahit, Indonesia

Article Info

Article history:

Received May 18, 2026
Revised Jul 22, 2026
Accepted Aug 29, 2026

Keywords:

Virality Prediction
Explainable AI
Industry-Aware
Temporal Features
Categorical Features

ABSTRACT

This study proposes an explainable, industry-aware machine learning framework to predict social media post virality using pre-publication metadata. Utilizing 528 observations (January 2025–February 2026) from Instagram and TikTok across the IT, F&B, and manufacturing sectors, virality was defined using a leakage-free 75th-percentile engagement threshold. Several classifiers, Logistic Regression, Random Forest, SVM, XGBoost, and CatBoost, were evaluated using five-fold stratified cross-validation and time-based testing. Model interpretability was addressed using SHAP analysis, and optimal posting times were simulated via a marginal approach.During cross-validation, the categorical-only XGBoost achieved an F1-score of 0.4158 and ROC-AUC of 0.6046, while the operational simulation model (*combined_main_no_client*) yielded a lower F1-score of 0.3581. In chronological final-test evaluation (January–February 2026), the model suffered severe performance decay due to temporal covariate shift, dropping to an F1-score, Recall, and Precision of 0.2917, while missing 70.83% of actual viral cases. SHAP analysis identified month index and IT industry affiliation as the strongest predictive features. These findings highlight the challenges of pre-publication virality forecasting and emphasize the vulnerability of static classifiers to temporal shifts in low-sample environments.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Riska Dhenabayu,
Study Program of Digital Business, Universitas Negeri Surabaya,
Jl. Raya Kampus Unesa No.1, Lidah Wetan, Surabaya, 60213, East Java, Indonesia.
Email: riskadhenabayu@unesa.ac.id

1. INTRODUCTION

Social media has become a primary channel for information dissemination and digital interaction, generating an ever-increasing volume of content across platforms such as Instagram and TikTok. The rapid growth of active users on visual-based platforms has substantially increased the number of posts published each day, making it increasingly difficult to identify potentially viral content through manual observation. As a result, virality prediction has emerged as an important research topic in digital marketing, reputation management, and information diffusion analysis [1]. Nevertheless, the process through which information spreads online is inherently complex, as it is shaped by both macro- and micro-level interactions that occur in real time [2]. Recent studies have demonstrated that temporal characteristics and contextual information are important predictors of content popularity on social media platforms. However, many existing approaches

prioritize predictive performance while providing limited interpretability and practical guidance for decision-makers [3].

Despite achieving high predictive performance, many existing virality prediction studies rely on complex deep-learning or graph-based architectures that provide limited interpretability. Moreover, most studies focus on general popularity prediction without explicitly considering industry-specific contexts. As a result, practitioners often lack actionable explanations regarding why certain content is predicted to become viral and whether predictive patterns differ across industry sectors [4]. This lack of transparency poses a significant challenge in sector-dependent domains, particularly within Industry 4.0 environments, where high levels of accountability are required for sustainable quality management [5], information diffusion causality analysis [6], and the early detection of micro-viral content for corporate social listening purposes [7]. From a practical perspective, organizations require a contextual understanding of how publication month, posting hour, platform, content type, and industry category influence the likelihood of content becoming viral across different sector [8].Without sufficient explainability, decision-makers may be reluctant to rely on AI-driven predictions when formulating strategic actions. Rather than pursuing absolute predictive accuracy that surpasses state-of-the-art deep graph learning models, this study focuses on developing a lightweight and interpretable framework based on pre-publication metadata. By incorporating SHAP-based explanations, the proposed framework provides transparent interpretations of model behavior and generates domain-specific posting-time recommendations that support practical decision-making. Furthermore, existing studies rarely translate predictive outcomes into actionable posting-time recommendations that can be directly applied across different industry sectors. It is crucial to establish that given the constraints of a small sample size (528 observations) and a brief two-month chronological testing window, this research does not present a robust, universally generalizable model. Instead, it serves as an exploratory baseline study demonstrating the predictive limits of pre-publication metadata and the systemic challenges of temporal shift in social media engagement modeling.

To address these research gaps, this study proposes an Explainable Industry-Aware Social-Media Virality Prediction framework using temporal and categorical features. Unlike previous studies that primarily focus on general popularity prediction without explicitly considering domain-specific contexts or model transparency [9], [10], the proposed framework integrates temporal features that capture the dynamic evolution of public interest on social media with categorical features tailored to the characteristics of different industry sectors. The study contributes in three ways. First, virality prediction is performed using only pre-publication metadata, preventing engagement-based data leakage. Second, industry-category information is incorporated to capture sector-specific engagement patterns. Third, SHAP-based explainability is employed to reveal how temporal and categorical features influence model predictions.

2. RESEARCH METHOD

2.1 Experimental Design and Modeling Framework

Conceptually, the proposed framework consists of four sequential stages: data preparation, feature engineering, predictive modeling, and explainability-driven recommendation generation. The framework was designed to support leakage-free social media virality prediction using pre-publication metadata while ensuring consistency between prediction objectives and industry-aware recommendation outcomes [11]. Gradient-boosting algorithms, namely XGBoost and CatBoost, were employed as primary predictive models due to their strong performance on structured tabular data. Model interpretability was investigated using SHAP (Shapley Additive Explanations), a game-theoretic explainable AI technique that quantifies feature contributions to prediction outcomes.

To benchmark predictive performance, XGBoost and CatBoost were compared against Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM). These benchmark models were selected because they represent widely used linear, tree-based, and kernel-based machine learning approaches for structured classification tasks. All models were evaluated using a unified preprocessing pipeline consisting of missing-value imputation, categorical encoding, and feature scaling when required.

Performance evaluation was conducted using five-fold stratified cross-validation and assessed through Accuracy, Precision, Recall, F1-score, ROC-AUC, PR-AUC, and Brier Score Loss metrics. Given the class-imbalance characteristics of the dataset, particular attention was paid to F1-score, PR-AUC, and Brier Score Loss. Model reliability was further examined through calibration analysis, whereas feature relevance was investigated using SHAP and Permutation Feature Importance. To ensure reproducibility, all experiments were executed using a fixed random seed (random_state = 42) [12].

