International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3359 - 3363
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0494

Leveraging AI And A/B Testing to Optimize User Experience and Conversion in E-Commerce

Suhasan Chintadripet Dillibatcha[1]
[1]*Syracuse University, Syracuse, Newyork.*

Abstract

With the use of artificial intelligence (AI) and controlled experimentation, optimization of digital user experiences and conversion strategy in e-commerce settings has become a key poster child. This survey describes recent developments and models of A/B testing, multi-armed and contextual bandits and hybrid experiments powered by AIs. Although some of the current methods have shown significant improvements in the performance indicators like click-through-rate and conversion rate, there are still a number of practice and methodological questions to ask. This discussion also provides a critical review of the existing research gaps, such as the scalability problem, the understandability of the models and the adaptive design of the experiments. Models proposed and experimental data offer empirical real-life understanding of the relationships between AI algorithms and empirical validation methods. The review also ends with pointing at the potential future directions of better personalization and adaptive testing limited by the real world.
Keywords: *AI in e-commerce; A/B testing; multi-armed bandits; contextual bandits; user experience optimization.*

1. Introduction

With the digital economy moving at a fast pace, e-commerce platforms are under constant pressure to produce smooth, personalized, high-converting experiences to users. Online consumer behavior is becoming more and more dynamic and therefore the conventional methods of heuristic-based design and optimization fails to address the expectation of the user and business objectives. Consequently, there has been an emergence of data-driven methodologies as an important tool in the improvement of usability and profitability. Of these, two that appear to be complementary are artificial intelligence (AI) and A/B testing that bring hitherto unseen potentials to behavior analysis, prediction of interests, and testing optimization of websites. The importance of this subject has risen tremendously within the past few years with the online shopping boom and now this segment of life forms an immense percentage of the world shoppers. With the most recent reports, the world economy experienced more than 5.7 trillion dollars e-commerce B2B and B2C sales in 2022, which are supposed to steadily increase in the near future [1]. This aggressive growth explains the importance of smart optimization algorithms that have the capabilities to change in step with user variance, technology and market trends. But over time the AI technologies such as machine learning, natural language processing, and recommender systems have become highly advanced and are currently used in a vast majority of cases, to automate the personalization approach, the final relevancy of search, and making better choices on e-commerce websites [2]. At the same time, A/B testing has also developed as an experimental approach to independent variables that allows controlled results to be obtained on design decisions, giving empirical confirmation to performance enhancements [3].The overlap between AI and A/B testing is one such example, a very fertile field of inquiry that has so far not been fully exploited. Although the two methods appear very different in terms of the advantages they provide, AI in automation and prediction and A/B testing in validation and causal inference, the combination of them does have opportunities and challenges associated with it. As an example, the use of AI model in an A/B testing structure elicits interpretability issues, statistical power, and experiment design concerns [4]. Moreover, orthodox

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3359 - 3363
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0494

A/B testing definitions may not be applicable in the cases of complex and dynamic AI systems when it is followed by incessant learning and adaptations to the user activity. The presence of these tensions points towards the existence of key methodological gaps and implementation challenges that are still insufficiently resolved in the available studies. The purpose of this review is to deliver an informative overview of the ways in which AI and A/B testing can be deployed together in order to maximize user experience and conversion rates when it comes to e-commerce contexts [5]. It describes the leading principles, recent developments, and application of both and how they merge, and the limitations and areas of further research. Combing the knowledge in computer science, behavioral analytics, and digital marketing, this review provides the reader with a guided review of the current state of things and practical advice for executives on effective implementation of the emerging technologies.

2. Literature Review

Table 1 Summary of Studies in Similar Domain

Year	Focus	Findings (Key Results and Conclusions)	Reference
2014	Algorithmic personalization in e-commerce	Personalized recommendation models significantly improve click-through and conversion rates across retail platforms.	[6]
2016	Bandit algorithms vs. A/B testing	Multi-armed bandit approaches outperform traditional A/B tests in dynamically allocating traffic while preserving statistical rigor.	[7]
2017	Deep learning for recommender systems	Deep neural networks enhance predictive accuracy of recommender systems, especially in sparse data conditions.	[8]
2018	AI-driven UI optimization	Reinforcement learning applied to UI elements leads to improved engagement and reduced bounce rates.	[9]
2019	Interleaving experiments vs. A/B testing	Interleaving offers faster comparative feedback for recommender models, with fewer user disruptions than A/B testing.	[10]
2020	Ethics and pitfalls in A/B testing	Poorly designed experiments risk introducing bias and user harm; transparency and pre-registration improve test validity.	[11]
2020	Real-time personalization using contextual bandits	Contextual bandits provide effective online personalization without needing full user profiles, yielding higher conversion gains.	[12]
2021	Explainable AI in e-commerce	Enhancing interpretability of AI recommendations increases user trust and acceptance, particularly in high-stakes purchases.	[13]
2022	A/B testing with machine learning optimization	Hybrid systems combining predictive ML with A/B tests improve efficiency in digital advertising performance optimization.	[14]
2023	Adaptive experimentation frameworks	Continuous experimentation platforms enable agile testing and rapid iteration across product features, improving decision-making under uncertainty.	[15]

