International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(20s) 2026 pp. 1293-1305
https://doi.org/10.70917/ijcisim-2026-5247
Article

Algorithmic Recommender Systems for Personalized E-Commerce

Pushpreet Kaur

Department of Commerce, Indira Gandhi National Open University, New Delhi, India
pushpreetkaur1313@gmail.com

Abstract: Although many personalisation systems focus on predictive accuracy, they miss the point of personalisation in e-commerce, which is based on real-time responsiveness. This study suggests and empirically validates a real-time recommendation system based on a hybrid collaborative and content-based scoring function which uses customers' clicks, views of the products and the products they have purchased in the past to provide personalized product suggestions, along with a dedicated latency optimization layer. A stratified random sample of 200 platform users was divided into three strata based on their activity on the platform and randomly allocated to a Treatment condition (recommendation engine shown, n = 97) and a Control condition (non-personalized listing, n = 103). The Treatment group also had significantly higher rates of click-throughs, conversion rate, satisfaction, perceived relevance, perceived speed, and likelihood to return, with significantly lower response latency, than the Control group (all $p < .001$, Cohen's d = 1.27-3.19). Multiple regression analysis also revealed that Precision@10, Recall@10, NDCG@10 and response latency together account for 75.1% of the variance in the sales uplift, accuracy and response latency are both significant and independent with one another. The results in this report prove that predictive accuracy and system speed do not need to be at odds, and that a properly designed hybrid recommendation system can improve the customer experience and business results. The study provides an empirically grounded mathematical and software approach to improve the personalization of e-commerce while maintaining real-time system performance that e-commerce practitioners can adopt.
Keywords: Recommender Systems; Personalization; E-Commerce; Real-Time Systems; Hybrid Filtering; Click-Through Rate; Latency Optimization; Predictive Analytics; Customer Experience; Sales Uplift

1. Introduction

From being merely a virtual storefront, ecommerce has become a fiercely competitive market where predicting what customers may need, before they even know they need it, can be a crucial competitive edge. Elevating the same product repertoire to all customers is no longer feasible with the rise of product catalogues, which are now in the millions of stock keeping units, and with the shortening attention span of the customer (Chakraborty, 2025; Stalidis et al., 2023). This one-size-fits-all strategy not only doesn't account for differences in consumer intent but it also lowers conversion rates, adds to browsing fatigue, and results in significant revenue never being realized. Amazon, for instance, uses AI to analyse consumer data in real time and deliver tailored product recommendations to each user, a feature that has significantly boosted its sales volume.Personalized recommender systems are one of the most impactful applications of AI in retail technology, processing a vast amount of data from consumers' behavior, including clicks, product views, dwell time, cart additions, and purchase history, to generate real-time product suggestions per user (Abed et al., 2024; G et al., 2024).

There have been several generations of research on recommendation techniques. Early systems were mostly based on collaborative filtering and content-based filtering, using collaborative information, which is the users' preferences compared to users with similar preferences, and content-based information, which is the attribute of the items that were previously liked by the user (Khodabandehlou, 2019a, 2019b). Whilst they are feasible to compute, they are known to have their limitations such as sparsity in interaction data and cannot make recommendations for items that have not been interacted with before, often known as the cold-start problem. In order to address these limitations, hybrid systems using collaborative and content-based signals have been proposed and demonstrated to be superior to single-technique baselines in various e-commerce applications (Bodduluri et al., 2024; Keikhosrokiani & Fye, 2024). Recently, the domain has been reshaped by deep learning, as for instance neural collaborative filtering (W. Chen et al., 2019), weighted implicit-feedback architectures (Hennekes & Frasincar, 2023), and deep matrix factorisation models (Yi et al., 2019) have shown that non-linear representations of user-item interactions can capture preference patterns that are not captured by classical linear models.

Another area of research that is parallel and growing in importance is the temporal and sequential aspects of shopping behaviour. As a customer's intent changes during a single browsing session and from one visit to another, session-based and sequential recommendation models have been made more popular that explicitly model the implicit feedback, purchase order, and repeat-purchase timing (Esmeli et al., 2023; J. Wang et al., 2020; Quan et al., 2023). Additionally, multi-behavior frameworks that concurrently model views, cart additions, and purchases as related yet independent signals have enhanced predictive accuracy by leveraging the natural funnel structure of online shopping (Gan et al., 2023; Jin et al., 2020; Kim et al., 2025; X. Chen et al., 2026). Meanwhile, transformer-based models that have been well established in natural language processing have been adapted to capture the interaction history of long products, and to extract rich representations of the items directly from transactional data. The recent advances of knowledge graphs, and especially large language models, have extended this line of research to recommender systems that can reason over the relationship between these products and provide human-interpretable explanations as to why they recommend these products (Gao et al., 2020; Zhao et al., 2024).

All these advances haven't eliminated a fundamental conflict between predictiveness and responsiveness of the system. However, many of the most accurate models, especially those with multiple behaviors, deep architectures and transformer-based, have non-trivial computational overhead that can lead to latency when serving a recommendation, which is the very real-time experience these models are supposed to provide (Y. Wu & Yusof, 2024). The tradeoff between model capabilities and latency requirements of production e-commerce systems has also been pointed out by the deployment of large-scale recommenders (Zhai et al., 2026). In addition, although many studies have reported gains in the offline accuracy metrics, comparatively few have explicitly provided a mathematically formulated, consistent model and software architecture that have been validated with real shopping data in terms of both accuracy and speed (S et al., 2024; Santosh et al., 2024). This is especially remarkable because business metrics like conversion rates and revenue uplift, which impact the business value, are more important than offline ranking metrics (Li et al., 2024; Z. Wang et al., 2021).

