Journal of Engineering, Project, and Production Management
2026, 16(5), 2026-0033

HNN-DIC: Spatio-Temporal Evaluation of SME Digital Innovation Capabilities Using Industrial Internet Platform Data

Ming Guo[1], Qi Kai[2], and Lijing Kang[3]
[1] Postgraduate Student, Business Administration, Inje University, Gimhae, South Gyeongsang Province, South Korea
[2] Postgraduate Student, Business Administration, Dong-eui University, Busanjin-gu, Busan, South Korea
[3] Ph.D., School of Civil Engineering and Architecture, Guangxi Minzu University, China, E-mail: lijing-kang@outlook.com (corresponding author).

Project Management
Received April 28, 2026; revised June 4, 2026; accepted July 3, 2026
Available online July 12, 2026

Abstract: In the context of the digital economy, scientifically assessing the Digital Innovation Capability (DIC) of Small and Medium-Sized Enterprises (SMEs) is essential for achieving sustainable competitive advantages. Traditional evaluation approaches often struggle to process multi-source, high-dimensional and nonlinear digital data and lack dynamic assessment capability and model interpretability. To address these limitations, this study proposes a data-driven intelligent evaluation framework that combines industrial internet platform data with deep learning techniques to enable dynamic, interpretable SME-DIC assessment. Specifically, a Hybrid Neural Network-based Digital Innovation Capability evaluation model (HNN-DIC) is developed by integrating a One-Dimensional Convolutional Neural Network (1D-CNN) and a Long Short-Term Memory (LSTM) network. The proposed framework simultaneously captures local spatial correlations and long-term temporal evolution patterns among digital innovation indicators. Based on DIC, a multidimensional evaluation system containing nine core indicators is constructed across four dimensions: technological embedding, dynamic adaptation, resource coordination, and innovation output. The proposed model is trained and validated using two years of panel data collected from 500 manufacturing SMEs operating on an industrial internet platform. Experimental results demonstrate that HNN-DIC achieves a coefficient of determination (R^2) of 0.931 on the test set, with a Mean Absolute Error (MAE) of 1.72 and a Root Mean Square Error (RMSE) of 2.28, significantly outperforming benchmark models such as Random Forest and XGBoost. Furthermore, SHapley Additive exPlanations (SHAP) analysis identifies “data analysis activity” and “external collaboration breadth” as the most influential drivers of SMEs digital innovation capability. The study contributes a domain-specific hybrid intelligent assessment framework for SMEs-DIC evaluation and provides an interpretable decision-support tool that can be deployed in cloud-based digital platform ecosystems to support enterprise self-diagnosis, strategic decision-making, and policy formulation.

Keywords: Digital innovation capability; small and medium-sized enterprises; digital platform; hybrid neural network.

DOI 10.32738/JEPPM-2026-0033

---

1. Introduction

Against the backdrop of the global digital economy, digital transformation has become a critical pathway for small and medium-sized enterprises (SMEs) to enhance competitiveness and achieve sustainable development. Digital Innovation Capability (DIC), as the comprehensive dynamic capacity for enterprises to integrate digital technologies, coordinate internal and external resources, respond to market changes, and ultimately achieve value creation, serves as the core metric for measuring the depth and effectiveness of their digital transformation. Scientifically and accurately assessing SMEs DIC not only helps enterprises identify weaknesses and optimize resource allocation but also provides policymakers with decision-making support to grasp industrial innovation trends and implement targeted empowerment measures.

However, traditional SME innovation capability assessments often rely on cross-sectional surveys, expert scoring, or financial indicator analysis (Liu, 2012). These methods reveal significant limitations when applied to the digital economy: First, they are highly subjective and rely on outdated data, making it difficult to capture the real-time, dynamic behavior of enterprises operating on digital platforms. Second, they lack the capacity to manage complex nonlinear relationships, as DIC formation results from the interplay of multidimensional factors, including technology, organization, and ecosystem factors, which traditional linear models struggle to adequately capture. Finally, they rely on coarse-grained data, often lacking fine-grained, time-series behavioral data generated from genuine business scenarios, leading to discrepancies between assessment results and actual conditions. Concurrently, digital infrastructure represented by industrial internet platforms is becoming increasingly prevalent. These platforms continuously generate massive, multi-source, high-dimensional operational data, providing an unprecedented data foundation for objectively and dynamically assessing enterprise DIC (Hu and Liao, 2025). Therefore, leveraging this real-time data alongside advanced computational intelligence technologies to construct an accurate, dynamic, and interpretable SMEs-DIC assessment model has become a critical academic and practical challenge requiring urgent resolution.

Existing research on this topic primarily unfolds across three dimensions. First, the essence and structure of SMEs digital innovation capabilities. Scholars generally recognize DIC as a multidimensional construct encompassing multiple levels, including technology adoption and application, process digital transformation, and even business model innovation (Nie, 2011; He, 2012; Yin et al., 2011; Liu et al., 2018). Second, the mechanism through which digital platforms influence innovation capabilities. Research indicates that platforms profoundly influence corporate innovation processes and effectiveness by providing plug-and-play infrastructure, reducing transformation costs, and facilitating knowledge spillovers and collaborative network formation (Yu and Ma, 2024; Fu, 2014; Xiao and Fu, 2023; Yang, 2022; Shan et al., 2023). Third, the application of neural networks in the assessment of complex systems. Particularly in fields like environmental monitoring and supply chain optimization, models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) have become effective tools for handling high-dimensional, nonlinear evaluation tasks due to their powerful nonlinear fitting and sequence modeling capabilities (Ni and Li, 2021; Kuang and Guo, 2025; Yu and Liu, 2020; Yu, 2020).

Despite the fruitful outcomes of existing research, studies that organically integrate these three strands remain scarce. Specifically, no study has systematically utilized real, time-series, multidimensional data from digital platforms to design a resolute hybrid neural network model. This model simultaneously captures both the spatial correlations (complex interactions among indicators at the same point in time) and temporal evolution (dynamic trends of indicators over time) inherent in DIC formation, thereby enabling dynamic, granular, and interpretable assessment of enterprise DIC.

To address this gap, this study formulates three explicit research questions. How can industrial internet platform data be operationalized into a multidimensional, dynamic evaluation system for SMEs-DIC grounded in dynamic capability theory? Does a hybrid neural network combining 1D-CNN (spatial features) and LSTM (temporal dependencies) achieve significantly higher predictive accuracy for SMEs-DIC compared to benchmark models? Which specific indicators most strongly drive SMEs-DIC, and can SHapley Additive exPlanations (SHAP)-based interpretability provide actionable diagnostic insights? To answer these questions, we construct a Digital Platform Data-Based Hybrid Neural Network SMEs-DIC Evaluation Model (HNN-DIC). Its primary contributions are theoretically proposing an integrated "Platform Environment, Capability Performance-Intelligent Evaluation" analytical framework and operationalizing dynamic capability theory into a quantifiable evaluation metric system, methodologically designing a hybrid 1D-CNN + LSTM architecture to simultaneously extract spatial and temporal dependency features, and practically employing real-world industrial internet platform panel data with SHAP analysis for interpretable diagnostics. This enables evaluation results not only to output scores but also to diagnose key drivers and improvement pathways, ultimately transforming into an intelligent decision-support tool for enterprises, governments, and platform operators.

2. Evaluation of Model Construction

2.1. Overall Framework

The SMEs-DIC assessment framework developed in this study follows a core logic of “data-driven, intelligent modeling, value output,” forming a closed-loop system comprising “multi-source data aggregation, hybrid neural network evaluation, result interpretation and application.”

