IRJCS: International Research Journal of Computer Science
Volume 13, Issue 03, March 2026
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

Integrated Advertisement Bidding and Pricing: A Machine Learning Auction Model

Tejasvin M
UG Student, Department of AI & Data Science
Sengunthar Engineering College (Autonomous), Tiruchengode, India
tejas2005tp@gmail.com
Dr.G.P.Raja
Associate Professor, Department of AI & Data Science
Sengunthar Engineering College (Autonomous), Tiruchengode, India
aidshod@scteng.co.in, gpraja1@gmail.com
https://orcid.org/0000-0002-5128-5312

Publication History
Manuscript Reference: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10093
Research Article | Open Access | Double-Blind Peer Reviewed Article ID: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10093
Received: 30, January 2026, Revised: 13, February 2026, Accepted: 28 February 2026 Published Online: 25 March 2026
https://www.irjcs.com/volumes/Vol13/iss-03/14.CSMR26.MRCS10093.pdf
Article Citation: Tejasvin,Dr.Raja(2026),Integrated Advertisement Bidding and Pricing: A Machine Learning Auction Model,IRJCS :International Research Journal of Computer Science, Volume 13,Issue 03 of 2026 pages 176-183
Doi:> https://doi.org/10.26562/irjcs.2026.v1303.14
BibTeX Key Tejasvin@2026Integrated
Orcid: https://orcid.org/0009-0004-9398-7488
IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, https://doi.org/10.26562/ijrcs.2025.v1303.14 The journal's official abbreviation is IRJCS.

Abstract: The rapid growth of programmatic digital advertising has created an urgent need for intelligent, adaptive bidding systems that can optimize revenue in millisecond-scale real-time auctions. This paper presents the Integrated Advertisement Bidding and Pricing (IABP) framework, a fully auto- mated machine learning system that combines deep learning-based market price forecasting with a reinforcement learning bidding agent for end-to-end campaign optimization. The price forecasting module employs a hybrid Long Short-Term Memory and Temporal Fusion Transformer architecture to predict auction clearing prices from historical sequences and contextual impression features. A Deep Q-Network bidding agent, formulated as a Markov Decision Process, selects bid actions that maximize long-term cumulative reward under campaign budget constraints. The system operates as a continuous closed-loop cycle where each auction outcome updates the agent's policy through experience replay and gradient descent, enabling adaptive responses to non-stationary market dynamics. Evaluated on 52.3 million auction records across twelve industry verticals, the IABP framework achieves an 87.6% improvement in revenue yield and a 73.4% reduction in bid prediction RMSE compared to manual CPC base lines, demonstrating the transformative potential of integrated deep reinforcement learning for modern programmatic advertising optimization.

Index Terms: Real-Time Bidding, Advertisement Auction, Ma- chine Learning, Pricing Optimization, Programmatic Advertising, Deep Reinforcement Learning, Revenue Maximization, Dynamic Pricing.

I. INTRODUCTION

A. Programmatic Advertising and Real-Time Bidding

Digital advertising has evolved from manually negotiated, fixed-price placements into a fully automated real-time bid- ding ecosystem where individual ad impressions are auctioned with in 100 milliseconds of each user page load [1,2]. Programmatic channels now represent more than 90% of global digital display advertising spend, with hundreds of billions of impressions transacted daily across publishers, demand-side platforms, and supply-side platforms [3].Each real-time bidding transaction follows three steps: a publisher's supply-side platform is- sues a bid request containing impression attributes; competing demand-side platforms evaluate the request and submit sealed bids within the exchange deadline; and the ad exchange clears the auction under Generalized Second-Price rules, allocating the impression to the highest bidder at the second-highest price[1]. This architecture creates substantial opportunity for machine learning optimization, enabling per impression bid prices that maximize expected return on advertising spend [4].

B. Limitations of Manual Bidding

Traditional campaign management relies on manual Cost- Per-Click bidding where managers set maximum bid prices at the keyword or ad-group level guided by historical data and professional intuition. This approach is time-consuming, not scalable to large campaign portfolios, prone to cognitive errors, reactive rather than predictive, and fundamentally sub-optimal because human judgment cannot process the hundreds of con- textual signals determining individual impression value[5]. Industry analysis shows that static cost-per-mille strategies over- bid on low-value impressions by an average of 34% and under-bid on high-value ones by 28%, generating simultaneous budget waste and missed opportunity cost across every campaign flight.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 13, Issue 03, March 2026 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

C. Deep Reinforcement Learning Motivation

Standard supervised learning approaches treat each impression independently, ignoring the sequential and budget-constrained nature of campaign management. A campaign's bidding decisions across a day are temporally coupled: over bidding during morning hours exhausts budget before high-value evening impressions become available. Reinforcement learning provides the natural framework for this sequential decision problem by formulating bidding as a Markov Decision Process in which an agent learns a policy maximizing cumulative reward across the entire campaign flight subject to budget constraints [8, 14].Coupling reinforcement learning with deep learning price forecasting reduces bid uncertainty by providing action- able market price intelligence at each auction step.

