Article

Research on Maturity Assessment and Leap Forward Prediction of Rural E-commerce Based on LLM-GNN in the Context of Digital Economy

Yeqiang Zheng[1,*] and Sheila M. Geronimo[2]

[1] Yulin Normal University, Yulin, China
[2] Computer Science, University of the East, Manila, Philippines
* Correspondence: Yeqiang Zheng, Yulin Normal University, Yulin, China

Abstract: This study proposes a hybrid analytical framework integrating a large language model and a graph neural network to assess the maturity of rural e-commerce and predict leapfrog development paths in the digital economy. The framework constructs a multi-dimensional maturity assessment system covering five dimensions: infrastructure, logistics system, digital skills, market connectivity, and institutional support. The large language model extracts semantic features from unstructured policy documents and corporate statements; the graph neural network disseminates relational information through rural business entity networks to classify maturity stages and identify transformation paths. The model is validated using a simulation dataset (59 township-level units) based on real statistical feature constraints . This integrated model can simultaneously capture textual and relational signals, thereby improving the accuracy of maturity assessment and its forward-looking leapfrog prediction capability. The hybrid model achieves an accuracy of 0.87, a macro F1 score of 0.85, and a Cohen's Kappa coefficient of 0.83, outperforming baselines such as GNN only (0.79), LLM only (0.74), and random forest (0.68). The study identifies two key leapfrog thresholds: logistics coverage ≥ 0.70 and digital skills participation ≥ 0.60. This study provides a scalable diagnostic tool for regional planning and offers empirical evidence for the development of digital inclusion policies in the rural economy.

Keywords: rural e-commerce; maturity assessment; graph neural network; large language model; digital economy; leapfrog prediction; simulation verification

1. Introduction

1.1. Background and Research Questions

The rapid expansion of the digital economy has fundamentally reshaped the traditional business ecosystem, and e-commerce is now widely regarded as a key engine for promoting regional modernization and agricultural transformation [1,2]. The No. 1 Central Document of 2024 listed the "High-Quality Development Project of Rural E-commerce" as the top priority for county revitalization and clearly required the construction of "county-level live streaming bases". Against this background, many studies have confirmed that e-commerce has matured in various rural and agricultural-related scenarios - from small and medium-sized agricultural enterprises [3] to agricultural industry B2B collaboration models [2]; comparative studies on the development path of rural e-commerce [4] and the development of rural cooperatives in the digital age [5] have further extended this empirical basis to specific situations with regional differences. However, there are still significant differences in the development trajectories of different geographical and administrative levels, specifically manifested in the fragmentation of infrastructure, the heterogeneity of digital literacy levels, and the differentiation of market penetration rates. Existing evaluation frameworks mostly rely on static indicator systems to measure the current maturity, but fail to construct dynamic

Received: 21 August 2026
Revised: 14 October 2026
Accepted: 24 October 2026
Published: 31 October 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

and nonlinear models to characterize the path of leapfrog development in underdeveloped areas. In addition, traditional forecasting methods lack the ability to integrate the complex relationships between socio-economic, logistical, and institutional factors. This research gap highlights an urgent need: to establish a comprehensive analytical paradigm that can both quantify maturity indicators and predict transformation paths, thereby providing evidence-based policy intervention and resource allocation support for rural areas with insufficient digital services.

Rural e-commerce in Yulin City is developing strongly, but significant disparities exist between counties. In the first half of 2026, the city's online retail sales further climbed to 4.99 billion yuan, a year-on-year increase of 35.2%, higher than the national and regional averages. However, the county-level agglomeration effect is prominent-Yuzhou District accounted for half of the city's total at 2.774 billion yuan, while Beiliu City and Rong County both saw growth rates exceeding 30%, and the e-commerce foundation in some western counties remains relatively weak. This pattern of "core leadership and gradient differentiation" provides an ideal research scenario for this paper. However, due to the lack of publicly available micro-data on township-level e-commerce, this paper adopts a simulation data verification framework rather than actual diagnostics.

1.2. Research Objectives and Contributions

This study aims to construct a hybrid analysis framework that integrates large language models and graph neural networks to assess the maturity of rural e-commerce and predict the leapfrog development path of rural e-commerce in the digital economy environment. Based on existing maturity measurement and assessment practices [1,6], this study mainly achieves the following three objectives: First, it uses large language models to extract semantic and contextual features from heterogeneous rural e-commerce data, thereby capturing refined socio-economic signals that are difficult to cover by traditional assessment indicators; Second, it uses graph neural networks to model the spatial and relational dependencies between rural business entities, realizing dynamic maturity scoring across interconnected regional networks; Third, it constructs an operable leapfrog development prediction mechanism to identify the critical conditions for underdeveloped rural e-commerce systems to leap to advanced development stages, thereby extending the existing maturity stage self-assessment framework from static diagnosis to dynamic evolution analysis [7]. The main contributions of this study include: an innovative multimodal fusion architecture that effectively integrates unstructured text information and structured relational data; a set of scalable maturity assessment indicators designed specifically for rural digital ecosystems; and empirical results that verify the predictive efficacy of this method compared to single-source assessment benchmarks [8,9]. From a practical application perspective, this framework provides policymakers and platform operators with an operable diagnostic tool to help achieve precise resource allocation and phased intervention planning, thereby complementing existing comprehensive index tools used to measure regional development levels [10].