The Multi-source Data Aggregation module serves as the foundational support layer of the assessment framework. Designed around the principles of “data comprehensiveness, controllable quality, and temporal continuity,” this module establishes a multidimensional data-collection and standardized-processing system to provide high-quality inputs for subsequent modeling. Leveraging the real-time data advantages of digital platforms and the complementary nature of multi-source data, three core data sources are identified: ① Core Data Sources (direct Application Programming Interface (API) connection to digital platforms): Captures real-time dynamic data from industrial internet platforms, including enterprise cloud service usage behavior, IoT device connectivity status, and cross-entity collaboration records, ensuring timeliness and objectivity. ② Supplementary Data Sources (enterprise public databases): Extracts authoritative public information such as digital-related patent applications and revenue share from new products, supplementing quantitative metrics for innovation output. ③ Supplementary Data Sources (Lightweight Targeted Surveys): Focusing on qualitative information difficult to collect automatically via platforms, such as user feedback adoption mechanisms and internal innovation process optimization, these surveys enable the quantitative conversion of qualitative data through structured questionnaires.

To eliminate data heterogeneity and noise interference, a standardized preprocessing workflow is applied: First, missing values are imputed using the K-Nearest Neighbors (KNN) algorithm, leveraging feature similarity among samples to ensure the plausibility of the imputed values. Second, outliers are identified and removed using the 3σ rule to prevent extreme data points from distorting evaluation results. Finally, Min-Max normalization maps all indicators to the [0,1] range, eliminating the impact of dimensional differences on model training. The preprocessed data form a panel data tensor with an eight-quarter time step (2021Q1-2022Q4) and nine core indicators as feature dimensions, meeting the hybrid neural network’s input data structure requirements.

The Hybrid Neural Network Evaluation Module, serving as the framework's core computational layer, focuses on the objectives of "deeply mining multidimensional features, accurately modeling nonlinear relationships, and stably outputting DIC scores." It employs a hybrid architecture of "parallel feature extraction followed by serial feature fusion" to comprehensively capture the spatiotemporal evolution characteristics of SMEs-DIC. The formation of SMEs-DIC depends both on the interaction of multiple indicators at the same point in time (spatial features) and on the dynamic trends of indicators over time (temporal features). To address this, the module employs a dual-branch parallel design. ① Spatial Feature Extraction Branch (1D-CNN): Utilizes a two-layer convolutional neural network structure. The first layer contains 32 convolutional kernels of size three, while the second layer contains 64 convolutional kernels of size three. The ReLU activation function enhances nonlinear fitting capabilities. A MaxPooling layer compresses feature dimensions while preserving key information. Its core function is to uncover local interactive dependencies and spatial correlation patterns among the nine evaluation metrics at the same time point. ② Time-Series Feature Modeling Branch (LSTM): Utilizes a two-layer Long Short-Term Memory network with 64 neurons per layer. Leveraging LSTM's gating mechanisms (input gate, forget gate, output gate), it captures the dynamic evolution patterns and long-term temporal dependencies of each indicator across eight quarters, addressing the long-term dependency challenges in sequence data that traditional models struggle to manage.

In the Concatenate layer, the spatial features from the 1D-CNN output and the temporal features from the LSTM output are dimension-concatenated to form a "spatial-temporal" fusion feature vector. Subsequently, two fully connected layers (containing 128 and 64 neurons respectively, with ReLU activation functions) achieve high-order nonlinear feature combination. Finally, a linear activation neuron outputs the enterprise's comprehensive DIC score at the current time point. Model training uses the Adam optimizer (initial learning rate 0.001) with Mean Squared Error (MSE) as the loss function. Early Stopping (patience threshold set to ten) monitors the validation set loss and halts training when the validation loss remains unchanged for ten consecutive epochs. This effectively prevents overfitting and ensures the generalization capability of evaluation results.

The results interpretation and application module serves as the value transformation layer of the framework. It achieves a closed-loop output of "quantitative evaluation, mechanism Interpretation, and multi-scenario application," ensuring both the scientific rigor of evaluation results and their practical applicability. Final outputs include: ① A quantitative scoring system comprising a composite DIC total score and sub-scores across four dimensions. Technical Embedding Depth (TE), Dynamic Adaptive Agility (DA), Resource Coordination Efficiency (RC), Innovation Output Performance (IO). Supporting enterprises in conducting comprehensive self-assessments and dimensional breakdown analyses. ② Visual diagnostic reports, presenting enterprise DIC performance positioning and strengths/weaknesses through radar charts (dimensional benchmarking), trend curves (temporal evolution), and feature importance heatmaps (key influencing factors). ③ Interpretable analysis results: Quantifies each indicator's contribution and directional impact on DIC scores using the SHAP (SHapley Additive exPlanations) method, revealing the intrinsic "indicator-capability" mechanism to provide precise improvement guidance.

2.2. Metric Framework

The target layer is Digital Innovation Capability (DIC), which represents an enterprise's comprehensive ability to integrate digital technologies, coordinate internal and external resources, respond to market changes, and deliver innovative outputs (Khurana, Ghura, and Dutta, 2025). The criterion layer (dimensions) is decomposed into four core dimensions based on the resource integration, dynamic adaptation, and the value output logic of dynamic capability theory. Technology Embedding Depth (TE) (the technological foundation of digital innovation, reflecting integration and application capabilities of technical resources), Dynamic Adaptation Agility (DA) (responsiveness to market changes, indicating levels of dynamic adjustment and iterative optimization), Resource Coordination Efficiency (RC) (synergy in leveraging internal and external resources, highlighting resource allocation effectiveness within platform ecosystems), and Innovation Output Performance (IO) (the ultimate value manifestation of DIC, serving as the target variable for model training). The indicator layer selects three core, quantifiable metrics for each dimension, totaling nine indicators, forming a three-tiered, progressive evaluation system: the objective layer, the guideline layer, and the indicator layer.

The specific definitions, measurement dimensions, and data sources for each indicator are as follows. Technology Embedding Depth (TE) measures the intensity of a company's absorption, application, and deployment of digital technologies, serving as the foundational support for digital innovation. Cloud Service Penetration Index (TE1) is defined as the breadth and depth of cloud service adoption in a company's digital transformation. It is measured as "number of cloud service types used (e.g., IaaS, PaaS, SaaS) × Proportion of core business operations migrated to the cloud." Data is sourced from direct Application Programming Interface (API) connections to industrial internet platforms. TE2 is a Data Analysis Activity that reflects the intensity of data-driven decision-making practices within enterprises. The measurement method is "Monthly BI (Business Intelligence) tool invocation count × Average daily viewing duration of data dashboards." Data is sourced from the platform's behavioral log database. TE3 is Device Connection Density, measuring the digital transformation level of production processes. The calculation method is "Total number of production line devices and sensors connected to the platform × Average daily device online rate." Data is collected in real-time by the platform's device management module.

Dynamic Adaptive Agility (DA) reflects an enterprise's ability to rapidly respond to market demands and user feedback while iteratively optimizing. DA1: Market Response Speed, defined as the timeliness of adjusting product strategies based on market dynamics. Calculated as "Average delay in days from identifying market demand changes via platform sentiment analysis tools to completing product strategy adjustments" (inverse indicator converted to positive by taking the reciprocal). Data combines platform sentiment analysis and records of corporate strategy adjustments. DA2 is the iteration update frequency, reflecting the continuous optimization capability of products/services. This measurement method is "the number of major functional updates released for software versions or hardware products within a year." Data is sourced from the company's product update logs and platform filing information. DA3 is and measures the effectiveness of user needs in driving innovation. Calculated as "the proportion of valid user feedback collected via the platform client that is verified as feasible and incorporated into product development." Data is matched and calculated using the platform's user feedback module and corporate R&D records.