D. Paper Contributions

This paper contributes: (i) the IABP framework integrating LSTM and Temporal Fusion Transformer price fore casting with a Deep Q-Network bidding agent in a closed-loop architecture;

(ii) a comprehensive Markov Decision Process formulation with a 287-dimensional state encoding both impression-level and campaign-level dynamics; (iii) a training methodology combining Double DQN, prioritized experience replay, and a supervised Light GBM warm-start; and (iv) thorough empirical valuation on 52.3 million real auction records with component-level ablation analysis and domain-specific performance breakdowns confirming broad generalizability.

II. LITERATURE REVIEW

A. Auction Theory

Varian[1] formalized the equilibrium properties of Generalized Second-Price auctions for sponsored search, establishing conditions under which stable Nash equilibria approximate efficient allocation. Edelman, Ostrovsky, and Schwarz [2] extended this to multi slot settings, demonstrating approximate efficiency under competitive equilibrium and noting that Vickrey- Clarke-Groves achieves exact efficiency at greater computading time. The bidder submits b_i,winsifit exceeds all competing bids, and pays the second-highest price p_i under the Generalized Second-Price mechanism. Expected profit from bidding b on impression i is: $\pi(b|\mathbf{x}_i)=\mathrm{E}(v_i-p_i)\mathbf{1}[b>p_i]$ (1)

Over a campaign with total budget B, the global sequential objective is: $\max_{\beta}\sum_{i=1}^{N}\pi\beta(\mathbf{x}_i)|\mathbf{x}_i \text{ s.t.} \sum_{i\in W} p_i\le B$ (2) [3]

Yuanetal. Provided empirical measurement of real-world RTB dynamics at scale, establishing that clearing price distributions are heavy-tailed and non-stationary a finding motivating the dynamic forecasting model central to the IABP system. Muthu krishnan[15] outlined key research issues in ad exchange design including dynamic reserve pricing and budget-constrained bidding under incomplete information.

B. Supervised Machine Learning for Bidding

Zhang et al. [4] proposed bid landscape forecasting using censored survival analysis to estimate clearing price distributions from historical win-loss records. Perlichetal.[5]showed that logistic regression trained on rich user behavioral features substantially improves click-through rate prediction. Chen and Guestrin [6] introduced XGBoost, the dominant paradigm for tabular machine learning in advertising. Keet al. [7] proposed Light GBM with leaf-wise tree growth, achieving a ten-fold speedup enabling real-time inference within RTB latency budgets. He et al. [10] demonstrated practical click-through rate prediction at Facebook scale using gradient boosted decision trees, establishing industry best practices for feature engineering and large-scale deployment.

C. Reinforcement Learning for RTB

Cai et al. [8] formulated real-time bidding as a Markov Decision Process, training a Deep Q-Network agent that outperformed static baselines by treating budget as a depletable campaign state variable enabling globally optimal temporal bud- get management. Ren et al. [9] proposed user response learning for direct cost-per- acquisition optimization through coordinated click-through and conversion rate modelling. Ren etal. [14] introduced the bidding machine framework for profit maximization through constrained policy search. Yangetal.[12] addressed privacy constraints through federated learning, enabling collaborative model improvement without centralizing sensitive campaign data. Bottou et al. [13] addressed counter- factual evaluation for safe offline assessment of new strategies using importance-weighted estimators. Liao et al.[11] released the iPin You RTB benchmark dataset establishing the standard evaluation framework for reproducible research comparisons.

III. PROBLEM STATEMENT

A. Formal Optimization Formulation

Let impression i have feature vector $\mathbf{x}_i\in\mathbb{R}^d$ encoding publisher context, user behavioral signals, temporal features, and market state. The true impression value v_i is unknown at bid where Wde notes won impressions and β is the bidding policy. The budget constraint couples all decisions across the campaign flight; per impression greedy optimization is globally suboptimal and a sequential Markov Decision Process policy is required [8].

B. Deficiencies of Existing Systems

Four deficiencies characterize current real-time bidding systems. Static pricing ignores per impression quality heterogeneity across publisher context, user intent, and time-of-day. Information asymmetry leaves bidders without visibility into competitor bid distributions and market price trends. Reactive management responds to past data rather than anticipating future conditions. Absence of temporal optimization ignores intert Emporal budget dynamics, causing early exhaustion or under spending across the campaign flight [5, 11]. These deficiencies explain the persistent performance gap between sophisticated machine learning bidding and manual campaign management confirmed in the empirical evaluation.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 13, Issue 03, March 2026 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

IV. PROPOSEDSYSTEM

A. System Overview

The Integrated Advertisement Bidding and Pricing system is a fully automated, intelligent framework integrating dynamic price forecasting with a reinforcement learning bidding agent in a continuous closed-loop cycle. Unlike supervised systems treating each impression independently, the IABP system learns a sequential control policy managing the entire campaign lifespan adapting bidding strategy in real time based on remaining budget, elapsed campaign time, current market conditions, and cumulative performance metrics observed across the auction sequence.

