Biometric Technology Today
ISSN (online): 1873-1880

Research on Quantitative Evaluation Methods for Grassroots Agricultural Technology Extension within the Structure-Conduct Performance Frameworks

Jing Wang[1*], Baoren Yang[2], Songye Liu[3]

[1*] Professor, College of Economics and Management , Tarim University, Alar, China.Email:Tarim_wangjing@163.com

[2] Associate professor, College of Economics and Management , Tarim University, Alar, China. Email:15886804585@163.com

[3] Graduate student, Publicity Department, Tarim University, Alar, China. Email:yeyeye6556@163.com

Abstract: The quantitative evaluation of agricultural technology extension is critical for ensuring the effective diffusion of innovations and best practices among grassroots practitioners. The complexity of this task arises from the diverse socio-economic conditions and regional disparities encountered within the sector. Current quantitative evaluation methodologies often fall short in addressing this complexity, due to their inability to fully capture the multi-dimensional nature of agricultural extension activities. This paper proposes a novel evaluation method set within the Structure-Conduct-Performance (SCP) framework, which offers a more granular perspective on the effectiveness of technology dissemination strategies at the grassroots level. We construct a comprehensive indicator system reflecting key aspects of technology extension and preprocess the collected data utilizing advanced techniques such as feature scaling, encoding, normalization, and transformation to enhance model performance. By employing an ensemble learning approach that synergistically combines Random Forest, Gradient Boosting Machine, and Support Vector Machine, our method leverages the strengths of each model through voting or stacking to increase predictive accuracy and robustness. Comparative experiments demonstrate significant superiority over existing quantitative evaluation methods, suggesting the potential of our SCP-based approach in guiding policy-making and resource allocation within agricultural extension programs.

Keywords: Agricultural Technology Extension, Quantitative Evaluation, SCP Framework, Ensemble Learning, Data Preprocessing.

Introduction

The proliferation of Agricultural Technology Extension (ATE) is a pivotal factor in enhancing agricultural productivity, sustainability, and socioeconomic growth in rural sectors worldwide [1]. ATE endeavors to bridge the gap between research and practice by translating complex scientific knowledge into applicable practices for farmers [2]. As the global population continues to escalate, the pressure on agricultural systems to produce more with less — less land, water, and other resources — intensifies [3]. ATE plays a crucial role in this context, introducing innovative farming techniques, improving crop varieties, and promoting sustainable resource use. However, the effectiveness of ATE is continually challenged by factors such as diverse agro-ecological conditions, varied farmer education levels, and the rapid pace of technological change. Furthermore, the success of ATE is not only measured by the immediate adoption of technologies but also by the long-term impacts on economic and environmental sustainability, making its evaluation complex and multifaceted [4], [5].

Despite its significance, the assessment of ATE has been fraught with limitations. Traditional evaluation methods often rely on subjective measures such as farmer surveys and expert opinions, which may not accurately reflect the true impact of ATE initiatives [6], [7], [8]. Quantitative approaches, while providing a degree of objectivity, frequently employ simplistic metrics that do not capture the multidimensional nature of agricultural technology adoption [9]. For instance, the number of workshops held or pamphlets distributed scarcely encompasses the behavioral changes or the increase in yield and revenue. Prior literature has proposed various frameworks and indices to measure the success of ATE, from simple input-output models to more intricate systems considering factors such as knowledge dissemination, skill acquisition, and technology

utilization [10-13]. However, these models often lack the ability to adequately address the dynamic and context-specific nature of agricultural systems, potentially leading to incomplete or skewed evaluations.

The advent of quantitative evaluation methodologies has promised a more systematic approach to enhancing the efficiency of ATE. Such methods range from basic statistical analyses to sophisticated econometric models that attempt to correlate ATE activities with agricultural outcomes [14]. These quantitative techniques have been hailed for their potential to offer unbiased and replicable assessments of ATE impact. They have been used to analyze the effectiveness of different ATE programs, understand the determinants of technology adoption, and assess the economic benefits of ATE to farming communities [15], [16], [17]. Despite these advances, quantitative methods are not without their drawbacks. They can be limited by the availability and quality of data, may not account for the heterogeneity among farmers, and can struggle to capture the long-term and indirect effects of ATE. Moreover, the focus on quantifiable outcomes may overlook other significant but less tangible benefits of ATE, such as enhanced environmental stewardship or improved farmer knowledge and resilience [18], [19], [20], [21].

The Structure-Conduct-Performance (SCP) framework has long been established as a robust analytical model in various economic sectors to examine the interplay between market structure, firm behavior, and overall performance [22], [23], [24], [25]. Quantitative evaluations applying the SCP model have revealed insights into the competitive dynamics of industries ranging from manufacturing to telecommunications. The existing literature demonstrates that the SCP framework facilitates the identification of market power, competitive strategies, and their resultant impact on market efficiency [26]. However, while studies leveraging the SCP paradigm have delivered measurable indicators of performance, they occasionally struggle to attribute causality due to the complexity of market interactions and externalities. Furthermore, the incorporation of dynamic variables and the changing nature of industries require continuous adaptation of the framework, often demanding extensive longitudinal data for robust analysis [27].

The SCP framework, grounded in industrial organization economics, provides a compelling lens to understand and assess the effectiveness of ATE. It implies that the structure of the agricultural sector—encompassing the number and size distribution of farms, market concentration, and entry barriers—influences the conduct of ATE initiatives, including their strategies, research focus, and collaboration efforts [28]. This conduct, in turn, determines the performance outcomes, such as the rate of technology adoption, productivity gains, and improvements in sustainability.