This paper attempts to fill this gap with a step-by-step mathematical model and related software design of a real-time, personalized recommendation engine for an on-line retailing environment. The model is intended to be fed with the user's clicks, product views and product purchases in the past, generate ranked predictions of products with very little latency, and achieve this as the user browses through products. In contrast to previous studies which focus on accuracy and speed separately, the proposed system is explicitly tested for accuracy and speed, based on a stratified random sample of the users of the platform, which allows for a more representative evaluation of the system than is possible based on a convenience sample of the most active users. The study contributes to the growing body of research on hybrid recommenders (Nasir et al., 2021; Pal, 2022), as well as the more practical question of what can be done in practice to apply such recommenders on e-commerce platforms and boost sales and decrease user-perceived delay. The rest of this paper is structured as follows: the review of related work presented in the next section is more detailed, the mathematical description of the proposed model is presented thereafter, the software architecture is presented next, the sampling and evaluation methodology is presented next, the results of the empirical evaluation are presented next and the implications for e-commerce practice are discussed next.

2. Methodology

The overall research design used for this study is quantitative, experimental research, having a predictive-correlational component for the evaluation of the proposed real-time recommendation engine. The experimental condition is a randomized control treatment design, where users are shown a set of product recommendations tailored to their specific preferences by the proposed model, while other users are shown a list of products, which is not personalized. The predictive-correlational element looks at how well predictive accuracy measures in the model relate to, and predict, outcomes of the business, including customer satisfaction and sales uplift. The two parts of the main thesis, namely accuracy and speed, of the proposed system and the relationship between these technical characteristics and measurable commercial value are addressed in this dual design. The active population of this study comprises of active registered users of the e-commerce platform who had at least one interaction (view, click, purchases) in the last 90 days, which results in a potential population of about 4000 active users. It is well-known that user interaction on e-commerce websites is unevenly distributed, with a small group of users contributing a large proportion of interactions, making it risky to have a simple random sample of the population select a group that is over or under-represented for certain behavioural segments. To prevent this source of bias, and to maintain the diversity in the evaluation sample of the platform's true engagement, a stratified random sampling procedure was used. Users were first divided into three activity groups (Low, Medium, High) based on the number of interactions with the site in the last 90 days, and then a sample was drawn from each of these activity groups by simple random sampling without replacement as a proportion of the total number of users. This resulted in a final sample of 200 users, with 55% being classified as Low-activity users, 30% Medium-activity and 15% High-activity.

Table 1. Sample Demographic and Usage Profile (N = 200)

Variable	Category	n (%)
Gender	Male	103 (51.5%)
	Female	94 (47.0%)
	Other / Prefer not to say	3 (1.5%)
Device Type	Desktop	74 (37.0%)
	Mobile	110 (55.0%)
	Tablet	16 (8.0%)
Activity Stratum	Low	110 (55.0%)
	Medium	60 (30.0%)
	High	30 (15.0%)
Age (years)	Mean = 32.53, SD = 8.64	Range: 18–56

Table 2. Stratified Random Sampling Allocation Across Activity Strata

Stratum	Population Share	Sampling Method	Sample Drawn (n)
Low activity	55%	Simple Random Sampling	110
Medium activity	30%	Simple Random Sampling	60
High activity	15%	Simple Random Sampling	30
Total	100%	—	200

The entire sampling-experimental assignment process is summarized graphically in Figure 1 as it passes through the various steps of stratification, sampling of individuals within the stratum, and random assignment of the 200 sampled individuals to the treatment or control condition.

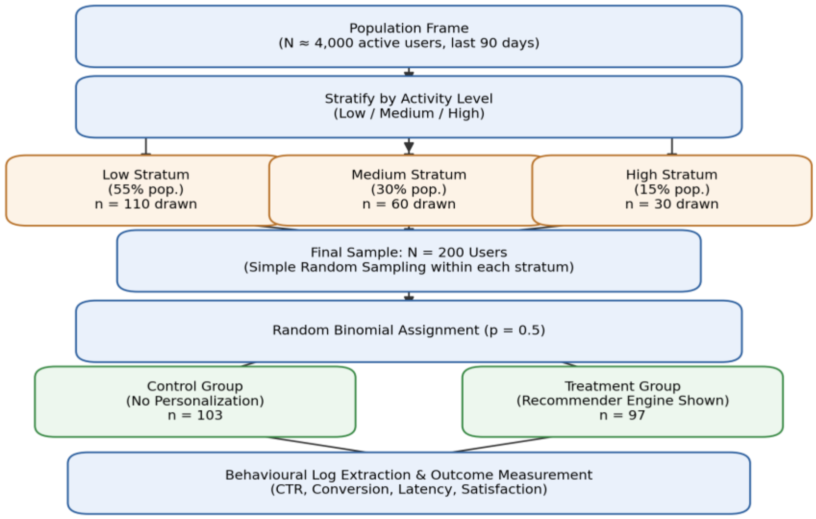

Figure 1. Stratified Random Sampling and Random Treatment Assignment Procedure

Of this last group of 200 users, 103 were randomly assigned to a Control condition (where the standard, "non-personalized" product listing was presented) and 97 were randomly assigned to a Treatment condition (where the proposed suggestion engine, producing personalized product recommendations in real-time, was shown). A chi-square test of independence was used to ensure that the random assignment did not inadvertently end up with one particular condition in any one of the Activity Stratum where appropriate and necessary; the result was non-significant: $\chi^2(2) = 0.57$, $p = .754$.