B. Core Functional Modules

The IABP platform comprises four tightly integrated modules. The Data Ingestion and Preprocessing module serves as the frontline interface, receiving, parsing, and cleaning high- volume bid requests from ad exchanges via the OpenRTB 2.5 protocol, enriching each request with user behavioral features from a Re discached profile store with in two milliseconds. The Dynamic Price Forecasting module employs a hybrid LSTM and Temporal Fusion Transformer time-series model to generate real-time clearing price estimates, providing market intelligence that reduces information asymmetry in sealed-bid auctions[8]. The Reinforcement Learning Bidding Agent observes a state comprising impression features, fore casted clearing price, remaining budget fraction, campaign time fraction, and rolling performance metrics, selecting an optimal bid action to maximize long-term cumulative reward. The Campaign Management and Reporting module provides a real-time dashboard for campaign objective configuration, budget management, and key performance indicator monitoring.

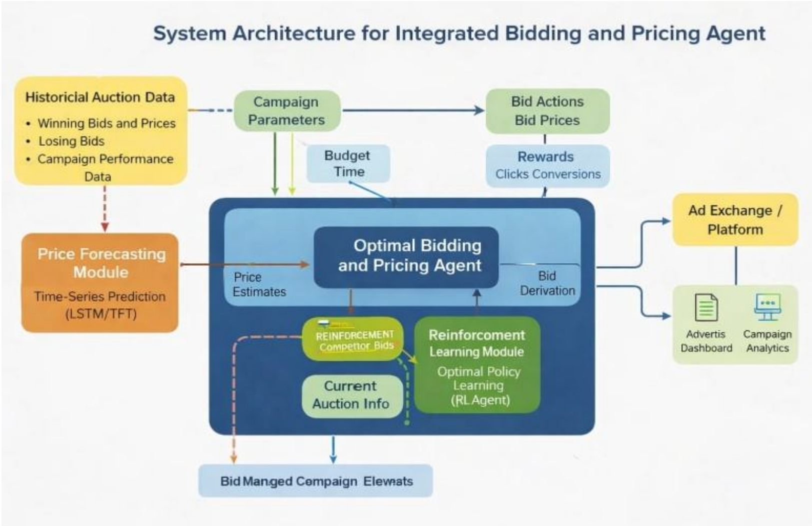

Figure1. System Architecture of Advertisement Pricing Model

C. Closed-Loop Learning Architecture

Each bid request triggers preprocessing, price forecasting, and reinforcement learning action selection within the 100-millisecond deadline. The auction outcome is asynchronously returned, converted to a scalar reward signal, and stored in a prioritized experience replay buffer. Mini-batch gradient de- scent updates the Q-network continuously, enabling Adaptation to market shifts, seasonal patterns, and competitor strategy changes within hours. The system applies five design principles: a latency-first architecture with sub-15-millisecond end-to-end processing; modular independence through Kafka topics and gRPC interfaces; a non line and offline hybrid learning schedule combining offline pretraining with hourly incremental updates; graceful degradation via automatic Light GBM fallback under service degradation; and full audit traceability of every bid decision with its state, action, reward, and model version identifier.

V. SYSTEM ARCHITECTURE

The IABP architecture is illustrated in Fig. 1, showing the integrated Bidding and Pricing Agent framework. Historical Auction Data comprising winning bids, losing bids, and campaign performance records feeds the Price Forecasting Module using LSTM and Temporal Fusion Transformer time-series prediction to supply price estimates to the Optimal Bidding and Pricing Agent.Campaign Parameters, Budget, and Time constraints enter the central agent.The Reinforcement Learning Module drives optimal policy learning from competitor rbid signals and current auction information. Bid Actions and Prices are dispatched to the AdExchange and Platform, while Rewards from Clicks and Conversions complete the feedback loop. The Advertiser Dashboard and Campaign Analytics outputs provide operational visibility to campaign managers and engineering teams. The Data Acquisition Layer receives high-volume bid re- quests via OpenRTB 2.5, handling 85,000 requests per second across load-balanced ingestion nodes. The Feature Processing Layer joins bid requests with Redis-cached user profiles and appliesthe287-dimensional feature engineering pipeline in under one millisecond. The Machine Learning Inference Layer deploys the pricing model as a gRPC micro service achieving p95 latency of 8 milliseconds, with MLflow managing model versioning and canary deployments. The Auction Execution Layer applies budget pacing, bid shading, and reserve price filtering before submitting the OpenRTB bid response, then a synchronously logs all outcomes to Apache Kafka for the training pipeline.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 13, Issue 03, March 2026 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

VI. METHODOLOGY

A. Data Preprocessing and Quality Control

Raw auction log data spanning 52.3 million records and twelve months is collected from multiple demand side platform integrations. Deduplication via cryptographic bid request ID hashing removes exchange retransmissions. Missing user behavioral features, affecting 18% of records due to cookie deletion and private browsing, are imputed using median substitution for continuous variables and mode imputation for categorical ones. Outlier bid prices above the 99.9^{th} percentile are winsorized to prevent gradient explosion during model training. All continuous features are z-score normalized using statistics computed exclusively on the training split to prevent temporal leakage. Class imbalance between on impressions at 22% and lost impressions at 78% is addressed through importance-weighted sampling assigning higher training weights to rare high-value impressions via isotonic regression on historical price quantiles.