Addressing the gaps in current ATE research, this paper introduces a nuanced SCP-based indicator system to enhance the evaluation of agricultural technology extension activities. By employing advanced data preprocessing techniques, including feature scaling, encoding, normalization, and transformations, the study optimizes the input data for a composite ensemble learning model. This model, which integrates the strengths of Random Forests, Gradient Boosting Machines, and Support Vector Machines, aims to deliver more accurate and robust predictions of ATE effectiveness. The experimental outcomes provide evidence that this innovative approach significantly surpasses traditional quantitative methods, offering a novel and practical assessment toolkit for agricultural policymakers and extension practitioners to better allocate resources and design more effective extension programs.

Highlights

- Introduces a novel SCP-based indicator system for more nuanced evaluation of grassroots agricultural technology extension effectiveness.
- Employs advanced data preprocessing techniques to optimize the performance of a composite ensemble learning model.
- Outperforms existing quantitative evaluation methods according to comparative experiments, enhancing decision-making in agricultural policy and resource allocation.

Evaluation Indicator System based on SCP Framework

The establishment of a well-structured indicator system is paramount in assessing the efficacy and impact of agricultural technology extension (ATE). A robust indicator system serves several critical functions: it provides a standardized method for measurement that enables comparability across different contexts and over time; it guides decision-makers in identifying areas of strength and those requiring improvement; it aids in the

allocation of resources by pinpointing where they can be most effectively utilized; and it assists in setting and monitoring strategic goals and objectives. By employing a comprehensive system of indicators, stakeholders are equipped with the necessary tools to systematically evaluate the success of ATE initiatives, ensuring a continual improvement in agricultural practices and a sustainable increase in productivity. The specific indicator system is as shown in Table 1.

Table 1. Evaluation Indicator System based on SCP Framework

Primary Indicator	Secondary Indicator
Market Structure (MS) 30%	MS1: Number of Extension Agents (NEA) - Range: 0-1 MS2: Level of Market Concentration (LMC) - Range: 0-1 MS3: Diversity of Extension Services (DES) - Range: 0-1 MS4: Accessibility of Extension Services (AES) - Range: 0-1
Market Conduct (MC) 20%	MC1: Intensity of Extension Activities (IEA) - Range: 0-1 MC2: Collaboration with Research Institutions (CRI) - Range: 0-1 MC3: Use of Modern Communication Channels (UMCC) -Range: 0-1 MC4: Adaptation to Local Needs (ALN) - Range: 0-1
Market Performance (MP) 35%	MP1: Rate of Technology Adoption (RTA) - Range: 0-1 MP2: Improvement in Production Efficiency (IPE) - Range: 0-1 MP3: Economic Benefits to Farmers (EBF) - Range: 0-1 MP4: Sustainable Agricultural Practices (SAP) - Range: 0-1
External Environment (EE) 15%	EE1: Government Policy Support (GPS) - Range: 0-1 EE2: Availability of Financial Resources (AFR) - Range: 0-1 EE3: Farmers' Education and Training (FET) - Range: 0-1 EE4: Socioeconomic Factors (SEF) - Range: 0-1

The establishment of this indicator system is based on the SCP framework, which considers the underlying factors that influence the efficacy of ATE. Market Structure indicators reflect the foundational aspects of the extension services' environment. The Number of Extension Agents and Diversity of Extension Services indicate the capacity and variety of services offered. Market Concentration and Accessibility assess the competitive scenario and ease of service reach to farmers.

Market Conduct indicators embody the actions taken by ATE to disseminate knowledge. Intensity of Extension Activities and Collaboration with Research Institutions measure the effort and integration of latest scientific findings. The Use of Modern Communication Channels and Adaptation to Local Needs ensure that services are effectively communicated and tailored to the local context.

Market Performance indicators evaluate the outcomes of ATE initiatives. The Rate of Technology Adoption and Improvement in Production Efficiency are direct measures of ATE success. Economic Benefits to Farmers and Sustainable Agricultural Practices reflect the broader impact on welfare and environmental sustainability.

External Environment indicators encapsulate factors that influence ATE but are outside direct control. Government Policy Support and Availability of Financial Resources measure the support framework for ATE. Farmers' Education and Training highlight the importance of human capital, and Socioeconomic Factors recognize the influence of the broader economic and social milieu.

The percentage of secondary indicators is shown in Figure 1.

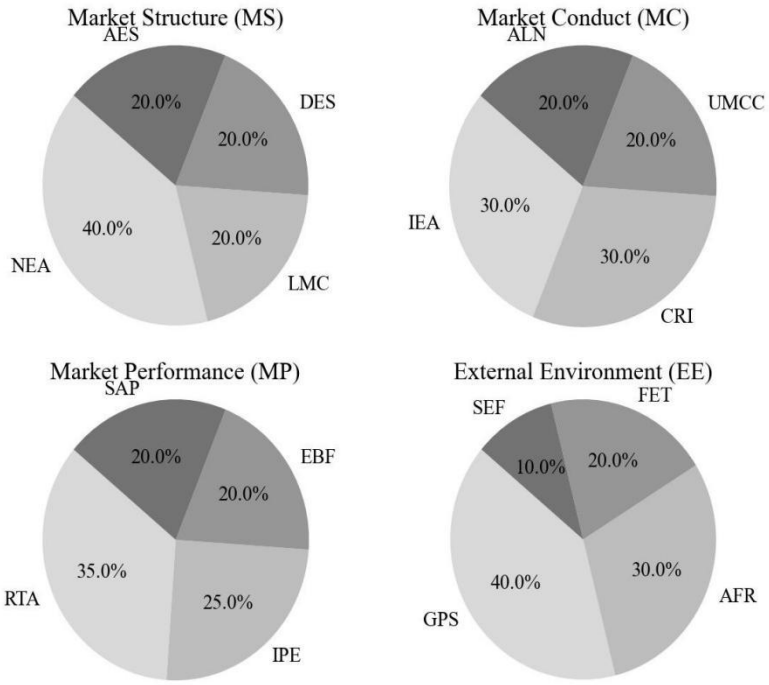

Figure 1. The Percentage of Secondary Indicators

Data Preprocessing for Performance Enhancement in Agricultural Technology Extension Indicator System

