Scientific Research Publishing

iBusiness, 2026, 18(2), 59-87
https://www.scirp.org/journal/ib
ISSN Online: 2150-4083
ISSN Print: 2150-4075

Digital-Intelligence-Driven Evaluation and Forecasting of First-Launch Economy Performance: Evidence from the Plush Toy Industry in Ankang City, China

Mo Shi[1*], Baixiao Tang[1], Ruiyu Li[2]

[1]School of Economics and Management, Ankang University, Ankang, China
[2]School of Arts, Ankang University, Ankang, China
Email: *shimo0204@outlook.jp

How to cite this paper: Shi, M., Tang, B. X., & Li, R. Y. (2026). Digital-Intelligence-Driven Evaluation and Forecasting of First-Launch Economy Performance: Evidence from the Plush Toy Industry in Ankang City, China. *iBusiness, 18,* 59-87.
https://doi.org/10.4236/ib.2026.182005

Received: January 11, 2026
Accepted: May 24, 2026
Published: May 27, 2026

Abstract

The first-launch economy has emerged as a critical policy instrument for driving industrial upgrading and rural revitalization in China's underdeveloped regions; however, quantitative frameworks to assess its industrial performance remain limited. This study proposes a novel digital-intelligence-driven framework to evaluate and forecast production expansion dynamics, utilizing the plush toy industry in Ankang City (2018-2026) as a representative longitudinal case study. Based on the integration of multi-source data from the National Bureau of Statistics, e-commerce platforms, and web-crawled records, a robust Extract-Transform-Load (ETL) pipeline is constructed. This process employed systematic time-series standardization and exhaustive aggregation to generate a monthly production value of industrial activity. To capture complex nonlinear trends, a Long Short-Term Memory (LSTM) neural network was developed to model efficiency evolution, incorporating the national export benchmark of toys as an exogenous variable. In this study, model performance was validated using the coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Bias Error (MBE), demonstrating high predictive fidelity. As the first-launched industry, analytical results in this study reveal a sustained efficiency trajectory in Ankang's plush toy sector post-2018, providing quantitative evidence of the efficacy of digital-intelligence-enabled interventions. The findings in this study offer a transferable, data-driven methodology for the following studies to monitor emerging regional industries and optimize resource allocation within the first-launch economy model.

Keywords

First-Launch Economy, Industrial Performance, LSTM Neural Network, Digital Intelligence, Rural Revitalization, Time-Series Forecasting

1. Introduction

The rapid development of the digital economy has become a pivotal driver of socioeconomic transformation worldwide, particularly in emerging and transitional contexts. As the study of Wang et al. (2025) indicates, defined broadly as economic activities enabled by digital technologies, digital infrastructure, and data analytics, the digital economy enhances productivity, reduces transaction costs, and reshapes traditional industrial structures by increasing connectivity and information flows. While its macro-economic role has been widely studied, the implications of digital transformation for rural industrial development and policy-driven economic models have received comparatively less scholarly attention. Many previous studies emphasize that this gap is notable in the context of China's rural revitalization strategy, where digitalization is increasingly viewed not only as an enabler of infrastructure and services but also as a catalyst for industrial upgrading and economic resilience (Liu et al., 2024a; Tian et al., 2023; Wang et al., 2023; Zhang et al., 2023).

According to the former study of Liu et al. (2024b), the rural revitalization has become a central policy paradigm in China in the post-poverty-alleviation era, aiming to achieve a prosperous industry, a livable ecology, cultural advancement, effective governance, and improved rural livelihoods. Within this framework, from e-commerce platforms to mobile digital finance, the digital tools are being mobilized to overcome traditional barriers to market access, optimize supply chains, and stimulate local entrepreneurship. Existing empirical evidence confirms a significant positive relationship between the development of the digital economy and dimensions of rural revitalization, including industrial prosperity and economic dynamism (Lu et al., 2024; Deng et al., 2024; Cen et al., 2022). However, most studies adopt macro or meso perspectives, employing aggregate indices or panel regression models to evaluate broad digital economy effects. Due to this, such approaches tend to overlook micro-level industrial dynamics and sector-specific efficiencies that emerge under targeted economic initiatives.

One such initiative of the first-launch economy describes the intentional stimulation and formation of nascent industries in designated counties through coordinated policy support, market linkages, and resource integration. As a localized, policy-oriented economic innovation, the first-launch economy requires more granular, data-driven evaluation frameworks to capture temporal performance and guide strategic enhancements (Liu, 2019; Luo et al., 2023; Paszkowski & Sarniak, 2019). Due to the traditional evaluation metrics often struggling with heterogeneous data quality, missing values, and low-frequency reporting typical of emerging rural in-

dustries, there is a methodological need for structured data acquisition, cleaning, and modeling approaches capable of integrating multi-source temporal data into coherent analytical streams suitable for performance assessment.

As a countermeasure, specifically deep learning models designed for time-series prediction, advances in machine learning offer promising tools for handling complex, nonlinear temporal patterns in economic indicators. As Siami-Namini & Namin (2018) indicate, the Long Short-Term Memory (LSTM) neural network architecture has demonstrated superior forecasting capability relative to traditional time-series models when capturing long-term dependencies in sequential data, while the application of LSTM in economic forecasting contexts, such as export dynamics and industrial output, highlights its potential to improve predictive accuracy and inform intervention strategies (Adamopoulos et al., 2025; Haider et al., 2019; Hamiane et al., 2024; Pyo, 2025; Zheng & Li, 2023). Nevertheless, the application of LSTM to evaluate the efficiency of policy-driven rural industrial transformations remains underexplored.

This study addresses these gaps by proposing a comprehensive, digital-intelligence-driven framework that integrates multi-source data acquisition, rule-based time-series cleaning, and LSTM-based forecasting to evaluate and predict the efficiency of the first-launch economy. Utilizing the plush toy industry in Ankang City, China (a representative case initiated in 2018), this research constructs annual industrial performance indicators by synthesizing official statistics and web-crawled enterprise data. By evaluating historical trends and forecasting future developments, the framework in this study provides both an empirical assessment of industry performance and actionable insights for policymakers seeking to optimize rural industrial support strategies. Upon the above, this research contributes to a more nuanced understanding of how digital intelligence can facilitate effective, evidence-based economic policies in the development of rural revitalization.

2. Data Processing and Reconstruction Methodology

2.1. Export Amount of the Toy Industry

According to many previous studies, China's toy export industry has experienced a long-term upward trajectory since the mid-1990s, reflecting both the country's manufacturing capacity expansion and its deep integration into global value chains (Chen et al., 2016; Kee & Tang, 2016). As shown in **Figure 1**, based on the data from the National Bureau of Statistics, which indicates that this evolution can be delineated into distinct phases: a period of steady accumulation followed by accelerated growth subsequent to China's accession to the World Trade Organization in 2001. This earlier phase was predominantly driven by cost advantages, economies of scale, and export-oriented industrial policies, solidifying China's status as the world's premier toy manufacturer. A significant structural shift is observable post-2015, where export values entered a rapid expansionary phase, culminating in a historical peak around 2021. This surge was underpinned by the diversification of toy categories and a transition toward higher value-added pro-

duction segments, demonstrating the sector's resilience despite temporary global economic shocks.

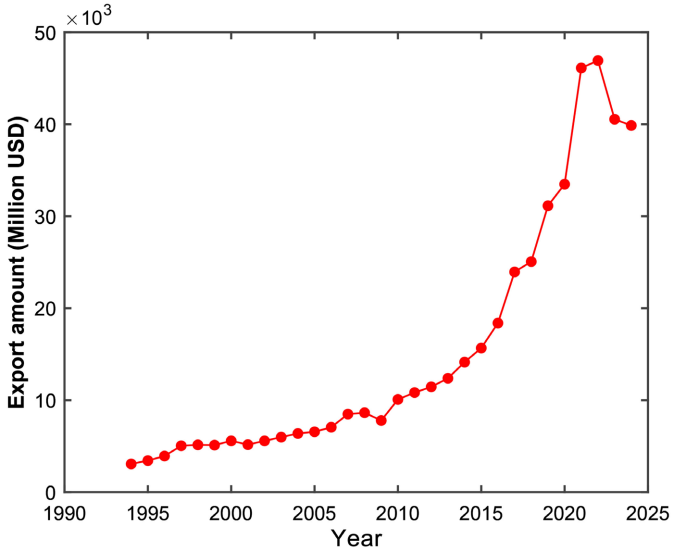


Figure 1. Economic data of China's toy export industry.

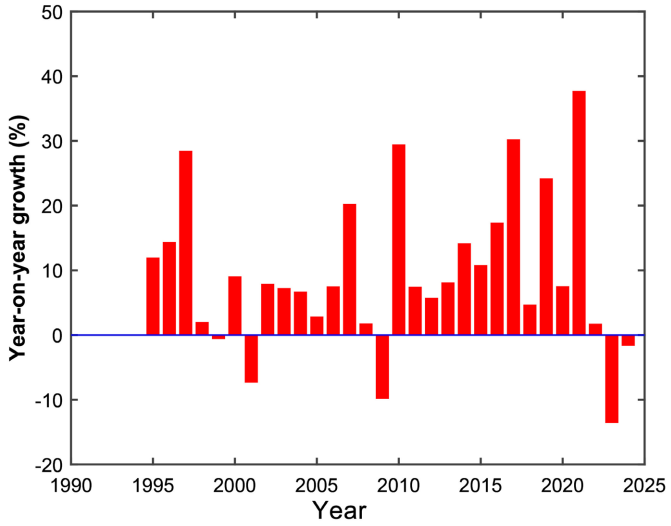


Figure 2. Year-on-year growth.

Synthesis of Figure 1 and Figure 2, while the aggregate trend remains positive, the results of year-on-year growth in Figure 2 reveal the pronounced cyclical volatility and structural constraints underlying this growth. The year-on-year growth series highlights that China's toy exports are highly sensitive to external macro shocks, due to early negative fluctuations aligning with regional disruptions such as the Asian financial adjustments, while significant volatility is evident during the 2008-2009 global financial crisis. Furthermore, the sharp correction observed after the 2021 peak (visible in both Figure 1's decline and Figure 2's growth deceleration) signals a departure from purely scale-driven expansion. As noted in

recent literature, this downturn reflects mounting external pressures, including rising labor costs, supply chain disruptions, and intensifying trade frictions (Gereffi, 2019; Kaplinsky & Morris, 2016). The dynamics above suggest that the industry is transitioning from a low-cost volume regime to a complex phase shaped by input cost inflation and geopolitical reconfiguration.