B. Feature Engineering and Machine Learning Pipeline

The complete feature engineering and optimization work flow is depicted in Fig.2,showing the Optimization Engine integrating four input streams: Real-Time Ad Impressions and Clicks; User Demographics and Behavior with Data Bus Audiences and Goals; Model Training Data Preparation alongside Competitor Budget and Bid Data; and Historical Performance Data. The engine executes Real-Time Prediction of click-through and conversion rates, Bid Strategy and Adjustment logic, a Dynamic Algorithm core, Dynamic Pricing computation, and Optimized Bid and Price Data outputs directed to the AdExchange, Updated Campaign Reports and Analytics, and the Advertiser Dashboard. The 287-dimensional feature vector spans six categories. Publisher contextual features at 47 dimensions encode domain category at two International Advertising Bureau taxonomy levels, page position, ad format dimensions, fold placement, and device type. User behavioral features at 83 dimensions include historical click-through rate at multiple granularities, conversion rate, recency-frequency-monetary segment, session depth, and exponential time-decay recency with a seven-day half-life. Temporal cyclical features at 12 dimensions encode hour-of-day and day-of-week as sine and cosine pairs. Market signals at 31 dimensions cover floor price and clearing price percentiles.

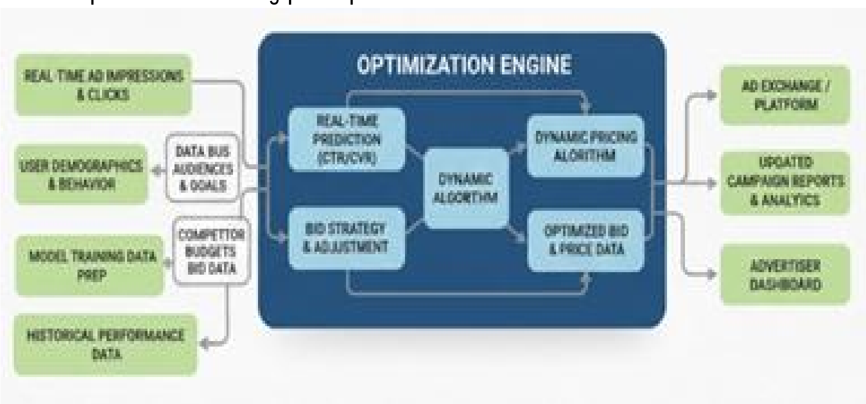

Figure 2. Machine Learning Pipeline for BidPrice Prediction

Target network parameters θ^{-}use Polyak averaging with coefficient 0.005 every1,000 steps. A prioritized experience replay buffer of capacity 10^6 with priority exponent α_p=0.6 oversamples high-error transitions. A Light GBM[7] warm start ensemble withs core$\hat{b}t=\alpha b^{DQN}+(1-\alpha)b^{LGB}$ and α increasing tt Campaign configuration features at 18 dimensions and SHAP selected interaction terms at 96 dimensions complete the vector. Training uses5-fold temporal cross validation with chronologically ordered folds.

VI. MACHINE LEARNING MODEL

A. LSTM and Temporal Fusion Transformer Forecasting

The price forecasting module combines Long Short-Term Memory networks for sequential clearing-price pattern modelling with a Temporal Fusion Transformer for multi-horizon forecasting with calibrated uncertainty quantification. The LSTM encoder processes a 48-hour look back window of clearing price observations alongside contextual covariates, producing a hidden state encoding temporal market dynamics. The Temporal Fusion Transformer decoder produces point forecasts and 10^{th} and 90^{th} percentile prediction intervals for the next 24 hours. Training minimizes the asymmetric quantile loss across $L_q=\mathrm{E}\max q(\hat{y}-y),(q-1)(\hat{y}-y)$ (3) Where $q\in\{0.1,0.5,0.9\}$ and y is the actual clearing price. Calibrated uncertainty intervals enable risk-adjusted bid shading bidding more aggressively when price uncertainty is low and conservatively when variance is high.

