

Factors Influencing Agricultural Economic Development in Guangdong Province

Jikun Zheng

Business School, Lingnan Normal University, Zhanjiang, Guangdong, China

E-mail: 13326519596@163.com

Received: August 21, 2026 Accepted: October 8, 2026 Published: October 14, 2026

doi:10.5296/jas.v15i1.24101

URL: <https://doi.org/10.5296/jas.v15i1.24101>

Abstract

As China's largest economic province and a major agricultural region, Guangdong has long ranked among the top in core indicators of agriculture, forestry, animal husbandry, and fishery. In 2024, the total output value of agriculture, forestry, animal husbandry, and fishery in Guangdong reached 971.861 billion yuan, with the total agricultural output value at 470.631 billion yuan, a year-on-year increase of 6.21%. Since the reform and opening up, leveraging its subtropical climate endowment, geographical proximity to Hong Kong and Macao, and continuous rural reform, Guangdong has gradually established a sound modern agricultural industrial system and steadily enhanced its comprehensive agricultural production capacity. To further promote the high-quality development of Guangdong's agricultural economy and identify the key influencing factors, this paper selects 17 core indicators across five dimensions: agricultural production inputs, agricultural water conservancy status, agricultural production scale, agricultural output, demand and income, and agricultural technology and development, covering the period from 2015 to 2024. Two machine learning (ML) models—support vector machines and neural networks—are constructed, and SHAP is employed to rank indicator importance and analyze explanatory power. The study ultimately identifies five core indicators that drive Guangdong's agricultural economic development. The findings provide a scientific basis for the precise formulation of agricultural economic policies in Guangdong. By clarifying the core influencing indicators, policymakers can rationally allocate regional advantageous resources and introduce targeted measures, such as optimizing the agricultural support and farmer subsidy system to improve input efficiency; advancing rural water supply facilities and intelligent water conservancy systems, and establishing a provincial-level digital management platform for water conservancy; and fostering leading agricultural enterprises, supporting large-scale production bases, and improving the level of large-scale agricultural operation. These measures can effectively enhance the precision of agricultural economic policy support and promote the high-quality development of Guangdong's agricultural economy.

Keywords: agricultural economic development, support vector machine, neural network

1. Introduction

As a core area of agricultural economic development in China and an important “vegetable basket” and “rice bag” security base for the country, Guangdong Province has established a full-chain modern agricultural industry system covering grain planting, specialty fruits and vegetables, southern medicinal materials, and deep processing of agricultural products, relying on the advantages of its subtropical monsoon climate and vast arable and mountainous resources. In recent years, accelerating the “cultivation of distinctive agricultural industrial clusters” to improving county revitalization has been highlighted in the 2025 Central Document No. 1 (The Central People's Government of the State Council, 2025). The Guangdong Province has implemented “Hundred Counties, Thousand Towns, Ten Thousand Villages High Quality Development Project”. The quality and efficiency of agriculture in the province have continued to improve. The total agricultural output value increased from 249.02 billion yuan (2015) to 470.631 billion yuan (2024). The average annual growth rate has reached to 7.33%, ranking among the top in the country for many years. The output of major agricultural products in Guangdong Province has increase by 31.4% from 67.0549 million tons (2015) to 88.1178 million tons (2024). Among them, the vegetable output in 2024 was 42.4501 million tons, and the output of horticultural fruits was 20.5619 million tons, both ranking among the top in the country, cultivating a large number of highly recognizable Lingnan characteristic agricultural product brands, such as lychee, longan, and pineapple. The ten large-scale agricultural product processing enterprises in Guangdong Province achieved an operating revenue of 1.63 trillion yuan in 2024. In the meantime the total number of leading agricultural enterprises in the province exceeding 5500, with an annual operating revenue of nearly one trillion yuan. The agricultural industry has become an important cornerstone supporting the coordinated development of Guangdong's regional economy, contributing irreplaceable strength to stabilizing the overall economic situation of the province and facilitating the growth of county-level farmers' earnings. However, with the continuous expansion of the agricultural scale in Guangdong Province, its development still faces some practical pain points, such as the overall small scale of characteristic industries, insufficient core competitiveness of the seed industry, difficulties in promoting agricultural machinery in hilly and mountainous areas, scattered agricultural management entities, and homogeneous product competition. To provide a scientific empirical basis for the precise formulation of subsequent agricultural support policies, it is urgent to make a profound analysis of the influencing factors of agricultural economic development in Guangdong Province around various resources, prioritizing the resolution of core shortcomings that restrict long-term agricultural development.

Currently, agricultural economic research mainly focuses regarding the factors that drive and shape agricultural economic growth and the application of ML models in agricultural economic research.

In the study of factors influencing agricultural economic development, Ruili Zheng et al. (2025) measured and analyzed the level of sustainable agricultural development (ASDL) in