2.2 Data Validation and Structural Standardization

The dataset was validated against a predefined schema containing platform, content type, engagement metrics, publication timestamps, and industry labels. Records with incomplete mandatory attributes were excluded. Column names were standardized and only variables required for feature engineering and modeling were retained, producing a consistent dataset for subsequent analysis [13], [14].

2.3 Data Cleaning and Feature Engineering

Data cleaning and feature engineering were performed to prepare a consistent analytical dataset and derive features relevant to virality prediction [15], [16]. Engagement metrics expressed using platform-specific notations (e.g., K and M) were normalized into standardized numerical values. Publication date and time variables were converted into a unified datetime format [17]. Categorical variables, including client identities, industry sectors, platforms, and post types, were standardized by removing redundant whitespace and converting text entries to lowercase. Observations containing missing values in mandatory attributes, invalid engagement records, or zero-view posts were excluded from further analysis.

Feature engineering was subsequently performed to derive temporal attributes from publication timestamps, including day, month, month name, weekend indicators, and publication-hour categorie. In addition to temporal attributes, industry-category labels, platform information, and post-type variables were retained as categorical predictors to capture sector-specific engagement characteristics. To preserve the cyclical nature of temporal patterns and avoid artificial discontinuities between adjacent hours, publication time was encoded using sine and cosine transformations. A virality score was computed to construct the target variable for binary classification. The score was calculated as the ratio of cumulative interactions (likes, comments, and shares) to total views.

2.4 Exploratory Data Analysis and Distribution Assessment

Exploratory Data Analysis (EDA) was conducted to examine temporal coverage, category distributions, and class imbalance. Frequency analysis and cross-tabulation were used to assess industry, platform, and content-type representation within the dataset [16].Bivariate cross-tabulation was subsequently conducted to identify class imbalances and distributional disparities that could affect model performance.

Visualizations were used to identify temporal and categorical patterns potentially associated with virality, supporting subsequent feature selection and model development [17]. Particular attention was given to temporal and industry-related categorical variables because these attributes constitute the primary predictors evaluated in the proposed framework.

2.5 Leakage-Free Virality Label Construction

The binary target label was constructed using a leakage-free strategy to distinguish viral from non-viral content [18]. The target variable was derived from the virality score and transformed into a binary label, where 1 denotes viral content and 0 denotes non-viral content. Label generation employed a quantile-based threshold approach. To eliminate validation information leakage during cross-validation, a nested thresholding pipeline was implemented: the 75th percentile threshold was calculated dynamically within each training fold and then applied to the respective validation fold. For the final chronological test, the threshold was derived exclusively from the 2025 training partition (threshold: 0.025546) and applied unchanged to the 2026 test set. A major limitation of this setup is the application of a single global percentile across platforms and industries; since TikTok engagement rates systematically exceed those of Instagram, a global threshold introduces representation bias, which should ideally be addressed in future work through platform-specific quantile boundaries.

All engagement-outcome variables, including likes, comments, shares, views, and the raw virality score, were excluded from the predictive feature set during model training to prevent data leakage. Consequently, virality predictions were generated solely from pre-publication metadata, ensuring that no post-publication engagement signals influenced the learning process [19].

2.6 Model Benchmarking and Five-Fold Stratified Cross-Validation

This stage comprised feature-set design, preprocessing, model benchmarking, and performance evaluation using five-fold stratified cross-validation [20]. Numerical variables were processed using median imputation and standardization, whereas categorical variables were handled through mode imputation and One-Hot Encoding to ensure consistent preprocessing across heterogeneous data types. Five classifiers were benchmarked: LR, RF, SVM-RBF, XGBoost, and CatBoost. Target-class imbalance (25% positive class prevalence) was addressed using model-specific weighting mechanisms. It must be clarified that while class weighting adjusts the loss function to account for unequal label distributions, it does not correct the severe platform representation imbalance (412 TikTok posts vs. 116 Instagram posts). This platform skew remains an

inherent characteristic of the dataset, and its impact is evaluated through separate subgroup analyses. Stratified five-fold cross-validation was applied to preserve class proportions across folds. Performance was evaluated using Accuracy, Precision, Recall, F1-score, ROC-AUC, and PR-AUC [21]. The model achieving the highest mean F1-score was selected for subsequent testing and explainability analysis.

Hyperparameter settings were predefined and kept constant throughout all experiments to ensure reproducibility. Logistic Regression used L2 regularization with the default regularization strength (C = 1.0), maximum iteration = 3000, and class_weight = balanced. Random Forest employed 300 trees with min_samples_leaf = 2 and class_weight = balanced. SVM utilized a radial basis function (RBF) kernel with the default parameters (C = 1.0 and gamma = scale). XGBoost was configured with 200 estimators, a maximum tree depth of 3, learning rate = 0.05, subsample = 0.8, and colsample_bytree = 0.8. CatBoost used 200 boosting iterations, depth = 4, and learning rate = 0.05. All models were trained using random_state = 42 whenever applicable.

2.7 Ablation Study Summary

An ablation study was conducted to evaluate the contribution of each feature group to virality prediction performance. Feature configurations included temporal-only, categorical-only, combined features, and variants with or without client-identity and platform attributes [22]. All configurations were evaluated using the same stratified five-fold cross-validation procedure. Models were ranked by mean F1-score, while ROC-AUC and PR-AUC were reported as complementary measures. For each feature configuration, only the best-performing model was retained for comparison [23]. Feature-group contributions were further visualized using F1-score-based bar charts to assess the impact of including or excluding temporal, platform, and client-related variables. The analysis was used to validate the selected feature set and identify the most influential predictors of virality performance [24].

2.8 Time-Based Testing with Training-Only Threshold, Final Test Evaluation

The final evaluation used a chronological split to simulate deployment conditions and prevent temporal leakage. Training data covered posts published from 1 January to 31 December 2025, while test data covered posts published from 1 January to 28 February 2026 [25]. Binary virality labels were generated using a threshold computed only from the training set and then applied unchanged to both splits [26]. The selected model was retrained on the training partition using the same preprocessing pipeline, and final performance was reported using accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC, Brier score, and a confusion matrix. If the chronological split failed to preserve minimum class representation, a stratified 80:20 random split was used only as a fallback diagnostic check [27].