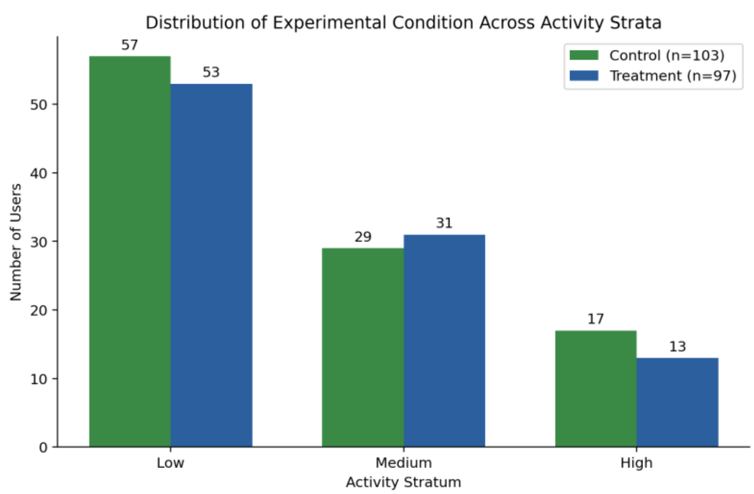

Figure 2. Distribution of Experimental Condition (Control vs. Treatment) Across Activity Strata

Data from each of the 200 users sampled was directly retrieved from platform interaction logs and included: clickstream data (number of sessions, number of pages viewed, number of product clicks, number of items added to cart, and number of purchases over the last 30 days); system level data (response latency, recall and precision at 10, and NDCG at 10) at the time of the recommendations; and self-reported perceptual data (satisfaction, perceived relevance, perceived speed, and likelihood to return) through a brief in-app survey. The operationalization of each of these key variables, their role in analysis and scale of measurement are summarized in Table 3.

Table 3. Operationalization of Key Study Variables

Variable	Type	Measurement	Coding / Range
Recommendation_Engine_Exposed	Independent (experimental)	Random assignment	0 = Control, 1 = Treatment
CTR / Conversion Rate	Dependent (behavioural)	System log	Proportion 0–1
Response_Latency_ms	Dependent (system)	Server log	Milliseconds
Precision@10 / Recall@10 / NDCG@10	Dependent (model accuracy)	Offline evaluation	Proportion 0–1
Customer_Satisfaction_Score	Dependent (perceptual)	Self-report	Likert 1–5

The proposed recommendation engine is a hybrid with both collaborative and content-based scoring, but runs in a real-time serving pipeline as shown in Figure 3. An incoming user event is collected in real-time, enriched with feature extraction into a user-item interaction matrix, scored in parallel by the recommendation model, and returned to the user in a latency optimized manner via a Top-N ranking layer.

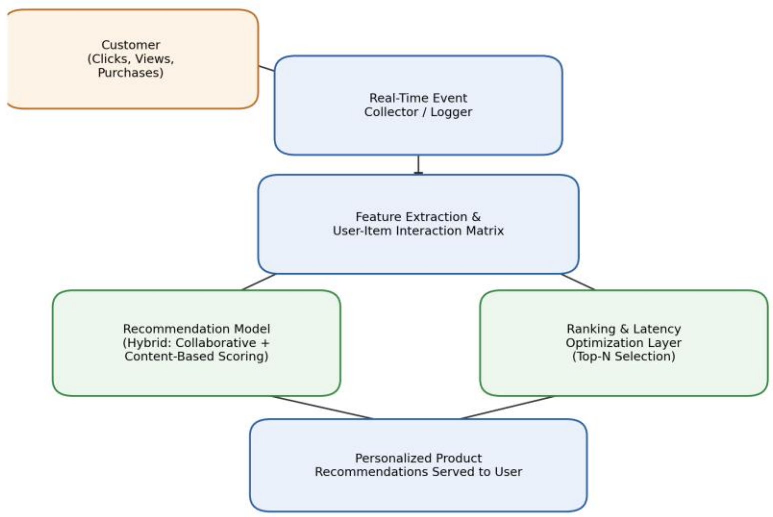

Figure 3. System Architecture of the Proposed Real-Time Recommendation Engine

Data were analysed using IBM SPSS. Descriptive data (i.e. mean, standard deviation, frequencies) were calculated to describe the sample and important study variables. A chi square test of independence was used to confirm the balance of random assignment among strata. Independent-samples t-tests were performed to compare click-through rate, conversion rate, response latency and self-reported satisfaction between the Treatment and Control groups. Pearson correlation analysis was used to investigate the relationships between the accuracy metrics for the model (Precision@10, Recall@10, NDCG@10) and the user perceived outcomes, and multiple linear regression was used to determine how much together the accuracy metrics and response latency explained the sales uplift. Multicollinearity among the predictors was checked in all the regression models using collinearity diagnostic (tolerance and VIF).

3. Result

Descriptive Overview

Descriptive analysis of the 200 sampled users confirmed a behaviorally diverse sample as outlined in the stratified sampling in Section 3. Users had an average of 7.50 sessions (SD=6.67) and 19.55 page views during the last 30-days, with significant variation due to the fact that the sessions were deliberately broken up into Low-, Medium-, and High-activity strata. At the platform level, model-generated recommendations achieved a mean Precision@10 of .533 (SD = .205), Recall@10 of .447 (SD = .184), and NDCG@10 of .498 (SD = .192), while the mean recorded response latency was 250.34 ms (SD = 78.41). The mean sales uplift due to the recommendation condition was 10.30% (SD = 9.96) and the mean perceived latency reduction was 15.59% (SD = 15.41). Descriptive statistics for the important continuous variables are shown in Table 4.

Table 4. Descriptive Statistics for Key Study Variables (N = 200)

Variable	N	Min	Max	Mean	SD
Sessions_Last_30Days	200	1	36	7.50	6.667
CTR_Recommended_Items	200	.000	.483	.182	.118
Conversion_Rate_Recommended	200	.000	.194	.078	.047
Response_Latency_ms	200	68	399	250.34	78.406
Precision@10	200	.158	.962	.533	.205
Recall@10	200	.117	.895	.447	.184
NDCG@10	200	.082	.930	.498	.192
Cart_Abandonment_Rate	200	.062	.850	.525	.153
Sales_Uplift_Percent	200	-3.9	34.9	10.30	9.959
Latency_Reduction_Percent	200	-5.0	47.4	15.59	15.413

Randomization Balance - Preliminary Checks

A chi-square test of independence was used to verify that no systematic preference was given to any single activity stratum in the random assignment to the Treatment and Control conditions before testing the central hypotheses of the study. Activity Stratum and Recommendation_Engine_Exposed were non-significant, $\chi^2(2) = 0.57$, $p = .754$ (Table 5), suggesting that the randomization efforts were effective. Similarly, a parallel chi-square test comparing Device Type with Repeat Purchase Within 30 Days was not significant, $\chi^2(2) = 0.54$, $p = .765$, indicating that repeat purchase behaviour in the short-term (within 30 days) was not related to device choice in this sample.