Shandong Province, and studied its key driving factors based on the XGBoost SHAP framework. It is concluded that the proportion of the primary industry, rural income, green land area, and the scale of fiscal revenue have a positive impact on ASDL, while mechanical power intensity and per capita GDP have a negative impact on ASDL (Zheng, 2025). Chenghan Guo, Rong Zhang, and Yuntao Zou (2023) measured the efficiency and changes of agricultural circular economy in 31 provinces and cities in China from 2017 to 2020 using Data Envelopment Analysis (DEA) and the Malmquist index model, and validated the correlation analysis using Tobit regression in the context of rural revitalization policies. The research results indicate that the rural revitalization policy has great influence on improving the efficiency of the agricultural circular economy, and government financial support significantly promotes the efficiency of the agricultural circular economy (Guo, Zhang, & Zou, 2023). Xiang Jingwei, Han Peng, and Chen Wanxu (2023) adopted the DEA method to explore key factors influencing cultivated land spatial morphology and the coordinated development efficiency of the agricultural economy (CECA) in 16 counties in western Hubei. According to the analysis results, the population density in southwestern Hubei presents the most significant effect on CECA, while the gross domestic product (GDP) in the northwestern region has the greatest impact on CECA (Xiang, Han, & Chen, 2023). Xiao Mei Wang (2011) believes that in terms of agricultural production factors, land, labor, and capital can be used as indicators, and a linear regression model can be used to analyze agricultural economy and its influencing factors (Wang, 2011). Zhang Mihong (2022) analyzed the influencing factors from the perspective of the interaction and cooperation between the “the Belt and Road” initiative (B&R) and the supply-side structural reform of agricultural economy, and proposed measures to adjust agricultural structure in line with development concepts, promote circular development in line with ecological concepts, integrate innovation mechanisms to enrich industrial structure, and integrate market demand to promote overseas agricultural economic cooperation (Zhang, 2022).

In agricultural economic research, Su Li, Changjun Yang, and Kexin Li (2026) used panel data from 1410 counties in China from 2014 to 2023, employing the entropy weight method and dual ML model to analyze the impact of traditional rural finance (TRF) on agricultural economic resilience (AER) and its potential mechanisms. The conclusion of the reaserch emphasize that TRF plays an important role in improving AER, while agricultural technology innovation (ATI) and agricultural socialized services (ASS) play a mediating role (Li, Yang, & Li, 2026). Jing Li, Zhen Han, and Hui Zhou (2026) used panel data from heavily polluting listed companies in China’s A-share market from 2011 to 2023, employing double machine learning methods to study the impact of digital green collaborative transformation (DGCT) of heavy polluting enterprises on AER. The reaserch suggests that the DGCT of heavy polluting enterprises has a great impact on improving AER (Li, Han, & Zhou, 2026). Bo Xu, Chuanjian Yi, and Kejun Lin (2026) used dual machine learning and spatial Durbin difference models to study the dynamic impact, potential mechanisms, heterogeneity, and spatial spillover effects of new urbanization on AER. The results show that the new urbanization policy (NUP) exerts significant spatial spillover effects on AER, and the improvement of AER is largely attributed to the optimization of industrial structure, progress of agricultural technology, and buildup of agricultural human capital (Xu, Yi, & Lin, 2026).

Based on panel data from 31 provinces in China from 2011 to 2022, Chun Yang, Wangping Liu, and Jiahao Zhou (2024) used machine learning methods such as random forest, GBRT, SHAP value methods, and ALE graphs to study the influencing factors of AER. The results show that among all factors, government investment, urbanization level, and planting structure have the most significant impact on AER (Yang, Liu, & Zhou, 2024). Deng Chuyao (2022) believes that based on the development and consolidation of the agricultural sector, the application of inclusive finance can effectively analyze and judge the credit needs of poor clients through artificial intelligence (AI) system platforms, identify and screen different credit data of clients, and ultimately screen out clients with superior qualifications for credit services, thereby effectively improving the overall economic income of clients with agricultural credit needs (Deng, 2022).

Due to ML models have been widely applied in reaserch of agricultural economic, this paper simulate the influence of 17 indicators on the agricultural economic development of Guangdong Province through construction of support vector machines and neural networks. Then, this paper ranks the importance of 17 indicators using the SHAP tool. It also analyzes their explanatory power and identifies the five most important factors influencing the agricultural economic development of Guangdong Province. This study can provide scientific policy references for policymakers to identify the core driving factors of agricultural economic development in Guangdong Province, optimize the spatial allocation and industrial layout of agricultural resources such as land, capital, technology, and labor within the province, and provide targeted decision-making references and practical suggestions for building a modern agricultural industrial system in Guangdong Province, ensuring the supply security of food and important agricultural products, and achieving agricultural progress characterized by superior quality and sustainability.

2. Construction of ML Model

In order to construct ML models and conduct feature importance analysis using SHAP tools, this paper first selects appropriate indicators as the dependent and feature variables of ML models. Then, this paper divides the data into training and testing sets, and performs standardized preprocessing on them. In order to evaluate the predictive ability of ML models, this paper needs to determine the corresponding model evaluation indicators. After completing the data collection and processing in the early stage and determining the evaluation indicators of ML models, this paper constructs two types of ML models and adjusts the parameters of the models. Finally, this paper uses SHAP tool to analyze and rank the importance of feature indicators of the ML models with good predictive ability, in order to select the top 5 indicators with the greatest impact on the development of agricultural economy in Guangdong Province.

2.1 Construction of Model Indicators System

This paper uses the total agricultural output value of Guangdong Province to represent the level of agricultural economic development in Guangdong Province and takes it as the dependent variable of the model. In selecting the model feature variables, this paper utilizes the “Guangdong Statistical Yearbook” to select 17 indicators representing five aspects of

agricultural production input, agricultural water conservancy status, agricultural production scale, agricultural output, demand and income, and agricultural technology and development from 2015 to 2024 as factors influencing the agricultural economic development of Guangdong Province (Table 1).