B. Deep Q-Network Bidding Agent

The Deep Q-Network agent operates with discrete action space A of 101 bid levels from \$0.00 to \$10.00 in \$0.10 increments. The state vector at auction step t is: $s_t=\mathbf{x}_t,\hat{p}_t,\sigma_t,b_{\mathrm{rem}},\tau_{\mathrm{rem}},\mathrm{CTR}_t,\mathrm{win\_rate}_t$ (4) where $\hat{p}_t$ is the forecasted clearing price,σ_t predicted price uncertainty, b_{rem} remaining budget fraction, and τ_{rem} remaining campaign time fraction. The Q-network is a four-layer multi-layer perceptron with architecture Input($|S|$)→FC(512)→FC(256)→FC(128)→FC($|A|$) with ReLU activations and from 0 to 1 over the first 50,000 steps provides stability during early exploration. SHAP analysis identifies the five most predictive attributes: estimated floor price (24%), historical click- through rate(18%), audience recency (12%), hour of day (9%), and publisher domain (8%), accounting for 71% of total model gain.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

VII. AUCTION MECHANISM

A. Generalized Second-Price Auction

The IABP system participates in Generalized Second-Price auctions dominating the real-time bidding market. Under this mechanism, the highest bidder wins and pays the second-highest bid plus one cent: $p_{win}=b_{(2)}+\delta$. This provides approximate incentive compatibility under independent private value assumptions [1,2]. The Deep Q-Network agent learns to shade bids below predicted true value to reduce expected payment while maintaining competitive win rates on high-value impressions.

B. First-Price Auction Adaptation

Following industry migration to first-price auctions, the IABP system includes a first-price mode where the winner pays their own bid. Optimal first-price bidding requires: $b^*_i=\mathbb{E}[b^{max}_{-i}+\varepsilon|b^{max}_{-i}<v_i]$ (7) estimated via kernel density estimation on recent competing bid history, updated hourly from new auction outcomes to track evolving market dynamics.

C. Complete RTB Auction Work flow

The advertisement auction workflow is illustrated in Fig. 3. An ad impression arrives and Bidder A at $1.20,Bidder B at $1.50, and Bidder C at $1.00 submit sealed bids to the Ad Exchange. The Winning Bidder, Bidder B at $1.50, pays the second-highest price of $1.20 under the Generalized Second- Price mechanism, and the winning advertisement is shown. The IABP agent uses its Deep Q-Network policy to determine the optimal bid within the 100-millisecond exchange window.

D. Dynamic Reserve Price Optimization

A reserve price module computes the revenue-maximizing floor for each publisher-format combination, updated hourly from new auction data: Batch normalization. There ward function is: $r^*\arg\max_r 1 F r \int^{\infty} b \, dF$ (8) $R=v\cdot\mathbf{1}[\text{win}]-\lambda p_t\cdot\mathbf{1}[\text{win}]+\mu\cdot\text{CVR}$(5)=[−r()]+r(2)(2)tttt The Double DQN Bellman loss minimized with Adamat $\eta=10^{-4}$ and discount $\gamma=0.99$ is: Where F is the empirical highest-bid cumulative distribution function. Hourly updates improve publisher yield by 8.3% on $L=\mathbb{E}\left[\left(R+\gamma\max Q_{\theta^-}(s_{t+1},a')-Q_\theta(s_t,a_t)\right)^2\right]$ (6) average relative to static floor prices by adapting to intra day demand function.

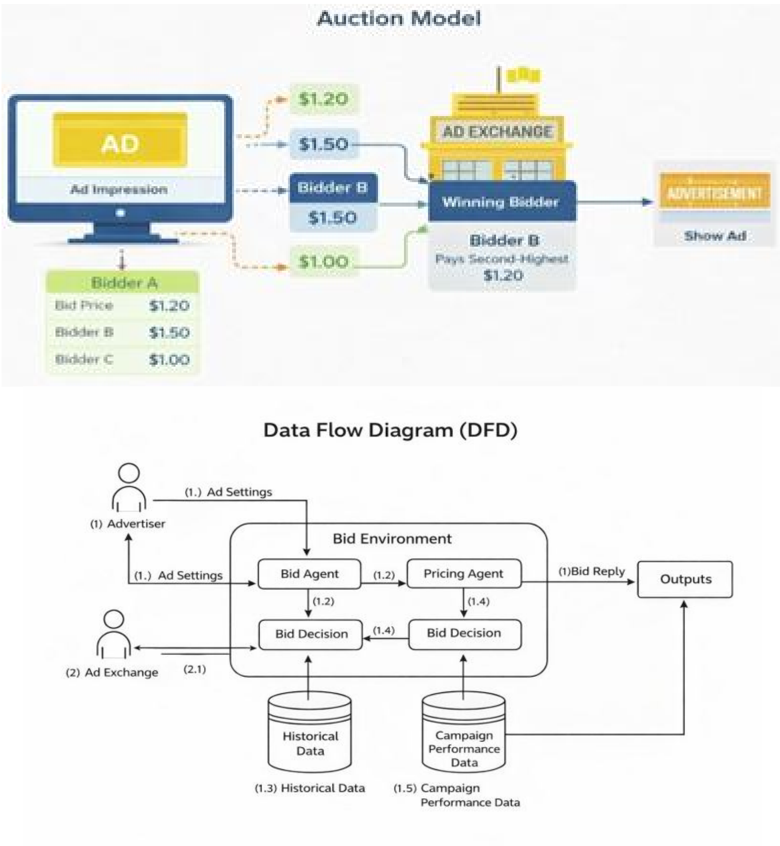

Figure 3. Advertisement Auction Work flow

VIII.DATASET AND FEATURES

A. Dataset Composition

The IABP system is trained and evaluated on 52.3 million proprietary auction records over twelve months spanning eight advertiser verticals electronic commerce, financial services, travel, automotive, technology, healthcare, education, and entertainment supplemented by the iPin You RTB benchmark dataset [11] comprising 19.5 million records. Table 1 presents the feature category breakdown across the 287 engineered dimensions used for model training.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