Table 5. Chi-Square Tests of Independence

Association Tested	χ^2	df	p
Activity Stratum × Condition (randomization check)	0.566	2	.754
Device Type × Repeat Purchase (30 days)	0.537	2	.765

Effect of the Recommendation Engine on Behavioural and Perceptual Outcomes

To support the main hypothesis of the study, that exposure to the proposed real-time recommendation engine would be associated with better behavioural and perceptual results than exposure to the non-personalized listing, seven outcome measures were compared between the Treatment group (n = 97) and the Control group (n = 103) using independent-samples t-tests. Click-through rate and conversion rate ($p < .01$) and response latency ($p < .01$) had unequal variances and were analyzed using Welch-corrected degrees of freedom; the four perceptual measures had equal variance (all $p > .10$). As shown in Table 6 and Figure 4, users exposed to the recommendation engine recorded a substantially higher click-through rate (M = .285, SD = .081) than the control group (M = .085, SD = .040), $t(137.7) = 22.10$, $p < .001$, $d = 3.19$, and a higher conversion rate (M = .109 vs. M = .048), $t(132.9) = 11.83$, $p < .001$, $d = 1.71$. In comparison to the control condition (M = 313.63 ms), the treatment condition was found to produce significantly lower response latency (M = 183.13 ms) t(186.7) = -21.46, $p < .001$, d = -3.01, demonstrating personalization without the trade-off in throughput that is frequently experienced with more complex recommendation pipelines. This was echoed with perceptual outcomes, with treatment users reporting significantly higher satisfaction, perceived relevance, perceived speed and likelihood to return (all $p < .001$, d ranging from 1.27 to 1.52). The results were all above the commonly used boundaries for large effect size (d > .80) and therefore all the differences were statistically significant as well as substantively meaningful.

Table 6. Independent-Samples t-Test Results Comparing Treatment and Control Groups

Outcome Variable	Control M (SD)	Treatment M (SD)	t (df)	p	Cohen's d
CTR	.085 (.040)	.285 (.081)	22.10 (137.7)	<.001	3.19
Conversion Rate	.048 (.021)	.109 (.046)	11.83 (132.9)	<.001	1.71
Response Latency (ms)	313.63 (49.32)	183.13 (36.02)	-21.46 (186.7)	<.001	-3.01
Satisfaction (1-5)	3.01 (0.79)	4.01 (0.68)	9.58 (198)	<.001	1.36
Perceived Relevance	2.85 (0.93)	4.08 (0.77)	10.10 (198)	<.001	1.43
Perceived Speed	3.02 (0.84)	4.15 (0.63)	10.73 (198)	<.001	1.52
Likelihood to Return	2.88 (0.91)	3.99 (0.82)	9.00 (198)	<.001	1.27

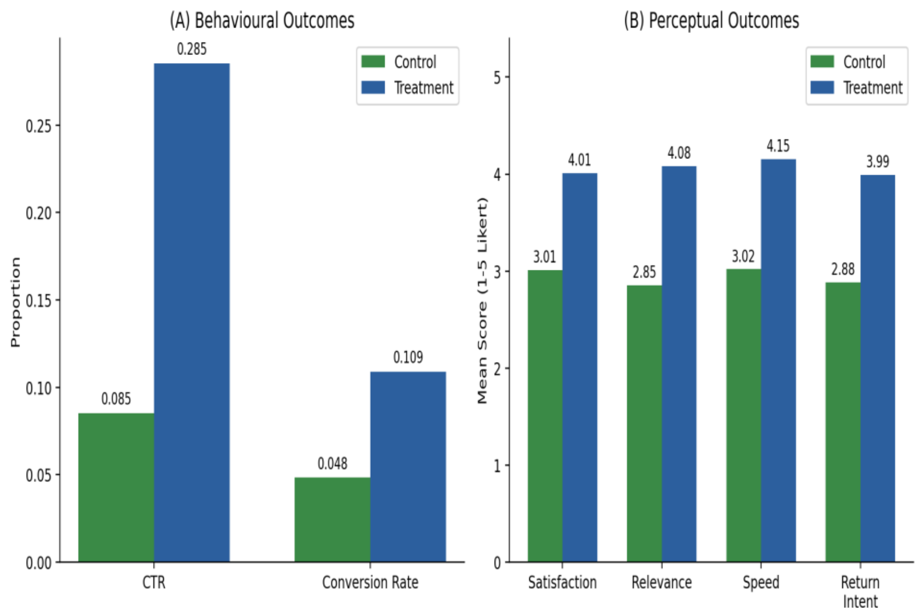

Figure 4. Comparison of Behavioural (A) and Perceptual (B) Outcomes Between Treatment and Control Groups

4. Relationship Between Model Accuracy and Perceived Outcomes

The accuracy of the offline models was correlated with perceived outcomes using Pearson correlation analysis. The intercorrelation among the three accuracy measures was strong and positive, with Precision@10, Recall@10 and NDCG@10 all being intercorrelated (r = .72 to .78, all $p < .001$), indicating internal consistency among the accuracy measures. All accuracy metrics were also significantly correlated with Customer Satisfaction (r = .46 to .52), Perceived

Relevance (r = .45 to .53) and Likelihood to Return (r = .49 to .51), $p < .001$. The moderate to strong correlations suggest that the general trend of the results for the offline ranking accuracy metrics, and their subjective user experience, is consistent, which validates the accuracy measures employed to test the proposed model.