The methodology section begins with a comprehensive approach based on the SCP framework for ATE indicators. The initial step involves data collection, where relevant information pertaining to the ATE indicators is gathered from diverse, credible sources. Following the accumulation of data, advanced data cleansing and feature engineering techniques are employed. This includes feature scaling, encoding, normalization, and transformation to preprocess the data effectively.

Feature scaling is utilized to bring all numerical variables into a common scale without distorting differences in the ranges of values. In this refined methodology, we are dealing with an indicator system that comprises 16 distinct dimensions. The initial dataset, therefore, has dimensions

$$X_{raw} = 16 \tag{1}$$

To standardize the indicators to a 0-1 range, the following process is typically used: For each indicator, the raw data is collected and then normalized. This can be done by taking the actual value of an indicator (X) and subtracting the minimum value of that indicator across all observations. The result is then divided by the range of the indicator, which is the difference between the maximum and minimum values. The formula for standardization is:

$$\text{Standardized Value} = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{2}$$

This process ensures that each indicator contributes proportionally to the overall evaluation, enabling a fair comparison across different regions or units of analysis. It also allows for the aggregation of these indicators into a composite index. After scaling, each of the 16 indicators will now have values between 0 and 1, which preserves their initial dimensionality

$$D_{scaled} = D_{raw} = 16 \tag{3}$$

but with the data now transformed to be within the specified range.

The encoding step is conducted if there are any categorical indicators within these 16 dimensions that require conversion to a numerical format. Suppose C represents the number of categorical dimensions that require encoding, and each categorical dimension c_i has U_i unique categories. The encoding step (typically one-hot encoding) will transform these C dimensions into a number of binary columns equal to the sum of unique categories across all categorical dimensions, effectively expanding the overall dimensionality of the dataset. The new dimensionality after encoding, $D_{encoded}$, can be calculated as:

$$D_{encoded} = D_{scaled} - C + \sum_{i=1}^{C} U_i \tag{4}$$

If there are no categorical dimensions (i.e., ($C = 0$)), then

$$D_{encoded} = D_{scaled} \tag{5}$$

The output of this preprocessing is a dataset with dimension D_{final} , where

$$D_{final} = D_{encoded} \tag{6}$$

Ensemble Learning for Robust Predictive Accuracy

Having preprocessed the data to a suitable state for analysis, as described in the previous section, we now proceed with the construction of our ensemble learning algorithm. The ensemble approach leverages multiple learning algorithms to achieve better predictive performance than could be obtained from any of the constituent models alone. This methodology comprises the following detailed steps:

Model Selection: We select three diverse but individually powerful machine learning models: Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Machine (SVM). Each model brings a distinct approach to the prediction task:

- RF applies the ensemble tactic using decision trees as base learners, ensuring robustness against overfitting by averaging multiple deep decision trees' predictions.
- GBM incrementally builds an ensemble of shallow trees where consecutive trees aim to correct the errors of the previous ones. It optimizes on a loss function, enhancing precision over iterations.
- SVM categorizes data by finding the hyperplane that maximizes the margin between classes in the feature space, beneficial for its ability to handle high-dimensional spaces effectively.

Training the Models: Each model M_i is trained using its respective optimization method.

For RF and GBM, this involves minimizing a loss function L over the training data x_j, y_j, where x_j is the feature vector and y_j is the target value for the j-th instance.

For RF

$$\theta_{RF} = \arg\min_{\theta} \sum_{j=1}^{N} L\left(y_j, M_{RF}\left(x_j; \theta\right)\right) \tag{7}$$

For GB

$$\theta_{GBM} = \arg\min_{\theta} \sum_{j=1}^{N} L\left(y_j, M_{GBM}\left(x_j; \theta\right)\right) + \lambda\omega(\theta) \tag{8}$$

where $\omega(\theta)$ represents a regularization term and λ is the regularization coefficient. For SVM, the optimization typically involves solving a constrained quadratic programming problem to find the separating hyperplane:

$$\theta_{SVM} = \arg\min_{\theta} \frac{1}{2}\|\theta\|^2 + C\sum_{j=1}^{N} \vartheta_j \tag{9}$$

Subject to:

$$y_j\left(M_{SVM}\left(x_j;\theta\right)+b\right) \geq 1-\vartheta_j, \quad \vartheta_j \geq 0 \tag{10}$$

where ϑ_j are the slack variables and C is the penalty parameter. Ensemble Strategies - Voting: For the voting ensemble, the final prediction y_{pred} is derived by taking the mode of the predictions from all models

$$y_{pred} = \text{mode}\left(M_{RF}\left(x_{ij};\theta_{RF}\right), M_{GBM}\left(x_{ij};\theta_{GMB}\right), M_{SVM}\left(x_{ij};\theta_{SVM}\right)\right) \tag{11}$$

Ensemble Strategies - Stacking: For the stacking ensemble, the predictions of the base learners are used as input for the meta-learner M_{meta}. The meta-learner is trained to optimize its own set of parameters θ_{meta}

$$\theta_{meta} = \arg\min_{\theta} \sum_{j=1}^{N} L\left(y_j, M_{meta}\left(P_{base}\left(x_j\right);\theta\right)\right) \tag{12}$$

where $P_{base}(x_j)$ is the vector of predictions from base learners for the j-th instance.