Table 1. Model Indicators System

Category	Indicator
Agricultural production input	Total agricultural output value (RMB hundred million) TAOV
	Total power of agricultural machinery (10,000 kW) TPAM
	Fertilizer application rate (10,000 tons) FAR
	Usage of agricultural plastic film (10,000 tons) UAPF
	Agricultural diesel usage (10,000 tons) ADU
	Employment in the primary industry (10,000 people) EPI
Agricultural water conservancy situation	Accumulated area of soil erosion control (1000 hectares) AASEC
	Irrigation area of cultivated land (1000 hectares) IACL
	Flood control area (1000 hectares) FCA
	Protection area of embankment (1000 hectares) PAE
Agricultural production scale	Total sowing area of crops (10,000 mu) TSAC
	Machine cultivated area (1000 hectares) MCA
	Machine broadcasting area (1000 hectares) MBA
Agricultural demand, output, and income	Main agricultural product output (10,000 tons) MAPO
	Per capita disposable income of rural resident households (RMB) PCDIRRH
	Per capita consumption of major agricultural products (kg) PCCMAP
Agricultural technology and development	Rural Professional and Technical Association (unit) RP&TA
	Provincial level scientific and technological award achievements in agriculture (item) S&TAAA

2.2 Sample Division and Data Preprocessing

The construction of ML models requires dividing the overall sample into two parts: the

training set and the test set. The training set data is mainly used to construct the ML model, while the test set data is primarily used to test the predictive ability of the ML model based on the training set on new samples. In terms of the division of the overall sample, this paper uses the year of the data as the basis for division, with the random seed fixed at 50 via Python's `random_state` function, randomly shuffling the order of the data, and then dividing 60% of the shuffled population sample into the training set, with the remaining 40% divided into the test set.

This paper standardizes all data to obtain values within the range of [0,1], which helps eliminate the influence of different dimensions between the data and accelerates the convergence of the ML model during the training process. It is applicable to algorithms sensitive to the distance between data, such as neural networks. Specifically, as shown in the following Formula 1.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

In Formula 1, X' represents the standardized data; X denotes the original data; X_{max} signifies the maximum value of the original data; X_{min} stands for the minimum value of the original data.

2.3 Selection of Model Evaluation Indicators

This paper selects RMSE as the metric to measure the model's predictive ability for the target variable in the test set, as shown in Formula 2 below:

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

In Formula 2, N represents the sample size of the test set; y_i represents the true value of the target variable from the test set; and $\hat{y}_i$ denotes the predicted value of the model for the target variable of the test set. Compared with the MSE index, the RMSE index does not produce the problem of inconsistent dimensional units due to the use of the square root symbol. The calculation results can directly reflect the predictive accuracy of the training model, without amplifying the difference between abnormal predicted values and true values due to the presence of squared terms, as shown by the MSE metric. The smaller the RMSE value, the closer it is to 0, indicating better predictive performance of the training model on the test set.

2.4 Overview of SVM

SVM is a supervised learning model that can be used for classification and regression analysis of datasets. Its core idea is to find a hyperplane that maximizes the separation of data points from different categories. In regression problems, SVM finds the optimal regression

function by minimizing the sum of squared prediction errors. This feature makes SVM widely applicable in predicting continuous values and solving regression problems. In order to handle different types of datasets, there are currently three main categories of SVM: linearly separable SVM, linear SVM, and kernel function SVM. For non-linear separable datasets, SVM maps the data to a higher dimensional space using kernel functions, making the data linearly separable in that high-dimensional space. Common kernel functions include radial basis function (RBF), polynomial kernel, and sigmoid kernel. By selecting appropriate kernel functions and parameters, SVM can handle various complex nonlinear problems.

2.5 Overview of NN

A NN is a computational model that simulates the biological nervous system, which achieves intelligent processing, learning, and prediction of data by simulating the complex interactions and information processing mechanisms between neurons. Compared with traditional machine learning models, NN have strong adaptability and robustness, and can handle complex and nonlinear problems. The basic unit of NN is neurons, which simulate the working mode of biological neurons. A neuron receives multiple input signals and processes them through weighted summation, activation functions, and other methods to generate an output signal. These neurons are organized together in a certain hierarchical structure, forming a complete neural network. The training process of NN is achieved by adjusting the connection weights between neurons. This process is usually accomplished using the backpropagation algorithm. During the training process, the NN will continuously adjust its weights based on the error between the output results and the actual results, until it reaches the preset error threshold or training frequency.

2.6 Overview of SHAP

The SHAP is a mathematical tool that can explain the prediction results of machine learning models. This method is mainly based on the Shapley value in game theory, which measures the importance of the impact of various input features on the predicted values of the model. The Shapley value aims to fairly distribute the benefits generated by cooperation, with its core being to consider the unique contributions of each participant. Specifically, this method first constructs a game model, treating each feature as an independent participant in the game. Next, evaluate the expected benefits that each feature can achieve when combined with other features, which are determined based on the values of other features and their impact on the prediction results. Finally, the Shapley value of each feature is calculated, which intuitively demonstrates the role of that feature in model prediction.

Suppose there is a machine learning model $f(x)$ with input $x = [x_1, x_2, \dots, x_n]$ and output $f(x)$. To explain the prediction result of $f(x)$ under a certain input x , it is necessary to calculate the Shapley value $\varphi_i(x)$ of each feature x_i , as shown in Formula 3 below:

$$\varphi_i(x) = \sum_{C \subseteq T} \frac{(n-|C|)!|C|!}{n!} [f(x_C \cup x_{-i}) - f(x_C)] \quad (3)$$

In Formula 3, T represents the set of all features, x_C denotes the input that only includes the features in set C , x_{-i} stands for the input features other than x_i . All possible feature subsets C are included in Formula 3. The contribution of adding the input feature x_i to these subsets can be calculated through Formula 3. The basic principle of SHAP method is shown in Figure 1.