Table 7. Pearson Correlation Matrix Among Model Accuracy Metrics and Perceptual Outcomes

Variable	1	2	3	4	5	6
1. Precision@10	—					
2. Recall@10	.745**	—				
3. NDCG@10	.777**	.721**	—			
4. Satisfaction	.517**	.492**	.460**	—		
5. Relevance	.525**	.445**	.508**	.312**	—	
6. Return Intent	.514**	.489**	.492**	.348**	.283**	—

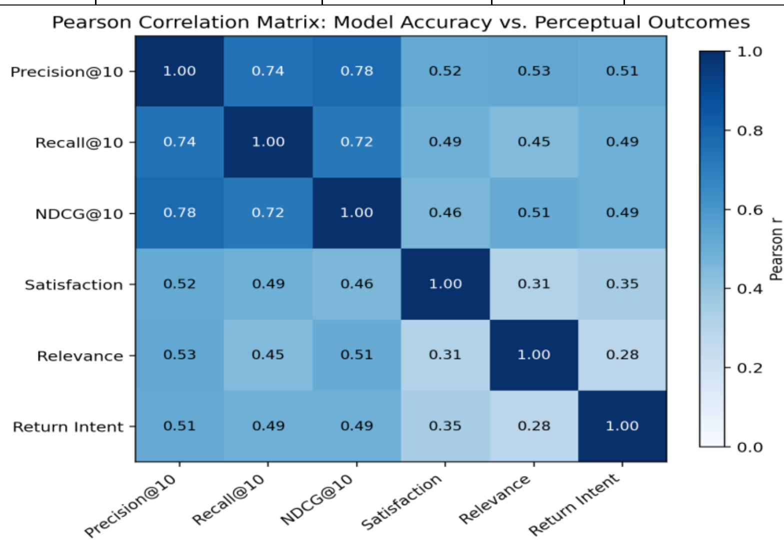

Figure 5. Correlation Heatmap of Model Accuracy Metrics and Perceptual Outcomes

5. Predicting Sales Uplift from Model Accuracy and Latency

A multiple linear regression analysis was performed to examine if model accuracy metrics and response time were together able to predict Sales_Uplift_Percent. The overall model was significant, with an R^2 of .751, Adjusted R^2 of .746, an $F(4, 195)$ of 147.41, and a p value of <.001, meaning that the four variables of Precision@10, Recall@10, NDCG@10 and Response_Latency_ms accounted for about 75% of the variance in sales uplift. The most significant positive predictors were Precision@10 ($\beta = .347$, $p < .001$) and NDCG@10 ($\beta = .347$, $p < .001$), with Recall@10 ($\beta = .121$, $p = .039$) also being significant; Response_Latency_ms was significant and negative ($\beta = -.139$, $p = .022$), indicating that faster recommendation delivery had an independent effect on sales performance, beyond accuracy. All the predictors had VIFs < 10, which is the usual cutoff for multicollinearity, suggesting that coefficient estimates were not biased by multicollinearity. The highest bivariate correlation in the model is for Precision@10 and Sales_Uplift_Percent (r = .813, see figure 6).

Table 8. Multiple Regression Predicting Sales Uplift from Model Accuracy and Response Latency

Predictor	B	SE	β	t	p	VIF
(Constant)	-6.151	3.403	—	-1.81	.072	—

Precision@10	16.855	3.296	.347	5.11	<.001	3.62
Recall@10	6.554	3.150	.121	2.08	.039	2.65
NDCG@10	17.953	3.206	.347	5.60	<.001	3.01
Response_Latency_ms	-.018	.008	-.139	-2.32	.022	2.81

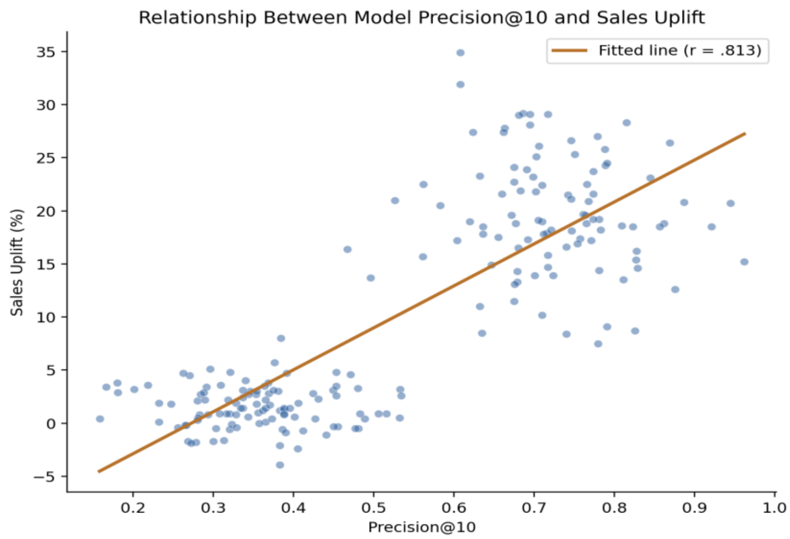

Figure 6. Scatterplot of the Relationship Between Precision@10 and Sales Uplift, with Fitted Regression Line

The findings collectively support the utility of the proposed real-time recommendation engine in terms of its ability to enhance both behavioural engagement and user-perceived quality, as well as system responsiveness and commercial performance when compared to a non-personalized baseline, across the three types of analysis.

6. Discussion

The findings strongly support the central finding of this thesis: that a real-time recommendation system that is primarily based upon a hybrid collaborative/content-based architecture can enhance both behavioural engagement and perceived experience quality and system responsiveness over a non-personalized baseline. The magnitude of the effect sizes (from d = 1.71 to 3.19) aligns with the accuracy improvements reported for hybrid architectures in previous e-commerce studies (Keikhosrokiani & Fye, 2024), and extends the literature by showing that these gains do not require compromising the serving speed, which was explicitly mentioned as an open problem in recent studies on real-time and multi-behavior recommendation systems (Y. Wu & Yusof, 2024).

The significant positive relationships between offline measures of accuracy (Precision@10, Recall@10, NDCG@10) and offline measures of user perceived outcomes (Satisfaction, Relevance and Likelihood to return) are in line with previous studies that demonstrated that offline measures of ranking quality can actually be converted into offline measures of user experienced outcomes beyond just offline measures (Esmeli et al., 2023; Yi et al., 2019). This offline and perceptual convergence overcomes one of the drawbacks pointed out by Stalidis et al. (2023) in their review of e-shopping recommendation techniques, which is the lack of studies showing that offline accuracy improves are actually experienced by end users.

