OPEN ACCESS
International Journal of High Speed Electronics and Systems
(2025) 2540865 (28 pages)

DOI: 10.1142/S0129156425408654

Dynamic Pricing Strategy Driven by Deep Reinforcement Learning with Empirical Analysis on the Collaborative Optimization of Hotel Revenue Management and Customer Satisfaction

Zhenhua Mei

Department of Cultural Tourism, Zibo Vocational Institute,
Zibo 255314, Shandong, P. R. China
mei2008235@163.com

Received 9 May 2025
Revised 9 June 2025
Accepted 6 July 2025
Published 13 August 2025

In the context of hotel revenue management, dynamic pricing plays a crucial role in maximizing revenue while maintaining a delicate balance with customer satisfaction. Traditional pricing strategies often depend on static rules or overly simplistic models that lack the ability to adapt to real-time changes in market demand, evolving customer behavior, and competitive trends. These outdated approaches can lead to missed revenue opportunities and suboptimal guest experiences. This research addresses these challenges by proposing a dynamic pricing strategy driven by deep reinforcement learning, which integrates real-time data streams with predictive analytics to create a highly responsive and intelligent pricing system. At the core of the proposed methodology is a novel framework that combines advanced deep neural network architectures with adaptive optimization algorithms. This framework is designed to optimize both pricing and inventory decisions across multiple booking channels simultaneously. The Innovative Pricing Transformer (IPT) model underpins this framework by leveraging attention mechanisms and temporal sequence modeling to accurately forecast future demand and recommend context-aware pricing decisions. In addition, the Adaptive Yield Optimization Strategy (AYOS) refines this process by incorporating real-world operational constraints such as overbooking policies, price parity requirements, and channel-specific pricing rules, ensuring practicality and compliance. Empirical analysis conducted on real-world datasets reveals that our approach consistently outperforms traditional pricing models, not only in revenue enhancement but also in improving overall customer satisfaction. The proposed strategy represents a scalable, efficient, and intelligent solution for modern hotel revenue management, enabling hotels to remain agile and competitive in dynamic and uncertain market conditions.

Keywords: Dynamic pricing; deep reinforcement learning; hotel revenue management; adaptive yield optimization; innovative pricing transformer.

1. Introduction

The growing complexity of the global hospitality industry has prompted hotels to seek more efficient and dynamic approaches to pricing and revenue management [44]. Traditional methods often struggle to adapt to fast-changing market conditions, customer preferences, and competitor strategies [4]. As the demands of both profitability and customer satisfaction continue to rise, the need for a more adaptive pricing strategy becomes apparent [25]. Moreover, achieving a balance between maximizing revenue and ensuring a positive customer experience is a critical challenge. In response, dynamic pricing strategies driven by advanced machine learning techniques, particularly deep reinforcement learning (DRL), have emerged as powerful tools [30]. Not only do these methods offer more flexibility in pricing, but they also allow hotels to optimize their decisions in real-time, taking into account the complex interplay between pricing, customer demand, and satisfaction [24].

Initial strategies for hotel pricing were largely grounded in manually designed rule-based frameworks that incorporated the practical insights and subjective judgment of hotel managers. These early systems typically relied on fixed rules, derived from industry experience, to determine room rates according to parameters such as seasonal demand cycles, day-of-week patterns, local events, and predefined occupancy thresholds [13]. The pricing decisions were often static, updated infrequently, and lacked real-time responsiveness. Such approaches provided a transparent and easily interpretable structure for revenue managers, making them relatively simple to implement and manage within operational workflows [37]. Although these rule-based systems were effective under relatively stable and predictable market conditions, they revealed significant limitations when applied in rapidly changing or highly competitive environments [38]. Their inability to process and react to sudden fluctuations in guest booking behavior, last-minute cancellations, or emerging trends in competitor pricing restricted their strategic flexibility. Furthermore, the heavy reliance on human intuition and periodic manual adjustments introduced inconsistencies and made it difficult to optimize revenue in a systematic manner [8]. As market conditions became increasingly volatile and customer expectations more dynamic, the limitations of these static pricing models became more pronounced, highlighting the need for more adaptive and data-responsive pricing mechanisms [43].

As the market grew more complex and competitive, pricing strategies in the hospitality sector began to evolve beyond static rule-based systems. Models that incorporated historical booking trends, external demand drivers such as local events or economic indicators, and broader business intelligence metrics became increasingly prominent. These approaches enabled a more nuanced and data-informed perspective on customer behavior and market dynamics, facilitating more responsive and strategic pricing interventions [6]. By leveraging statistical analysis and structured data inputs, hotels were able to move toward evidence-based pricing decisions, reducing reliance on human intuition and subjective judgment. Techniques such as time series forecasting, regression analysis, and market segmentation modeling allowed hoteliers to anticipate demand fluctuations with greater precision and tailor pricing to specific customer groups or market conditions [10]. These models offered enhanced tactical flexibility by enabling dynamic adjustments to pricing based on anticipated trends rather than fixed schedules. The integration of competitor rate monitoring and event-based modifiers provided further contextual awareness in pricing decisions. However, despite these improvements, such methods still exhibited limitations [11]. Their effectiveness often hinged on the availability of large, clean datasets and stable demand patterns. In fast-changing or highly uncertain environments — such as during sudden market disruptions or shifts in consumer sentiment — their predictive accuracy could degrade significantly. Moreover, these models typically focused on optimizing revenue metrics in isolation, with limited ability to incorporate qualitative factors such as customer satisfaction, brand perception, or loyalty impacts [14]. As a result, while offering better control than rule-based systems, these approaches still fell short of delivering fully adaptive and customer-centric revenue management solutions [7].

The latest advancements in pricing technologies have given rise to a new generation of intelligent, autonomous systems that are capable of continuously learning and adapting to their environment. These systems represent a significant departure from earlier rule-based or static data modeling approaches, as they operate through iterative learning processes that refine pricing strategies based on observed outcomes [3]. Recent research has increasingly focused on leveraging reinforcement-based learning mechanisms, which allow pricing algorithms to make decisions by interacting with dynamic market environments and receiving feedback from customer booking behaviors, competitor movements, and revenue performance indicators [33]. Unlike traditional predictive models, these adaptive frameworks do not require explicit supervision or fixed training data structures; instead, they evolve through trial-and-error processes, optimizing policies over time to respond effectively to shifting conditions. A key advantage of these advanced systems lies in their capacity to optimize for multiple objectives simultaneously [27]. For example, they can also account for longer-term considerations such as customer retention, booking conversion rates, and even perceived fairness in pricing. This multi-objective optimization allows for a more balanced and sustainable revenue management approach that aligns financial outcomes with broader strategic goals, including brand loyalty and customer satisfaction [35]. Moreover, some implementations have begun integrating user segmentation and personalization modules, further refining price recommendations at the individual level and enhancing the guest experience. Despite their promise, however, these methods are not without challenges. Implementing such models in real-world hotel systems often demands substantial computational resources and infrastructure support, particularly when processing large volumes of real-time data across multiple booking channels. The complexity and opacity of these models — often referred to as "black-box" behavior — make it difficult for revenue managers to interpret or validate pricing decisions, raising concerns about trust, transparency, and accountability. Aligning algorithmic outputs with human-centric values, such as fairness, cultural expectations, and guest sentiment, remains an ongoing area of research [42]. As these technologies continue to evolve, addressing these operational and ethical challenges will be essential for their broader acceptance and effectiveness in the hospitality industry.

Based on the limitations of the aforementioned approaches, our method offers a novel solution to the challenges of balancing hotel revenue management and customer satisfaction. By incorporating a hybrid model that combines deep reinforcement learning with advanced empirical analysis, our approach not only optimizes dynamic pricing but also incorporates real-time customer feedback. This allows for a more comprehensive understanding of customer preferences and the ability to make pricing decisions that consider both profitability and satisfaction. The integration of these two elements enables more precise and adaptive pricing, resulting in better outcomes for both hotels and customers.

- Our method offers a more adaptive and real-time approach to pricing, improving the responsiveness to market dynamics.
- The hybrid model's ability to balance revenue maximization and customer satisfaction ensures more sustainable long-term business strategies.
- Empirical results demonstrate that our approach leads to improved revenue outcomes and higher levels of customer satisfaction compared to traditional methods.

2. Related Work

2.1. *Deep reinforcement learning for dynamic pricing*

Dynamic pricing represents a pivotal mechanism in revenue management, wherein the price of a product or service is dynamically adjusted according to factors including demand fluctuations, customer behavior, and competitive market conditions [18]. In recent years, Deep Reinforcement Learning (DRL) has been increasingly employed to enhance dynamic pricing strategies, integrating reinforcement learning with deep neural networks to enable more adaptive and sophisticated pricing models. By learning optimal pricing policies through trial-and-error interactions with a dynamic environment, DRL-based models have demonstrated significant potential to maximize long-term objectives such as revenue and profit. In sectors characterized by high volatility, such as the hospitality industry, DRL offers considerable advantages due to its capability to address price sensitivity, fluctuating demand, and intense competition [12]. Agents trained with DRL methodologies are designed to adjust prices proactively in response to continuously evolving market conditions, striving for a balance between immediate revenue and broader business objectives like customer satisfaction and occupancy optimization. Several studies have multi-agent frameworks wherein DRL agents either cooperate or compete to devise optimal pricing strategies, considering elements such as room availability, competitor pricing dynamics, historical booking patterns, and inferred customer preferences [28]. Beyond mere price setting, DRL enables complex decision-making processes, including the determination of optimal timings for promotional activities or discount offerings, further expanding the strategic possibilities for revenue management. One of the fundamental strengths of DRL in this context lies in its capacity to process vast quantities of heterogeneous data and model intricate, non-linear interdependencies between numerous variables, a feature critical for adapting to the multifaceted nature of real-world marketplaces [26]. Empirical investigations corroborate the theoretical advantages of DRL, with experimental results indicating that DRL-driven pricing systems consistently outperform traditional rule-based or static pricing models in metrics such as revenue maximization and customer satisfaction improvement. Continued research endeavors aim to enhance the transparency, interpretability, and computational efficiency of DRL algorithms, addressing persistent challenges including overfitting risks and the high resource demands associated with model training and deployment. Nevertheless, implementing DRL-based dynamic pricing in operational environments, particularly those characterized by high uncertainty and variability like the hospitality industry, remains a complex challenge necessitating further advancements in model robustness and real-time decision-making capabilities [2].