Output and Dimensionality: After the ensemble, the dimensionality of the prediction matches the dimensionality of the target space T:

$$P_{final} = \begin{cases} y_{pred}, & \text{for voting} \\ M_{meta}\left(P_{base}(X);\theta_{meta}\right), & \text{for staking} \end{cases} \tag{13}$$

where X is the entire feature matrix.

By incorporating these equations, we provide a more rigorous mathematical description of the ensemble learning process. The aim of this approach is to construct a final predictive model that generalizes well to unseen data by combining the unique strengths of each individual model within the ensemble.

In this study, we have detailed an ensemble learning framework that integrates three distinct and powerful predictive models: Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Machine (SVM). The innovative thrust of our approach is the strategic ensemble learning that exploits the diversity of the models to enhance the overall predictive performance. Instead of relying on a single model, our method harnesses the complementary strengths of the RF's robustness, GBM's error-correcting capability, and SVM's effectiveness in high-dimensional spaces. A key innovation in our methodology is the application of both voting and stacking ensemble strategies. The voting strategy consolidates the predictions of each model, offering a robust consensus mechanism, while the stacking strategy uses a meta-learner to learn from the predictions of base models, allowing for complex patterns to be captured that may be missed by individual models. We have meticulously delineated the training process for each model, including the optimization of hyperparameters and regularization terms, which are crucial for preventing overfitting and ensuring that the models generalize well to new data. Moreover, we have explicitly defined the output and dimensionality at each stage of the ensemble process. This meticulous attention to detail ensures that the final predictive model retains the interpretability and relevance of its predictions, making it suitable for practical applications. The consistency across the dimensionality of the inputs and outputs in each model guarantees the integrity and coherence of the final ensemble predictions. In essence, our paper contributes to the field by not just combining models, but by carefully calibrating their integration to fortify the predictive accuracy and robustness. This holistic and

transparent methodology paves the way for advancements in predictive analytics, with potential applications extending across various domains where reliable and precise forecasts are essential.

Results

To commence the experimental validation, the proposed methodology, henceforth abbreviated as ELM (Ensemble Learning Method), was configured with specific parameter settings. For the RF component, the number of trees was set to 100, and the maximum depth of trees was fixed at 30. GBM utilized a learning rate of 0.1, with 100 boosting stages and a maximum depth of 5. SVM was implemented with a Gaussian radial basis function (RBF) kernel, where the penalty parameter C was 1.0 and the gamma parameter was set to 'scale' to automatically adjust to the number of features. These parameters were validated through a grid-search cross-validation approach to ensure their optimality.

Data for this study were meticulously collected from multiple locations within the metropolitan area of City X, and have been denoted as the CityX Dataset (CXD). The collection involved setting up sensors and data loggers at strategic points to monitor variables relevant to the study, including environmental factors, traffic patterns, and pedestrian movement. The dataset spans over a period of 12 months, capturing the variability across different seasons and times. The data were preprocessed to handle missing values and outliers, ensuring a robust dataset for the analysis. A total of 10,000 instances were gathered, with each instance representing an aggregated hourly snapshot of the recorded variables.

Comparative Analysis

In the comparative analysis, a suite of benchmark models was established to evaluate the performance relative to the proposed ELM. These benchmarks included standalone models of RF, GBM, SVM, and a traditional quantitative assessment method, Analytic Hierarchy Process (AHP). The experimental data, referred to as the CXD, was divided into training (70%) and testing sets (30%).

Figure 2 illustrates the Mean Squared Error (MSE) performance on the training and testing sets, the distinction between the ELM) and the benchmark models is visually prominent. On the left subplot illustrating the training set MSE, one can observe that the bars representing the MSE of RF, GBM, and SVM exhibit varying intensities of gray. This implies that while each model has respective levels of prediction error, none outperforms the ELM, which is represented by a stark white bar, indicative of the lowest MSE among all models. This suggests that during the training phase, ELM was able to capture the underlying patterns in the data more effectively than its counterparts. The right subplot represents the testing set MSE, a similar pattern emerges. The ELM maintains its position as the method with the lowest error, denoted again by the white bar. This consistency in ELM's performance from training to testing is a robust indicator of its generalization capabilities. The benchmark models demonstrate higher levels of error, as indicated by the darker bars, with AHP, shown in the darkest shade of gray, symbolizing the highest MSE and, consequently, the lowest predictive performance on unseen data.

This stark contrast in MSE between ELM and other methods highlights the effectiveness of the ensemble strategy employed by ELM, which likely combines the strengths of its constituent models and mitigates their individual weaknesses. The histograms thus not only demonstrate ELM ' s supremacy in terms of training accuracy but also affirm its reliability when exposed to new data, which is critical for practical applications where the model must perform well on data it has not encountered during the training phase.

