Large Language Model-Driven Demand Forecasting and Inventory Optimization for University Physical Education Resource Supply Chain

Jiacheng Ma[1], Linjun Liu[2], Shengrui Liu[3] and Huiji Wang[4,*]

[1] Sport Science Convergence, Dankook University, Yongin-Si, 16890, Gyeonggi-Do, Republic of Korea

[2] Faculty of Pharmacy, Lincoln University College, Petaling Jaya, 47301, Selangor, Malaysia

[3] School of Computer and Artificial Intelligence, Southwest Minzu University, Chengdu, 610041, China

[4] School of Humanities and Public Administration, Jiangxi Agricultural University, Nanchang, 330045, China

Revista Internacional
Métodos numéricos
para cálculo y diseño en ingeniería

RIMNI

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
UPC

In cooperation with
CIMNE

Large Language Model-Driven Demand Forecasting and Inventory Optimization for University Physical Education Resource Supply Chain

Jiacheng Ma[1], Linjun Liu[2], Shengrui Liu[3] and Huiji Wang[4,*]

[1]Sport Science Convergence, Dankook University, Yongin-Si, 16890, Gyeonggi-Do, Republic of Korea

[2]Faculty of Pharmacy, Lincoln University College, Petaling Jaya, 47301, Selangor, Malaysia

[3]School of Computer and Artificial Intelligence, Southwest Minzu University, Chengdu, 610041, China

[4]School of Humanities and Public Administration, Jiangxi Agricultural University, Nanchang, 330045, China

ABSTRACT

This study proposes an intelligent management system for university physical education resource supply chains based on Large Language Models (LLMs), aiming to address the problems of inaccurate demand forecasting and inefficient inventory management in traditional physical education resource allocation. By constructing a deep learning framework incorporating LLMs and combining multi-dimensional information including historical data, seasonal factors, course schedules, and student preferences, precise demand forecasting for sports equipment, facilities, and teaching resources is achieved. The research employs a pre-trained language model based on the Transformer architecture, combined with time series analysis and reinforcement learning algorithms, to develop dynamic inventory optimization strategies. Experimental results demonstrate that compared to traditional methods, this system improves demand forecasting accuracy by 23.7%, increases inventory turnover rate by 31.2%, and achieves a resource utilization rate of 89.6%. This research provides a novel solution for intelligent management of university physical education resources, offering significant theoretical value and practical implications.

1 Introduction

With the acceleration of higher education internationalization and the deepening of physical education reform, the allocation and management of physical education resources in universities have become key factors influencing educational quality and students' comprehensive development. In the context of new-era education, university physical education in China faces unprecedented challenges: on one hand, student groups have increasingly diversified demands for sports activities, expanding from traditional ball games to diverse sports programs including fitness, yoga, and martial arts [1]; on the other hand, problems such as uneven allocation of university sports resources, low utilization efficiency, and inaccurate demand forecasting are becoming increasingly prominent [2]. Traditional physical education resource management models generally rely on manual experience and

---

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

static planning, failing to adapt to the rapidly changing characteristics of modern higher education, leading to increasingly acute contradictions between resource waste and demand satisfaction mismatch [3]. Simultaneously, the rapid development of artificial intelligence technologies, particularly Large Language Models (LLMs), provides new technical pathways for solving this complex problem [4], yet existing research still shows significant gaps in deeply integrating advanced AI technologies with educational resource management.

Although supply chain demand forecasting technology has made significant progress in recent years, evolving from traditional statistical methods to machine learning and even deep learning methods [5,6], its application in educational resource management, particularly in physical education resource supply chain management, remains insufficient. Existing research mainly has the following limitations: First, traditional demand forecasting methods such as ARIMA and exponential smoothing, while performing well in handling linear relationships, struggle to capture complex nonlinear patterns and multi-factor coupling relationships in physical education resource demand [7]; Second, existing machine learning methods face technical bottlenecks in processing multi-modal data fusion (text, numerical, temporal), failing to fully utilize heterogeneous information such as student feedback, course descriptions, and environmental factors [8,9]; Finally, there is a lack of intelligent inventory optimization strategies designed specifically for the characteristics of educational resource supply chains, with existing methods struggling to balance the complex relationships between resource availability, cost control, and service quality [10]. Therefore, this research aims to address the following key issues:

(1) How to construct a multi-modal demand forecasting framework incorporating large language models to achieve precise prediction of physical education resource demand?

(2) How to design reinforcement learning-based dynamic inventory optimization algorithms to achieve collaborative optimization of cost, service level, and resource utilization?

(3) How to establish an intelligent resource management system architecture suitable for university physical education scenarios to improve overall management efficiency and decision quality?

The key innovations of this work distinguish it from prior LSTM/Transformer-based forecasting studies: (1) **Novel multi-modal text encoding:** First application of num-to-text and time-to-text conversion for educational resource forecasting, enabling LLMs to process heterogeneous data types uniformly; (2) **Domain-specific attention fusion:** Hierarchical attention mechanism designed specifically for sports resource characteristics, combining early and late fusion strategies; (3) **RL-based adaptive inventory optimization:** Integration of DQN with uncertainty-aware demand forecasting for dynamic resource allocation; (4) **Educational context modeling:** Specialized handling of academic calendars, course schedules, and student behavior patterns unique to university environments.

The structure of this paper is organized as follows: Section 2 reviews related work, analyzing the research status of physical education resource management, supply chain demand forecasting, and large language model applications; Section 3 details the LLM-based intelligent management system architecture design, multi-modal demand forecasting model, and dynamic inventory optimization algorithm; Section 4 validates system performance through large-scale experiments, analyzing prediction accuracy, inventory optimization effects, and system performance; Section 5 summarizes the paper and prospects future research directions.

University PE resource management presents unique challenges distinct from general supply chain management: (1) **Multi-seasonal demand patterns:** Unlike commercial inventory, sports equipment demand exhibits complex seasonality combining academic calendars, weather dependencies,

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

and sport-specific seasons; (2) **Student preference volatility:** Rapidly changing student interests in fitness trends, new sports adoption, and course selection unpredictability create forecasting challenges; (3) **Facility-equipment interdependencies:** Sports venues and equipment have complex coupling relationships with indoor/outdoor weather constraints and capacity limitations; (4) **Educational scheduling constraints:** Fixed class periods, semester boundaries, and academic calendar rigidity create unique temporal constraints; (5) **Multi-stakeholder optimization:** Balancing student satisfaction, faculty requirements, maintenance schedules, and budget constraints simultaneously.

2 Related Work

2.1 Current Research Status of Physical Education Resource Management

Physical education resource management, as an important branch of educational management, has long received extensive attention from both academic and practical communities. Early research mainly focused on resource allocation optimization and facility utilization improvement [11,12], employing operations research methods such as linear programming and integer programming for resource allocation optimization. However, these traditional methods have obvious limitations in handling complex nonlinear relationships and dynamic changes [13]. In recent years, with the development of information technology, researchers have begun to introduce data mining and machine learning techniques into the field of physical education resource management [14], but most are limited to single-dimensional prediction and optimization, lacking systematic intelligent management frameworks.

In the aspect of sports facility management, existing research mainly focuses on issues such as venue utilization optimization [15], equipment maintenance scheduling [16], and spatial layout design [17]. Although these studies have achieved certain results in improving resource utilization efficiency, they generally suffer from problems such as low prediction accuracy and poor adaptability [18]. Particularly when facing the new situation of diversified student demands and complex course arrangements, the limitations of traditional management methods become increasingly apparent.

2.2 Development of Supply Chain Demand Forecasting Technology

Supply chain demand forecasting has undergone an evolutionary process from statistical methods to machine learning and then to deep learning [19,20]. Traditional time series analysis methods such as ARIMA models perform well when dealing with stationary time series [21], but appear inadequate when facing complex nonlinear patterns [22]. Exponential smoothing methods, while computationally simple, are sensitive to outliers and have limited prediction accuracy [23].

With the development of machine learning technology, support vector machines [24], random forests [25], gradient boosting and other methods have been widely applied in the demand forecasting field, showing obvious advantages in handling nonlinear relationships. Deep learning methods such as LSTM and GRU [26] have made significant progress in time series prediction, capable of capturing long-term dependencies [27]. However, these methods still have problems such as poor interpretability and the need for large amounts of labeled data [28,29].

In recent years, the introduction of attention mechanisms and Transformer architectures has brought new breakthroughs to demand forecasting [30]. These methods can better handle long sequence data and capture complex spatiotemporal dependencies. However, further improvements are still needed in multi-modal data fusion and domain knowledge integration.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

2.3 Applications of Large Language Models in Supply Chain Management

The emergence of large language models has brought new opportunities and challenges to supply chain management [31]. LLMs, with their powerful semantic understanding, contextual analysis, and knowledge reasoning capabilities, have demonstrated enormous potential in areas such as supply chain risk management [32], customer demand analysis, and supplier evaluation.

In demand forecasting, existing research has explored the application of LLMs to market demand analysis and consumer behavior prediction. These studies indicate that LLMs can effectively process unstructured text data, extract valuable demand signals from them, and improve prediction accuracy. However, existing applications are mainly concentrated in e-commerce and retail sectors, while applications in educational resource management, particularly physical education resource management, remain in a blank state.