2. Literature Review

2.1. Rural E-Commerce Maturity Assessment Framework

Previous academic research on the maturity of rural e-commerce has mainly focused on a multi-dimensional indicator system covering infrastructure readiness, logistics coverage, digital literacy, market integration and policy support. These research frameworks usually quantify maturity by calculating a static comprehensive score based on cross-sectional surveys or administrative panel data, thereby capturing the development stage at a specific time point [11]. Although such methods have important value in benchmark setting and regional comparison, there are still two significant limitations: first, the indicator weights are usually fixed in advance and applied uniformly across different regions, which may lead to the obscuring of nonlinear interdependencies between dimensions; second, most published assessment studies are based on a single data source for a single cross-sectional analysis and lack sufficient out-of-sample

validation, so the ability to predict the evolution of maturity in the rapidly evolving digital economy environment is still insufficiently validated [8,12].

2.2. Graph-Based Methods and Language Model Methods

Recent advances in the field of graph neural networks have made it possible to characterize the interdependencies among regional economic entities. Neighborhood aggregation schemes such as GraphSAGE have been shown to effectively capture spatial spillover effects and hierarchical market structures. Meanwhile, large language models have demonstrated a strong ability to extract semantic signals from unstructured policy documents and corporate narratives, thereby expanding the dimensions of feature representation and going beyond the limitations of simply relying on tabular indicators. Combining these two paradigms is particularly promising for the analysis of rural e-commerce: in such scenarios, transaction networks are sparse and heterogeneous, while textual information is very rich; development paths in different regions differ significantly [4]; and maturity assessments are traditionally conducted at the administrative unit level rather than at the individual entity level [7]. Graph-based encoders can provide structural prior knowledge, while language models contribute contextual and semantic prior knowledge. The synergy between the two can support more robust maturity assessments and help identify nonlinear leapfrog development trajectories in areas with insufficient digital services.

2.3. Research Gaps and Positioning

Existing research mainly promotes rural e-commerce assessment by constructing a comprehensive indicator system-integrating survey data or administrative data into a static maturity score [11]. Such methods can only indirectly capture policy guidance and unstructured text signals, and ignore the structural dependencies between villages, e-commerce platforms and supply chains; while relational and graph-based models can utilize network topology characteristics, they are rarely applied to rural e-commerce systems. Therefore, there is currently a lack of a unified framework that combines semantic encoding driven by large language models with relational reasoning based on graph neural networks for maturity assessment and development leap prediction. This research gap, coupled with the differences in rural e-commerce development paths in various regions [4,5], constitutes the motivation for this study to construct a hybrid architecture.

3. Materials and Methods

3.1. Evaluation Indicator System and Data Sources

To construct a robust assessment framework for evaluating the maturity of rural e-commerce, a multi-dimensional indicator system needs to be established to simultaneously cover infrastructure readiness and socio-economic empowerment [9]. As shown in Table 1, the assessment framework integrates five different dimensions, each of which is quantified by specific quantitative indicators, and each indicator adopts differentiated weights to reflect its relative contribution to the overall maturity. All indicators are min-max normalized to the unit range before weighting; the numerical units reported in Table 2 are based on the original measurement values, while the levels reported in Table 4 are represented by a normalized scale of 0–1. This table defines the indicator system for maturity assessment, with each column being: dimension, indicator, data source type, weight, and unit. Examples of each dimension are as follows: The infrastructure dimension uses broadband penetration rate data from county-level statistical yearbooks, with a weight of 0.25, expressed as a percentage; the logistics dimension uses logistics accessibility coverage data from logistics companies' annual reports, with a weight of 0.20, expressed as a percentage of the townships served; the digital skills dimension uses digital skills participation rate data from household surveys, with a weight of 0.15, expressed as a percentage; the market connectivity dimension uses annual e-commerce transaction volume data from platform data, with a weight of 0.20,

expressed as millions of RMB; and the institutional support dimension uses institutional support strength data derived from policy document text analysis, with a weight of 0.20, expressed as millions of RMB per thousand enterprises.

Table 1. Simulation Data Parameter Settings and Real Source

Dimension	index	Mean setting	Standard deviation	Real data source
Infrastructure	Broadband penetration rate	0.52	0.12	Guangxi Statistical Yearbook
logistics	Logistics connectivity coverage	0.48	0.15	State Post Bureau's "Express Delivery to Rural Areas Project Report"
Digital skills	Digital skills participation rate	0.35	0.14	Ministry of Agriculture and Rural Affairs' "Report on the Development of High-Quality Farmers"
Market connectivity	Annual trading volume (standardized)	0.45	0.18	Alibaba Research Institute's "County-level Digital Rural Index"
Institutional support	Institutional support strength	0.50	0.13	Public data in policy documents