Within this macroeconomic context, the national toy export trend provides a critical benchmark for evaluating emerging regional industries under the first-launch economy framework. Compared with the long-established and export-oriented national toy industry, newly developed regional clusters (such as the plush toy industry in Ankang City) operate at a significantly smaller scale and are primarily driven by domestic demand, policy support, and localized industrial linkages. Due to this, incorporating national toy export data into the analytical framework enables a relative performance assessment, allowing regional industrial performance to be examined against broader market dynamics. This comparative perspective is essential for distinguishing endogenous growth effects generated by first-launch economy policies from exogenous macroeconomic fluctuations, thereby enhancing the robustness of subsequent efficiency evaluation and forecasting analyses.

2.2. Technical Route and Data Processing Framework

In this study, the economic data are extracted spanning the period from 2018 to 2026 to evaluate the first-launch economy performance. As shown in **Figure 3**, the dataset integrates multi-source information from the National Bureau of Statistics alongside enterprise-level records scraped via standardized APIs from Baidu, Qichacha, Tianyancha, and Aiqicha. Across the multi-source-based dataset, the primary variables collected include enterprise count, production scale, and registered capital. Moreover, this study duplicates records generated by overlapping platform sources, which were removed using time-based primary keys, and imputes missing temporal observations using the Last Observation Carried Forward (LOCF) method to ensure data integrity. Because the annual production totals for 2024 (6.1 billion CNY) and 2025 (7.5 billion CNY) are verified, directly observed anchors sourced from local government and national media reports, it is critical to distinguish between directly observed data and synthesized data in this study. Upon the approach notes above, all monthly data points (a total of 96 anchor observations) are reconstructed and allocated based on these verified annual totals.

To ensure the reliability and analytical validity of the subsequent efficiency evaluation and forecasting, this study constructs a robust Extract-Transform-Load (ETL) pipeline by integrating multi-source data from the National Bureau of Statistics, major e-commerce and business information platforms, and web-crawled records. These data sources differ substantially in temporal resolution, data structure, and statistical standards, ranging from officially published macroeconomic indicators to firm-level registration and transaction information. As **Figure 3** illustrates, the ETL process begins with data extraction through standardized API access and automated web crawling, followed by unified schema mapping to har-

monize variable definitions and time formats. This integration strategy enables heterogeneous raw data to be transformed into a coherent time-series dataset reflecting industrial activity intensity, thereby mitigating the inconsistencies commonly observed in single-source datasets.

```
>>> # config.py
... 
... # Location List
... LOCATIONS = [
...     "陕西", "安康"
... ]
... 
... # Target industry keywords
... KEYWORD = "毛绒玩具,月度生产总值"
... 
... # API endoints
... API_NBS = "https://data.stats.gov.cn/index.htm"
... API_Baidu = "https://www.baidu.com/"
... API_Google = "https://www.google.com/"
... API_QCC = "https://www.qcc.com/"
... API_TYC = "https://www.tianyancha.com/"
... API_AIQI = "https://aiqicha.baidu.com/"
... 
... # Header + Token
... HEADERS_TEMPLATE = {
...     "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64)",
...     "Accept": "application/json",
...     # "Authorization": "Bearer YOUR_API_TOKEN"
... }
... 
... # Time range
... START_YEAR = 2018
... END_YEAR = 2026  # Latest
```

Figure 3. Structure of the crawled system (example code).

Regarding the transformation stage, systematic time-series standardization and rule-based cleaning algorithms are applied to address missing values, temporal misalignment, and duplicate records. Specifically, all timestamps are normalized to a monthly frequency to balance data availability and temporal granularity. Due to LOCF preventing look-ahead bias in predictive modeling, and being suitable for low-frequency, gradually evolving industrial statistics, missing observations in this study are handled using forward-filling with this approach. Formally, for a time series x_t, missing values are imputed as:

$$x_t = x_{t-1}, \text{ if } x_t \text{ is missing} \tag{1}$$

Duplicate records arising from overlapping data sources are resolved using time-based primary keys, ensuring that each monthly observation is unique. This standardized temporal structure provides a stable foundation for subsequent aggregation and modeling procedures.

To construct a representative indicator of industrial activity, an exhaustive aggregation strategy is employed within the ETL pipeline. Such as enterprise count, production scale, and registered capital, through brute-force enumeration across multidimensional parameter spaces, monthly production value indicators are generated using group-by aggregation operations. Conceptually, the aggregated production value PV_t at month T can be expressed as:

$$PV_t = \sum_{i=1}^{N_t} w_i \cdot v_{i,t} \tag{2}$$

where $v_{i,t}$ denotes the activity measure of firm i at time t, and w_i represents its corresponding weight derived from enterprise attributes. The exhaustive tra-

versal in the data processing framework ensures that all valid parameter combinations are incorporated without subjective weighting bias, thereby enhancing the robustness of the constructed time series. The resulting monthly production value serves as a key input for efficiency evaluation and LSTM-based forecasting, enabling the model to capture nonlinear temporal dynamics and structural changes in the first-launch economy. In this study, the data-driven technical route in the data processing framework bridges raw multi-source information and advanced predictive analytics, providing a transparent and replicable methodological foundation for assessing regional industrial performance.

2.3. MPV Reconstruction and Analysis

To construct a monthly time series of production value (MPV) for Ankang's plush-toy cluster, this study anchors the reconstruction to publicly reported annual totals. Let A_y denote the official/anchored annual production value for year y (in CNY). For this study, the key anchors are $A_{2024}=6.1$ (billion CNY) and $A_{2025}=7.5$ (billion CNY), as reported by local government and national media sources (Ankang Plush Toy Industry Inspection Group, 2018; Zhang, 2024). The monthly baseline allocation $m_{y,m}$ for month m in year y is constructed by applying an interpretable seasonal weight vector $s=[s_1,\cdots,s_{12}]$ (Discussions in this study indicate that the weights reflect the established "8 - 11 month export/production crest" characteristic of the plush toy global supply chain) and normalizing so that $\sum_{m=1}^{12} s_m=1$. The calculate method can be illustrated as **Equation (3)** shown below:

$$\tilde{m}_{y,m}=A_y\cdot s_m,\ s_m\geq 0,\ \sum_{m=1}^{12} s_m=1 \tag{3}$$

Because 3% stochastic perturbation is introduced to simulate realistic month-to-month production/order noise, which prevents the neural network from overfitting to a smooth artificial curve, a small stochastic perturbation $\varepsilon_{y,m}\sim\mathcal{N}\left(0,\left(\alpha\tilde{m}_{y,m}\right)^2\right)$ with $\alpha=0.03$ is added to each month in **Equation (3)** to reflect realistic month-to-month production/order noise: $m_{y,m}=\tilde{m}_{y,m}\left(1+\eta_{y,m}\right)$, $\eta_{y,m}\sim\mathcal{N}\left(0,\alpha^2\right)$. This anchoring, combined with seasonal allocation, yields an initial, explainable monthly series that is consistent with verified annual public anchors and sectoral seasonality.

Following the study by Shi et al. (2026), a multilayer perceptron regressor (MLP Regressor) is fitted as a data-driven smoother to produce a continuous, smooth, non-linear monthly series amenable to downstream modeling. The MLP input vector for time index t (months since 2018-01) is $x_t=\left[t,\sin\left(2\pi m/12\right),\cos\left(2\pi m/12\right)\right]$ to capture linear trend and annual seasonality. The MLP defines a parametric mapping $\hat{y}_t=f_\theta\left(x_t\right)$ where f_θ is a feed-forward network with parameters θ. Training minimizes regularized mean squared error, which is shown in **Equation (4)**:

$$\mathcal{L}\left(\theta\right)=\frac{1}{N}\sum_{t=1}^{N}\left(y_t-\hat{y}_t\right)^2+\lambda\left\|\theta\right\|_2^2 \tag{4}$$

where y_t are the anchored noisy monthly targets, λ is an L2 penalty (e.g., $\lambda = 10^{-4}$), and optimization is performed with Adam (learning rate lr, early stopping on validation loss). Practical hyperparameters used for reproducibility: two hidden layers (64.32 neurons), ReLU activations, solver = "adam", batch size = 32, max epochs = 500, early_stopping = True, validation_fraction = 0.2. The choice of MLP (rather than sequence models) in this study is justified by the constructive nature of the monthly targets (anchored, not dense historical observations) and by the goal, as Goodfellow et al. (2016) indicate, that smoothing and non-linear interpolation, rather than pure sequence extrapolation. Incorporating the USD/CNY exchange rate analysis in **Table 1**, the model produces the blue series (MPV (MLP)) shown in **Figure 4** and provides a smooth, interpretable monthly trajectory consistent with yearly anchors.

Table 1. Official PBOC/CFETS USD/CNY exchange rate (2025-2026).