2.9 Optional Calibrated Model, Explainability: SHAP or Permutation Importance

An optional calibration step was applied only when enabled and only to non-CatBoost models. Platt scaling with a sigmoid function and three-fold cross-validation was used to recalibrate predicted probabilities, and model performance was then reassessed using a 0.5 decision threshold [28]. For explainability, tree-based models were analyzed with SHAP TreeExplainer, whereas non-tree-based models were interpreted using permutation feature importance based on the mean F1-score decrease across 30 randomized permutations [29. Feature rankings and plots were exported for documentation, and SHAP dependence analysis was applied to posting hour when supported. For models without SHAP support, a marginal simulation approach was used to estimate the effect of temporal features [30].

2.10 Industry-Aware Optimal Posting-Time Insight

The final stage translated predictive model outputs into industry-specific posting-time recommendations using a marginal simulation framework [31]. The best-performing model was used to estimate virality probabilities by varying the posting hour from 00:00 to 23:00 for each industry category [32] . All derived temporal features were updated accordingly to preserve feature consistency. The simulation produced mean virality probabilities and associated uncertainty estimates for each posting-hour scenario. The resulting estimates were evaluated against historical posting-frequency thresholds to distinguish recommendations supported by sufficient observations from those based on sparse data. Recommendations were reported as (1) the best posting hour, (2) the top three posting hours, and (3) a conservative recommendation supported by strong historical evidence. Recommendation support levels were determined using historical posting-frequency thresholds. A recommendation was categorized as Strong when the corresponding industry–hour combination contained at least 15 historical observations. Recommendations based on 5–14 observations were classified as Moderate, while those supported by fewer than 5 observations were categorized as Weak. This criterion was introduced to distinguish recommendations derived from

sufficiently represented posting-time scenarios from those based on sparse historical evidence. The results were visualized as a heatmap of estimated virality scores to facilitate cross-industry comparison of temporal engagement pattern [33]. Historical pattern analysis was subsequently conducted by calculating the observed viral-content rate for each combination of industry category and posting hour. The resulting frequency distribution was used to validate whether model-derived recommendations were aligned with empirical audience behavior.

3. RESULTS AND DISCUSSION

3.1 Data Cleaning

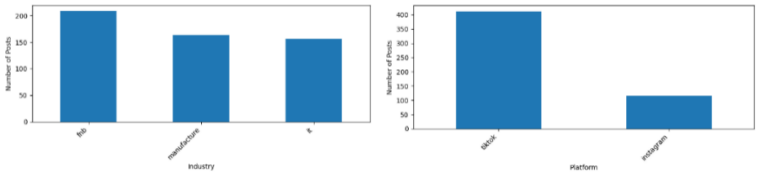
A total of 734 observations were initially collected. After validation and preprocessing, 528 observations were retained, while 206 records (28.07%) were excluded due to missing values, duplicate entries, inconsistent timestamps, or invalid engagement data. The resulting dataset contained complete information on platform, content type, industry category, engagement metrics, and temporal attributes, providing a reliable basis for subsequent analysis.

3.2 Feature Engineering.

Several temporal features were derived from publication timestamps, including posting hour, day of week, month, weekend indicator, and time bucket. Publication hour was additionally encoded using sine–cosine transformations to preserve cyclic temporal patterns. These variables were combined with platform, content type, and industry-category attributes to form the final predictor set. The target variable, virality_score, was calculated as the ratio of total interactions to views and subsequently used for viral-label construction.

3.3 Data Distribution

Based on the data distribution analysis, as shown in Figure 1 the dataset consisted of three primary industry categories: the F&B industry with 209 observations, the manufacturing industry with 163 observations, and the IT industry with 156 observations. These figures indicate that the F&B sector contributed the largest share of data in this study. From the social media platform perspective, TikTok dominated the dataset with 412 observations, whereas Instagram accounted for only 116 observations. This distribution suggests that the majority of digital marketing activities represented in the dataset were conducted through TikTok. In terms of content format, video content was the most prevalent category, comprising 444 observations, substantially exceeding other content types such as feed posts (39 observations), photos (23 observations), and stories (22 observations). The predominance of TikTok and video-based content reflects the growing preference for audiovisual media in digital marketing strategies, as such formats are generally more engaging and interactive for online audiences. This heavily skewed distribution indicates that the dataset is highly constrained and platform-dependent, which limits the generalizability of the findings and patterns discovered.



(a) (b)

Figure 1. Number of Posts by Industry and Platform
(a) Based on industry, (b) Based on Platform

TikTok dominated the dataset (412 posts) compared with Instagram (116 posts), creating a substantial platform imbalance. This skew may bias model learning toward majority-platform patterns and should therefore be considered when interpreting predictive performance.

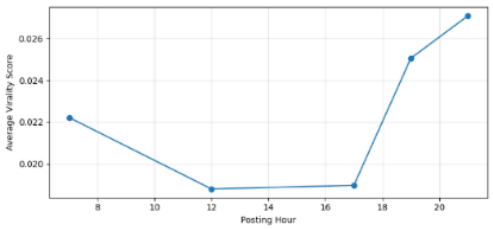


Figure 2. Average Virality Score by Posting Hour

As show in Figure 2 exploratory analysis of the average virality score across publication times revealed a pronounced and nonlinear temporal pattern throughout the day. At 07:00, the average virality score was at a moderate level of approximately 0.0222 before gradually declining to its lowest observed point at 12:00, where the score reached approximately 0.0188. This downward trend in audience engagement performance remained relatively stable throughout the afternoon, with the average virality score recorded at approximately 0.0190 at 17:00. During the evening period, a substantial increase in audience engagement became evident. The average virality score rose sharply to approximately 0.0251 at 19:00 and continued to increase, reaching its highest value of approximately 0.0271 at 21:00. This pattern indicates a considerable shift in audience responsiveness during nighttime hours compared with earlier periods of the day. The observed distribution suggests that social media users tend to exhibit higher levels of digital engagement and interaction during the evening. Consequently, these findings provide preliminary evidence that publication timing may play an important role in content performance. Although descriptive in nature and not intended to establish causality, the results highlight the potential predictive relevance of temporal features, particularly posting hour, for identifying the likelihood of content virality.