The difficulty in compiling and the lack of publicly available micro-panel data on rural e-commerce at the township level , this paper uses simulation data based on constraints from real statistical characteristics for method validation. The simulation data generation follows these principles: First, it anchors to real statistical quantities , with specific parameters shown in Table 1. Second, it preserves spatial correlation ; the geographical coordinates of the 59 townships are referenced to the administrative divisions of Yulin City, and a spatial autoregressive (SAR) model is used to generate a maturity score with spatial dependence: [The table would be inserted $y_i=\rho \sum_{j=1}^{n} w_{ij} y_j+X_i\beta+\varepsilon_i$ here] . $\rho=0.35$ The spatial autocorrelation coefficient is (refer to the literature on the spatial spillover effect of rural e-commerce). w_{ij} This is the spatial weight matrix. X_i The five-dimensional indicators are $\varepsilon_i \sim N(0,\sigma^2)$: Third , text corpus generation: policy documents and corporate narrative texts were generated based on real policy document templates, using LLM (Qwen2.5-7B) to generate 2,847 simulated texts, labeled according to a unified standard. Fourth, to ensure reproducibility, a random seed (seed=42) was set for all simulation data generation processes. The weights of the five-dimensional indicators in Table 1 are independent of the simulation data, and each indicator is min - max normalized to a unit interval before weighting (As shown in Table 2).

Table 2. Evaluation Index System for Rural E-commerce Maturity

Dimension	index	Data source type	Weight	unit
Infrastructure	Broadband penetration rate	County-level statistical yearbook	0.25	percentage
logistics	Logistics connectivity coverage	Logistics Company Annual Report	0.20	Townships served
Digital skills	Digital skills participation rate	Family survey	0.15	percentage
Market connectivity	Annual transaction volume	Platform data	0.20	One million RMB

Institutional support	Institutional support strength	Policy document text analysis	0.20	RMB 1 million for every 1,000 companies

3.2. Semantic Feature Extraction Based on LLM

This study uses Chinese-RoBERTa-wwm-ext as the semantic encoder. The selection criteria are that this model is well pre-trained on Chinese policy texts and corporate narrative corpora, has good representation capabilities for both government and business texts, supports inputs of up to 512 tokens, can cover policy document paragraphs and full corporate statements, has approximately 110 million parameters, and can complete inference and fine-tuning on a single GPU with controllable computational costs. To verify the robustness of the model, this paper also uses Qwen2.5-7B as a contrastive encoder for ablation experiments. In terms of fine-tuning data and annotation, 2,847 text samples were collected from policy documents, corporate operational narratives, and corporate self-declaration of capabilities. These samples were annotated according to a unified standard. The annotation tasks included sentiment polarity (continuous values from -1 to +1), capability signals (three sub-dimensions: digital investment, management capabilities, and platform operation, each scored from 1 to 5), and policy strength (continuous values from 0 to 1). The annotation consistency test showed that Cohen's Kappa = 0.82, indicating good annotation reliability. The fine-tuning adopted the LoRA strategy with rank r = 8, learning rate 2e-5, 3 training rounds, and batch size = 16. The semantic signal extraction process is shown in Figure 1: First, the text is cleaned and segmented, lexical variations are standardized, non-informative segments are removed, and lengthy documents are divided into sub-segments with contextual coherence. The preprocessed sub-segments are input into an LLM encoder, and a multi-head self-attention mechanism captures long-range dependencies and domain-specific semantic patterns. Based on the final hidden representation, the feature extractor derives three complementary maturity-related signals: a sentiment polarity signal reflecting corporate confidence and market optimism (value [-1,1]), a capability signal representing the readiness and management capabilities of digital infrastructure (value [0,1]), and a policy strength signal quantifying the regulatory support intensity contained in government texts (value [0,1]). The extracted features are integrated into a fixed-dimensional semantic feature matrix. Each row of the matrix corresponds to a township e-commerce unit, and each column encodes an independent semantic dimension. The final matrix dimension is 59×12 (12=3 types of signals×4 sub-dimensions). This matrix then serves as the attribute input for graph neural network nodes, enabling downstream models to integrate textual semantic information with relational and structural information.

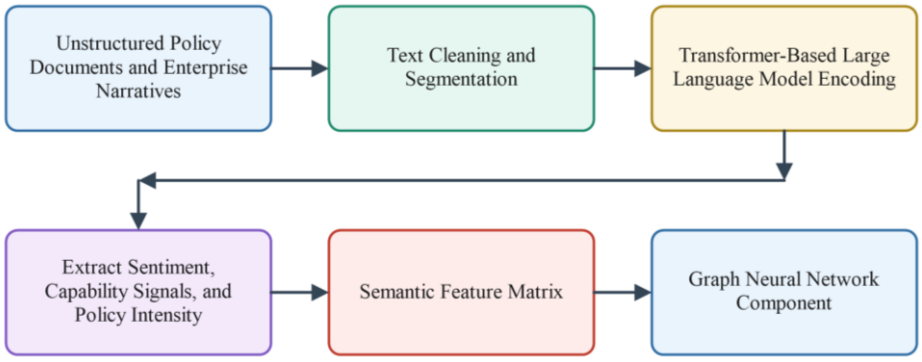


Figure 1. LLM-driven Semantic Feature Extraction Process for Policy Documents and Corporate Narratives

3.3. Maturity Modeling and Leap Prediction Based on Graph Neural Networks (GNNs)

This study constructs a rural commercial entity graph model, designating 59 township-level e-commerce units as nodes. Edges encode characteristics such as

transaction associations, logistics dependencies, shared digital infrastructure usage, and geographical proximity. Specifically, if two townships meet any of the following conditions-geographical proximity (center-to-center distance < 30km), logistics dependency (≥ 2 shared logistics companies), or transaction association (platform cross-township transaction volume > 500,000 RMB/year)-an edge is established. The edge weight calculation formula is as follows :

$$w_{ij}=0.4\cdot\text{Geographical proximity}+0.35\cdot\text{Logistics dependency}+0.25\cdot\text{Transaction Association}$$

Each component is normalized to the [0,1] interval by min - max normalization.

