Reinforcement Learning for Budget and Bid Optimization in Online Ad Auctions: Methods and Applications

Arun Thomas
Purdue University, Indiana, USA

doi: https://doi.org/10.37745/ejcsit.2013/vol13n442436 Published June 26, 2025

Citation: Thomas A. (2025) Reinforcement Learning for Budget and Bid Optimization in Online Ad Auctions: Methods and Applications, *European Journal of Computer Science and Information Technology*, 13(44),24-36

Abstract: *Reinforcement learning has emerged as a transformative solution for optimizing bidding strategies and budget allocation in online advertising auctions. The dynamic nature of these auctions, characterized by rapid market changes and complex user behaviors, necessitates sophisticated decision-making mechanisms beyond traditional rule-based systems. By leveraging advanced machine learning techniques, including contextual bandits, Deep Q-learning networks, and actor-critic architectures, modern advertising platforms can achieve significant improvements in campaign performance and return on investment. The implementation of these systems requires careful consideration of practical challenges, including reward shaping, delayed feedback handling, and counterfactual estimation. Through effective feature engineering and model architecture optimization, these challenges can be addressed while maintaining computational efficiency and system reliability. The integration of emerging technologies, such as multi-agent systems and transfer learning, continues to push the boundaries of what's possible in automated advertising optimization, promising even greater improvements in targeting accuracy and campaign effectiveness.*

Keywords: Real-time bidding optimization, reinforcement learning automation, delayed feedback mechanisms, multi-agent advertising systems, computational efficiency optimization

INTRODUCTION

The dynamic nature of online advertising auctions represents a complex ecosystem where traditional optimization approaches often fall short of achieving optimal performance. According to recent research in recommender systems, the integration of deep reinforcement learning (DRL) with online advertising has demonstrated remarkable improvements in click-through rate (CTR) prediction and conversion rate optimization. Studies have shown that DRL-based approaches can achieve up to 44.7% improvement in rewards compared to traditional methods, with specific implementations showing a 3.2% increase in CTR and a 2.8% enhancement in conversion rates [1].

The complexity of real-time bidding (RTB) environments presents unique challenges that demand sophisticated solution approaches. Research has demonstrated that reinforcement learning strategies in RTB scenarios can effectively optimize both immediate advertising performance and long-term budget allocation. Experimental results from large-scale online advertising platforms have shown that RL-based bidding strategies can achieve up to 23.4% improvement in advertising performance compared to conventional bidding strategies. These improvements are particularly significant given that the RTB market processes billions of ad impression opportunities daily, with each auction requiring bid decisions within milliseconds [2].

The advancement in DRL methodologies has introduced novel approaches to handling the exploration-exploitation dilemma in online advertising. Recent implementations utilizing Deep Q-Network (DQN) architectures have shown promising results in balancing immediate rewards with long-term benefits. Studies indicate that DRL models can process complex state spaces incorporating over 16 different features, including user demographics, historical behavior patterns, and contextual information. These models have demonstrated the ability to maintain stable performance across varying budget constraints, with some implementations showing consistent performance improvements of 20-35% in terms of return on investment (ROI) compared to traditional methods [1].

Real-world deployments of RL systems in advertising contexts have revealed significant challenges in handling delayed rewards and attribution. Research has shown that sophisticated reward modeling approaches can effectively address these challenges, with some implementations reducing attribution errors by up to 18.7%. The integration of these systems with existing advertising infrastructures has required careful consideration of computational efficiency, with successful implementations processing up to 10,000 queries per second while maintaining response times under 100 milliseconds [2].

The scale of modern advertising platforms presents unique opportunities for reinforcement learning applications. Studies have shown that RL-based systems can effectively manage daily budgets ranging from thousands to millions of dollars, with some implementations demonstrating the ability to optimize spending across more than 100 concurrent campaigns. The most successful implementations have shown the ability to adapt to market changes within 2-3 hours, compared to traditional systems that might require days or weeks for optimization [1].

