iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IX Month of publication: September 2025

DOI: https://doi.org/10.22214/ijraset.2025.74411

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

Comparative Study of Traditional vs. AI-Enhanced A/B Testing for UX Optimization

Dr. S. Dhanalakshmi[1], Selvavignesh M[2], Syedfarith C[3]
[1]Assosiate Professor, Department of Software Systems, Sri Krishna Arts and Science College Coimbatore, India
[2], [3] Students, Department of Software Systems, Sri Krishna Arts and Science College Coimbatore, India

Abstract: In modern digital platforms, optimizing user experience (UX) is crucial for user engagement and business success. Traditional A/B testing methods are widely used but can be time-consuming, require a lot of traffic, and struggle to adjust dynamically. To tackle these issues, we propose an AI-enhanced A/B testing framework that combines machine learning models and adaptive decision-making algorithms to optimize UX more efficiently. Our approach uses predictive modeling to estimate design performance with smaller datasets, which shortens the duration of experiments. We also include a multi-armed bandit strategy that reallocates user traffic to better-performing design variants in real time, reducing the costs of poor- performing options. The system incorporates detailed behavioral analytics, like cursor movements, scroll depth, hesitation patterns, and engagement metrics. This provides deeper insights into user interactions beyond standard conversion rates. This AI-driven approach speeds up decision-making and lowers experimental overhead, ensuring ongoing adjustment to changing user behavior. By connecting UX research with AI-driven analytics, our framework gives organizations a smart, scalable way to improve UX iteratively.

Keywords: A/B Testing, User Experience (UX), Artificial Intelligence (AI), Machine Learning, Multi- Armed Bandit, Behavioral Analytics, Human- Centered Design.

I. INTRODUCTION

User experience (UX) is now a key factor in the success of digital products like web apps, e-commerce sites, and mobile apps. Customer satisfaction, engagement, and business results can all be improved with a well-designed user interface. Traditionally, evaluating and improving UX has relied on A/B testing. This method shows users two or more design options and uses statistical methods to find the best one. However, conventional A/B testing has several limitations: 1. Time-consuming: A large number of users must take part before reaching statistical significance. 2. Traffic inefficiency: Evenly sharing traffic can waste resources on poor-performing options. 3. Limited adaptability: Static testing methods do not change with user behavior. To address these issues, researchers a nd professionals are looking into adding artificial intelligence (AI) to A/B testing processes. By using machine learning models, multi- armed bandit algorithms, and behavioral analytics, UX testing can become faster, more adaptable, and more insightful. This paper suggests an AI-enhanced A/B testing framework that speeds up UX optimization while lowering experimental costs.

II. WORK

Several studies have looked into how AI can improve web optimization and adaptive interfaces. Multi- armed bandit algorithms are used in recommendation systems. They show the benefits of adapting allocation strategies. Predictive modeling is also used in marketing to predict conversion rates. However, not many studies have applied these methods directly to UX-focused A/B testing. Most current systems focus only on conversion metrics like click-through rates. They overlook behavioral signals such as scroll depth, hesitation, and cursor movement, which could indicate user frustration or satisfaction. Our proposed framework fills this gap by including behavioral analytics in the testing process, making UX evaluation more complete.

III. METHODOLOGY

A. *Framework Overview*

The proposed system has four main components:

1) Design Variant Input: Two or more UX designs, such as landing pages or navigation flows.
2) Predictive Modeling Layer: Early performance estimation using small user samples.
3) Multi-Armed Bandit Engine: Dynamic reallocation of traffic to better variants.
4) Behavioral Analytics Module: Collection of detailed user interaction data.

B. *Predictive Modeling*
1) A supervised machine learning model, like logistic regression or random forest, predicts the chance of conversion after initial exposure to a small sample size.
2) Session length, scroll depth, bounce rate, and hesitation analytics are among the features.
3) Output: Early confidence score of variant performance.

C. *Multi-Armed Bandit Algorithm*
1) Uses Thompson Sampling or Upper Confidence Bound algorithms.
2) Adjusts traffic allocation in real time; less traffic goes to poor variants while promising ones get more exposure.
3) Ensures quicker progress to the best design.