Resource Coordination (RC) measures an enterprise's efficiency in integrating external resources and conducting collaborative innovation within the digital platform ecosystem. RC1: External Collaboration Breadth, Reflects participation in open innovation. Measured as "the number of digital-related joint projects collaboratively undertaken with other platform entities (universities, research institutions, suppliers, customers) within the year." Data sourced from the platform's collaborative project management database. RC2 measures the intensity of knowledge acquisition and the company's capacity to absorb external digital knowledge. It is calculated as "the number of times industry solutions were downloaded from the platform × cumulative time spent participating in online technical communities," sourced from the platform's knowledge service module logs. RC3 is the digital asset reuse rate and measures the ecological value and the extent of sharing of the enterprise's digital innovation outcomes. The measurement method is "cumulative number of times the enterprise's self-developed digital algorithms, models, and APIs were invoked by other enterprises within the platform." Data is collected from the platform's interface invocation statistics module.

Innovation Output (IO) serves as the direct value representation of DIC and is the target variable for model training, measuring the economic and operational benefits generated by digital innovation. IO1 represents the proportion of revenue from new products, gauging the market value of digital innovation. The calculation method is: "Revenue generated by new digital products/services launched within the past twelve months as a percentage of the company's total operating revenue." Data is sourced from corporate financial statements and platform revenue filing information. IO2 represents the number of digital-related patents, reflecting the accumulation of technological innovation achievements. It is calculated as "the number of invention patents directly related to digital technologies (e.g., artificial intelligence, Internet of Things, big data) that the enterprise applied for and disclosed within the year." Data is sourced from the national patent database and publicly disclosed enterprise information. IO3 measures digital process efficiency improvements and quantifies operational innovation benefits. The methodology is defined as "the percentage reduction in time for key business processes (e.g., order fulfillment, production scheduling, inventory management) after digital transformation relative to the original process duration." Data is calculated by comparing performance records before and after the transformation process.

2.3. HNN-DIC Model Design

The digital innovation capability of SMEs is shaped by the combined effects of static structural coordination and dynamic temporal evolution (Li, 2025). At the static level, this manifests as nonlinear coupling across core dimensions, such as technology embedding and resource coordination, at a single point in time. At the dynamic level, it reflects long-term dependencies among dimensional indicators that evolve with environmental changes. Traditional single-architecture models struggle to fully capture this complex mechanism. The proposed hybrid neural network model employs a dual-branch architecture: "spatial feature extraction, temporal feature modeling, higher-order fusion." The model receives three-dimensional tensor inputs as shown in Eq. (1).

$$X \in R^{B\times T\times N} \tag{1}$$

B=32 represents the batch size, T=8 denotes the time lag (corresponding to eight quarters), and N=9 indicates the feature dimension (i.e., 9 core indicators). During data preprocessing, Min-Max normalization is applied to eliminate dimensional differences. The normalization formula is shown in Eq. (2).

$$x_{norm}=\frac{x_{ori}\text{-min}(x)}{\max(x)\text{-min}(x)} \tag{2}$$

Input tensors are normalized to the range [0,100] (for intuitive interpretation of DIC scores) using the following normalization formula as shown in Eq. (3).

$$x_{ori}=x_{norm}\cdot(\max(x)\text{-min}(x))+\min(x) \tag{3}$$

The x_{ori} denotes the raw indicator value, x_{norm} represents the normalized value, while max(x) and min(x) denote the maximum and minimum values of the indicator across the entire sample, respectively. Data loading employs an "enterprise-independent partitioning" strategy: the eight quarterly time-series data points from the same enterprise are assigned exclusively to the training, validation, or test set, thereby preventing overestimation of generalization capabilities due to cross-set data leakage.

The spatial feature extraction branch focuses on local interactions and correlation dependencies among multiple indicators within the same time step. It employs the classic feature extraction chain: "convolution, batch normalization, activation, pooling." The first convolutional layer (Conv1D-1) uses a convolutional kernel count of count1=32 and a kernel size of k=3 (determined by the empirical rule "number of indicators = 9, optimal local window size = 3") to effectively capture interactions between adjacent indicators. The stride is set to s=1, and the padding mode is "same" (ensuring the output sequence length matches the input). Three is the optimal local window size" and cross-validation, effectively capturing interactions between adjacent indicators, stride s = 1, and padding mode "same" (ensuring output sequence length matches input). The parameter count formula for the convolutional layer is shown in Eq. (4).

$$\text{Params}_{Conv}=k\cdot c_{in}\cdot c_{out}+c_{out} \tag{4}$$

Where c_{in} denotes the number of input channels (here $c_{in} = 9$), count represents the number of output channels, and the final term is the bias parameter. Substituting yields the number of parameters for Conv1D-1: 3×9×32 + 32 = 896.

After the convolutional layer, the input data is normalized (μ = 0) to accelerate model convergence and mitigate the impact of internal covariate shift on training.

The ReLU function is employed.

$f(x)=\max(0,x)$

This function effectively addresses the vanishing gradient problem while introducing nonlinear fitting capabilities to accommodate complex relationships between indicators.

The pooling layer uses a pooling window Kpool=2, stride Spool=1, and padding method "same", with output dimensions (32, 8, 32). Downsampling preserves key features, reduces computational load, and prevents overfitting.

The second convolutional layer (Conv1D-2) has 64 convolutional kernels (cout2 = 64), with all other parameters identical to Conv1D-1 (k=3, s=1, padding "same"). Parameters: 3×32×64 + 64 = 6208; Input dimension (32, 8, 32), output dimension (32, 8, 64), further extracting high-order local correlation features. The spatial feature vector dimension is (32, 8, 64), corresponding to 64-dimensional static interaction features for each time step.

To address the dynamic evolution of DIC, the LSTM's gating mechanism (input gate, forget gate, output gate) captures long-term temporal dependencies in indicators, thereby avoiding the vanishing gradient problem in traditional recurrent neural networks (RNNs). The design is as follows: The first LSTM hidden layer has $d_{hid} = 64$ neurons (optimized via grid search within 32-128, yielding minimum validation loss at $d_{hid} = 64$), with input dimensions (32, 8, 9). Return sequences=True is set to retain outputs at each time step for subsequent feature fusion. The LSTM layer parameter count is calculated, as shown in Eq. (5).

$$\mathrm{Params}_{LSTM}=4\cdot(d_{in}\cdot d_{hid}+d_{hid}^{2}+d_{hid}) \tag{5}$$

The d_{in} represents the input dimension (9 in this case), and the coefficient 4 corresponds to the 4 sets of weights for the gating mechanism (input gate, forget gate, output gate, cell state update). Substituting yields the first LSTM layer parameters: 4×(9×64 + 64² + 64) = 18,944; output dimensions (32, 8, 64). A Dropout layer (dropout rate=0.2) is added after the LSTM layer, randomly deactivating 20% of neurons to suppress overfitting. The second LSTM layer maintains d_{hid}=64 hidden neurons, with input dimensions (32, 8, 64) (d_{in}=64) and return sequences=True. Parameters: 4×(64×64 + 64² + 64) = 32,640; Output dimension (32, 8, 64), enabling high-order modeling of temporal features. The output temporal feature vector has dimensions (32, 8, 64), corresponding to 64-dimensional dynamic evolutionary features at each time step.

Feature fusion and the fully connected layer perform a nonlinear mapping from the fused features to DIC scores. The concatenation layer concatenates the CNN branch output (32, 8, 64) with the LSTM branch output (32, 8, 64) along the last dimension, yielding a fused feature vector (32, 8, 128) that comprehensively integrates static interaction and dynamic evolution information. The three-dimensional fused features are converted to a two-dimensional vector (32, 8×128) = (32, 1024) to meet the fully connected layer's input requirements.