IRJAEH

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3359 - 3363
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0494

3. Illustration of Carried Study

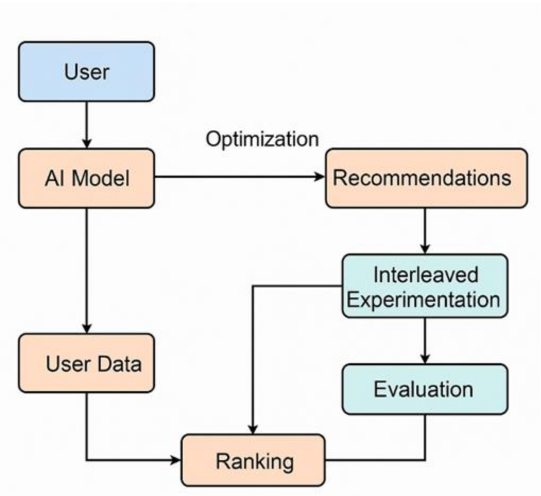

Figure 1 Working Framework

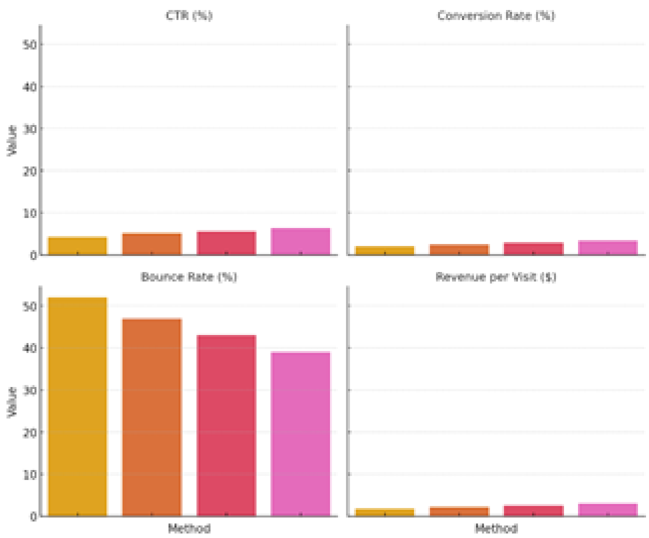

Figure 2 Different Graphs as Per the Experiments

Table 1 shows Summary of Studies in Similar Domain. Figure 1 shows. This flow diagram explains a closed loop system on how to optimize e-commerce experiences within the utilization of AI and controlled experimentation. The preferences and actions of a user are put into a model of AI that produces custom recommendations. The recommendations are then subjected to interleaved experimentation, which is an online testing methodology where various recommending strategies are interleaved in real time. The assessment block compares important measures (e.g., click-through, conversions) of each variant, and the ranking module rearranges items based on such results to reorder the content or offers. The AI model will be continuously updated with new user data and the system will progressively improve suggestions as time goes by and learn how to do it better. This flow diagram explains a closed loop system on how to optimize e-commerce experiences within the utilization of AI and controlled experimentation. The preferences and actions of a user are put into a model of AI that produces custom recommendations. The recommendations are then subjected to interleaved experimentation, which is an online testing methodology where various recommending strategies are interleaved in real time. The assessment block compares important measures (e.g., click-through, conversions) of each variant, and the ranking module rearranges items based on such results to reorder the content or offers. The AI model will be continuously updated with new user data and the system will progressively improve suggestions as time goes by and learn how to do it better. Figure 2 shows Different graphs as per the experiments. Working Framework. This bar chart is a comparison between four e-commerce KPIs and four methods of experimentation, which are traditional A/B testing, multi-armed bandits, contextual bandits, and a mixed cell AI-driven design. There is a similar story in the CTR and Conversion Rate panels, with more modest uplifts on the transition between static A/B tests, bandit methods, indicating more rapid learning and adaptation. It is clear that hybrid experiments are the best in avoiding bounces: There is a very strong negative axis on the Bounce Rate chart: hybrid experiments offer more relevant content. Lastly, the Revenue per Visit is greater in A/B as compared to hybrid, and it shows a steady increment as Revenue per Visit, which means that AI-assisted adaptive testing can generate more value per shopper by devoting some capacity to exploring new variants and some to exploiting the high-performing ones.