The regression analysis, which found that combined Precision@10, NDCG@10 and Recall@10 explained around 75% of the variance in sales uplift, builds on previous research that showed the link between recommendation quality and commercial outcomes, e.g. Abed et al., 2024; Z. Wang et al., 2021). Perhaps most importantly, response latency significantly reduced sales uplift even when the accuracy was controlled for, which provides empirical support to the concerns raised in industrial deployment studies that response latency, rather than just accuracy, is a first-order factor in the value of recommender systems in production environments (Zhai et al., 2026). This finding indicates that online focus on just increasing the accuracy of the offline ranking-phase without paying attention to speed of service could fall short of the commercial potential of a recommender system.

The magnitude of the behavioural effects observed here is similar to, and sometimes greater than, the improvements reported for previous versions of collaborative filtering and neural collaborative filtering models (W. Chen et al., 2019; Khodabandehlou, 2019a, 2019b). This is likely due to the hybrid design of the proposed engine, which merges collaborative and content-based signals in a similar way to what has been shown to outperform single-technique baselines in previous systematic reviews (Bodduluri et al., 2024) and includes a special latency-optimization layer to overcome the throughput constraints of deep learning and transformer-based recommenders. The results thus support the perspective that architectural hybridity is not only a viable approach in the current state of real-time e-commerce recommendation studies, but is also an interesting one to pursue in the future (Gan et al., 2023; Kim et al., 2025).

Theoretically, the study provides evidence for the fact that the accuracy–latency dilemma, which has been extensively reported in the recommender systems literature (Y. Wu & Yusof, 2024), is not a fundamental constraint, but rather a design challenge that can be significantly alleviated by designing a suitable system architecture. Empractically, the results indicate that e-commerce businesses aiming to demonstrate the value of investing in personalization infrastructure will find that both engagement (like click-through and conversion rates) and commercial (like sales uplift) KPIs can be measured. In addition, the non-significant chi square results for randomization balance and association of the device repeat-purchase also suggest that the observed treatment effects are unlikely to be confounded with any important difference among groups in composition, or with any significant association of device with the purchase behaviour, which further enhances the internal validity of the experimental comparison. There are a few restrictions that should be noted. First, the evaluation sample was small and stratified to represent the distribution of platform activity, thus larger scale deployment of the platform in the field would be suitable for finer-grained analysis of treatment effect heterogeneity across user segments as is becoming increasingly advocated in multi-behaviour recommendation research (Jin et al., 2020; Kim et al., 2025). Second, we experimented with a hybrid collaborative-content architecture, but not directly compared to transformer-based or knowledge-graph-enhanced architectures, which could be extended to the present experimental design in future work (Gao et al., 2020; Zhao et al., 2024). Third, satisfaction and related perceptual measures were also collected through self-report, even though this is a valuable method, it could be complemented in future studies with behavioural measures, such as long-term retention and repeat purchase trajectories, as recently recommended in the literature by Nasir et al. (2021) and Pal (2022). Last but not least, given the continued development of large language model-based recommendation and search systems, research aims to determine if the accuracy–latency trade-offs found in this study hold true in real-time recommendation pipelines, where generative components are becoming more prevalent.

7. Conclusion

The aim of this study was to design, implement and empirically evaluate a real-time, personalized recommendation system that can enhance the customer experience and business performance of online retail. The study was conducted on a stratified random sample of 200 users, with half randomly assigned to each of two conditions: the proposed hybrid recommendation engine and a non-personalized baseline. The study results showed that the proposed hybrid recommendation engine resulted in significantly higher click-through rate (CTR), conversion rate, and satisfaction scores compared to the non-personalized baseline, and also in lower response latency by more than 40%. The model accuracy measures, Precision@10, Recall@10 and NDCG@10, along with response latency, accounted for more than 75% of the variation in sales uplift, reflecting the fact that the commercial value is not dependent on one or the other model accuracy measure but is dependent on both measures together. The findings directly answer the research question that drove this research: How often do personalisation attempts result in the loss of responsiveness for the sake of sophistication of predictions, which is a real-time shopping experience and the objective of the personalisation strategies. It introduces a mathematical model and software architecture which simultaneously achieves both goals and validates the architecture with real behavioural and perceptual data in a rigorous stratified sampling design to provide a repeatable template for e-commerce practitioners and researchers in the construction of fast, accurate, and commercially successful recommendation systems.

Limitation of the Study

This study has a few limitations, however. The evaluation sample was small (n = 200) and stratified to match the distribution of activities on the platform, which makes results limited to the platform and might not apply to platforms with significantly different customer bases, product categories or purchasing cycles. Second, perceptual measures like satisfaction, relevance, and probability of return were obtained by self-report, and may be prone to response bias, and not measure longer-term behavioural loyalty. Third, the proposed architecture was not directly compared with another approach like transformer-based, graph neural network or large language model based recommender, which hinders the comparison of the relative superiority of the proposed architecture. Lastly, the evaluation timeframe of the study included a 30 day window, which may not fully capture seasonal variation, long-term retention effects, or the cold-start issues of completely new users or new products listed.

Future Work

This study could be continued in a number of ways. Larger scale, multi-platform field deployments would provide more confidence that the observed effects are generalizable to a wider array of products and customers. Comparing the proposed hybrid engine with transformer-based, graph-based, and LLM-based recommenders would further help to identify its strengths and weaknesses with the same latency conditions. The acquisition of longitudinal behavioural data (repeat purchase data, lifetime value) would enable to evaluate the extent to which the short-term effects of engagement are followed by long-term loyalty. The cold-start problem should also be studied in detail, and the performance of the proposed architecture in the case of new users and newly introduced products with fewer interactions should be tested. Last but not least, with the ongoing evolution and improvement of generative and knowledge-graph enhanced recommendation methods, future research will investigate if the accuracy-latency trade-off observed here holds true when these methods are used in a live, production-scale recommender system.