The first fully connected layer (Dense-1) has n1 = 128 neurons, uses the ReLU activation function, and incorporates L2 regularization to suppress overfitting by penalizing excessively large weights. Parameters: 1024 × 128 + 128 = 131,200; output dimension (32, 128). Second Dense Layer (Dense-2): Number of neurons n2 = 64, ReLU activation function. Parameters: 128 × 64 + 64 = 8,256; Output dimension (32, 64), further compressing feature dimensions and enhancing high-order nonlinear combinations.

Since DIC scores are continuous variables (range [0,100]), the model inherently performs regression. Thus, the output layer is designed as a single linear activation neuron with input dimension (32, 64) and output dimension (32, 1). Parameters: 64×1 + 1 = 65; directly outputs the enterprise's current comprehensive DIC score without additional activation function conversion (avoiding score range distortion).

To ensure model convergence, stability, and generalization, the following training strategy employs Mean Squared Error (MSE) to effectively penalize larger prediction errors, as shown in Eq. (6).

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{6}$$

The y represents the true DIC score for the i-th sample, $\hat{y}_i$ denotes the model-predicted score, and n indicates the number of samples.

The total number of trainable parameters in HNN-DIC is approximately 198.2k, including all convolutional, LSTM, and fully connected layers.

3. Experiments and Results Analysis

3.1. Experimental Setup

Panel data from 500 manufacturing SMEs (500 firms × 8 quarters) were randomly partitioned by firm ID. The training set comprises 70% (350 firms, 2,800 observations), with 15% allocated to the validation set (75 firms, 600 observations) and 15% to the test set (75 firms, 600 observations). The eight-quarter time-series data for each firm was fully assigned to a single dataset (training/validation/test). Consistency with the population distribution was verified using a chi-square test (industry distribution) and a one-way ANOVA (regional distribution, indicator mean values). Results showed: Industry distribution chi-square value = 3.26 ($p=0.35>0.05$), and F-value for mean variance analysis of indicators = 1.21 ($p=0.28>0.05$). This indicates that there are no significant differences in industry, region, or indicator distributions between the training, validation, and test sets and the overall population, ensuring fairness in experimental comparisons. Four types of baseline models were selected, as shown in Table 1.

Table 1. Overview of the baseline model and parameter settings

Model Categories	Model Name	Core Parameter Settings	Training Configuration
Traditional Econometric Models	Multiple Linear Regression (MLR)	No hyperparameters; ordinary least squares (OLS) used to solve coefficients	Implemented using Scikit-learn 1.0, with a convergence threshold of 1e-4
Classical Machine Learning Models	Random Forest (RF)	n_estimators=100 (number of decision trees), max_depth=10 (maximum tree depth), min_samples_split=5 (minimum sample size for node splitting)	Implemented using Scikit-learn 1.0, with grid search optimization ranges: n_estimators ∈ [32, 200], max_depth ∈ [5, 20]
	XGBoost	learning_rate=0.1 (learning rate), n_estimators=100 (number of weak learners), max_depth=6 (maximum tree depth), subsample=0.8 (sample subsampling rate)	Implemented using XGBoost 1.6.1, objective function set for regression task (objective="reg:squarederror")
Foundational Neural Network Models	Backpropagation Neural Network (BPNN)	Network architecture: Input layer (9-dimensional) → Hidden layer 1 (128 neurons, ReLU activation) → Hidden layer 2 (64 neurons, ReLU activation) → Output layer (1-dimensional, linear activation)	Implemented using PyTorch 1.12, Adam optimizer (lr=0.001), MSE loss function, 50 training epochs
Component Comparison Models	CNN-only Model	Identical to the CNN branch of HNN-DIC (2 layers of Conv1D, 32/64 convolution kernels, MaxPooling1D, BatchNorm), followed by fully connected layers (128→64→1)	Training configuration identical to HNN-DIC (Adam optimizer, MSE loss, early stopping)
	LSTM-only Model	Identical to the LSTM branch of HNN-DIC (2 LSTM layers, 64 neurons, Dropout=0.2), followed by fully connected layers (128→64→1)	Training configuration identical to HNN-DIC

The coefficient of determination R^2 measures the degree of fit between a model's predicted values and actual values. Its range is [0,1], with values closer to 1 indicating a better fit. The formula is shown in Eq. (7).

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{7}$$

The yi denotes the true DIC score for the i-th sample, $\hat{y}_i$ represents the model's predicted score, $\bar{y}$ is the sample mean of the true scores, and n denotes the number of samples.

The Mean Absolute Error (MAE) reflects the average absolute deviation between predicted and true values. A smaller value indicates greater prediction accuracy and stronger robustness to outliers. The formula is shown in Eq. (8).

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{8}$$

Root Mean Square Error (RMSE) is more sensitive to outliers, reflecting the stability of predictions. A lower value indicates reduced variability in model forecasts. The formula is shown in Eq. (9).

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

3.2. Implementation Details

All experiments were conducted on a dedicated server equipped with an Intel Xeon Gold 6248 CPU @ 2.50 GHz, 128 GB RAM, and four NVIDIA Tesla V100 (32 GB) GPUs. The operating system was Ubuntu 20.04 LTS. The HNN-DIC model was implemented using PyTorch 1.12.0, CUDA 11.6, and cuDNN 8.3.2. Python 3.8.10 was used. Baseline models (MLR, Random Forest, XGBoost, BPNN) were implemented using Scikit-learn 1.0.2 and XGBoost 1.6.1.

The Adam optimizer was used with an initial learning rate of 0.001, β_1=0.9, β_2=0.999, and weight decay of 1e^{-5}. The batch size (B) was set to 32 (as defined in Eq. (1)). The model was trained for a maximum of 200 epochs, but early stopping was applied with a patience of 10 epochs: training terminated if the validation loss did not decrease for 10 consecutive epochs, and the model checkpoint with the lowest validation loss was restored. The loss function was Mean Squared Error (MSE, Eq. (6)). Gradient clipping was applied at a maximum norm of 1.0 to prevent gradient explosion.

Following Eq. (2), all nine indicator features were normalized to the range [0, 1] using min-max scaling calculated from the training set only (to avoid data leakage). The scaling parameters (min, max) were then applied to the validation and test sets. The DIC scores (target variable) were normalized to [0, 100] using the same min-max approach for model training; predictions were then inverse-transformed to the original scale for metric calculation. Missing values were imputed using KNN imputation with k=5 neighbors, implemented via sklearn.impute.KNNImputer. Outliers beyond ±3 standard deviations were capped at the boundary values.

All random processes (data splitting, model initialization, dropout, and GPU operations) were fixed using a seed of 42. random.seed (42), np.random.seed (42), torch.manual_seed (42), and torch.cuda.manual_seed_all (42). This ensures exact reproducibility of the reported results.