Understanding the Challenge

Online ad auctions operate in an environment of exceptional volatility where market conditions demand sophisticated real-time decision-making capabilities. Recent industry analysis reveals that programmatic advertising platforms process an average of 30 billion bid requests daily, with peak volumes reaching up to 1.5 million queries per second. Machine learning approaches to real-time bidding (RTB) have demonstrated the ability to reduce Cost Per Acquisition (CPA) by up to 32% while increasing conversion rates by 27% compared to traditional rule-based systems. These improvements become particularly significant when considering that bid prices must be calculated within 100 milliseconds to participate in the auction effectively [3].

The complexity of modern advertising ecosystems is further highlighted by the challenges in viewability optimization and inventory quality assessment. Industry data shows that advertisers using real-time viewability data can achieve up to 75% viewability rates, compared to the industry average of 50%. This significant improvement is attributed to the ability to make instantaneous adjustments to bidding strategies based on real-time performance metrics. Furthermore, campaigns utilizing advanced optimization techniques have shown the potential to reduce wasted ad spend by up to 40% through better targeting and bid adjustments [4].

A critical aspect of understanding RTB challenges involves examining how user engagement and conversion metrics change based on the time window following ad impressions. As shown in Table 1, engagement rates are highest (58%) within the first 24 hours after an impression, with similar patterns observed for conversion rates (2.4%) and view-through rates (75%). The efficiency of budget utilization also follows this pattern, with optimal spending efficiency (65%) occurring within the immediate 24-hour window.

This time-sensitive nature of user response presents significant challenges for RTB systems, as delayed feedback mechanisms must account for both immediate and longer-term engagement patterns. The sharp decline in engagement metrics after the 72-hour mark (dropping to 20% engagement rate and 0.8% conversion rate) underscores the importance of timely optimization and the difficulties in attributing long-tail conversions to specific impressions [3, 4].

Table 1.Time-Based Performance Metrics in Real-Time Bidding Environments [3, 4]

Time Window	Engagement Rate (%)	Conversion Rate (%)	Budget Utilization (%)	View-through Rate (%)
0-24 hours	58	2.4	65	75
24-72 hours	22	1.8	45	60
72+ hours	20	0.8	35	40

Reinforcement Learning Approaches

Contextual Bandits

Contextual bandit algorithms have emerged as a powerful initial approach to reinforcement learning in digital marketing, demonstrating significant improvements in campaign performance metrics. Implementation data shows that these systems can effectively process and utilize up to 20 different contextual features, including user behavior patterns, demographic information, and historical interaction data. Recent studies indicate that contextual bandit implementations have achieved improvements of 12-15% in click-through rates (CTR) compared to traditional rule-based systems, while maintaining an exploration rate of approximately 10% to balance new opportunity discovery with exploitation of known successful strategies [5].

The adaptability of contextual bandits in handling dynamic marketing environments has proven particularly valuable in campaign optimization. Research shows that these systems can effectively manage marketing budgets ranging from $5,000 to $50,000 per month, with the ability to automatically adjust bidding strategies based on real-time performance metrics. The implementation of contextual bandits has demonstrated a 20% improvement in customer engagement rates and a 15% reduction in customer acquisition costs across various digital marketing channels [5].

Deep Q-Learning Applications

Deep Q-learning has revolutionized real-time bidding strategies in sponsored search advertising, with remarkable improvements in both efficiency and effectiveness. Empirical studies demonstrate that DQN-based systems can process auction data containing over 1 billion impressions and 8.2 million clicks, with the ability to handle bid requests at a rate of 10,000 queries per second. These implementations have shown consistent improvements in key performance metrics, achieving a 20.3% increase in auction wins while maintaining cost-per-click (CPC) targets and improving return on investment (ROI) by up to 17.8% [6].
The sophisticated architecture of DQN systems enables complex pattern recognition in user behavior and market dynamics. Research indicates that these systems can effectively manage daily budgets exceeding $100,000 while maintaining stable performance across different market conditions. Implementation data shows that DQN-based approaches have achieved a 15.4% reduction in average CPC while increasing click-through rates by 18.7% compared to traditional bidding strategies. The systems have demonstrated particular strength in handling seasonal variations, with performance improvements of up to 23.5% during high-traffic periods [6].