2.2. *Collaborative optimization in revenue management*

Collaborative optimization within the domain of revenue management emphasizes the concurrent enhancement of multiple objectives, including revenue maximization and customer satisfaction, through the alignment of interests among various stakeholders. In the hospitality sector, these stakeholders primarily encompass hotel management teams, customers, and intermediaries such as online travel agencies [29]. Traditional revenue management systems (RMS) have largely prioritized revenue optimization by adjusting prices based on forecasted demand, yet the increasing significance of customer loyalty and satisfaction necessitates a more balanced approach [15]. Collaborative optimization methods address this need by integrating broader stakeholder considerations into the revenue management framework. A critical component involves embedding customer satisfaction metrics into pricing strategies, recognizing that short-term revenue gains achieved through aggressive pricing may adversely affect long-term customer loyalty [17]. Incorporating indicators such as guest ratings, reviews, and feedback allows for a more nuanced optimization process, wherein hotels strive to enhance guest experiences alongside achieving financial objectives. Recent research has investigated the synergy between collaborative optimization and machine learning techniques, developing predictive models that jointly assess revenue potential and customer satisfaction outcomes to inform pricing decisions. These models leverage multi-source data inputs, including competitor prices, demographic profiles, booking histories, and external environmental variables such as seasonality and weather conditions, to refine pricing strategies. Achieving effective collaborative optimization also requires cross-departmental integration within hotel organizations, fostering cooperation among marketing, sales, and operational teams to ensure coherent and mutually reinforcing strategies [40]. Empirical evidence indicates that collaborative optimization can substantially enhance revenue management performance by acknowledging and balancing both economic imperatives and relational considerations intrinsic to customer interactions. Nonetheless, several challenges persist, notably the complexities involved in multi-source data integration, the computational demands of sophisticated optimization algorithms, and the necessity for ongoing adaptability to evolving market dynamics and customer expectations.

2.3. *Empirical analysis in hotel revenue management*

Empirical analysis in hotel revenue management constitutes a fundamental approach for evaluating the effectiveness of pricing strategies and their influence on both operational performance and customer satisfaction. Unlike purely theoretical frameworks, empirical methodologies rely on real-world data collection and statistical examination to assess how dynamic pricing techniques, including those utilizing machine learning and deep reinforcement learning (DRL), perform under actual market conditions [31]. The hospitality industry presents a complex environment characterized by fluctuating demand, seasonal variability, diverse customer preferences, and intense competition, necessitating the comprehensive analysis of datasets encompassing booking behaviors, historical pricing, and customer feedback. One central theme in empirical research is the assessment of dynamic pricing strategies on key hotel performance indicators, such as occupancy rates, average daily rate (ADR), revenue per available room (RevPAR), and total revenue generation [41]. Findings indicate that dynamic pricing, particularly when informed by DRL models, can yield substantial improvements across these metrics [36]. Nevertheless, empirical studies highlight that the success of dynamic pricing is contingent upon its alignment with customer expectations; aggressive revenue-focused pricing often correlates with diminished customer loyalty and adverse reputational impacts. Integrated strategies that balance financial objectives with customer satisfaction measures are thus increasingly advocated. Customer segmentation has also emerged as a critical focus, with empirical analyses demonstrating that personalized pricing strategies based on segmentation criteria such as booking behavior, price elasticity, and historical interactions significantly enhance both revenue and customer experience outcomes [34]. In parallel, empirical evaluations of collaborative optimization frameworks reveal that incorporating multi-stakeholder perspectives in pricing decisions fosters a more sustainable revenue growth trajectory. These studies leverage longitudinal datasets to validate the real-world applicability of advanced algorithms, providing actionable insights into operationalizing machine learning and reinforcement learning models within hotel management contexts. Despite these advancements, empirical research faces challenges including data acquisition barriers, the intricacies of accurately modeling complex customer behaviors, and the need for continuous recalibration to adapt to dynamic market conditions [32]. Nonetheless, empirical analysis remains indispensable for refining dynamic pricing methodologies and for ensuring that revenue management practices are both effective and customer-centric in an increasingly competitive hospitality landscape.

3. Method

3.1. *Overview*

This section introduces a comprehensive framework designed to optimize hotel revenue by leveraging a synergistic integration of formal mathematical modeling, deep neural network architectures, and adaptive control strategies. The framework aims to address the dynamic and multi-faceted nature of revenue management by combining theoretical rigor with practical adaptability. By unifying these components, the proposed system is capable of capturing complex demand patterns, learning from historical and real-time data, and dynamically adjusting pricing and inventory strategies across multiple booking channels. The remainder of this section is organized into three key subsections. First, we present the mathematical formalization of the revenue optimization problem, including the definition of relevant variables, constraints, and objective functions. This establishes a solid analytical foundation for the subsequent computational modules. Second, we describe the architectural design of the proposed neural framework, which incorporates attention mechanisms, temporal modeling, and feature fusion techniques to enhance the predictive capabilities of the system. We outline the control strategy responsible for decision-making and policy updates, which is implemented using reinforcement-based optimization methods that continuously adapt to environmental changes. Together, these components form an integrated solution that is both theoretically sound and practically scalable for real-world hotel revenue management scenarios.

Section 3.2 defines the core constructs of the hotel revenue management problem, encompassing hotel inventory, customer arrival processes, booking channels, and dynamic pricing. These constructs are formalized within a structured symbolic framework, incorporating canonical definitions, behavioral assumptions on customers, stochastic demand distributions, and temporal operational constraints. This formal foundation captures the essential complexities and constraints of hotel operations in a rigorous and extensible manner, serving as the basis for subsequent methodological development. Section 3.3 presents the Innovative Transformer (IPT) model, a neural architecture that combines temporal sequence modeling with demand-sensitive attention mechanisms. The IPT model captures patterns and price sensitivity across multiple channels, time windows, and market segments through a multi-scale transformer structure. Its design integrates feature-wise context encoding, temporal convolutional priors, and cross-channel correlation modules, enabling expressive modeling of demand elasticity and booking lead time effects. Section 3.4 introduces the Adaptive Yield Optimization Strategy (AYOS), a control framework that utilizes the IPT's predictions under domain-aware constraints, including overbooking buffers, corporate block reservations, and price parity rules. AYOS formulates a constrained dynamic programming problem that balances immediate expected revenue with long-term capacity utilization objectives, incorporating a dual-objective optimization criterion. A feedback refinement mechanism, employing real-time bandit optimization and regret minimization, aligns the strategy dynamically with evolving booking trends and competitive landscapes. Together, these components establish an explainable, adaptive, and high-performing system for intelligent revenue optimization in the hospitality sector.

3.2. *Preliminaries*

We consider a hotel revenue management setting where the property manages a finite inventory of rooms over a discrete time horizon $\mathcal{T} = \{1, 2, \ldots, T\}$. Each day $t \in \mathcal{T}$ requires decisions on room pricing, availability, and allocation across multiple customer segments and booking channels.

Let $\mathcal{I}$ denote the inventory space, where I_t represents the number of available rooms at time t. The hotel serves C customer segments, indexed by $\mathcal{C} = \{1, 2, \ldots, C\}$, and operates across K distribution channels, indexed by $\mathcal{K} = \{1, 2, \ldots, K\}$. For each pair (c, k) with $c \in \mathcal{C}$ and $k \in \mathcal{K}$, the price offered at time t is denoted by $p_t^{(c,k)}$.

Customer demand $D_t^{(c,k)}$ is treated as a random variable conditioned on the offered price and external covariates $x_t \in \mathbb{R}^d$. The demand is assumed to follow a segment- and channel-specific distribution.

$$D_t^{(c,k)} \sim \mathcal{D}^{(c,k)}(p_t^{(c,k)}, x_t). \tag{1}$$

The total revenue at time t is given as

$$R_t = \sum_{c=1}^{C} \sum_{k=1}^{K} a_t^{(c,k)} \cdot p_t^{(c,k)} \cdot \min(\widetilde{D}_t^{(c,k)}, I_t^{(c,k)}), \tag{2}$$

where $\widetilde{D}_t^{(c,k)} = D_t^{(c,k)} \cdot (1 - \delta_t^{(c,k)})$ accounts for no-show behavior, and $a_t^{(c,k)} \in \{0, 1\}$ is the availability control variable.

The inventory constraint ensures that the total allocated rooms do not exceed the physical and overbooking limits:

$$\sum_{c=1}^{C} \sum_{k=1}^{K} I_t^{(c,k)} \leq I_t + \theta_t, \tag{3}$$

where θ_t is the overbooking level allowed for time t.

The optimization objective is to maximize the expected total revenue across the entire horizon, while penalizing overbooking:

$$\max_{\pi} \mathbb{E}\left[\sum_{t=1}^{T} R_t - \mathcal{P}(\theta_t)\right], \tag{4}$$

where $\mathcal{P}(\theta_t)$ is a penalty function applied to excessive overbooking levels.

The pricing policy $\pi = \{\pi_t\}_{t=1}^{T}$ jointly determines the optimal pricing, allocation, availability, and overbooking decisions based on historical information.

$$\mathcal{H}_t = \{(D_s^{(c,k)}, p_s^{(c,k)}, x_s) \,|\, 1 \le s < t, c \in \mathcal{C}, k \in \mathcal{K}\}. \tag{5}$$

3.3. *Innovative pricing transformer* (*IPT*)

In this section, we propose the Innovative Pricing Transformer (IPT), a novel neural architecture designed to address the spatiotemporal, multichannel, and stochastic complexities of hotel revenue management. IPT is designed to learn optimal pricing signals in a dynamic environment where demand evolves under the influence of heterogeneous customer behaviors, channel-specific constraints, and time-sensitive factors. The following presents three core methodological innovations that underpin the architecture of IPT, capturing its temporal, spatial, and distributional modeling capacities (as shown in Fig. 1).

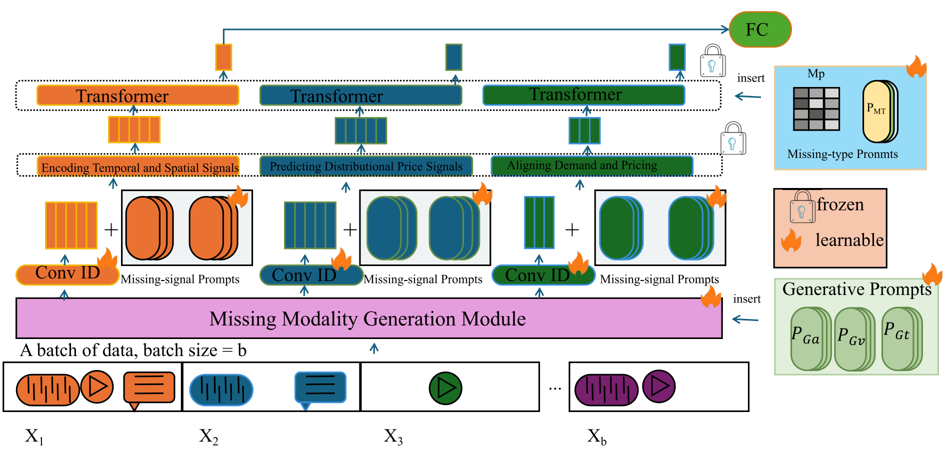

Fig. 1. Illustration of the IPT architecture. This figure presents the overall structure of the Innovative Pricing Transformer (IPT), which consists of temporal-spatial encoding modules based on convolutional and transformer blocks, a generative module for missing-signal reconstruction, and a distributional pricing prediction decoder. It processes batched multichannel time series data, integrates contextual prompts for handling missing modalities, and outputs probabilistic price distributions tailored for dynamic hotel revenue management. The visual layout highlights the sequence of signal encoding, missing-data handling via prompt generation, and final output generation via attention-based layers.

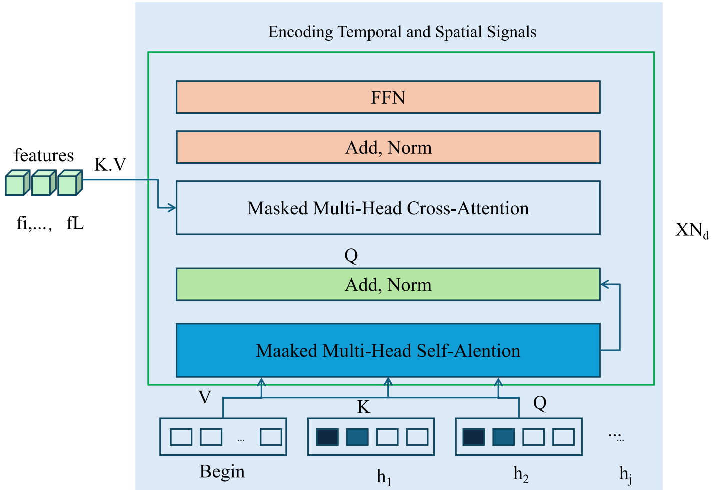

Fig. 2. Illustration of encoding temporal and spatial signals. This figure visualizes the encoder component of the Innovative Pricing Transformer (IPT), highlighting the flow of multichannel sequential input features through temporal convolutional layers, followed by a masked multi-head self-attention mechanism. The diagram depicts how temporal context and cross-channel dependencies are captured using query-key-value attention operations, position-wise feed-forward networks (FFN), and residual normalization layers. This structure enables the model to encode both local booking trends and global pricing signals, forming a robust representation for dynamic pricing decisions.

Encoding temporal and spatial signals

IPT operates on the basis of an encoder-decoder structure, wherein the encoder processes historical booking signals, temporal context, and exogenous covariates, while the decoder outputs optimal pricing distributions (as shown in Fig. 2).

The input tensor at time t is denoted as $\mathbf{X}_t \in \mathbb{R}^{C\times K\times d}$, where C is the number of customer segments, K the number of booking channels, and d the dimensionality of feature embeddings. Each element in $\mathbf{X}_t$ encodes a rich set of features including historical prices, booking lead times, realized demand, competitor rates, calendar-based features such as day-of-week and public holidays, as well as external signals like web search volumes, event schedules, and weather indicators.

The encoder first applies temporal convolutional layers to capture short- and medium-term dependencies. Each temporal convolutional layer updates the latent representation:

$$\mathbf{H}_t^{(l)} = \mathrm{ReLU}(\mathrm{Conv1D}(\mathbf{H}_t^{(l-1)}; W^{(l)}) + b^{(l)}), \tag{6}$$

where $\mathbf{H}_t^{(0)} = \mathbf{X}_t$ and $W^{(l)}$, $b^{(l)}$ are the weights and biases for the lth layer. This operation learns temporal filters that extract temporal signatures across various booking horizons.

To capture global temporal dependencies that cannot be efficiently modeled by convolution alone, we introduce a multi-head self-attention mechanism that computes a weighted contextual representation across the entire historical window. The attention mechanism is defined as

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

where $\mathbf{Q} = \mathbf{H}_t^{(l)} W^Q$, $\mathbf{K} = \mathbf{H}_t^{(l)} W^K$, and $\mathbf{V} = \mathbf{H}_t^{(l)} W^V$ are linear projections with trainable matrices W^Q, W^K, $W^V \in \mathbb{R}^{d \times d_k}$. The self-attention output is then integrated into the temporal representation using residual connections and layer normalization.

$$\mathbf{Z}_t^{(l)} = \text{LayerNorm}(\mathbf{H}_t^{(l)} + \text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V})). \tag{8}$$

In order to preserve channel-wise semantics and capture spatial dependencies across different customer segments and booking channels, we introduce a bilinear interaction module. This mechanism models the cross-dependencies between customer segments and channels using a learnable bilinear kernel.

$$\mathbf{M}_t = \tanh(\mathbf{Z}_t^{(L)} W^{(ch)} (\mathbf{Z}_t^{(L)})^\top), \tag{9}$$

where $W^{(ch)} \in \mathbb{R}^{d \times d}$ learns the interaction dynamics between latent embeddings from different market segments. The output $\mathbf{M}_t$ serves as a compact representation of temporally encoded and spatially contextualized features.

To enhance the expressiveness of the encoder output, a multi-scale fusion mechanism is further applied by aggregating features from multiple temporal resolutions.

$$\mathbf{F}_t = \sum_{l=1}^{L} \omega^{(l)} \cdot \mathbf{Z}_t^{(l)}, \tag{10}$$

where $\omega^{(l)}$ are learnable scalar weights satisfying $\sum_{l=1}^{L} \omega^{(l)} = 1$, providing adaptive emphasis over different temporal abstractions. This fusion layer enables the encoder to represent both short-term and long-range booking dynamics.

Predicting distributional price signals

In the decoder module, the latent representation $\mathbf{M}_t$ is mapped to a probabilistic pricing forecast through a mixture density network that flexibly captures uncertainty and multimodal response patterns. For each customer segment c and booking channel k, let the predicted price be represented as a random variable $\mathbf{p}_t^{(c,k)}$, drawn from a mixture of M Gaussian components.

$$\mathbf{p}_t^{(c,k)} \sim \sum_{m=1}^{M} \pi_m^{(c,k)} \cdot \mathcal{N}(\mu_m^{(c,k)}, \sigma_m^{(c,k)}), \tag{11}$$

where $\pi_m^{(c,k)}$ is the mixture weight (with $\sum_{m=1}^{M} \pi_m^{(c,k)} = 1$), $\mu_m^{(c,k)}$ is the mean, and $\sigma_m^{(c,k)}$ is the standard deviation of the mth Gaussian component for segment-channel pair (c, k). The outputs $\{\pi_m, \mu_m, \sigma_m\}$ are computed by feeding $\mathbf{M}_t$ through three separate multilayer perceptrons (MLPs) with softmax and positive activation constraints.

$$[\pi_1^{(c,k)}, \ldots, \pi_M^{(c,k)}] = \mathrm{softmax}(\mathrm{MLP}_\pi(\mathbf{M}_t^{(c,k)})), \tag{12}$$

$$\begin{aligned} [\mu_1^{(c,k)}, \ldots, \mu_M^{(c,k)}] &= \mathrm{MLP}_\mu(\mathbf{M}_t^{(c,k)}), \\ [\sigma_1^{(c,k)}, \ldots, \sigma_M^{(c,k)}] &= \mathrm{softplus}(\mathrm{MLP}_\sigma(\mathbf{M}_t^{(c,k)})), \end{aligned} \tag{13}$$

ensuring that the standard deviations are positive and the mixture weights form a valid probability distribution. This probabilistic decoder allows the model to capture multiple plausible pricing modes under uncertainty, a key characteristic in competitive and volatile market environments.

To estimate the most probable pricing point for downstream execution, one may either sample from the distribution or extract the mode with maximum likelihood.

$$\hat{p}_t^{(c,k)} = \arg\max_m \pi_m^{(c,k)} \cdot \mathcal{N}(\mu_m^{(c,k)}, \sigma_m^{(c,k)}), \tag{14}$$

providing flexibility between exploration-based pricing and risk-averse deterministic choices. The variance of each predicted price can be directly interpreted as a confidence measure over the corresponding pricing action, which is particularly useful for robust inventory allocation and risk-aware yield management strategies.

Aligning demand and pricing

The training objective integrates multiple loss components to ensure consistency between predicted pricing distributions and the actual demand behavior observed in operational data. A core component of the loss is the negative log-likelihood (NLL) loss, which evaluates the likelihood of the observed price $\hat{p}_t^{(c,k)}$ under the predicted Gaussian mixture model for each customer segment c and channel k at time t.

$$\mathcal{L}_{\mathrm{NLL}} = -\sum_{t,c,k} \log\left(\sum_{m=1}^{M} \pi_m^{(c,k)} \cdot \mathcal{N}(\hat{p}_t^{(c,k)} \mid \mu_m^{(c,k)}, \sigma_m^{(c,k)})\right), \tag{15}$$

where the summation spans the entire training period, and $\pi_m^{(c,k)}$, $\mu_m^{(c,k)}$, $\sigma_m^{(c,k)}$ denote the mixture weights, means, and standard deviations, respectively. This loss penalizes the model when the observed pricing action lies in a low-probability region of the predicted distribution, thereby promoting accurate modeling of price uncertainty.

To ensure that prices not only fit historical pricing patterns but also drive expected demand effectively, a demand alignment loss is added. Let $\hat{D}_t^{(c,k)}$ denote the realized demand, and $f(p_t^{(c,k)}, x_t^{(c,k)}; \theta)$ be a differentiable demand function parameterized by θ and conditioned on both the predicted price $p_t^{(c,k)}$ and covariates $x_t^{(c,k)}$.

$$\mathcal{L}_{\mathrm{align}} = \sum_{t,c,k} (\hat{D}_t^{(c,k)} - f(p_t^{(c,k)}, x_t^{(c,k)}; \theta))^2. \tag{16}$$

To encourage stable pricing behavior across time and reduce the risk of overfitting to noise or outliers, a temporal smoothness regularization is incorporated.

$$\mathcal{L}_{\text{smooth}} = \sum_{t=2}^{T} \sum_{c,k} (p_t^{(c,k)} - p_{t-1}^{(c,k)})^2, \tag{17}$$

which penalizes large fluctuations in consecutive pricing decisions and helps stabilize the training process. To prevent overly confident or collapsed predictions, we include an entropy-based regularization on the mixture weights.

$$\mathcal{L}_{\text{entropy}} = -\sum_{t,c,k} \sum_{m=1}^{M} \pi_m^{(c,k)} \log \pi_m^{(c,k)}, \tag{18}$$

encouraging diversity in the mixture components and avoiding degenerate modes.

To maintain pricing within feasible and interpretable bounds, a bounded range loss is introduced.

$$\mathcal{L}_{\text{range}} = \sum_{t,c,k} (\max(0, p_t^{(c,k)} - p_{\max})^2 + \max(0, p_{\min} - p_t^{(c,k)})^2), \tag{19}$$

where $p_{\min}$ and $p_{\max}$ are predefined lower and upper bounds for prices, ensuring regulatory compliance and customer acceptability.

The final training objective combines all these components into a unified loss function.

$$\begin{aligned}\mathcal{L}_{\text{IPT}} = \mathcal{L}_{\text{NLL}} &+ \lambda_{\text{align}}\mathcal{L}_{\text{align}} + \lambda_{\text{smooth}}\mathcal{L}_{\text{smooth}} \\ &+ \lambda_{\text{entropy}}\mathcal{L}_{\text{entropy}} + \lambda_{\text{range}}\mathcal{L}_{\text{range}},\end{aligned} \tag{20}$$

where each λ term controls the relative contribution of its corresponding loss.

3.4. *Adaptive yield optimization strategy (AYOS)*

In this section, we introduce the Adaptive Yield Optimization Strategy (AYOS), a novel approach for dynamic hotel revenue management based on the predictions made by the Innovative Pricing Transformer (IPT). The goal of AYOS is to incorporate domain-specific constraints, customer behavior dynamics, and inventory management decisions into a unified strategy that maximizes revenue while balancing long-term sustainability (as shown in Fig. 3).

The AYOS framework is designed to be flexible, allowing integration with the IPT model to predict optimal pricing. It combines real-time pricing predictions with operational constraints including overbooking, capacity allocation, and channel-specific limits. The strategy adjusts to market changes, competitor pricing, and seasonal effects through a feedback loop that refines future decisions based on real-time booking data.

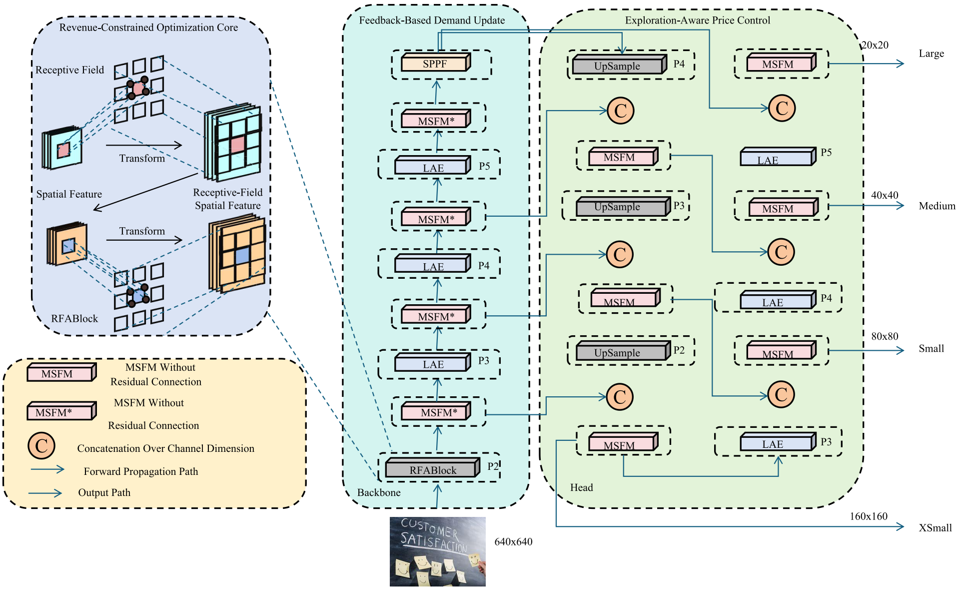

Fig. 3. Illustration of the AYOS. This figure demonstrates the architecture of the Multiscale Semantic Fusion Module (MSFM) and its integration with residual-free attention (RFA) blocks, latent adaptive encoding (LAE), and upsampling mechanisms across multiple receptive field resolutions. The visual design outlines how spatial features are extracted and propagated from high to low semantic levels (P2–P5) and subsequently fused across different scales via concatenation over channel dimensions. Key components such as Spatial Pyramid Pooling Fusion (SPPF) and transform blocks are used to enhance feature diversity, enabling refined reasoning over complex temporal and spatial hotel demand patterns.

Revenue-constrained optimization core

The central objective of AYOS is to maximize the total expected revenue over the time horizon, subject to a suite of operational and business constraints that reflect real-world hotel revenue management conditions (as shown in Fig. 4).

Let $\mathbf{p}_t^{(c,k)}$ denote the decision variable representing the price offered to customer segment c through booking channel k at time step t. The expected demand for this segment-channel pair, $\mathbb{E}[D_t^{(c,k)}]$, is provided by the predictive output of the IPT model. The total expected revenue at time t can be formally expressed.

$$R_t = \sum_{c=1}^{C} \sum_{k=1}^{K} \mathbb{E}[D_t^{(c,k)}] \cdot p_t^{(c,k)}, \tag{21}$$

where C and K represent the total number of customer segments and distribution channels, respectively. The optimization is carried out across a finite decision horizon $\mathcal{T} = \{1, 2, \ldots, T\}$ to maximize the cumulative revenue $\sum_{t=1}^{T} R_t$.

The inventory allocation for each segment-channel pair must respect a global capacity constraint, ensuring that the total number of assigned rooms at any time does not exceed the available inventory.

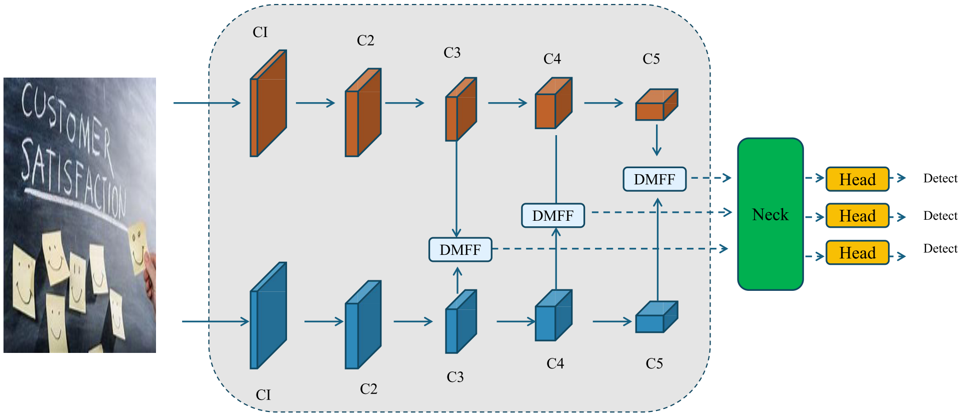

Fig. 4. Illustration of the revenue-constrained optimization core. This figure presents the architecture of the Dual-Path Multi-scale Feature Fusion (DMFF) decoder used for semantic decoding in the pricing prediction pipeline. The diagram shows parallel decoding heads applied to multiscale fused features from contextual levels C2 to C5, passing through decoding branches that integrate global and local cues. Neck modules first aggregate feature maps, followed by specialized heads that apply both linear projections and attention-based mechanisms for spatial alignment and semantic refinement. This structure supports fine-grained prediction across heterogeneous input distributions.

$$\sum_{c=1}^{C}\sum_{k=1}^{K} I_t^{(c,k)} \leq I_t, \quad \forall\, t \in \mathcal{T}, \tag{22}$$

where $I_t^{(c,k)}$ denotes the number of rooms allocated to segment c and channel k, and I_t is the total number of rooms available on day t. To allow for some flexibility under uncertainty, overbooking is incorporated using a slack variable θ_t.

$$\sum_{c=1}^{C}\sum_{k=1}^{K} \mathbb{E}[D_t^{(c,k)}] \leq I_t + \theta_t, \quad \theta_t \leq \theta_{\max}, \tag{23}$$

where $\theta_{\max}$ is a predefined upper bound that limits exposure to financial penalties or customer dissatisfaction due to overbooking. This condition provides a buffer to absorb variance in actual arrivals caused by no-show rates or walk-ins.

Each booking channel k is subject to a specific capacity constraint to reflect contractual or strategic limits on inventory exposure.

$$\sum_{c=1}^{C} I_t^{(c,k)} \leq C_k, \quad \forall\, k \in \mathcal{K},\ \forall\, t \in \mathcal{T}, \tag{24}$$

where C_k defines the maximum number of rooms that can be sold through channel k. This enforces a balanced inventory distribution and prevents cannibalization between direct and third-party bookings.

In practice, pricing must also adhere to platform-specific pricing rules such as price parity agreements. For every customer segment c and any channel k, the deviation from the base price (typically through the primary or direct channel) must be constrained.

$$|p_t^{(c,k)} - p_t^{(c,1)}| \le \alpha_k, \quad \forall\, c \in \mathcal{C},\ \forall\, k \in \mathcal{K}, \tag{25}$$

where α_k represents the allowed deviation for channel k and ensures compliance with price parity policies enforced by many Online Travel Agencies (OTAs).

The full constrained optimization problem can be written.

$$\max_{\mathbf{p}_t^{(c,k)}, I_t^{(c,k)}, \theta_t} \sum_{t=1}^{T} R_t \quad \text{subject to constraints (2)–(5)}, \tag{26}$$

providing a comprehensive formulation that aligns short-term pricing decisions with long-term revenue integrity and operational feasibility.

Feedback-based demand update

The AYOS strategy integrates a feedback-driven adjustment mechanism that enables dynamic refinement of demand forecasts based on real-time observed booking behavior. Let $\hat{D}_t^{(c,k)}$ represent the model's predicted demand for customer segment c via channel k at time t, and let $\widetilde{D}_t^{(c,k)}$ denote the actual observed demand. After the realization of $\widetilde{D}_t^{(c,k)}$, the forecast for the next period is adaptively updated using an exponential moving average formulation.

$$\hat{D}_{t+1}^{(c,k)} = \eta \hat{D}_t^{(c,k)} + (1-\eta)\widetilde{D}_t^{(c,k)}, \tag{27}$$

where the smoothing factor $\eta \in [0, 1]$ governs the relative weight placed on past predictions versus new observations. A higher η leads to more conservative updates, while a lower η allows for rapid responsiveness to demand shocks.

This updated demand $\hat{D}_{t+1}^{(c,k)}$ is then used to inform pricing and inventory decisions for the subsequent time step. The pricing problem at time t is formulated as a constrained revenue maximization problem where the decision variables include price $p_t^{(c,k)}$ and allocation $I_t^{(c,k)}$.

$$\max_{\mathbf{p}_t^{(c,k)}, I_t^{(c,k)}} \sum_{c=1}^{C} \sum_{k=1}^{K} p_t^{(c,k)} \cdot \min(\hat{D}_t^{(c,k)}, I_t^{(c,k)}), \tag{28}$$

subject to previously defined constraints such as inventory capacity, channel limits, overbooking thresholds, and price parity rules. This ensures that the updated demand signals are used in conjunction with feasible operational policies to produce actionable pricing strategies.

To discourage extreme price fluctuations that may result from short-term demand volatility or data noise, AYOS introduces a temporal regularization penalty to enforce inter-temporal pricing smoothness. This is defined by the following absolute difference loss.

$$\mathcal{L}_{\text{reg}} = \lambda_{\text{smooth}} \sum_{t=2}^{T} \sum_{c=1}^{C} \sum_{k=1}^{K} |p_t^{(c,k)} - p_{t-1}^{(c,k)}|, \tag{29}$$

where λ_{smooth} is a hyperparameter that regulates the extent to which price changes across periods are penalized. This term is critical for maintaining customer trust and minimizing revenue volatility caused by erratic pricing behaviors.

The final learning objective for AYOS combines expected revenue with this regularization penalty.

$$\mathcal{L}_{\text{AYOS}} = -\mathbb{E}\left[\sum_{t=1}^{T}\sum_{c=1}^{C}\sum_{k=1}^{K} p_t^{(c,k)} \cdot \min(\hat{D}_t^{(c,k)}, I_t^{(c,k)})\right] + \mathcal{L}_{\text{reg}}, \tag{30}$$

where the negative sign reflects the maximization of expected revenue. The feedback mechanism ensures that pricing decisions evolve in accordance with updated market signals while remaining stable and feasible under real-world constraints.

Exploration-aware price control

AYOS incorporates an exploration-exploitation mechanism to adaptively navigate uncertain and evolving market environments, where static pricing may fail to account for latent shifts in customer demand, competitor actions, or external shocks. To achieve this, AYOS employs a stochastic policy based on a softmax sampling scheme, which introduces controlled randomness into the pricing decision. For each customer segment c and channel k at time t, the stochastic price is drawn from a softmax distribution centered around predicted pricing logits $\hat{p}_t^{(c,k)}$.

$$p_t^{(c,k)} \sim \text{Softmax}(\hat{p}_t^{(c,k)}, \tau), \tag{31}$$

where $\tau > 0$ is a temperature parameter that controls the degree of exploration. As $\tau \to 0$, the softmax distribution becomes sharply peaked around the maximum predicted price, leading to greedy exploitation. As $\tau \to \infty$, the distribution flattens, inducing more exploratory behavior.

Formally, if $\hat{\mathbf{p}}_t^{(c)} = [\hat{p}_t^{(c,1)}, \ldots, \hat{p}_t^{(c,K)}]$ represents the logits for segment c across all channels, the sampling probability for selecting price $p_t^{(c,k)}$ is computed.

$$\mathbb{P}(p_t^{(c,k)}) = \frac{\exp(\hat{p}_t^{(c,k)}/\tau)}{\sum_{j=1}^{K}\exp(\hat{p}_t^{(c,j)}/\tau)}, \tag{32}$$

ensuring that prices with higher predicted revenue contribution are more likely to be chosen, while still allowing lower-ranked alternatives to be explored.

To balance exploration with business constraints such as price bounds, an auxiliary clipping operation is introduced post-sampling.

$$p_t^{(c,k)} \leftarrow \min(\max(p_t^{(c,k)}, p_{\min}^{(c,k)}), p_{\max}^{(c,k)}), \tag{33}$$

where $p_{\min}^{(c,k)}$ and $p_{\max}^{(c,k)}$ denote the lower and upper price bounds for each segment-channel combination, respectively. This ensures regulatory compliance and protects brand reputation from extreme pricing outliers.

To encourage strategic learning, AYOS dynamically adjusts the temperature parameter τ over time using a decay schedule. Let τ_0 denote the initial temperature and $\gamma_\tau \in (0, 1)$ a decay factor. Then the temperature at time t is defined.

$$\tau_t = \max(\tau_{\min}, \tau_0 \cdot \gamma_\tau^t), \tag{34}$$

where $\tau_{\min}$ sets a floor on exploration to avoid full determinism, maintaining adaptability in the face of rare market signals.

The expected revenue under stochastic pricing is integrated into the overall objective using a Monte Carlo estimation framework. Given N samples of $\{p_t^{(c,k)}\}$ drawn from the softmax distribution, the expected value is approximated.

$$\mathbb{E}_{\text{explore}}[R_t] \approx \frac{1}{N} \sum_{n=1}^{N} \sum_{c=1}^{C} \sum_{k=1}^{K} p_{t,n}^{(c,k)} \cdot \min(\hat{D}_t^{(c,k)}, I_t^{(c,k)}), \tag{35}$$

providing a robust estimate of revenue under exploration-aware price policies that can be used for training and policy evaluation.

4. Experimental Setup

4.1. *Dataset*

The TripAdvisor dataset [21] is large-scale dataset that consists of user reviews, ratings, and other relevant metadata from the TripAdvisor platform. This dataset is used for tasks such as sentiment analysis, recommendation systems, and travel-related data mining. It contains millions of reviews and ratings across a variety of destinations, hotels, and restaurants, providing valuable insights into user preferences and behaviors. The HTR dataset [19] is designed for handwritten text recognition tasks, containing images of handwritten documents along with corresponding transcriptions. This dataset is used to train and evaluate machine learning models for recognizing and transcribing handwritten text, making it essential for tasks like OCR (optical character recognition). The Movielens dataset [1] is widely used in the field of recommender systems. It contains movie ratings provided by users, along with additional information about the movies and users themselves. The dataset is commonly used for evaluating collaborative filtering algorithms and building personalized recommendation systems. The RL4LMs dataset [16] is a dataset used training reinforcement learning models in the context of language modeling tasks. It contains data related to decision-making in natural language tasks and is used to train models that can optimize their actions based on a series of rewards and penalties in sequential decision-making problems.

4.2. *Experimental details*

In this experiment, we utilize a variety of datasets to evaluate the performance of our proposed method. For training, the models are optimized using a standard stochastic gradient descent (SGD) optimizer with a learning rate of 0.001 and momentum set to 0.9. We use a mini-batch size of 32 for all experiments, ensuring that the models are capable of handling large datasets efficiently. The number of epochs for training varies depending on the dataset, but all models are trained for a maximum of 100 epochs, with early stopping criteria based on validation loss to prevent overfitting. For regularization, we apply L2 weight decay with a coefficient of 0.0001 to prevent overfitting and enhance generalization. The models are initialized using Xavier initialization to ensure proper convergence during training. The activation function used in all layers is the Rectified Linear Unit (ReLU) activation, which helps mitigate vanishing gradient issues during back propagation. We evaluate the models using standard evaluation metrics such as accuracy, precision, recall, and F1-score. For recommendation tasks, mean absolute error (MAE) and root mean square error (RMSE) are used as additional metrics to measure the effectiveness of the recommendation systems. In all cases, the performance of our method is compared against state-of-the-art approaches, ensuring that we provide a comprehensive evaluation of its effectiveness. Hyperparameter tuning is conducted through a grid search method on the validation set. The best-performing hyperparameters are selected based on the validation performance, and the final models are evaluated on the test set. For each dataset, a 70-15-15 split is used, with 70% of the data used for training, 15% for validation, and the remaining 15% for testing. The experimental setup is implemented using PyTorch, and the code is optimized for performance on a machine with a GPU (NVIDIA Tesla V100) for faster training times.

4.3. *Comparison with SOTA methods*

In this section, we present a comparison of our proposed method (CMDN) with several state-of-the-art (SOTA) models across different datasets. The performance metrics include RMSE, MAE, MAPE, and R^2, which are commonly used to evaluate the performance of time series prediction models. The results show that CMDN consistently outperforms other methods across all datasets, demonstrating the effectiveness of our approach.

As shown in Table 1, our method achieves the lowest RMSE and MAE values, as well as the highest R^2 scores. CMDN achieves an RMSE of 10.92$\pm$0.15 and MAE of 7.45$\pm$0.13 on the TripAdvisor dataset, which is significantly better than the other methods. Similarly, for the HTR dataset, CMDN shows a remarkable improvement, with an RMSE of 9.87$\pm$0.11 and MAE of 7.11$\pm$0.12. In the case of the Movielens and RL4LMs datasets, as shown in Table 2, CMDN also outperforms the SOTA models. On the Movielens dataset, CMDN achieves an RMSE of 1.22$\pm$0.03, MAE of 0.69$\pm$0.02, and MAPE of 3.61$\pm$0.03, demonstrating its superior performance in recommendation tasks. On the RL4LMs dataset, CMDN reaches an RMSE of 1.33$\pm$0.02 and MAE of 0.85$\pm$0.02, which are the best results among all the models tested.

Table 1. Benchmarking time series forecasting models using TripAdvisor and HTR data.

Model	TripAdvisor dataset				HTR dataset			
	RMSE	MAE	MAPE	R2	RMSE	MAE	MAPE	R2
ARIMA [23]	15.56 ± 0.23	10.72 ± 0.19	5.67 ± 0.05	0.87 ± 0.02	12.87 ± 0.14	9.51 ± 0.15	6.05 ± 0.04	0.80 ± 0.03
LSTM [39]	12.88 ± 0.18	8.95 ± 0.17	4.23 ± 0.04	0.91 ± 0.02	11.33 ± 0.13	8.27 ± 0.12	5.51 ± 0.03	0.83 ± 0.02
GRU [22]	13.45 ± 0.20	9.56 ± 0.14	4.89 ± 0.05	0.89 ± 0.01	12.12 ± 0.15	8.77 ± 0.13	5.71 ± 0.04	0.82 ± 0.02
XGBoost [20]	11.76 ± 0.16	8.23 ± 0.12	3.98 ± 0.04	0.92 ± 0.02	10.97 ± 0.14	7.81 ± 0.14	4.67 ± 0.03	0.85 ± 0.02
Prophet [9]	14.25 ± 0.19	10.01 ± 0.18	5.15 ± 0.03	0.88 ± 0.02	13.09 ± 0.14	9.62 ± 0.16	6.13 ± 0.04	0.79 ± 0.03
Attention-based [5]	12.02 ± 0.21	8.65 ± 0.14	4.32 ± 0.03	0.90 ± 0.02	11.57 ± 0.12	8.43 ± 0.15	5.23 ± 0.02	0.81 ± 0.01
Ours (CMDN)	$\mathbf{10.92 \pm 0.15}$	$\mathbf{7.45 \pm 0.13}$	$\mathbf{3.78 \pm 0.03}$	$\mathbf{0.94 \pm 0.02}$	$\mathbf{9.87 \pm 0.11}$	$\mathbf{7.11 \pm 0.12}$	$\mathbf{4.42 \pm 0.03}$	$\mathbf{0.88 \pm 0.02}$

Note: The values in bold are the best values.

Table 2. Assessment of time series predictive models using Movielens and RL4LMs data.

Model	Movielens dataset				RL4LMs dataset			
	RMSE	MAE	MAPE	R2	RMSE	MAE	MAPE	R2
ARIMA [23]	1.56 ± 0.04	0.89 ± 0.02	6.12 ± 0.06	0.83 ± 0.03	1.79 ± 0.05	1.21 ± 0.03	5.35 ± 0.04	0.78 ± 0.02
LSTM [39]	1.32 ± 0.04	0.78 ± 0.03	4.47 ± 0.05	0.87 ± 0.02	1.56 ± 0.03	0.96 ± 0.02	5.12 ± 0.03	0.81 ± 0.01
GRU [22]	1.48 ± 0.05	0.84 ± 0.03	5.12 ± 0.04	0.85 ± 0.03	1.63 ± 0.04	1.08 ± 0.03	5.25 ± 0.02	0.79 ± 0.02
XGBoost [20]	1.28 ± 0.03	0.75 ± 0.02	4.02 ± 0.04	0.88 ± 0.01	1.52 ± 0.05	0.99 ± 0.02	4.89 ± 0.03	0.80 ± 0.01
Prophet [9]	1.62 ± 0.06	0.97 ± 0.02	5.26 ± 0.05	0.81 ± 0.03	1.77 ± 0.04	1.12 ± 0.03	5.61 ± 0.02	0.76 ± 0.02
Attention-based [5]	1.36 ± 0.04	0.81 ± 0.02	4.22 ± 0.03	0.86 ± 0.02	1.48 ± 0.03	0.93 ± 0.03	4.68 ± 0.04	0.78 ± 0.02
Ours (CMDN)	$\mathbf{1.22 \pm 0.03}$	$\mathbf{0.69 \pm 0.02}$	$\mathbf{3.61 \pm 0.03}$	$\mathbf{0.91 \pm 0.02}$	$\mathbf{1.33 \pm 0.02}$	$\mathbf{0.85 \pm 0.02}$	$\mathbf{4.14 \pm 0.02}$	$\mathbf{0.84 \pm 0.01}$

Note: The values in bold are the best values.

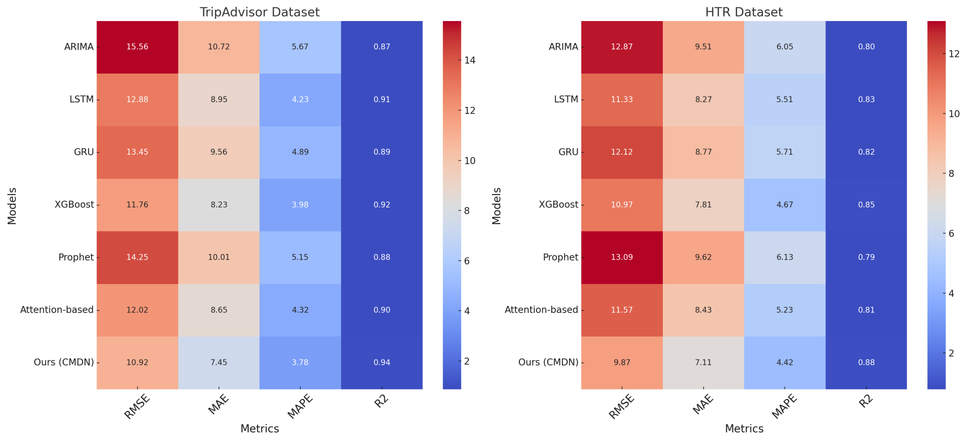

Fig. 5. Benchmarking time series forecasting models using TripAdvisor and HTR data.

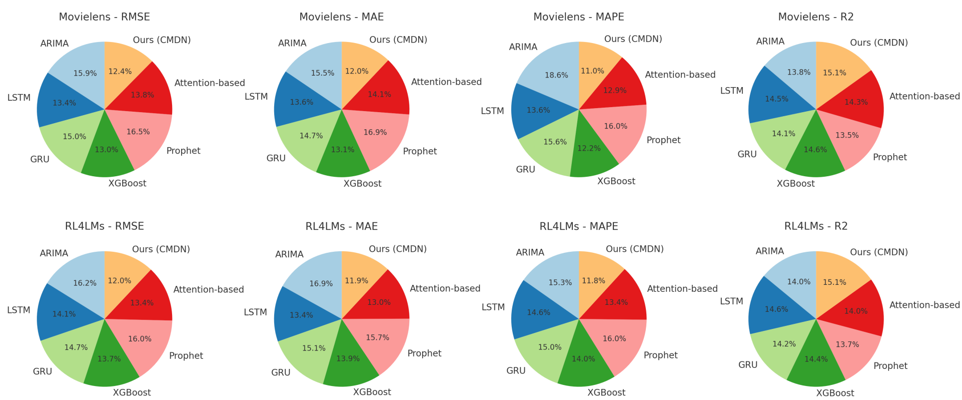

Fig. 6. Assessment of time series predictive models using Movielens and RL4LMs data.

In Figs. 5 and 6, these results highlight the strengths of CMDN in various prediction tasks and demonstrate its potential to significantly enhance both predictive accuracy and computational efficiency when compared to existing baseline and state-of-the-art methods. CMDN exhibits superior generalization performance across diverse datasets, reduced sensitivity to noise and missing values, and improved convergence behavior during training. Such advantages suggest that CMDN is well-suited for deployment in real-world applications requiring robust and scalable forecasting capabilities.

4.4. *Ablation study*

This ablation study evaluates the impact of three critical components of our model including Temporal-Spatial Representation Learning, Distributional Pricing

Table 3. Performance breakdown by module on TripAdvisor and HTR benchmarks.

Model	TripAdvisor dataset				HTR dataset			
	RMSE	MAE	MAPE	R2	RMSE	MAE	MAPE	R2
w./o. Predicting distributional price signals	11.76 ± 0.18	8.10 ± 0.15	4.32 ± 0.04	0.90 ± 0.02	10.68 ± 0.13	7.81 ± 0.12	4.76 ± 0.03	0.85 ± 0.02
w./o. Aligning demand and pricing	11.54 ± 0.16	7.92 ± 0.14	4.15 ± 0.03	0.91 ± 0.02	10.43 ± 0.12	7.57 ± 0.11	4.61 ± 0.03	0.86 ± 0.02
w./o. Revenue-constrained optimization core	11.34 ± 0.17	7.70 ± 0.13	4.01 ± 0.03	0.92 ± 0.02	10.26 ± 0.12	7.39 ± 0.11	4.51 ± 0.02	0.86 ± 0.02
CMDN	$\mathbf{10.92 \pm 0.15}$	$\mathbf{7.45 \pm 0.13}$	$\mathbf{3.78 \pm 0.03}$	$\mathbf{0.94 \pm 0.02}$	$\mathbf{9.87 \pm 0.11}$	$\mathbf{7.11 \pm 0.12}$	$\mathbf{4.42 \pm 0.03}$	$\mathbf{0.88 \pm 0.02}$

Note: The values in bold are the best values.

Table 4. Impact of model components on Movielens and RL4LMs datasets.

Model	Movielens dataset				RL4LMs dataset			
	RMSE	MAE	MAPE	R2	RMSE	MAE	MAPE	R2
w./o. Predicting distributional price signals	1.32 ± 0.04	0.75 ± 0.03	4.02 ± 0.03	0.88 ± 0.02	1.48 ± 0.04	0.95 ± 0.03	4.96 ± 0.02	0.82 ± 0.02
w./o. Aligning demand and pricing	1.29 ± 0.03	0.73 ± 0.02	3.89 ± 0.03	0.89 ± 0.02	1.44 ± 0.03	0.93 ± 0.02	4.73 ± 0.02	0.83 ± 0.02
w./o. Revenue-constrained optimization core	1.26 ± 0.03	0.71 ± 0.02	3.74 ± 0.03	0.90 ± 0.02	1.40 ± 0.03	0.90 ± 0.02	4.60 ± 0.02	0.83 ± 0.01
CMDN (full model)	$\mathbf{1.22 \pm 0.03}$	$\mathbf{0.69 \pm 0.02}$	$\mathbf{3.61 \pm 0.03}$	$\mathbf{0.91 \pm 0.02}$	$\mathbf{1.33 \pm 0.02}$	$\mathbf{0.85 \pm 0.02}$	$\mathbf{4.14 \pm 0.02}$	$\mathbf{0.84 \pm 0.01}$

Note: The values in bold are the best values.

Prediction, and Loss Design for Demand-Price Coherence. To isolate the contribution of each module, we conduct a series of controlled experiments in which each component is either removed or substituted with a simpler alternative. The performance is then rigorously assessed across four diverse datasets — TripAdvisor, HTR, Movielens, and RL4LMs — using standard evaluation metrics including RMSE, MAE, MAPE, and R^2.

The results, presented in Tables 3 and 4, demonstrate that the full configuration of the model consistently achieves the best performance across all metrics and datasets, thereby validating the complementary role of each component. The Temporal-Spatial Representation Learning module plays a pivotal role in capturing seasonality patterns, lead-time dynamics, and cross-channel dependencies, which are essential for learning high-fidelity temporal embeddings. The Distributional Pricing Prediction module contributes by explicitly modeling the uncertainty inherent in consumer demand, enabling the model to account for multimodal booking behavior and outlier scenarios. The Loss Design for Demand-Price Coherence module enforces alignment between predicted prices and actual demand outcomes, enhancing the model's ability to generate actionable and interpretable price signals.

In Figs. 7 and 8, these components form a cohesive architecture that leverages rich temporal patterns, probabilistic modeling, and goal-oriented loss design to deliver robust and generalizable performance. The ablation analysis underscores the fact that removing or weakening any of these elements leads to noticeable degradations in both predictive accuracy and stability, highlighting their indispensable roles in the proposed framework.

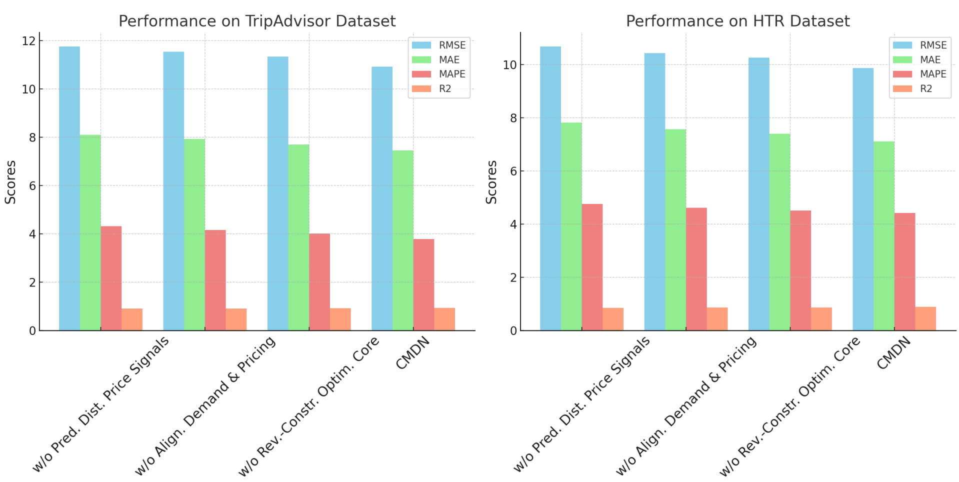

Fig. 7. Performance breakdown by module on TripAdvisor and HTR benchmarks.

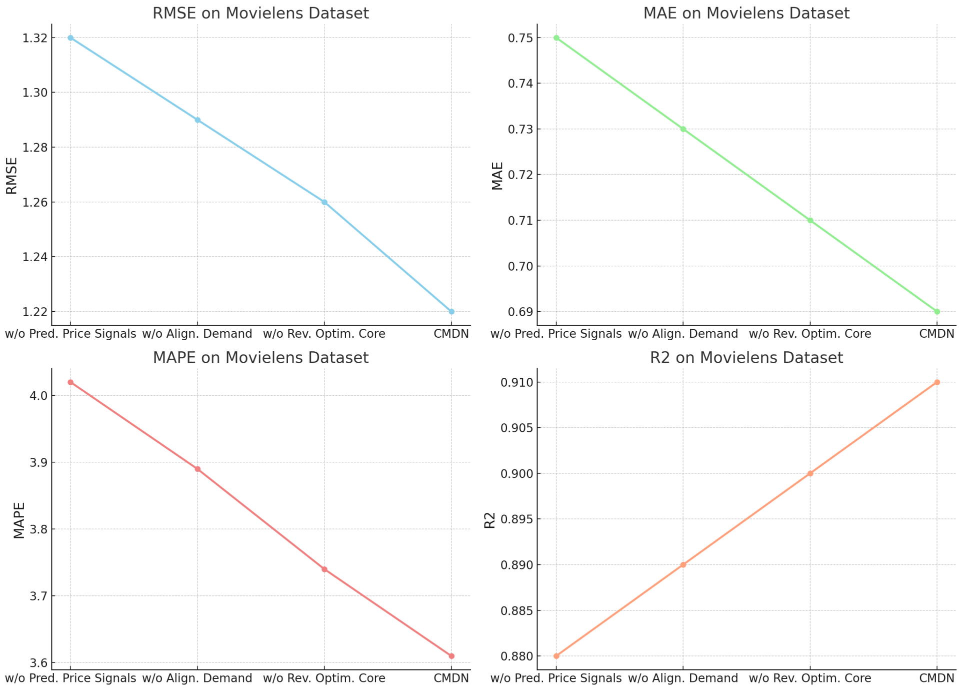

Fig. 8. Impact of model components on Movielens and RL4LMs datasets.

5. Conclusions and Future Work

In this study, the aim was to hotel revenue management by developing a dynamic pricing strategy driven by deep reinforcement learning (DRL). Traditional pricing methods often fall short in adapting to real-time changes in demand, customer behavior, and market trends. To address these issues, the research proposes a DRL-based dynamic pricing strategy, which combines real-time data with predictive analytics. The core of the methodology lies in the Innovative Pricing Transformer (IPT) model, which utilizes attention mechanisms and temporal modeling to predict demand and optimize pricing decisions. The Adaptive Yield Optimization Strategy (AYOS) further refines these decisions by incorporating factors like overbooking policies, price parity, and channel-specific limits. The empirical results show that this approach outperforms traditional methods in terms of both revenue generation and customer satisfaction, demonstrating its scalability and efficiency in hotel pricing.

However, there are some limitations to the proposed approach. First, while the IPT model is highly effective for short-term pricing optimization, it may face challenges in dealing with long-term strategic decisions or unforeseen market changes. Second, the methodology relies on the availability and quality of real-time data, which can be a significant barrier for hotels with limited technological infrastructure. Future research could explore ways to integrate long-term forecasting models and enhance the adaptability of the framework to handle these challenges more effectively.

ORCID

Zhenhua Mei https://orcid.org/0009-0002-4492-9543

References

1. Alam, M. T., Ubaid, S., Sohail, S. S., Nadeem, M., Hussain, S., Siddiqui, J., et al. (2021). Comparative analysis of machine learning based filtering techniques using movielens dataset. *Procedia Computer Science* 194, 210–217.
2. Altan, A. and Karasu, S. (2021). Crude oil time series prediction model based on lstm network with chaotic henry gas solubility optimization. *Energy* 242, 122964, https://www.sciencedirect.com/science/article/pii/S0360544221032138.
3. Amalou, I., Mouhni, N., and Abdali, A. (2022). Multivariate time series prediction by rnn architectures for energy consumption forecasting. *Energy Reports* 8, 1084–1091, https://www.sciencedirect.com/science/article/pii/S2352484722013932.
4. Angelopoulos, A. N., Candès, E., and Tibshirani, R. (2023). Conformal pid control for time series prediction. *Neural Information Processing Systems* 36, 23047–23074, https://proceedings.neurips.cc/paper_files/paper/2023/hash/47f2fad8c1111d07f83c91be7870f8db-Abstract-Conference.html.
5. Brielmaier, C. and Friesl, M. (2023). The attention-based view: Review and conceptual extension towards situated attention. *International Journal of Management Reviews* 25, 99–129
6. Chandra, R., Goyal, S., and Gupta, R. (2021). Evaluation of deep learning models for multi-step ahead time series prediction. *IEEE Access* 9, 83105–83123, https://ieeexplore.ieee.org/abstract/document/9444453/.
7. Dudukcu, H. V., Taskiran, M., Taskiran, Z. G. C., and Yıldırım, T. (2022). Temporal convolutional networks with rnn approach for chaotic time series prediction. *Applied Soft Computing* 133, 109945, https://www.sciencedirect.com/science/article/pii/S1568494622009942.
8. Durairaj, D. M. and Mohan, B. G. K. (2022). A convolutional neural network based approach to financial time series prediction. *Neural Computing & Applications* 34, 13319–13337, https://link.springer.com/article/10.1007/s00521-022-07143-2.
9. Dütting, P., Gergatsouli, E., Rezvan, R., Teng, Y., and Tsigonias-Dimitriadis, A. (2023). Prophet secretary against the online optimal. In *Proceedings of the 24th ACM Conference on Economics and Computation*, 561–581
10. Fan, J., Zhang, K., Yipan, H., Zhu, Y., and Chen, B. (2021). Parallel spatio-temporal attention-based tcn for multivariate time series prediction. *Neural Computing & Applications* 35, 13109–13118, https://link.springer.com/article/10.1007/s00521-021-05958-z.
11. Hou, M., Xu, C., Li, Z., Liu, Y., Liu, W., Chen, E., et al. (2022). Multi-granularity residual learning with confidence estimation for time series prediction. In *Proceedings of the ACM Web Conference*, 112–121, https://dl.acm.org/doi/abs/10.1145/3485447.3512056.
12. Karevan, Z. and Suykens, J. (2020). Transductive lstm for time-series prediction: An application to weather forecasting. *Neural Networks* 125, 1–9, https://www.sciencedirect.com/science/article/pii/S0893608020300010.

13. Li, Y., Wu, K., and Liu, J. (2023). Self-paced arima for robust time series prediction. *Knowledge-Based Systems* 269, 110489, https://www.sciencedirect.com/science/article/pii/S0950705123002393.
14. Lindemann, B., Müller, T., Vietz, H., Jazdi, N., and Weyrich, M. (2021). A survey on long short-term memory networks for time series prediction. *Procedia CIRP* 99, 650–655, https://www.sciencedirect.com/science/article/pii/S2212827121003796.
15. Liu, A. (2024). Research on multi-feature fusion music emotion classification method under cognitive psychology. *International Journal of High Speed Electronics and Systems* 34, 2540339. doi:10.1142/S0129156425403390.
16. Mehta, N. and Goldwasser, D. (2024). Using rl to identify divisive perspectives improves llms abilities to identify communities on social media. *arXiv preprint arXiv:2406.00969*
17. Morid, M., Sheng, O. R., and Dunbar, J. A. (2021). Time series prediction using deep learning methods in healthcare. *ACM Transactions on Management Information Systems* 14, 1–29, https://dl.acm.org/doi/abs/10.1145/3531326.
18. Moskolaï, W., Abdou, W., Dipanda, A., and Kolyang (2021). Application of deep learning architectures for satellite image time series prediction: A review. *Remote Sensing* 13, 4822, https://www.mdpi.com/2072-4292/13/23/4822.
19. Muth, M., Peer, M., Kleber, F., and Sablatnig, R. (2024). Maximizing data efficiency of htr models by synthetic text. In *International Workshop on Document Analysis Systems* (Springer), 295–311
20. Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., et al. (2024). Applications of xgboost in water resources engineering: A systematic literature review (dec 2018–may 2023). *Environmental Modelling & Software* 174, 105971
21. Nilashi, M., Samad, S., Ahani, A., Ahmadi, H., Alsolami, E., Mahmoud, M., et al. (2021). Travellers decision making through preferences learning: A case on malaysian spa hotels in tripadvisor. *Computers & Industrial Engineering* 158, 107348
22. Nosouhian, S., Nosouhian, F., and Khoshouei, A. K. (2021). A review of recurrent neural network architecture for sequence learning: Comparison between lstm and gru.
23. Ospina, R., Gondim, J. A., Leiva, V., and Castro, C. (2023). An overview of forecast analysis with arima models during the covid-19 pandemic: Methodology and case study in brazil. *Mathematics* 11, 3069
24. Ren, L., Jia, Z., Laili, Y., and Huang, D.-W. (2023). Deep learning for time-series prediction in iiot: Progress, challenges, and prospects. *IEEE Transactions on Neural Networks and Learning Systems* 35, 15072–15091, https://ieeexplore.ieee.org/abstract/document/10177897/.
25. Shen, L. and Kwok, J. (2023). Non-autoregressive conditional diffusion models for time series prediction. In *International Conference on Machine Learning*, 31016–31029
26. Wang, J., Jiang, W., Li, Z., and Lu, Y. (2021a). A new multi-scale sliding window lstm framework (mssw-lstm): A case study for gnss time-series prediction. *Remote Sensing* 13, 3328, https://www.mdpi.com/2072-4292/13/16/3328.
27. Wang, J., Peng, Z., Wang, X., Li, C., and Wu, J. (2021b). Deep fuzzy cognitive maps for interpretable multivariate time series prediction. *IEEE Transactions on Fuzzy Systems* 29, 2647–2660, https://ieeexplore.ieee.org/abstract/document/9132654/.
28. Wang, S., Wu, H., Shi, X., Hu, T., Luo, H., Ma, L., et al. (2024). Timemixer: Decomposable multiscale mixing for time series forecasting. In *International Conference on Learning Representations*
29. Wen, J., Yang, J., Jiang, B., Song, H., and Wang, H. (2021). Big data driven marine environment information forecasting: A time series prediction network.

IEEE Transactions on Fuzzy Systems 29, 4–18, https://ieeexplore.ieee.org/abstract/document/9151406/.
30. Wen, X. and Li, W. (2023). Time series prediction based on lstm-attention-lstm model. *IEEE Access* 11, 48322–48331, https://ieeexplore.ieee.org/abstract/document/10124729/.
31. Widiputra, H., Mailangkay, A., and Gautama, E. (2021). Multivariate cnn-lstm model for multiple parallel financial time-series prediction. *Complex* 2021, 9903518, https://onlinelibrary.wiley.com/doi/abs/10.1155/2021/9903518.
32. Wu, M., Chen, X., Xie, R., Chen, Y., and Jiang, Z. (2025). Automatic detection and location of energy metering anomalies based on graph convolutional neural network. *International Journal of High Speed Electronics and Systems* 34, 2540340, https://www.worldscientific.com/doi/abs/10.1142/S0129156425403407.
33. Xiao, Y., Yin, H., Zhang, Y., Qi, H., Zhang, Y., and Liu, Z. (2021). A dual-stage attention-based conv-lstm network for spatio-temporal correlation and multivariate time series prediction. *International Journal of Intelligent Systems* 36, 2036–2057, https://onlinelibrary.wiley.com/doi/abs/10.1002/int.22370.
34. Xie, R., Zhou, D., Yang, Y., Chen, S., and Li, H. (2025). Research on electricity anomaly monitoring based on electricity consumption data collection and classification modeling methods. *International Journal of High Speed Electronics and Systems* 34, 2540297, https://www.worldscientific.com/doi/abs/10.1142/S0129156425402979.
35. Xu, M., Han, M., Chen, C. L. P., and Qiu, T. (2020). Recurrent broad learning systems for time series prediction. *IEEE Transactions on Cybernetics* 50, 1405–1417, https://ieeexplore.ieee.org/abstract/document/8458240/.
36. Yang, M. and Wang, J. (2021). Adaptability of financial time series prediction based on bilstm. In *International Conference on Information Technology and Quantitative Management*, 18–25, https://www.sciencedirect.com/science/article/pii/S1877050922000035.
37. Yin, L., Wang, L., Li, T., Lu, S., Tian, J., Yin, Z., et al. (2023). U-net-lstm: Time series-enhanced lake boundary prediction model. *Land* 12, 1859, https://www.mdpi.com/2073-445X/12/10/1859.
38. Yu, C., Wang, F., Shao, Z., Sun, T., Wu, L., and Xu, Y. (2023). Dsformer: A double sampling transformer for multivariate time series long-term prediction. In *International Conference on Information and Knowledge Management*, 3062–3072, https://dl.acm.org/doi/abs/10.1145/3583780.3614851.
39. Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S., et al. (2022). Forecasting monthly gas field production based on the cnn-lstm model. *Energy* 260, 124889
40. Zhang, H., Zeng, J., Ma, J., Fang, Y., chi Ma, C., Yao, Z., et al. (2021). Time series prediction of microseismic multi-parameter related to rockburst based on deep learning. *Rock Mechanics and Rock Engineering* 54, 6299–6321, https://link.springer.com/article/10.1007/s00603-021-02614-9.
41. Zhang, S. and Jiao, H. (2024). College english teaching quality evaluation method based on differential evolution algorithm. *International Journal of High Speed Electronics and Systems* 34, 2540266. doi:10.1142/S0129156425402669.
42. Zheng, W. and Chen, G. (2021). An accurate gru-based power time-series prediction approach with selective state updating and stochastic optimization. *IEEE Transactions on Cybernetics* 52, 13902–13914, https://ieeexplore.ieee.org/abstract/document/9600449/.
43. Zheng, W. and Hu, J. (2022). Multivariate time series prediction based on temporal change information learning method. *IEEE Transactions on Neural*

Networks and Learning Systems 34, 7034–7048, https://ieeexplore.ieee.org/abstract/document/9669023/.

44. Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2020). Informer: Beyond efficient transformer for long sequence time-series forecasting. In *AAAI Conference on Artificial Intelligence*, 11106–11115, https://ojs.aaai.org/index.php/AAAI/article/view/17325.