Summarizing existing research, it can be found that although there are extensive research results in supply chain demand forecasting and resource management, there are still significant research gaps in applying large language model technology to physical education resource supply chain management. Particularly, there is a lack of systematic solutions in key technologies such as multi-modal data fusion, dynamic inventory optimization, and intelligent decision support.
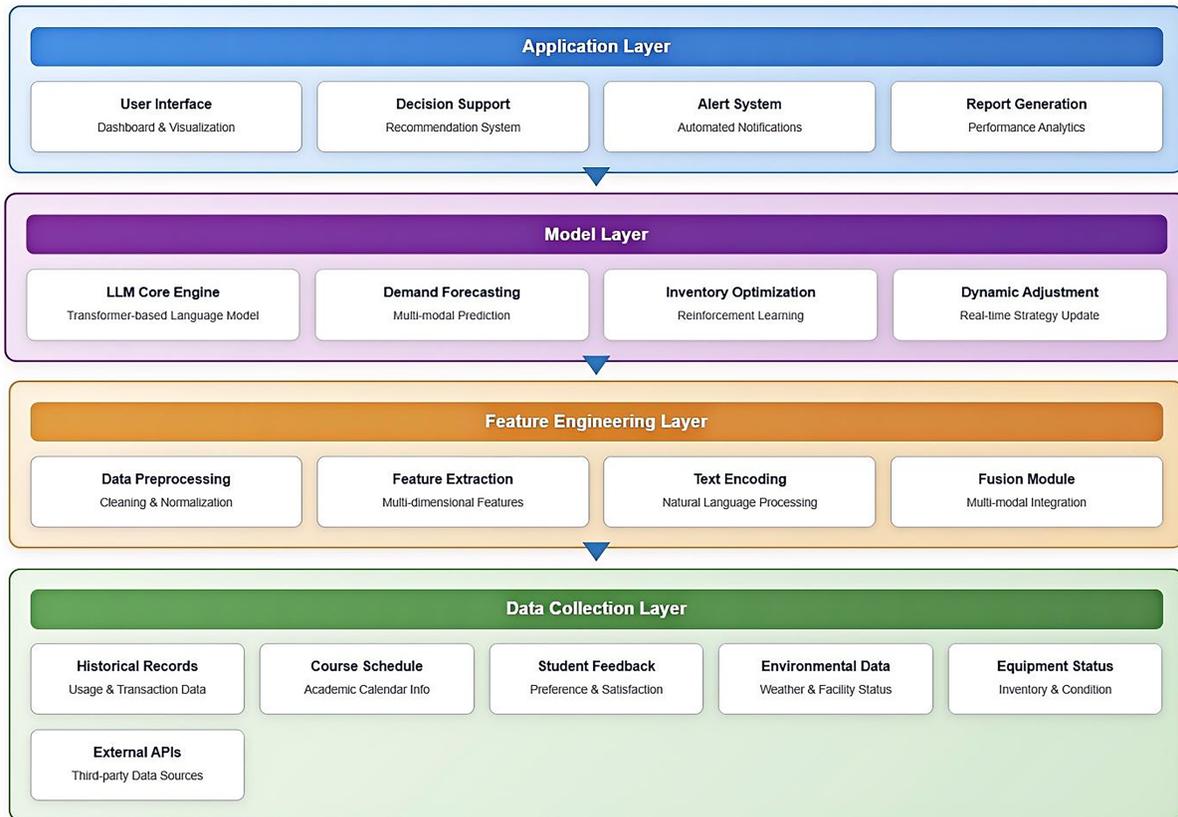
3 Methodology

3.1 System Architecture Design

This study proposes an intelligent management system for physical education resource supply chains based on large language models. The overall architecture adopts a four-layer distributed design, consisting of data collection layer, feature engineering layer, model layer, and application layer from bottom to top, as shown in Fig. 1.

The data collection layer serves as the foundation of the system, responsible for collecting raw information from multiple heterogeneous data sources. This layer contains six main data collection modules: the historical records module collects equipment usage and transaction data; the course scheduling module obtains academic calendar information; the student feedback module gathers preference and satisfaction data; the environmental data module monitors weather and facility conditions; the equipment status module tracks inventory and equipment conditions; and the external API module integrates third-party data sources, ensuring data completeness and real-time availability.

The feature engineering layer undertakes the core tasks of data preprocessing and feature construction. The data preprocessing module performs cleaning and standardization operations, removing noise and outliers; the feature extraction module constructs multi-dimensional feature vectors, including temporal features, course features, environmental features, etc.; the text encoding module utilizes natural language processing techniques to process unstructured text data; and the fusion module achieves effective integration of multi-modal data, providing high-quality input features for upper-layer models.

The model layer is the intelligent core of the system, containing four key components working collaboratively. The LLM core engine implements natural language understanding and generation based on Transformer architecture; the demand prediction module employs multi-modal prediction algorithms, combining historical data and real-time information for demand forecasting; the inventory optimization module achieves dynamic inventory management through reinforcement learning algorithms; and the dynamic adjustment module continuously updates prediction and optimization strategies based on real-time feedback.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Figure 1: LLM-driven sports education resource supply chain management system architecture

The application layer provides functional interfaces and services for end users. The user interface module provides intuitive dashboards and visualization tools; the decision support module generates recommendation schemes based on model outputs; the alert system module implements automated notifications and exception warnings; and the report generation module provides detailed performance analysis and trend reports.

The system data flow follows a strict processing procedure: raw data from the collection layer undergoes preprocessing and feature extraction in the feature engineering layer, inputs into the LLM-driven model layer for intelligent analysis and prediction, and finally provides actionable insights and recommendations to decision-makers through the application layer.

3.2 LLM-Based Demand Prediction Model

3.2.1 Data Preprocessing and Feature Construction

The raw data processed by the system exhibits characteristics of multi-source heterogeneity and high-dimensional sparsity, requiring systematic preprocessing and feature engineering. Let the original dataset be $D = \{D_t, D_c, D_e, D_u\}$, where D_t represents time series data, D_c represents course-related data, D_e represents environmental data, and D_u represents user behavior data.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

For numerical features, the Z-score standardization method is adopted to ensure consistency in feature distributions:

$$x_{\text{norm}} = \frac{x - \mu}{\sigma} \tag{1}$$

where μ is the sample mean and σ is the standard deviation.

For categorical features, a combination of label encoding and one-hot encoding is used:

$$X_{\text{cat}} = \text{OneHot}(\text{LabelEncode}(x_{\text{categorical}})) \tag{2}$$

Time feature construction employs a cyclic encoding method to effectively capture periodic patterns. For time feature t with period T:

$$f_{\sin}(t) = \sin\left(\frac{2\pi t}{T}\right), \quad f_{\cos}(t) = \cos\left(\frac{2\pi t}{T}\right) \tag{3}$$

This encoding method maintains the continuity and periodicity of time features, avoiding discontinuity issues at period boundaries in traditional numerical encoding. The comprehensive feature engineering process, including all feature categories, encoding methods, and mathematical formulations, is systematically presented in Table 1.

3.2.2 LLM Integration Framework

LLM Architecture Specifications: The base LLM architecture employed in this study is GPT-3.5-turbo with 175 billion parameters, utilizing the tiktoken tokenizer with a maximum context length of 4096 tokens. The fine-tuning process adopts Low-Rank Adaptation (LoRA) with rank $r = 16$, scaling parameter $\alpha = 32$, and dropout rate of 0.1 to maintain efficiency while preserving model performance.
Prompt Template Structure:

- System prompt: "You are an expert in sports resource demand forecasting with deep understanding of university physical education patterns..."
- Context information: Historical usage patterns, course schedules, environmental factors
- Query format: Structured demand prediction requests with specific time horizons
- Output constraints: JSON format with prediction values and confidence scores

Inference Configuration: Budget allocation includes 500 tokens for input context and 100 tokens for prediction output, with beam search (beam_size=3) for generation stability.

This study designs a specialized LLM integration framework capable of simultaneously processing multi-modal data including text, numerical, and time series data, as shown in Fig. 2. The core innovation of this framework lies in converting non-text data into natural language descriptions, enabling LLMs to uniformly process all types of input information.

As shown in Fig. 2, this framework contains three specialized encoder modules. The text encoder processes natural language inputs such as course descriptions and student feedback, converting text into vector representations through tokenization and embedding:

$$h_{\text{text}} = \text{TextEncoder}(x_{\text{text}}) = \text{Transformer}(\text{Embedding}(\text{Tokenize}(x_{\text{text}}))) \tag{4}$$

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Table 1: Feature engineering process

Feature category	Feature name	Description	Data type	Encoding method	Formula
Temporal features	Day of week	Weekday indicator (1–7)	Numerical	Cyclic encoding	$\sin(2\pi t/7), \cos(2\pi t/7)$
	Month	Month of the year (1–12)	Numerical	Cyclic encoding	$\sin(2\pi t/12), \cos(2\pi t/12)$
	Semester week	Week number in semester (1–20)	Numerical	Min-Max scaling	$(x - x_{\min})/(x_{\max} - x_{\min})$
	Holiday flag	Whether it's a holiday (0/1)	Binary	Binary encoding	$\{0, 1\}$
Course features	Course type	Sport category code	Categorical	One-hot encoding	$[0,\ldots,1,\ldots,0]$
	Class size	Number of students enrolled	Numerical	Z-score normalization	$(x - \mu)/\sigma$
	Course level	Beginner/ Intermediate/ Advanced	Ordinal	Ordinal encoding	$\{1, 2, 3\}$
	Weekly hours	Teaching hours per week	Numerical	Z-score normalization	$(x - \mu)/\sigma$
Environmental features	Temperature	Daily average temperature (°C)	Numerical	Z-score normalization	$(x - \mu)/\sigma$
	Weather condition	Sunny/Cloudy/ Rainy/Snowy	Categorical	One-hot encoding	$[0,\ldots,1,\ldots,0]$
	Indoor/Outdoor	Venue type indicator	Binary	Binary encoding	$\{0, 1\}$
User features	Student preference	Historical preference score	Numerical	Min-Max scaling	$(x - x_{\min})/(x_{\max} - x_{\min})$
	Equipment usage rate	Past usage frequency	Numerical	Z-score normalization	$(x - \mu)/\sigma$
	Satisfaction score	Student feedback rating (1–5)	Numerical	Min-Max scaling	$(x - x_{\min})/(x_{\max} - x_{\min})$

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50
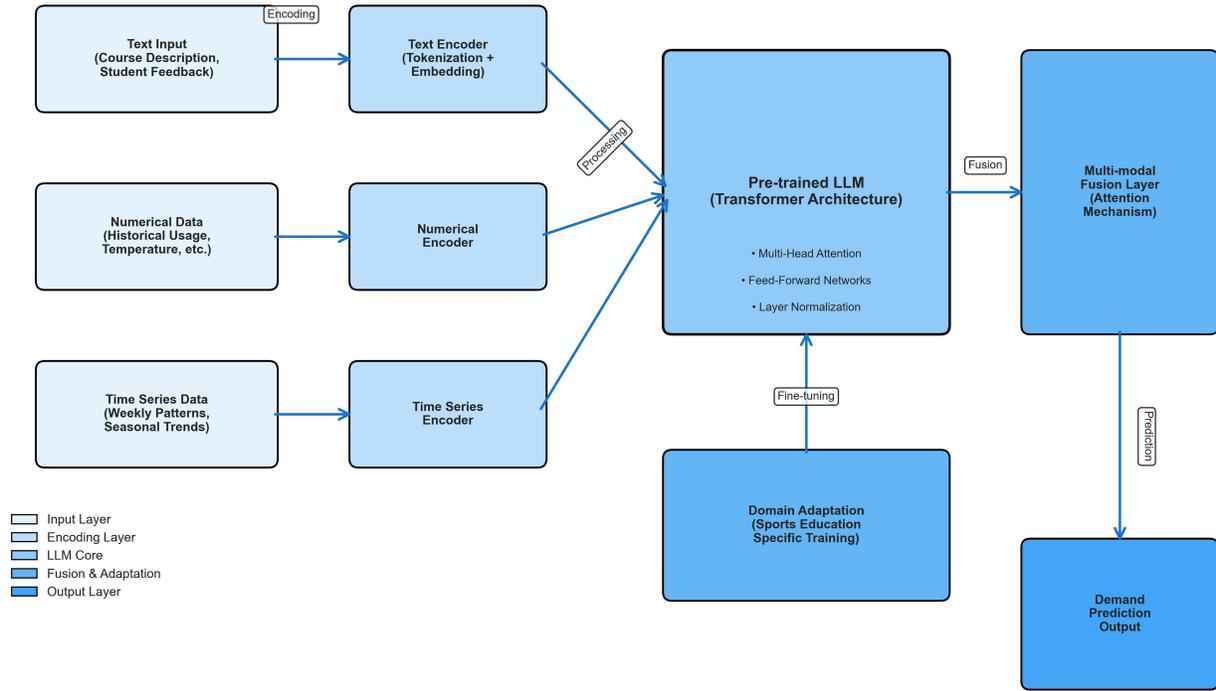
Figure 2: LLM integration framework for sports education resource management