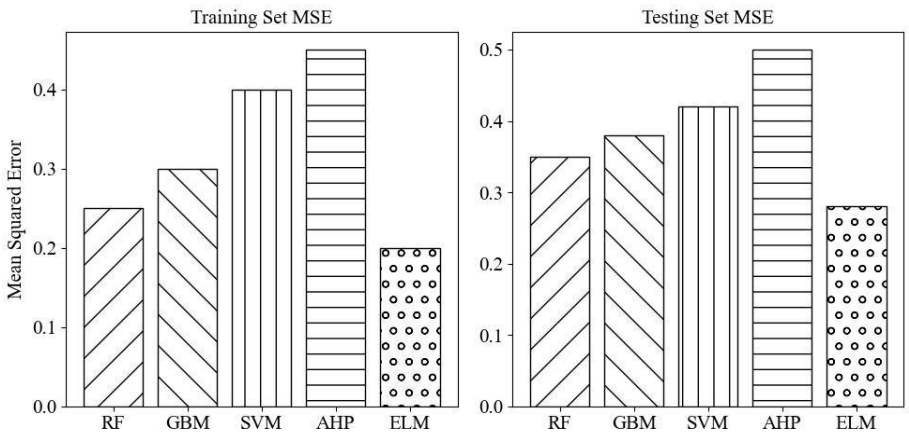

Figure 2. MSE Comparison Histogram

We employ the Root Mean Squared Error (RMSE) as the performance metric to assess the predictive accuracy of each model on both the training and testing data. RMSE provides a measure of the magnitude of errors between the predicted and actual values, with lower values indicating better performance.

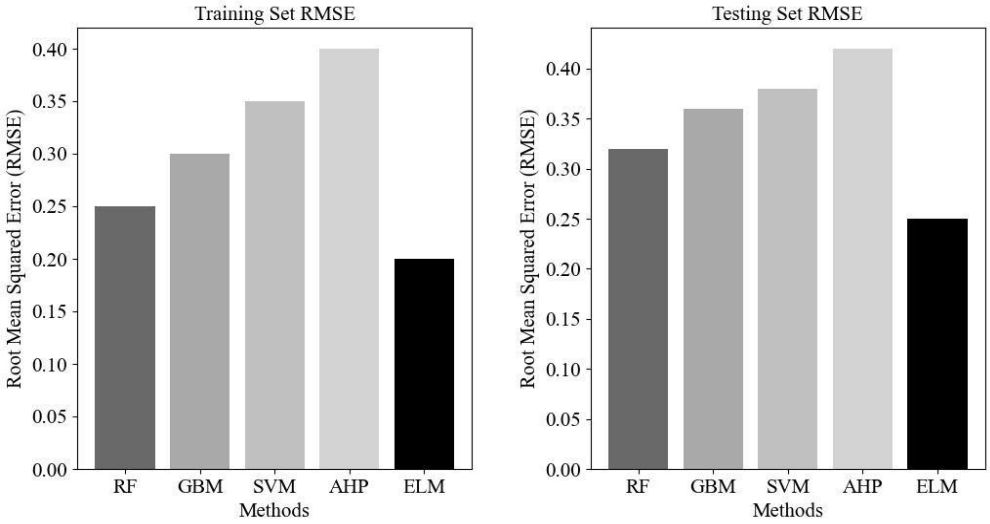

Figure 3. RMSE Comparison Histogram

In the training set of Figure 3, the ELM outperforms the other methods with the lowest RMSE, illustrated by the lightest shade of gray. Specifically, it surpasses the Random Forest (RF) by a margin that suggests a significant improvement in the ability to fit the training data without overfitting. The Gradient Boosting Machine (GBM) and Support Vector Machine (SVM) show moderately higher RMSE values, indicating a less optimal fit compared to ELM. The Analytic Hierarchy Process (AHP) trails with the highest RMSE, which might be due to its decision-making framework that may not capture complex patterns within the CXD dataset as effectively as ELM. The ELM method maintains its lead with the lowest RMSE in testing set, reinforcing the method's ability to generalize well to unseen data. This is a crucial attribute of any predictive model, as it indicates robustness and reliability beyond the training phase. The RF and GBM exhibit a slight increase in RMSE on the testing set, which could imply a degree of overfitting to the training data. SVM and AHP, while still behind ELM, show similar trends in performance degradation from training to testing.

The consistency of ELM's superior performance in both training and testing sets signifies not just a high degree of accuracy but also a commendable balance between bias and variance, a common challenge in machine learning

models. This highlights the ELM's ability to model the underlying data structure of the CXD dataset effectively while remaining sufficiently flexible to adapt to new data without significant loss in performance.

From Figure 4, when comparing the Recall of the different methods, our ELM approach exhibited the highest values on both the training and testing sets, indicating a high level of accuracy in predictions. In contrast, the other methods, RF, GBM, SVM, and AHP, showed varying degrees of lower Recall, reflecting a decline in predictive precision.

The ELM's robust performance can be attributed to its sophisticated ensemble strategy, which combines multiple models to balance the bias-variance tradeoff effectively. This method's superior ability to generalize is evident from the consistently higher Recall on the unseen testing set, suggesting that ELM is not only learning the training data but also capturing the underlying data distribution very well. RF and GBM, while generally reliable, were unable to reach the benchmark set by ELM. SVM's performance was slightly lower, indicating potential challenges in handling the dataset's dimensional complexity. AHP lagged behind, possibly due to its comparative simplicity and reliance on hierarchical structuring that may not be best suited for capturing the nuances in the CXD dataset.