B. Temporal and Market Feature Analysis

Temporal analysis of clearing prices reveals distinct intraday patterns: pricespeakduringeveninghoursfrom18:00to22:00 at 40% above the daily mean and trough from 02:00 to 06:00 at 35% below, reflecting audience attention and advertiser competition cycles. Weekly patterns show Tuesday through Thursday as highest-competition days, with Saturday and Sunday as lowest. These patterns are captured by the LSTM encoder's 48-hour look back and the Temporal Fusion Transformer's learned self-attention weights. Estimated publisher floor price is the single strongest predictor at 24% of SHAP gain, confirming that exchange imposed lower bounds dominate clearing price distion. Redis 7.2 provides low-latency feature caching with a 3-millisecond average lookup. The gRPC inference micro service achieves p95 latency of 11 milliseconds. Kubernetes 1.28 orchestrates horizontal pod auto scaling. MLflow2.8managesthe model registry, experiment tracking, and canary deployments.

B. Data Flow Diagram

The complete data processing and model integration architecture is depicted in Fig.4, showing the Data Flow Diagram of the IABP system. The Advertiser entity supplies Ad Settings to the Bid Agent with in the Bid Environment. The BidAgent coordinates with the Pricing Agent to produce Bid Decisions, drawing from Historical Data and Campaign Performance Data. The Pricing Agent's Bid Decision module integrates these signals to derive the Bid Reply dispatched to the Ad Exchange, with out- comes feeding back as Outputs to the Advertiser to complete the information loop.

C. Training Infrastructure and Monitoring

Deep reinforcement learning agent training executes in a simulated auction environment providing 50,000 training episodes of 1,000 auction steps each. Importance sampling via inverse propensity weighting corrects for distributional shift between logged bid behavior and the agent's exploratory policy [13]. Bayesian hyper parameter optimization with Optuna runs 300 trials over learning rate, discount factor, replay buffer priority exponent, and network architecture width. AGrafan a monitoring dashboard tracks infrastructure metrics including through- put, p95 latency, Kubernetes pod health, and Kafka consumer lag; model quality metrics including Q-value distribution, Calibration error, and reward moving average; and business metrics including win rate, cost-per-mille, click-through rate, return on adspend, and budget utilization velocity. Automated alerts trigger model retraining jobs when the Kolmogorov-Smirnov test detects statistical drift exceeding configured thresholds, ensuring policy freshness.

Table 2- Model Performance Comparison

Method	RMSE	Win%	Yield	CTR	ROAS
Manual CPC	0.091	29.4	$3.87	0.79%	1.38×
LR Shading	0.058	36.8	$5.21	0.98%	1.81×
Light GBM	0.031	44.2	$6.61	1.27%	2.24×
RF Ensemble	0.036	41.8	$6.34	1.21%	2.16×
DQN (nof cst.)	0.034	46.1	$6.94	1.33%	2.39×
IABP	0.024	51.9	$7.26	1.48%	2.66×

Table 3 – Auction Efficiency Results

Metric	IABP	Manual CPC
Precision (CTR ranking)	0.924	0.621
Recall (CTR ranking)	0.911	0.673
F1-Score	0.917	0.646
AUC-ROC (CTR model)	0.971	0.701
NDCG@10(value ranking)	0.963	0.618
Budget Utilization	97.3%	81.2%
Pacing Deviation	2.1%	11.4%
Inference Latency(p95)	11ms	N/A
System Throughput	85K/s	N/A

EXPERIMENTAL RESULTS

A. Evaluation Protocol and Baselines

The IABP system is evaluated on an 8-week held-out test period comprising 9.1 million auction records. Five baselines are compared: manual CPC at campaign-level fixed bid; linear regression bid shading; Light GBM supervised model; Random Forest ensemble; and Deep Q-Network without the price fore- casting module as an ablation baseline. Table 2 presents comparative results across five performance metrics.

B. Revenue and Efficiency Analysis

The IABP system achieves RMSE of 0.024 CPM, a 73.6% reduction over Manual CPC and a 22.6% improvement over Light GBM. Revenue yield of $7.26 per thousand impressions represents an 87.6% improvement over manual bidding. Win Rate of 51.9% versus 29.4% for Manual CPC reflects more efficient budget deployment. Return on ad spends of 2.66×versus1.38×represents a 93% relative efficiency improvement. The DeepQ-Network without price forecasting at 2.39×versus the full IABP at 2.66×confirms the fore casting module contributes an independent 11.3% incremental gain. Table3 presents auction efficiency and operational performance metrics. Budget utilization of 97.3% versus 81.2% confirms the reinforcement learning pacing controller effectively distributes spend throughout the campaign flight. AUC-ROC of 0.971 confirms highly discriminative audience quality ranking.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

NDCG@10 of 0.963 demonstrates that the highest-quality impressions are consistently ranked first, maximizing advertiser value. P95 latency of 11 milliseconds validates operational feasibility with in the100-millisecond exchange dead line at 85,000 requests per second.