3.3. Performance Comparison Analysis

The experiment was conducted on an independent test set (75 enterprises, 600 observations), with performance metrics as shown in Table 1. HNN-DIC achieved a coefficient of determination (R^2) of 0.931 ± 0.004, significantly surpassing all baseline models. Compared to the second-best performer, XGBoost, R^2 improved by 1.6%, indicating superior explanatory power for DIC scores. HNN-DIC's Mean Absolute Error (MAE) was 1.72±0.08, and Root Mean Square Error (RMSE) was 2.28±0.10, representing reductions of 13.0% and 14.0%, respectively, compared to XGBoost, reflecting superior predictive accuracy. A paired t-test confirmed differences in prediction errors between HNN-DIC and XGBoost, yielding $t=4.26$ ($p<0.01$) for MAE and $t=3.89$ ($p<0.01$) for RMSE, confirming statistically significant performance disparities. Regarding baseline model performance. Traditional Multiple Linear Regression (MLR) failed to capture non-linear and temporal dependencies among indicators, resulting in the poorest performance with an R^2 of 0.723. The basic Backpropagation Neural Network (BPNN), lacking an architecture for spatio-temporal features, demonstrated lower prediction accuracy than HNN-DIC. Component models, such as CNN/LSTM, which extract only single-type features, also exhibited lower accuracy than HNN-DIC, which integrates spatio-temporal features, confirming the necessity of hybrid architectures.

Table 2. Performance comparison of models on the test set (mean and standard deviation)

Model	R^2(↑)	MAE(↓)	RMSE(↓)	Training duration (seconds)
HNN-DIC (Proposed Model)	0.931±0.004	1.72±0.08	2.28±0.10	335
XGBoost	0.916±0.005	1.98±0.10	2.65±0.12	42
Random Forest	0.902±0.006	2.15±0.11	2.98±0.14	28
LSTM Only	0.887±0.007	2.41±0.12	3.22±0.15	290
Convolutional Neural Network Only	0.872±0.008	2.58±0.13	3.45±0.16	255
Fully Connected Neural Network	0.855±0.010	2.80±0.15	3.70±0.18	205
Multiple Linear Regression	0.723±0.015	4.35±0.25	5.82±0.30	<1

To further analyze the trade-off between 'prediction accuracy, training efficiency, and parameter complexity' in the model, this study employs parallel coordinate plots and radar charts for visualization. All metrics are standardized such that 'higher values indicate superior performance' (where MAE, RMSE, training time, and parameter count undergo inverse normalization).

Fig. 1 presents a parallel coordinate diagram illustrating the multi-metric contributions of the models. The horizontal axis encompasses five core evaluation metrics (R^2 (norm), MAE (inv-norm), RMSE (inv-norm), training time (inv-norm), and number of parameters (inv-norm)), while the vertical axis displays the normalized contribution values of each metric. Regarding predictive accuracy, HNN-DIC (blue curve) exhibits values close to 1.0 across R^2, MAE (inv-norm), and RMSE (inv-norm) dimensions, positioning it within the optimal range among all models and directly corroborating its high-precision advantage as indicated by quantitative results. Regarding efficiency and complexity: Whilst HNN-DIC's training time (inverse norm) and number of parameters (inverse norm) are lower than those of Random Forest and XGBoost (tree models inherently exhibit high training efficiency and minimal parameter count), they significantly outperform LSTM/CNN models alone. This demonstrates that HNN-DIC maintains accuracy without unduly increasing model complexity, rendering it feasible for practical deployment.

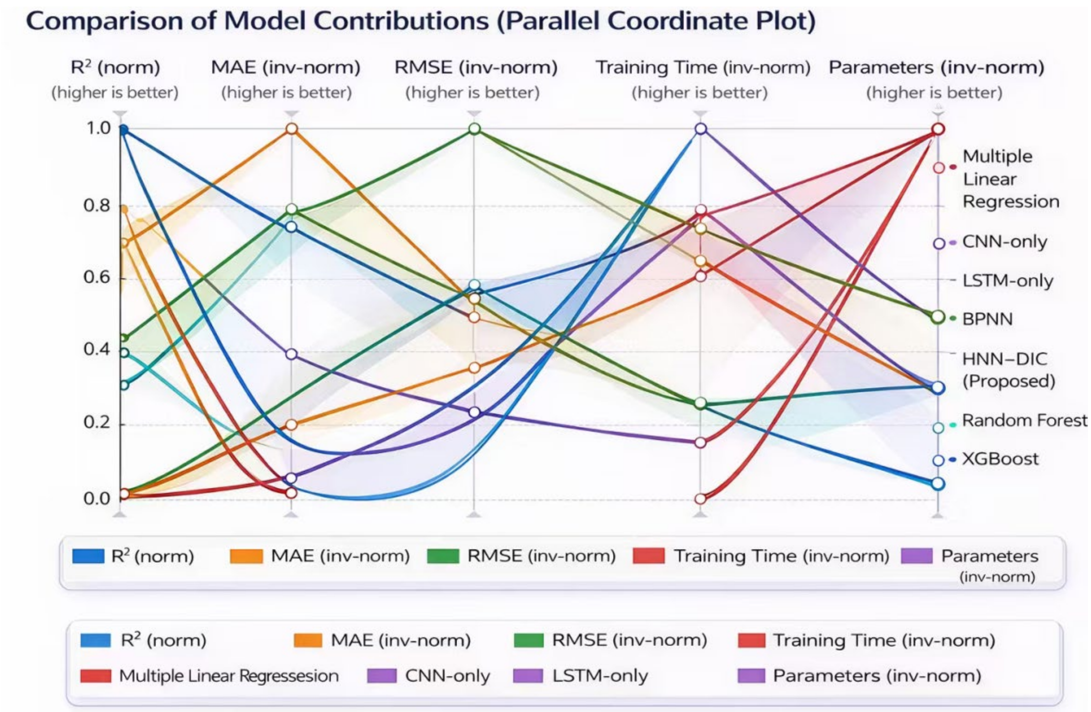

Fig. 1. Comparison of normalized performance across five metrics for each evaluation model

Fig. 2 presents a simplified normalized performance parallel coordinate plot, offering a clearer illustration of each model's performance contours. HNN-DIC (blue curve) demonstrates significantly superior performance across R^2, MAE (inverse norm), and RMSE (inverse norm) metrics compared to other models, with only a marginal decrease in training time (inverse norm). This reflects its design philosophy of 'prioritizing high accuracy with a moderate trade-off in efficiency.' Tree-based models (XGBoost, Random Forest), whilst advantageous in training time and parameter count dimensions, demonstrate markedly weaker predictive accuracy compared to HNN-DIC. Only CNN/LSTM models exhibit overall performance profiles below HNN-DIC, further validating the necessity of the 'spatio-temporal feature fusion' architecture.

Fig. 3 presents a radar chart illustrating the comprehensive performance of the models, with five evaluation metrics serving as dimensions. It visually demonstrates each model's overall capabilities. The coverage area of HNN-DIC (deep blue contour) is significantly larger than that of other models. It approaches maximum values in the dimensions of R^2, inverse MAE, and inverse RMSE, while only slightly lower in the dimensions of training time (inverse) and parameter count (inverse). This reflects the combined advantages of "high accuracy, moderate efficiency, manageable complexity." Traditional MLR and BPNN exhibit the smallest contour areas, reflecting their performance limitations. Only the CNN/LSTM model's contour area is weaker than HNN-DICs, further validating the hybrid architecture's superior ability to capture spatio-temporal features comprehensively.

3.4. Ablation Studies and Feature Importance

To validate the necessity of the 'spatial features (CNN branch), temporal features (LSTM branch), and fusion' architecture, this study designed two ablation variant models listed below.

Variant one (HNN-DIC w/o CNN) removed the 1D-CNN spatial feature branch, retaining only the LSTM temporal feature branch. Variant two (HNN-DIC w/o LSTM) removed the LSTM temporal feature branch, retaining only the 1D-CNN spatial feature branch.