D. *Behavioral Analytics Integration*
Collects detailed signals beyond conversions:
1) Cursor heatmaps
2) Repeated clicks on dormant elements are known as "rage clicks."
3) Scroll completion percentage
4) Navigation hesitation time

E. *Workflow Representation*
The overall workflow of the proposed AI-enhanced A/B testing framework is illustrated in **Fig. 1**. The process begins with design variants and user traffic as input. Then, predictive modeling estimates early performance. Traffic is then sent to variants with higher performance via the multi-armed bandit engine. The behavioral analytics module captures detailed user interaction data to support traditional conversion metrics. Finally, the framework checks for significant improvements. If there are none, the system retrains models and reallocates traffic. If there are improvements, the winning UX design is deployed, and reports are generated for stakeholders.

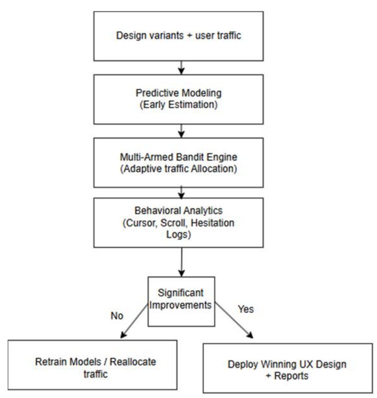


Fig. 1

IV. EXPERIMENTAL SETUP

A. *Dataset and Environment*
Experiments were conducted on a simulated e-commerce platform with two checkout page variants.
1) Variant A: Traditional multi-step checkout.
2) Variant B: Simplified one-page checkout.
3) Dataset: 10,000 simulated user sessions.

B. *Baseline (Traditional A/B Testing)*
1) Equal traffic split (50/50).
2) Conversion rates measured after 10,000 users.
3) Time to reach statistical significance: 3 weeks.

C. Proposed Method (AI-Enhanced)
1) Initial 50/50 traffic allocation.
2) Predictive model evaluated after 1,000 users.
3) Multi-armed bandit algorithm reallocated 70% traffic to Variant B after early results.
4) Convergence achieved after 3,000 users (about 1 week).

D. Evaluation Metrics
To measure the effectiveness of the proposed AI-enhanced A/B testing framework, multiple quantitative and qualitative metrics were used:
1) Time to Decision: The duration needed to identify a statistically reliable winning variant.
2) Traffic Efficiency: The proportion of users exposed to underperforming variants, showing resource use.
3) Conversion Uplift: The percentage improvement in conversion rate achieved by the winning design compared to the baseline.

V. RESULTS AND DISCUSSION

Metric	Traditiona l A/B	AI-Enhanced A/B	Improvement
Time to decision	3 weeks	1 week	66% faster
Users exposed to loser	5000	1200	76% lower
Conversion Rate (Winner)	23% (B)	24% (B)	+1% uplift

Analysis
1) The AI-powered approach significantly cut down time to decision and reduced wasted traffic on ineffective variants.
2) Behavioral analytics revealed that drop-offs in scroll depth were a major reason for lower engagement in Variant A.
3) The system also offered natural language summaries for stakeholders, such as "Variant B improved conversions by 4% because of less navigation friction."

VI. SYSTEM ARCHITECTURE AND IMPLEMENTATION
This section details the practical setup and architecture of the proposed framework. It outlines how each component—predictive modeling, the multi-armed bandit engine, and behavioral analytics— interacts within the system.
1) Data Ingestion Layer: Collects user interaction logs, session data, and real-time feedback from the application interface.
2) Modeling Layer: Implements supervised learning models to produce early confidence scores for design variants.
3) Decision Engine: Hosts the multi-armed bandit algorithms that dynamically update traffic allocation based on continuous learning.
4) Analytics Dashboard: Presents stakeholders with visualizations of results, including conversion rates, engagement metrics, and natural-language summaries
5) Feedback Loop: Ensures the system is adaptive, with retraining of models as new data accumulates to account for evolving user behavior.

This architecture ensures modularity, scalability, and real-time responsiveness, making the framework suitable for large-scale deployments in e-commerce and mobile applications.

VII. ETHICAL CONSIDERATIONS IN ADAPTIVE UX TESTING
The integration of AI into A/B testing brings several ethical challenges that must be addressed to ensure fair, transparent, and responsible experimentation.
1) Bias in Machine Learning Models : Predictive models may unintentionally favor certain user groups if the training data is unbalanced, such as focusing more on desktop users than mobile users. This could lead to misleading results, where design variants perform well for one group but poorly for others. Ethical practice requires diverse training datasets and bias detection methods to ensure inclusivity.