Actor-Critic Methods

Actor-critic architectures have demonstrated exceptional capability in handling the continuous action space inherent in digital marketing optimization. These systems have shown the ability to process and analyze user interaction data across multiple touchpoints, with studies indicating successful handling of up to 100,000 daily user interactions. The dual-network approach has enabled more precise budget allocation, with documented improvements of 18-22% in campaign ROI compared to single-network implementations [5].

The implementation of actor-critic methods in sponsored search bidding has yielded particularly impressive results in long-term value optimization. Research shows that these systems can effectively predict and optimize for lifetime value metrics, with improvements of up to 26.7% in long-term ROI compared to traditional methods. The critical network's evaluation capabilities have enabled more efficient learning, with studies showing a 30% reduction in the time required to achieve optimal bidding strategies. Performance data indicates that actor-critic systems maintain consistent effectiveness even when handling complex campaigns with daily impression volumes exceeding 10 million, while achieving a 21.4% improvement in conversion rates [6].

Table 2. Reinforcement Learning Approach Effectiveness [5, 6].

Learning Method	CTR Improvement (%)	Conversion Rate (%)	Cost Reduction (%)	ROI Improvement (%)
Contextual Bandits	41	18	27	34
Deep Q-Learning	56	31.2	23	47
Actor-Critic	38	25	31	26.7

Real-World Implementation Challenges

Reward Shaping in Digital Advertising

The implementation of reward shaping mechanisms in digital advertising presents complex challenges, particularly in handling delayed feedback signals. Research conducted across major advertising platforms shows that traditional immediate-reward systems capture only 65% of the true engagement value, with delayed interactions accounting for up to 35% of total campaign performance metrics. Studies indicate that approximately 40% of all meaningful user engagements occur between 1 and 24 hours after the initial ad impression, with an additional 25% occurring within a 7-day window. Implementation of sophisticated reward shaping mechanisms has demonstrated improvements in engagement rates ranging from 20% to 35%, with particularly strong performance in mobile advertising channels where delayed engagement rates are highest [7].

Advanced reward architectures incorporating time-decay models have shown significant promise in capturing long-term value. Platforms implementing these systems have reported improvements in prediction accuracy by up to 42% compared to traditional immediate-reward models. The integration of hierarchical reward structures has enabled more efficient learning, with some implementations reducing the training time required for optimal performance by up to 28% while maintaining robust performance across diverse campaign objectives [7].

Delayed Feedback Management

Recent studies in recommendation systems have revealed that effectively managing delayed feedback can lead to substantial performance improvements. Implementation data shows that sophisticated bidding systems can handle up to 125,000 queries per second while processing delayed feedback signals from multiple time windows. These systems have demonstrated the ability to improve click-through rates by 27.3% and conversion rates by 31.2% compared to traditional approaches that don't account for delayed feedback patterns. The analysis of user behavior patterns indicates that approximately 58% of conversions occur within the first 24 hours, while 22% occur between 24-72 hours, and the remaining 20% span up to 15 days post-impression [8].

The integration of advanced feedback handling mechanisms has shown particular promise in addressing the temporal aspects of user engagement. Systems implementing sophisticated delay-aware algorithms have achieved improvements of up to 24.8% in ranking accuracy while simultaneously reducing bid price volatility by 18.3%. Research indicates that these improvements are most pronounced in campaigns with average conversion delays exceeding 48 hours, where traditional systems typically struggle to maintain consistent performance [8].

Counterfactual Learning and Estimation

Counterfactual learning approaches have demonstrated significant potential in optimizing both bidding and ranking strategies. Studies show that integrated systems can evaluate up to 1,000 potential bidding strategies per second while maintaining estimation accuracy within 5% of actual performance metrics. Implementation data reveals that these systems can reduce exploration costs by up to 45% compared to traditional A/B testing approaches, while achieving a 31.5% improvement in overall advertising efficiency [7].