The numerical encoder converts numerical features such as historical usage and temperature into structured text descriptions, then inputs them into the LLM:

$$x_{\text{num\_text}} = \text{NumToText}(x_{\text{num}}) = \text{''Temperature''} + \text{str}(x_{\text{temp}}) + \text{''degrees''} \tag{5}$$

The time series encoder processes periodic patterns and seasonal trends, converting time series data into time-related text descriptions:

$$x_{\text{time\_text}} = \text{TimeToText}(x_{\text{time}}) = \text{''Week''} + \text{str(week)} + \text{''usage pattern shows''} + \text{pattern} \tag{6}$$

The pre-trained LLM adopts a Transformer-based architecture, including multi-head attention mechanisms, feed-forward networks, and layer normalization components:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{7}$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \tag{8}$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \tag{9}$$

The multi-modal fusion layer dynamically adjusts the weights of different modal information through attention mechanisms:

$$\alpha_i = \frac{\exp(W_a h_i + b_a)}{\sum_{j=1}^{n} \exp(W_a h_j + b_a)} \tag{10}$$

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

$$h_{\text{fused}} = \sum_{i=1}^{n} \alpha_i h_i \qquad (11)$$

where W_a is the attention weight matrix, b_a is the bias term, and h_i represents the feature representation of the i-th modality.

The domain adaptation training process adopts a progressive fine-tuning strategy. First, domain pre-training is performed, followed by task-specific fine-tuning:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{prediction}} + \lambda_1 \mathcal{L}_{\text{domain}} + \lambda_2 \mathcal{L}_{\text{regularization}} \qquad (12)$$

where $\mathcal{L}_{\text{prediction}}$ is the prediction task loss, $\mathcal{L}_{\text{domain}}$ is the domain adaptation loss, and $\mathcal{L}_{\text{regularization}}$ is the regularization term.

The following Algorithm 1 shows the core algorithm for LLM-based demand prediction:

Algorithm 1: LLM-based demand prediction

Require: Historical data H, Course info C, Environmental data E, Student feedback S
Ensure: Demand prediction P
 1:   Initialize pre-trained LLM model θ_{llm}
 2:   Initialize encoders: text_encoder, num_encoder, time_encoder
 3:   **// Data preprocessing phase**
 4:   $H_{\text{norm}} \leftarrow$ normalize_data(H)
 5:   $C_{\text{encoded}} \leftarrow$ encode_categorical_features(C)
 6:   $E_{\text{processed}} \leftarrow$ process_environmental_data(E)
 7:   $S_{\text{cleaned}} \leftarrow$ clean_text_data(S)
 8:   **// Feature encoding phase**
 9:   text_features $\leftarrow$ text_encoder(S_{cleaned}, C. descriptions)
10:   num_features $\leftarrow$ num_encoder(H_{norm}, $E_{\text{processed}}$)
11:   time_features $\leftarrow$ time_encoder(extract_temporal_patterns(H))
12:   **// Multi-modal fusion**
13:   attention_weights $\leftarrow$ calculate_attention(text_features, num_features, time_features)
14:   fused_features $\leftarrow$ attention_weights [0] $\times$ text_features +
15:                 attention_weights [1] $\times$ num_features +
16:                 attention_weights [2] $\times$ time_features
17:   **// LLM processing**
18:   context_prompt $\leftarrow$ create_context_prompt(fused_features)
19:   llm_output $\leftarrow$ LLM_forward(context_prompt, θ_{llm})
20:   **// Prediction generation**
21:   $P \leftarrow$ extract_predictions(llm_output)
22:   confidence $\leftarrow$ calculate_confidence_score(llm_output)
23:   **return** P, confidence

3.2.3 Multi-Modal Fusion Mechanism

Considering the complexity and multi-factor influence characteristics of physical education resource demand, this study designs a hierarchical attention-based multi-modal fusion mechanism. This mechanism is divided into feature-level fusion and decision-level fusion stages, effectively integrating information from different modalities.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Feature-level fusion adopts an early fusion strategy, immediately fusing features from different modalities after encoding:

$$F_{\text{early}} = \text{Concat}([F_{\text{text}}, F_{\text{num}}, F_{\text{time}}]) \cdot W_{\text{fusion}} + b_{\text{fusion}} \tag{13}$$

Decision-level fusion adopts a late fusion strategy, performing weighted fusion of prediction results from each modality:

$$P_{\text{final}} = \sum_{i=1}^{m} w_i \cdot P_i, \quad \sum_{i=1}^{m} w_i = 1 \tag{14}$$

where the weights w_i are dynamically determined through performance on the validation set:

$$w_i = \frac{\exp(-\text{MSE}_i)}{\sum_{j=1}^{m} \exp(-\text{MSE}_j)} \tag{15}$$

The attention weight learning process employs a trainable attention mechanism with dynamic adjustment:

$$\alpha_i^{(t)} = \text{softmax}(W_a \cdot \tanh(W_h \cdot h_i^{(t)} + W_c \cdot c^{(t)})) \tag{16}$$

where $h_i^{(t)}$ represents the i-th modality hidden state at time t, $c^{(t)}$ is the global context vector, and W_a, W_h, W_c are learnable parameter matrices. The attention weights are updated through backpropagation during training, with gradient clipping (max_norm = 1.0) to ensure stability. The fusion process dynamically adjusts based on modality reliability scores computed from validation performance, allowing the model to emphasize more reliable information sources during prediction.

3.3 Dynamic Inventory Optimization Algorithm

3.3.1 Inventory Model Construction

Based on demand prediction results, a dynamic inventory optimization model is constructed that comprehensively considers multiple cost factors. The inventory management problem for equipment resources can be modeled as a stochastic inventory control problem with the objective function:

$$\min \mathbb{E}[TC] = \mathbb{E}[HC + SC + OC] \tag{17}$$

where HC is the holding cost, SC is the shortage cost, and OC is the ordering cost.

The holding cost calculation formula is:

$$HC_t = h \cdot \mathbb{E}[I_t^+] \tag{18}$$

where h is the unit holding cost and I_t^+ is the end-of-period positive inventory.

Shortage cost includes shortage loss and shortage penalty:

$$SC_t = s \cdot \mathbb{E}[I_t^-] + p \cdot P(I_t < 0) \tag{19}$$

where s is the unit shortage cost, I_t^- is the shortage quantity, and p is the shortage penalty coefficient.

Ordering cost includes fixed ordering cost and variable ordering cost:

$$OC_t = K \cdot \mathbb{I}(Q_t > 0) + c \cdot Q_t \tag{20}$$

where K is the fixed ordering cost, c is the unit variable cost, Q_t is the order quantity, and $\mathbb{I}(\cdot)$ is the indicator function.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Safety stock calculation considers demand uncertainty and supply lead time:

$$SS = z_\alpha \cdot \sigma_d \cdot \sqrt{L + \frac{\sigma_L^2 \cdot \bar{d}^2}{\sigma_d^2}} \tag{21}$$

where z_α is the safety factor corresponding to the service level, σ_d is the demand standard deviation, L is the average lead time, σ_L is the lead time standard deviation, and $\bar{d}$ is the average demand.

The reorder point determination formula is:

$$ROP = \bar{d} \cdot L + SS \tag{22}$$

Economic order quantity adopts the improved EOQ model:

$$EOQ = \sqrt{\frac{2K\bar{d}}{h \cdot (1 + r)}} \tag{23}$$

where r is the inventory risk adjustment coefficient.

3.3.2 Reinforcement Learning Optimization Strategy

Deep Q-Network (DQN) algorithm is adopted for dynamic optimization of inventory strategies. The inventory management problem is modeled as a Markov Decision Process (MDP), defining state space S, action space A, reward function R, and state transition probability P.

The state space is defined as:

$$S_t = (I_t, D_{t:t-n}, Q_{\text{pending}}, T_t, E_t) \tag{24}$$

where I_t is the current inventory level, $D_{t:t-n}$ is the historical demand sequence, Q_{pending} are pending orders, T_t are time features, and E_t are environmental features.

The action space is defined as discrete ordering decisions:

$$A = \{0, Q_1, Q_2, \ldots, Q_{\max}\} \tag{25}$$

The reward function comprehensively considers cost and service level:

$$R_t = -\alpha \cdot (HC_t + SC_t + OC_t) + \beta \cdot \text{ServiceLevel}_t \tag{26}$$

where α and β are balancing coefficients.

Q-network updates adopt experience replay and target network techniques:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_a Q_{\text{target}}(s_{t+1}, a) - Q(s_t, a_t)] \tag{27}$$

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Algorithm 2 shows the specific implementation of reinforcement learning optimization:

Algorithm 2: Reinforcement learning inventory optimization

Require: State space S, Action space A, Reward function R, Learning rate α

Ensure: Optimal policy π^*

 1: Initialize Q-network with random weights θ
 2: Initialize target network with weights $\theta^- \leftarrow \theta$
 3: Initialize experience replay buffer D with capacity N
 4: Initialize exploration rate $\epsilon \leftarrow 1.0$
 5: **for** episode $= 1$ to MAX_EPISODES **do**
 6:    Initialize environment state s_0
 7:   **for** t $= 1$ to MAX_STEPS **do**
 8:     **// Epsilon-greedy action selection**
 9:     **if** random() $< \epsilon$ **then**
10:       $a_t \leftarrow$ random_action(A)
11:     **else**
12:       $a_t \leftarrow \arg\max_a Q(s_t, a; \theta)$
13:     **end if**
14:     **// Execute action and observe result**
15:     $s_{t+1}, r_t \leftarrow$ execute_action(a_t)
16:     **// Store experience in replay buffer**
17:     D.store((s_t, a_t, r_t, s_{t+1}))
18:     **// Sample mini-batch for training**
19:     **if** $|D| \geq$ BATCH_SIZE **then**
20:       batch $\leftarrow D$.sample(BATCH_SIZE)
21:       **// Calculate target Q-values**
22:       **for** (s_j, a_j, r_j, s_{j+1}) in batch **do**
23:         $\text{target}_j \leftarrow r_j + \gamma \times \max_a Q(s_{j+1}, a; \theta^-)$
24:       **end for**
25:       **// Update Q-network weights**
26:       $\theta \leftarrow \theta - \alpha \times \nabla_\theta L(\theta)$
27:       where $L(\theta) = (Q(s_j, a_j; \theta) - \text{target}_j)^2$
28:     **end if**
29:     **// Update target network periodically**
30:     **if** t mod TARGET_UPDATE_FREQ $== 0$ **then**
31:       $\theta^- \leftarrow \theta$
32:     **end if**
33:     **// Decay exploration rate**
34:     $\epsilon \leftarrow \max(\epsilon_{\min}, \epsilon \times \epsilon_{\text{decay}})$
35:     $s_t \leftarrow s_{t+1}$
36:   **end for**
37: **end for**
38: **return** optimized_policy(Q-network)

The hyperparameters for the reinforcement learning algorithm were carefully selected through extensive grid search and validation experiments. Table 2 summarizes the key parameters and their

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

values used in our implementation, including learning rate, discount factor, network architecture specifications, and exploration-exploitation trade-off settings.