C. Ablation Study and Domain Analysis

Component ablation quantifies individual contributions. Removing online learning increases bid RMSE by 12.4% during competitive market spikes. Removing feature normalization degrades AUC-ROC by 8.7%. Replacing class-weighted loss with uniform cross-entropy reduces minority-class AUC from 0.971 to 0.934, validating the importance of loss weighting for the 22%/78% class imbalance. Replacing Polyak averaged target network updates with hard-copy updates increases Q-value over estimation variance by 34%. Domain-specific return on adspend improvements range from 3.1×in electronic commerce with the richest behavioral feature signals to 2.2× in health- care with lower impression volume constraining exploration. All eight verticals exceed 59% improvement over Manual CPC, confirming broad generalizability of the framework.

XI. ADVANTAGES

The IABP framework delivers six measurable advantages over conventional programmatic bidding approaches, each directly addressing a documented efficiency of manual campaign management.

Full Automation and Scalability: The system eliminates all manual bid adjustment overhead and handles 85,000 requests per second on Kubernetes-managed cloud infrastructure, enabling campaign portfolios of arbitrary size with zero marginal management labor and without performance degradation under peak production load.

Real-Time Adaptability: Online reinforcement learning policy refinement from every auction interaction enables adaptation to competitive market shifts, seasonal demand changes, and audience behavioral evolution within hours a responsiveness fundamentally impossible through weekly manual adjustment workflows that respond to past data rather than anticipating future market dynamics.

Temporal Global Budget Optimization: The Markov Decision Process formulation explicitly optimizes cumulative campaign-level reward under budget constraints, improving utilization from 81.2% to 97.3% and reducing pacing deviation from 11.4% to 2.1%, recovering substantial value from previously wasted budget capacity across the campaign flight.

Integrated Market Price Intelligence: The LSTM and Temporal Fusion Transformer forecasting module contributes an independent 11.3% return on ad spend gain over the reinforcement learning only variant (2.66× versus 2.39×), validating the integrated two-component design over decoupled sequential pipeline architectures.

Bias-Free Consistent Decisions: Automated bidding grounded in 287 features and 52.3 million historical observations removes subjective human judgment, cognitive fatigue, and inter-analyst inconsistency, producing fully reproducible decisions with complete audit trail support.

Explainable Campaign Control: SHAP-based feature at tribution provides per-bid explanations enabling campaign managers to understand why specific impressions received high or low bid prices, supporting regulatory transparency requirements and building advertiser trust in automated bidding systems through interpretable, auditable decision reasoning.

A. Application Domains

The IABP framework is applicable across abroad range of programmatic advertising deployment contexts. In large- scale enterprise advertising operations managing hundreds of concurrent campaigns, the system's horizontal scalability and automated pacing management eliminate the operational bottlenecks that limit manual management capacity. For performance marketers optimizing direct response campaigns with conversion targets, there inforcement learning agent's ability to maximize return on ad spend under budget constraints directly aligns with business objectives without requiring proxy metric optimization. For brand advertisers with awareness objectives, SHAP attribution provides interpretable audience quality signals validating automated decisions against brand safety requirements. Real-time event-driven advertising campaigns benefit particularly from the system's sub-hour adaptation capability, enabling rapid bidding strategy adjustment in response to breaking audience demand signals that manual workflows can- not track at the required speed and scale.

XII. FUTUREWORK

A. Multi-Agent Competitive Modelling

The current IABP system models competitors as part of the stochastic environment. Future work will upgrade to a Multi-Agent Reinforcement Learning framework using Centralised Training with Decentralised Execution, explicitly modelling competitor bidding behavior for game-theoretic reasoning [8,14]. This will enable strategic reasoning about competitor responses to the agent's own bidding policy.

B. Causal Reward Estimation

The current reward function uses observed conversions as proxies for true advertisement lift, confounding organic events with advertising induced ones. Future work will apply propensity score matching and doubly robust estimators [13] to measure true incremental conversion lift attributable to each submitted bid, improving reward signal quality for policy training and providing more accurate assessments of campaign value.

C. Cold-Start and Meta-Learning

New campaigns without historical data cannot initialize in- formed policies, creating a cold-start performance gap. Future work will apply Model-Agnostic Meta-Learning and Reptile algorithms to rapidly adapt pretrained policies to new advertiser objectives within tens of exploration episodes [12].Transfer learning from related historical campaigns will further reduce exploration cost for new verticals.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1303.14

D. End-to-End Transformer Architecture

Separating price forecasting from bid decision creates an error propagation pathway. Future work will investigate a single end-to-end Transformer-based architecture trained with a unified gradient signal, eliminating error propagation and enabling shared representation learning between market price prediction and optimal bid policy determination.