The joint optimization of bidding and ranking has emerged as a particularly effective approach in modern advertising systems. Research indicates that unified systems can process over 2.5 million impressions per hour while maintaining real-time performance requirements. These implementations have shown improvements in conversion rates of up to 35.2% while reducing average cost-per-acquisition by 22.7%. The integration of sophisticated estimation techniques has enabled more efficient resource allocation, with studies showing that advanced systems can achieve optimal performance using only 40% of the data required by traditional approaches [8].

Table 3. Real-Time Bidding System Performance Optimization Metrics [7, 8].

Optimization Metric	Before Optimization	After Optimization	Improvement (%)	Business Impact
Response Time (ms)	100	25	75	Enables participation in more auctions, increasing win rates
Queries/Second (K)	250	350	40	Higher throughput capacity for peak traffic periods
Error Rate (%)	5.3	0.3	94	Fewer failed bids, resulting in more consistent campaign delivery
Resource Usage (%)	100	59	41	Reduced infrastructure costs and improved system scalability
Average Latency (ms)	85	22	74	More competitive in time-sensitive auction environments
Processing Efficiency (ops/sec/core)	15K	38K	153	Higher operational efficiency with existing hardware

Practical Considerations

Real-Time Pacing

The implementation of optimal real-time bidding strategies requires sophisticated pacing mechanisms that can adapt to dynamic market conditions. Research shows that optimal bidding strategies using stochastic control approaches can improve campaign performance by up to 30% compared to static bidding methods. Mathematical modeling of bidding behaviors indicates that dynamic pacing algorithms can maintain spending efficiency within a 5% variance while achieving up to 25% higher engagement rates compared to fixed-price strategies. Studies demonstrate that advanced pacing systems can effectively manage daily budget utilization patterns, maintaining spending rates within optimal ranges 92% of the time while adapting to market price fluctuations that can vary by up to 40% during peak periods [9].

Advanced pacing mechanisms incorporating fluid limit approximations have shown particular promise in large-scale implementations. These systems demonstrate the ability to process and optimize campaigns with daily impression volumes exceeding 1 million while maintaining stable performance metrics. Research indicates that sophisticated pacing algorithms can reduce cost per acquisition by up to 18% through intelligent budget allocation, while ensuring delivery goals are met across different time zones and market conditions. Implementation data shows that these systems can effectively manage price dynamics across multiple auction types, with some platforms reporting performance improvements of up to 22% in terms of overall campaign efficiency [9].

QPS Limitations

The management of query-per-second constraints in real-time bidding systems presents significant technical challenges, particularly in fraud detection and prevention. Recent studies indicate that modern RTB platforms must process and validate up to 400,000 bid requests per second while maintaining response times under 100 milliseconds. Implementation data shows that advanced fraud detection systems can identify and filter out suspicious traffic patterns with 99.7% accuracy while adding only 15-20 milliseconds to overall processing time. These systems have demonstrated the capability to reduce fraudulent impressions by up to 85% while maintaining normal bidding operations [10].

Research in real-time fraud detection mechanisms has revealed that efficient processing architectures can significantly improve system performance. Studies show that optimized validation systems can reduce average response times from 95ms to 25ms while maintaining detection accuracy above 98%. Implementation of sophisticated caching mechanisms has demonstrated the ability to handle sustained loads of up to 350,000 QPS with 99.95% availability, while reducing computational resource requirements by up to 40%. Advanced traffic pattern analysis systems have shown the capability to identify and block suspicious activities within 50 milliseconds of detection, preventing potential fraud losses estimated at 15-20% of advertising budgets [10].

Safe Policy Rollout

The deployment of new bidding strategies requires careful consideration of safety mechanisms and fraud prevention measures. Recent research indicates that progressive rollout strategies incorporating real-time fraud detection can reduce negative campaign impact by up to 75% compared to traditional deployment approaches. Implementation data shows that carefully managed traffic validation, starting with 10% of inventory and gradually increasing based on performance and safety metrics, can achieve optimal results while maintaining fraud rates below 0.5% of total impressions [9].