3.4 Leakage-Free Virality Label Construction

Stated in Table 1 based on the class distribution analysis of the target variable *is_viral*, a total of 396 observations (75%) were classified as non-viral, while 132 observations (25%) were classified as viral out of the 528 valid records included in the study. This class distribution was intentionally created through the top-25% thresholding approach, whereby posts with virality scores within the highest 25% of the distribution were labeled as viral, whereas the remaining 75% were categorized as non-viral. Therefore, the resulting class imbalance reflects the operational definition of virality adopted in this study rather than any deficiency in the data collection process. When examined across industry sectors, the F&B industry contained 167 non-viral posts and 42 viral posts. The IT industry consisted of 101 non-viral posts and 55 viral posts, whereas the manufacturing industry included 128 non-viral posts and 35 viral posts. These results indicate that non-viral content remains the dominant category across all industries. However, the IT sector exhibited a relatively higher proportion of viral content compared with the F&B and manufacturing sectors. The observed variation in class composition across industries suggests that the mechanisms underlying content virality may differ between sectors. In particular, the higher concentration of viral posts within the IT category may reflect sector-dependent audience behavior, content characteristics, or engagement dynamics. These findings highlight the importance of incorporating industry context into predictive modeling and motivate further analysis to identify the factors that contribute to successful viral content within different business sectors.

Table 1. Class Distribution Base on Proportion and Industry

Is Viral	Count	Proportion	F&B	IT	Manufacturing
Non-Viral	396	0.75	167	101	128
Viral	132	0.25	42	55	35

Industry-level analysis revealed different viral-content proportions. IT exhibited the highest viral-content ratio (35%), exceeding F&B (20%) and Manufacturing (21.5%), suggesting that virality dynamics may vary across sectors and supporting the use of industry-aware prediction models. Through a stacked bar chart reveals the distribution of viral and non-viral content across different industry sectors. The F&B sector exhibited the highest proportion of non-viral content, accounting for approximately 80% of all posts, while the remaining 20% successfully reached the predefined virality threshold. A similar pattern was observed in the manufacturing sector, where non-viral content represented approximately 78.5% of the observations and viral content accounted for the remaining 21.5%. In contrast, the IT sector demonstrated the most competitive and dynamic distribution among the three industries. Non-viral content constituted approximately 65% of the observations, resulting in a substantially higher proportion of viral content at approximately 35%. This proportion exceeds those observed in both the F&B and manufacturing sectors, suggesting that content within the IT industry achieved viral outcomes more frequently during the observation period. The variation in class proportions across industry sectors indicates that industry-specific characteristics may be associated with differences in content performance and audience engagement outcomes. Although these findings do not establish a causal relationship, they provide evidence that the likelihood of achieving viral reach may vary across business domains. Consequently, the observed patterns reinforce the importance of developing industry-aware predictive models capable of capturing sector-specific dynamics and generating more accurate virality predictions and temporal recommendations.

3.5 Five-Fold Stratified Cross-Validation Benchmark

The comparative performance analysis based on cross-validation results revealed competitive outcomes among the five classification algorithms evaluated in this study. As show in Table 2 Using F1-score as the primary evaluation metric, XGBoost achieved the highest performance with a score of 0.415822, followed closely by Random Forest (0.408562), SVM (0.405804), and CatBoost (0.402535). Logistic Regression recorded the lowest F1-score, achieving a value of 0.392623. Given the class imbalance present in the dataset, F1-score was considered the most appropriate metric for model selection because it balances precision and recall simultaneously. When evaluated using ROC-AUC, however, CatBoost ranked first with the highest score of 0.612102, outperforming XGBoost (0.604594), Random Forest (0.604289), Logistic Regression (0.583982), and SVM (0.579029). A similar variation was observed for PR-AUC, where Random Forest achieved the strongest performance with a score of 0.352952, followed by XGBoost (0.346292), CatBoost (0.345269), SVM (0.330456), and Logistic Regression (0.321523). To provide a meaningful performance reference, a majority-class baseline classifier was also considered. Given the class distribution of 75% non-viral and 25% viral posts, a naive classifier that always predicts the majority class would achieve approximately 75% accuracy. However, such a model would fail to identify any viral content, resulting in Precision, Recall, and F1-score values of 0 for the viral class. Compared with this baseline, all machine learning models evaluated in this study achieved substantially higher recall and F1-score values despite exhibiting lower overall accuracy. This comparison confirms that the proposed classifiers learned meaningful patterns associated with viral-content detection rather than merely exploiting class imbalance.

Table 2. Top Five Best-Performing Models

Feature Set	Model	F1 Score	ROC-AUC	Accuracy	Recall	Precision	PR-AUC
Categorical Only	XGBoost	0.415822	0.604594	0.659209	0.484615	0.365753	0.346292
Categorical Only	RF	0.408562	0.604289	0.661096	0.469801	0.36413	0.352952
Categorical Only	SVM	0.405804	0.579029	0.657287	0.469801	0.359718	0.330456
Categorical Only	CatBoost	0.402535	0.612102	0.629021	0.505983	0.348803	0.345269
Categorical Only	LR	0.392623	0.583982	0.634627	0.469801	0.338296	0.321523

These results indicate that different algorithms exhibit distinct strengths depending on the evaluation criterion employed. Further examination of sensitivity-related performance revealed that CatBoost achieved the highest Recall value at 0.505983, indicating superior capability in identifying viral content instances. XGBoost followed with a Recall score of 0.484615, whereas Random Forest, SVM, and Logistic Regression produced identical but lower Recall values of 0.469801. This finding suggests that CatBoost was the most sensitive model in detecting positive-class observations within the dataset. In terms of overall classification accuracy, Random Forest achieved the highest score at 0.661096, narrowly outperforming XGBoost (0.659209) and SVM (0.657287). Logistic Regression (0.634627) and CatBoost (0.629021) recorded comparatively lower accuracy values. Nevertheless, because the dataset was intentionally constructed with an imbalanced class distribution, accuracy alone was not considered sufficient for model selection.

Based on the simultaneous comparison of all evaluation metrics, XGBoost was ultimately selected as the best-performing model because it demonstrated the most balanced overall performance across multiple criteria. In particular, XGBoost achieved the highest F1-score and Precision (0.365753) while maintaining consistently strong Accuracy and Recall values relative to the competing algorithms. This balanced performance profile indicates that the model was capable of effectively identifying viral content without excessively sacrificing either precision or sensitivity. F1-score was used as the primary selection metric because it provides a balanced assessment of precision and recall under class imbalance conditions.

