Research on the Prediction of Overall Monthly Active Users of WeChat Mini Programs Based on Linear Regression

Lidong Liu[a], Jiajia Hu[a], Yantao He[a*], Sijia Liu[b], Shuhua Li[a], Yongbin Luo[a], Rong Zhang[a], Ziyu Chen[a]

[a]Department of Computer Science, Guangdong University of Science and Technology, No. 2, Songshan Lake Section, 523083, China
[b]School of Mechanical Engineering, Hunan University of Science and Technology, Taoyuan Road, 411201, China

* Corresponding author: heyantao@gdust.edu.cn

Abstract. In recent years, the user base of WeChat mini programs has experienced a consistent growth. This study employs a linear regression approach to forecast the future monthly active user count for these mini programs. Throughout our research, we utilized two distinct methods to independently determine the slope, thereby mitigating potential discrepancies in results. Our analysis, based on the derived linear regression model, indicates a projected upward trend in the overall monthly active user count for future mini programs. Such insights offer robust decision support for both developers and marketers, fostering the swift advancement and widespread adoption of WeChat mini programs.

Keywords: WeChat Mini Programs market, linear regression, number of users

1 INTRODUCTION

Linear regression, a potent statistical analysis tool, seeks the optimal solution of an objective function by constructing a series of linear inequality or equality constraints. This method's strength lies in its conciseness and efficiency, enabling rapid identification of regularities within vast data sets and offering robust support for decision-making[1-2]. Within the realm of machine learning, linear regression is favored due to its unique optimization capability, finding extensive application in classification, regression, and clustering tasks. Compared to the neural network model, the linear regression structure primarily offers several benefits: it is relatively succinct, making it an effective choice for certain application scenarios that do not necessitate complex models; due to its simplicity, the linear regression model demands fewer computational resources for training and prediction[3]. When dealing with limited data, this model is less likely to encounter issues such as overfitting or significant fluctuations in results[4],

A. Haldorai et al. (eds.), *Proceedings of the 2024 3rd International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2024)*, Atlantis Highlights in Intelligent Systems 11,
https://doi.org/10.2991/978-94-6463-490-7_31

intelligence, WeChat Mini Programs have rapidly permeated all sectors of society due to their lightweight nature and ease of dissemination. As market competition intensifies, WeChat Mini Programs have emerged as a crucial arena for enterprises to acquire users. Through these programs, enterprises can target user groups with greater precision, offering personalized products and services, which enhances user satisfaction and loyalty. Simultaneously, WeChat Mini Programs serve as a key instrument for enterprise digital transformation, aiding enterprises in distinguishing themselves amidst intense market competition.

Numerous researchers have dedicated their efforts to the exploration of linear regression algorithms and the applications of WeChat mini programs. Nemati[5] proposed an optimization method for the economic operation and energy dispatching of microgrids by integrating genetic algorithms with mixed integer linear regression. This approach offers a robust optimization strategy for managing microgrid operations, thereby enhancing energy efficiency and reducing operational expenses. Li[6] examined the service functions and attributes of WeChat mini programs within academic social network applications. By examining user behavior and their requirements for these applications, the study highlighted the potential of WeChat mini programs in addressing user needs and improving user experience. Bakawu[7] utilized the theory of possibilities to construct regression models, thereby enhancing the management of uncertainty and ambiguity. By analyzing and modeling economic data from Nigeria, this study offers a forecast of future GDP fluctuations and assesses the efficacy and precision of the possibility linear regression model in this forecasting endeavor. Bemporad[8] introduces a piecewise linear regression and classification algorithm, which is applied to the learning and predictive control of hybrid systems. This algorithm enhances the model's accuracy and predictive capacity by segmenting data processing and employing distinct linear models for regression and classification within each segment. The study emphasizes the application of this algorithm in hybrid system learning and control, as well as its influence on system performance and stability. Hao[9] delved into the evolution of WeChat mini programs, emphasizing their convenience, cross-platform capabilities, and diverse application scenarios as drivers of their rapid growth. The article further explored the broad application of mini programs in e-commerce, social networking, tools, and other domains, offering insights into future trends. Tang[10] approached the adoption of government WeChat mini programs from a trust-risk perspective and innovation diffusion theory. The study integrated the impacts of trust and risk on the adoption of these programs, using innovation diffusion theory to elucidate their dissemination and adoption trajectory. These findings offer profound insights into the determinants influencing the adoption of government WeChat mini programs.

Linear regression serves as a pivotal tool in forecasting the user base of WeChat Mini Programs, this facilitates the precise prediction of future user numbers, thereby offering robust decision-making support for Mini Program developers and marketers. By leveraging these predictive outcomes, they can devise more efficacious operational strategies and resource allocation plans, enhance the user activity and retention rate of Mini Programs, and ultimately achieve sustainable development. This research employs the linear regression algorithm to forecast the future trajectory of the cumulative

monthly active user count for WeChat mini programs. The derived predictions offer valuable insights into prevailing market trends and pivotal industry advancements.

2 MODEL BUILDING

2.1 Theoretical model derivation

By extracting data from QuestMobile's panoramic ecological traffic database[11], the overall monthly active user scale (in billions) of mini programs from May 2022 to January 2024 was obtained, as shown in Figure 1. The quantity of data available is limited, rendering it unsuitable for prediction via a neural network model.

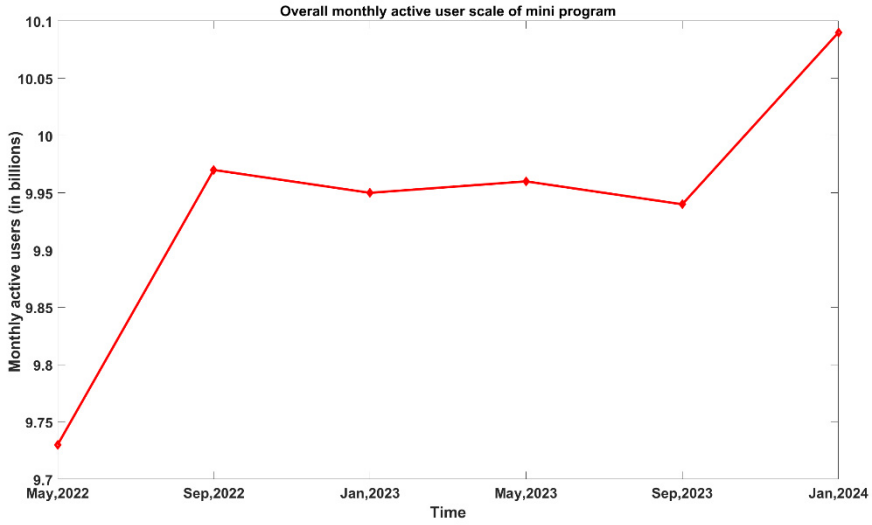

Fig. 1. Overall monthly active user scale of mini programs.

Assuming time is $t_i, i = 1 \cdots n$, and y_i is the overall monthly active user scale of the WeChat mini programs, n independent observations and analyses were conducted based on the existing data in the Figure 1, and the equation is as follows.

$$y_i = \beta_0 + \beta_i t_i + \varepsilon_i \tag{1}$$

In order to represent the Equation (1) using a matrix, the matrix is established as follows, and ε_i is set as an uncorrelated random variable.

$$Y = t\beta + \varepsilon \tag{2}$$

where β is the regression coefficient. Then the Equation (3) is used to solve the regression coefficient by least square method.

$$\beta = \left(t^T t\right)^{-1} t^T Y \tag{3}$$

where $\left(t^T t\right)^{-1} t^T$ represents the pseudo-inverse of t, and the regression coefficient $\beta = 0.761$ is obtained, which can be used to judge that t has a strong influence on y.

2.2 Establishment of calculation model

Aiming at the slope $u(b)$ of the linear regression equation, two methods are used , and to judge whether the results are consistent or not, which can be used as mutual verification to avoid the chance of the results. Equation 4 and Equation 5 show these two solving methods,

$$u(b) = \frac{\sigma}{\sqrt{\sum\left(x - \bar{x}\right)^2}} \tag{4}$$

where σ is the standard deviation,

$$u(b) = \sqrt{\frac{S_{yy} - bS_{xy}}{(n-2)S_{xx}}} = \sqrt{\frac{\frac{S_{yy}}{S_{xx}} - b^2}{(n-2)}} \tag{5}$$

in Equation 5, S_{yy} and S_{xx} represent the variance of the prediction function and the variance of time, respectively. S_{xy} is the covariance of x and y.

By substituting the data, the $u(b)$ of the above two methods is 0.121, which proves that the slope data is reliable.

Then the intercept $u(a) = -235.246$ is obtained by using Equation 6.

$$u(a) = u(b)\sqrt{\frac{\sum x_i^2}{n}} \tag{6}$$

$$y = -235.246 + 0.121 * t \tag{7}$$

Based on the above derivation, a linear regression model for the overall monthly active user scale of future WeChat mini programs was ultimately solved in Equation 7. Figure 2 is based on this model to predict the overall number of monthly active users of mini program in the future. As can be seen from the picture and data, the number of users as a whole shows an upward trend.

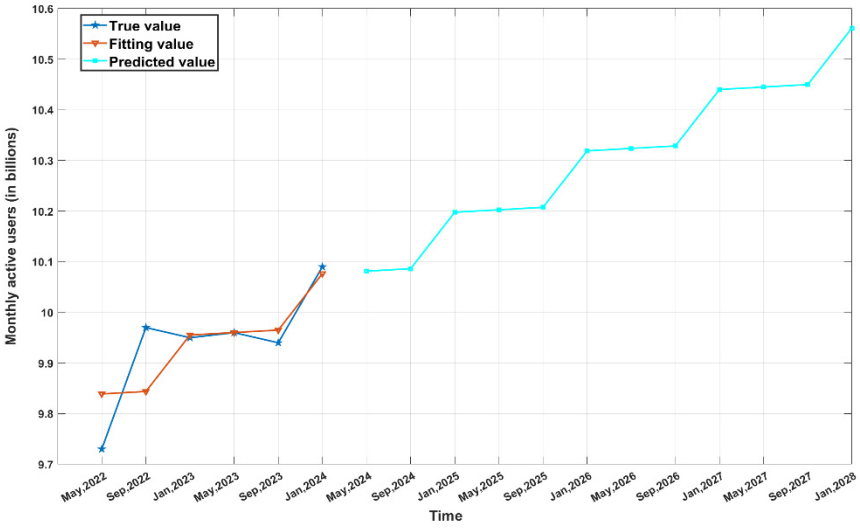

Fig. 2. Prediction of overall monthly active users for WeChat mini programs.

3 CONCLUSION

This study presents a linear regression model designed to forecast the projected monthly active users of WeChat Mini Programs in future periods. The findings suggest that there will be an upward trajectory in the number of WeChat Mini Program users. Comprehending this growth trend is crucial for businesses as it aids in assessing their market position and brand value, thereby facilitating more informed decision-making regarding collaboration opportunities. Furthermore, such insights can steer product development and function enhancement to align with user requirements and expectations, optimize resource allocation efficiency, and ensure enterprises retain their competitive edge in a dynamic market landscape.

ACKNOWLEDGEMENTS

This research was funded by the Quality Engineering Project of Guangdong University of Science and Technology under Grant GKY-2019CQYJ-9 and Grant GKY-2023KYYBK-2.

REFERENCES

1. Shaikh T A, Rasool T, Lone F R. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming[J]. Computers and Electronics in Agriculture, 2022, 198: 107119.

2. Alloghani M A. Harnessing AI for Sustainability: Applied AI and Machine Learning Algorithms for Air Quality Prediction[M]//Artificial Intelligence and Sustainability. Cham: Springer Nature Switzerland, 2023: 1-32.
3. Shams S R, Jahani A, Kalantary S, et al. The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO2 concentration[J]. Urban Climate, 2021, 37: 100837.
4. Bejani M M, Ghatee M. A systematic review on overfitting control in shallow and deep neural networks[J]. Artificial Intelligence Review, 2021, 54(8): 6391-6438.
5. Nemati M, Braun M, Tenbohlen S. Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear regression[J]. Applied energy, 2018, 210: 944-963.
6. Li J, Zhang G, Lin J, et al. Research on WeChat Mini Program APP Service of Academic Social Network Application[C]//2020 15th International Conference on Computer Science & Education (ICCSE). IEEE, 2020: 227-231.
7. Bakawu M A, Buba A T, Malik A. Forecasting the Real GDP in Nigeria Using Possibilistic Linear Regression Model[J]. Benin Journal of Statistics, 2023, 6: 59-69.
8. Bemporad A. A piecewise linear regression and classification algorithm with application to learning and model predictive control of hybrid systems[J]. IEEE Transactions on Automatic Control, 2022, 68(6):3194-3209.
9. Hao L, Wan F, Ma N, et al. Analysis of the development of WeChat mini program[C]//Journal of Physics: Conference Series. IOP Publishing, 2018, 1087(6): 062040.
10. Tang Z, Zhou Z, Xu F, et al. Apps within apps: predicting government WeChat mini-program adoption from trust–risk perspective and innovation diffusion theory[J]. Information Technology & People, 2022, 35(3): 1170-1190.
11. Quest Mobile. (2024). Mini Program Market report. https://www.questmobile.com.cn/products/plus.