Advanced fraud prevention systems have demonstrated remarkable effectiveness in ensuring safe policy rollout, with studies showing that automated monitoring systems can detect and prevent up to 94% of potentially fraudulent bidding patterns within milliseconds. These systems typically maintain bid price validation within three standard deviations of historical averages while allowing for natural market adaptation. Research indicates that implementing comprehensive monitoring systems with automated fallback mechanisms can reduce exposure to fraudulent activities by up to 88%, with typical response times under 45 milliseconds for critical anomaly detection [10].

Performance Optimization

Feature Engineering

Feature engineering in real-time bidding systems represents a critical process of transforming raw data into meaningful predictive signals. Research shows that effective feature selection and transformation can improve model performance by up to 35% compared to basic implementations. The process typically involves analyzing historical data spanning 90-day windows, with studies indicating that properly engineered temporal features can capture seasonal patterns with 87% accuracy. Implementation data reveals that successful feature engineering approaches can reduce dimensionality by up to 60% while maintaining 95% of the original information content, significantly improving model efficiency and training speed [11]. Advanced feature engineering techniques incorporating domain knowledge have shown particular promise in RTB systems. Studies demonstrate that engineered features combining user behavioral patterns with contextual information can improve prediction accuracy by up to 28%. The implementation of sophisticated feature selection methods has enabled the identification of key predictive variables from among thousands of potential features, with some platforms reporting that as few as 50 well-engineered features can capture up to 85% of the variance in bidding outcomes. Research indicates that proper feature scaling and normalization techniques can reduce model training time by up to 45% while improving convergence stability by 30% [11].

Model Architecture Optimization

The optimization of bidding strategies through reinforcement learning and machine learning techniques has demonstrated significant improvements in campaign performance. Implementation studies show that optimized model architectures can achieve up to 42% improvement in return on advertising spend (ROAS) compared to traditional approaches. Research indicates that carefully designed neural network architectures with 4-6 hidden layers and appropriate batch normalization can reduce inference time by up to 55% while maintaining prediction accuracy above 91%. The integration of advanced optimization techniques has enabled processing of up to 150,000 bid requests per second with average response times under 20 milliseconds [12].

Real-world deployments of optimized bidding systems have shown remarkable efficiency gains through strategic architecture design. Studies demonstrate that implementing efficient parameter sharing techniques can reduce model size by up to 65% while maintaining performance within 96% of non-compressed models. Advanced research in RTB optimization reveals that hybrid approaches combining reinforcement learning with traditional machine learning methods can improve campaign performance by up to 38% while reducing computational overhead by 25%. Implementation data shows that optimized systems can achieve up to 99.95% uptime while processing daily volumes exceeding 10 million bid requests [12].

Performance Impact Analysis

The synergistic effect of comprehensive feature engineering and model optimization has demonstrated significant improvements in real-world advertising scenarios. Analysis shows that properly engineered features combined with optimized model architectures can reduce cost per acquisition (CPA) by up to 33% while improving conversion rates by 28%. Implementation studies reveal that systems incorporating both optimized features and efficient architectures can maintain stable performance even during high-traffic periods, with some platforms reporting sustained throughput improvements of up to 45% during peak loads [11].

Advanced optimization techniques have shown particular effectiveness in managing complex bidding scenarios. Research indicates that properly optimized systems can achieve click-through rate (CTR) improvements of up to 47% while maintaining bid prices within optimal ranges. The implementation of sophisticated monitoring and adaptation mechanisms has demonstrated the ability to automatically adjust bidding strategies based on real-time market conditions, with some systems showing the capability to reduce average bid prices by 25% while maintaining or improving campaign performance metrics. Studies show that comprehensive optimization approaches can improve overall campaign ROI by up to 52% compared to baseline implementations [12].

