

Forecasting New Follower Growth Using ARIMAX, LSTM, and Stacking Fusion

Yunhao Han*

School of Computer Science and Technology, Shandong University of Technology, Zibo
255000, China

*Corresponding Author

Abstract

Against the backdrop of intensifying competition in social media content, new follower growth better reflects the conversion from content exposure to sustained engagement than metrics such as views, likes, and comments. This study uses daily operational data from influencers on a social media platform as the research sample. It incorporates metrics such as views, likes, comments, and new follower growth, and further constructs the interaction rate and day-of-week variables. ARIMAX, LSTM, a static weighted fusion model, and a Stacking fusion model are constructed to predict next-day new follower growth. Model performance is evaluated using mean squared error, root mean squared error, and mean absolute error. The results indicate that while ARIMAX can capture some linear relationships, its overall predictive performance is weak; LSTM significantly outperforms ARIMAX; the static weighted fusion model does not yield additional improvements; and the Stacking fusion model performs best. The study demonstrates that fusion methods can more effectively integrate information from different models to improve the accuracy of new follower growth forecasting, providing valuable insights for optimizing influencer operations and identifying creators on the platform.

Keywords

ARIMAX; LSTM; Fusion Forecasting; Time Series; Stacking Fusion.

1. Introduction

On social media platforms, competition among influencers extends beyond content exposure to whether users are willing to continue following an account. Views, likes, and comments reflect the reach and engagement of content at a specific point in time, but these metrics primarily capture immediate feedback. In contrast, new follower growth more directly reflects the conversion from “seeing content” to “deciding to continue following,” making it a more informative indicator for assessing account growth, identifying creator potential, and screening candidates for brand collaborations.

However, new follower growth is not stable and does not follow a simple linear accumulation pattern. It is typically influenced by a combination of factors such as historical growth, engagement performance, content appeal, and day-of-week effects, and therefore often exhibits clear temporal dependence, fluctuation, and nonlinearity. For such problems, single linear models are easy to interpret but limited in their ability to capture complex dynamics, whereas deep learning models are better at nonlinear fitting but may suffer from insufficient stability when the sample is short.^[1-3] Against this background, combining statistical models with deep learning models and further examining whether ensemble methods can generate additional gains is worth investigating.^[4-7]

Based on the above considerations, this study uses daily operational data from influencers on a social media platform as the research sample. In addition to basic variables such as views,

likes, comments, and new follower growth, an interaction rate and day-of-week variables are introduced. ARIMAX, LSTM, a static weighted fusion model, and a Stacking fusion model are then constructed to forecast next-day new follower growth, and their performance is compared under short-sample conditions. This study aims to examine the applicability of fusion methods to this task and provide a reference for identifying influencer growth potential and optimizing platform operations.

2. Data Sources and Research Methods

2.1. Data Sources and Sample Description

This study uses user behavior data from a social media platform for analysis. The sample period spans from July 11, 2024, to July 20, 2024, and is used to forecast new follower growth on July 21, 2024. The raw data contain fields such as user ID, behavior type, influencer ID, and behavior timestamp. The behavior types include views, likes, comments, and follows.

During data preprocessing, we first standardized behavior codes, removed records with missing key fields, and converted behavior timestamps into a standard date format. We then filtered the data within the study period and aggregated them by influencer and date to obtain influencer-level daily metrics, including views, likes, comments, and new follower growth. To prevent missing dates from affecting subsequent modeling, we further completed all date-influencer combinations and assigned a value of 0 to missing behavior counts.

Because time series forecasting depends on the chronological order of observations, this study adopts a one-step rolling forecasting scheme with an expanding window. For any target date, the model is trained on all historical samples prior to that date and then generates a one-step-ahead forecast for new follower growth on that date. Backtesting starts on July 19, 2024; therefore, the test dates are July 19 and July 20, 2024. On this basis, all samples up to July 20, 2024, are further used to forecast new follower growth on July 21, 2024.

2.2. Variable Definition and Feature Engineering

In this study, next-day new follower growth is taken as the prediction target and denoted by y_t . Behavioral indicators such as views V_t , likes L_t , and comments C_t are selected as input features. Considering that differences in interaction quality may lead to significantly different growth outcomes even when accounts have similar exposure levels, this study further constructs an interaction rate indicator:

$$E_t = \frac{L_t + C_t}{V_t + \epsilon} \quad (1)$$

where E_t denotes the interaction rate, and ϵ is a small constant introduced to prevent the denominator from being zero.