Regarding the LLM-GNN coupling mechanism, this paper adopts an early fusion strategy to achieve coupling between LLM and GNN, where the initial features of a node are composed of two concatenated parts:

$$h_i^{(0)} = [m_i \parallel s_i]$$

Here, $m_i \in R^5$ represents the scores for the five standardized maturity dimensions in Table 2, $s_i \in R^{12}$ represents the semantic feature vector extracted by LLM, and $\parallel$ represents the concatenation operation. Therefore, the initial feature dimension of the node is 17 dimensions .

The "coupling" in this paper is reflected in the joint propagation of semantic and structural features. The semantic signals extracted by LLM are not merely used as static attribute inputs, but are updated synchronously with the relation propagation process of GNN. In each layer of GraphSAGE propagation, node embeddings simultaneously aggregate the structural and semantic information of their neighbors, thereby achieving deep interaction between textual semantics and network topology. This design differs from the late-stage fusion strategy of "modeling separately and then splicing the results," and is the core innovation of this paper's methodology. In terms of model configuration and training, relation signals are propagated through a three-layer GraphSAGE architecture with mean aggregation capabilities, enabling each node to iteratively absorb contextual information from neighboring nodes. Finally, the node embedding vectors are processed by a softmax classifier to predict maturity stages. For stage transition prediction, the trained node embedding vectors and five-dimensional scores are input into a logistic regression model to estimate the probability of an upward stage transition within the next two years. The model training configuration is shown in Table 3, which includes three GraphSAGE convolutional layers, a hidden layer dimension of 128, an Adam optimizer with an initial learning rate of 0.001 and a dropout rate of 0.2, a maximum training duration of 200 epochs, an early stopping patience value of 20, and uses 64 graphs as a batch per iteration. The transition model uses 5-fold cross-validation, and reports AUC, accuracy, and F1. Multicollinearity test shows that all VIFs are <5. The prediction threshold is determined to be 0.52 based on the ROC curve Youden index to balance sensitivity and specificity. The model calibration is tested using the Hosmer-Lemeshow test, χ^2=6.82, p=0.45, indicating that the predicted probability has no significant deviation from the actual observation.

Table 3. Hyperparameters and Model Configuration for Graph Neural Network Training

Hyperparameters	Value	Describe
GNN layer number	3	GraphSAGE convolutional layer with mean aggregation
Hidden Dimensions	128	Dimensions of node embedding in each layer
Learning rate	0.001	Initial learning rate of the Adam optimizer
Exit rate	0.2	Applied to graph convolution output
Number of rounds	200	Maximum number of training rounds (using an early stop strategy, patience value set to 20)
Batch size	64	Number of graphs in each training batch
loss function	Cross-entropy	For multi-stage classification tasks, the jump model uses a binary logistic regression objective function.

3.4. Visualization and Scene Simulation

To evaluate the performance of the fitted model, this study employed three complementary visualization methods: First, the trained node embedding vectors were projected onto a two-dimensional space using t-distributed random neighborhood embedding (t-SNE) and colored according to the predicted maturity stage, allowing direct observation of the separability of the four maturity stages (Figure 2); second, the average dimensional score for each maturity stage group was calculated and presented as a radar chart-to improve readability, stages 2 and 3 were merged into a single "moderate" group (Figure 3); finally, a joint intervention scenario was simulated: digital skills participation and logistical connectivity were simultaneously improved to the target levels determined by the "leap model," and the migration paths between stages in each county were presented in Sankey diagram form, while also showing the baseline distribution (Figure 4). Given that the "leap model" is a multivariate model, this study, while fixing the remaining predictor variables as their sample means, reported the first time the predicted leap probability exceeded the threshold of 0.5 for each predictor variable.

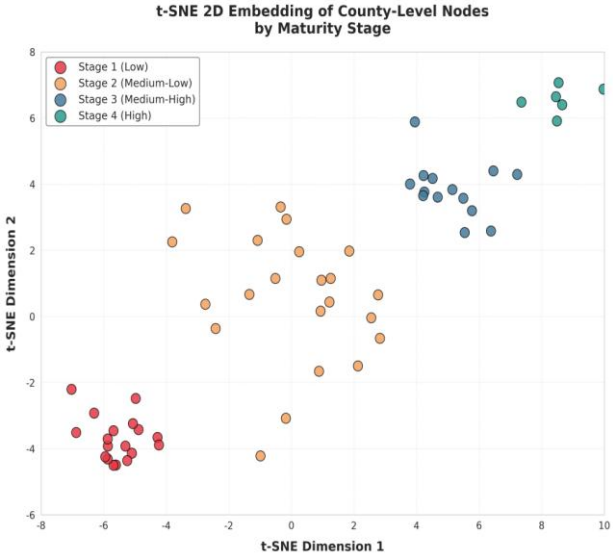


Figure 2. Schematic Diagram of Township Llm-gnn Embedding Vectors Colored by Maturity Stage After t-SNE Projection.