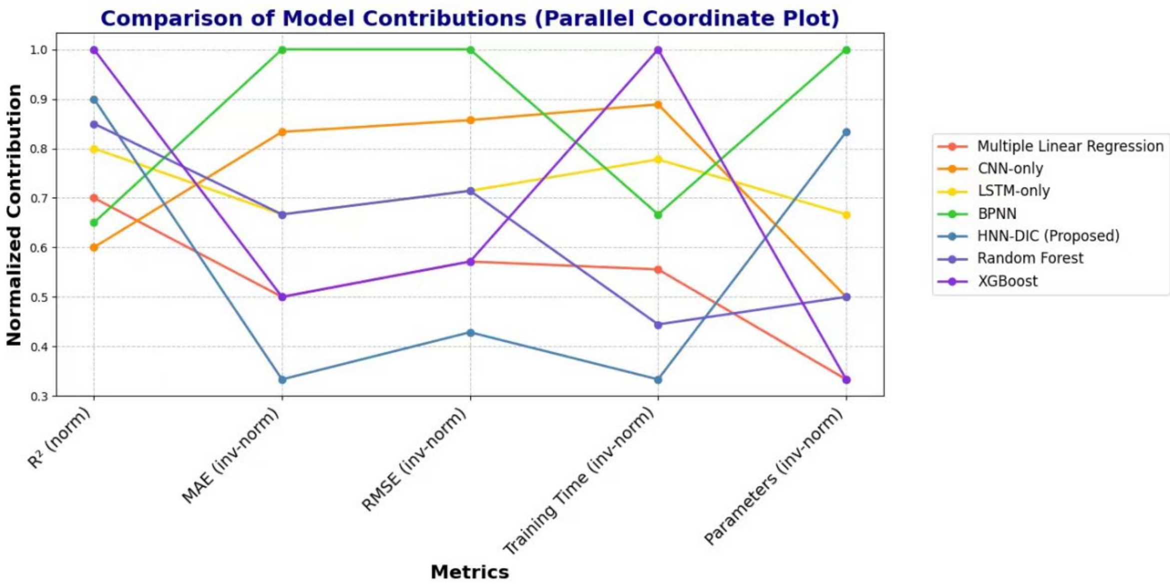

Fig. 2. Core performance profile of HNN-DIC and benchmark models

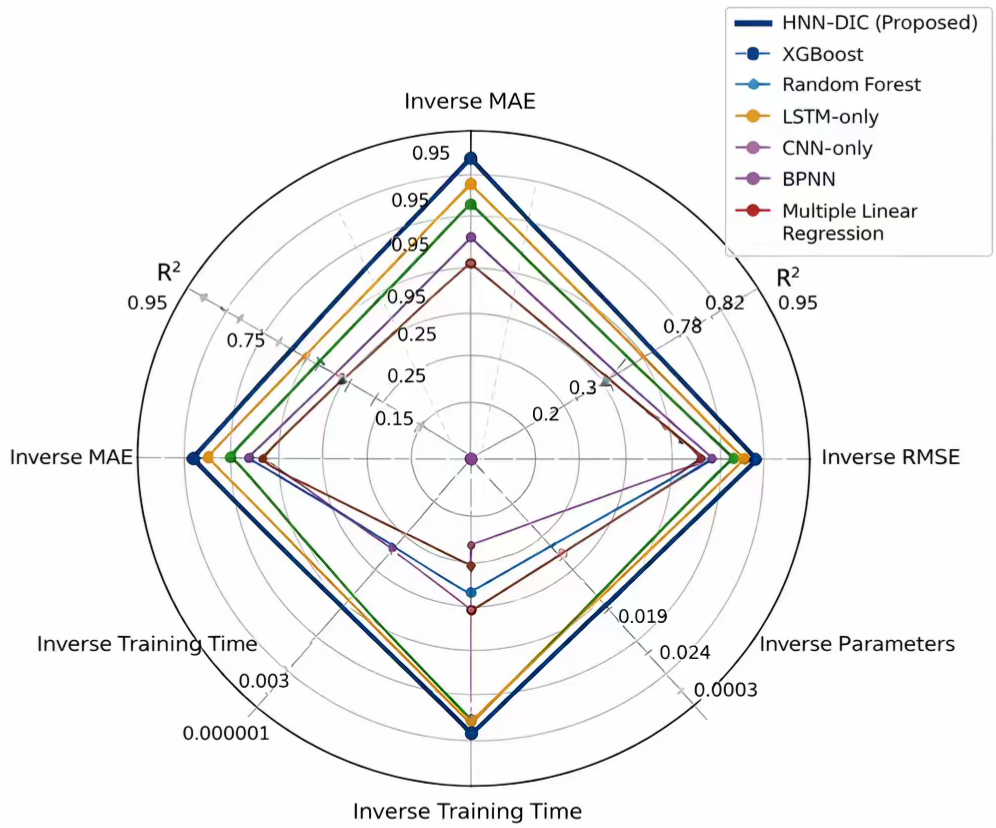

Fig. 3. Multidimensional comparison of the comprehensive performance of various evaluation models

The ablation experiment results on the test set are shown in Table 3. After removing the CNN branch, the model's R^2 decreased from 0.931 to 0.884 (a 4.7% reduction), while MAE increased to 2.35 (a 36.6% increase). Upon removing the LSTM branch, the model's R^2 decreased from 0.931 to 0.870 (a 6.3% reduction), while RMSE improved to 3.12 (a 36.8% increase). Paired t-tests revealed statistically significant performance differences between both variants and the original HNN-DIC model, with $p<0.01$ for both comparisons.

Table 3. Performance comparison of ablation experiments (test set)

Model	R^2(↑)	MAE(↓)	RMSE(↓)
HNN-DIC (Original Model)	0.931	1.72	2.28
HNN-DIC without CNN	0.884	2.35	2.95
HNN-DIC without LSTM	0.870	2.48	3.12

These results indicate that spatial (local interactions between indicators) and temporal (dynamic evolution of indicators) characteristics are both indispensable for DIC assessment, with the hybrid architecture serving as the core design element, ensuring high model accuracy.

The SHAP (SHapley Additive exPlanations) method was employed to calculate the global contribution of each indicator to the DIC score (where a higher SHAP value signifies a stronger positive influence of that indicator on the DIC score), with the results presented in Fig. 4.

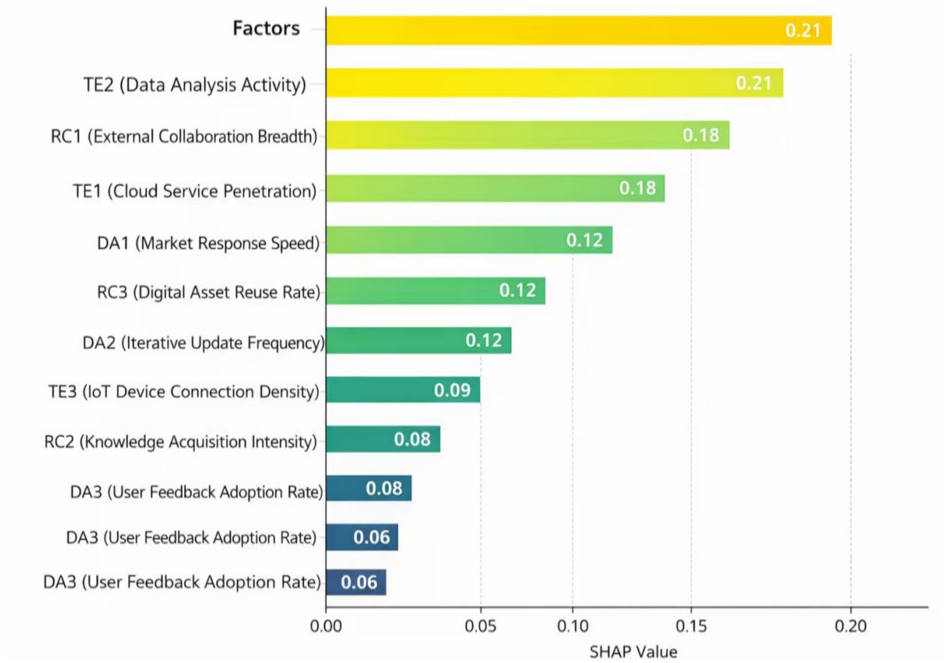

Fig. 4. Global feature importance ranking based on SHAP values (mean | SHAP value |)

The average SHAP value for 'Data Analysis Activity (TE2)' reached 0.21, making it the most influential indicator on DIC scores; 'External Collaboration Breadth (RC1)' and 'Cloud Service Penetration Index (TE1)' both recorded SHAP values of 0.18, ranking in the second tier. This indicates that data-driven decision-making capabilities, open collaboration within platform ecosystems, and deep application of cloud services constitute the core levers of SMEs' digital innovation capacity.