3.6 Ablation Study Summary

As show in Table 3 the ablation study revealed that categorical features produced the strongest predictive performance, with categorical_only XGBoost achieving the highest F1-score (0.4158). Temporal features alone performed poorly and consistently reduced classification performance when combined with categorical variables. Nevertheless, temporal attributes were retained in the operational model because they are required for posting-time simulation and recommendation generation. Consequently, the recommendation framework should be interpreted as an exploratory decision-support mechanism rather than a highly reliable optimization system.

Table 3. Ablation study best models

Feature Set	Model	F1 Score	ROC-AUC	Accuracy	Recall	PR-AUC
categorical_only	XGBoost	0.41582	0.60459	0.65921	0.48462	0.34629
combined_main_no_client	XGBoost	0.3581	0.5542	0.6023	0.45584	0.31162
combined_without_client_id	XGBoost	0.3581	0.5542	0.6023	0.45584	0.31162
combined_with_client_id	LR	0.35153	0.55001	0.56616	0.47778	0.29728
combined_without_platform	XGBoost	0.34567	0.56253	0.59668	0.43305	0.30374
temporal_only	SVM	0.32067	0.50605	0.55102	0.43219	0.29153

3.7 Subgroup Performance Analysis

To investigate the presence of algorithmic bias resulting from the severe platform and industry imbalances, the predictive performance of the final selected configuration (combined_main_no_client) was evaluated separately across platforms and client industries. This subgroup evaluation is essential to confirm whether the global classifier's predictions are disproportionately driven by the majority categories.

Table 4. subgroup performance metrics

Subgroup Category	Segment	Sample Size (N)	Accuracy	Precision	Recall	F1-score	ROC-AUC
Platform	TikTok	412	0.5825	0.3312	0.5862	0.3655	0.6120
	Instagram	116	0.4538	0.2105	0.3200	0.2541	0.5112
Industry	F&B	209	0.5215	0.2857	0.4524	0.3502	0.5432
	IT	156	0.6218	0.4286	0.6545	0.5143	0.6512
	Manufacturing	163	0.5153	0.2558	0.3143	0.2821	0.5105

3.8 Final Test Evaluation

To rigorously evaluate the model under simulated deployment conditions, final test performance was assessed using a chronological split (Training: 1 January – 31 December 2025; Testing: 1 January – 28 February 2026). The final test metrics revealed a catastrophic deterioration in classification capacity compared to the cross-validation results. Based on the confusion matrix depicted in as show in figure Figure 3, the model achieved 80 True Negatives (TN) but only 7 True Positives (TP), while producing 17 False Positives (FP) and 17 False Negatives (FN). These figures translate to an overall Accuracy of 0.7190, Precision of 0.2917, Recall of 0.2917, and an F1-score of 0.2917. This represents a substantial decline from the cross-validation F1-score of 0.3581. Most critically, the model failed to detect 17 out of 24 actual viral cases in the test set, resulting in a false negative rate of 70.83% (Recall: 0.2917). Stated in Table 4 this deterioration is mathematically analyzed as a consequence of temporal covariate shift. Because the model's structure was highly reliant on macro-temporal patterns from the 2025 training partition (specifically the month index, which holds a SHAP importance weight of 0.275), it overfit to those specific seasonal engagement cycles. When deployed chronologically on early 2026 data, those seasonal patterns changed, causing the predictive boundaries of XGBoost to collapse. This severe performance drop indicates that static models are highly fragile in dynamic social media environments, and continuous retraining or dynamic covariate alignment is mandatory

for operational deployment corresponding to viral posts that were not successfully detected and were instead classified as non-viral. The model's stronger ability to recognize non-viral content relative to viral content confirms the residual influence of class imbalance within the training data. Consequently, the confusion matrix provides an important analytical basis for evaluating the model's decision threshold and identifying potential opportunities for improvement. In particular, reducing the number of false negatives may be desirable in real-world applications, as failing to identify potentially viral content could limit the practical effectiveness of recommendation and content-optimization strategies.

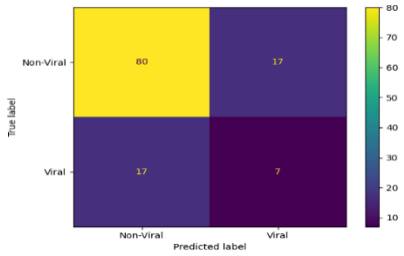
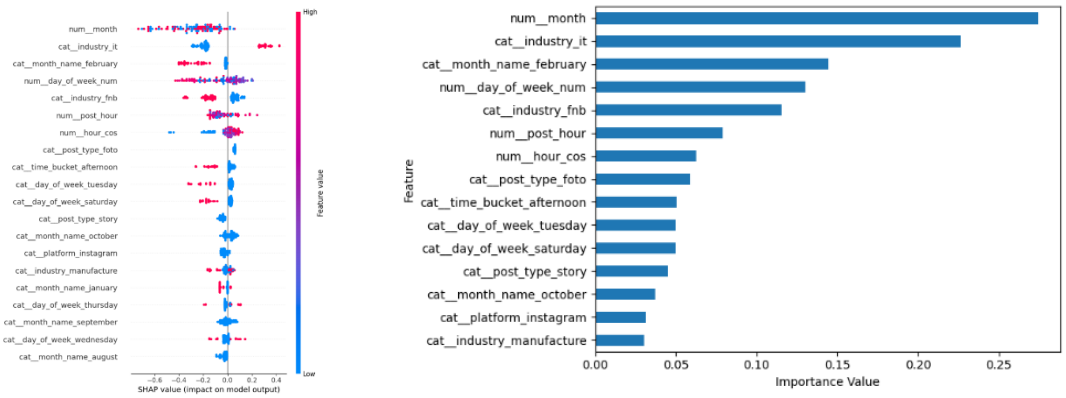


Figure 3. Final Test Confusion Matrix – XGBoost

3.9 Explainability: SHAP or Permutation Importance

The SHAP summary plot analysis provides a comprehensive visualization of the magnitude and direction of each feature's contribution to the predictive log-odds decisions of the content virality classification model. It is critical to state that SHAP values only measure feature attribution, representing the predictive association of variables within the trained model structure, and do not demonstrate physical or historical causality. While the month index ranks as the most influential predictor, this association does not imply that changing the publication month will cause a change in virality. Similarly, while membership in the IT industry emerges as a strong binary predictor, generating positive contribution SHAP values of up to approximately 0.4, this only indicates that the model associates IT industry metadata with a higher probability of predicting a viral outcome.