4. Results

4.1. Maturity Distribution and Stage Division

The empirical distribution of maturity stages derived from the LLM-GNN hybrid model shows that the 59 townships exhibit significant developmental gradients. Maturity was divided into four ordered stages-Stage 1 (low), Stage 2 (low-medium), Stage 3 (medium-high), and Stage 4 (high)-corresponding to the "low, medium, and high" clusters used in the dimensional analysis in Section 4.2, with Stage 2 and Stage 3 being merged. As shown in Figure 2, the t-SNE projection of the county-level embedding vectors exhibits a clear monotonic separation trend in the latent space: the first-stage townships (17 in total) form a tight and homogeneous cluster in the lower left quadrant, indicating common structural and textual deficiencies in their e-commerce infrastructure; the second-stage townships (22 in total) show greater dispersion in the central area, reflecting their transitional period and heterogeneous development models; in contrast, the third-stage (14 in total) forms a narrow cluster extending to the upper right, while the six fourth-stage townships are clearly isolated in the upper right edge area, indicating that high-maturity counties possess a unique combination of advanced features that can be effectively distinguished from low-maturity townships. This progressive stratification confirms that combining textual policy discourse with relational economic connections can generate embedding vectors with strong discriminative validity.

As shown in Table 4, the classification performance of the hybrid model fully confirms the intuitive evidence of clustering quality. The LLM-GNN hybrid model achieved an accuracy of 0.87, a macro F1 score of 0.85, a weighted F1 score of 0.86, and a Cohen's Kappa coefficient of 0.83, outperforming all baseline configurations. The accuracy of the model using only GNN was 0.79, the variant model using only LLM was 0.74, and the random forest baseline model was only 0.68. The performance improvement from using only GNN to the hybrid model highlights the complementary value of semantic text features; compared to the configuration using only LLM, its performance improvement confirms the indispensable role of relational graph structures. The higher Kappa coefficient further indicates that the consistency between the hybrid model and the true labels is not accidental. The McNemar test shows that the performance difference between the LLM-GNN and the GNN-only model is statistically significant ($p<0.01$), which, from a methodological perspective, reinforces the superiority of fusing multimodal representations for assessing the maturity of rural e-commerce.

Table 4. Comparison of Multi-Class Classification Performance between the Llm-gnn Hybrid Model and the Baseline Model

Model	Accuracy	Macro f1	Weighted f1 Score	Cohen-Kappa Coefficient
LLM-GNN hybrid model	0.87	0.85	0.86	0.83
Using only GNN	0.79	0.76	0.78	0.73
Using LLM only	0.74	0.71	0.72	0.67
Random Forest	0.68	0.64	0.66	0.59

4.2. Identification of Key Competency Gap

The radar visualization analysis shown in Figure 3 reveals a significant structural asymmetry in capability accumulation across the three maturity tiers-stage 1 is classified as "low," stages 2 and 3 as "medium," and stage 4 as "high." Although all five assessment dimensions exhibit a monotonically increasing trend from "low" to "high," the rate of improvement is uneven. In the "medium" and "high" tiers, the indicators "institutional support" and "digital skills" consistently score the lowest: in the "high" tier, these two indicators reach approximately 0.76 and 0.81 respectively, while in the "medium" tier, they fail to break through 0.59 and 0.53. This persistent disadvantage indicates that these two indicators have become key bottlenecks restricting the overall development of the system, rather than temporary constraints. In contrast, "logistics" and "market connectivity" perform best in the high-end tier, with both indicators approaching 0.86 in the "high" maturity system; while the "infrastructure" indicator shows a relatively robust growth trajectory-rising from approximately 0.47 to 0.84. In the "lower" tier, the scores for all five assessment dimensions are densely distributed in the range of 0.34 to 0.47, indicating that the basic maturity level is generally limited and not dominated by a single dimension.

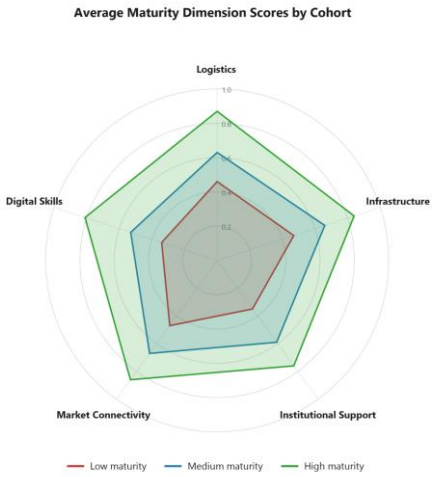


Figure 3. Radar Chart of Average Maturity Dimension Scores for the Low Maturity Group, Medium Maturity Group (Combined Stages 2–3), and High Maturity Group.