Table 4. Performance Optimization Results [11, 12]

Optimization Area	Performance Gain (%)	Resource Reduction (%)	Time Saved (%)	Accuracy (%)
Feature Engineering	35	60	45	87
Model Architecture	42	65	55	91
Combined Approach	52	75	65	94

Future Directions

Advanced Real-Time Bidding Systems

The evolution of real-time bidding systems represents a significant leap forward in advertising technology, with modern platforms processing over 10 billion daily ad impressions. Implementation studies show that advanced RTB systems can reduce average cost-per-click (CPC) by up to 40% while improving targeting accuracy by 35%. These systems demonstrate remarkable efficiency in audience targeting, with some platforms reporting improvements of up to 45% in audience matching accuracy compared to traditional programmatic advertising methods. The integration of sophisticated bidding algorithms has enabled advertisers to achieve up to 60% higher ROI through more precise bid optimization and audience targeting capabilities [13].

Artificial Intelligence and Machine Learning Integration

The integration of AI and machine learning technologies in advertising platforms has revolutionized campaign optimization capabilities. Research indicates that AI-powered systems can process and analyze user behavior patterns across multiple channels simultaneously, leading to improvements of up to 32% in conversion rates. Implementation data shows that these advanced systems can reduce campaign optimization time by up to 75% while improving audience targeting precision by 28%. The adoption of machine learning algorithms has enabled more sophisticated audience segmentation, with some platforms reporting up to 40% improvement in targeting accuracy for specific demographic groups [14].

Programmatic Direct and Private Marketplaces

The advancement of programmatic direct advertising and private marketplaces has introduced new opportunities for campaign optimization. Studies show that these specialized platforms can achieve premium inventory access while maintaining cost efficiency, with some implementations reporting up to 25% lower costs compared to traditional direct buying methods. The integration of automated negotiation systems has demonstrated the ability to reduce transaction times by up to 85% while improving inventory quality scores by 30%. Research indicates that programmatic direct implementations can achieve up to 50% higher viewability rates compared to open marketplace transactions [13].

Cross-Channel Integration and Attribution

The development of sophisticated cross-channel attribution models has enabled more effective campaign optimization across multiple platforms. Implementation data shows that advanced attribution systems can track and analyze user interactions across up to 12 different channels simultaneously, leading to improvements of up to 38% in conversion attribution accuracy. These systems have demonstrated the ability to reduce attribution errors by up to 45% while providing more granular insights into customer journey patterns. Research indicates that integrated attribution models can improve overall campaign efficiency by up to 33% through better budget allocation across channels [14].

Privacy-Focused Technologies

The evolution of privacy-preserving advertising technologies represents a critical development in the industry. Studies show that advanced privacy-compliant systems can maintain targeting effectiveness while reducing reliance on third-party cookies, with some implementations reporting only a 5-10% decrease in performance metrics during the transition. Implementation data demonstrates that privacy-focused targeting solutions can achieve up to 85% of traditional targeting capabilities while fully complying with modern data protection regulations. The integration of first-party data strategies has shown particular promise, with some platforms reporting improvements of up to 25% in engagement rates compared to third-party data approaches [13].

Emerging Technologies and Future Applications

The integration of emerging technologies such as augmented reality (AR) and virtual reality (VR) in advertising platforms presents exciting opportunities for engagement improvement. Research shows that AR-enhanced advertising campaigns can achieve up to 70% higher engagement rates compared to traditional display advertising. Implementation studies demonstrate that interactive advertising formats can improve brand recall by up to 45% while increasing purchase intent by 32%. The development of immersive advertising experiences has shown particular promise, with some platforms reporting improvements of up to 55% in user engagement metrics for AR/VR-enhanced campaigns [14].

CONCLUSION

The implementation of reinforcement learning in online advertising auctions represents a significant advancement in automated decision-making capabilities. The successful integration of sophisticated bidding strategies has demonstrated substantial improvements in campaign performance across multiple metrics, including conversion rates, return on investment, and budget utilization efficiency. The adoption of advanced techniques such as contextual bandits and Deep Q-learning has enabled more precise targeting and dynamic adaptation to market conditions, while actor-critic architectures have proven particularly effective in handling continuous action spaces. The development of robust solutions for handling delayed feedback and implementing safe policy rollout mechanisms has enhanced the practical applicability of these systems in production environments. Feature engineering and model architecture optimization have emerged as critical factors in achieving optimal performance while maintaining computational efficiency. Looking forward, the continued evolution of multi-agent systems, transfer learning, and meta-learning techniques promises to further enhance the capabilities of automated advertising systems. The integration of privacy-preserving technologies and emerging platforms suggests a future where advertising optimization becomes increasingly sophisticated while remaining respectful of user privacy and regulatory requirements. These advancements point toward a future where advertising platforms can deliver increasingly personalized and effective campaigns while maintaining operational efficiency and user trust.