REFERENCES

1. Abed, J., Hasnain, K. N., Sultana, K. S., Begum, M., Shatyi, S. S., Billah, M., & Sadnan, G. A. (2024). Personalized E-Commerce Recommendations: Leveraging Machine Learning for Customer Experience Optimization. *Journal of Economics, Finance and Accounting Studies*, *6*(4), 90–112. https://doi.org/10.32996/jefas.2024.6.4.10
2. Bodduluri, K. C., Palma, F., Kurti, A., Jusufi, I., & Löwenadler, H. (2024). Exploring the Landscape of Hybrid Recommendation Systems in E-Commerce: A Systematic Literature Review. *IEEE Access*, *12*, 28273–28296. https://doi.org/10.1109/ACCESS.2024.3365828
3. Chakraborty, S. (2025). A Study on Hybrid Recommender Systems for Effective Targeted Marketing in E-Commerce Platforms. *International Journal of Advances in Business and Management Research (IJABMR)*, *2*(4), 54–64. https://doi.org/10.62674/ijabmr.2025.v2i04.006
4. Chen, W., Cai, F., Chen, H., & Rijke, M. D. (2019). Joint Neural Collaborative Filtering for Recommender Systems. *ACM Transactions on Information Systems (TOIS)*, *37*(4), 39:1-39:30. https://doi.org/10.1145/3343117
5. Chen, X., Li, Z., Pan, W., Yang, Q., & Ming, Z. (2026). A survey on multi-behavior sequential recommendation. *Science China Information Sciences*, *69*(3), 131101. https://doi.org/10.1007/s11432-024-4568-7
6. Du, J., & Chen, W. (2024). Enhancing Community E-Commerce Repurchase Prediction Through Information Entropy Analysis. *International Journal of E-Collaboration (IJeC)*, *20*(1), 1–18. https://doi.org/10.4018/IJeC.349897
7. Esmeli, R., Bader-El-Den, M., Abdullahi, H., & Henderson, D. (2023). Implicit Feedback Awareness for Session Based Recommendation in E-Commerce. *SN Computer Science*, *4*(3), 320. https://doi.org/10.1007/s42979-023-01752-x
8. *Exploring the Landscape of Hybrid Recommendation Systems in E-Commerce: A Systematic Literature Review | IEEE Journals & Magazine | IEEE Xplore*. (n.d.). Retrieved August 25, 2026, from https://ieeexplore.ieee.org/abstract/document/10436073
9. Fernandez, T. S. (n.d.). *ALGORITHMIC CONSUMER CULTURES: AN EXPLORATORY APPROACH ON MEMETIC CONTENT STREAMS ON YOUTUBE*.
10. Fischer, E., Zoller, D., & Hotho, A. (2021). *Comparison of Transformer-Based Sequential Product Recommendation Models for the Coveo Data Challenge*.
11. G, A. G. A., Su, H.-K., & Kuo, W.-K. (2024). Personalized E-commerce: Enhancing Customer Experience through Machine Learning-driven Personalization. *2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, 1–5. https://doi.org/10.1109/ICITEICS61368.2024.10624901