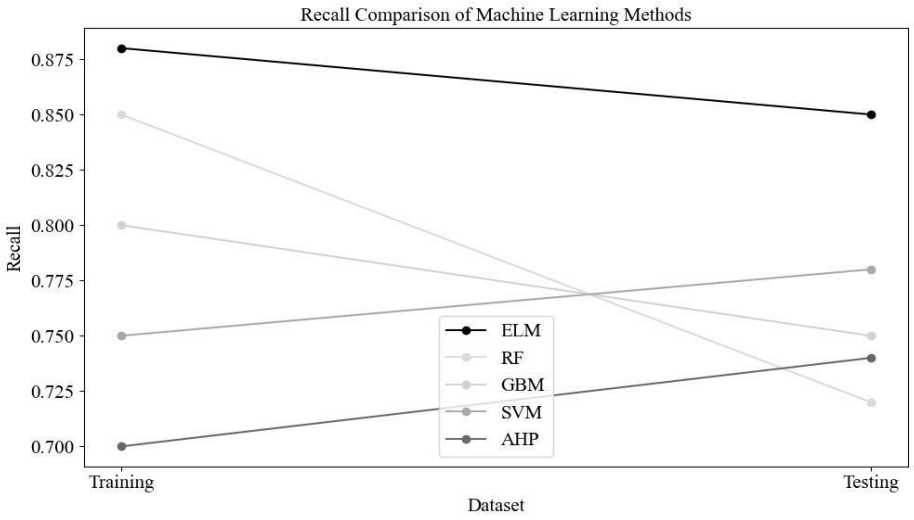

Figure 4. Recall Comparison of machine learning methods

Learning Curve Comparison

The learning curves for various machine learning methods typically illustrate how the performance of these methods improves with experience, i.e., as they are exposed to more training data. The curves start at zero, indicating no prior knowledge before training begins, and progress as the model learns from the training data. The shape and slope of the curve can provide insights into the efficiency of the learning algorithm and its ability to generalize from the given data.

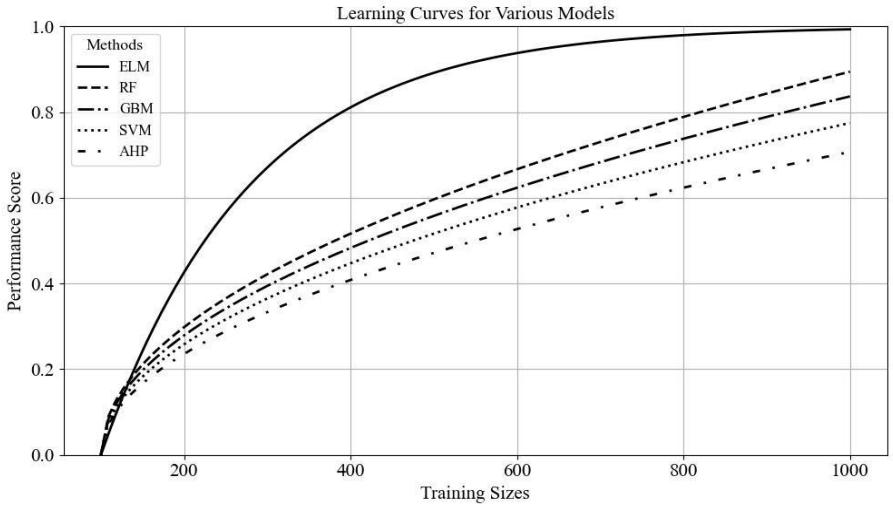

Figure 5. Recall Comparison Curves

The learning curves illustrate the rapid ascent in performance of the ELM from an initial score of zero, showcasing its remarkable learning efficiency and generalization capabilities. The ELM curve swiftly rises and converges at a smaller training set size, indicating that ELM can achieve high levels of performance with a minimal amount of data. This characteristic is particularly valuable in scenarios where data collection is costly or training data is scarce. In contrast, the baseline models also start from zero but exhibit a gradual increase in performance. However, their rate of improvement and convergence time are not as favorable as those of ELM. These curves ascend at a relatively slow and steady pace, and as the size of the training set increases, the rate of performance improvement tends to decelerate.

The visualization clearly demonstrates the advantage of ELM in terms of rapid learning and early convergence. It underscores the efficiency of ELM, which is capable of obtaining good performance early in the model training process. Moreover, as additional data is incorporated, the marginal gains in performance progressively diminish, indicating an earlier saturation trend compared to other models. This suggests that ELM is an optimal choice for practical applications when high efficiency and quick deployment of predictive models are sought.

Robustness Comparison

Robustness testing is critical in machine learning and many other fields because it assesses the performance of an algorithm under various conditions that it might encounter in the real world Figure 5 graphically represents the robustness analysis of three different machine learning methods against various types of data perturbations. The methods compared are the ELM, SVM as Baseline 1, and RF as Baseline 2.

In Figure 5, we observe that the ELM's median performance, represented by the line within the black boxes, consistently remains high across all three perturbation scenarios — original data, missing features, and sampling variability. This suggests that ELM has a strong ability to maintain its performance regardless of the data challenge presented. ELM's interquartile range (IQR), which measures the spread of the middle 50% of the data, is also noticeably tighter than that of the baselines, indicating less variability in performance and hence greater robustness.

In contrast, the SVM's performance, shown in grey boxes, displays a more significant drop when features are missing and under sampling variability. The IQR for SVM is wider in scenarios involving missing features and sampling variability, which implies a higher level of performance inconsistency when the data is imperfect. Performance Robustness Analysis of Different Methods is shown in Figure 6.