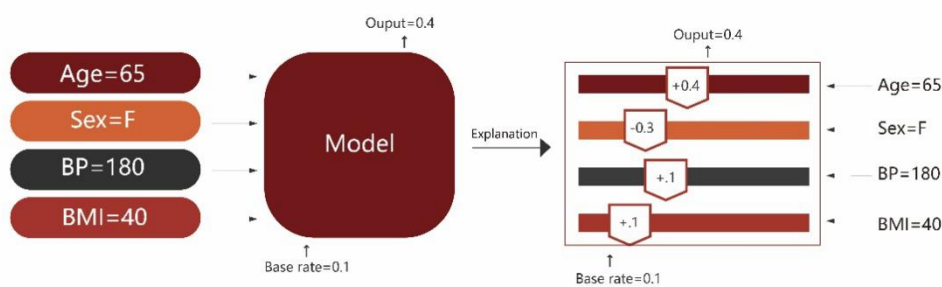


Figure 1. Basic Schematic Diagram of SHAP

Figure 1 shows the split results of the contribution or explanatory power of a single sample with different feature inputs to the model results under the SHAP method. In the left half of Figure 1, the base value of the model is 0.1, and there are four input features: Age, Sex, BP, and BMI. The output of the model is 0.4. The core concept of the SHAP method is to visualize how the four input features of the model drive the model from a base value of 0.1 to an output value of 0.4, while breaking down the contribution of different features. In the right half of Figure 1, the SHAP method calculates the Shapley values of different features and displays the contribution or explanatory power of each feature in the form of a bar chart. The explanatory power of BMI, BP, Sex and Age are 0.1, 0.1, -0.3 and 0.4 respectively. The combined effect of these indicators pushes the base value of the model from the original 0.1 to 0.4. From the perspective of the absolute value of Shapely values, the relative importance of the above four features can be ranked. The feature with the highest explanatory power is Age, followed by Sex, and the indicators BP and BMI are tied last.

3. Analysis of Explanatory Power of Influencing Factor Indicators

3.1 Results of SVM Model Parameter Adjustment

This article mainly selects three types of kernel models, namely RBF, Poly, and Linear, for construction, and compares the predictive capability of these three models using the RMSE metric, based on which the best SVM model is chosen. In terms of model parameter settings, this article adopts a grid search parameter traversal method to debug the “C”, “epsilon”, and “gamma” parameters of the SVM model. Specifically, the “C” parameter is set to take 7 values from the list [0.001, 0.01, 0.1, 1, 10, 100, 1000]; the “gamma” parameter takes 7

values from the list [0.001, 0.01, 0.1, 1, 10, 100, 1000]; the “epsilon” takes 20 values, evenly spaced from 0.01 to 1. After parameter traversal, the results show that the best parameter combination for the RBF kernel model is “C” = 100, “epsilon” = 0.01, and “gamma” = 0.001, with an RMSE value of 0.0549; the best parameter combination for the Poly kernel model is “C” = 0.001, “epsilon” = 0.01, and “gamma” = 1000, with an RMSE value of 0.1118; and the best parameter combination for the Linear kernel model is “C” = 1, “epsilon” = 0.01, and “gamma” = 0.001, with an RMSE value of 0.0552. Therefore, the best SVM model is the RBF kernel model. The true values of the target variable in the test set and the predicted values of the target variable by the RBF kernel model in the test set are shown in Table 2. This article will analyze the explanatory power of the feature variables based on the RBF kernel model.

Table 2. RBF Kernel Model Prediction Results

True value of the target variables	Predicted value of the target variables
4308.23	4449.18
4430.96	4612.78
3951.14	3945.91
2889.97	2810.25

3.2 Results of NN Model Parameter Adjustment

There are two types of neural network models: single-hidden-layer and two-hidden-layer, being constructed in this paper. This paper uses a parameter traversal method to traverse all integers between the number of hidden layer neurons i from 1 to 99. Among them, for the two-hidden-layer model, this article assumes that the number of neurons i in the two hidden layers before and after it is equal. This article sets the random seed “random_date” to 42 for both types of neural network models, and the maximum iteration count “max_iter” to 10000. This helps ensure that the trained models can fully converge, thus avoiding underestimating model performance due to insufficient iterations. Other parameters are set to the default values in Python as mentioned in this article. The single hidden layer model is designated as Model 1, and the two-hidden-layer model is designated as Model 2. The RMSE values of the two models under different numbers of hidden layer neurons are shown in Figures 2 and 3.