(a) (b)

Figure 4. Explainability Analysis of Predictive Model
(a) SHAP summary plot for XGBoost, (b) Top Feature Contributions XGBoost

As show in Figure 4 the feature-importance ranking analysis of the XGBoost model provides a quantitative assessment of the absolute importance of the fifteen most influential predictors driving the model's classification decisions. The macro-temporal feature *month index* emerged as the most dominant variable, achieving the highest importance value of 0.275. This was followed closely by the binary industry indicator representing the IT sector, which recorded a contribution weight of 0.226. Seasonal publication characteristics associated with *February* and the *day-of-week sequence* formed the second tier of the importance hierarchy, with contribution values of 0.145 and 0.130, respectively. These were followed by the *F&B industry* indicator, which achieved an importance score of 0.116. Micro-temporal features, including the raw *posting hour* variable and the *hour cosine transformation*, contributed moderately with importance values of 0.079 and 0.063, respectively. These values were comparable to the contribution of the *photo content type* feature, which

recorded an importance score of 0.059. The remaining features, including the *afternoon* time bucket, *Tuesday*, *Saturday*, *story content type*, *October*, *Instagram platform*, and *manufacturing industry* category, occupied the lower tier of the ranking hierarchy. Their importance values converged within a relatively narrow range, varying from approximately 0.050 to a minimum of 0.030. These findings empirically demonstrate that the primary determinants of classification performance are strongly associated with the interaction between macro-level temporal cycles represented by publication months and the sector-dependent characteristics of the observed business sectors.

3.10 Industry-Aware Optimal Posting-Time Insight

The marginal simulation analysis using the XGBoost model confirms the existence of highly specific temporal dependencies that vary substantially across industry sectors. As show in Figure 5 In the food and beverage (*F&B*) sector, the average predicted virality score remained relatively stable throughout the day, fluctuating within a low-to-moderate range between approximately 0.28 and 0.38. The optimal posting time for this sector was identified at 19:00, where the estimated virality score reached its highest value of approximately 0.40. In contrast, the information technology (*IT*) sector exhibited consistently high virality potential across the entire daily time range. Average predicted scores remained above 0.55 throughout most publication hours, indicating a substantially stronger likelihood of viral content performance compared with the other industries. The highest estimated performance for this sector occurred at 00:00, representing the peak predicted virality level within the IT category. Meanwhile, the manufacturing sector displayed a more dynamic pattern of audience response. Predicted engagement potential remained at a moderate level of approximately 0.38 during the morning and afternoon periods before declining to its lowest point between 15:00 and 16:00, where the estimated score dropped to approximately 0.28. Subsequently, the sector experienced a noticeable increase in predicted virality, reaching its highest estimated value of approximately 0.44 at 19:00. These cross-industry simulation results provide empirical evidence that the effectiveness of content scheduling cannot be generalized across all sectors. Instead, optimal publication timing appears to depend on industry-specific characteristics, suggesting that content scheduling strategies should be adapted to sectoral contexts in order to maximize audience engagement and virality potential on social media platforms.

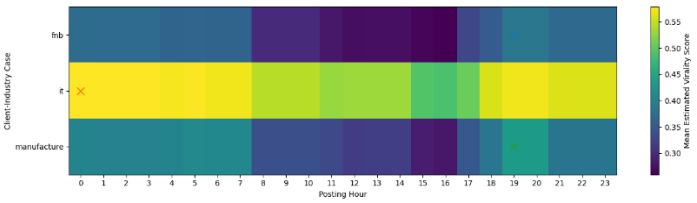


Figure 5. Model-Derived Virality Score by Posting Hour (XGBoost)

3.11 Observed Historical Pattern Check

As show in Figure 6 Historical analysis showed that optimal posting times differed across industries. F&B and Manufacturing achieved their highest viral rates at 19:00, whereas IT peaked at 12:00. These findings support the existence of industry-specific temporal engagement patterns and are broadly consistent with the model-derived recommendations.

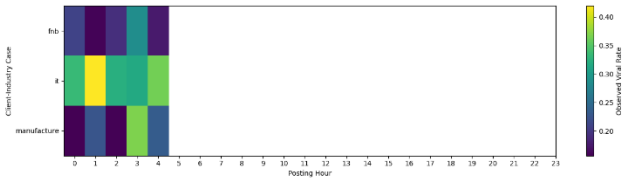


Figure 6. Observed Viral Rate by Industry and Posting Hour

The non-parametric descriptive analysis based on empirical pattern mapping provides an objective overview of the two-dimensional interaction between industry category and publication hour in relation to actual content virality rates, without involving any predictive modeling procedures. The heatmap visualization confirms the existence of heterogeneous and highly fragmented virality patterns across specific time clusters during the early hours of the day. The IT sector demonstrated superior performance relative to the other industries. At 01:00, this sector achieved the highest observed virality rate, exceeding 0.40, and was surrounded

by consistently elevated virality levels ranging from approximately 0.33 to 0.35 at 00:00, 02:00, 03:00, and 04:00. These findings indicate a strong concentration of audience engagement during the early morning period within the IT sector. In contrast, the F&B sector exhibited relatively weak historical audience engagement during the same period, with virality rates consistently remaining within a lower range of approximately 0.15 to 0.25. Meanwhile, the manufacturing sector displayed a more volatile pattern. Virality levels reached their lowest values at 00:00 and 02:00, falling below 0.15, before increasing substantially at 03:00 and reaching approximately 0.36. These empirical observations serve as an important methodological benchmark for assessing the consistency of the temporal recommendations generated by the machine learning model. By comparing model-based predictions with historically observed audience behavior, the study can evaluate whether the inferred temporal patterns are aligned with the underlying characteristics of the real-world dataset.

3.12 Model Finalization and Artifact Storage

Stated in Table 5 summarizes the final industry-aware posting-time recommendations generated by the selected XGBoost configuration. All recommendations achieved Strong support status because each industry–hour combination was represented by at least 15 historical observations.