Month	Opening (Approx)	Closing (Approx)	Monthly High	Monthly Low
2025-01	7.2981	7.1809	7.3315	7.1809
2025-02	7.2878	7.2828	7.2888	7.2500
2025-03	7.2340	7.2567	7.2567	7.2340
2025-04	7.3388	7.2860	7.3388	7.2860
2025-05	7.2706	7.2037	7.2706	7.2037
2025-06	7.1802	7.1636	7.1802	7.1636
2025-07	7.1738	7.1764	7.1764	7.1600
2025-08	7.1828	7.1530	7.1828	7.1530
2025-09	7.1414	7.1338	7.1414	7.1129
2025-10	7.1190	7.1169	7.1264	7.1000
2025-11	7.1186	7.1022	7.1186	7.0867
2025-12	7.0717	6.9961	7.0717	6.9961
2026-01	7.0051	6.9678	7.0051	6.9678

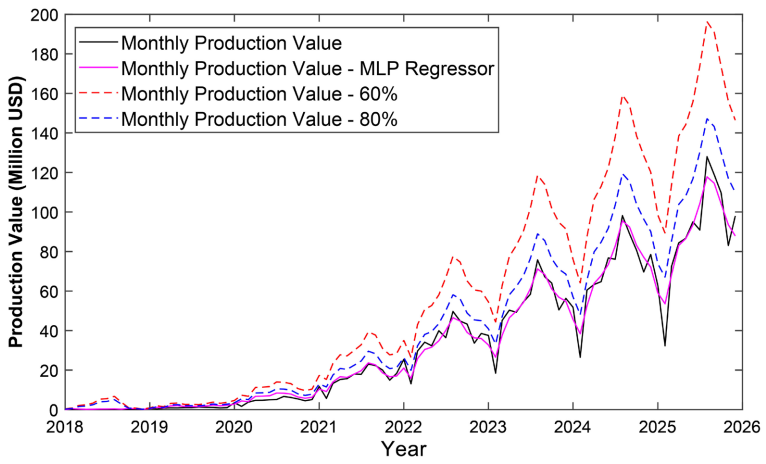


Figure 4. Production value of the plush toy industry in Ankang City.

Through Figure 4, province-level scenario construction follows a transparent algebraic rule to express uncertainty in Ankang's share of provincial output. If p denotes the assumed Ankang share of Shaanxi provincial plush-toy output, the implied province annual production is $P_{year}^{province} = A_y / p$. Due to the parameters $p = 0.60$ and $p = 0.80$ being selected to establish rigorous conservative and optimistic boundary conditions for Ankang's output relative to the entire Shaanxi province, two situations were examined in this study, as shown in Figure 4. For illustration, $A_{2024} = 6.1$ (billion CNY) implies $P_{2024}^{province} = 6.1(\text{billion CNY})/0.6 = 10.167(\text{billion CNY})$ under $p = 0.60$ and $P_{2024}^{province} = 6.1(\text{billion CNY})/0.8 = 7.625(\text{billion CNY})$ under $p = 0.80$. Conversion from RMB to USD uses the reference exchange rate $r = 7.1697$ (USD/CNY), which can be calculated based on the USD/CNY exchange rate data shown in Table 1. Thereby, the value of MPV in million USD is computed as:

$$MPV_t(\text{million USD}) = \frac{m_{y,m}}{10^6} \cdot \frac{1}{r} \tag{5}$$

Using **Equation (5)** above, the 2024 annual anchor $A_{2024} = 6.1$ (billion CNY) corresponds to ≈850.9 (million USD) annually and an average monthly value ≈ 70.9 (million USD) numbers that align with the MPV seasonal profile shown in Figure 4.

Referring to the results in Figure 4, interpreting the MPV trajectories through the First-Launch Economy lens emphasizes policy causality and practical implications. The reconstructed series in Figure 4 exhibits negligible activity in 2018, stepwise growth during 2019-2021, and rapid expansion in 2022-2024, including industrial relocation, community factories, targeted subsidies, and logistics support, such behavior consistent with a policy-driven, cluster-forming process. Based on the discussion above, the quantitative reconstructions provide empirical validation for narrative reports (local government and media) that Ankang's cluster matured rapidly after programmatic intervention. From a policy perspective, the MPV results justify continued prioritization of supply-side measures (workforce training, quality upgrade, design/branding support) and demand-side facilitation (export channels, trade fairs) to convert scale into sustainable, higher-value growth rather than transient volume increases.

3. LSTM Predictive Modeling Methodology

3.1. Time-Series Reconstruction

Before establishing the LSTM structure, this study constructs a high-resolution daily time series $y_{hr}(t)$ ($N = 2921$ daily points spanning 2018-2026) from a sparse set of 96 monthly anchor observations previously generated via an MLP-based reconstruction. This densification is motivated by the architectural requirements of LSTM neural networks, which rely on long, densely sampled sequences to reliably capture short- and long-range dependencies, intra-year seasonal cycles, and high-frequency volatility that sparse anchors fail to expose. To construct a continuous mathematical space across the full timeline, this study employs a hy-

brid approach that 1) establishes a shape-preserving baseline via piecewise cubic Hermite interpolation (PCHIP), 2) injects controlled, locally-weighted high-frequency details, and 3) enforces alignment at anchor indices to preserve verified totals. As noted by Fawaz et al. (2018), the approach balances fidelity to verified anchors with the synthetic enrichment necessary for the LSTM training process. Since the densification and normalization processes are applied to the entire timeline, the LSTM is trained to extrapolate the structural dynamics of this anchor-constrained synthetic series rather than forecasting raw, independent future observations.

As step 3, which is illustrated in **Figure 5**, the codes in MATLAB use PCHIP to obtain $y_{base}\left(x_{hr}\right)$, a monotonicity-preserving cubic Hermite interpolation. PCHIP computes, for each interval $\left[x_i, x_{i+1}\right]$, a Hermite cubic polynomial $H_i\left(t\right)$ that matches function values and chosen node slopes m_i, while the calculation can be expressed as **Equation (6)** below:

$$H_i\left(t\right) = y_i h_{00}\left(\tau\right) + m_i h_{10}\left(\tau\right) h + y_{i+1} h_{01}\left(\tau\right) + m_{i+1} h_{11}\left(\tau\right) h \tag{6}$$

with $\tau = \left(x - x_i\right)/h$ and Hermite basis functions h_{00}, h_{10}, h_{01}, h_{11}. PCHIP is preferred over natural cubic splines because it prevents overshoot and preserves the qualitative shape (no spurious oscillations) due to a desirable property when anchors represent real economic aggregates that interpolation artifacts should not violate (Fritsch & Carlson, 1980). Across this step, mathematical guarantees of monotonicity and shape-preservation reduce the risk of introducing implausible trends that could mislead the LSTM.

```
%% 3. Monotonic Cubic Interpolation (Hard Constraints)
% Ensures the base curve strictly passes through all original data points.
y_base = pchip(x, y, x_hr);

%% 4. Zero-at-Knot Weighting Function
% Construct a weight function that vanishes at original sample points to maintain integrity.
dx = min(abs(x_hr - x'), [], 2);    % Euclidean distance to the nearest original point
w  = (dx ./ max(dx)).^2;            % Normalized weighting: 0 at knots, peaking in between

%% 5. Synthetic Residual Generation (Heuristic Detail Synthesis)
% Generates high-frequency features modulated by local curvature and knot proximity.
curv = abs(gradient(gradient(y_base)));   % Local second-order derivative (curvature)
amp  = 0.01 * mean(y) * curv / max(curv); % Scale amplitude based on local dynamics
rng(20260206);                            % Reproducibility seed

% Synthesize multi-harmonic noise as a proxy for learned high-frequency features
detail = amp .* w .* ...
    (sin(2*pi*5*x_hr/96) + 0.4*sin(2*pi*11*x_hr/96) + 0.2*randn(N,1));

%% 6. Final Signal Composition
% Superimpose the generated residuals onto the interpolated base signal.
y_hr = y_base + detail;

%% 7. Knot Alignment Enforcement (Numerical Safety)
% Explicitly overwrite interpolated values at knot indices to prevent floating-point drift.
idx = round((x-1)/(96-1)*(N-1)) + 1;
y_hr(idx) = y;
```

Figure 5. Code structure of time-series reconstruction.

Also, referring to **Figure 5**, to confine synthetic details away from anchor points, the code computes the distance $d\left(t\right) = \min_i \left|x_{hr}\left(t\right) - x_i\right|$ to the nearest anchor and defines a quadratic weight as shown in **Equation (7)** below:

$$\omega\left(t\right) = \left(\frac{d\left(t\right)}{\max d}\right)^2 \tag{7}$$

where $\omega = 0$ at anchors and increases between anchors. This multiplicative mask ensures exact anchor preservation. The detail amplitude is modulated by a curvature proxy $c(t) \approx \left| y''_{base}(t) \right|$, computed numerically by two successive gradient operators. The amplitude scaling can be expressed as **Equation (8)** below:

$$amp(t) = \alpha \bar{y} \frac{c(t)}{\max c} \tag{8}$$

with $\alpha = 0.01$ controlling overall magnitude and $\bar{y}$ the mean of the original anchors. This curvature weighting means higher curvature receives larger high-frequency augmentation, which is consistent with empirical observations that real economic series often show stronger short-term variability around structural inflection points.

The codes in Figure 5 are treated as the learnable residual that an LSTM might reproduce, while the synthetic residual is formed as:

$$detail(t) = amp(t)\omega(t)\left[\sin\left(2\pi \frac{f_1 t}{T} \right) + \beta \sin\left(2\pi \frac{f_2 t}{T} \right) + \gamma \varepsilon_t \right] \tag{9}$$

with $f_1 = 5$, $f_2 = 11$, $T = 96$, $\beta = 0.4$, $\gamma = 0.2$ and $\varepsilon_t \sim \mathcal{N}(0,1)$. These deterministic sinusoidal components embed chosen mid-range spectral energy (periodicities approximating sub-annual cycles when mapped to daily resolution), while Gaussian noise supplies stochastic variability. The combination creates a controlled broadband signal envelope whose local amplitude follows economic curvature and whose fine structure is confined away from anchors by A. As Goodfellow et al. (2016) emphasized, setting a fixed seed (rng) ensures reproducibility, which is essential for scientific verification and for repeated model selection runs.

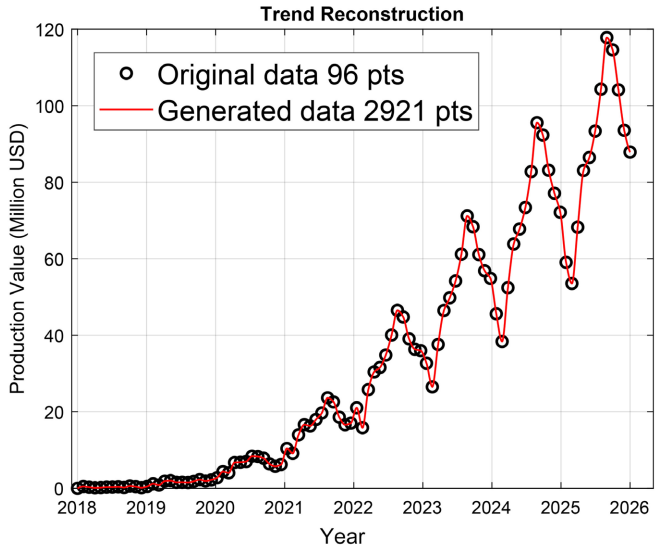


Figure 6. Time-series reconstruction results.

The synthesized series is $y_{hr}(t) = y_{base}(t) + detail(t)$, with enforced replacement $y_{hr}(t_i) = y_i$ at anchor indices t_i, while the calculated results are shown in

Figure 6. According to the curve in Figure 6, the hard alignment prevents long-term drift from known validated points and preserves policy-anchored annual totals or verified monthly observations. For LSTM training, the dataset is then windowed into supervised examples using a sliding window of length L as each sample $X^{(i)} = [y_{t-i-L+1}, \cdots, y_{t-i}]$ predicts y_{t-i+1} (one-step) or the next h steps (multi-step). Prior to training, standard preprocessing (z-score normalization $z = (y - \mu)/\sigma$, optional differencing $\Delta_{y_t} = y_t - y_{t-1}$, and seasonal deseasonalization if desired) stabilizes variance and accelerates convergence. Referring to the prior study of Iwana & Uchida (2021), the dense daily sampling yields many overlapping windows, increasing the effective sample size for gradient-based optimization and allowing the LSTM to learn both intrayear seasonal patterns and long-range dependencies.

3.2. LSTM Neural Network Structure

According to many prior studies, Long Short-Term Memory (LSTM) networks were designed to overcome the principal limitations of vanilla recurrent neural networks (RNNs) for long-sequence modeling, namely the vanishing and exploding gradient problems that prevent learning of long-range dependencies (Hochreiter & Schmidhuber, 1997; Pascanu et al., 2013). **Figure 7** summarizes the per-time-step architecture used in this study that includes an explicit cell state C_t (the "memory channel") that runs through time with near-linear interactions, while three multiplicative gates (forget, input, output) control what information is removed, written, and exposed. Specifically, to maintain consistency with the defined LSTM neural network, the input vector at time t, denoted as x_t, is explicitly defined as a 2D sequence as $x_t = [MPV_t, NE_t]$, where MPV_t is the normalized Monthly Production Value and NE_t is the corresponding normalized national export benchmark. Upon the description above, the structure of the LSTM neural network in this study creates a parameterized pathway that can preserve gradients over many time steps when the forget gate takes values near one, enabling learning of slowly evolving trends (long-term MPV growth) concurrently with shorter-term dynamics (seasonality, shocks) embedded in the dense daily series.

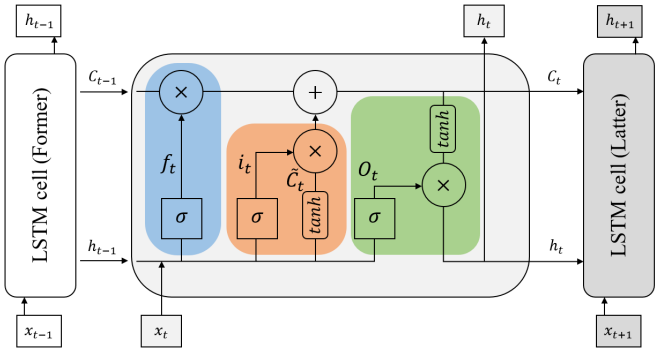


Figure 7. Cell structure of LSTM neural networks.

This study adopts the standard vectorized LSTM equations at time t, while concatenation $[h_{t-1}, x_t]$ denotes appending the previous hidden state and the current input, as **Equations (10-15)** illustrate below:

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{10}$$

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{11}$$

$$\tilde{C}_t = \tanh\left(W_C\left[h_{t-1}, x_t\right] + b_C\right) \tag{12}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{13}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{14}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{15}$$

As illustrated above, the **Equations (10-15)** implement three separate roles: f_t gates the retention of past memory, i_t gates the incorporation of candidate updates $\tilde{C}_t$, and o_t gates how much of the cell content is revealed to the hidden state. For the MPV forecasting process, the input x_t can be univariate (normalized MPV) or multivariate (MPV with national export benchmarks, calendar encodings sin/cos, and lagged national export in this study). The readout (forecast) is typically produced by applying a linear layer to h_t.

From the learning dynamics viewpoint, the multiplicative of LSTM cell gates directly influences gradient propagation during backpropagation through time (BPTT). This study also considers the derivative of the loss $\mathcal{L}$ with respect to the previous cell state as **Equation (16)** shows below:

$$\frac{\partial \mathcal{L}}{\partial c_{t-1}} = \frac{\partial \mathcal{L}}{\partial c_t} \odot f_t + \left(\text{terms from paths via } h_t \text{ and gates}\right) \tag{16}$$

If f_t is close to 1, the factor $\odot f_t$ preserves the magnitude of $\partial \mathcal{L} / \partial c_t$, thereby mitigating vanishing gradients; conversely, adaptive f_t values allow selective forgetting. Referring to the study of Greff et al. (2016), the mathematical property is the principal reason LSTM can carry long-range signals (trend, anchor constraints) while still learning to respond to short-term inputs (seasonal peaks) present in the synthesized daily MPV data.

Algorithmically, training an LSTM for MPV forecasting follows supervised sequence learning with sliding windows in this study. Let window length L and forecast horizon H be chosen according to domain temporal scales (such as $L = 90$ - 365 days to capture intra-annual seasonality and medium-term dynamics). Construct input windows $X^{(n)} = \left[x_{t-L+1}, \cdots, x_1\right]$ with target $Y^{(n)} = \left[y_{t+1}, \cdots, y_{t+H}\right]$. Optimize parameters $\theta = \left\{W_*, b_*\right\}$ by minimizing a regularized loss, typically mean squared error (MSE) plus L2 penalty, by **Equation (17)** below:

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{H}\left(y_{t^{(n)}+k} - \hat{y}_{t^{(n)}+k}\right)^2 + \lambda\left\|\theta\right\|_2^2 \tag{17}$$

Based on **Equation (17)** above, by using gradient-based optimizers of Adam,

apply gradient clipping (clip global norm to τ) to control exploding gradients (Pascanu et al., 2013). Regularization strategies (dropout on inputs and recurrent connections, early stopping on walk-forward validation) are critical because the training data here include synthetic high-frequency components created by the preceding densification step.

Based on the specific production value data of the plush toy industry in Ankang City, the design of the LSTM structure is begin with a single or stacked LSTM architecture (1 - 2 layers) with hidden dimensionality A in the range 32 - 256 depending on dataset size and compute budget; use batch normalization or layer normalization if needed, recurrent dropout (0.1 - 0.3) for regularization, and normalize inputs (z-score) to stabilize gate activations. Moreover, when training on the 2921-point daily series, experiment with one-step and multi-step forecasting setups, and use walk-forward validation (rolling origin) to simulate realistic forecasting and detect regime sensitivity. Referring to the study of Gal & Ghahramani (2016), for probabilistic or uncertainty-aware forecasts, ensemble multiple trained LSTMs or use Monte Carlo dropout to obtain predictive intervals.

Furthermore, evaluation must report policy-relevant metrics ($R^2, RMSE, MAE, MBE$) and perform sensitivity checks specific to the First-Launch Economy context: 1) how forecasts respond to alternative synthetic seeds and detail amplitudes; 2) whether the model respects annual anchor constraints (reconciling predicted aggregated monthly/daily totals with known yearly anchors); and 3) the contribution of exogenous covariates (national export benchmark) to predictive skill (use ablation studies and attribution tools such as SHAP for interpretability). Together, the gate mechanics, training protocols, and evaluation practices described above provide a rigorous theoretical and practical foundation for inserting the reconstructed high-resolution MPV data into LSTM networks for robust forecasting of First-Launch Economy performance.

4. Results and Discussion

4.1. Implementation and Configuration

As delineated in the comprehensive workflows presented in **Figures 8-11**, the coding system implemented in this study serves as the computational technology of the proposed LSTM neural network, systematically operationalizing the transition from raw signal processing to predictive modeling. Specifically, the coding framework addresses the dimensionality and temporal complexity of the reconstructed high-resolution MPV series by employing a rigorous sliding window strategy. The step restructures the univariate time-series data into a supervised learning format, creating input-output pairs (X_t, Y_{t+1}) that allow the model to learn temporal mapping functions rather than static correlations.

Furthermore, the neural network model explicitly incorporates the LSTM's distinct gating mechanisms (forget, input, and output gates) within the hidden layers. As verified by the training dynamics shown in **Figures 8-11**, the LSTM structural design in this study enables the system to selectively retain or discard information

over long sequences, thereby mitigating the vanishing gradient problem inherent in standard Recurrent Neural Networks (RNNs) when processing high-frequency data.

```
%% 2. Data analysis & Windowing
num_samples = length(original_data);   % Should be 2921 based on the original data
kim = 30;                              % Input Sequence Length (Lookback window)
zim = 1;                               % Predictor step (Forecast horizon)

% Arrange data (Sliding Window Construction)
res = [];
for i = 1: num_samples - kim - zim + 1
    res(i,:) = [reshape(original_data(i:i + kim - 1), 1, kim), original_data(i + kim + zim -1)];
end
```

```
%% 3. Divide train data and test data
total_obs = size(res, 1);
train_ratio = 0.75;
num_train = floor(total_obs * train_ratio);

% Partitioning
P_train_raw = res(1:num_train, 1:kim)';
T_train_raw = res(1:num_train, kim+1)';
P_test_raw  = res(num_train+1:end, 1:kim)';
T_test_raw  = res(num_train+1:end, kim+1)';

M = size(P_train_raw, 2); % Train sample size
N = size(P_test_raw, 2);  % Test sample size
```

Figure 8. Data analysis & Data divide.

```
%% 4. Data Normalization (Z-Score as requested)

mu_X = mean(P_train_raw, 2);
sigma_X = std(P_train_raw, 0, 2);

% Z-score Standardization for Inputs
P_train = (P_train_raw - mu_X) ./ sigma_X;
P_test = (P_test_raw - mu_X) ./ sigma_X;

% Z-score Standardization for Targets (Optional but recommended for convergence)
mu_Y = mean(T_train_raw);
sigma_Y = std(T_train_raw);

t_train = (T_train_raw - mu_Y) ./ sigma_Y;
t_test_norm = (T_test_raw - mu_Y) ./ sigma_Y;
```

```
%% 5. Data Format Transformation for LSTM
% Convert to Cell Array format which is standard for MATLAB LSTM sequences
p_train_cell = cell(M, 1);
p_test_cell = cell(N, 1);

for i = 1 : M
    p_train_cell{i, 1} = P_train(:, i);
end
for i = 1 : N
    p_test_cell{i, 1} = P_test(:, i);
end
```

Figure 9. Data Normalization (Z-Score as requested).

```
%% 6. Structure of the model
% Architecture Reference:
% - Hidden dim: 32-256 (Selected 128)
% - Batch Normalization: Yes
% - Dropout: 0.1-0.3 (Selected 0.2)

inputSize = kim; % 10 features/time steps

layers = [
    sequenceInputLayer(inputSize, 'Name', 'Input')

    % LSTM Layer: Hidden units = 128 (Range 32-256)
    lstmLayer(128, 'OutputMode', 'last', 'Name', 'LSTM')

    % Batch Normalization (as suggested in text)
    batchNormalizationLayer('Name', 'BN')

    % ReLU Activation
    reluLayer('Name', 'ReLU')

    % Recurrent Dropout / Dropout (0.1 - 0.3)
    dropoutLayer(0.2, 'Name', 'Dropout')

    fullyConnectedLayer(1, 'Name', 'FC')
    regressionLayer('Name', 'Output')];
```

Figure 10. Structure of the model.

```
%% 7. Parameter setting
options = trainingOptions('adam', ...
    'MaxEpochs', 1000, ...                          % Epochs
    'MiniBatchSize', 64, ...                        % Batch size
    'InitialLearnRate', 0.005, ...                  % Slightly lower initial rate for stability
    'LearnRateSchedule', 'piecewise', ...
    'LearnRateDropFactor', 0.5, ...
    'LearnRateDropPeriod', 100, ...
    'GradientThreshold', 1, ...                     % Prevent exploding gradients
    'Shuffle', 'every-epoch', ...
    'Plots', 'training-progress', ...
    'Verbose', false);
```

Figure 11. Parameter setting.

As illustrated in the comprehensive workflow in Figure 8, the coding system orchestrates the full cycle of data processing, beginning with the assimilation of the expanded daily series (comprising 2921 pts). To render the univariate time series amenable to the LSTM architecture, the system employs a sliding window technique with a specific lookback period of 30 days. This mechanism systematically transforms the sequential data into a supervised learning format, mapping a historical sequence of length *kim* to a subsequent target value, thereby capturing short-term temporal dependencies.

Especially, a pivotal design element of this coding system is the rigorous chronological stratification of the dataset into training and testing subsets (75%/25%). Unlike random cross-validation, which violates temporal order, this study adheres to a strict temporal partition: the training set encompasses the historical horizon from 2018 to 2024 (representing the first 75% of the data flow) to optimize model parameters, while the test set is strictly reserved for the period from 2024 to 2026 (representing the subsequent 25% of the data flow). This design rationale is critical for simulating a realistic forecasting scenario, effectively preventing data leakage (future information inadvertently influences training) and ensuring that the model's performance is evaluated strictly on unseen, out-of-sample data.

As further detailed in the preprocessing workflow depicted in Figure 9, the coding system incorporates a critical data standardization stage utilizing Z-score normalization prior to network ingestion. This procedure is fundamental for stabilizing the optimization landscape; by rescaling the input features to possess zero mean and unit variance, the system effectively prevents the saturation of activation functions (tanh or sigmoid) and facilitates more efficient gradient descent convergence within the LSTM layers. Following the statistical conditioning, the script executes a structural transformation, converting the numerical tensors into MATLAB's proprietary cell-array sequence format. The final formatting step in Figure 9 is essential to satisfy the input requirements of the trainNetwork solver, ensuring seamless integration with the deep learning computational graph.

As synthesized in the structure of the neural network model in Figure 10 and the hyperparameter configuration in Figure 11, the core of the coding system is an engineered LSTM-based regression network. Referring to Figure 10, the design of 128 hidden units within a single LSTM layer strikes a balance between model expressive capacity and computational parsimony, enabling the capture of high-dimensional temporal dependencies in the MPV series without incurring prohibitive training costs. To enhance the stability and generalizability of the recurrent framework, the code structure integrated a Batch Normalization layer and a ReLU activation function, which collectively accelerate convergence by normalizing the internal distribution of activations and mitigating the gradient vanishing effect. Moreover, a Dropout rate of 0.2 was implemented as a stochastic regularization technique to prevent co-adaptation of neurons, thereby reducing the risk of overfitting on historical training data. According to Figure 11, the training dynamics are further governed by the Adam optimizer, utilizing a piecewise learning-rate schedule to navigate the complex loss landscape. By incorporating a Gradient

Threshold ($\tau = 1$) for gradient clipping, the coding system effectively suppresses the gradient explosion phenomena common in Recurrent Neural Networks (RNNs) during the backpropagation through time (BPTT) process. Also, the combination of a high epoch budget of 1000 and a robust optimization strategy ensures that the model achieves a stable global minimum, providing a reliable foundation for the subsequent long-term predictive analysis.

Following the convergence of the training phase, the coding system executes a dual-phase evaluation protocol (in-sample fidelity and out-of-sample generalizability) to rigorously quantify the model's predictive performance. A critical post-processing step involves the inverse transformation (de-normalization) of the network outputs, ensuring that subsequent error metrics (*RMSE* and *MAE*) are interpreted on the original physical scale of the MPV series, while the R^2 assesses the model's capacity to explain observed variance. To further scrutinize the residual structure, the implementation incorporates diagnostic visualizations, where time-series trajectory plots facilitate local error analysis and error histograms verify whether the residuals approximate a Gaussian distribution, which is a prerequisite for ensuring that no significant non-linear patterns remain uncaptured.

4.2. Training and Testing Results

The layer structure of the LSTM in Table 2 summarizes a compact yet expressive LSTM-based regression pipeline, whose dimensions and parameterization match those of standard gated-recurrent formulations. At its core is a single LSTM layer with $H = 128$ hidden units, as shown in Table 2. Also, because an LSTM computes four separate gate/transform vectors per time step, the learned linear projections are grouped as blocks of size $4H$.

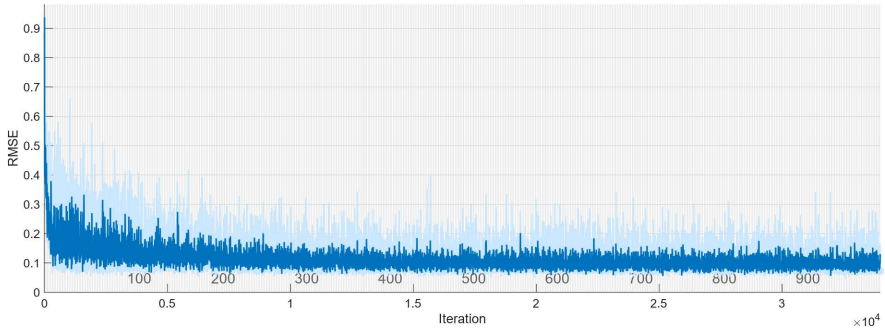
Table 2. Layer information of LSTM.

Layer	Name	Type	Activations	Learnable Sizes		State Sizes	
1	Input Sequence input 2 dimensions	Sequence Input	2(C) × 1(B) × 1(T)	-		-	
2	LSTM LSTM with 128 hidden units	LSTM	128(C) × 1(B)	Input Weights Recurrent Weights Bias	512 × 2 512 × 128 512 × 1	Hidden State Cell State	128 × 1 128 × 1
3	BN Batch normalization with 128 channels	Batch Normalization	128(C) × 1(B)	Offset Scale	128 × 1 128 × 1	Trained Mean Trained Variance	128 × 1 128 × 1
4	ReLU ReLU	ReLU	128(C) × 1(B)	-		-	
5	Dropout 20% dropout	Dropout	128(C) × 1(B)	-		-	
6	FC 1 fully connected layer	Fully Connected	1(C) × 1(B)	Weights Bias	1 × 128 1 × 1	-	
7	Output Mean-squared-error with response "Response"	Regression Output	1(C) × 1(B)	-		-	

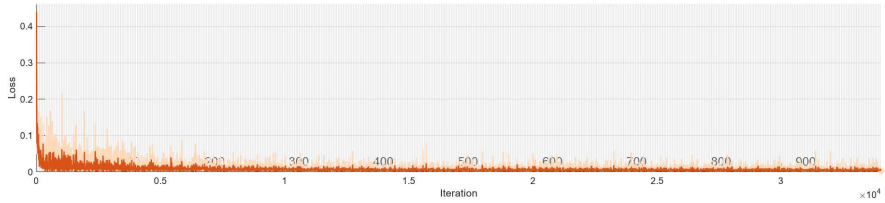
According to Table 2, the learnable weight blocks have shapes in matrix notation as shown in **Equation (18)** below:

$$W_x \in \mathbb{R}^{4H\times D},\ W_h \in \mathbb{R}^{4H\times H},\ b \in \mathbb{R}^{4H} \tag{18}$$

where D is the input dimensionality (the table lists a 2-dimensional sequence input, hence $4H = 512$ rows produce the reported input weight size 512 × 2, recurrent weight size 512 × 128, and bias size 512 × 1). The stated state sizes (hidden state and cell state both 128 × 1) follow directly. Downstream, batch normalization (128 channels) and a ReLU activation are applied before a dropout layer (20%) and a single fully connected output neuron, producing a scalar regression output with MSE loss. As Srivastava et al. (2014) emphasized, this layout combines the LSTM's capacity to capture non-linear temporal dependencies with common deep-learning stabilizers (batch norm, dropout) and a compact readout, while a configuration that balances representational power and computational tractability for the MPV forecasting problem.