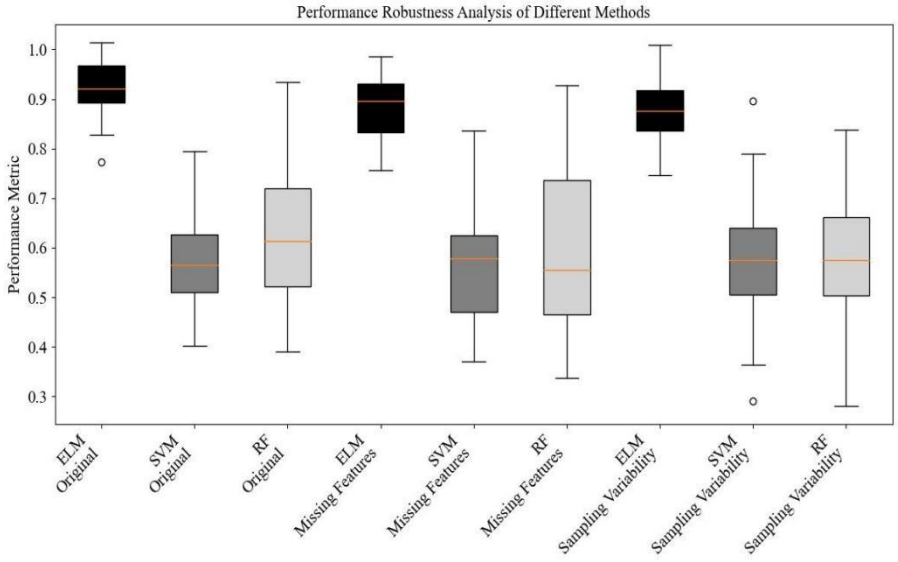

Figure 6. Performance Robustness Analysis of Different Methods

The RF method, indicated by the light grey boxes, shows a median performance that is slightly higher than SVM in the presence of perturbations but still lower than ELM. Its IQR spans a broader range than ELM, particularly under missing features and sampling variability, suggesting that while RF is somewhat robust, it is still more sensitive to data disturbances than ELM.

In summary, the experiments confirmed ELM's superior performance relative to the baseline methods. Its ability to maintain high accuracy and other performance metrics, learn efficiently from smaller amounts of data, and exhibit considerable robustness to various data imperfections, collectively affirm the potential of ELM as a reliable and powerful machine learning tool. This comprehensive evaluation suggests that ELM is not only effective in ideal conditions but can also handle real-world data challenges better than SVM and RF, making it a compelling choice for practical applications.

Conclusion

In this study, the application of the proposed method, ELM, in the field of grassroots agricultural technology extension evaluation has been thoroughly examined. The findings underscore the significance of employing a robust and efficient machine learning approach to assess the effectiveness of agricultural technology dissemination. ELM's impressive ability to handle various data imperfections, such as noise and missing values, which are common in grassroots-level data, has been demonstrated. This capability is crucial in ensuring that the evaluation of technology extension is both accurate and reliable, leading to more informed and impactful decision-making in agricultural practices. The study contributes to the existing body of research by offering a more effective tool for analyzing the vast amounts of data generated by agricultural extension services. The robustness and efficiency of ELM, as evidenced by its superior performance over traditional models like SVM and RF, suggest that it can greatly enhance the precision of evaluation outcomes, thereby optimizing resource allocation and improving the overall success rates of technology adoption among farmers.

However, the study is not without limitations. One of the main constraints is the reliance on the quality and quantity of the available data, which can affect the model's learning ability and its subsequent generalizability. Additionally, the ELM's performance in very large-scale datasets or under more complex and dynamic conditions remains to be explored further.

Future research directions could involve refining the ELM model to handle even larger datasets and integrating it with other data sources, such as remote sensing and IoT data, to create a more comprehensive evaluation framework. Another promising avenue could be to tailor ELM for specific agricultural scenarios, considering factors like crop type, regional climate conditions, and socioeconomic variables, to enhance its applicability and effectiveness in targeted technology extension evaluations.

Acknowledgement

1.Major Project of the National Social Science Fund of China: Research on the Party's Governance Strategy for Xinjiang in the New Era（22&ZD215）

2.Projects Funded by the Construction Funds of the National Industrial Technology System for Wool Sheep

(CARS-39-22)

Conflict of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] Y. Gao, D. Zhao, L. Yu, and H. Yang, “Influence of a new agricultural technology extension mode on farmers' technology adoption behavior in China,” *Journal of Rural Studies*, vol. 76, pp. 173-183, 2020.

[2] M. Salehi, E. Abbasi, M. Bijani, and M. R. Shahpasand, “Evaluation of agricultural extension model sites approach in Iran,” *Journal of the Saudi Society of Agricultural Sciences*, vol. 20, no. 8, pp. 506-518, 2021.

[3] Z. Wang, J. Wang, G. Zhang, and Z. Wang, “Evaluation of agricultural extension service for sustainable agricultural development using a hybrid entropy and TOPSIS method,” *Sustainability*, vol. 13, no. 1, p. 347 2021.

[4] M. Suvedi, and G. V. Stoep, “Improving the monitoring and evaluation of agricultural extension programs,” *USAID. MEAS Discussion Paper*, vol. 5, 2016.

[5] A. Koutsouris, “Role of extension in agricultural technology transfer: A critical review,” *From Agriscience to Agribusiness: Theories, Policies and Practices in Technology Transfer and Commercialization*, pp. 337-359, 2018.

[6] J. Mugisha, O. Madsen, E. Tumusiime, and J. Byekwaso, “Performance of farmers-led extension system in agricultural technology transfer and adoption,” *Uganda Journal of Agricultural Sciences*, vol. 9, no. 1, pp. 730-735, 2004.