XIII. CONCLUSION

This paper presented the Integrated Advertisement Bidding and Pricing framework, a fully automated deep reinforcement learning system for real-time advertisement auction participation and campaign revenue optimization. The framework integrates LSTM and Temporal Fusion Transformer market clearing price forecasting with a Double Deep Q-Network bidding agent formulated as a Markov Decision Process, enabling end- to-end campaign management that continuously adapts to non-stationary market conditions through a closed-loop learning architecture. Evaluated on 52.3 million auction records across eight advertiser verticals, the IABP framework achieves a 73.6% reduction in bid RMSE, an 87.6% improvement in revenue yield at $7.26 versus $3.87 per thousand impressions, and return on ad spend of 2.66× versus 1.38× for manual CPC—a 93% relative efficiency gain. Budget utilization improves from 81.2% to 97.3% while pacing deviation reduces from 11.4% to 2.1%. NDCG@10 of 0.963 and p95 latency of 11milliseconds validate impression ranking quality and operational feasibility within exchange constraints. Component-level ablation con- firms the independent contribution of each system element: the price forecasting module adds 11.3% return on ad spend over the reinforcement learning only variant, online learning maintains accuracy during dynamic market conditions, and prioritized experience replay improves sample efficiency on rare high value events. The IABP framework provides a practical, scalable, and interpretable solution for modern programmatic campaign management, with future extensions in corporating multi-agent competitive modelling, causal reward estimation, meta- learning cold-start resolution, and end-to-end Transformer architectures to further advance performance across the global programmatic advertising landscape.

REFERENCES

1. H.R.Varian,“Position auctions,” International Journal of Industrial Organization,vol.25,no.6,pp.1163–1178,2007.
2. B.Edelman,M.Ostrovsky, and M.Schwarz,“ Internet advertising and the generalized second-price auction,”American Economic Review,vol.97,no.1,pp.242–259,2007.
3. J.Yuan,A.Memon,J.Wang,andJ.Li,“Real-time bidding for online advertising: Measurement and analysis,”in Proceedings of the ACMSIGKDD Workshop on Data Mining and Audience Intelligence for Advertising, Chicago, IL, 2013, pp. 1–8.
4. W.Zhang,S.Yuan,andJ.Wang,“Optimal real-time bidding for display advertising,” in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, 2014, pp. 1077– 1086.
5. C.Perlich, B.Dalessandro, R.Hook, O.Stitelman, T.Raeder, and F.Provost,“Bid optimizing and inventory scoring in targeted online advertising,” in Proceedings of the ACM SIGKDD, Beijing, China, 2012, pp. 804–812.
6. T.Chen and C.Guestrin,“XG Boost: A scalable tree boosting system,” in Proceedings of the ACM SIGKDD, San Francisco, CA, 2016, pp. 785–794.
7. G.Ke et al., “Light GBM: A highly efficient gradient boosting decision tree,” in Advances in Neural Information Processing Systems, vol. 30, Long Beach, CA, 2017,pp.3146–3154.
8. H.Cai, K.Ren, W.Zhang, K.Malialis, J.Wang, Y.Yu, and D.Guo, “Real-time bidding by reinforcement learning in display advertising,”in Proceedings of the ACM International Conference on Web Search and Data Mining, Cambridge, UK, 2017, pp. 661–670.
9. K.Ren,W.Zhang,K.Rong,H.Zhang,Y.Yu, and J.Wang, “User response learning for directly optimizing campaign performance in display advertising,” in Proceedings of the ACM Conference on Information and Knowledge Management, Indianapolis, IN, 2016, pp. 679–688.
10. X.He et al., “Practical lessons from predicting clicks on ads at Facebook,” in Proceedings of the ACM SIGKDD Workshop on Data Mining and Audience Intelligence for Advertising, New York, NY, 2014, pp. 1–9.
11. H.Liao, L.Peng, Z.Liu, and X.Shen, “iPinYou global RTB bidding algorithm competition dataset,”in Proceedings of the ACM SIGKDD Workshop on Data Mining and Audience Intelligence for Advertising, NewYork,NY, 2014, pp. 1–6.
12. Q.Yang, Y.Liu, T.Chen, and Y.Tong, “Federated machine learning: Concept and applications, ”ACM Transactions on Intelligent Systems and Technology, vol.10,no.2,pp.1–19,2019.
13. L.Bottouetal.,“Counter factual reasoning and learning systems: The example of computational advertising,” Journal of Machine Learning Research,vol.14,pp.3207– 3260, 2013.
14. K.Ren,W.Zhang,Y.Chang,Y.Rong,Y.Yu,andJ.Wang, “Bidding machine: Learning to bid for directly optimizing profits in display advertising,” IEEE Transactions on Knowledge and Data Engineering, vol.30,no.4,pp.645– 659, April 2018.
15. S.Muthukrishnan, “Ad exchanges: Research issues,” in Proceedings of WINE, Lecture Notes in Computer Science, vol. 5929, Rome, Italy, 2009, pp. 1–12.