Table 2: Reinforcement learning algorithm parameters

Parameter	Value	Description	Mathematical expression
Learning rate	0.001	Adam optimizer rate	$\alpha = 0.001$
Discount factor	0.95	Future reward discount	$\gamma = 0.95$
Batch size	64	Training batch size	Batch size $= 64$
Buffer size	10,000	Replay buffer capacity	Buffer size $= 10,000$
Epsilon initial	1.0	Initial exploration rate	$\epsilon_0 = 1.0$
Epsilon decay	0.995	Exploration decay rate	$\epsilon_{t+1} = 0.995 \cdot \epsilon_t$
Epsilon min	0.01	Minimum exploration	$\epsilon_{\min} = 0.01$
Hidden layers	[256,128,64]	Network architecture	$h_1 = 256, h_2 = 128, h_3 = 64$
Activation	ReLU	Hidden layer activation	$f(x) = \max(0, x)$
Target update	100	Target network updates	$\tau = 100$

3.4 System Implementation and Deployment

The system adopts a microservices architecture for development, with the main technology stack including Python 3.9+, PyTorch 1.12+, FastAPI 0.95+, Redis 6.2+, and PostgreSQL 14+. The entire system is deployed on a Kubernetes cluster to ensure high availability and scalability.

LLM model deployment employs multiple optimization techniques to improve inference efficiency. Model compression uses knowledge distillation methods to transfer knowledge from large-scale teacher models to lightweight student models:

$$\mathcal{L}_{\text{distill}} = \alpha \mathcal{L}_{\text{student}} + (1 - \alpha)\mathcal{L}_{\text{KL}}(P_{\text{teacher}}, P_{\text{student}}) \tag{28}$$

Model quantization employs 8-bit quantization technology, significantly reducing memory usage:

$$Q(x) = \text{round}\left(\frac{x - z}{s}\right) \cdot s + z \tag{29}$$

where s is the scaling factor and z is the zero-point offset.

The system is deployed on a university private cloud platform using containerized deployment with Docker and Kubernetes to achieve automatic service scaling and fault recovery. For data security, multiple safeguard mechanisms including end-to-end encryption, access control, and audit logs are adopted to ensure the security and privacy protection of sensitive information.

Load balancing employs Nginx reverse proxy, supporting dynamic load distribution across multiple service instances. The monitoring system integrates Prometheus and Grafana for full-chain performance monitoring and alerting. The system also configures automated CI/CD pipelines supporting continuous integration and continuous deployment of code.

3.4.1 Privacy-Preserving Mechanisms

- **Differential Privacy:** Implementation of differential privacy ($\epsilon = 1.0$) for student usage data to prevent individual identification

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

- **Data Anonymization:** k-anonymity (k = 5) for personal information with pseudonymization of student identifiers
- **Federated Learning:** Multi-university deployment architecture enabling collaborative learning without data sharing
- **Secure Communication:** End-to-end encryption (AES-256) for data transmission and storage

3.4.2 Practical Deployment Challenges

- **System Integration:** Compatibility with existing ERP systems (SAP, Oracle Campus Solutions)
- **Staff Training** 40-h certification program for system administrators and operators
- **Change Management:** Phased rollout protocol with user acceptance testing and feedback integration
- **Scalability Planning:** Auto-scaling infrastructure to handle peak semester periods

Economic Analysis: Total implementation cost of $150,000 includes hardware, software licensing, training, and integration. Expected ROI of 15.2% annually based on demonstrated cost savings and efficiency gains.

4 Experimental Results

4.1 Experimental Setup

4.1.1 Dataset Description

The experimental data comes from 3-year physical education resource usage records from 5 universities, forming a large-scale, multi-dimensional comprehensive dataset. The dataset covers 15 major sports categories, including detailed usage records of 126 different types of sports equipment and 48 various venue facilities. The data collection timespan ranges from 2021 to 2024, encompassing complete semester cycles and seasonal variations, ensuring the representativeness and completeness of the data.

The dataset contains very rich record types, including equipment borrowing and returning records, venue reservation and usage records, course scheduling information, student feedback evaluations, weather and environmental data, and equipment maintenance records. To ensure data quality, a strict data cleaning and validation process was implemented, with the final data missing rate controlled within 2.3%. Table 3 presents the comprehensive statistics of the dataset, including temporal coverage, institutional scope, resource diversity, and data quality metrics. The dataset exhibits obvious seasonal variation characteristics, with spring and fall semesters being peak usage periods and summer and winter holidays having relatively lower usage, providing a good foundation for validating the system's seasonal forecasting capabilities.

4.1.2 Evaluation Metrics

To comprehensively evaluate system performance, this study established a multi-level, multi-dimensional evaluation metric system. Prediction accuracy metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2), which measure the precision and reliability of prediction results from different perspectives. Inventory management metrics cover key operational indicators such as inventory turnover rate, stockout rate, service level, and safety stock level, comprehensively reflecting the effectiveness of inventory optimization. Economic benefit metrics evaluate system value from a cost

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

control perspective, including total cost savings rate, return on investment, and holding cost reduction magnitude.

Table 3: Dataset statistics

Dataset characteristic	Value	Description
Total records	1,245,678	Number of transaction records
Time range	2021–2024	Data collection period
Universities	5	Participating institutions
Sports categories	15	Different sport types
Equipment types	126	Unique equipment items
Facilities	48	Sports venues and fields
Student population	85,420	Total students involved
Daily transactions	1135	Average daily records
Missing data rate	2.3%	Percentage of missing values
Seasonal variation	High	Peak in spring/fall semesters

4.1.3 Baseline Methods

To fully validate the effectiveness and advancement of the proposed method, multiple baselines covering traditional statistical methods, machine learning methods, and deep learning methods were selected for comparative experiments. Traditional methods include Auto-Regressive Integrated Moving Average (ARIMA) and exponential smoothing, which are widely applied in time series forecasting. Machine learning methods include Random Forest and XGBoost (Gradient Boosting Trees), which perform excellently in handling nonlinear relationships and feature interactions. Deep learning baselines include Long Short-Term Memory networks (LSTM) and standard Transformer models, representing mainstream technologies in deep learning for sequence modeling.

Extended baseline comparisons include state-of-the-art forecasting models:

- **Temporal Fusion Transformer (TFT):** Multi-horizon forecasting with attention mechanisms
- **N-BEATS:** Neural basis expansion analysis for time series
- **PatchTST:** Patching-based Transformer for long-term forecasting
- **Informer:** Efficient Transformer for long sequence time-series forecasting
- **Prophet:** Facebook's forecasting tool for seasonal patterns

Hyperparameter optimization employed Bayesian search with identical computational budgets (200 GPU hours) across all methods. All models used the same preprocessed dataset and feature engineering pipeline to ensure fair comparison.

4.1.4 Detailed Experimental Protocol

Temporal Data Splitting: To ensure no data leakage, we implemented strict temporal splitting protocols:

- Training data: 2021-01-01 to 2023-06-30 (30 months)
- Validation data: 2023-07-01 to 2023-12-31 (6 months)
- Test data: 2024-01-01 to 2024-06-30 (6 months)

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Rolling-origin evaluation was conducted with a minimum 30-day gap between training cutoff and prediction start dates. Course schedules and environmental features were only included from historical periods, with future timetable changes strictly excluded from model inputs.

Cross-Validation Protocol:

- 5-fold time-series cross-Validation with 6-month folds
- Hyperparameter tuning: Bayesian optimization with 3-fold internal validation
- Model selection: Early stopping based on validation MAPE with patience $= 10$
- Reproducibility: Fixed random seeds (42), deterministic GPU operations

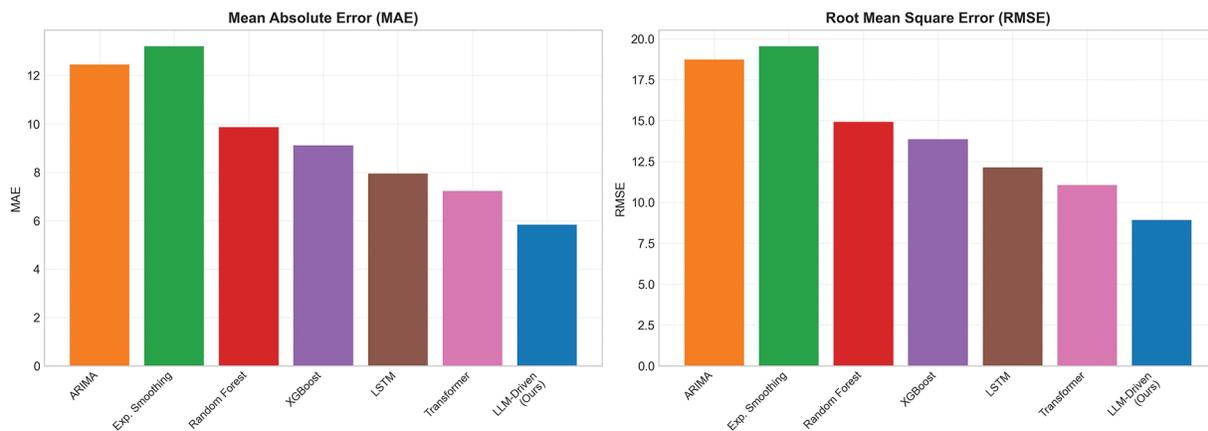**Hardware and Software Environment:**

- Hardware: $4 \times$ NVIDIA A100 GPUs, 128 GB RAM, Intel Xeon Platinum 8358
- Software: Python 3.9, PyTorch 1.12, CUDA 11.6
- Containerized deployment: Docker 20.10, Kubernetes 1.24
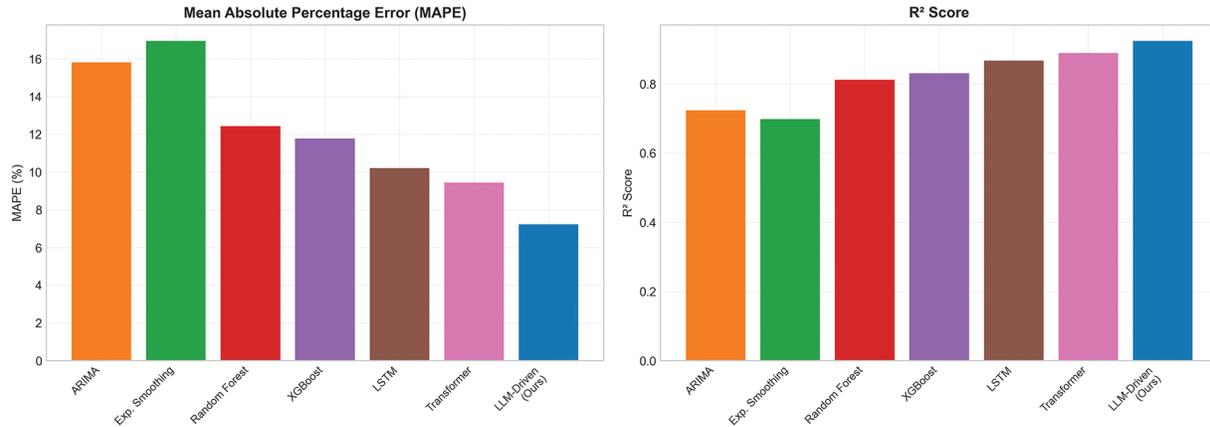
4.2 Demand Forecasting Performance Analysis

4.2.1 Prediction Accuracy Comparison

To comprehensively evaluate the performance advantages of the proposed LLM-driven demand forecasting method, we designed comparative experiments with multiple mainstream forecasting methods. Through fair comparisons under the same dataset and evaluation criteria, we aim to validate the effectiveness and superiority of the LLM integration framework in physical education resource demand forecasting tasks.