2) Fair Traffic Allocation : Multi-armed bandit algorithms dynamically assign more traffic to better- performing variants. While this is efficient, it could mean that some users are disproportionately exposed to less effective or even frustrating designs during the early stages of testing. To maintain fairness, systems should implement minimum exposure thresholds and user safeguards to prevent prolonged poor experiences.
3) Transparency and User Consent : Users often do not realize they are part of an adaptive A/B testing process. Although consent is not always necessary for minor UI experiments, lacking transparency can raise ethical concerns if experiments significantly change user interactions. A best practice is to maintain disclosure policies or user-focused design guidelines that respect user choice while still allowing for improvement.
4) Data Privacy and Behavioral Analytics : Behavioral signals, such as cursor movements, hesitation patterns, and rage clicks, are sensitive indicators of cognitive and emotional states. Collecting this data requires strict privacy measures, including anonymization, GDPR compliance, and clear limits on data retention.

VIII. FUTURE WORK

While the proposed AI-enhanced A/B testing framework shows significant improvements in speed, efficiency, and depth of UX evaluation, there are several opportunities for further research and development:

1) Real-World Deployment : The current study is based on simulated user sessions. Future work will involve deploying the framework on large-scale, real- world platforms, such as e-commerce, ed-tech, healthcare, and finance, to validate performance under diverse conditions.
2) Integration with Design Tools : Incorporating this framework into popular UX/UI design platforms like Figma and Adobe XD will allow designers to conduct AI-driven testing directly during the design process. This will improve collaboration between researchers and practitioners.
3) Multi-Variant and Multivariate Testing : Expanding the system to handle A/B/n testing (multiple design variants) and multivariate testing (testing multiple UI elements simultaneously) will provide richer insights for complex design decisions.
4) Adaptive Personalization : Future enhancements may look into user-level personalization, where the system adjusts interface elements for individuals or groups instead of selecting a single global winner.
5) Ethical and Privacy Considerations : As behavioral analytics gather detailed user interaction data, such as cursor movements, scroll depth, and hesitation, future research must address ethical concerns, privacy protection, and transparent data usage policies to ensure responsible AI use.
6) Explainable AI Integration : Adding explainability layers to predictive models will help stakeholders understand why certain variants perform better than others, which will increase trust in AI-driven decision-making.

IX. CONCLUSION

This paper presented an AI-powered A/B testing framework for UX optimization. It integrates predictive modeling, adaptive allocation, and behavioral analytics. The results showed faster convergence, reduced traffic inefficiency, and more detailed insights compared to traditional methods.

By combining AI with human-centered design principles, this framework provides a promising direction for future UX optimization systems.

Furthermore, the experimental evaluation highlighted that the integration of predictive modeling allowed meaningful results to be obtained with smaller datasets. This reduced the time to decision-making while still maintaining statistical reliability. The use of multi-armed bandit algorithms further ensured that user traffic was dynamically shifted toward better- performing variants, minimizing the opportunity cost associated with exposing users to weak designs.

The inclusion of behavioral analytics proved particularly valuable, as it provided fine-grained insights into user interactions beyond conversion rates. Signals such as hesitation, scrolling depth, and cursor patterns allowed the framework to capture user engagement more holistically. These qualitative measures not only explained performance differences between design variants but also supported design teams in identifying specific usability bottlenecks.

Taken together, these findings suggest that AI- enhanced A/B testing represents a scalable and adaptive solution for digital platforms seeking to optimize user experience in real time. Its ability to balance efficiency, accuracy, and interpretability makes it a viable foundation for next-generation UX optimization tools, with applications extending across e-commerce, web applications, and mobile platforms.

REFERENCES

[1] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," Foundations and Trends in Machine Learning, vol. 5, no. 1, pp. 1–122, 2012.

[2] J. Xu, A. A. Lee, and J. Lee, "A machine learning approach to web A/B testing," ACM Transactions on Intelligent Systems and Technology, vol. 9, no. 2, pp. 1–21, 2018.

[3] D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison, "Hidden technical debt in machine learning systems," in Advances in Neural Information Processing Systems (NIPS), 2015.

[4] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in Proc. 19th Int. Conf. on World Wide Web (WWW), pp. 661–670, 2010

[5] G. Box, W. G. Hunter, and J. S. Hunter, Statistics for Experimenters: Design, Innovation, and Discovery, 2nd ed. Hoboken, NJ, USA: Wiley, 2005.