Market Responsiveness (DA1), digital asset reuse rate (RC3), and Iterative Update Frequency (DA2) each yielded SHAP values of 0.12, indicating significant supporting dimensions for DIC and reflecting the positive impact of dynamic adaptability on innovation performance.

The SHAP values for 'IoT device connection density (TE3)', 'knowledge acquisition intensity (RC2)', and 'user feedback adoption rate (DA3)' were relatively low (0.06-0.09), yet these remain foundational elements for DIC formation. Their lower contribution may be due to shallow IoT deployment and insufficient knowledge conversion among some SMEs.

3.5. Robustness Analysis under Different Conditions

To further evaluate the robustness of HNN-DIC, we conduct additional experiments under varying conditions. We repeat the experiment 5 times with different random seeds (42, 123, 456, 789, 1011). The mean R^2 remains 0.930±0.005, MAE=1.74±0.09, RMSE=2.31±0.11, indicating low sensitivity to data split. Besides, we train HNN-DIC using 4-quarter, 6-quarter, and 8-quarter windows. Results show that longer windows improve R^2 (0.902 for 4 quarters, 0.921 for 6 quarters, 0.931 for 8 quarters), confirming that longer historical data benefits DIC prediction. Grid search over batch sizes {16, 32, 64} and learning rates {0.0005, 0.001, 0.002} shows that the original settings (batch=32, lr=0.001) yield the best performance, with variations within ± 0.005 R^2. Furthermore, when splitting the 500 firms into small (≤50 employees, 210 firms) and medium (51-250 employees, 290 firms) groups, HNN-DIC achieves R^2=0.918 for small firms and R^2=0.939 for medium firms, suggesting better predictive performance for larger SMEs with stable digital behavior.

These experiments demonstrate that HNN-DIC maintains impressive performance across various conditions, and the reported results are robust.

4. Discussion

4.1. Summary of Findings

Based on dynamic capability theory and the characteristics of digital platform ecosystems, this study constructs and empirically validates a Hybrid Neural Network-based Digital Innovation Capability (HNN-DIC) integrating spatio-temporal features. It systematically reveals the assessment methodology, core driving mechanisms, and practical application value of SMEs Digital Innovation Capability (SMEs-DIC). Methodologically, the spatio-temporal hybrid architecture achieves high-precision DIC assessment. The proposed HNN-DIC model employs a dual-branch architecture. '1D-CNN spatial feature extraction combined with LSTM temporal feature modeling' effectively addresses the limitations of traditional evaluation methods in handling multi-source high-dimensional data, complex non-linear relationships, and temporal evolution characteristics. Empirical results demonstrate that the model achieves a coefficient of determination (R^2) of 0.931 on an independent test set. Its MAE = 1.72 and the RMSE = 2.28 significantly outperform Multiple Linear Regression (MLR), Random Forest (RF), XGBoost, and standalone CNN/LSTM models. Compared to the second-best performer, XGBoost, R^2 improved by 1.6%, and MAE decreased by 13.0%, with differences statistically significant at the $p<0.01$ level. Ablation experiments further confirmed that removing either the CNN branch (spatial features) or the LSTM branch (temporal features) resulted in decreases in R^2 of 4.7% and 6.3%, respectively. This confirms that 'spatio-temporal feature fusion' is the core design element ensuring assessment accuracy, providing a methodological innovation for the dynamic and precise evaluation of SMEs-DIC. At the mechanism level, data-driven and open collaboration are core levers for enhancing DIC. Based on SHAP interpretability analysis, this study identified key drivers and their contribution strengths for SMEs-DIC. Data analysis activity (TE2, average SHAP value = 0.21) and external collaboration breadth (RC1, average SHAP value = 0.18) rank as the top two core indicators influencing DIC, followed by dimensions such as cloud service penetration index (TE1) and market responsiveness (DA1). This finding empirically substantiates the core tenets of dynamic capability theory, namely, 'resource coordination and integration' and 'dynamic adaptation.' Within digital platform ecosystems, SMEs digital innovation capabilities no longer rely solely on internal technological investment. Instead, they increasingly manifest as the combined effects of 'data-driven decision-making efficiency' (deep application of data analytics tools) and 'open collaborative ecosystem engagement' (cross-entity joint innovation and resource sharing). This provides empirical support for theoretically deciphering the mechanisms underlying the formation of SMEs innovation capabilities in the digital economy era. Unlike traditional 'score-only' assessment methods, the HNN-DIC model delivers dual value through 'high-precision evaluation' and 'interpretable diagnostics. First, it provides comprehensive capability profiles by quantifying overall DIC scores and sub-scores across four dimensions: technological embedding, dynamic adaptation, resource coordination, and innovation output. Second, it employs SHAP feature importance analysis and visual diagnostic reports (such as radar chart comparisons) to pinpoint specific innovation shortcomings (e.g., the 'insufficient digital asset reuse rate' issue observed in low-DIC enterprises within the case study), while offering actionable improvement pathways. Furthermore, the model's lightweight design (total parameters approximately equal to 198.2k) enables encapsulation as a cloud-based SaaS service prototype. This accommodates SMEs self-diagnostic needs while furnishing government departments with regional innovation vitality dashboards and platform operators with ecosystem empowerment metrics, thereby achieving a closed-loop system of 'theoretical methodology, tool implementation, and multi-stakeholder benefits.

To make managerial implications actionable, the HNN-DIC framework suggests four specific changes to decision processes. (i) Shift from intuition-based to SHAP-guided resource allocation. (ii) Adopt quarterly DIC monitoring instead of annual reviews. (iii) Use dimensional sub-scores for targeted capability diagnosis and (iv) Quantitatively justify platform collaboration investments based on the high SHAP value of external collaboration breadth (Pham et al., 2024).

4.2. Limitations and Future Work

Although the proposed HNN-DIC framework demonstrated strong predictive performance and interpretability, several limitations remain regarding generalizability to different datasets, data diversity, and real-world deployment.

First, the research data were collected exclusively from a major domestic industrial internet platform, where most registered enterprises are concentrated in manufacturing-oriented sectors such as equipment manufacturing and the electronic information industry. In addition, the sampled enterprises are primarily located in economically developed eastern regions. Although the dataset contains two years of panel data from 500 SMEs across eight quarters, the industry diversity and geographical representativeness of the sample remain relatively limited. Consequently, the proposed framework may not fully capture the digital innovation characteristics of SMEs operating in traditional consumer goods, light manufacturing, or service-oriented industries. Furthermore, the current study does not incorporate complementary data sources from consumer internet platforms (e.g., e-commerce transaction data), fintech platforms (e.g., supply chain finance data), or public innovation databases, which may restrict the comprehensiveness of digital innovation capability assessment across market access, financial support, and ecosystem collaboration dimensions.