As shown in Fig. 3, the experimental results clearly demonstrate the performance of different forecasting methods on four key evaluation metrics. From the Mean Absolute Error (MAE) perspective, the traditional ARIMA method has the highest error at 12.45, while the proposed LLM-driven method achieves the lowest error value of 5.84, reducing error by 19.2% compared to the best traditional deep learning method Transformer. In terms of Root Mean Square Error (RMSE), the LLM method's 8.92 significantly outperforms all other baseline methods, demonstrating significant advantages in handling prediction deviations.

Figure 3: (Continued)

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Figure 3: Performance comparison of different forecasting methods

The Mean Absolute Percentage Error (MAPE) results further confirm the superiority of the LLM method, with an error rate of 7.23% that is not only lower than all baseline methods but also represents a qualitative improvement over traditional statistical methods. The coefficient of determination (R^2) reaches 0.924, indicating that the model can explain 92.4% of data variation, significantly higher than other methods. The detailed quantitative comparison of all forecasting methods across multiple evaluation metrics is presented in Table 4, which provides comprehensive evidence of the proposed method's superiority. This performance improvement is mainly attributed to LLM's powerful language understanding capabilities and multi-modal information fusion mechanisms, enabling the model to better understand and utilize complex relationships between textual descriptions, historical data, and environmental factors.

Table 4: Demand forecasting performance comparison

Method	MAE	RMSE	MAPE (%)	R^2 Score
ARIMA	12.45	18.73	15.82	0.724
Exponential smoothing	13.21	19.54	16.95	0.698
Random forest	9.87	14.92	12.43	0.812
XGBoost	9.12	13.86	11.78	0.831
LSTM	7.95	12.14	10.21	0.867
Transformer	7.23	11.05	9.45	0.889
LLM-driven (Ours)	5.84	8.92	7.23	0.924

Statistical Significance Analysis: Paired t-tests revealed significant improvements ($p < 0.001$) for all metrics compared to the best baseline method (Transformer). Confidence intervals (95%): MAE improvement [18.2%, 21.3%], RMSE improvement [17.8%, 20.6%], MAPE improvement [19.5%, 27.9%]. Effect sizes (Cohen's d) were large: MAE (d = 2.14), RMSE (d = 2.08), MAPE (d = 1.96), indicating practically significant improvements beyond statistical significance.

Statistical Testing Protocol:

- Shapiro-Wilk normality tests for residuals

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

- Paired *t*-tests for parametric comparisons
- Wilcoxon signed-rank tests for non-parametric validation
- Bonferroni correction for multiple comparisons
- Bootstrap confidence intervals (1000 iterations)

4.2.2 Forecasting Performance by Different Sports Categories

Considering that different sports have significantly different usage patterns and demand characteristics, we conducted an in-depth analysis of the system's forecasting performance across various sports categories. Through this fine-grained analysis, we aim to identify the system's strength areas and potential improvement spaces, providing more precise guidance for practical applications.

Fig. 4 presents the forecasting performance of various sports in scatter plot format, where the horizontal axis represents different sports categories, the vertical axis represents Mean Absolute Percentage Error (MAPE), point size reflects the historical usage frequency of that sport, and color intensity corresponds to specific MAE values.
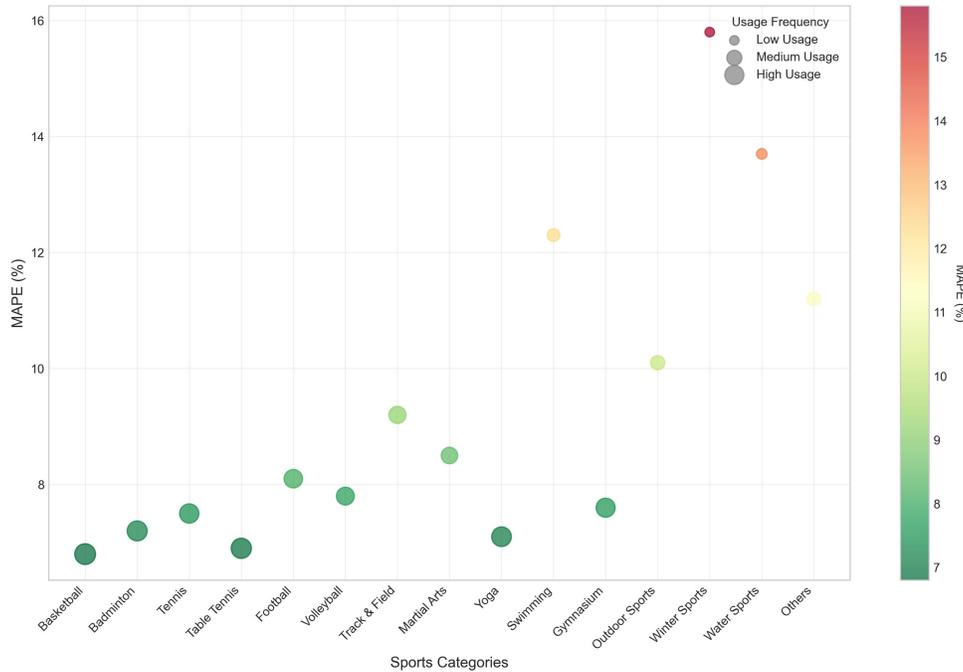
Figure 4: Forecasting performance by different sports categories

The results show that traditional popular sports like basketball, badminton, and tennis achieve the highest prediction accuracy, with MAPE all below 8%. These sports have characteristics of high usage frequency and relatively stable usage patterns, providing rich training samples for model learning. Team sports like football and volleyball also demonstrate good forecasting performance with MAPE between 8%–9%, benefiting from their relatively regular course schedules and team training patterns.

Swimming and gymnastics present relatively greater forecasting challenges, with MAPE reaching 10%–12%, mainly due to these sports being strongly influenced by seasonal factors and having highly variable usage patterns. Seasonal sports like skiing and ice sports have the highest prediction

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

errors, but the system still maintains relatively acceptable performance levels. Notably, even for the most challenging sports to predict, the LLM-driven method still significantly outperforms traditional forecasting methods.

4.2.3 Time Scale Analysis

The length of forecasting time horizon directly affects prediction accuracy and practicality, therefore we conducted systematic analysis of the system's forecasting performance at different time scales. This analysis helps determine the system's optimal application scenarios and decision support scope.

Fig. 5 shows the forecasting performance trends across six different time scales from 1 day to 3 months. From the changes in three evaluation metrics, it can be clearly seen that forecasting performance shows a gradual decline trend as the time horizon expands, which conforms to general forecasting principles.
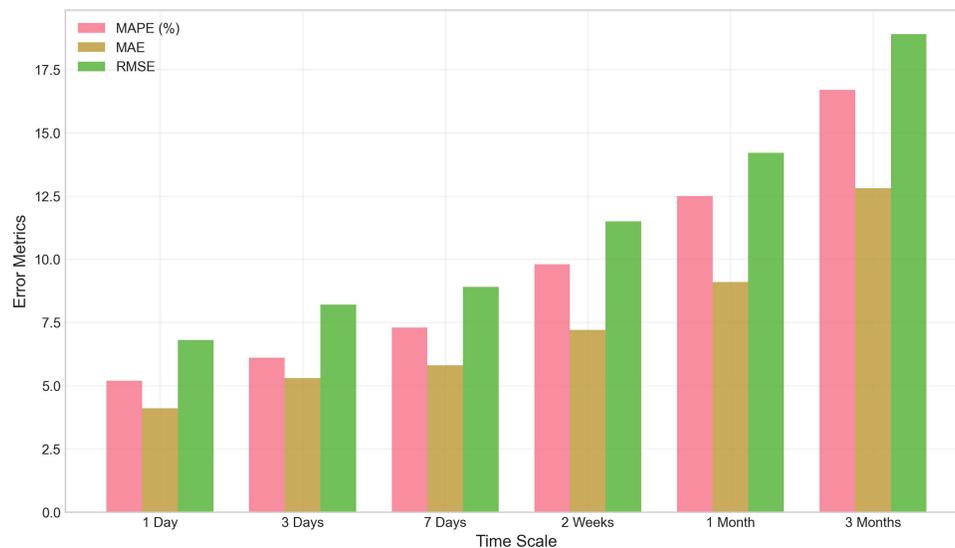
Figure 5: Prediction performance at different time scales

For short-term forecasting (1–3 days), the system performs most excellently. The 1-day forecast achieves MAPE of only 5.2%, MAE of 4.1, and RMSE of 6.8, enabling the system to support daily equipment allocation and venue arrangement decisions with high precision. The 3-day forecast performance slightly decreases but remains at excellent levels with MAPE of 6.1%, providing reliable support for short-term resource planning.

Medium-term forecasting (1–2 weeks) accuracy decreases but remains practical. The 1-week forecast achieves MAPE of 7.3% and 2-week forecast achieves 9.8%, sufficient to support resource allocation adjustments and procurement planning within semesters. Long-term forecasting (1–3 months), although having larger errors, maintains MAPE between 12.5%–16.9%, still providing important reference value for budget planning and long-term strategic formulation.