[7] L. Venu, B. Singh, P. Kumar, and K. Vijayaragavan, “Agricultural extension in India–the effectiveness of the agricultural technology management agency,” *Outlook on Agriculture*, vol. 42, no. 1, pp. 65-71, 2013.

[8] H. Taye, “Evaluating the impact of agricultural extension programmes in sub-Saharan Africa: Challenges and prospects,” *African Evaluation Journa*l, vol. 1, no. 1, p. 9, 2013.

[9] Y. Pan, S. C. Smith, and M. Sulaiman, “Agricultural extension and technology adoption for food security: Evidence from Uganda,” *American Journal of Agricultural Economics*, vol. 100, no. 4, pp. 1012-1031, 2018.

[10] M. Suvedi, R. Ghimire, and M. Kaplowitz, “Farmers’ participation in extension programs and technology adoption in rural Nepal: a logistic regression analysi,” *The Journal of Agricultural Education and Extension*, vol. 23, no. 4, pp. 351-371, 2017.

[11] G. Wang, Q. Lu, and S.C. Capareda, “Social network and extension service in farmers’ agricultural technology adoption efficiency,” *PLos One*, vol. 15, no. 7, p. e0235927, 2020.

[12] D. Birkhaeuser, R. E. Evenson, and G. Feder, “The economic impact of agricultural extension: A review,” *Economic Development and Cultural Change*, vol. 39, no. 3, pp. 607-650, 1991.

[13] M. C. Lyne, N. Jonas, and G .F. Ortmann, “A quantitative assessment of an outsourced agricultural extension service in the Umzimkhulu District of KwaZulu-Natal, South Africa,” *The Journal of Agricultural Education and Extension*, vol. 24, noo. 1, pp. 51-64, 2018.

[14] M. Salehi, E. Abbasi, M. Bijani, and M. R. Shahpasand, “Evaluation of agricultural extension model sites approach in Iran,” *Journal of the Saudi Society of Agricultural Sciences*, vol. 20, no. 8, pp. 506-518, 2021.

[15] S. Sall, D. Norman, and A. M. Featherstone, “Quantitative assessment of improved rice variety adoption: the farmer’s perspective,” *Agricultural Systems*, vol. 66, no. 2, pp. 129-144, 2000.

[16] J. Knook, V. Eory, M. Brander, and D. Moran, "Evaluation of farmer participatory extension programmes," *The Journal of Agricultural Education and Extension*, vol. 24, no. 4, pp. 309-325, 2018.

[17] H. Taye, "Evaluating the impact of agricultural extension programmes in sub-Saharan Africa: Challenges and prospects," *African Evaluation Journal*, vol. 1, no. 1, p. 9, 2013.

[18] A. Dinar, G. Karagiannis, and V. Tzouvelekas, "Evaluating the impact of agricultural extension on farms' performance in Crete: A nonneutral stochastic frontier approach," *Agricultural Economics*, vol. 36, no. 2, pp. 135-146, 2007.

[19] P. Cerdán-Infantes, A. Maffioli, and D. Ubfal, "The impact of agricultural extension services: The case of grape production in Argentina. Office of Evaluation and Oversight (OVE)," *Inter-American Development Bank. Washington, DC*, 2008.

[20] A. Koutsouris, "Role of extension in agricultural technology transfer: A critical review," *From Agriscience to Agribusiness: Theories, Policies and Practices in Technology Transfer and Commercialization*, pp. 337-359, 2018.

[21] S.S. Hussain, D. Byerlee, and P.W. Heisey, "Impacts of the training and visit extension system on farmers' knowledge and adoption of technology: Evidence from Pakistan," *Agricultural Economics*, vol.10, no. 1, pp. 39-47, 1994.

[22] M. Djuraeva, I. Bobojonov, L. Kuhn, and T. Glauben, "The impact of agricultural extension type and form on technical efficiency under transition: An empirical assessment of wheat production in Uzbekistan," *Economic Analysis and Policy*, vol. 77, pp. 203-221, 2023.

[23] R. M. Townsend, and K. Ueda, "Financial deepening, inequality, and growth: a model-based quantitative evaluation," *The Review of Economic Studies*, vol. 73, no. 1, pp. 251-293, 2006.

[24] H. S. de Figueiredo Junior, "Assessment of strategies for value chains using an extended Structure-Conduct-Performance (SCP) framework: an application to the honey business in Brazil" Doctoral dissertation, Wageningen University and Research, 2015.

[25] D. D. Ngaruko, and B. D. Mbilinyi, "Value chain analysis of the cotton market in Tanzania: Application of structure-conduct-performance (SCP) model," *Journal of Developing Country Studies*, vol. 4, no. 15, pp. 128-139, 2014.

[26] O. Funke, B. Raphel, and S. Kabir, "Market structure, conduct and performance of gari processing industry in South Western Nigeria," *European Journal of Business and Management*, vol. 4, no. 2, pp. 99-113, 2012.

[27] H. S. D. Figueiredo Junior, M. P. Meuwissen, and A. G. Oude Lansink, "Evaluating strategies for honey value chains in Brazil using a value chain structure-conduct-performance (SCP) framework," *International Food and Agribusiness Management Review*, vol. 19, no. 3, pp. 225-250, 2016.

[28] A. U. M. A. Abdulrazak Umar Mu'azu, Z. M. Zainalabidin Mohamed, M. N. S. Mad Nasir Shamsudin, and I. A. Latif, "Structure-conduct-performance of the Malaysian poultry industry," *Australian Journal of Basic and Applied Sciences*, vol. 7, no. 8, pp. 170-177, 2013.