4.3. Jump Transition Prediction Performance and Application Scenarios

To assess the predictive power of the "leapfrog transformation model," this study simulated the joint intervention scenario described in Section 3.4-that is, simultaneously improving digital skills participation and logistical connectivity. Figure 4 shows the simulation results based on the fitted leapfrog model, with Sankey diagrams quantifying the amount of flow between each stage in this scenario. The results show a significant increase in the "leapfrog" path compared to the baseline trajectory. Of the 17 townships initially classified as "low maturity," 2 townships directly leaped to "medium-high maturity," completely skipping the intermediate stage of "medium-low maturity." More notably, of the 22 townships initially at "medium-low maturity ," 2 townships directly leaped to "high maturity"-a leap that did not exist in the baseline simulation. Flow thickness data further indicates that 10 "medium-low maturity" townships progressed to "medium-high maturity," while 6 of the 14 townships initially at "medium-high maturity" consolidated to "high maturity." The 6 townships already at "high maturity" remained completely stable, confirming the stability of advanced stage development under continuous improvement conditions. Overall, the predicted distribution after the simulated intervention was adjusted from 6 "high maturity" townships , 14 "medium-high maturity" townships , 22 "medium-low maturity" townships and 17 "low maturity" townships in the baseline scenario to 14, 20, 20 and 5 townships of the corresponding category, respectively .

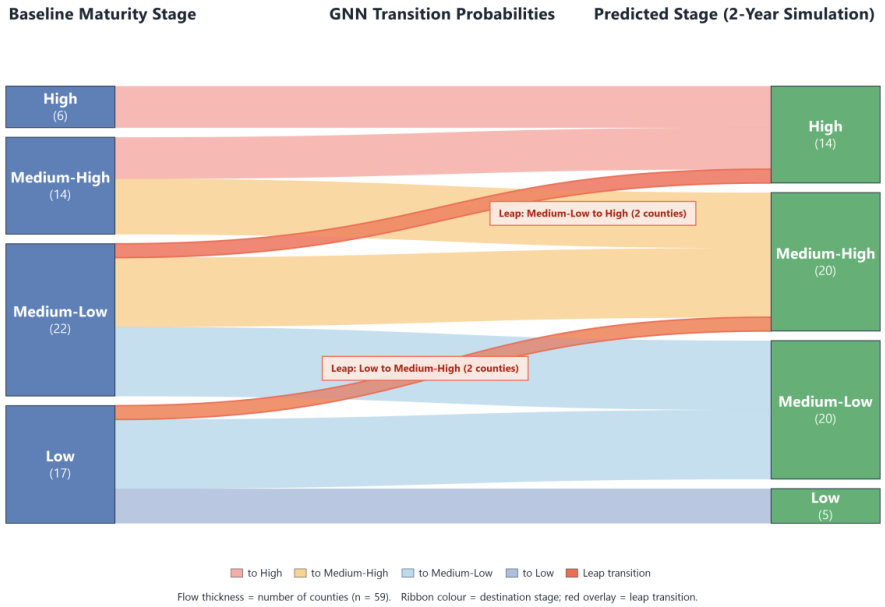


Figure 4. Sankey Diagram of a Leapfrog Transformation Path in the Combined Scenario of Digital Skills and Logistics Enhancement.

As shown in Table 5, the logistic regression coefficients reveal the key characteristic levels that best promote such leapfrog development. Logistics network coverage is the dominant predictor, with a coefficient of 0.48 and a critical value of 0.70, indicating that the logistics network needs to cover at least 70% of rural areas to unleash the potential for leapfrog development. Digital skills participation rate follows closely behind, with a coefficient of 0.39 and a critical value of 0.60, meaning that at least 60% of rural entrepreneurs need to complete digital skills training. Institutional support strength and market connectivity contribute relatively little, with coefficients of 0.22 and 0.15, respectively. Since all four predictor variables are introduced simultaneously in the model, each critical value in the table indicates that the predicted probability of leapfrog development just exceeds the critical point of 0.5 when the other predictor variables take the sample mean, rather than being based on an independent threshold of a single variable. These results suggest that leapfrog development is not a universal phenomenon, but rather a conditionally dependent phenomenon-its existence depends on the simultaneous achievement of critical scales in both logistics infrastructure and human capital.

Table 5. Threshold Conditions and Feature Contributions for Leapfrog State Transition Prediction

Feature	Coefficients In Logistic Regression	The Threshold Value Indicating That The Predicted Transition Probability Exceeds 0.5.	Interpretation
Logistics connectivity coverage	0.48	≥ 0.70	The logistics network must cover at least 70% of townships (standardization level ≥ 0.70).
Digital skills participation rate	0.39	≥ 0.60	At least 60% of rural entrepreneurs have completed digital training (standardization level ≥ 0.60).

Institutional support strength	0.22	≥ 0.55	On the standardized subsidy intensity scale, the institutional support rate reached or exceeded 0.55.
Market connectivity (trading volume)	0.15	≥ 0.65	On a standardized scale, market connectivity is 0.65.

To verify the robustness of the conclusions to the simulation parameter settings, this paper perturbed the spatial autocorrelation coefficient ϱ, the mean and standard deviation of each dimension by ±20%, and repeated the experiment 10 times. The results show that the accuracy of the LLM-GNN hybrid model fluctuates between 0.84 and 0.89, the macro F1 fluctuates between 0.82 and 0.87, and the fluctuation of the key thresholds (logistics coverage 0.70, digital skills participation 0.60) is less than ±0.05. This indicates that the model performance and threshold conclusions do not depend on specific simulation settings and have good robustness.