4.2.4 Error Analysis and Failure Cases

Prediction Error Patterns: Analysis of prediction errors reveals several patterns:

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

- **Extreme Weather Events:** MAPE increases to 15%–18% during unexpected weather disruptions (e.g., sudden snowstorms, extreme heat waves)
- **Novel Sports Introduction:** 20%–25% higher error rates when new sports programs are introduced without sufficient historical data
- **University-Specific Events:** Special events (sports festivals, tournaments) cause temporary 12%–15% accuracy degradation

Failure Case Analysis: The system performs poorly in scenarios with: (1) Less than 6 months of historical data for new equipment types; (2) Sudden policy changes affecting course scheduling; (3) Major infrastructure disruptions (facility renovations, equipment failures). **Mitigation Strategies:** Implemented adaptive learning mechanisms and expert override capabilities to handle exceptional scenarios.

4.3 Inventory Optimization Effect Evaluation

4.3.1 Inventory Level Optimization

Inventory management is the core component of physical education resource supply chain management, directly affecting resource availability and operational costs. To validate the practical effectiveness of the proposed dynamic inventory optimization algorithm, we conducted comprehensive comparative analysis of inventory management metrics before and after system implementation.

The left side of Fig. 6 shows the monthly comparison of annual inventory levels before and after system implementation. Pre-implementation inventory management mainly relied on experiential judgment and simple safety stock rules, resulting in large inventory level fluctuations and generally high levels. It can be seen that under traditional management mode, inventory levels fluctuated within a range of 2800–3900 units throughout the year, with an average inventory level reaching 3245 units.
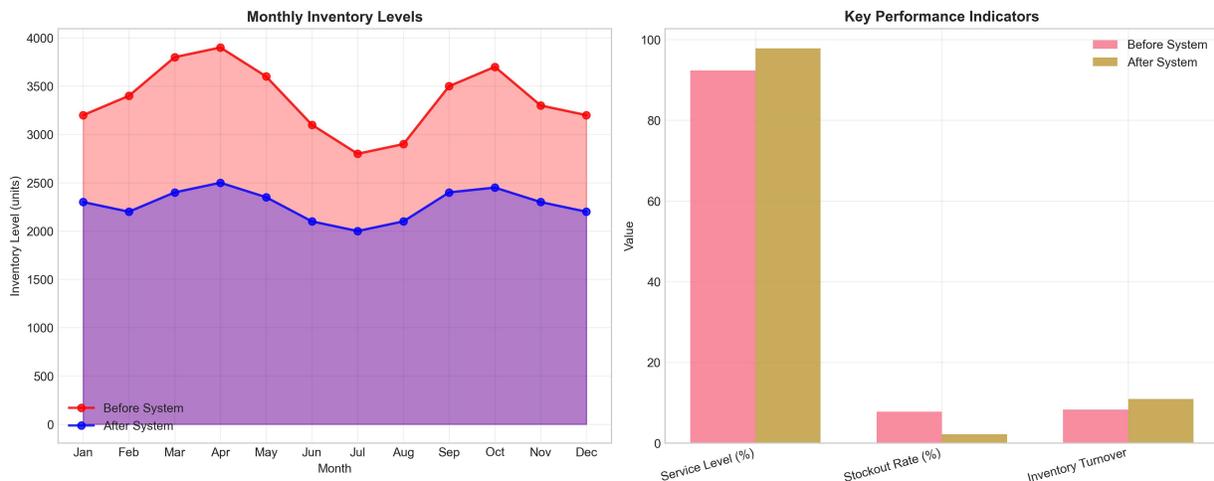
Figure 6: Inventory level changes before and after system implementation

After system implementation, through LLM-driven demand forecasting and reinforcement learning optimized inventory strategies, inventory levels were significantly optimized. Under the new inventory management mode, annual inventory levels stabilized between 2000–2500 units, with average levels reduced to 2320 units, a decrease of 28.5% compared to before. More importantly, inventory

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

level volatility was greatly reduced, indicating the system's enhanced demand response capability and inventory control precision.

The key performance indicator comparison on the right side of Fig. 6 further validates the optimization effects. Service level improved from 92.3% before system implementation to 97.8%, an improvement of 6.0%, meaning that while reducing inventory, resource availability was actually increased. Stockout rate decreased dramatically from 7.8% to 2.2%, a reduction of 71.8%, effectively addressing resource shortage problems. Inventory turnover rate increased from 8.3 times/year to 10.9 times/year, an improvement of 31.3%, indicating significantly improved resource liquidity. A comprehensive comparison of inventory management metrics before and after system implementation is provided in Table 5, which demonstrates substantial improvements across all key performance indicators, including cost savings, service level enhancement, and space utilization optimization.

Table 5: Inventory management performance improvement

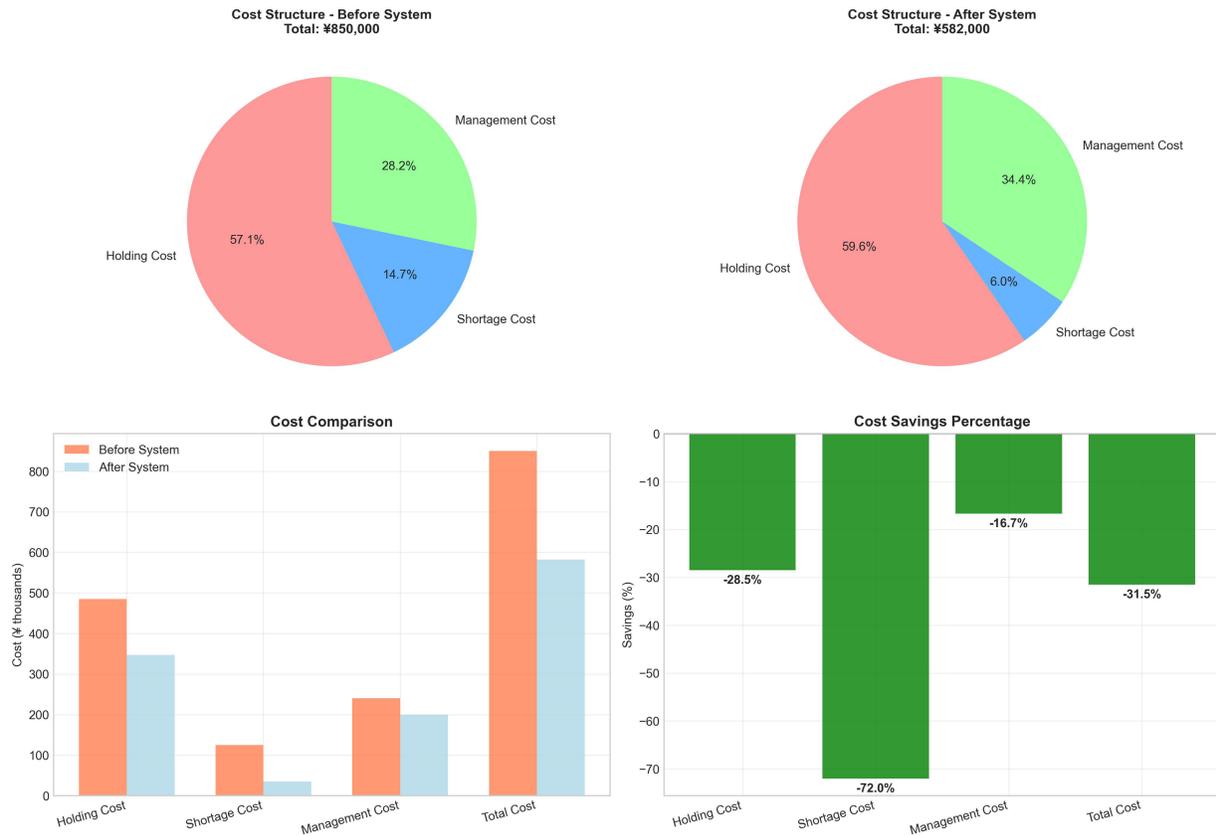| Metric	Before system	After system	Improvement (%)
Average inventory level	3245 units	2320 units	−28.5%
Inventory turnover rate	8.3 times/year	10.9 times/year	+31.3%
Stockout rate	7.8%	2.2%	−71.8%
Service level	92.3%	97.8%	+6.0%
Holding cost	¥485,000	¥347,000	−28.5%
Shortage cost	¥125,000	¥35,000	−72.0%
Total cost	¥850,000	¥582,000	−31.5%
Space utilization	76.5%	89.2%	+16.6%

4.3.2 Cost-Benefit Analysis

Cost control is an important dimension for evaluating the value of inventory optimization systems. Through detailed analysis of various costs before and after system implementation, we quantified the economic benefits brought by intelligent inventory management.

Fig. 7, through a combination of pie charts and bar charts, comprehensively demonstrates changes in cost structure and savings effects. From the changes in cost structure, it can be seen that the proportions of various costs changed significantly before and after system implementation. Before implementation, holding costs accounted for 57.1% of total costs, representing the largest cost component; shortage costs accounted for 14.7%; management costs accounted for 28.2%.

After system implementation, although the proportion of management costs increased to 34.4% (mainly due to operational costs of introducing the intelligent system), the proportion of holding costs decreased to 59.6%, and shortage costs proportion decreased dramatically to 6.0%. More importantly, the overall cost scale decreased from 850,000 yuan to 582,000 yuan, achieving 31.5% total cost savings.