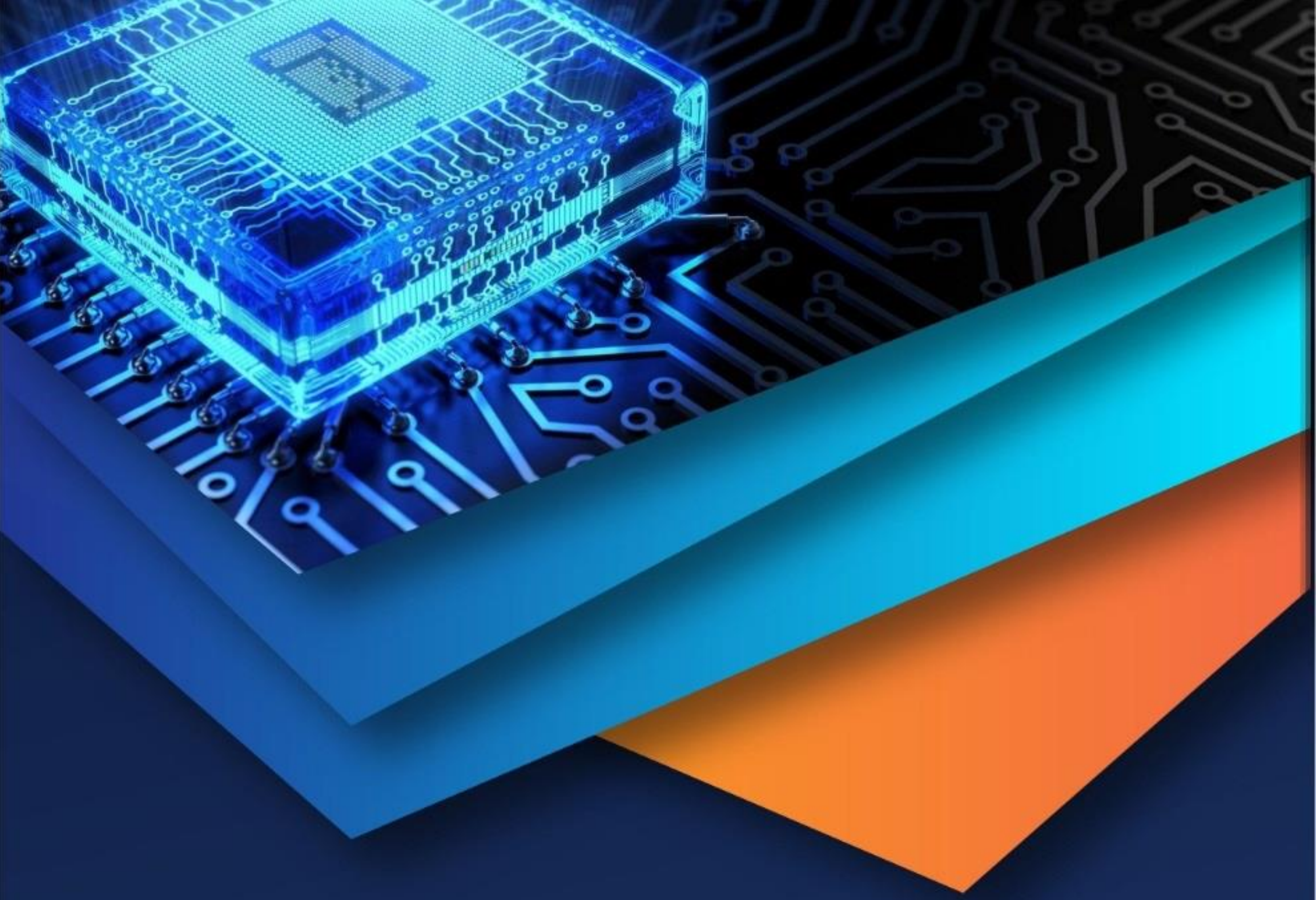
[6] J. Nielsen, Usability Engineering. San Francisco, CA, USA: Morgan Kaufmann, 1993.

[7] A. Toubia, J. Garcia, and E. G. Feinberg, "Probabilistic A/B testing: Experiment design and analysis," Marketing Science, vol. 38, no. 1, pp. 1–20, 2019.

[8] K. P. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press, 2012.

[9] J. Brooke, "SUS: A 'quick and dirty' usability scale," in Usability Evaluation in Industry, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, U.K.: Taylor & Francis, 1996, pp. 189–194.

[10] C. G. Bowsher and H. O. Stone, "Adaptive experimentation for web interfaces using bandit algorithms," Journal of Statistical Software, vol. 87, no. 5, pp. 1–23, 2018.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY