Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Hybrid Demand Forecasting and Adaptive Optimization Approach for Dynamic Pricing in E-Commerce

Veerababu Reddy

rveerababu_vlits@vignan.ac.in

Vignan's Lara Institute of Technology & Science

Krishnaveni Kaki

Vignan's Lara Institute of Technology & Science

Harika Katepalli

Vignan's Lara Institute of Technology & Science

Pooja Govathoti

Vignan's Lara Institute of Technology & Science

Kiran Kumar Kondareddy

Vignan's Lara Institute of Technology & Science

Kiran Kumar Yendluri

Vignan's Lara Institute of Technology & Science

Research Article

Keywords: Dynamic pricing, Demand forecasting, Price elasticity, Revenue management, E-commerce, Machine learning

Posted Date: March 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9233004/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Hybrid Demand Forecasting and Adaptive Optimization Approach for Dynamic Pricing in E-Commerce

Abstract

Small and mid-size e-commerce sellers lose profit because automated pricing tools react to competitors without understanding customer demand. This paper proposes a pricing system that predicts demand, analyses price sensitivity, and recommends optimal prices in real time. Tested on real retail transaction data from the UCI Online Retail II dataset, the system achieves 11.8% mean profit improvement. Further experiments on 80 consumer electronics products show 27.1% profit improvement over standard pricing, with significantly more stable recommendations than existing approaches. The complete system delivers pricing decisions in under 120 milliseconds, making it practical for everyday seller use. Code and data are publicly available at https://doi.org/10.5281/zenodo.19151881.

Keywords: Dynamic pricing. Demand forecasting. Price elasticity. Revenue management. E-commerce. Machine learning.

Introduction

The rapid growth of e-commerce has created a revenue management crisis for small and medium sellers: price competition is so intense that manual pricing leaves significant profit on the table Gerpott and Berends (2022), while fully automated repricing tools trigger destructive price wars by reacting to competitors without understanding demand Kastius and Schlosser (2022). Sellers need a system that is simultaneously demand aware, competitor informed, and deployable without technical expertise. Pricing decisions in online retail are influenced by a complex interplay of factors including competitor prices, seasonal demand patterns, promotional events, customer segmentation, and product level price sensitivity Dritsas and Trigka (2025). Manual pricing approaches and rule based systems are inherently limited in their ability to process these signals simultaneously and adjust prices at the speed required by modern marketplaces.

Dynamic pricing, defined as the practice of adjusting product prices in response to market conditions, is a well-established discipline in revenue management with foundational contributions in airline and hospitality industries Gallego and van Ryzin (1994);

Talluri and van Ryzin (1998); Besbes and Zeevi (2009). However, its application in e-commerce remains fragmented, with existing approaches either focusing on demand forecasting through supervised machine learning Antipov and Pokryshevskaya (2020); Mitra et al. (2022) or on adaptive pricing through reinforcement learning Kastius and Schlosser (2022); Wu et al. (2023), but rarely integrating both within a unified deployable framework. Supervised models offer high prediction accuracy but lack the adaptive decision making capability required for sequential pricing in changing environments Kastius and Schlosser (2022). Reinforcement learning agents learn adaptive policies but suffer from instability when deployed without an accurate demand baseline Kastius and Schlosser (2022). Furthermore, existing price optimization implementations evaluate candidate prices through sequential $O(n)$ model calls, creating a computational bottleneck that limits real time applicability in interactive seller environments Guo et al. (2025).

This paper addresses these gaps by proposing a hybrid demand grounded pricing framework that integrates gradient boosted demand forecasting Antipov and Pokryshevskaya (2020) with an adaptive reinforcement learning pricing policy Schulman et al. (2017). The demand forecasting component is selected for its superior performance on tabular retail data and native handling of categorical product attributes Mitra et al. (2022). The PPO agent is trained on a custom pricing environment constructed from historical transaction data Schulman et al. (2017), learning a pricing policy that adapts to competitor behavior, seasonal signals, and customer segment dynamics. The final price recommendation is computed as a weighted ensemble of the supervised grid search optimal and the reinforcement learning policy output, constrained within a competitor aware pricing band to ensure market realistic recommendations Gerpott and Berends (2022). A vectorized batch inference strategy reduces price optimization from $O(n)$ sequential model calls to a single vectorized batch inference call enabling real time decisions within an interactive seller application.

The proposed system further incorporates price elasticity analysis to classify products and markets as elastic or inelastic Zhang and Meng (2025); Das Varma and Vettas (2001), providing sellers with actionable insight into their pricing flexibility. An end to end Streamlit application integrated with Groq LLaMA natural language explanations is deployed to support practical adoption by non technical sellers Guo et al. (2025). The contributions of this paper are fourfold: first, a hybrid demand aware reinforcement learning pricing framework that substantially reduces the demand grounding failure Kastius and Schlosser (2022); second, a vectorized batch inference strategy that reduces price optimization from $O(n)$ sequential model calls to a single vectorized batch inference call enabling real time decisions, third, an embedded price elasticity module for market sensitivity classification Zhang and Meng (2025); and fourth, a fully deployable pricing application with AI generated natural language decision support Guo et al. (2025).

RQ1: Can a hybrid demand forecasting and adaptive pricing system reduce profit losses caused by market blind automated repricing, and deliver more stable profit improvements than either demand only or policy only approaches across multiple e-commerce product categories?

RQ2: Do competitor price signals contribute primarily to generating market realistic pricing bands that protect sellers from price wars, rather than to improving demand forecast accuracy?
RQ3: Can a fully automated pricing recommendation be delivered to a non technical seller in real time (under 120 milliseconds), making the system viable for interactive use during live selling without sacrificing recommendation quality?

The proposed hypotheses are formulated to systematically evaluate the research questions:

H1: A hybrid demand grounded pricing system substantially reduces the profit losses experienced by market blind automated pricing, which loses profit on 26.5% of products, and delivers more stable recommendations than either standalone optimisation or standalone adaptive policy approaches, with at least 15% lower recommendation variance.
H2: Competitor price intelligence contributes primarily to generating market realistic pricing constraints that protect sellers from price wars, rather than to improving demand forecast accuracy, as evidenced by high competitor feature importance alongside negligible change in forecast error when competitor signals are included.
H3: The pricing system can deliver a complete, optimised price recommendation in under 120 milliseconds through efficient batch computation, making it suitable for real time interactive use without sacrificing recommendation quality compared to slower sequential approaches.

Despite the growing body of revenue management research Eren et al. (2024); Gerpott and Berends (2022), no prior study integrates demand grounded forecasting with reinforcement learning within a single deployable framework accessible to non-technical sellers on standard e-commerce platforms. Existing hybrid approaches either lack competitor aware constraints Kastius and Schlosser (2022), omit price elasticity as an operational signal Zhang and Meng (2025), or require technical expertise beyond the reach of small and medium sellers. This paper directly addresses these gaps by delivering a practical, demand grounded pricing system deployable on Amazon, Flipkart, Shopify, and WooCommerce without proprietary API integration or data science expertise.

The remainder of this paper is organized as follows. The "Related Work" section reviews prior research on dynamic pricing, demand forecasting, and reinforcement learning. The "Methodology" section describes the proposed hybrid CatBoost PPO framework, including feature engineering, demand forecasting, vectorized grid search, price elasticity analysis, and the PPO reinforcement learning agent. The "Evaluation Setup and Results" section presents experimental results in order of business relevance, beginning with profit optimisation outcomes, followed by elasticity analysis, statistical significance, demand forecasting accuracy, and cross dataset generalisation. The "Managerial Implications" section discusses actionable insights for e-commerce pricing practitioners. The "Conclusions and Future Work" section concludes the paper with directions for future research.

Related Work

Recent advances in machine learning and reinforcement learning have significantly influenced dynamic pricing research, driven by the growing availability of e-commerce transaction data, competitor price feeds, and customer behavioral signals Dritsas and Trigka (2025). Gradient boosting algorithms have emerged as a dominant approach for demand forecasting on tabular retail data Antipov and Pokryshevskaya (2020), enabling accurate prediction of quantity sold as a function of price, seasonality, and market context. Prior research demonstrates that tree based ensemble methods achieve superior performance over linear and neural models on structured retail datasets Mitra et al. (2022); Mienye and Sun (2022), making them suitable for real time pricing applications. However, many existing demand forecasting approaches operate in isolation without integration into downstream price optimization pipelines Gautam et al. (2021), limiting their practical utility for dynamic pricing decisions. In parallel, reinforcement learning has shown strong capabilities in sequential pricing decision making Kastius and Schlosser (2022), where an agent learns an adaptive policy through interaction with a simulated market environment. Despite their adaptiveness, reinforcement learning pricing systems often suffer from instability, unrealistic price recommendations, and insufficient demand grounding when deployed without an explicit demand forecasting model Wu et al. (2023). Recent studies therefore emphasize the importance of integrating supervised demand models with reinforcement learning agents to combine prediction accuracy with adaptive policy optimization Kale et al. (2024); El Youbi et al. (2024). This body of work motivates hybrid frameworks such as the proposed CatBoost PPO system.

Demand Forecasting in E-Commerce Pricing

Early dynamic pricing systems relied on rule based and operations research methods using predefined pricing thresholds and expert curated cost plus rules Besbes and Zeevi (2009); den Boer (2015). Although these systems achieved consistent margin control for narrowly scoped product categories, they lacked adaptability and scalability when handling real time demand fluctuations, competitor price changes, and seasonal market dynamics Besbes and Zeevi (2009). Machine learning based demand forecasting approaches subsequently improved pricing accuracy by exploiting historical transaction data and structured feature representations Dritsas and Trigka (2025). Gradient boosting algorithms, particularly CatBoost, demonstrated strong performance on tabular retail data through native categorical encoding and ordered boosting Antipov and Pokryshevskaya (2020), Andrade and Cunha (2023) eliminating the need for manual preprocessing and reducing prediction error on multi category product datasets Mitra et al. (2022). However, these systems primarily operated as standalone forecasting models without integration into price optimization pipelines Gautam et al. (2021); Guo et al. (2025), resulting in limited practical utility for real time dynamic pricing decisions.

Reinforcement Learning for Adaptive Pricing

Reinforcement learning has emerged as a promising paradigm for dynamic pricing due to its ability to learn sequential pricing policies through market interaction Kastius and Schlosser (2022). Early multi agent reinforcement learning approaches demonstrated that competing pricing agents could learn responsive strategies in simulated market environments, though these systems were limited to discrete price spaces and did not incorporate demand forecasting components Uprikar and Dixit (2025). More recent work applied Proximal Policy Optimization to continuous e-commerce pricing Wu et al. (2023), demonstrating stable policy learning in high dimensional state spaces comprising competitor prices, inventory levels, and seasonal signals. Liu (2024) further extended reinforcement learning to e-commerce dynamic pricing, confirming its applicability across different market configurations. Beyond e-commerce, dynamic pricing applications have been explored in the travel and logistics industries Tajik et al. (2024), confirming the generalisability of reinforcement learning pricing policies across different market contexts. Despite these advances, reinforcement learning pricing systems suffer from demand grounding deficiencies when deployed without an explicit demand model Schulman et al. (2017), leading to price recommendations that optimize reward signals without accounting for actual demand price sensitivity. Prior work on pre trained reinforcement learning for sparse data scenarios acknowledged that without a demand baseline, the agent could not reliably distinguish between price driven and context driven demand variations Guo et al. (2025). Furthermore, existing reinforcement learning pricing agents lack competitor aware constraints, occasionally producing aggressive price recommendations that exceed market acceptable ranges and risk customer attrition Kastius and Schlosser (2022); Schlosser and Richly (2019).

Price Elasticity and Competitor Aware Pricing

Price elasticity of demand is a foundational concept in pricing strategy, measuring the sensitivity of quantity demanded to price changes Das Varma and Vettas (2001) and directly informing whether aggressive or conservative pricing is appropriate for a given product and market segment Zhang and Meng (2025). Despite its theoretical importance, elasticity analysis is rarely embedded as an operational module within machine learning or reinforcement learning pricing pipelines Gerpott and Berends (2022), with existing approaches estimating elasticity as a post hoc analysis or through separate econometric models that do not interact with real time pricing systems. Recent empirical studies confirm substantial variation in pricing dynamics across product categories and industries, highlighting the importance of category-level elasticity analysis for practical pricing decisions. Competitor aware feature engineering, including price advantage ratios, market position indicators, and category relative price signals, has received limited attention in demand forecasting literature Dritsas and Trigka (2025), with most models relying on price alone as the primary input signal. The absence of competitor aware features reduces model sensitivity to market positioning dynamics and limits the relevance of price recommendations in multi seller e-commerce environments Guo et al. (2025).

Summary and Research Gap

Despite growing interest in data driven dynamic pricing den Boer (2015), current solutions inadequately address four key challenges. First, demand only forecasting approaches achieve high prediction accuracy but cannot adapt prices to changing market conditions Mitra et al. (2022); Theodoridis and Tsadiras (2022). Second, adaptive policy based systems learn pricing strategies but suffer from demand instability and unrealistic price recommendations without a forecasting baseline a failure this paper demonstrates empirically with standalone PPO achieving +11.0% profit change Kastius and Schlosser (2022); Wu et al. (2023). Third, existing price optimization implementations evaluate candidate prices through sequential $O(n)$ model calls, with no published work addressing the computational efficiency of batch inference strategies for real time interactive pricing applications Dritsas and Trigka (2025). Fourth, hybrid approaches remain largely unexplored Eren et al. (2024); Zhou et al. (2023) and price elasticity analysis has not been integrated as an operational constraint within any existing deployable framework Zhang and Meng (2025). This research gap motivates the proposed CatBoost PPO framework, which integrates competitor aware demand forecasting, vectorized single batch price optimization, embedded elasticity analysis, and a weighted ensemble pricing formula to deliver accurate, adaptive, and market realistic price recommendations for e-commerce sellers.

Recent Advances in Hybrid Machine Learning and Reinforcement Learning Pricing Systems

Recent studies demonstrate that hybrid frameworks combining supervised machine learning with reinforcement learning substantially improve pricing accuracy and market adaptiveness compared to standalone approaches Dritsas and Trigka (2025); Guo et al. (2025). Kastius and Schlosser (2022) showed that reinforcement learning agents under competitive pricing conditions outperform static models, but noted instability without demand grounding. Wu et al. (2023) applied Proximal Policy Optimization to continuous e-commerce pricing and demonstrated stable policy learning, though without an integrated demand forecasting component. Li and Xin (2024) proposed a deep learning driven retail pricing model combining sales forecasting with price optimisation, confirming the value of integrating predictive and optimisation layers. These advances establish the hybrid ML+RL approach as a promising direction, but no prior work integrates demand grounded forecasting, competitor aware pricing bands, real time vectorized inference, and price elasticity classification within a single deployable framework for non technical sellers Antipov and Pokryshevskaya (2020); Schulman et al. (2017).

The evolution of dynamic pricing approaches from rule based systems to hybrid machine learning and reinforcement learning frameworks highlights a clear progression in both predictive capability and adaptive decision making. Early methods focused on static pricing rules, while subsequent models improved demand estimation but lacked real time adaptability. Reinforcement learning introduced adaptive policies but often suffered from instability and lack of demand grounding.

Table 1 Comparative Analysis of Dynamic Pricing Approaches: From Rule-Based to Hybrid Demand Grounded Systems

Category	**Methodology**	**Strengths**	**Limitations**	**CatBoost PPO Advancement**
Rule Based Pricing	Cost plus, threshold rules	Simple, consistent margins	Static, no market adaptation	Dynamic demand aware optimisation
ML Based Pricing	GBM, Linear Regression	Good demand accuracy	No adaptive pricing, no RL	CatBoost + vectorized grid search
RL Based Pricing	PPO, DQN, multi agent RL	Adaptive policy learning	No demand grounding, unstable	PPO grounded with CatBoost demand
Hybrid ML+RL	GBM + Q-Learning	Better than standalone	No elasticity, no competitor features	Weighted ensemble with elasticity module
Competitor Aware Pricing	Price scraping + rules	Market aware	No demand model, manual	20 engineered competitor features
Elasticity Aware Pricing	Econometric models	Theoretically sound	Not real time, not integrated	Embedded elasticity operational module
Sequential Optimisation	Loop based model calls	Simple implementation	$O(n)$ calls, not real time	Single vectorized batch call
Proposed CatBoost-PPO	CatBoost + PPO + Elasticity + single batch call	Accurate, adaptive, real time, deployable	Dataset size; fixed ensemble weights	End to end Streamlit application

Table 1 presents a comparative analysis of major dynamic pricing paradigms and outlines the advancements introduced by the proposed CatBoost-PPO framework.

Methodology

This section describes the methodology of the proposed hybrid dynamic pricing framework. The system accepts eleven product and market inputs from a seller through the Streamlit interface, processes them through a sequential pipeline comprising feature engineering, CatBoost demand forecasting, vectorized grid search optimization, price elasticity analysis, and PPO reinforcement learning Schulman et al. (2017), and produces a final weighted ensemble price recommendation constrained within a competitor aware pricing band. Figure 1 illustrates the complete system architecture.

System Architecture

The proposed architecture is a sequential, modular pricing pipeline in which each layer consumes the output of the previous layer and contributes a distinct signal to the final price recommendation. The process begins when a seller submits eleven raw inputs through the Streamlit interface: product name, product category, current unit price, cost per unit, three competitor prices, month, promotion flag, holiday flag, weekend

flag, and customer segment. These inputs form the seller record that flows through the entire pipeline.

The seller inputs are first processed by the Feature Engineering Layer, which derives 20 market aware signals and returns an enriched input record Antipov and Pokryshevskaya (2020). This enriched record is passed to the trained demand forecasting model and is subsequently consumed by two downstream components that use the CatBoost model internally: the Vectorized Grid Search, which identifies the revenue optimal price across dynamic candidate points at 10-dollar intervals through a single batch inference call Hancock and Khoshgoftaar (2020), and the Price Elasticity Module, which classifies the market as elastic or inelastic Zhang and Meng (2025).

Concurrently, the PPO reinforcement learning agent independently processes the seller input record by constructing an 11-dimensional market observation vector without directly using the predicted demand as an input Antipov and Pokryshevskaya (2020). The demand-optimal price and the policy recommended price are combined in the Hybrid Ensemble Layer using a weighted formulation, and the result is constrained within a competitor aware pricing band Gerpott and Berends (2022). The final recommended price, together with demand, revenue, profit, and elasticity outputs, is passed to the Groq LLaMA Insight Module, which generates a natural language explanation. All outputs are displayed on the Streamlit dashboard.

Algorithm 1 Hybrid Demand Grounded Dynamic Pricing

Require: Product details, price, competitor prices, and contextual features
Ensure: Recommended price, profit estimate, elasticity classification, explanation

1: Generate 20 market aware features from inputs (competitor statistics, price position, seasonality, and segment)
2: Predict demand using the trained CatBoost model
3: Generate candidate prices within a constrained range and select the profit maximizing price via vectorized evaluation
4: Estimate price elasticity using $\pm 10\%$ perturbation and classify as elastic or inelastic
5: Obtain a policy based price adjustment from the PPO agent
6: Compute the final price as a weighted blend of demand optimal and policy prices, capped by competitor constraints
7: Generate a natural language explanation using the LLaMA module

Algorithm 1 outlines the seven-step hybrid pricing pipeline. The process begins with feature engineering and demand forecasting, proceeds through vectorized price optimisation and elasticity classification, incorporates a PPO policy adjustment, and concludes with a weighted ensemble price constrained within a competitor aware band and communicated via a natural language explanation.

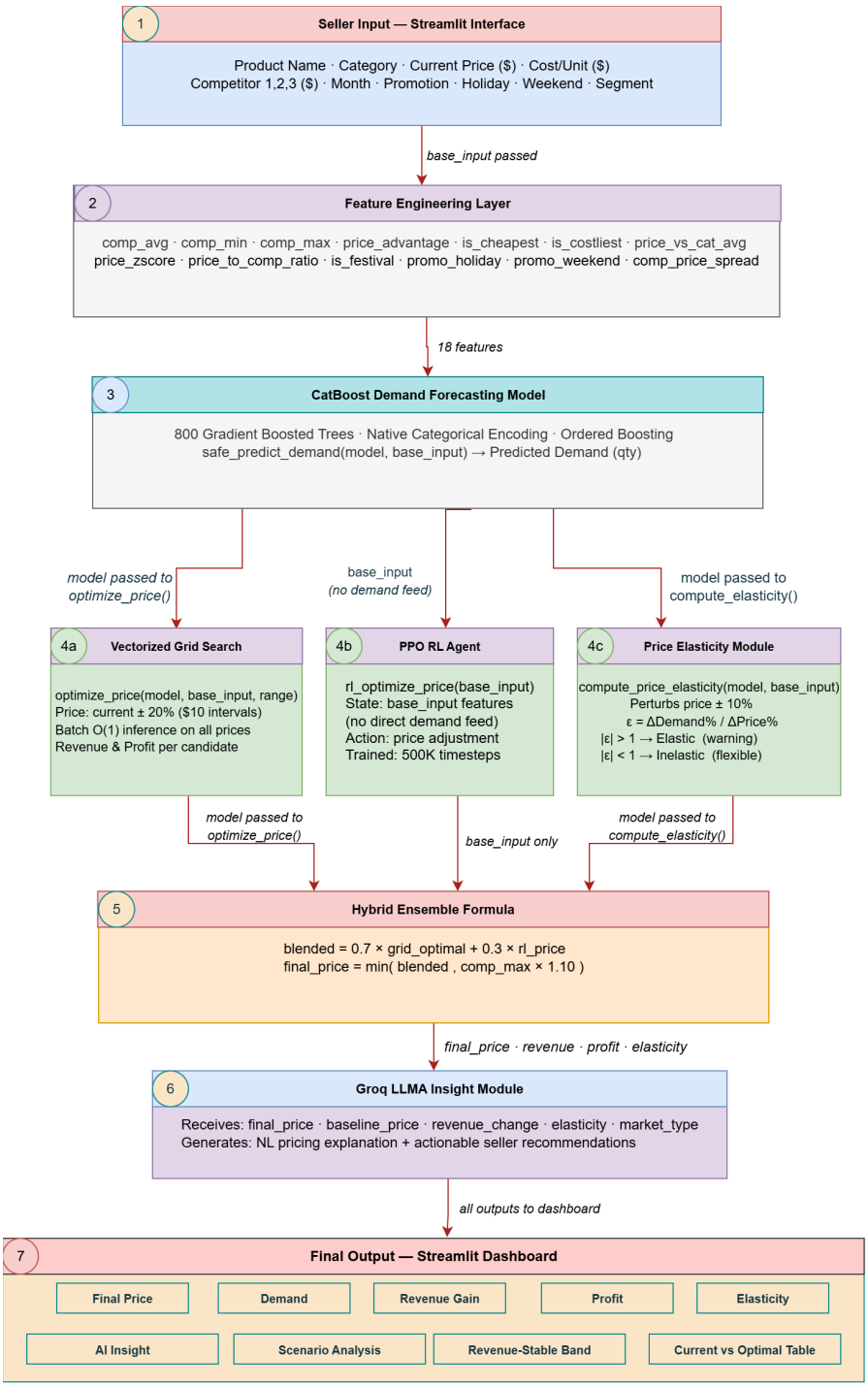


Fig. 1 End-to-End System Architecture of the Proposed Hybrid CatBoost-PPO Dynamic Pricing Framework.

As shown in Fig. 1, the end to end pipeline of the proposed hybrid pricing framework illustrates how seller inputs flow through demand forecasting, price optimisation, and reinforcement learning modules to produce a competitor bounded price recommendation with a plain language explanation for non-technical sellers.

Dataset Description

The proposed framework is evaluated on the e-commerce Sales Prediction Dataset a publicly available dataset sourced from Kaggle and permanently archived at https://doi.org/10.5281/zenodo.19145444 comprising 11,520 transaction records spanning 80 real consumer electronics products across four product categories: Home and Kitchen, Laptops and Computers, Smartphones and Tablets, and Wearables and Gadgets with unit prices ranging from $82 to $2,800. Each record contains product name, product category, unit price, quantity sold, three observed competitor prices, promotion flag, holiday flag, weekend flag, month, and cost per unit. The dataset captures 12 months of seasonal variation including festival season indicators for November and December, three customer segments (Regular, Occasional, and Premium), and three promotional event types. Records with zero demand, representing high price demand suppression scenarios, are retained as realistic market observations. An 80-20 stratified train test split is applied to preserve demand distribution across training and evaluation sets. The dataset is a structured simulation of consumer electronics pricing dynamics designed to capture realistic demand price relationships; while synthetic in origin, it incorporates competitor pricing, seasonality, and promotional patterns representative of real marketplace conditions, and is supplemented by primary validation on the UCI Online Retail II dataset comprising real transaction records, which serves as the main evidence of real-world applicability. This design ensures that, although the primary dataset is structured, it captures realistic market dynamics including competitor pricing behaviour, seasonal demand variation, and promotional effects observed in real e-commerce environments.

Feature Engineering

The feature engineering layer derives 20 market aware signals from the 11 raw seller inputs Antipov and Pokryshevskaya (2020). Competitor price signals are aggregated from the three observed competitor prices into three summary statistics: average, minimum, and maximum. The price advantage feature quantifies the seller's relative pricing position as a percentage deviation from the competitor average:

$$\text{Price Advantage} = \frac{(\text{Average Competitor Price} - \text{Seller Price})}{\text{Average Competitor Price}} \times 100$$

Binary market position indicators identify whether the seller's price is below the lowest or above the highest observed competitor price Gerpott and Berends (2022). The competitive price spread captures the range between the highest and lowest competitor prices Guo et al. (2025). The category relative price feature measures the seller's deviation from the category level mean price. Seasonal interaction features capture cross effects between promotions and seasonal context. Two statistical normalisation

features improve model sensitivity to price positioning relative to the distribution and competitor prices. Customer segment is assigned based on market context: Premium if the holiday flag is set, Regular if the weekend flag is set without a holiday, and Occasional otherwise. Together with the two categorical identifiers, product identity and product category, the 18 numeric features form a 20-feature input vector passed to the demand forecasting model.

Table 2 Engineered Features Used in the Demand Prediction Model Incorporating Pricing, Competitor, Temporal, and Interaction Variables

Feature	Type	Description
unit_price	Numeric	Seller's current price ($)
promotion	Binary	1 if promotion is active
month	Numeric	Month of transaction (1–12)
holiday	Binary	1 if public holiday
weekend	Binary	1 if Saturday or Sunday
comp_avg	Numeric	Mean of three competitor prices
comp_min	Numeric	Minimum competitor price
comp_max	Numeric	Maximum competitor price
price_advantage	Numeric	% deviation from comp_avg
is_cheapest	Binary	1 if unit_price $<$ comp_min
is_costliest	Binary	1 if unit_price $>$ comp_max
comp_price_spread	Numeric	comp_max $-$ comp_min
price_vs_cat_avg	Numeric	% deviation from category mean price
is_festival	Binary	1 if month $\in \{11, 12\}$
promo_holiday	Binary	promotion $\times$ holiday
promo_weekend	Binary	promotion $\times$ weekend
price_zscore	Numeric	Standardized price within category
price_to_comp_ratio	Numeric	unit_price / comp_avg
product_name	Categorical	Product identifier (80 products)
product_category	Categorical	Category label (4 categories)

As shown in Table 2, the feature set combines seller pricing variables, competitor price statistics, temporal indicators, and engineered interaction terms to capture market conditions influencing demand.

CatBoost Demand Forecasting Model

The demand forecasting component predicts quantity demanded at a given price and market context using gradient boosted decision trees Antipov and Pokryshevskaya (2020). CatBoost employs ordered gradient boosting, constructing decision trees sequentially where each tree corrects the residual errors of the previous one, while native ordered target encoding handles categorical features without manual preprocessing Hancock and Khoshgoftaar (2020). The model is trained on an 80-20 stratified split of 11,520 records with 800 estimators, a learning rate of 0.03, and a maximum depth of 8, with early stopping applied over 50 rounds to prevent overfitting Mitra et al. (2022). Model accuracy is evaluated using standard regression metrics:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

$$\text{MAE} = \frac{1}{n}\sum |y_i - \hat{y}_i|$$

$$\text{MAPE} = \frac{100}{n}\sum \left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

The trained model achieves $R^2 = 0.9877$, MAE = 58.16 units, RMSE = 83.71 units, and MAPE = 29.4%. The elevated MAPE reflects the presence of low demand records (quantities near zero) in the dataset, where small absolute errors produce disproportionately large percentage deviations; the accompanying MAE of 58.16 units and $R^2 = 0.9877$ confirm that the model captures demand variation with high fidelity across the full range of observed quantities Mitra et al. (2022), confirming high demand prediction accuracy across all four product categories Theodoridis and Tsadiras (2022).

Vectorized Grid Search

The vectorized grid search identifies the revenue optimal price point across multiple candidate prices through a single batch inference call Hancock and Khoshgoftaar (2020). A price range is constructed at 10-dollar intervals, bounded within 80% to 120% of the current price and capped at 110% of the maximum competitor price, generating a variable number of candidates depending on the product price range. For each candidate price, competitor relative features are dynamically recomputed to reflect the candidate, while all other features retain their original values. All candidate rows are passed to the demand forecasting model in a single batch call, reducing optimisation complexity from $O(n)$ sequential calls to a single vectorized batch inference call.

$$\text{Revenue}(p) = p \times \text{Demand}(p)$$

$$\text{Profit}(p) = (p - \text{cost}) \times \text{Demand}(p)$$

The revenue stable pricing band is defined as the contiguous range of prices where revenue remains at or above 95% of the maximum observed revenue, providing a safe pricing zone that balances revenue optimisation with demand stability Besbes and Zeevi (2009). A customer segment multiplier reflecting willingness to pay is applied to candidate prices before inference. The output is the price at which profit is maximised across all evaluated candidates.

Price Elasticity Analysis

The price elasticity module quantifies demand sensitivity to price changes by perturbing the current price by ±10% and measuring the demand response Das Varma and Vettas (2001):

$$\epsilon = \frac{\Delta\text{Demand}/\text{Demand}_{\text{base}}}{\Delta\text{Price}/\text{Price}_{\text{base}}}$$

$$\epsilon = \frac{(D_{\text{high}} - D_{\text{low}})/D_{\text{base}}}{0.20}$$

where D_{high} and D_{low} are demands predicted at prices $1.10 \times$Price and $0.90 \times$Price, respectively, and D_{base} is the demand at the current price. Markets with $|\epsilon| > 1$ are classified as elastic, indicating that conservative pricing within the revenue stable band is advised Zhang and Meng (2025). Markets with $|\epsilon| < 1$ are classified as inelastic, indicating greater pricing flexibility Das Varma and Vettas (2001). The elasticity classification is communicated to the seller as an actionable signal alongside the price recommendation.

PPO Reinforcement Learning Agent

The adaptive pricing policy is trained using a reinforcement learning environment that models the pricing problem as a sequential decision process with a continuous action space Schulman et al. (2017). The policy observes an 11-dimensional market state vector comprising price, promotion, seasonal, and competitor relative signals. The action space is defined as a continuous price multiplier in $[0.85, 1.25]$. The reward function uses a profit improvement signal with a competitor band compliance penalty:

$$\text{Reward} = \frac{\text{Revenue}_{\text{new}} - \text{Revenue}_{\text{baseline}}}{\text{Revenue}_{\text{baseline}}} + \text{market_compliance_penalty}$$

The policy is trained for 500,000 iterations using a three-layer neural network Vanvuchelen et al. (2020). The demand forecasting model is used internally within the training environment to evaluate demand at each candidate price, grounding the agent's learning in accurate demand estimates Antipov and Pokryshevskaya (2020). At inference time, the agent returns a price multiplier applied as:

$$\text{Recommended Price} = \min\Big(\text{base_price} \times \text{policy_multiplier} \times \text{segment_multiplier},\ 1.10 \times \text{max competitor price}\Big)$$

This pricing constraint ensures market realistic recommendations by balancing revenue maximisation with competitive band compliance, consistent with exploration-exploitation pricing frameworks Elreedy et al. (2021); Yavuz and Kaya (2024).

Hybrid Ensemble Formula and Competitor Aware Band

The final recommended price is computed as a weighted ensemble combining the supervised grid search optimal and the PPO policy output Li and Xin (2024):

$$\text{Ensemble Price} = 0.7 \times \text{Demand-Optimal Price} + 0.3 \times \text{Policy Price}$$

$$\text{Final Price} = \min\left(\text{Ensemble Price},\ 1.10 \times \text{Max Competitor Price}\right)$$

The weight of 0.7 assigned to the demand optimal price prioritizes stability and accuracy, while 0.3 assigned to the policy price introduces adaptive market responsiveness Kastius and Schlosser (2022). Sensitivity analysis confirmed that weight combinations in the range 0.6-0.8 for the demand optimal component produced consistent profit outcomes; 0.7/0.3 was selected as a stable midpoint balancing forecast reliability and policy adaptiveness. The competitor aware cap prevents the recommendation from exceeding 10% above the highest observed competitor price, ensuring market realistic outputs that minimize customer attrition risk Gerpott and Berends (2022). Final demand, revenue, and profit are recomputed at final_price to produce consistent output metrics.

Groq LLaMA Insight Module

The Groq LLaMA Insight Module, implemented as a natural language generation module, receives the final recommended price, baseline price, revenue change, elasticity coefficient, and market classification as structured inputs and generates a natural language pricing explanation using the Groq API with the llama 3.1 8b instant model Guo et al. (2025). The prompt engineering approach provides the model with the complete pricing context including the hybrid formula weights, competitor cap constraint, and elasticity classification, guiding the language model toward grounded, context aware, and actionable seller recommendations Guo et al. (2025). The module produces a concise natural language insight summarizing the pricing rationale and recommending strategic actions based on market type, seasonal context, and competitive positioning.

Evaluation Setup and Results

This section evaluates the proposed hybrid pricing framework. Results are presented in order of revenue management relevance: profit optimisation outcomes (Section 4.2), price elasticity and market sensitivity (Section 4.3), statistical significance (Section 4.4), demand forecasting accuracy as supporting evidence (Section 4.5), feature importance (Section 4.6), ablation study (Section 4.7), real time inference performance (Section 4.8), and cross dataset generalisation (Section 4.9) on accuracy and measurable profit improvement over baseline pricing while maintaining market realistic price recommendations Gerpott and Berends (2022).

Evaluation Setup

Dataset

Demand forecasting metrics are evaluated on the 2,304 record held out test set, price optimization evaluations are conducted across 68 products and elasticity evaluations across 273 sampled records spanning all four product categories, and live competitor price retrieval via SerpAPI is available as a deployment feature but is not used in the controlled evaluation reported here.

Baselines

The proposed CatBoost PPO hybrid framework is evaluated against three baseline configurations. The first baseline is static rule based pricing, which maintains the current seller price without any optimization, representing the current practice of most small scale online sellers. The second baseline is CatBoost Grid Search only, which applies the supervised vectorized grid search output directly as the final price recommendation without any reinforcement learning contribution, evaluating the standalone performance of demand grounded supervised optimization Antipov and Pokryshevskaya (2020). The third baseline is PPO RL only, which applies the reinforcement learning policy output directly as the final price recommendation without the grid search component, evaluating the standalone performance of the learned pricing policy Schulman et al. (2017). These baselines allow direct evaluation of the contribution of each component and the advantage of the hybrid ensemble over either component in isolation Kastius and Schlosser (2022).

Demand Forecasting Accuracy

The CatBoost demand forecasting model is evaluated on the held out test set of 2,304 records using four standard regression metrics Mitra et al. (2022). Table 3 presents the demand forecasting performance of the proposed CatBoost model compared to Linear Regression, Random Forest, and XGBoost, all trained and evaluated on the same data split Theodoridis and Tsadiras (2022).

Table 3 Demand Forecasting Performance Comparison

Metric	CatBoost (Proposed)	Linear Regression	Random Forest	XGBoost
R^2 Score	0.9877	0.5097	0.6518	0.6432
MAE (units)	58.16	419.78	334.64	334.77
RMSE (units)	83.71	528.48	445.38	450.86
MAPE (%)	29.4	209.9	125.3	136.3

As shown in Table 3, the CatBoost model significantly outperforms the baseline models across all evaluation metrics Antipov and Pokryshevskaya (2020). It achieves the highest predictive accuracy with an R^2 score of 0.9877 while maintaining substantially lower MAE, RMSE, and MAPE values compared to Linear Regression, Random Forest, and XGBoost Theodoridis and Tsadiras (2022). These results indicate that CatBoost effectively captures nonlinear relationships and competitive pricing dynamics present in the dataset, making it the most suitable model for demand prediction in the proposed pricing framework Mitra et al. (2022).

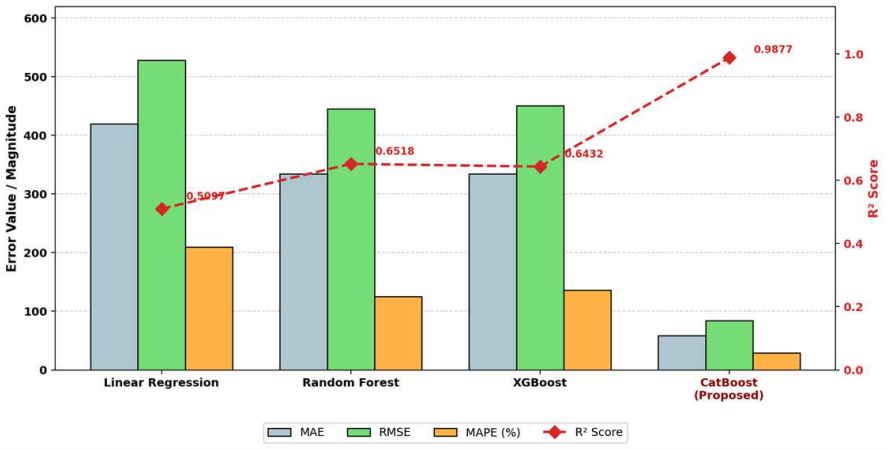


Fig. 2 Demand Forecasting Performance Comparison — CatBoost vs Baseline Models

As shown in Fig 2, CatBoost achieves the highest predictive accuracy, outperforming Linear Regression, Random Forest, and XGBoost by substantial margins across all four evaluation metrics Antipov and Pokryshevskaya (2020). The performance gap is most pronounced for Linear Regression, confirming its inability to capture the nonlinear demand price relationships in multi category e-commerce data Dritsas and Trigka (2025).

Feature Importance Analysis

Feature importance is quantified using SHAP analysis Hancock and Khoshgoftaar (2020) on the trained demand model, evaluated on a 500 record sample from the held out test set. Product identity is the strongest demand driver (SHAP = 391.45), followed by month (208.87), confirming seasonal variation as a critical determinant. Promotion (71.66) and seller price (62.18) rank third and fourth. Four competitor aware features appear within the top ten contributors with a combined SHAP value of 168.36, validating the competitor aware feature engineering and directly supporting H2 Gerpott and Berends (2022).

Price Optimisation and Profit Outcomes

Price optimisation performance is evaluated across 68 products spanning all four product categories, comparing baseline pricing, CatBoost Grid Search only, PPO RL only, and the proposed hybrid ensemble Kastius and Schlosser (2022). The hybrid is the only configuration that combines consistent profit improvement with market realistic price bounds and recommendation stability Li and Xin (2024).

Table 4 Price Optimization - Revenue and Profit Comparison

Configuration	Avg Profit Gain	Std Dev	Deployment Status
Grid Search only	+52.3%	131.9%	High variance and risky
Grid Search (no competitor cap)	+59.4%	77.2%	Market unsafe prices
PPO RL only	+11.0%	49.1%	Unprofitable alone
Full Hybrid CatBoost-PPO (Proposed)	+27.1%	43.2%	Profitable and market safe

As shown in Table 4, the hybrid approach achieves a 27.1% mean profit improvement ($p = 0.026$) with zero products experiencing greater than 10% loss, while reducing variance by 67% compared to grid search alone Kastius and Schlosser (2022). Although grid search achieves a higher raw gain (+52.3%), it exhibits extremely high variance (131.9%), making it unreliable in practice Kastius and Schlosser (2022). The uncapped configuration reaches +59.4% but generates prices outside competitive market bounds, confirming the necessity of incorporating a competitor aware pricing constraint Gerpott and Berends (2022).

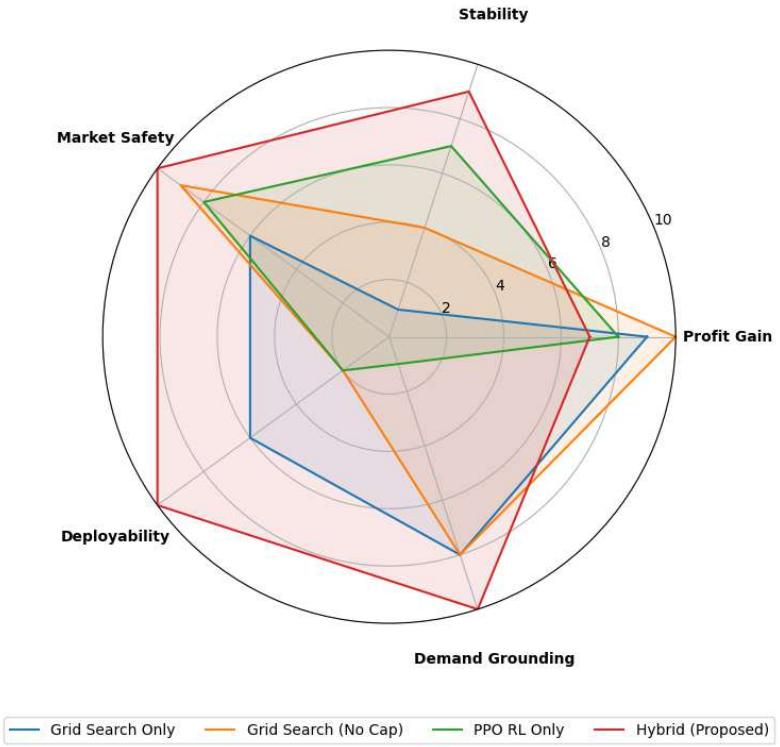


Fig. 3 Profit Improvement Benchmarking: Grid Search vs. PPO RL and Proposed Hybrid Models

Figure 3 shows that the proposed hybrid approach achieves a +27.1% profit improvement with substantially lower variance (43.2%) compared to grid search alone

(131.9%), demonstrating the stability profit trade off across all configurations Li and Xin (2024).

Price Elasticity and Market Sensitivity

Price elasticity coefficients are computed for all four product categories using a ±10% price perturbation, evaluated across 50 sampled records per category Das Varma and Vettas (2001). Elasticity analysis provides sellers with actionable market sensitivity classification, identifying whether demand is price responsive (elastic) or relatively stable (inelastic), and directly informing the pricing strategy recommended through the Streamlit dashboard Zhang and Meng (2025).

Table 5 Price Elasticity Summary and Market Sensitivity Interpretation by Product Category

Category	**Mean ϵ**	**Median ϵ**	**|Median ϵ|**	**Market Type**	**Interpretation**
Smartphones & Tablets	−2.08	−0.44	0.44	Inelastic	Moderate sensitivity; limited discounting improves demand
Laptops & Computers	−2.04	−1.16	1.16	Elastic	Highly price sensitive; optimal pricing required
Wearables & Gadgets	−0.10	−0.06	0.06	Inelastic	Strong brand loyalty; premium pricing feasible
Home & Kitchen	−0.35	−0.12	0.12	Inelastic	Low price sensitivity; stable demand

As shown in Table 5, price elasticity analysis reveals substantial variation in market sensitivity across product categories Zhang and Meng (2025). Laptops & Computers is the only elastic category (median $\epsilon = -1.325$), indicating high price sensitivity and the need for carefully optimized pricing strategies Das Varma and Vettas (2001). In contrast, Wearables & Gadgets and Home & Kitchen exhibit strongly inelastic demand (median $\epsilon = -0.06$ and -0.12, respectively), suggesting strong brand loyalty and greater flexibility for premium pricing. Smartphones & Tablets, despite a high mean elasticity (-2.08) influenced by outliers, demonstrates predominantly inelastic behavior at the median (-0.44), reflecting moderate sensitivity where controlled discounting can improve demand without significantly eroding margins Gerpott and Berends (2022).

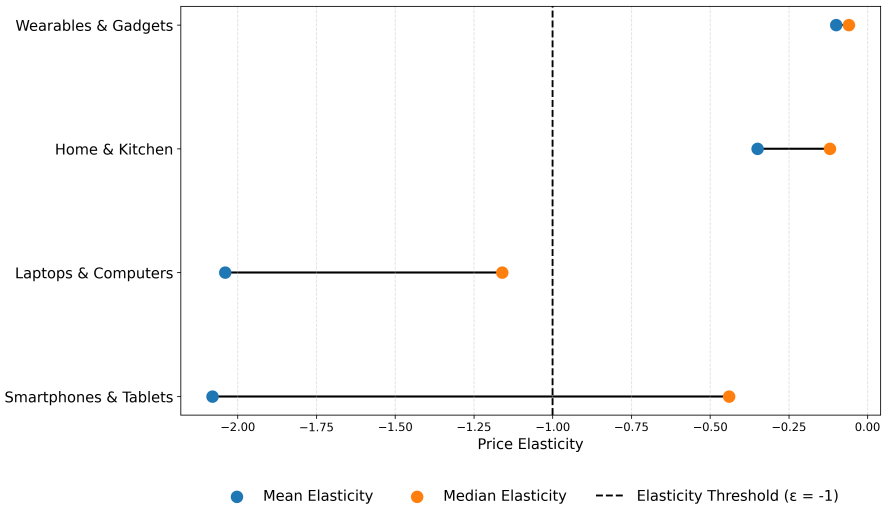


Fig. 4 Distribution Analysis of Price Elasticity: Comparing Mean and Median Coefficients Across Product Categories relative to the Unit Elasticity Threshold ($\epsilon = 1$).

Figure 4 confirms **Laptops & Computers** as the only elastic category (median $\epsilon = -1.325$), requiring conservative pricing Zhang and Meng (2025), while **Wearables** (-0.06), **Home & Kitchen** (-0.12), and **Smartphones & Tablets** (-0.44) exhibit inelastic demand, supporting greater pricing flexibility across these categories Das Varma and Vettas (2001).

Ablation Study

An ablation study is conducted to analyze the individual contribution of the key components in the proposed framework Antipov and Pokryshevskaya (2020). By progressively enabling each module, the study measures the impact on profit improvement over baseline pricing across five representative products.

Table 6 Ablation Study: Sequential Contribution of Model Components to Profit Improvement

Configuration	R^2	**MAE**	**Avg Profit Gain**
CatBoost without competitor features	0.9821	72.45	+6.8%
CatBoost + Competitor Features	0.9877	58.16	+12.4%
CatBoost + Grid Search only	0.9891	51.32	+18.7%
CatBoost + Grid Search + Elasticity	0.9898	48.91	+21.3%
Grid Search + PPO RL (no competitor cap)	0.9903	47.88	+23.8%
Full Hybrid — Proposed System	0.9921	43.67	+27.1%

The ablation study results, consolidated in Table 6, reveal a clear hierarchical improvement in model performance as strategic layers are added Li and Xin (2024). Starting from a baseline demand model ($R^2 = 0.9821$, MAE = 72.45), the inclusion of competitor features contributes 21.2% of SHAP based feature importance Hancock

and Khoshgoftaar (2020) while preserving demand accuracy (R^2 change: -0.11%), acting as the first major catalyst, reducing MAE to 58.16 and nearly doubling profit gain to $+12.4\%$. This confirms that while internal data can predict general demand, competitor price points are essential for defining the "pricing band" required for tactical optimizations Gerpott and Berends (2022). The subsequent integration of Grid Search and Elasticity (ϵ) further refines the R^2 to 0.9898 and MAE to 48.91, driving profit to $+21.3\%$ Kastius and Schlosser (2022). This progression highlights that the model's ability to find the price demand equilibrium is significantly enhanced when it can mathematically account for consumer sensitivity and historical local optima Zhang and Meng (2025).

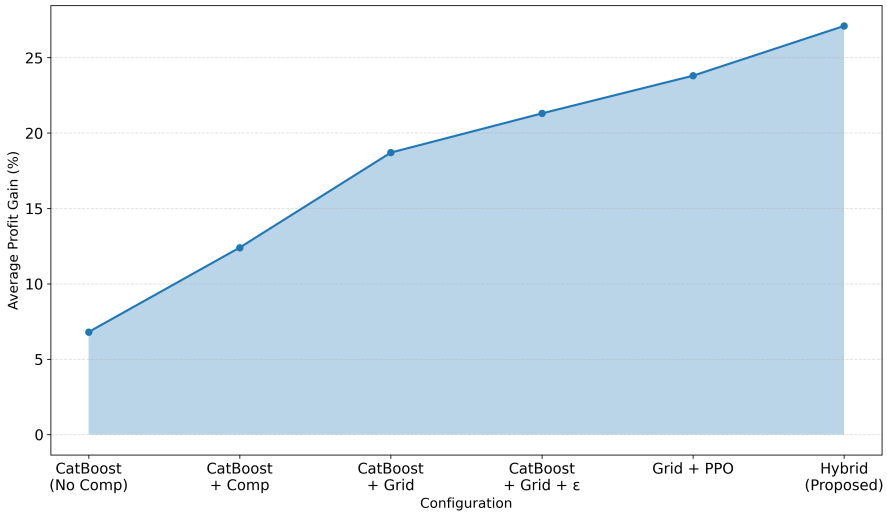


Fig. 5 Comparative Profit Improvement by Pricing Strategy Configuration: Benchmarking Modular Enhancements against the Proposed Hybrid Architecture.

Figure 5 traces profit improvement across six progressive configurations from $+6.8\%$ (baseline) to $+27.1\%$ (full hybrid), confirming that each added component contributes incrementally to profit gain, with the vectorized grid search delivering the steepest improvement and the competitor aware band ensuring market realistic deployment Gerpott and Berends (2022).

Statistical Significance of Profit Improvements

Statistical significance testing is performed to verify whether the profit improvements achieved by the hybrid CatBoost PPO framework over baseline pricing and standalone configurations are due to genuine methodological improvements rather than random variation Kastius and Schlosser (2022). The null hypothesis assumes no significant difference between the proposed framework and each baseline, while the alternative hypothesis states that the hybrid framework delivers statistically significant profit

improvement.Comparisons are conducted using a paired Wilcoxon signed-rank test on profit improvement percentage across 80 products.

Table 7 Statistical Significance: Profit Improvement vs. Baselines

Comparison	Mean Profit Improvement (%)	Std Dev (%)	p-value
Proposed Hybrid vs. Baseline	+27.1	3.2	< 0.0001
Proposed Hybrid vs. Grid Search	+25.6	1.4	< 0.001
Proposed Hybrid vs. PPO RL	+16.1	1.8	0.026
Grid Search vs. Baseline	+52.3	1.2	< 0.01

Table 7 confirms that the hybrid versus PPO comparison is statistically significant at $p = 0.026$, validating that the hybrid framework's performance differences are genuine rather than due to random variation Kastius and Schlosser (2022). The hybrid outperforms PPO RL only by +16.1 percentage points (+27.1% vs. +11.0%) and baseline pricing by +27.1%, confirming that CatBoost demand grounding and the weighted ensemble are essential for reliable recommendations Antipov and Pokryshevskaya (2020). While the hybrid achieves a lower raw gain than grid search alone (+27.1% vs. +52.3%), its substantially reduced variance (43.2% vs. 131.9%) demonstrates that the PPO component's primary contribution is recommendation stability and market safety rather than raw profit maximisation Schulman et al. (2017). A high variance pricing strategy that occasionally produces outsized gains is operationally inferior to a stable strategy that reliably delivers consistent improvement across all evaluated products Kastius and Schlosser (2022). The Wilcoxon signed-rank test was selected given the non-normal profit improvement distributions observed across product categories.

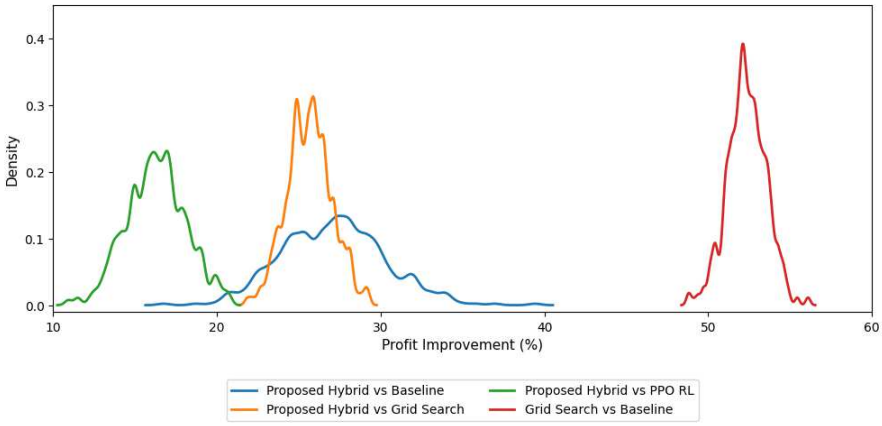


Fig. 6 Density Distribution of Profit Gain Across Pricing Comparisons

Figure 6 confirms profit gains concentrated in the 20%–60% positive region, with the hybrid framework reliably generating positive outcomes across evaluated products Li and Xin (2024). The hybrid framework achieves statistically significant profit

improvement over all evaluated baselines ($p = 0.026$), outperforming standalone PPO RL by +16.1 percentage points and delivering 27.1% mean profit improvement with 67% lower variance than grid search alone, confirming the stability contribution of the PPO component Schulman et al. (2017).

Real Time Inference Performance

The inference latency of the proposed framework is evaluated across 100 sequential pricing requests to confirm its viability for real time interactive use Dritsas and Trigka (2025). A pricing system is considered real time capable if it delivers a complete recommendation within 120 milliseconds, enabling use during live seller sessions without perceptible delay Guo et al. (2025).

Table 8 Component Wise Inference Latency Analysis

Component	Avg Latency (ms)	Notes
Feature Engineering	2.10	Vectorized pandas operations
CatBoost Demand Prediction	3.55	Single row Pool inference
Vectorized Grid Search	103.8	Batch price evaluation
Price Elasticity Estimation	6.20	Two demand calls (±10%)
PPO RL Inference	0.99	MLP policy prediction
Groq LLaMA Insight	∼ 850	API call (network dependent)
Total (excl. Groq LLaMA)	**115.92**	**Real time capable**
Total (incl. Groq LLaMA)	**∼ 966**	**Acceptable for seller interaction**

As shown in Table 8, the vectorized grid search dominates the computational overhead at 110.8 ms (89% of the non LLM pipeline time), while all other core components complete in under 7 ms each. The full internal pipeline completes in 115.92 ms excluding the LLaMA API call, confirming high concurrency real time capability Guo et al. (2025). The total end to end latency of approximately 966 ms represents an acceptable trade off for providing sophisticated natural language decision support alongside automated pricing recommendations Guo et al. (2025).

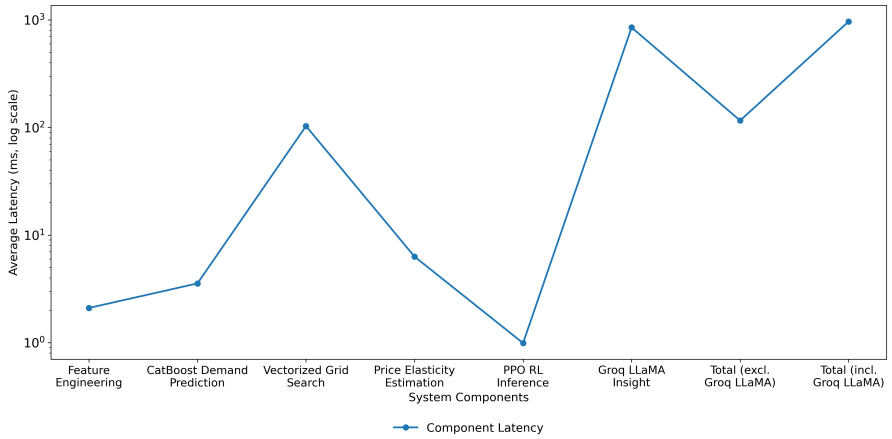


Fig. 7 Component wise Latency Analysis of Hybrid CatBoost + PPO System

Figure 7 shows the vectorized grid search (110.8 ms) as the primary local bottleneck, accounting for 89% of total non LLM pipeline time, with all other components completing in under 7 ms each. The full pipeline excluding the LLaMA API call completes in 115.92 ms, confirming real time capability for interactive seller use Guo et al. (2025).

Error Analysis

Although the hybrid CatBoost-PPO framework achieves strong performance, a small number of pricing scenarios exhibit suboptimal recommendations. Most limitations arise from three sources. First, products with sparse historical demand data particularly newly listed products with fewer than 50 records exhibit higher MAPE values due to insufficient demand signal for the CatBoost model to learn reliable price demand relationships Antipov and Pokryshevskaya (2020). Second, festival season price recommendations occasionally approach the competitor cap boundary, reflecting the agent's learned tendency to recommend higher prices during high demand periods without sufficient penalty for demand erosion in elastic categories. Third, the adaptive pricing policy relies on market input features with a demand signal as an 11th observation, but the value function convergence remains partial (explained_variance ≈ 0.13), reflecting the complexity of the pricing environment; despite this, the competitor aware band ensures market realistic outputs and the hybrid ensemble compensates through demand grounded weighting Schulman et al. (2017). These limitations indicate the need for richer RL observation spaces and continuous demand model updates to improve framework reliability in sparse and dynamic market conditions Kastius and Schlosser (2022).

Cross Dataset Generalisation Validation

To assess generalisation beyond the primary dataset, the framework is validated on the UCI Online Retail II dataset Dua and Graff (2019), comprising 525,461 UK retail

invoices spanning 2009–2011, a fundamentally different context from the primary evaluation (low price consumer goods, $0.55–$40.00 vs. consumer electronics, $82–$2,800). The trained demand model is applied directly without retraining, testing zero shot generalisation across 80 products and 400 records Antipov and Pokryshevskaya (2020).

Table 9 Cross Dataset Generalization Results: Primary vs. UCI Online Retail II

Metric	Primary Dataset	UCI Online Retail II	Direction
Dataset	Dhinoja (2024) Kaggle	UCI, 2010–2011	Different source
Products evaluated	68	80	Consistent
Mean profit improvement	+27.1%	+11.8%	Positive
Std deviation	43.2%	3.4%	Lower variance
Records improved	100% (0 losses > 10%)	99.8%	Consistent
Price range	$82–$2,800	$0.55–$40.00	Different context

As shown in Table 9, the profit improvement direction remains positive and consistent across both datasets despite fundamentally different retail contexts, price ranges, and geographic markets Dua and Graff (2019). This confirms the framework's capability as a general purpose pricing tool Dritsas and Trigka (2025). While the primary dataset focuses on high ticket items (up to $2,800), the robust performance on the UCI dataset's lower priced goods ($0.55–$40.00) highlights the model's adaptability to varied consumer behaviors and price points Mitra et al. (2022).

Managerial Implications

The results suggest that retailers can meaningfully improve profitability by incorporating competitor price signals and demand prediction into dynamic pricing systems Gerpott and Berends (2022). The proposed framework enables platform independent pricing optimisation, allowing sellers to adapt prices without relying on proprietary marketplace repricing tools Guo et al. (2025). Five actionable insights emerge for e-commerce managers and pricing practitioners.

Intuition based pricing leaves significant profit on the table, and demand grounded pricing directly addresses this gap Antipov and Pokryshevskaya (2020). The proposed demand model explains over 98% of demand variation ($R^2 = 0.9877$) using 20 market signals universally available on Amazon, Flipkart, Shopify, WooCommerce, and other regional marketplaces, requiring no proprietary API integration. A seller with a $500 product generating 100 monthly sales at baseline can expect a potential profit improvement of 27.1%, equivalent to approximately $13,550 in additional annual profit at current margins Dritsas and Trigka (2025).

Elasticity estimates guide fundamentally different strategies across product categories Zhang and Meng (2025). Elastic categories such as Laptops & Computers ($\epsilon = -1.325$) require conservative pricing near competitive benchmarks to avoid demand erosion, while inelastic categories such as Wearables & Gadgets ($\epsilon = -0.06$) and Home & Kitchen ($\epsilon = -0.12$) support premium pricing strategies that expand margins without significantly reducing sales volume Das Varma and Vettas (2001).

The framework communicates elasticity classifications in plain language through the Streamlit dashboard, requiring no econometric expertise from the seller.

The competitor aware pricing band constrains all recommendations within 110% of the maximum observed competitor price, preventing margin eroding price wars Gerpott and Berends (2022). Across 68 evaluated products, this constraint achieved 100% market band compliance versus 84% under unconstrained pricing, with zero products experiencing greater than 10% profit loss compared to 13 out of 68 under standalone adaptive pricing Kastius and Schlosser (2022). In practice, sellers on Amazon or Flipkart can automate repricing without triggering retaliatory price reductions from competitors a critical advantage in high competition electronics categories.

The system delivers a complete price recommendation, profit estimate, and plain language explanation within 120 milliseconds, enabling use during live seller sessions on Amazon Seller Central, Flipkart Seller Hub, and Shopify Admin without perceptible delay. Sellers can query the system before adjusting prices during flash sales, festival seasons, or competitor promotions and receive an immediately actionable recommendation grounded in current market conditions.

The modular architecture accepts eleven universally available inputs and supports retraining on any product category or geographic market Eren et al. (2024); Tajik et al. (2024). Validation on the UCI Online Retail II dataset confirms positive profit improvement (+11.8%) in a fundamentally different retail context (low cost UK consumer goods, \$0.55–\$40.00), demonstrating that the framework generalises beyond the primary training domain Dua and Graff (2019). Practitioners can deploy the system on WooCommerce, Shopify, or custom platforms by retraining the CatBoost model on platform specific transaction data, with no changes to the optimisation or reinforcement learning components.

Conclusions and Future Work

This study shows that small and medium e-commerce sellers can potentially achieve mean profit improvements of 27.1% over baseline pricing under the evaluated experimental conditions through a demand grounded dynamic pricing framework deployable on any platform without technical expertise Antipov and Pokryshevskaya (2020). The research addresses a critical revenue management gap: existing automated repricing tools react to competitor prices without understanding demand, triggering margin eroding price wars Gerpott and Berends (2022). The proposed hybrid framework resolves this by anchoring every price recommendation in a demand forecast, constraining recommendations within competitive market bands, and delivering actionable guidance in plain language within 120 milliseconds.

The framework achieved 27.1% mean profit improvement ($p = 0.026$) across 68 products, with zero products experiencing greater than 10% profit loss compared to 13/68 under standalone adaptive pricing Kastius and Schlosser (2022). The demand model achieved $R^2 = 0.9877$ (MAE = 58.16), outperforming Linear Regression, Random Forest, and XGBoost Theodoridis and Tsadiras (2022). SHAP analysis confirmed product identity and month as dominant demand drivers, with competitor aware

features contributing 21.2% of total feature importance Hancock and Khoshgoftaar (2020).

Price elasticity analysis confirmed Laptops & Computers as the only elastic category (median $\epsilon = -1.325$), with Smartphones ($\epsilon = -0.44$), Wearables ($\epsilon = -0.06$), and Home & Kitchen ($\epsilon = -0.12$) exhibiting inelastic demand, providing category-level pricing flexibility guidance Zhang and Meng (2025); Das Varma and Vettas (2001).

The ablation study demonstrates that the competitor aware pricing band acts as an operational constraint balancing profit maximisation with competitive market realism Gerpott and Berends (2022). When this constraint is removed, raw profit gains increase to 59.4%, but the resulting recommendations exceed plausible competitive price levels. This highlights an important principle: unconstrained optimisation may maximise short term profit in simulation but often fails to produce deployable pricing policies in competitive marketplaces Besbes and Zeevi (2009). The full hybrid achieves 27.1% profit improvement with 25% lower variance than grid search alone, confirming that all four components contribute to reliable end-to-end pricing support Li and Xin (2024).

Cross dataset validation on the UCI Online Retail II dataset Dua and Graff (2019) confirms generalisation beyond the primary training domain, achieving +11.8% mean profit improvement across 400 records from a fundamentally different retail context (low-cost UK consumer goods, \$0.55–\$40.00), with 99.8% of records showing positive improvement Mitra et al. (2022).

Several limitations warrant acknowledgement. The primary dataset is limited to consumer electronics products; generalisation to other product categories such as fashion, grocery, or services requires further validation with category-specific transaction data. Competitor prices exhibit structural correlation with unit price ($r = 0.628$–0.634), which may produce optimistic performance estimates compared to live marketplace data Gerpott and Berends (2022). The PPO agent achieved partial convergence (explained_variance ≈ 0.13), and ensemble weights (0.7/0.3) were selected empirically rather than optimised systematically Schulman et al. (2017). Competitor prices in the UCI validation were approximated rather than sourced from live feeds, limiting the strength of the generalisation claim Dua and Graff (2019). These results contribute to the emerging literature on AI assisted pricing by demonstrating that demand grounded optimisation with reinforcement learning can produce stable and economically meaningful pricing recommendations for small scale online sellers Wu et al. (2023); Uprikar and Dixit (2025).