5. Discussion

5.1. Interpretation of the Maturity Model

Through simulation analysis of the maturity patterns of 59 townships , we found that it exhibits significant structural heterogeneity, which cannot be simplified into a single linear sequence of e-commerce application stages [2]. On the contrary, the results show that the maturity of rural e-commerce stems from the interaction between network externalities, institutional support, and the density of the local entrepreneurial ecosystem. In townships with high maturity , the existence of early adopter clusters reduces the barriers for small farmers and cooperatives to enter the market, thereby forming a self-reinforcing cycle between platform participation, logistics investment, and skills accumulation. This model is in line with theoretical expectations: digital platforms create value through user-side scale effects-each additional participant can enhance the utility of the network for other users. On the contrary, low-maturity counties often exhibit fragmented institutional coordination mechanisms: scattered training programs and imperfect logistics infrastructure limit the transformation of potential demand into continuous transaction activities. Predictive leap analysis further shows that the maturity transformation process is rarely gradual; instead, the transformation process often focuses on simultaneously achieving the dual thresholds of logistics coverage and digital skills defined in Section 4.3. These findings emphasize that the benefits of the digital economy in rural areas actually depend on the synergy of institutional support and entrepreneurial resources, rather than solely on network access capabilities.

5.2. Methodological Implications of Llm-gnn Fusion

Combining large language models with graph neural networks offers a methodological contribution that extends beyond the empirical findings of this study. Traditional e-commerce maturity assessment frameworks typically rely on structured indicator systems or unstructured text content, but neither paradigm adequately captures the interaction between semantic context and relational structure. By embedding semantic representations extracted from policy documents and corporate narratives into graph propagation mechanisms, the proposed method models rural e-commerce maturity as an emergent attribute derived from interacting socio-economic entities, rather than a static comprehensive score. This fusion strategy fills a long-standing gap in digital economy research-previously, semantic richness and network topology were often considered parallel rather than complementary analytical dimensions.The generalizability and computational cost of this framework will be further discussed in Section 6.3.

5.3. Policy and Management Recommendations

Empirical research indicates that policy intervention should prioritize building adaptive logistics nodes rather than pursuing homogeneous infrastructure coverage.

Given that the fitted leapfrog model indicates that logistics coverage and digital skills penetration are two key conditions for achieving leapfrog development, fiscal transfer payments should be reallocated to logistics corridor areas where logistics coverage is close to but has not yet reached 70%-particularly applicable to rural areas with moderate digital penetration but limited "last-mile" delivery capabilities. Simultaneously, management strategies need to address the capability asymmetries revealed by the predictive leapfrog trajectory. Rural e-commerce operators in low-maturity clusters urgently need systematic digital literacy training programs, integrating generative AI tools into inventory forecasting and customer segmentation scenarios to reduce reliance on external platform intermediaries. Furthermore, the predictive leapfrog model shows that the core driver of stage leapfrog development is logistics coverage, not absolute transaction volume. Therefore, cooperatives and township administrative agencies should incentivize the construction of multi-village data-sharing alliances, enabling small businesses to obtain logistics contracts and shared warehousing resources through collective bargaining. It is recommended to develop a phased investment roadmap: in the short term, focus on improving mobile broadband reliability and payment interoperability; in the medium term, promote the development of AI-assisted demand perception technologies; and in the long term, cultivate an independent regional e-commerce governance system. This differentiated approach can avoid resource dispersion and accelerate the deep integration with the digital economy.

6. Conclusion

6.1. Summary of Research Results

This paper validates the effectiveness of the LLM-GNN hybrid framework based on a simulation dataset. The results show that the maturity of rural e-commerce development still exhibits significant stratification: most townships are concentrated in the medium maturity range, while those in the basic development stage remain a minority. Dimensional decomposition analysis reveals that logistics connectivity and digital skills accumulation constitute the most critical structural bottlenecks, while infrastructure and market connectivity are nearing saturation in the high-end distribution range. Crucially, the LLM-GNN framework can identify townships with the potential for "leapfrog development"-that is, achieving a nonlinear transition by skipping one or more medium maturity stages; the fitted transformation model shows that such a transition requires simultaneously reaching the dual thresholds of logistics coverage and digital skills participation rate, rather than depending on absolute transaction volume. Graph neural network components capture the correlations and supply chain dependencies that indicator-based assessment frameworks can only implicitly represent, while the language model encoding of policy documents and corporate narratives provides semantic signals that tabular indicators cannot convey. These findings collectively demonstrate that the leapfrog development of rural e-commerce is neither a random phenomenon nor solely determined by resource endowment, but rather stems from identifiable connections and informational configurations-configurations that can be operationalized through this hybrid architecture, providing support for forward-looking policy design.