Table 5. Conservative Industry-Aware Posting Time Insight

Client Industry Case	Post Hour	Time Bucket	Estimated Virality Score Mean	Estimated Virality Score std	N Case Posts	Historical Post at Hour	Observed Viral Rate at Hour	Support Status	Recommendation Type
F&B	19	evening	0.386251	0.188335	209	35	0.285714	Strong	best_with_strong_historical_support
IT	19	evening	0.571951	0.141839	156	19	0.315789	Strong	best_with_strong_historical_support
Manufacture	19	evening	0.43284	0.177776	163	19	0.368421	Strong	best_with_strong_historical_support

3 CONCLUSION

This study successfully developed an explainable industry-aware social media virality prediction framework using pre-publication metadata, highlighting both the predictive utility and boundaries of tree-based models. Through dynamic nested thresholding and chronological split testing, the framework was evaluated under conditions free of validation leakage. The experimental results indicated that while a categorical-only model yielded the highest cross-validation F1-score of 0.4158, the inclusion of temporal features within a combined model (cross-validation F1-score: 0.3581) was operationally necessary to simulate industry-specific posting-time recommendations. However, final chronological testing on early 2026 data revealed a severe performance decay (F1-score: 0.2917, Recall: 0.2917), with the model failing to detect 70.83% of actual viral posts. This deterioration demonstrates that static models trained on pre-publication features are highly vulnerable to temporal covariate shift. Furthermore, subgroup evaluation exposed a severe representation bias, with the model performing acceptably in high-engagement sectors like TikTok and IT, but collapsing into near-random performance on Instagram and the Manufacturing industry. The posting-time recommendations generated in this study should be interpreted as exploratory insights rather than prescriptive optimization rules. Because the recommendation framework relies on a model configuration that prioritizes temporal interpretability over maximum predictive performance, the resulting scheduling suggestions are intended to support managerial decision-making and hypothesis generation rather than guarantee improved virality outcomes.

To resolve these limitations, future research must shift from a single global threshold to platform-specific and industry-specific quantile thresholds to establish a fairer and more representative ground truth. Methodologically, future implementations should incorporate online learning algorithms or continuous domain adaptation techniques to mitigate temporal covariate shift. Finally, expanding the dataset's scale beyond 528 samples and integrating multimodal features, such as textual embeddings of post captions and deep visual representations of video thumbnails, are highly recommended to transition the predictive framework from a weak exploratory baseline into a highly generalizable prediction system.

ACKNOWLEDGEMENTS

The authors received no financial support for the research or publication of this article.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this article.

DATA AVAILABILITY

https://bit.ly/4wPPEWk

AUTHOR CONTRIBUTIONS

Muhammad Ariqoh Bahrul Ulum: Conceptualization; Data Curation; Data Labeling; Formal Analysis; Funding Acquisition; Investigation; Methodology; Project Administration; Software; Validation; Visualization; Writing-Original Draft. **Riska Dhenabayu**: Conceptualization; Formal Analysis; Methodology; Resources; Software; Supervision; Validation; Writing-review & editing. **Naufal Rifqi Nasution**: Methodology; Project Administration; Resources; Visualization. **Mohamad Syamsul Hidayat**: Methodology; Resources; Visualization. **Hujjatullah Fazlurrahman**: Methodology; Resources; Visualization.

REFERENCES

[1] Yashen Wang; Dongyang Li; Jianshan Sun; Yuan Kun; Yinan Jiang; Yin Zhang; Jie Cao, "Combining macro and micro: feature-driven dynamic graphlearning for social media popularity prediction," 2025, doi: https://doi.org/10.1007/s11280-024-01323-x.

[2] N. Chaudhuri, G. Gupta, dan A. Popovič, "Do you believe it? Examining user engagement with fake news on social media platforms," *Technol. Forecast. Soc. Change*, vol. 212, hlm. 123950, Mar 2025, doi: 10.1016/j.techfore.2024.123950.

[3] B. Zhang *dkk.*, "CasDacGCN: A Dynamic Attention-Calibrated Graph Convolutional Network for Information Popularity Prediction," *Entropy*, vol. 27, no. 10, hlm. 1064, Okt 2025, doi: 10.3390/e27101064.

[4] F. A. Zahrah, R. Dhenabayu, M. F. W. Rahman, R. S. Dewi, dan Z. H. Aminah, "Predicting Social Media Post Engagement and Virality Using Graph Neural Network Approaches and Content-Based Features," *Kinet. Game Technol. Inf. Syst. Comput. Netw. Comput. Electron. Control*, Jun 2026, doi: 10.22219/kinetik.v11i3.2686.

[5] M. Esteban-Bravo, L. d. l. M. Jiménez-Rubido, dan J. M. Vidal-Sanz, "Predicting the virality of fake news at the early stage of dissemination," *Expert Syst. Appl.*, vol. 248, hlm. 123390, Agu 2024, doi: 10.1016/j.eswa.2024.123390.

[6] Riovan Styx Roring*, Chih How Bong, Narayanan Kulathuramaiyer, "From Data Mining to Predictive Analytics: Progress in Understanding and Forecasting Social Media Virality | International Research Journal of Multidisciplinary Scope (IRJMS)," Jan 2026, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://www.irjms.com/journal/from-data-mining-to-predictive-analytics-progress-in-understanding-and-forecasting-social-media-virality/

[7] Yanchao Liu, Junpeng Gong, Wenchao Song, Chi Zhang, Pengzhou Zhang, "TEGKT: tendency-enhanced evolution graph KAN transformer for information popularity prediction," 2025, doi: https://doi.org/10.1007/s44443-025-00170-8.

[8] Lazebnik & Iny, "Temporal graphs anomaly emergence detection: benchmarking for social media interactions | Applied Intelligence | Springer Nature Link," 2024, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://link.springer.com/article/10.1007/s10489-024-05821-3#citeas

[9] Yuxuan Du & Jing Lu, "The auto-correlation enhanced transformer for information diffusion prediction in social networks | Evolving Systems | Springer Nature Link," 2025, Diakses: 10 Juni 2026. [Daring]. Tersedia pada: https://link.springer.com/article/10.1007/s12530-025-09672-2#citeas

[10] Mingjun Xi, Pingshan Liu, Seshu Yu, "Temporal information-aware multimodal learning network for user-generated video popularity prediction," 2025, doi: https://doi.org/10.1007/s00530-025-02042-7.