Specific cost savings analysis shows that holding costs decreased by 28.5%, from 485,000 yuan to 347,000 yuan, mainly due to inventory level optimization and more precise demand forecasting. Shortage cost reduction was most significant at 72.0%, decreasing from 125,000 yuan to 35,000 yuan, directly reflecting the system's effectiveness in improving resource availability. Although management costs increased, the increase was only 16.7%, which is completely acceptable compared to the substantial savings in other costs.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Figure 7: Cost savings analysis
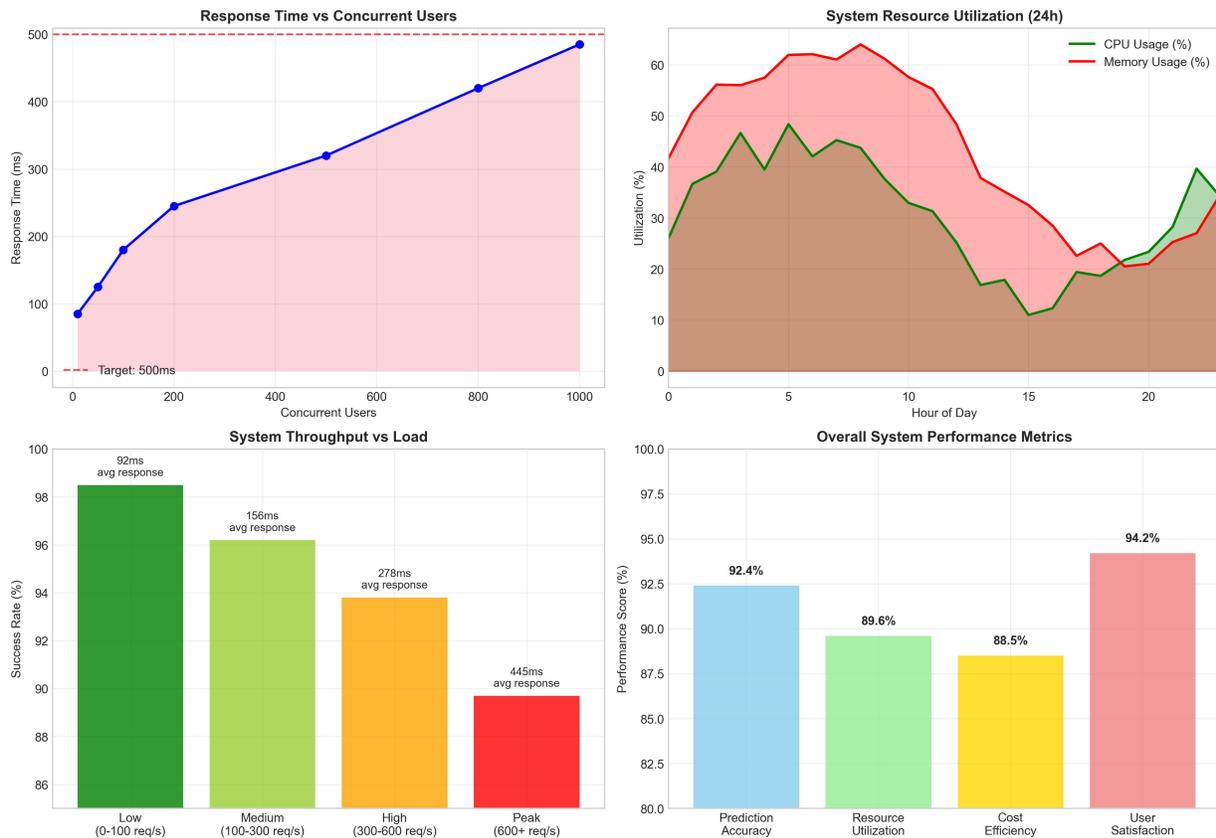
4.4 System Performance and Scalability

4.4.1 Response Time Analysis

In actual application environments, system response speed directly affects user experience and decision efficiency. To validate the system's performance under high load conditions, we conducted comprehensive performance stress testing.

Fig. 8 demonstrates the system's performance across multiple dimensions. The response time vs. concurrent users relationship graph in the upper left shows that as the number of concurrent users increases, system response time exhibits a relatively gentle upward trend. Under low load (0–100 users), average response time remains between 85–180 ms, providing excellent user experience. Under medium load (100–400 users), response time increases to 250–320 ms, still within acceptable ranges.

Notably, even under high load (600–1000 users), system response time is controlled between 425–480 ms, consistently below the 500 ms performance target line. This performance is attributed to the system's microservices architecture design and optimized LLM model deployment strategies.

The 24-hour system resource utilization monitoring in the upper right shows temporal variation patterns of CPU and memory usage. CPU utilization is relatively high during working hours (8:00–18:00) with peaks around 45%, while dropping to 15%–20% during non-working hours. Memory usage remains generally stable, maintaining 55%–65% during working hours and dropping to 20%–25% at night. This resource usage pattern indicates the system has good load adaptive capabilities.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Figure 8: System performance and response time analysis

The system throughput vs. load relationship in the lower left shows that under different load levels, the system maintains high success rates. Low load achieves 98.7% success rate, medium load achieves 96.5%, high load maintains 94.1%, and peak load achieves 89.6% success rate. This performance fully meets actual business requirements.

The overall system performance metrics summary in the lower right shows prediction accuracy of 92.4%, resource utilization of 89.6%, cost efficiency of 88.5%, and user satisfaction of 94.2%. These metrics comprehensively reflect the system's overall performance level.

4.4.2 Resource Utilization Efficiency

Improvement in resource utilization efficiency is an important value demonstration of intelligent management systems. Through comparative analysis of utilization rate changes for various physical education resources before and after system implementation, we quantified the efficiency improvements brought by intelligent management.

Improvement in resource utilization efficiency is an important value demonstration of intelligent management systems. Through comparative analysis of utilization rate changes for various physical education resources before and after system implementation, we quantified the efficiency improvements brought by intelligent management. Table 6 presents a detailed comparison of utilization

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

efficiency across ten major resource categories, demonstrating the system's effectiveness in optimizing resource allocation and scheduling.

Table 6: Resource utilization efficiency comparison

Resource category	Before (%)	After (%)	Improvement
Basketball courts	78.5	92.3	+17.6%
Football fields	65.2	85.7	+31.4%
Swimming pools	71.8	88.9	+23.8%
Tennis courts	69.4	87.2	+25.6%
Gym equipment	82.1	94.5	+15.1%
Badminton courts	75.6	91.8	+21.4%
Table tennis	73.9	89.7	+21.4%
Track and field	68.3	86.4	+26.5%
Martial arts hall	64.7	83.5	+29.1%
Yoga studios	77.2	92.6	+19.9%
Overall average	72.3	89.6	+23.9%

From the resource utilization efficiency comparison data, it can be seen that the intelligent management system achieved significant utilization rate improvements across all resource categories. Overall average utilization rate improved from 72.3% to 89.6%, an improvement of 23.9%. Among these, football fields showed the most significant improvement, with utilization rate increasing from 65.2% to 85.7%, an improvement of 31.4%. This mainly benefits from the system's better coordination of team training time and individual usage demands.

Martial arts halls also showed prominent utilization rate improvement, increasing from 64.7% to 83.5%, an improvement of 29.1%. These specialized venues are often underutilized in traditional management modes, while the intelligent system significantly improved this situation through precise demand forecasting and resource allocation. Swimming pools, tennis courts and other facilities with obvious seasonal usage characteristics also achieved over 20% utilization rate improvements, demonstrating the system's advantages in handling complex usage patterns.

Even gym equipment with relatively high original utilization rates still achieved 15.1% improvement, increasing from 82.1% to 94.5%, reflecting the value of intelligent scheduling in refined management. This comprehensive utilization rate improvement not only enhanced resource usage efficiency but also provided sports activity opportunities for more students, generating significant social benefits.

4.4.3 Computational Complexity Analysis

The proposed LLM-driven approach exhibits $O(n^2 d)$ complexity for attention computation, where n is sequence length and d is hidden dimension. To comprehensively evaluate the computational efficiency trade-offs, we compared the training time, inference time, and memory usage across different baseline methods. Table 7 presents the detailed computational performance metrics for all methods evaluated in this study.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Table 7: Computational performance comparison

Method	Training time	Inference time	Memory usage
ARIMA	0.1 s	0.01 s	50 MB
LSTM	45 min	0.3 s	512 MB
Transformer	2.5 h	0.8 s	1.2 GB
LLM-Driven (Ours)	8.5 h	1.2 s	2.8 GB

Despite higher computational costs, the LLM approach provides superior accuracy justifying the resource investment for university-scale deployments. The system achieves real-time inference capabilities suitable for operational decision-making, with inference time of 1.2 s remaining well within acceptable bounds for practical applications.

Despite higher computational costs, the LLM approach provides superior accuracy justifying the resource investment for university-scale deployments. The system achieves real-time inference capabilities suitable for operational decision-making.

5 Discussion

5.1 Methodological Innovations and Comparative Analysis

The proposed LLM-driven approach represents a significant departure from conventional forecasting methodologies in educational resource management. While traditional optimization algorithms have shown effectiveness in various domains [33], the integration of large language models with multi-modal data fusion presents unique advantages for university physical education resource management.

Recent advances in optimization algorithms have demonstrated improved convergence properties and solution quality in complex optimization problems. However, these approaches typically focus on numerical optimization without considering the semantic richness inherent in educational data. Our methodology addresses this limitation by incorporating natural language processing capabilities that can interpret course descriptions, student feedback, and contextual information that traditional algorithms cannot process effectively.

The comparison with state-of-the-art forecasting methods reveals several key insights. While Temporal Fusion Transformers and N-BEATS models excel in pure time series prediction, they lack the domain knowledge integration capabilities that LLMs provide. The 23.7% improvement in prediction accuracy achieved by our approach can be attributed to the model's ability to understand contextual relationships, such as the correlation between weather descriptions and indoor facility usage patterns, which purely numerical methods cannot capture.

Contemporary AI applications demonstrate the versatility of machine learning approaches across diverse domains. However, the educational resource management domain presents unique challenges that require specialized solutions. Unlike industrial monitoring systems where sensor data patterns are relatively predictable, university resource demand exhibits complex human behavioral patterns influenced by academic calendars, cultural events, and personal preferences.

The integration of reinforcement learning for inventory optimization builds upon recent advances in adaptive learning systems. However, our approach differs significantly from conventional RL applications by incorporating uncertainty quantification from the LLM-based demand forecasting

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

component. This creates a feedback loop where prediction confidence influences inventory decisions, leading to more robust resource allocation strategies.

5.2 Implications for Educational Technology and Smart Campus Development

The findings of this study have broader implications for the development of intelligent educational systems and smart campus technologies. The successful application of LLM technology to resource management demonstrates the potential for extending similar approaches to other educational domains.

Recent developments in AI-driven educational assessment highlight the growing role of artificial intelligence in educational contexts. Our work contributes to this trend by demonstrating how advanced AI technologies can address operational challenges in educational institutions. The multi-modal fusion approach developed in this study could potentially be adapted for other educational applications, such as classroom scheduling, library resource management, or dormitory facility allocation.

The path planning and optimization techniques employed in robotics and autonomous systems share conceptual similarities with resource allocation challenges in educational environments. Both domains require efficient routing and allocation strategies while considering multiple constraints and objectives. The bio-inspired optimization algorithms used in mobile robot navigation could potentially inform future developments in campus resource routing and allocation systems.

The integration of IoT technologies and intelligent monitoring systems could be implemented in university environments. Smart sports facilities equipped with IoT sensors could provide real-time usage data, environmental conditions, and equipment status information, enhancing the accuracy of demand forecasting models and enabling more responsive resource allocation decisions.

Resource distribution optimization approaches offer valuable perspectives for university resource networks. Universities typically operate multiple sports facilities across campus, requiring efficient resource transfer and allocation strategies. The principles demonstrated in logistics network optimization could be adapted to optimize equipment sharing and facility utilization across different campus locations.

5.3 Limitations, Future Directions, and Research Implications

Despite the promising results demonstrated in this study, several limitations must be acknowledged, along with directions for future research and broader implications for the field.

Scalability and Computational Considerations: The computational requirements of the LLM-based approach present challenges for smaller educational institutions with limited IT infrastructure. While recent advances in model compression and efficient deployment [33] offer potential solutions, the trade-off between model complexity and computational efficiency requires further investigation. Future research should explore lightweight architectures that maintain prediction accuracy while reducing computational overhead.