(a) Process results of RMSE



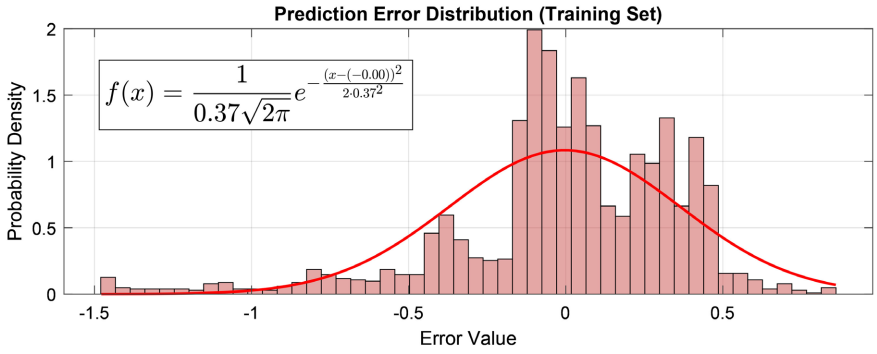
(b) Process results of Loss

Figure 12. Process results (1000 epochs).

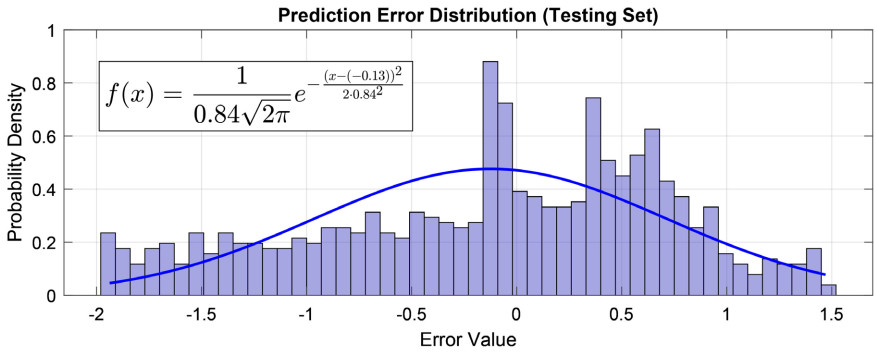
The results in **Figure 12(a)** and **Figure 12(b)** together indicate a classical deep-learning training dynamics, while illustrating a rapid improvement during the initial phase of optimization, followed by a slower convergence and a low-amplitude stationary regime. In both aspects, the early iterations show large, frequently spiking *RMSE* and loss values that decline steeply. After this transient, the traces settle into a narrower band of values with only modest fluctuations, where *RMSE* stabilizes around a substantially lower level, and the training loss oscillates at a low baseline. The high initial variability and the gradually shrinking amplitude of spikes reflect 1) the optimizer reducing coarse structural error (trend and seasonality that are strong in the MPV series), 2) the effect of batch stochasticity (mini-batch gradients produce high variance early on), and 3) the impact of regularizers (dropout,

batch-norm) and gradient clipping that suppress large parameter excursions and thereby reduce loss volatility.

The training-error histogram in this study is shown in Figure 13(a), while the Gaussian curve for the training set illustrates that the fitted parameters are $\mu_{train} \approx 0.00$ and $\sigma_{train} \approx 0.37$. Based on the Gaussian curve indicated in Figure 13(a), the near-zero mean indicates that, on the training data, the LSTM produces essentially unbiased point forecasts ($error = actual - predicted$ is centered at zero), consistent with effective in-sample fitting of the reconstructed MPV trend and seasonal pattern. Also, the relatively small standard deviation (0.37) implies moderate residual dispersion in the training set, which indicates that the model captures the dominant variance components (long-run trend and primary seasonal peaks) but leaves higher-frequency, lower-amplitude residuals. Additionally, visual inspection shows some departure from a perfect Gaussian (slight asymmetry and localized modes), which is expected given the injected synthetic high-frequency detail and the mixture of deterministic sinusoidal residuals plus stochastic noise used during densification.



(a) Training Set

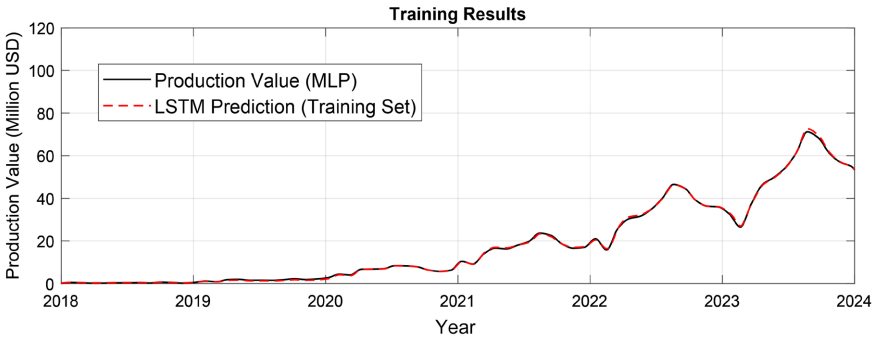


(b) Testing Set

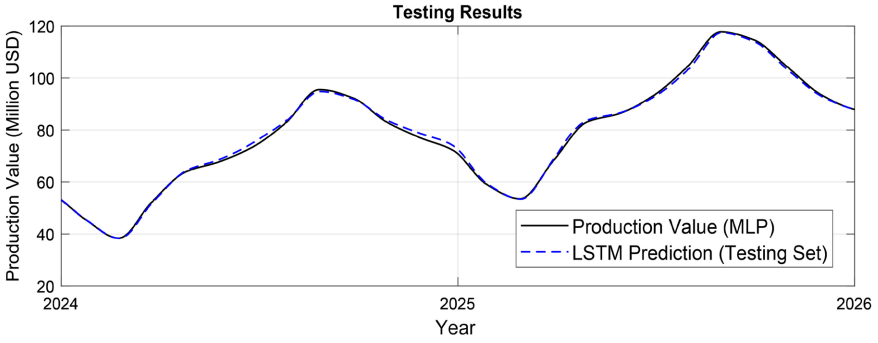
Figure 13. Prediction error distribution.

The testing distribution is overlaid with the Gaussian curve in Figure 13(b), with $\mu_{test} \approx -0.13$ and $\sigma_{test} \approx 0.84$. Compared with the training set, including a larger dispersion (σ roughly doubled) and a small negative bias ($\mu_{test} \approx -0.13$), the test errors show two important differences. The larger variance indicates de-

graded generalization, which indicates the model's predictive uncertainty increases on unseen periods (2024-2026) likely because the synthetic augmentation injects high-frequency patterns whose statistics differ between training and test segments, and real regime shifts (as the rapid industrial scaling, policy changes or post-pandemic demand shifts of the plush toy industry in Ankang City) alter data generating dynamics in the later horizon. According to the Gaussian Curve illustrated in Figure 13(b), the negative mean implies a systematic tendency to overpredict (predictions exceed realized values) on the test set under the current error sign convention.



(a) Results of the training set



(b) Results of the testing set

Figure 14. Results of LSTM prediction.

As shown in Figure 14(a), the graph of the training set illustrates a near-perfect overlap between the LSTM predictions (dashed red) and the anchor-constrained MLP production series (solid black), indicating that the neural network of LSTM successfully captured the dominant long-run trend, inter-annual growth profile, and seasonal phasing embedded in the reconstructed MPV series. The close in-sample fit of the training set in Figure 14(a) is consistent with the neural network of LSTM capacity (single LSTM layer with 128 units), the dense sliding-window supervision derived from the 2921-point sequence, and the preprocessing pipeline (z-score normalization, batch normalization, and dropout) that stabilized optimization. From a methodological perspective, such conformity demonstrates that the LSTM has learned the low-frequency components (trend and annual seasonality) and much of the injected high-frequency structure used during densification.

Moreover, the hard alignment of synthesized daily values to original monthly/annual anchors prevents long-term drift and forces the model to respect policy-relevant totals.

On the other hand, the testing set in Figure 14(b) shows that the LSTM retains good out-of-sample fidelity due to the LSTM predictions (dashed blue) following the anchor-constrained MLP production series (solid black) closely, reproducing the timing and amplitude of the major seasonal peak and trough cycles while exhibiting only modest deviations at some local extrema. Compared with the training period, the test errors are somewhat larger and show slightly broader residual dispersion (as seen previously in error-distribution diagnostics), which indicates reduced certainty when extrapolating to later periods that may include structural shifts (rapid industrial scaling, demand changes, or policy adjustments). Based on the discussion of Figure 14(a) and Figure 14(b), the prediction results in the training set and testing set imply that the LSTM learned robust, policy-relevant dynamics of the First-Launch Economy (e.g., the cluster's expansion pattern) but remains sensitive to regime changes and to the statistical assumptions embedded in the synthetic high-frequency augmentation.

According to Table 3, this study summarizes four complementary error metrics used to evaluate the LSTM's forecasting performance on the MLP anchor-constrained MPV series: the coefficient of determination (R^2), root-mean-square error (*RMSE*), mean absolute error (*MAE*), and mean bias error (*MBE*). These metrics are computed on denormalized predictions (MLP production value series) and together provide a balanced view of explanatory power (variance explained), typical absolute deviations, sensitivity to large errors, and systematic bias, all of which are important when forecasts are used for policy decisions in a First-Launch Economy setting. First of all, R^2 quantifies the fraction of target variance explained by the model and is computed as:

$$R^2 = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2}{\sum_i \left(y_i - \bar{y}\right)^2} \tag{19}$$

Based on **Equation (19)** illustrated, $R^2_{train} = 0.9997$ and $R^2_{test} = 0.9983$ in this study, while both are exceedingly close to 1.0. Such values indicate that the LSTM captures almost all variance present in the reconstructed MPV series on both sets, confirming that the model learns the dominant trend and seasonal components driven by anchor constraints and the densification procedure.

Table 3. Error index.