REFERENCES

[1] Xiangyu Zhao et al., "Deep Reinforcement Learning for Online Advertising in Recommender Systems," ResearchGate, 2019. [Online]. Available: https://www.researchgate.net/publication/335713358_Deep_Reinforcement_Learning_for_Online_Advertising_in_Recommender_Systems
[2] Han Cai et al., "Real-Time Bidding by Reinforcement Learning in Display Advertising," ACM Digital Library, 2017. [Online]. Available: https://dl.acm.org/doi/10.1145/3018661.3018702#:~:text=Thus%2C%20it%20is%20of%20great,of%2Dthe%2Dart%20methods.
[3] Sofía Pérez, "Real-Time Bidding: A machine learning approach to finding optimal bidding prices," Marvik, 2022. [Online]. Available: https://blog.marvik.ai/2022/08/30/real-time-bidding-a-machine-learning-approach-to-finding-optimal-bidding-prices/
[4] IAS, "Real-time viewability insights for real-time optimization," 2019. [Online]. Available: https://integralads.com/insider/real-time-viewability-insights-for-real-time-optimization/
[5] Mrinal Singh, "All You Need To Know About Reinforcement Learning For Digital Marketing," Analytics Vidhya, 2023. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/06/all-you-need-to-know-about-reinforcement-learning-for-digital-marketing/
[6] Jun Zhao et al., "Deep Reinforcement Learning for Sponsored Search Real-time Bidding," arxiV, 2018. [Online]. Available: https://arxiv.org/abs/1803.00259
[7] Justina Chen and Ira Ktena, "Improving engagement on digital ads with delayed feedback," X Engineering, 2019. [Online]. Available: https://blog.x.com/engineering/en_us/topics/insights/2019/improving-engagement-on-digital-ads-with-delayed-feedback

[8] Geng Ji et al., "Learning to bid and rank together in recommendation systems," Springer Nature Link, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s10994-023-06444-4
[9] Joaquin Fernandez-Tapia, Tapia Olivier Guéant and Jean-Michel Lasry "Optimal Real-Time Bidding Strategies," ResearchGate, 2015. [Online]. Available: https://www.researchgate.net/publication/284721659_Optimal_Real-Time_Bidding_Strategies
[10] Everton Gomede, "Guarding the Gates: Detecting Fraudulent Activity in Real-Time Bidding Systems," Medium, 2025. [Online]. Available: https://medium.com/@evertongomede/guarding-the-gates-detecting-fraudulent-activity-in-real-time-bidding-systems-10248fccf52b
[11] Harshil Patel, "Feature Engineering Explained," Buitl-in, 2024. [Online]. Available: https://builtin.com/articles/feature-engineering#:~:text=Feature%20Engineering%20Definition-,Feature%20engineering%20is%20the%20process%20of%20selecting%2C%20manipulating%20and%20transforming,exploratory%20data%20analysis%20and%20benchmarking.
[12] Arti Jha et al., "Optimizing Real-Time Bidding Strategies: An Experimental Analysis of Reinforcement Learning and Machine Learning Techniques," ScienceDirect, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050924008676
[13] Camphouse, "Real-Time Bidding Advertising: How RTB Works & Benefits," 2025. [Online]. Available: https://camphouse.io/blog/real-time-bidding
[14] Quynh Pham, "New Advertising Technology: 5 Emerging Advertising Technology Trends," Orient, 2024. [Online]. Available: https://www.orientsoftware.com/blog/new-advertising-technology/