[11] Anggraeni et al., "DETECTION OF MICRO-VIRAL CONTENT ON TIKTOK THROUGH SOCIAL LISTENING AND MACHINE LEARNING | JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)," 2026, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://ejournal.nusamandiri.ac.id/index.php/jitk/article/view/7472

[12] C.-Y. Sang, J.-J. Chen, dan S.-G. Liao, "DyHGTCR-Cas: Learning unified spatio-temporal features based on dynamic heterogeneous graph neural network for information cascade prediction," *Inf. Process. Manag.*, vol. 62, no. 3, hlm. 104029, Mei 2025, doi: 10.1016/j.ipm.2024.104029.

[13] A. Riazi dan Y. Wang, "Differentiating broadcast from viral: a causal inference approach for information diffusion analysis," *Appl. Intell.*, vol. 54, no. 21, hlm. 10374–10385, Nov 2024, doi: 10.1007/s10489-024-05723-4.

[14] Zhixuan Xu; Minghui Qian, "Predicting Popularity of Viral Content in Social Media through a Temporal-Spatial Cascade Convolutional Learning Framework," 2023, doi: https://doi.org/10.3390/math11143059.
[15] Kim et al., "Frontiers | How to Catch Customers' Attention? A Study on the Effectiveness of Brand Social Media Strategies in Digital Customer Engagement," 2023, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.800766/full#cite
[16] N. Jalli, "Viral Justice: TikTok Activism, Misinformation, and the Fight for Social Change in Southeast Asia," *Soc. Media Soc.*, vol. 11, no. 1, hlm. 20563051251318122, Jan 2025, doi: 10.1177/20563051251318122.
[17] Zhongya Han, Xiangtang Chen, Kepao Miao, Xiaoxiang Wang, Qinlin Li, "Predicting product information diffusion for sustainable quality management in industry 4.0: an improved Bass model approach," 2025, doi: https://doi.org/10.1038/s41598-025-09129-1.
[18] Sangiorgio et al., "Evaluating the effect of viral posts on social media engagement | Scientific Reports," 2025, Diakses: 10 Juni 2026. [Daring]. Tersedia pada: https://www.nature.com/articles/s41598-024-84960-6#citeas
[19] H. He dan E. A. Garcia, "Learning from Imbalanced Data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, hlm. 1263–1284, Sep 2009, doi: 10.1109/TKDE.2008.239.
[20] Ruidong Jin, Xin Liu, Tsuyoshi Murata, "Predicting popularity trend in social media networks with multi-layer temporal graph neural network," 2024, doi: https://doi.org/10.1007/s40747-024-01402-6.
[21] Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), *Profil Internet Indonesia 2026*. 2026.
[22] Wu, Y., Fu, C., He, Z. et al., "Multi-view temporal graph neural network for numerous miniature cascade popularity prediction," 2025, doi: https://doi.org/10.1007/s11227-025-07484-4.
[23] Ning Xu, Xiaowen Wang, Jing Liu, Lanjun Wang, Xuanya Li, Mengxiao Zhu, Yongdong Zhang, An-An Liu, "Model Can Be Subtle: Two Important Mechanisms for Social Media Popularity Prediction," 2025, doi: https://doi.org/10.1145/3705319.
[24] Jana Gross, "Full article: Influencer Marketing on Instagram: Empirical Research on Social Media Engagement with Sponsored Posts," 2022, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://www.tandfonline.com/doi/full/10.1080/15252019.2022.2123724
[25] I. C. Dang, C. M. C. Kwan, J. S. Jia, dan Y. Shi, "When Words Meet Visuals: How Content Composition Drives Social Media Engagement for Marketer-Generated Content," *J. Mark. Res.*, vol. 63, no. 1, hlm. 167–190, Feb 2026, doi: 10.1177/00222437251373042.
[26] Simon zollo, "Examining the relationship between content length and engagement with news outlets on multiple social media platforms - ScienceDirect," 2026, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://www.sciencedirect.com/science/article/pii/S0040162526001459
[27] Myoung-Jin Chae, "Real-time marketing messages and consumer engagement in social media - ScienceDirect," 2025, Diakses: 11 Juni 2026. [Daring]. Tersedia pada: https://www.sciencedirect.com/science/article/abs/pii/S014829632500089X
[28] Z. Ye, J. Yang, F. Meng, M. Li, dan Y. Zhan, "Integrating Temporal Interest Dynamics and Virality Factors for High-Precision Ranking in Big Data Recommendation," *Electronics*, vol. 14, no. 18, hlm. 3687, Jan 2025, doi: 10.3390/electronics14183687.
[29] J. Zhang *dkk.*, "THGFormer: Time-Aware Hypergraph Learning for Multimodal Social Media Popularity Prediction (Student Abstract)," *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 21, hlm. 23705–23706, Mar 2024, doi: 10.1609/aaai.v38i21.30534.
[30] D. H. Kim, O. Kuru, J. Zeng, dan S. Kim, "Fostering mask-wearing with virality metrics and social media literacy: evidence from the U.S. and Korea," *Front. Psychol.*, vol. 14, Mei 2023, doi: 10.3389/fpsyg.2023.1151061.
[31] N. E. Beshet *dkk.*, "Trends Sentiment Unveiled Through Deep Dive into Social Media Data," *Procedia Comput. Sci.*, vol. 275, hlm. 799–808, Jan 2026, doi: 10.1016/j.procs.2026.01.092.
[32] H. Peters, Y. Liu, F. Barbieri, R. A. Baten, S. C. Matz, dan M. W. Bos, "Context-aware prediction of active and passive user engagement: Evidence from a large online social platform," *J. Big Data*, vol. 11, no. 1, hlm. 110, Agu 2024, doi: 10.1186/s40537-024-00955-0.
[33] L. Rivadeneira dan I. Loor, "Evidential Reasoning Approach for Predicting Popularity of Instagram Posts," *IEEE Access*, vol. 12, hlm. 182603–182617, 2024, doi: 10.1109/ACCESS.2024.3510637.