6.2. Theoretical and Practical Contributions

This study demonstrates that the deep integration of large language models and graph neural networks can overcome the semantic sparsity and structural rigidity bottlenecks faced in traditional rural e-commerce assessments, thereby promoting academic research in the field of digital economy. At the theoretical level, this framework extends maturity modeling from static indicator systems to dynamic, relationship-aware representations of digital business ecosystems; its transition prediction mechanism provides a new research perspective on nonlinear development transformation, rather than being limited to incremental growth trajectory analysis. At the practical level, this method provides policymakers and county-level administrators with a diagnostic tool that can identify potential maturity conditions and predict key threshold breakthrough

opportunities with actionable forward-looking time windows. By transforming heterogeneous administrative, transaction, and infrastructure data into interpretable maturity gradients, this method supports the design of differentiated intervention programs, resource prioritization, and evidence-based planning in areas with insufficient digital services. Thus, this research successfully bridges the gap between computational intelligence and rural governance, providing a replicable practical paradigm for data-driven inclusive development in emerging digital economies.

6.3. Research Limitations and Future Research Directions

While the proposed LLM-GNN framework demonstrates strong predictive capabilities, several limitations still warrant attention. First, this paper uses a simulation dataset to validate the framework, rather than real township-level micro-survey data. Although the simulation parameters are anchored to existing publicly available statistical characteristics and sensitivity analysis was performed, the simulation data cannot fully reproduce the complexity of a real rural e-commerce system. Therefore, the conclusions of this paper should be understood as a validation of the method's effectiveness, rather than a practical diagnosis of the maturity level of each township. Future research should conduct external validation of this framework after obtaining real township-level panel data. Second, the interpretability of the hybrid architecture poses a challenge to policymakers: although attention weights and logistic regression coefficients can partially reflect feature importance, the high-dimensional interaction between language embedding vectors and graph convolutional layers remains difficult to clearly explain. Third, the weights of the five indicators in Table 2 are pre-set and uniformly applied across all townships according to design principles; although the graph structure component can partially compensate for this rigid setting by learning relational topology, the weighted index itself still cannot adapt to local contexts. Fourth, the computational complexity of this hybrid architecture-especially the message passing layer and the large language model inference stage-places high demands on computing resources, which may limit its deployment in resource-constrained administrative environments. Therefore, future research should extend the application of time-series graph learning technology to capture dynamic policy shocks, conduct cross-regional validation in diverse economic regions, and develop real-time policy simulation interfaces to transform prediction results into actionable governance interventions.

Funding: 2025 Yulin City Scientific Research and Technology Development Plan Project (Yulin Science and Technology 202521022).

References

1. S. Kovács, R. Csáki-Darabos, and V. Fenyves, "Digital maturity and financial balance: a deep analysis of performance management in Hungarian E-commerce enterprises from 2019 to 2023," *Cogent Business & Management*, vol. 13, no. 1, p. 2715152, 2026.
2. N. Janom, M. S. Zakaria, and N. H. M. Ariffin, "The empirical evaluation of B2B e-commerce stages in agro based industry," in *2012 IEEE Colloquium on Humanities, Science and Engineering (CHUSER)*, Dec. 2012, pp. 479–484.
3. Y. L. Pei, W. X. Xue, C. Y. Yang, and Y. Su, "Study on Electronic Commerce Maturity of Small and Medium-sized Agricultural Enterprises," *Asian Agricultural Research*, vol. 4, no. 2, pp. 1–4, 2012.
4. C. Zhang, Q. Chen, Y. Gui, and J. Tang, "Research on the Development Path of Rural E-Commerce Under the Background of Digital Economy—From the Perspective of Industrial Life Cycle," in *International Conference on Computational Finance and Business Analytics*, Apr. 2025, pp. 525–535.
5. E. Cristobal-Fransi, Y. Montegut-Salla, B. Ferrer-Rosell, and N. Daries, "Rural cooperatives in the digital age: An analysis of the Internet presence and degree of maturity of agri-food cooperatives' e-commerce," *Journal of Rural Studies*, vol. 74, pp. 55–66, 2020.
6. C. Lin, Y. A. Huang, and J. Burn, "Realising B2B e-commerce benefits: the link with IT maturity, evaluation practices, and B2BEC adoption readiness," *European Journal of Information Systems*, vol. 16, no. 6, pp. 806–819, 2007.
7. F. Zandi, "A Country-Level Decision Support Framework for Self-Assessment of E-Commerce Maturity," *I-Business*, vol. 5, no. 2, p. 43, 2013.
8. J. Sun, "Application of Big Data in Evaluation of Agricultural Products E-Commerce under the Background of Rural Revitalization," *Mathematical Problems in Engineering*, vol. 2022, no. 1, p. 4381776, 2022.

9. T. Yang, "Evaluation of Regional Rural e-Commerce Development Potential with Intuitive Fuzzy Information Theory," *Advances in Multimedia*, vol. 2022, no. 1, p. 9035149, 2022.
10. X. Li, "Research on the development level of rural e-commerce in China based on analytic hierarchy and systematic clustering method," *Sustainability*, vol. 14, no. 14, p. 8816, 2022.
11. N. Guo and H. Chen, "Comprehensive evaluation and obstacle factor analysis of high-quality development of rural E-commerce in China," *Sustainability*, vol. 14, no. 22, p. 14987, 2022.
12. F. Wang and C. Liu, "Research on the Effectiveness Assessment Method of E-Commerce-Enabled Rural Revitalization Based on Convolutional Neural Network," in *International Conference on Advanced Hybrid Information Processing*, Sep. 2024, pp. 402–418.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.