	Training Set	Testing Set	Ratio
R^2	0.9997	0.9983	−0.14%
RMSE	0.3679	0.8462	56.52%
MAE	0.2658	0.6804	60.93%
MBE	0.0034	0.1251	97.28%

The root-mean-square error (*RMSE*) illustrated in Table 3 quantifies the standard deviation of prediction errors in the original units of MPV, and also provides a direct measure of the typical magnitude of large deviations between MLP anchor-constrained MPV series and predicted series. As shown in **Equation (20)**, *RMSE* is computed as the square root of the mean squared forecasting error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{20}$$

Upon the calculation with **Equation (20)** above, *RMSE* increased from 0.3679 on training to 0.8462 on testing with relative increase of 56.5%. This rise indicates that, although the model explains variance (high R^2), its absolute predictive uncertainty grows out-of-sample.

Moreover, Mean Absolute Error (*MAE*) and Mean Bias Error (*MBE*) complement *RMSE* by respectively quantifying the typical absolute forecast deviation and any systematic directional error, while both are illustrated in Table 3 and computed as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{21}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) \tag{22}$$

As shown in **Equation (21)**, the evaluation value of *MAE* indicates the average magnitude of errors regardless of sign, and in this study, *MAE* rose from 0.2658 (train) to 0.6804 (test), with 60.93% increase as shown in Table 3, reflecting the *RMSE* trend but being less sensitive to large individual errors. The *MAE* change indicates that typical point-forecast deviations roughly doubled when the model was applied to unseen periods, which suggests that the model's expected absolute error under operational conditions is materially larger than suggested by in-sample diagnostics. Referring to Hyndman & Athanasopoulos (2018), for decision-makers in a First-Launch Economy context, *MAE* is often more interpretable than *RMSE* because it maps directly to the average (absolute) production-value forecasting error and can be compared to tolerable policy thresholds.

Furthermore, according to Table 3, the value of *MBE* increases from a near-zero 0.0034 on the training set to 0.1251 on the testing set, with a relative increase of 97% based on **Equation (22)**. A near-zero training *MBE* confirms unbiased in-sample fitting, but the positive testing *MBE* signals a systematic offset in the held-out period (the sign convention determines whether this is over- or under-prediction). Although the absolute magnitude of 0.1251 must be interpreted relative to the MPV scale, its growth highlights a persistent directional error that can compound when aggregated monthly/annually, which is critical when forecasts inform resource allocation or policy evaluation.

In summary, the four different metrics show a model that explains variance extremely well but exhibits substantial growth in absolute error and bias out-of-sample. This pattern is consistent with the study's reliance on synthetic densifica-

tion (anchor-constrained daily augmentation) to create training sequences, due to the LSTM successfully learns the anchor-consistent structure but being sensitive to differences between synthetic residual statistics and real later behavior.

4.3. Summarization

As shown in Figure 15, the complete prediction cycle of the LSTM neural network demonstrates that the forecasts with LSTM (red dashed line) track the anchor-constrained MLP production series (black line) across both the training (2018-2024) and test (2024-2026) horizons. The LSTM neural network in this study captures the long-run growth trend associated with the First-Launch Economy ramp-up in Ankang, distinct inter- and intra-annual seasonal peaks (notably the 8 - 11-month export/production crests), and medium-frequency local fluctuations introduced by the synthetic densification procedure. The alignment in this study is corroborated by the high in-sample explained variance and low training *RMSE*, reflecting the capacity of LSTM to preserve long-term temporal dependencies through its gated cell state while flexibly responding to short-term inputs.

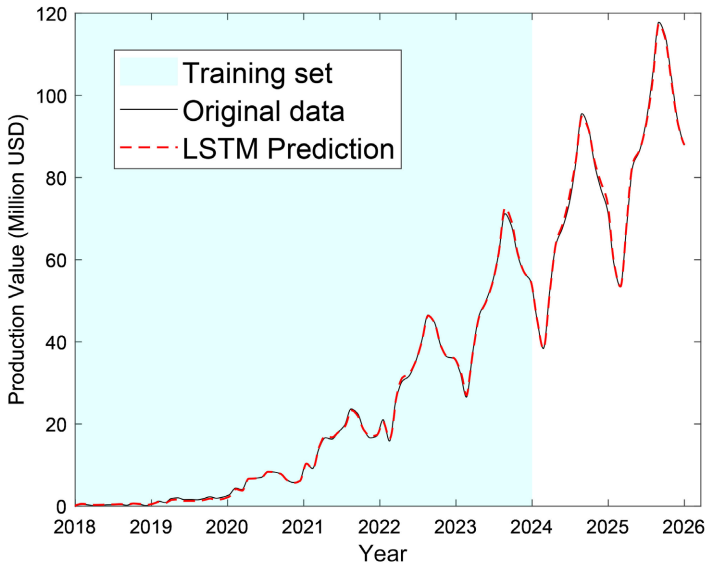


Figure 15. Full trend comparison with LSTM prediction.

Despite the accurate extrapolation observed in the test set, error metrics indicate a measurable increase in out-of-sample pointwise uncertainty, characterized by elevated *RMSE* and *MAE* alongside a slightly lower R^2_{test} relative to the training phase. This discrepancy between structural tracking and pointwise accuracy is driven by the potential for real-world regime shifts (e.g., accelerated scaling, policy changes, or demand shocks) during 2024-2026, and the reliance on synthetic high-frequency residuals during training that may not match the true stochastic structure of the future horizon. Furthermore, because the evaluation relies on a synthetic target constructed via the anchors, the out-of-sample metrics primarily assess the robustness of LSTM in extrapolating learned nonlinear trends

and seasonal distributions within this synthetic manifold, rather than its capacity to predict independent, unseen macroeconomic shocks.

From a First-Launch Economy policy perspective, the tight agreement between prediction and anchor-constrained series validates that the digital-intelligence pipeline (ETL → MLP monthly reconstruction → densification → LSTM forecasting) produces policy-coherent, temporally granular projections that preserve officially reported annual anchors while supplying actionable monthly/daily estimates.

Drawing upon public reports regarding the plush toy industry in Ankang City (Department of Human Resources and Social Security of Shaanxi Province, 2025; Hong & Liang, 2023; Xu, 2024; Yu, 2023; Zou, 2025), **Table 4** elucidates a sustained, monotonic expansion across four key structural indicators (enterprise quantity, annual production value, employment, and export volume) from 2018 to 2025. This growth trajectory aligns with the developmental stages of a policy-driven industrial cluster, characterizing the evolution from initial capacity formation (2018-2019) and agglomeration-based scale-up (2020-2022) to rapid output expansion during the maturation phase (2022-2025). These observed economic behaviors corroborate the First-Launch Economy of the plush toy industry in Ankang City, which is underpinned by industrial relocation and community factory deployment, thereby providing a robust empirical foundation that justifies the data reconstruction via MLP and the economic analysis with prediction using LSTM neural networks.

Table 4. Annual economic data of the Ankang plush toy industry (Department of Human Resources and Social Security of Shaanxi Province, 2025; Hong & Liang, 2023; Xu, 2024; Yu, 2023; Zou, 2025).

Year	Number of industries	Annual production value	Employment	Exports
		Million USD	Person	Million USD
2018	Initial stage	-	52	-
2021	569	13.92	12,000	-
2022	826	27.27	18,000	-
2023	788 - 807	80.51	19,000	9.61
2024	805 - 806	105.24	21,770	7.95
2025	800 - 811	139.77	-	-

A rigorous comparison between the publicly published annual production values in **Table 4** and the LSTM-predicted trajectories in **Figure 15** reveals a high degree of alignment, validating the efficacy of the pipeline's hard-anchoring mechanism. By preserving MLP-derived annual benchmarks during densification, this procedure effectively constrains the LSTM to replicate policy-relevant aggregates while simultaneously capturing granular intra-year dynamics. Consequently, the high coefficient of determination (R^2) and low in-sample errors reported in **Table**

3 signify that the model has successfully learned anchor-constrained trends and seasonality, rather than succumbing to arbitrary overfitting. Furthermore, the moderate increase in out-of-sample error metrics (*RMSE*/*MAE*) underscores the inherent limitations of forecasting via synthetic augmentation. The results indicate that even if the LSTM generalizes the structural trend robustly, it exhibits increased pointwise uncertainty when extrapolating into future horizons, potentially subject to regime shifts or unmodeled exogenous shocks. Upon the discussion above, the statistical concordance between the ETL aggregates and LSTM outputs confirms the framework's practical feasibility for generating temporally granular, anchor-consistent forecasts essential for First-Launch Economy evaluation.

5. Conclusion

By synthesizing export data across the National Bureau of Statistics, customs benchmarks, business registries, e-commerce records, and cross-validated against annual anchors from local government and media reports, this study employed an ETL (Extract-Transform-Load) pipeline to reconstruct coherent annual and monthly production series for Ankang's plush toy industry. The resulting aggregates demonstrate internal consistency across multiple dimensions, including enterprise counts, employment, export volumes, and total production value, thereby validating the capacity of ETL-based data fusion to recover policy-relevant economic signals in the absence of high-frequency official observations. Consequently, the findings in this study establish that a digital-intelligence workflow, grounded in official macro-statistics and enriched by multi-source data extraction, offers a robust and empirically grounded framework for analyzing emergent, policy-driven First-Launch Economy clusters in data-scarce environments.

Leveraging the ETL-derived time series, the LSTM-based forecasting framework developed in this study demonstrates a robust capacity to replicate anchor-constrained Monthly Production Value (MPV) trajectories and extend structural trends into short-term out-of-sample horizons. This predictive efficacy is fundamentally rooted in the model's recurrent gate architecture (comprising forget, input, and output gates), which enables the simultaneous preservation of long-term anchor signals and the learning of complex intra-annual seasonality and high-frequency residual patterns introduced via controlled data densification. Quantitatively, the model achieves high explanatory power (R^2) and minimal in-sample error, and it also yields defensible out-of-sample forecasts characterized by manageable increases in *RMSE*, *MAE*, and *MBE*, with minor biases that remain amenable to mitigation through ensemble or calibration techniques. Due to this, the results in this study substantiate the suitability of LSTM architectures for predictive modeling in contexts where periodic long-term anchors are available but dense, high-frequency longitudinal observations are sparse.