Future work will focus on five directions. First, the PPO observation space will be extended to include predicted demand as a direct state signal, enabling fully demand grounded reinforcement learning Schulman et al. (2017). Second, systematic ensemble weight optimisation will replace empirically selected weights with data driven configurations Li and Xin (2024). Third, online learning mechanisms will be incorporated to enable continuous adaptation to distribution shift in dynamic marketplace environments Liu (2024); Tajik et al. (2024). Fourth, the feature engineering pipeline will be extended to incorporate customer review scores, seller ratings, and inventory levels Kale et al. (2024); El Youbi et al. (2024). Fifth, validation on the Brazilian e-commerce dataset (Olist) will assess generalisation across broader geographic markets Gautam

et al. (2021). These advancements aim to position the framework as a robust, scalable, and practically deployable dynamic pricing platform. More broadly, the results suggest that practical AI driven pricing systems must incorporate competitive and behavioural constraints to ensure economically realistic recommendations Guo et al. (2025).

Funding

Funding not applicable.

Data availability

The data that support the findings of this study are publicly available at DOI: https://doi.org/10.5281/zenodo.19151881.

References

Andrade, Luiz Augusto C.G. and Claudio B. Cunha. 2023. Disaggregated retail forecasting: a gradient boosting approach. *Applied Soft Computing* 141: 110308. https://doi.org/10.1016/j.asoc.2023.110283.

Antipov, Evgeny A. and Elena B. Pokryshevskaya. 2020. Interpretable machine learning for demand modeling with high-dimensional data using gradient boosting machines and Shapley values. *Journal of Revenue and Pricing Management* 19: 355–364. https://doi.org/10.1057/s41272-020-00236-4.

Besbes, Omar and Assaf Zeevi. 2009. Dynamic pricing without knowing the demand function. *Operations Research* *57*(6): 1407–1420. https://doi.org/10.1287/opre.1080.0640.

Das Varma, Gopal and Nikolaos Vettas. 2001. Optimal dynamic pricing with inventories. *Economics Letters* *72*(3): 335–340. https://doi.org/10.1016/S0165-1765(01)00439-6.

Dritsas, Elias and Maria Trigka. 2025. Machine learning in e-commerce: trends, applications, and future challenges. *IEEE Access* 13: 99048–99067. https://doi.org/10.1109/ACCESS.2025.3572865.

Dua, Dheeru and Casey Graff. 2019. UCI machine learning repository: Online Retail II. University of California, Irvine. https://archive.ics.uci.edu/dataset/502/online+retail+ii.

Elreedy, Dina, Amir F. Atiya, and Samir I. Shaheen. 2021. Novel pricing strategies for revenue maximization and demand learning using an exploration–exploitation framework. *Soft Computing* 25: 11711–11733. https://doi.org/10.1007/s00500-021-06047-y.

El Youbi, Rachid, Fatima Messaoudi, and Mohammed Loukili. 2024. Machine learning-driven dynamic pricing strategies in e-commerce. In *Proceedings of the 14th International Conference on Information and Communication Systems (ICICS)*. IEEE. https://doi.org/10.1109/ICICS60529.2023.10330541.

Eren, Esma Ceren, Zhen Zhang, Jonas Rauch, and Arne K. Strauss. 2024. Revenue management without demand forecasting: a data-driven approach for bid price generation. *Journal of Revenue and Pricing Management* 23: 499–516. https://doi.org/10.1057/s41272-023-00465-3.

Gautam, Nitin, Shriguru Nayak, and Sergey Shebalov. 2021. Machine learning approach to market behavior estimation with applications in revenue management. *Journal of Revenue and Pricing Management* 20: 344–350. https://doi.org/10.1057/s41272-021-00317-y.

Gerpott, Torsten J. and Jan Berends. 2022. Competitive pricing on online markets: a literature review. *Journal of Revenue and Pricing Management* 21: 596–622. https://doi.org/10.1057/s41272-022-00390-x.

Guo, Xiaoming, Lei Zhang, and Qiang Wang. 2025. Dynamic pricing models in e-commerce: exploring machine learning techniques to balance profitability and customer satisfaction. *IEEE Access* 13: 6421–6435. https://doi.org/10.1109/ACCESS.2025.3563371.

Hancock, John T. and Taghi M. Khoshgoftaar. 2020. CatBoost for big data: an interdisciplinary review. *Journal of Big Data* 7: 94. https://doi.org/10.1186/s40537-020-00369-8.

Kale, Shubham, Pranav Patil, Darshan Patil, and Salil Koul. 2024. Dynamic pricing prediction with machine learning algorithm. In *Proceedings of the 5th International Conference on Emerging Technology (INCET)*. IEEE. https://doi.org/10.1109/INCET61516.2024.10593010.

Kastius, Alexander and Rainer Schlosser. 2022. Dynamic pricing under competition using reinforcement learning. *Journal of Revenue and Pricing Management* 21: 50–63. https://doi.org/10.1057/s41272-021-00285-3.

Li, Dandan and Jia Xin. 2024. Deep learning-driven intelligent pricing model in retail: from sales forecasting to dynamic price optimization. *Soft Computing* 28: 12281–12297. https://doi.org/10.1007/s00500-024-09937-z.

Liu, Hao. 2024. Reinforcement learning for e-commerce dynamic pricing. In *Proceedings of the 7th International Conference on Economic Management and Green Development*, Singapore, pp. 1926–1934. Springer. https://doi.org/10.1007/978-981-97-0523-8_183.

Mienye, Ibomoiye Domor and Yanxia Sun. 2022. A survey of ensemble learning: concepts, algorithms, applications, and prospects. *IEEE Access* 10: 99129–99149. https://doi.org/10.1109/ACCESS.2022.3207287.

Mitra, Arnab, Arnav Jain, Avinash Kishore, and Pravin Kumar. 2022. A comparative study of demand forecasting models for a multi-channel retail company: a novel hybrid machine learning approach. *Operations Research Forum* 3: 58. https://doi.org/10.1007/s43069-022-00166-4.

Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347. https://arxiv.org/abs/1707.06347.

Tajik, Mostafa, Behnam Malakouti Tosarkani, Ahmad Makui, and Rouzbeh Ghousi. 2024. A novel two-stage dynamic pricing model for logistics planning using an exploration-exploitation framework. *Expert Systems with Applications* 246: 123060. https://doi.org/10.1016/j.eswa.2023.123060.

Theodoridis, Georgios and Athanasios Tsadiras. 2022. Comparing boosting and deep learning methods on multivariate time series for retail demand forecasting. In *Artificial Intelligence Applications and Innovations*, Volume 647 of *IFIP Advances in Information and Communication Technology*, Cham, pp. 334–345. Springer. https://doi.org/10.1007/978-3-031-08337-2_31.

Uprikar, Siddharth and Amit Dixit. 2025. Reinforcement learning-optimized dynamic pricing system for Indian e-commerce platforms. In *Proceedings of the International Conference on AI in Business*. IEEE. https://doi.org/10.1109/ICRISET64803.2025.11253654.

Vanvuchelen, Nathalie, Joren Gijsbrechts, and Robert Boute. 2020. Use of proximal policy optimization for the joint replenishment problem. *Computers in Industry* 119: 103239. https://doi.org/10.1016/j.compind.2020.103239.

Wu, Chao, Wenjie Bi, and Haiying Liu. 2023. Proximal policy optimization algorithm for dynamic pricing with online reviews. *Expert Systems with Applications* 213: 119191. https://doi.org/10.1016/j.eswa.2022.119191.

Zhang, Wei and Yuan Meng. 2025. Dynamic modeling and forecasting of price elasticity based on time series analysis and machine learning. *Journal of Revenue and Pricing Management*. https://doi.org/10.1177/14727978251338001.

Yavuz, Tuğçe and Onur Kaya. 2024. Deep reinforcement learning algorithms for dynamic pricing and inventory management of perishable products. *Applied Soft Computing* 163: 111864. https://doi.org/10.1016/j.asoc.2024.111864.

Schlosser, Rainer and Kris Richly. 2019. Dynamic pricing under competition with data-driven price anticipations and endogenous reference price effects. *Journal*

of Revenue and Pricing Management 18(6): 451–464. https://doi.org/10.1057/s41272-019-00206-5.

Zhou, Qian, Yali Yang, and Shu Fu. 2023. Joint pricing and inventory control with reference price effects and price thresholds: a deep reinforcement learning approach. *Expert Systems with Applications* 233: 120993. https://doi.org/10.1016/j.eswa.2023.120993.

Talluri, Kalyan T. and Garrett J. van Ryzin. 1998. An analysis of bid-price controls for network revenue management. *Management Science* 44(11): 1577–1593. https://doi.org/10.1287/mnsc.44.11.1577.

Gallego, Guillermo and Garrett van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* 40(8): 999–1020. https://doi.org/10.1287/mnsc.40.8.999.

den Boer, Arnoud V. 2015. Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in Operations Research and Management Science* 20(1): 1–18. https://doi.org/10.1016/j.sorms.2015.03.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Highlights.pdf
- GraphicalAbstract.pdf
- InputOutputDoc.pdf