Second, the current HNN-DIC framework primarily focuses on static assessment, generating DIC scores at a specific observation period. Although the model effectively captures spatiotemporal characteristics in historical panel data, it has not yet achieved dynamic prediction and intelligent early warning capabilities. For example, the current framework cannot proactively identify declining DIC trends driven by continuously declining IoT device adoption rates or by interrupted collaborative innovation activities, nor can it quantitatively predict the lagged effects of DIC improvement on future innovation performance indicators, such as new product revenue growth.

Future research will focus on improving both the generalizability and practical applicability of the proposed framework. Specifically, subsequent studies will integrate multi-source heterogeneous data from industrial internet platforms, consumer internet ecosystems, fintech platforms, and government public databases to construct a multidimensional "technology-market-capital" data system. Meanwhile, cross-industry and cross-regional samples will be expanded to validate the robustness and universality of the proposed model under different economic and industrial conditions.

In addition, future work will extend the current HNN-DIC framework into a multi-task intelligent decision-support system capable of simultaneously performing DIC assessment, early warning of decline, and innovation performance forecasting. This will include adding additional LSTM-based forecasting branches for multi-quarter DIC trend prediction, integrating anomaly detection algorithms to identify innovation risk, and establishing lagged regression relationships between DIC scores and enterprise innovation outcomes to further enhance the practical deployment value of the proposed framework (Civelek et al., 2023).

5. Conclusion

This study addresses limitations of traditional SME digital innovation capability (DIC) assessment, subjectivity, static snapshots, and poor interpretability by proposing a data-driven framework grounded in dynamic capability theory. Using two-year panel data from 500 manufacturing SMEs on an industrial internet platform, we develop the HNN-DIC model that fuses 1D-CNN (spatial features) and LSTM (temporal dynamics). The model achieves R^2=0.931, MAE=1.72, RMSE=2.28, significantly outperforming XGBoost (R^2=0.916, MAE=1.98, RMSE=2.65), Random Forest, and single-architecture networks. Ablation studies confirm that removing CNN or LSTM branches reduces R^2 by 4.7%-6.3%. SHAP analysis identifies data analysis activity and breadth of external collaboration as core DIC drivers. The framework provides a cloud-based self-diagnostic tool for SMEs and a policy dashboard for governments. A limitation is that the dataset is restricted to manufacturing SMEs from a single platform, limiting generalizability to service industries or other platforms. Future work will expand to cross-industry, cross-regional datasets and extend HNN-DIC to dynamic prediction and early warning of innovation decline.

Author Contributions

Ming Guo contributed to the conception and design of the manuscript, material preparation, data collection and analysis. Kai Qi contributed to material preparation and data analysis and prepared the first draft. Lijing Kang contributed to conception and design, and to manuscript review and revision.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors used AI-assisted language editing tools solely for improving grammar, readability, and language fluency during manuscript preparation and revision. No AI tools were used for generating scientific ideas, research methodology, experimental design, data analysis, result interpretation, conclusions, or scholarly arguments. All academic content and scientific contributions were independently developed and verified by the authors.

References

Civelek, M., Krajčík, V., and Ključnikov, A. (2023). The impacts of dynamic capabilities on SMEs' digital transformation process: The resource-based view perspective. *Oeconomia Copernicana,* 14(4), 1367-1392.

Fu, Y. (2014). Research on aggregation mechanisms of platform-based e-commerce (Doctoral dissertation). Northwest University.

He, M. (2012). A study on the mechanism and model of social capital's influence on cross-organizational information system absorption (Doctoral dissertation). Harbin Institute of Technology.

Hu, Y. and Liao, X. (2025). Does "playful labor" create value? *Abstracts of Humanities and Social Sciences in Higher Education,* 42(3), 121-122.

Khurana, I., Ghura, A. S., and Dutta, D. K. (2025). Dynamic capabilities and digitalization in entrepreneurial SMEs: A knowledge-based view. *Journal of Business Research*, 188, 115074.

Kuang, H. and Guo, Q. (2025). A short-term wind power forecasting method based on multi-feature extraction convolutional neural networks and long short-term memory networks. *Power Generation Technology,* 46(1).

Li, W. (2025). Causal Inference Framework Based on CNN-LSTM-AM-DID: The Impact of Establishing Digital Economy Innovation Development Pilot Zones on Enterprise Digitalization. 2025 *IEEE Conference on Digital Economy Innovation*, 1-8.

Liu, L., Su, X., and Cheng, D. (2018). Factors influencing innovation behavior in university research teams supported by communication technology platforms: The cross-level moderating role of team leader support. *Science and Technology Progress and Policy,* 35(10), 7.

Liu, P. (2012). Research on the financial risk assessment system for overseas mergers and acquisitions by Chinese private enterprises (Doctoral dissertation). Shandong University of Finance and Economics.
Ni, H. and Li, H. (2021). A battery state-of-charge prediction method based on the combination of one-dimensional convolutional neural networks and long short-term memory networks. *Computer Applications*.
Nie, L. (2011). A study on governance structures and mechanisms of mobile business alliances (Doctoral dissertation). Huazhong University of Science and Technology.
Pham, T. H., Dinh, V. H., Tran, T. K. M., and Nguyen, T. H. (2024). Digital capabilities and SMEs' performance under technological uncertainty: empirical evidence from an emerging market. *International Journal of Innovation and Learning,* 36(3/4), 404-424.
Shan, D., Shi, L., and Tan, K. (2023). Bridge damage identification using a deep network integrating convolutional neural networks and long short-term memory. *Bridge Construction,* 53(4), 41-46.
Xiao, X. and Fu, H. (2023). Digital entrepreneurship growth: Literature review and research prospects. *Foreign Economics and Management,* 45(6), 118-136. https://doi.org/10.16538/j.cnki.fem.20221027.401
Yang, X. (2022). Research on spatial spillover effects and influence mechanisms of information infrastructure on economic growth (Doctoral dissertation). University of Chinese Academy of Social Sciences.
Yin, Y., Wu, H., and Zhang, T. (2011). Constructing a technology adoption model for third-party payment. *Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)*, (3).
Yu, D. and Ma, L. (2024). Digital transformation, platform-based change, and corporate innovation performance: An analysis based on the "technology–organization–innovation" paradigm. *Reform*, (2), 55-74.
Yu, S. T. (2020). Research on PM2.5 concentration prediction based on long short-term memory network and convolutional neural network (LSTM-CNN) (Doctoral dissertation). Shanghai Jiao Tong University.
Yu, S. and Liu, P. (2020). PM2.5 concentration prediction in Beijing based on LSTM-CNN. *Environmental Engineering*, 38(6), 6.

Guo Ming pursued her postgraduate studies in Business Administration at INJE University, South Korea, in 2023. Her research interests focus on the digital transformation of Small and Medium-Sized Enterprises (SMEs), green innovation and artificial intelligence. She has published several papers in conference proceedings. Her core research interests cover digital transformation and the low-carbon green transition of SMEs.

Qi Kai has been pursuing his postgraduate degree in Business Administration at Dong-eui University, South Korea, since 2023. He specializes in the digital transformation of service-oriented enterprises and the digitalization of tourism. He has published several papers in academic conference proceedings, and his research interests include the digital upgrading of tourism enterprises and the digital development of the tourism industry.

Lijing Kang received the Ph.D. degree in Civil Engineering from Hunan University in 2022. She is currently with Guangxi Minzu University. Her research interests include the seismic performance of prefabricated structures, lightweight structural inspection based on machine vision, and structural impact and protection.