In addition, considering that user activity may vary cyclically across the days of the week, this study introduces a weekend indicator variable to capture weekly temporal effects. The variable takes the value 1 if a given day falls on Saturday or Sunday, and 0 otherwise. Specifically, we construct the current-day weekend indicator W_t and the target-day weekend indicator W_{t+1} .

2.3. ARIMAX Model

The ARIMAX model is an extension of the ARIMA model that incorporates exogenous variables. It is used to capture the linear temporal dependence in new follower growth as well as the effects of external factors. Its general form is:

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t + \beta^\top z_t \quad (2)$$

where B is the lag operator, d is the order of differencing, $\phi(B)$ and $\theta(B)$ denote the autoregressive and moving-average polynomials, respectively, c is a constant term, ε_t is the random error term, $\beta^\top$ is the coefficient vector of the exogenous variables, and z_t is the exogenous variable vector.

In this study, current-day new follower growth, the interaction rate, and the target-day weekend indicator are selected as exogenous variables, i.e.,

$$z_t = [F_t, E_t, W_{t+1}] \quad (3)$$

The model order is selected by searching over candidate order combinations, with the minimum Akaike Information Criterion (AIC) used as the selection criterion.

2.4. LSTM Model

LSTM is effective in capturing nonlinear dynamic patterns in time series and is therefore well suited to the task of forecasting new follower growth. In this study, next-day new follower growth is also used as the prediction target. The input to the model is a sequence formed by the daily feature vectors from the previous three consecutive days, where

$$x_t = [E_t, F_t, W_t, W_{t+1}] \quad (4)$$

In terms of model architecture, we employ one LSTM layer, one Dropout layer, and one fully connected output layer, with 32 hidden units. The model outputs the predicted value of next-day new follower growth. During training, the input features and the target variable are normalized separately. The Adam optimizer is adopted, and mean squared error is used as the loss function.

2.5. Static Weighted Fusion Model

To examine the complementarity between the forecasting results of different models, this study constructs a static weighted fusion model that linearly combines the predictions generated by ARIMAX and LSTM, as follows:

$$\hat{y}_t = w\hat{y}_t^{(A)} + (1-w)\hat{y}_t^{(L)}, \quad 0 \leq w \leq 1 \quad (5)$$

where $\hat{y}_t$ denotes the final prediction of the static weighted fusion model at time t , $\hat{y}_t^{(A)}$ and $\hat{y}_t^{(L)}$ denote the predictions of the ARIMAX and LSTM models for the target variable y_t , respectively, and w is the weight assigned to the ARIMAX model, with $1-w$ being the weight assigned to the LSTM model.

In this study, the optimal weight is determined by grid search. Specifically, w is searched over the interval $[0,1]$ with a step size of 0.01, and the weight yielding the smallest validation error is selected as the optimal value.

2.6. Stacking Fusion Model

Stacking is a multi-level ensemble learning method whose core idea is to use the predictions of multiple base models as new input features and then build a second-level model through a

meta-learner, thereby improving final predictive performance. In this study, ARIMAX and LSTM are used as the first-level base models, and their predictions serve as the inputs to the second layer. The final prediction is produced by the meta-learner, as follows:

$$\hat{y}_t^{(S)} = g(\hat{y}_t^{(A)}, \hat{y}_t^{(L)}) \quad (6)$$

where $\hat{y}_t^{(S)}$ denotes the final prediction of the Stacking fusion model, and $g(\cdot)$ denotes the meta-learner. In this study, Ridge regression is used as the meta-learner, with the following optimization objective:

$$\min_{\beta} \sum_{i=1}^n (y_i - x_i^T \beta)^2 + \alpha \|\beta\|_2^2 \quad (7)$$

where β is the regression coefficient vector, x_i is the input feature vector of the i -th sample, y_i is the corresponding true value, and α is the regularization parameter. In this study, $\alpha = 1.0$ is used to improve estimation stability under small-sample conditions and reduce the risk of overfitting.