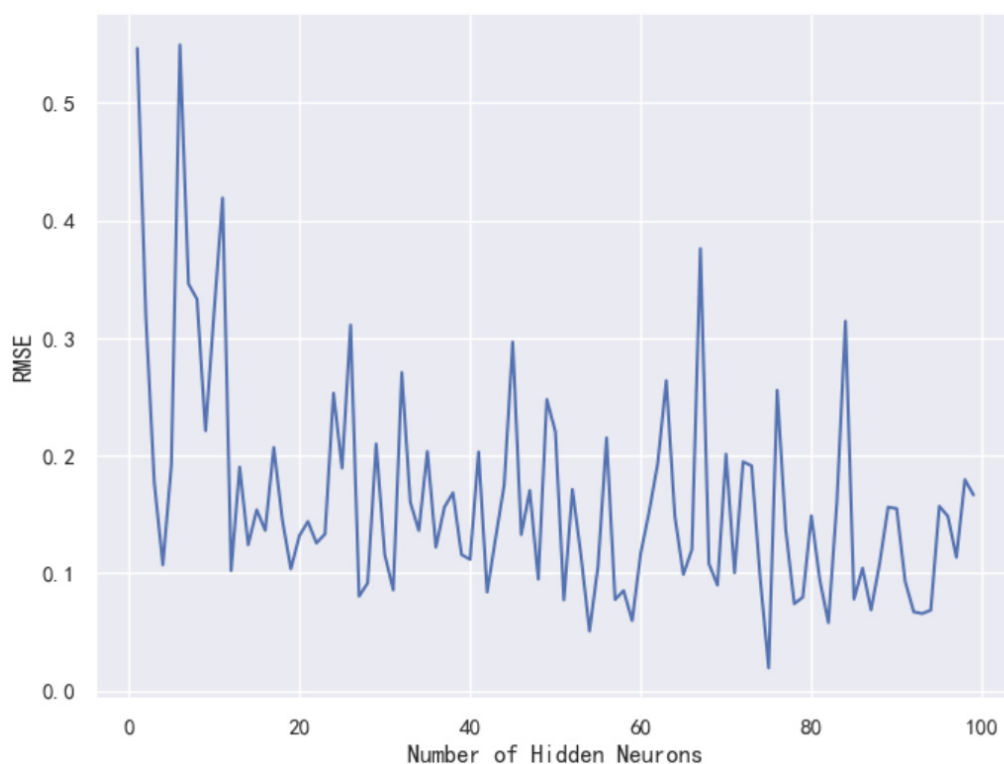


Figure 2. Line Graph of the Number of Neurons and RMSE in Model 1

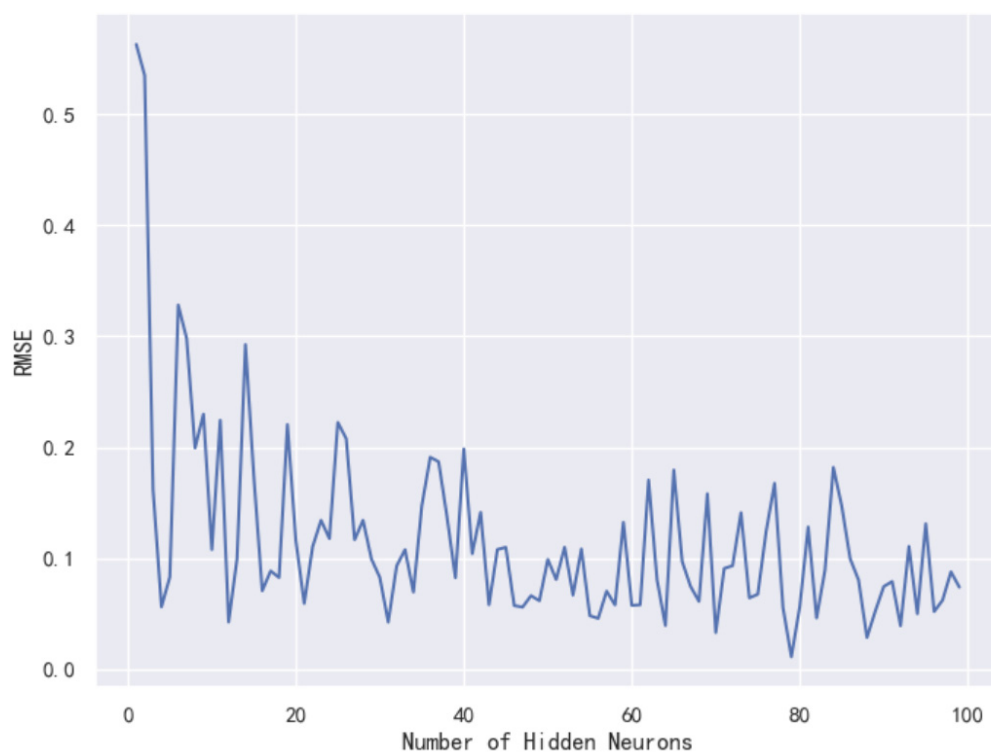


Figure 3. Line Graph of the Number of Neurons and RMSE in Model 2

According to the above results, it can be found that when the number of neurons i equals to 75, the single hidden layer model performs the best on the RMSE (0.0196). Moreover, when

the number of neurons i equals to 79, the two-hidden-layer model performs the best on the RMSE (0.0113). As the number of neurons i equaling to the optimal values, the predicted values and true values of the target variable in the test set for both models are shown in Table 3.

Table 3. NN Models Prediction Results

True value of the target variables	Predicted value of Model 1	Predicted value of Model 2
4308.23	4234.05	4355.32
4430.96	4395.18	4444.42
3951.14	3925.96	3960.12
2889.97	2900.15	2893.23

Due to the smaller RMSE of the two-hidden-layer model, it has better predictive performance for the target variables in the test set. Therefore, this paper will analyze the explanatory power of the feature variables based on the two-hidden-layer model.

4. Analysis on the Importance of Indicators for Influencing Factors

The SHAP tool is utilized in this research to analyze the importance of indicators, focusing on the RBF kernel model and the two-hidden-layer model, which exhibit good predictive capability. The SHAP tool estimates the Shapley value corresponding to every feature of each sample to represent its contribution to the model's prediction results. When the Shapley value is positive, the feature promotes the growth of the predicted outcome; by contrast, a negative Shapley value implies the feature lowers the predicted value. The Shapley values calculated directly using the SHAP tool are presented in matrix form, where each row represents a sample, and each column represents a feature. Since the study examines the contribution of each feature to the prediction results, this article will first take the absolute value of the Shapley values of each feature across all samples and all sample dimensions of the Shapley matrix, and then calculate the average. This article uses SHAP bar charts to display the level of explanatory power of each feature on the model's prediction results.