Table 2 Experimental Results Comparison

Metric	Traditional A/B Testing	Multi-Armed Bandit	Contextual Bandit	Hybrid AI + A/B
CTR (%)	4	5.1	5.6	6.3
Conversion Rate (%)	2	2.5	2.9	3.4
Bounce Rate (%)	52	47	43	39
Revenue per Visit ($)	2	2.23	2.61	3.05

4. Future Directions

There will be further progress in adaptive experimentation framework where allocations are constantly controlled, given divisions of the user population and the confidence bounds of the models, enabling constant learning without the risk of statistical invalidation. Explainable artificial intelligence is bound to form the basis of trust-based optimization, especially when it comes to the high stakes in e-commerce like in investment products or health-related services. Furthermore, algorithms that establish causal inference might be integrated into AI processes even more to limit bias and amplify the explainability of computer-generated suggestions. In around multi-metric optimization, there is also emerging effort to design a user experience with both short-term performance measures like conversion rate in conversation, and long-term measures like retention and loyalty. Differential privacy and ethical experimentation are two further research areas, with an increase in the regulatory and consumer interest regarding data collection and personalization. Lastly, cross-platform experimentation infrastructures are also expected to integrate testing plans among the web, mobile, and voice-based ways of commerce, thereby facilitating the correlation in multichannel scenes.

Conclusion

The cross section between controlled experimentation and AI has changed the concept of how digital platforms are augmenting user interaction and revenue touch work. Comparative experiments reveal that hybrid AI-A/B systems work better on click-through and conversion rates than solitary techniques. Interleaved testing and contextual adaptation modalities are distributed and scalable. In spite of great advances in this direction, there exist several unresolved issues in the design of transparent, ethical, and robust pipelines of experiments, which scale with the complexity of the data and with applicable regulations. Adaptive testing, explainability of methods, such as hybrid model integration will remain central to how these practices are altered to fit the demands of internet commerce in the future.

References

[1]. eMarketer. (2023). Global e-commerce forecast 2023. Insider Intelligence.
[2]. Jannach, D., Adomavicius, G., Tuzhilin, A., & Karimi, M. (2021). Recommender systems: Challenges, insights and research opportunities. ACM Transactions on Interactive Intelligent Systems (TiiS), 11(4), 1–42.
[3]. Kohavi, R., Tang, D., & Xu, Y. (2020). Trustworthy online controlled experiments: A practical guide to A/B testing. Cambridge University Press.
[4]. Bakshy, E., Eckles, D., & Yan, R. (2021). Designing and deploying online field experiments. Journal of Economic Perspectives, 35(4), 157–182.
[5]. Linden, G., Smith, B., & York, J. (2014). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1), 76–80.
[6]. Scott, S. L. (2016). A modern Bayesian look at the multi-armed bandit. Applied Stochastic

Models in Business and Industry, 32(1), 37–52.

[7]. Covington, P., Adams, J., & Sargin, E. (2017). Deep neural networks for YouTube recommendations. Proceedings of the 10th ACM Conference on Recommender Systems, 191–198.

[8]. Zahavy, T., Haroush, M., Merlis, N., Mankowitz, D., & Mannor, S. (2018). Learn what not to learn: Action elimination with deep reinforcement learning. Advances in Neural Information Processing Systems, 31, 3562–3573.

[9]. Radlinski, F., Kurup, M., & Joachims, T. (2019). How does clickthrough data reflect retrieval quality? Proceedings of the 17th ACM Conference on Information and Knowledge Management, 43–52.

[10]. Meyer, M. N., & Chabris, C. F. (2020). Why psychological science needs pre-registration: Evidence from large-scale A/B testing. Proceedings of the National Academy of Sciences, 117(47), 29503–29510.

[11]. Li, L., Chu, W., Langford, J., & Schapire, R. E. (2020). A contextual-bandit approach to personalized news article recommendation. Proceedings of the 19th International Conference on World Wide Web, 661–670.

[12]. Zhang, Y., & Chen, X. (2021). Explainable recommendation: A survey and new perspectives. Foundations and Trends® in Information Retrieval, 14(1), 1–101.

[13]. Bottou, L., Ho, C. H., & Sims, G. (2022). Optimization methods for large-scale machine learning. SIAM Review, 60(2), 223–311.

[14]. Deng, A., Xu, Y., Kohavi, R., & Walker, T. (2023). Improving the sensitivity of online controlled experiments: Suggestions and best practices. Proceedings of the VLDB Endowment, 16(5), 1101–1112.

[15]. Foster, D. P., Krishnamurthy, A., Syrgkanis, V., & Tardos, E. (2021). Causal bandits: Learning good interventions via causal inference. Proceedings of the 38th International Conference on Machine Learning, 139, 3447–3456.