From an economic and methodological perspective, the integration of ETL-based reconstruction with LSTM forecasting significantly enhances the analytical capacity to evaluate the efficiency of First-Launch Economy clusters within key

assistance counties. By operationalizing qualitative policy narratives (such as industrial relocation, the expansion of community factories, and targeted developmental support) into quantitative, time-resolved indicators, this framework provides a robust empirical basis for optimizing supply-side upgrades, workforce allocation, and export facilitation strategies. More significantly, the maintenance of annual anchors ensures policy coherence, as the model-derived predictions remain strictly reconciled with officially reported aggregates, while simultaneously delivering the high-frequency temporal granularity essential for precise policy interventions, including pre-season capacity scaling and mid-year strategic investments. Upon the discussion above, this study demonstrates how AI-enabled predictive modeling can be harmonized with traditional economic measurement to generate decision-relevant intelligence for regional development and industrial policy optimization.

In conclusion, this study provides both empirical evidence and a transferable methodological template, demonstrating that the synergy of ETL-based data fusion and LSTM deep learning architectures can effectively evaluate and forecast the performance of First-Launch Economy clusters. Based on the Ankang plush toy industrial ecosystem, the integrated pipeline in this study yields high-resolution, anchor-consistent projections that not only validate trajectories of rapid industrial expansion but also offer actionable strategic insights for regional policy formulation. Macroscopically, the proposed framework in this study establishes a rigorous theoretical and practical foundation for AI-enhanced economic monitoring and decision-support systems, offering a scalable paradigm for analyzing emergent industrial clusters in data-constrained or nascent economic environments.

Acknowledgements

This work was supported by the 2025 science and technology project of Ankang Municipal Science and Technology Bureau, "AI-based Plush Toy Image Generation and Design Using Deep Learning Neural Networks", under Grant AK2025-GY-24.

Additionally, this work benefited significantly from the comprehensive resources provided by the Ankang Science & Technology Innovation Center (Qinchuangyuan Platform), which facilitated the execution and provided technical assistance for this scientific investigation.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Adamopoulos, I., Valamontes, A., Karantonis, J. T., Tropaitis, P. P., & Syrou, N. (2025). Enhancing Economic Growth Time Series for UAE Forecasting with Deep Learning: A Seq2Seq and Attention-Driven LSTM Approach. *EDRAAK, 2025,* 62-75.

https://doi.org/10.70470/edraak/2025/009

Ankang Plush Toy Industry Inspection Group (2018). *Investigation Report on the Plush Toy Industry*. Xinhua News Agency. https://www.ankang.gov.cn/Content-1537332.html

Cen, T., Lin, S., & Wu, Q. (2022). How Does Digital Economy Affect Rural Revitalization? The Mediating Effect of Industrial Upgrading. *Sustainability, 14,* Article No. 16987. https://doi.org/10.3390/su142416987

Chen, D., Wei, W., Hu, D., & Muralidharan, E. (2016). Survival Strategy of OEM Companies: A Case Study of the Chinese Toy Industry. *International Journal of Operations & Production Management, 36,* 1065-1088. https://doi.org/10.1108/ijopm-04-2015-0212

Deng, X., Huang, M., & Peng, R. (2024). The Impact of Digital Economy on Rural Revitalization: Evidence from Guangdong, China. *Heliyon, 10,* e28216. https://doi.org/10.1016/j.heliyon.2024.e28216

Department of Human Resources, & Social Security of Shaanxi Province (2025). *Ankang: Systematically Building a New Integrated Development Paradigm of "Plush Toys + AI + Cultural Tourism"*. https://rst.shaanxi.gov.cn/sy/ztzl/rdzt/rsgzjyq/jyjl/202512/t20251231_3601171.html

Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2018). *Data Augmentation Using Synthetic Data for Time Series Classification with Deep Residual Networks*.

Fritsch, F. N., & Carlson, R. E. (1980). Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis, 17,* 238-246. https://doi.org/10.1137/0717021

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *International Conference on Machine Learning* (pp. 1050-1059). PMLR.

Gereffi, G. (2019). Global Value Chains, Development, and Emerging Economies. In P. Lund-Thomsen, *et al.* (Eds.), *Business and Development Studies* (pp. 125-158). Routledge. https://doi.org/10.4324/9781315163338-6

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems, 28,* 2222-2232. https://doi.org/10.1109/tnnls.2016.2582924

Haider, S. A., Naqvi, S. R., Akram, T., Umar, G. A., Shahzad, A., Sial, M. R. et al. (2019). LSTM Neural Network Based Forecasting Model for Wheat Production in Pakistan. *Agronomy, 9,* Article No. 72. https://doi.org/10.3390/agronomy9020072

Hamiane, S., Ghanou, Y., Khalifi, H., & Telmem, M. (2024). Comparative Analysis of LSTM, ARIMA, and Hybrid Models for Forecasting Future GDP. *Ingénierie des Systèmes d Information, 29,* 853-861. https://doi.org/10.18280/isi.290306

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9,* 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hong, Y., & Liang J. (2023). Plush Toys Produced in Small City in NW China's Shaanxi Reach over 80 Countries, Regions. *People's Daily*. https://en.people.cn/n3/2023/0921/c98649-20074698.html

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts.

Iwana, B. K., & Uchida, S. (2021). An Empirical Survey of Data Augmentation for Time Series Classification with Neural Networks. *PLOS ONE, 16,* e0254841. https://doi.org/10.1371/journal.pone.0254841

Kaplinsky, R., & Morris, M. (2016). Thinning and Thickening: Productive Sector Policies in the Era of Global Value Chains. *The European Journal of Development Research, 28,*

625-645. https://doi.org/10.1057/ejdr.2015.29

Kee, H. L., & Tang, H. (2016). Domestic Value Added in Exports: Theory and Firm Evidence from China. *American Economic Review, 106,* 1402-1436. https://doi.org/10.1257/aer.20131687

Liu, Z., Wu, H., & Badar, I. H. (2024a). Impact of Rural Digital Economy Development on Rural Revitalisation—Evidence from China. *Sustainability, 16,* Article No. 4573. https://doi.org/10.3390/su16114573

Liu, F., Yu, Z. C., & Deng, M. L. (2024b). An Empirical Study on the Impact of Digital Economy Development on the Quality of Rural Revitalization. *Science & Technology Progress and Policy, 41,* 47-57.

Liu, Y. (2019). Event and Sustainable Culture-Led Regeneration: Lessons from the 2008 European Capital of Culture, Liverpool. *Sustainability, 11,* Article No. 1869. https://doi.org/10.3390/su11071869

Lu, Y., Meng, Y., Chen, L., & Zhang, Y. (2024). Can the Digital Economy Contribute to Rural Revitalization? A Case of from China? *PLOS ONE, 19,* e0310313. https://doi.org/10.1371/journal.pone.0310313

Luo, G., Yang, Y., & Wang, L. (2023). Driving Rural Industry Revitalization in the Digital Economy Era: Exploring Strategies and Pathways in China. *PLOS ONE, 18,* e0292241. https://doi.org/10.1371/journal.pone.0292241

Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the Difficulty of Training Recurrent Neural Networks. In *International Conference on Machine Learning* (pp. 1310-1318). PMLR.

Paszkowski, S., & Sarniak, Ł. (2019). Differences in Rural Development Potential across Polish Regions. *Journal of Agribusiness and Rural Development, 51,* 59-68. https://doi.org/10.17306/j.jard.2019.01102

Pyo, D. J. (2025). Enhancing GDP Growth Forecasting with LSTM, GRU, and Hybrid Model: Evidence from South Korea. *Sage Open, 15,* No. 3. https://doi.org/10.1177/21582440251359828

Shi, M., Yuan, H. T., & Chen, X. A. (2026). Digital-Intelligence Empowerment for Rural Revitalization: Estimating First-Launch Economy Productivity Efficiency via ETL-MLP Time-Series Reconstruction. *Open Journal of Social Sciences, 14,* 340-367. https://doi.org/10.4236/jss.2026.143020

Siami-Namini, S., & Namin, A. S. (2018). Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. https://arxiv.org/abs/1803.06386

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research, 15,* 1929-1958.

Tian, Y., Liu, Q., Ye, Y., Zhang, Z., & Khanal, R. (2023). How the Rural Digital Economy Drives Rural Industrial Revitalization—Case Study of China's 30 Provinces. *Sustainability, 15,* Article No. 6923. https://doi.org/10.3390/su15086923

Wang, S., Peng, T., Du, A. M., & Lin, X. (2025). The Impact of the Digital Economy on Rural Industrial Revitalization. *Research in International Business and Finance, 76,* Article ID: 102878. https://doi.org/10.1016/j.ribaf.2025.102878

Wang, Y., Peng, Q., Jin, C., Ren, J., Fu, Y., & Yue, X. (2023). Whether the Digital Economy Will Successfully Encourage the Integration of Urban and Rural Development: A Case Study in China. *Chinese Journal of Population, Resources and Environment, 21,* 13-25. https://doi.org/10.1016/j.cjpre.2023.03.002

Xu, Q. (2024). Toys Help Ankang Residents to Lead a Plush New Life. *China Daily*. https://www.chinadailyhk.com/hk/article/232819?utm

Yu, H. C. (2023). *Economic Watch: Made-in-China Stuffed Bears Spread Joy of Christmas Overseas*. Xinhua News Agency. https://eng.yidaiyilu.gov.cn/p/0M6INNO1.html

Zhang, T. (2024). How Small Toys Form Large Industrial Clusters (Onsite Commentary)—Acting as Steadfast Practitioners and Doers in Advancing Chinese-Style Modernization. *People's Daily*. https://cpc.people.com.cn/BIG5/n1/2024/0509/c64387-40231976.html

Zhang, Z., Sun, C., & Wang, J. (2023). How Can the Digital Economy Promote the Integration of Rural Industries—Taking China as an Example. *Agriculture, 13,* Article No. 2023. https://doi.org/10.3390/agriculture13102023

Zheng, C., & Li, H. (2023). The Prediction of Collective Economic Development Based on the PSO-LSTM Model in Smart Agriculture. *PeerJ Computer Science, 9,* e1304. https://doi.org/10.7717/peerj-cs.1304

Zou, L. (2025). Proposal on Further Advancing the Upgrading and Development of the Plush Toy Industry. *Ankang Municipal People's Congress*. https://rd.ankang.gov.cn/Content-2874919.html