4.1 Ranking of the importance of RBF Kernel Model's Indicators

The analysis and ranking of the explanatory power of the RBF kernel model for the influencing factors of agricultural economic development in Guangdong Province are illustrated in Figure 4.

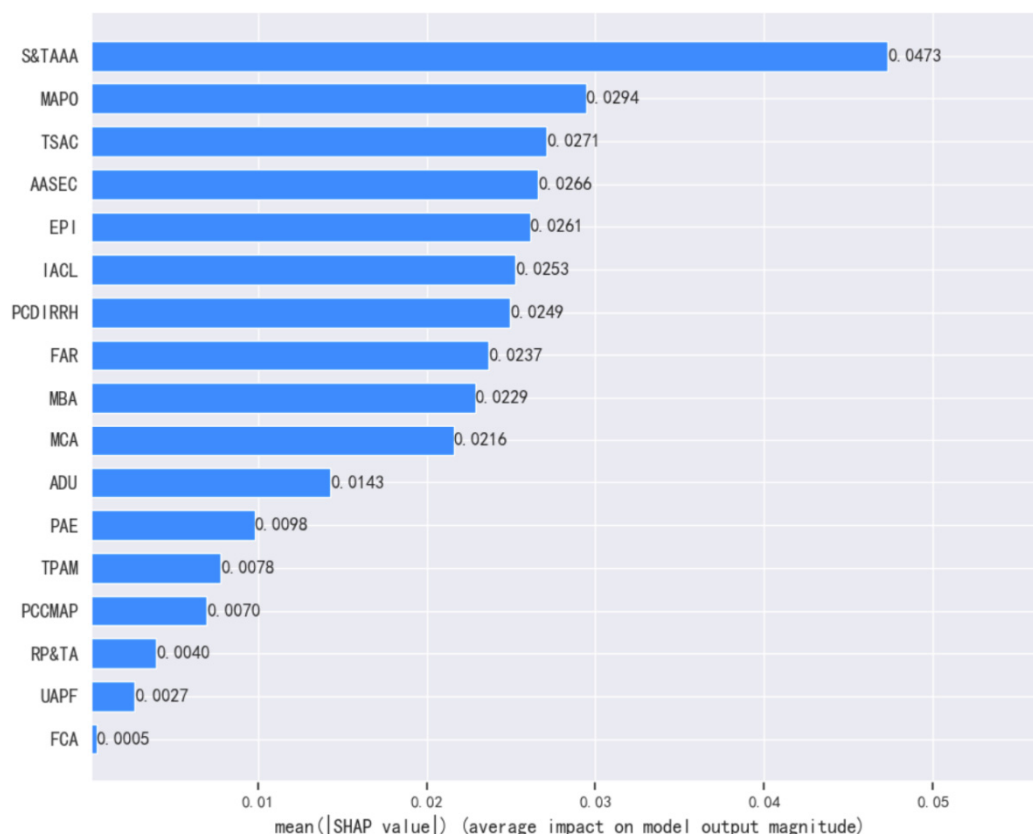


Figure 4. Bar Chart of SHAP for RBF Kernel Model

According to Figure 4, the top five factors that have the greatest impact on agricultural economic development in Guangdong Province in the RBF kernel model are: provincial agricultural science and technology award achievements, major agricultural product output, total crop sowing area, cumulative area of soil erosion control, and employment in the primary industry. The absolute average SHAP values of the above five factors are 0.0473, 0.0294, 0.0271, 0.0266 and 0.0261 respectively. These five indicators belong to the categories of agricultural technology and development, agricultural demand, output and income, agricultural production scale, agricultural water conservancy status, and agricultural production inputs.

4.2 Ranking of the importance of Two-Hidden-Layer Model's Indicators

The analysis and ranking of the explanatory power of indicators for the two-hidden-layer model are shown in Figure 5.

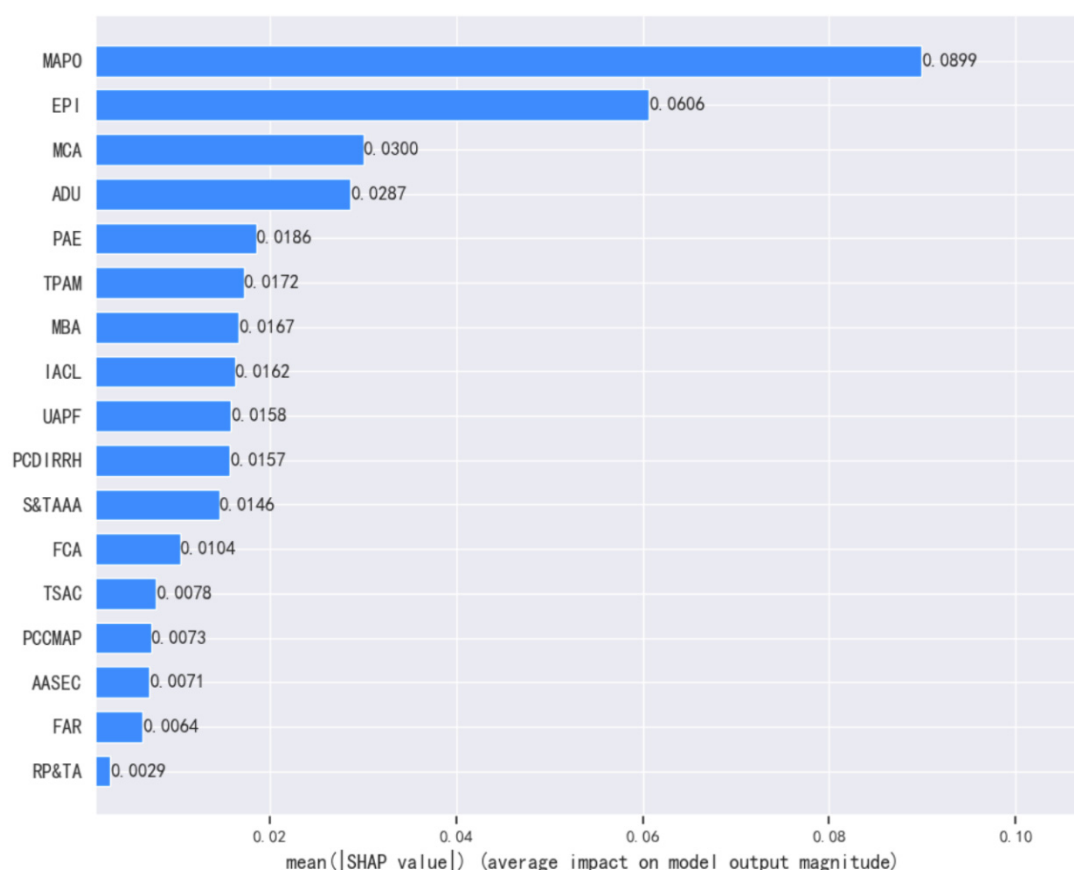


Figure 5. Bar Chart of SHAP for Two-Hidden-Layer Model

According to Figure 5, the top five factors that have the greatest impact on the agricultural economic development of Guangdong Province in the two-hidden-layer model are: the output of major agricultural products, employment in the primary industry, the area of mechanized farming, the use of agricultural diesel, and the embankment protection area. The absolute average SHAP values of the above five factors are 0.0899, 0.0606, 0.0300, 0.0287 and 0.0186 respectively. Among the five indicators mentioned above, 2 belong to the category of agricultural production inputs; 1 belongs to the category of agricultural demand, output and income; 1 belongs to the category of agricultural production scale, and 1 belongs to the category of agricultural water conservancy status

4.3 Statistical Results of the Ranking of ML Models's Indicators

Based on the results of Figures 4 and 5, it is possible to statistically analyze the frequency of each indicator and its category appearing in the top five sequences of SHAP value ranking for the two types of ML models' indicators, as shown in Table 4. The remaining indicators that did not appear in Table 4 indicate that these indicators did not appear in the top five sequences of SHAP value ranking for any type of ML model.

Table 4. Statistical Table of Ranking of ML Models' Indicators

Indicator	number of times	Category	number of times
MAPO	2	Agricultural production input	3
EPI	2	Agricultural water conservancy situation	2
S&TAAA	1	Agricultural production scale	2
TSAC	1	Agricultural demand, output, and income	2
AASEC	1	Agricultural technology and development	1
MCA	1		
ADU	1		
PAE	1		

According to the Table 4, both MAPO and EPI have been selected twice as the top 5 most important indicators for both types of ML models. S&TAAA, TSAC, AASEC, MCA, ADU, and PAE have only appeared once in the ranking of the top 5 most important indicators for two types of ML models. In terms of the categories to which the indicators belong, the indicators representing agricultural production input have appeared 3 times totally in the top 5 rankings of the two types of ML models. The indicators belonging to the categories of agricultural water conservancy situation, agricultural production scale, and agricultural demand, output, and income have only appeared twice each. The indicators belonging to the category of agricultural technology and development have only appeared once.

5. Conclsnion and Suggestions

5.1 Conclusion

Based on the statistical ranking results of the ML Models' indicators, it can be found that the two indicators with the greatest impacts on the agricultural economic development of Guangdong Province are employment in the primary industry and the output of major agricultural products, as these two indicators rank among the top five in both types of ML models. Among the six indicators, including provincial agricultural science and technology award achievements, total crop sowing area, cumulative area of soil erosion control, mechanized farming area, agricultural diesel use, and dike protection zones, each of which ranks top five in only one of the two ML models. Therefore, these six indicators have a strong impact on the agricultural economic development of Guangdong Province.

Regarding the importance of indicator categories on the agricultural economy of Guangdong Province, the ranking results of ML models' indicators shows that the indicators in the agricultural production input category have appeared three times in the ranking of explanatory power analysis in both ML models; the indicators in the agricultural water conservancy status category, agricultural production scale category and agricultural demand, output, and income category have appeared only twice; and the indicators in the agricultural technology and development category appeared only once. Therefore, agricultural production inputs have the highest importance for the development of the agricultural economy in Guangdong Province; agricultural water conservancy status, agricultural production scale,

and agricultural demand, output, and income have relatively high importance for the agricultural economic development of Guangdong Province; while agricultural technology and development categories show relatively low importance for the agricultural economic development of Guangdong Province.

5.2 Suggestions

The ranking results above clarify the importance of different influencing factor indicators and their categories for the agricultural economic development of Guangdong Province. To promote the high-quality and sustainable development of the agricultural economy in Guangdong Province and to provide a scientific decision-making basis for relevant policymakers, this paper proposes the following suggestions.

Research has found that the output of major agricultural products and the total sowing area of crops have a significant impact on the development of the agricultural economy. It is suggested that Guangdong Province, while strictly adhering to the red line of arable land, scientifically plan the structure of crop planting, stabilize the sowing area of grain and important agricultural products, and ensure a steady increase in the output of major agricultural products. At the same time, it is recommended to optimize the layout of agricultural regions based on local conditions, promote large-scale and standardized production of specialty agricultural products, and enhance the market competitiveness and added value of agricultural products, thereby driving the overall growth of the agricultural economy.