2.7. Evaluation Criteria

To compare the predictive performance of different models, this study uses mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) as evaluation metrics:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

where y_i denotes the true value, $\hat{y}_i$ denotes the predicted value, and n denotes the sample size. Smaller values of these metrics indicate better predictive performance.

3. Empirical Results and Analysis

3.1. Time-Series Characteristics of New Follower Growth

Figure 1 shows the aggregated daily trend of new follower counts across all influencers in the sample during the study period.

As shown in Figure 1, platform-wide new follower growth remained at a relatively high level during the sample period, although fluctuations across dates were still evident. New follower growth was relatively low on July 11, 2024, then increased rapidly and reached a relatively high level around July 18. After a brief decline on July 19, it rebounded on July 20. These results

suggest that new follower growth exhibits clear day-to-day fluctuations, which provides empirical support for the subsequent construction of time-series forecasting models.

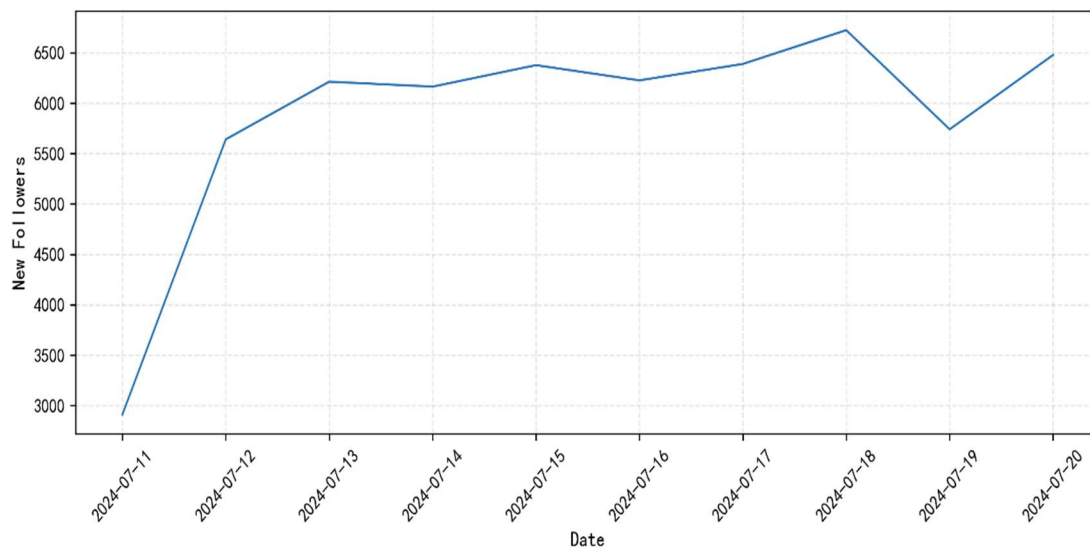


Figure 1. Time-Series Trend of New Follower Growth

3.2. Comparison of Forecasting Performance Across Models

Based on the backtesting results, a comparison of the predictive performance of the four models is shown in Table 1.

Table 1. Forecasting Performance of Different Models

Model	MSE	RMSE	MAE
ARIMAX	28783.885	169.658	99.135
LSTM	2434.848	49.344	35.329
Static weighted fusion	2434.848	49.344	35.329
Stacking fusion	1101.113	33.183	25.537

As shown in Table 1, the forecasting errors of the ARIMAX model are substantially higher than those of the other models, indicating that a model relying mainly on linear temporal structure and a limited number of exogenous variables is insufficient to fully capture the complex dynamics of new follower growth. By contrast, the LSTM model achieves markedly lower MSE, RMSE, and MAE values than ARIMAX, demonstrating a stronger ability to model nonlinear temporal relationships.

The static weighted fusion model yields exactly the same results as the LSTM model. This is because, under the current sample setting, the optimal weights obtained by grid search are 0.00 for ARIMAX and 1.00 for LSTM. This indicates that ARIMAX does not provide effective incremental information in the static linear combination, and the static weighted fusion model effectively reduces to the LSTM model.

In contrast, the Stacking fusion model achieves the lowest MSE, RMSE, and MAE values, at 1101.113, 33.183, and 25.537, respectively, indicating the best overall forecasting performance. This suggests that secondary modeling through a meta-learner can more effectively integrate the predictive information from ARIMAX and LSTM, thereby improving overall forecasting accuracy.

3.3. Comparison of Actual and Predicted Values for the Optimal Model

According to Table 1, the Stacking fusion model achieved the best performance on the validation set and was therefore selected as the final optimal model. A scatter plot comparing the actual and predicted values is shown in Figure 2.

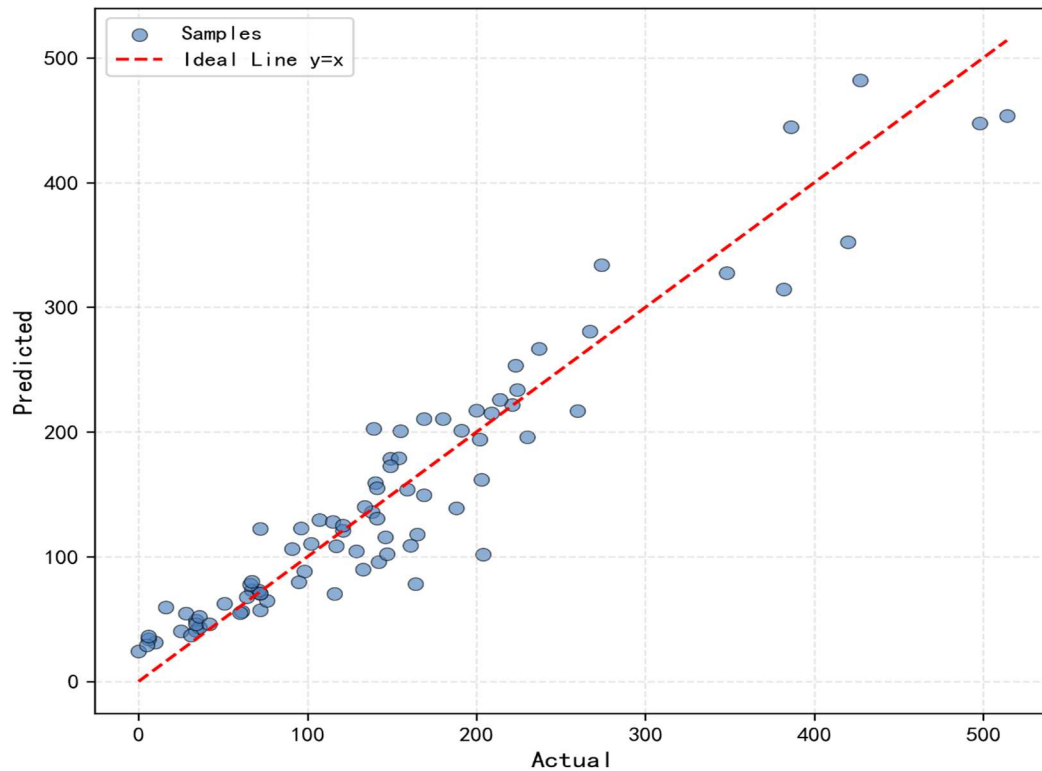


Figure 2. Scatter Plot of Actual and Predicted Values Under the Stacking Fusion Model

As shown in Figure 2, most sample points are distributed near the ideal line, indicating that the Stacking fusion model is generally able to fit the actual variation trend of new follower growth well. In particular, in the low-to-medium range, the predicted values are closely aligned with the actual values. However, in the higher-value range, some deviations still remain, suggesting that the model still has room for improvement in fitting extremely high-value samples. Overall, the pattern shown in Figure 2 is consistent with the results in Table 1, further validating the superiority of the Stacking fusion model in this study.

3.4. Analysis of Optimal Model Applications

After identifying the Stacking fusion model as the optimal model on the validation set, this study further used it to predict new follower growth for each influencer on July 21, 2024. The top five influencers ranked by predicted values were then selected, and the results are shown in Table 2.

Table 2. Top 5 Influencers Ranked by Predicted New Follower Growth

Rank	Influencer ID	Predicted New Follower Growth	Model Used
1	B21	450.388	Stacking fusion
2	B5	419.831	Stacking fusion
3	B60	357.512	Stacking fusion
4	B15	321.016	Stacking fusion
5	B13	262.573	Stacking fusion

As shown in Table 2, influencers B21, B5, and B60 exhibit relatively high potential for new follower growth during the forecast period. These results indicate that predictions based on the fusion model are useful not only for comparing forecasting errors, but also for supporting practical platform operations, such as identifying high-potential influencers, allocating platform resources, and screening potential business partners.