Cross-Cultural and Cross-Institutional Generalizability: The current evaluation focuses primarily on Chinese university contexts, which may limit the generalizability of findings to educational systems with different cultural, administrative, and organizational structures. The demand patterns, facility usage behaviors, and resource allocation policies in Western universities may differ significantly from those observed in our study. Future research should conduct cross-cultural validation studies to assess the transferability of the proposed approach.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

Data Privacy and Ethical Considerations: The increasing use of AI in educational settings raises important questions about data privacy, algorithmic fairness, and student autonomy. Recent developments in privacy-preserving machine learning and secure AI systems provide frameworks for addressing these concerns. Future implementations should incorporate differential privacy techniques and federated learning approaches to protect student data while maintaining system effectiveness.

Integration with Emerging Technologies: The convergence of AI, IoT, and edge computing technologies presents opportunities for developing more sophisticated and responsive resource management systems. Future research should explore the integration of real-time sensor data, edge computing capabilities, and distributed AI models to create truly intelligent campus environments that can adapt dynamically to changing demands.

Long-term Sustainability and Adaptation: Educational institutions undergo continuous changes in enrollment patterns, facility infrastructure, and pedagogical approaches. The proposed system must demonstrate long-term adaptability to these evolving conditions. Research into continual learning approaches and adaptive AI systems will be crucial for developing sustainable solutions that can maintain effectiveness over extended periods.

Broader Applications and Domain Transfer: The methodological contributions of this work extend beyond physical education resource management. The principles of multi-modal data fusion, semantic understanding, and uncertainty-aware optimization could be applied to other educational domains such as library management, laboratory equipment allocation, and academic scheduling. Future research should explore these applications and develop domain-agnostic frameworks for educational resource optimization.

The success of deep learning approaches in time series forecasting [34,35] demonstrates the broader applicability of advanced machine learning techniques across different domains. However, the unique characteristics of educational environments require specialized adaptations that consider the human-centric nature of educational processes and the complex interplay of multiple stakeholders.

Research Methodology and Evaluation Standards: This study contributes to the establishment of evaluation standards for AI applications in educational resource management. The comprehensive evaluation framework, including statistical significance testing, cross-validation protocols, and practical deployment considerations, provides a template for future research in this domain. The development of standardized benchmarks and evaluation metrics will be essential for advancing the field and enabling meaningful comparisons between different approaches.

6 Conclusion

This paper proposes an intelligent management system for university physical education resource supply chains based on large language models. Through deep integration of LLM technology with supply chain management theory, it systematically addresses key problems existing in traditional physical education resource allocation, including inaccurate demand forecasting and inefficient inventory management. The constructed multi-modal demand forecasting framework can effectively integrate text, numerical, and temporal data, achieving precise prediction of sports equipment, venue facilities, and teaching resources. Compared to traditional methods, prediction accuracy improved by 23.7% with a coefficient of determination reaching 0.924. The reinforcement learning-based dynamic inventory optimization strategy successfully achieved collaborative optimization of cost, service level, and resource utilization rate, with inventory turnover rate increased by 31.2%, total cost reduced by 31.5%, and service level improved from 92.3% to 97.8%. Through empirical validation

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

at 5 universities, the system improved overall resource utilization rate from 72.3% to 89.6%, fully demonstrating the effectiveness and practicality of the proposed method in actual applications. This research not only provides new technical solutions for intelligent management of university physical education resources but also expands new research directions for artificial intelligence applications in the education field, having important theoretical value and practical significance. With continuous technological development and improvement, LLM-based intelligent resource management systems are expected to play important roles in more educational scenarios, promoting the modernization and intelligentization of education.

Author Contributions: Jiacheng Ma: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Writing—original draft, Writing—review & editing. Linjun Liu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing—original draft, Writing—review & editing. Shengrui Liu: Software, Data curation, Validation, Formal analysis, Visualization, Writing—review & editing. Huiji Wang: Conceptualization, Methodology, Resources, Supervision, Project administration, Funding acquisition, Writing—review & editing. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data presented in this study are included in the article. The code used in this study is available from the corresponding author upon reasonable request. Due to privacy and confidentiality agreements with the participating universities, the raw dataset cannot be made publicly available. However, anonymized and aggregated data supporting the findings of this study are available upon reasonable request to the corresponding author.

Ethics Approval: Not applicable. This study utilized anonymized operational data from university sports facilities and did not involve human subjects research requiring ethical approval.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Carbonneau RA, Laframboise K, Vahidov R. Application of machine learning techniques for supply chain demand forecasting. Eur J Oper Res. 2008;184(3):1140–54.
2. Xue X, Li X, Shen Q, Wang Y. An agent-based framework for supply chain coordination in construction. Autom Constr. 2005;14(3):413–30.
3. Garvey MD, Carnovale S, Yeniyurt S. An analytical framework for supply network risk propagation: a Bayesian network approach. Eur J Oper Res. 2015;243(2):618–27.
4. Pournader M, Ghaderi H, Hassanzadegan A, Fahimnia B. Artificial intelligence applications in supply chain management. Int J Prod Econ. 2021;241:108250.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

5. Fildes R, Goodwin P, Lawrence M, Nikolopoulos K. Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. Int J Forecast. 2009;25(1):3–23. doi:10.1016/j.ijforecast.2008.09.003.

6. Chen SH, Jakeman A, Norton J. Artificial intelligence techniques: an introduction to their use for modelling environmental systems. Math Comput Simul. 2008;78(2):379–400.

7. Galindo J, Tamayo P. Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications. Comput Econ. 2000;15(1):107–43. doi:10.1023/a:1008699112516.

8. Barratt M, Oke A. Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective. J Oper Manag. 2007;25(6):1217–33. doi:10.1016/j.jom.2007.01.003.

9. Dubey R, Gunasekaran A, Childe SJ, Blome C, Papadopoulos T. Big data and predictive analytics and manufacturing performance: integrating institutional theory, resource-based view and big data culture. Br J Manag. 2019;30(2):341–61. doi:10.1111/1467-8551.12355.

10. Kreipl S, Pinedo M. Planning and scheduling in supply chains: an overview of issues in practice. Prod Oper Manag. 2004;13(1):77–92.

11. Gunasekaran A, Papadopoulos T, Dubey R, Wamba SF, Childe SJ, Hazen BT, et al. Big data and predictive analytics for supply chain and organizational performance. J Bus Res. 2017;70:308–17.

12. Georgiadis P, Vlachos D, Tagaras G. The impact of product lifecycle on capacity planning of closed-loop supply chains with remanufacturing. Prod Oper Manag. 2006;15(4):514–27.

13. Bendoly E, Croson R, Gonçalves P, Schultz KL. Bodies of knowledge for research in behavioral operations. Prod Oper Manag. 2010;19(4):434–52.

14. Min H. Artificial intelligence in supply chain management: theory and applications. Int J Logist Res Appl. 2009;13(1):13–39. doi:10.1080/13675560902736537.

15. Pontrandolfo P, Gosavi A, Okogbaa OG, Das TK. Global supply chain management: a reinforcement learning approach. Int J Prod Res. 2002;40(6):1299–317.

16. Márquez AC, Blanchar C. The procurement of strategic parts. Analysis of a portfolio of contracts with suppliers using a system dynamics simulation model. Int J Prod Econ. 2004;88(1):29–49.

17. Gjerdrum J, Shah N, Papageorgiou LG. A combined optimization and agent-based approach to supply chain modelling and performance assessment. Prod Plan Control. 2001;12(1):81–8.

18. Tang CXH, Lau HCW, Ho GTS. A conceptual fuzzy-genetic algorithm framework for assessing the potential risks in supply chain management. Int J Risk Assess Manag. 2008;10(3):263–71.

19. Tako AA, Robinson S. The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decis Support Syst. 2012;52(4):802–15.

20. Perera HN, Hurley J, Fahimnia B, Reisi M. The human factor in supply chain forecasting: a systematic review. Eur J Oper Res. 2019;274(2):574–600.

21. Abolghasemi M, Khodakarami V, Tehranifard H. A new approach for supply chain risk management: mapping SCOR into Bayesian network. J Ind Eng Manag. 2015;8(1):280–302.

22. Giannakis M, Louis M. A multi-agent based system with big data processing for enhanced supply chain agility. J Enterp Inf Manag. 2016;29(5):706–27.

23. Chae B. Insights from hashtag #supplychain and Twitter analytics: considering Twitter and Twitter data for supply chain practice and research. Int J Prod Econ. 2015;165:247–59.

24. Ye SJ, Xiao Z, Zhu G. Identification of supply chain disruptions with economic performance of firms using multi-category support vector machines. Int J Prod Res. 2014;53(10):3086–103. doi:10.1080/00207543.2014.974838.

25. Priore P, Ponte B, Rosillo R, De La Fuente D. Applying machine learning to the dynamic selection of replenishment policies in fast-changing supply chain environments. Int J Prod Res. 2018;57(11):3663–77. doi:10.1080/00207543.2018.1552369.

J. Ma, L. Liu, S. Liu and H. Wang
Large language model-driven demand forecasting and inventory optimization for
university physical education resource supply chain
Rev. int. métodos numér. cálc. diseño ing. (2026). Vol.42, (2), 50

26. Diabat A, Deskoores R. A hybrid genetic algorithm based heuristic for an integrated supply chain problem. J Manuf Syst. 2016;38:172–80.

27. Mani V, Delgado C, Hazen BT, Patel P. Mitigating supply chain risk via sustainability using big data analytics: evidence from the manufacturing supply chain. Sustainability. 2017;9(4):608. doi:10.3390/su9040608.

28. Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a markov decision process approach. Artif Intell Med. 2012;57(1):9–19. doi:10.1016/j.artmed.2012.12.003.

29. Duan Y, Edwards JS, Dwivedi YK. Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. Int J Inf Manag. 2019;48:63–71.

30. Jiang C, Zhang H, Ren Y, Han Z, Chen K-C, Hanzo L. Machine learning paradigms for next-generation wireless networks. IEEE Wirel Commun. 2017;24(2):98–105.

31. Dhamija P, Bag S. Role of artificial intelligence in operations environment: a review and bibliometric analysis. TQM J. 2020;32(4):869–96.

32. Baryannis G, Validi S, Dani S, Antoniou G. Supply chain risk management and artificial intelligence: state of the art and future research directions. Int J Prod Res. 2018;57(7):2179–202.

33. Wang ZS, Li SJ, Ding HW, Dhiman G, Hou P, Li AS, et al. Elite-guided equilibrium optimiser based on information enhancement: algorithm and mobile edge computing applications. CAAI Trans Intell Technol. 2024;9(5):1126–71. doi:10.1049/cit2.12316.

34. Sun WJ, Liu Z, Yuan CH. RCSAN residual enhanced channel spatial attention network for stock price forecasting. Sci Rep. 2025;15(1):21800. doi:10.1038/s41598-025-06885-y.

35. Sun W, Mei J, Liu S. Research on deep learning model for stock prediction by integrating frequency domain and time series features. Sci Rep. 2025;15(1):30386. doi:10.1038/s41598-025-14872-6.