Research shows that agricultural water conservancy indicators such as the area protected by embankments and the cumulative area of soil erosion control have a high explanatory power for the agricultural economy. Guangdong Province is located in a coastal area where extreme weather events such as typhoons and rainstorms are frequent, making the construction of water conservancy infrastructure particularly critical. It is recommended to continue increasing investment in the construction and maintenance of embankment projects and drainage facilities, and to improve flood control and drainage standards. In addition, it is recommended to continue promoting comprehensive management of soil erosion, optimize the agricultural ecological environment, enhance the ability of agriculture to withstand natural disasters, and provide solid water conservancy guarantees for stable and efficient agricultural production.

Agricultural demand, output, and income indicators are also of significant importance for the expansion and advancement of agricultural economy. It is suggested that Guangdong Province actively expand market demand for agricultural products, improve the circulation system for agricultural products and the cold chain logistics network, and promote the construction of agricultural product brands. At the same time, it is recommended to deepen structural reforms on the supply side of agriculture, develop new business formats such as deep processing of agricultural products and rural e-commerce, extend the agricultural industry chain, enhance the added value of agricultural products, and effectively increase farmers' operating and property income, achieving a virtuous cycle between agricultural output and farmers' income.

Research conclusions indicate that the explanatory power of agricultural technology and development indicators in the two models is relatively low, reflecting that the driving potential of agricultural technology in Guangdong's agricultural economy has not been fully released. It is suggested that Guangdong Province further increase investment in agricultural science and technology research and development, strengthen the cooperation between agricultural research institutions and agricultural production entities, and accelerate the research and promotion of new agricultural varieties, technologies, and equipment. At the same time, Guangdong Province also needs to improve the mechanism for the transformation of agricultural scientific and technological achievements, enhance the leading demonstration effect of provincial agricultural science and technology awards, and promote the transformation of agricultural development from factor-driven to innovation-driven growth, injecting lasting momentum into the high-quality development of the agricultural economy.

Acknowledgments

Not applicable.

Author contributions

Not applicable.

Funding

Not applicable.

Competing interests

Not applicable.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Macrothink Institute.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Cardoso, V. A., Lourenzani, A. E. B. S., Caldas, M. M., Bernardo, C. H. C., & Bernardo, R. (2022). The benefits and barriers of geographical indications to producers: A review. *Renewable Agriculture and Food Systems*, 37(6), 707–719. <https://doi.org/10.1017/S174217052200031X>
- Deng, C. (2022). Research on the development and integration of inclusive finance in agricultural economy based on the artificial intelligence machine learning algorithm. *Wireless Communications and Mobile Computing*, 2022, 1758983. <https://doi.org/10.1155/2022/1758983>
- Guo, C., Zhang, R., & Zou, Y. (2023). The efficiency of China's agricultural circular economy and its influencing factors under the rural revitalization strategy: A DEA–Malmquist–Tobit approach. *Agriculture*, 13(7), 1454. <https://doi.org/10.3390/agriculture13071454>
- Li, J., Han, Z., & Zhou, H. (2026). Digital-green collaborative transformation and agricultural economic resilience: Evidence from a double machine learning model. *Food and Energy Security*, 15(3), e70264. <https://doi.org/10.1002/fes3.70264>
- Li, S., Yang, C., & Li, K. (2026). Research on traditional rural finance, digital finance, and agricultural economic resilience: Causal inference based on double machine learning. *Sustainability*, 18(13), 6585. <https://doi.org/10.3390/su18136585>
- Wang, X. M. (2011). Study on relation between agriculture production factors and agricultural economy growth. *Advanced Materials Research*, 219–220, 1629–1632. <https://doi.org/10.4028/www.scientific.net/AMR.219-220.1629>
- Xiang, J., Han, P., & Chen, W. (2023). Coordinated development efficiency between cultivated land spatial morphology and agricultural economy in underdeveloped areas in China: Evidence from western Hubei province. *Journal of Geographical Sciences*, 33(4), 801–822. <https://doi.org/10.1007/S11442-023-2107-9>
- Xu, B., Yi, C., & Lin, K. (2026). Research on the new urbanization's impact mechanism on agricultural economic resilience: A quasi-natural experiment via double machine learning and spatial diff-in-diff approaches. *Frontiers in Sustainable Food Systems*, 10, 1724896. <https://doi.org/10.3389/fsufs.2026.1724896>

Yang, C., Liu, W., & Zhou, J. (2024). The role of digital finance in shaping agricultural economic resilience: Evidence from machine learning. *Agriculture*, 14(10), 1834. <https://doi.org/10.3390/agriculture14101834>

Zhang, M., & Daud, Z. B. M. (2022). Study on influencing factors the Belt and Road Initiative and supply side structural reform of China's agricultural economy. In Y. Jiang, Y. Shvets, & H. Mallick (Eds.), *Proceedings of the 2022 2nd International Conference on Economic Development and Business Culture (ICEDBC 2022)* (pp. 575–583). Atlantis Press. https://doi.org/10.2991/978-94-6463-036-7_85

Zheng, R., Liu, X., Gao, Z., Ma, Z., Zhang, M., Sun, Z., & Yan, D. (2025). Level measurement and influencing factor analysis of agricultural sustainable development in Shandong Province, China. *Environment, Development and Sustainability* (prepublish), 1–22. <https://doi.org/10.1007/S10668-025-07161-9>