4. Conclusion

This study examines the daily operational data of multiple influencers on a social media platform and develops four models for forecasting new follower growth: ARIMAX, LSTM, a static weighted fusion model, and a Stacking fusion model. Their forecasting performance is then compared and analyzed. The results indicate that new follower growth exhibits clear temporal dependence, fluctuation, and nonlinearity. Although traditional linear models can reflect part of these variation patterns, they are unable to fully capture the complex dynamic process.

Specifically, the ARIMAX model generated the highest prediction error, indicating that its adaptability to complex growth mechanisms is limited within a linear modeling framework. The LSTM model significantly outperformed ARIMAX, suggesting that deep learning methods have a stronger advantage in uncovering nonlinear temporal relationships. The static weighted fusion model failed to achieve additional gains beyond the better-performing single model under the current sample conditions, indicating that a fixed linear combination of weights has limited utility in leveraging complementary information from different models. The Stacking fusion model achieved the best predictive performance, indicating that re-learning the outputs of base models through a meta-learner can, to some extent, more effectively integrate information from different models and improve predictive accuracy.

Further ranking analysis shows that the Stacking fusion model not only performs best in terms of error metrics, but also has practical value in identifying the growth potential of different accounts. Overall, fusion forecasting methods, especially the Stacking approach, demonstrate promising application potential in predicting new follower growth for social media influencers, and can provide useful references for creator operation optimization, platform resource allocation, and the screening of potential business partners.

It should be noted that the sample period in this study only covers July 11, 2024, to July 20, 2024. The overall time span is relatively short, and the model backtesting is mainly concentrated on the two target dates of July 19 and July 20, 2024. Therefore, the conclusions of this study should be regarded more as a preliminary validation of the feasibility of the method than as a sufficient generalization of the longer-term patterns of new follower growth. Future studies may further extend the time span, enrich the feature system, and conduct validation across multiple platforms and creator scenarios to enhance the robustness and generalizability of the findings.

Acknowledgments

The author would like to express sincere gratitude to the teachers and students who participated in the discussion and provided valuable suggestions for improvement.

References

- [1] Zou Y . Forecasting Apple Inc. Stock prices: A comparative analysis of ARIMA, LSTM, and ARIMA-LSTM models[J].Advances in Operation Research and Production Management,2025,4(1):66-74.DOI:10.54254/3029-0880/2025.23870.

- [2] Cibicong, Zhang Pinyi. Financial Time Series Forecasting Based on ARIMA-LSTM Model[J]. Statistics and Decision, 2022, 38(11): 145-149. DOI: 10.13546/j.cnki.tjyjc.2022.11.029
- [3] Wang Yingwei, Ma Shucai. Time Series Forecasting Based on ARIMA and LSTM Hybrid Model[J]. Computer Applications and Software, 2021, 38(02): 291-298
- [4] Alain N ,Danrong Z ,Yiqing G , et al. Improving wastewater discharge forecasting in Lake Tanganyika utilizing hybrid ARIMA-LSTM attention mechanism, East Africa[J].Journal of Hydroinformatics,2025,27(11):1745-1770.DOI:10.2166/HYDRO.2025.058.
- [5] Zhu Shaomin, Xia Hong, Lü Xinzhi, et al. Prediction of Main Pump Status in Nuclear Power Plants Based on ARIMA and LSTM Combined Model[J]. Nuclear Power Engineering, 2022, 43(02): 246-253. DOI: 10.13832/j.jnpe.2022.02.0246
- [6] Dongyan F ,Hai S ,Jun Y , et al. Well production forecasting based on ARIMA-LSTM model considering manual operations[J]. Energy,2021,220(publish):119708-. DOI:10.1016/J.ENERGY. 2020.119708.
- [7] Jing Nan, Hu Yi, Han Xishuang. Research on the Trend of COVID-19 Online Attention Based on ARIMA and LSTM[J]. Journal of Safety Science of China, 2020, 30(12): 37-42. DOI: 10.16265/j.cnki.issn1003-3033.2020.12.006