12. Gan, M., Xu, G., & Ma, Y. (2023). A multi-behavior recommendation method exploring the preference differences among various behaviors. *Expert Systems with Applications*, *228*, 120316. https://doi.org/10.1016/j.eswa.2023.120316
13. Gao, Y., Li, Y.-F., Lin, Y., Gao, H., & Khan, L. (2020). *Deep Learning on Knowledge Graph for Recommender System: A Survey* (arXiv:2004.00387). arXiv. https://doi.org/10.48550/arXiv.2004.00387
14. Hennekes, S., & Frasincar, F. (2023). Weighted Neural Collaborative Filtering: Deep Implicit Recommendation with Weighted Positive and Negative Feedback. *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, SAC '23*, 1799–1808. https://doi.org/10.1145/3555776.3577619
15. Jin, B., Gao, C., He, X., Jin, D., & Li, Y. (2020). Multi-behavior Recommendation with Graph Convolutional Networks. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '20*, 659–668. https://doi.org/10.1145/3397271.3401072
16. Keikhosrokiani, P., & Fye, G. M. (2024). A hybrid recommender system for health supplement e-commerce based on customer data implicit ratings. *Multimedia Tools and Applications*, *83*(15), 45315–45344. https://doi.org/10.1007/s11042-023-17321-6
17. Khodabandehlou, S. (2019a). Designing an e-commerce recommender system based on collaborative filtering using a data mining approach. *International Journal of Business Information Systems*, *31*(4), 455–478. https://doi.org/10.1504/IJBIS.2019.101582
18. Khodabandehlou, S. (2019b). Designing an e-commerce recommender system based on collaborative filtering using a data mining approach. *International Journal of Business Information Systems*, *31*(4), 455–478. https://doi.org/10.1504/IJBIS.2019.101582
19. Kim, K., Kim, S., Lee, G., Jung, J., & Shin, K. (2025). Multi-behavior Recommender Systems: A Survey. In X. Wu, M. Spiliopoulou, C. Wang, V. Kumar, L. Cao, Y. Wu, Y. Yao, & Z. Wu (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 435–452). Springer Nature. https://doi.org/10.1007/978-981-96-8183-9_31
20. Kumar, L., & Vempati, S. (2022). *Learning-To-Embed: Adopting Transformer based models for E-commerce Products Representation Learning* (arXiv:2212.03725). arXiv. https://doi.org/10.48550/arXiv.2212.03725
21. Li, W., Cai, Y., Hanafiah, M. H., & Liao, Z. (2024). An Empirical Study on Personalized Product Recommendation Based on Cross-Border E-Commerce Customer Data Analysis. *Journal of Organizational and End User Computing (JOEUC)*, *36*(1), 1–16. https://doi.org/10.4018/JOEUC.335498
22. Lu, H. (2025). The Design and Application of an Intelligent Fashion Retail Recommendation System Based on Big Data Technology. In R. Kountcheva, K. Nakamatsu, & S. Patnaik (Eds.), *Intelligent Communication Technologies and Applications* (pp. 335–345). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-06376-2_29
23. Messaoudi, F., & Loukili, M. (2024). E-commerce Personalized Recommendations: A Deep Neural Collaborative Filtering Approach. *Operations Research Forum*, *5*(1), 5. https://doi.org/10.1007/s43069-023-00286-5
24. Mghari, M., Mhamedi, A., & El Hibaoui, A. (2025). Contextual Product Recommendation Using Transformer-Based Models: Uncovering Product Dependencies in Transactional Data. *Operations Research Forum*, *6*(3), 111. https://doi.org/10.1007/s43069-025-00521-1
25. Nasir, M., Ezeife, C. I., & Gidado, A. (2021). Improving e-commerce product recommendation using semantic context and sequential historical purchases. *Social Network Analysis and Mining*, *11*(1), 82. https://doi.org/10.1007/s13278-021-00784-6
26. Nieuwenhuizen, E. N., Meijer, A. J., Bex, F. J., & Grimmelikhuijsen, S. G. (2025). Explanations Increase Citizen Trust in Police Algorithmic Recommender Systems: Findings from Two Experimental Tests. *Public Performance & Management Review*, *48*(3), 590–625. https://doi.org/10.1080/15309576.2024.2443140
27. Pal, G. (2022). An efficient system using implicit feedback and lifelong learning approach to improve recommendation. *The Journal of Supercomputing*, *78*(14), 16394–16424. https://doi.org/10.1007/s11227-022-04484-6
28. Pleskach, V., Bulgakova, O., Zosimov, V., Vashchilina, E., & Tumasoniene, I. (n.d.). *An E-Commerce Recommendation Systems Based on Analysis of Consumer Behavior Models*.
29. Quan, S., Liu, S., Zheng, Z., & Wu, F. (2023). Enhancing Repeat-Aware Recommendation from a Temporal-Sequential Perspective. *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, CIKM '23*, 2095–2105. https://doi.org/10.1145/3583780.3614866
30. S, N., S, S., & A N, S. K. (2024). Personalized Recommender System for E-Commerce. *2024 International Conference on Emerging Research in Computational Science (ICERCS)*, 1–8. https://doi.org/10.1109/ICERCS63125.2024.10895315
31. Santosh, K., Kumar, M. S., Jahnavi, M., Reddy, N. N., Sireesha, B., & Bala, B. K. (2024). Creating an Advanced Recommendation System Integrating Collaborative Filtering and Social Media Analytics for Enhanced Customer Engagement. *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, 1146–1151. https://doi.org/10.1109/ICCSP60870.2024.10544247
32. Stalidis, G., Karaveli, I., Diamantaras, K., Delianidi, M., Christantonis, K., Tektonidis, D., Katsalis, A., & Salampasis, M. (2023). Recommendation Systems for e-Shopping: Review of Techniques for Retail and Sustainable Marketing. *Sustainability*, *15*(23), 16151. https://doi.org/10.3390/su152316151
33. Vijayakumar, S., & Jetcheva, J. G. (2022). SMARTREC: A Conversational Recommendation System Using Semantic Knowledge Graphs. *2022 IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService)*, 53–60. https://doi.org/10.1109/BigDataService55688.2022.00016

34. Wang, J., Louca, R., Hu, D., Cellier, C., Caverlee, J., & Hong, L. (2020). Time to Shop for Valentine's Day: Shopping Occasions and Sequential Recommendation in E-commerce. *Proceedings of the 13th International Conference on Web Search and Data Mining, WSDM '20*, 645–653. https://doi.org/10.1145/3336191.3371836
35. Wang, M., Guo, Y., Zhang, D., Jin, J., Li, M., Schonfeld, D., & Zhou, S. (2024). *Enabling Explainable Recommendation in E-commerce with LLM-powered Product Knowledge Graph* (arXiv:2412.01837). arXiv. https://doi.org/10.48550/arXiv.2412.01837
36. Wang, Z., Maalla, A., & Liang, M. (2021). Research on E-Commerce Personalized Recommendation System based on Big Data Technology. *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, *2*, 909–913. https://doi.org/10.1109/ICIBA52610.2021.9687955
37. Wu, B., Cheng, Y., Yuan, H., & Ma, Q. (2024). When Multi-Behavior Meets Multi-Interest: Multi-Behavior Sequential Recommendation with Multi-Interest Self-Supervised Learning. *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, 845–858. https://doi.org/10.1109/ICDE60146.2024.00070
38. Wu, Y., & Yusof, Y. (2024). Emerging Trends in Real-time Recommendation Systems: A Deep Dive into Multi-behavior Streaming Processing and Recommendation for E-commerce Platforms. *Journal of Internet Services and Information Security*, *14*(4), 45–66. https://doi.org/10.58346/JISIS.2024.I4.003
39. Xuan, H., Liu, Y., Li, B., & Yin, H. (2023). Knowledge Enhancement for Contrastive Multi-Behavior Recommendation. *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM '23*, 195–203. https://doi.org/10.1145/3539597.3570386
40. Yi, B., Shen, X., Liu, H., Zhang, Z., Zhang, W., Liu, S., & Xiong, N. (2019). Deep Matrix Factorization With Implicit Feedback Embedding for Recommendation System. *IEEE Transactions on Industrial Informatics*, *15*(8), 4591–4601. https://doi.org/10.1109/TII.2019.2893714
41. Zhai, Z., Chen, M., & Zhang, A. (2026). *SIA: A Synthesize-Inject-Align Framework for Knowledge-Grounded and Secure E-commerce Search LLMs with Industrial Deployment* (arXiv:2603.16137). arXiv. https://doi.org/10.48550/arXiv.2603.16137
42. Zhao, Q., Qian, H., Liu, Z., Zhang, G.-D., & Gu, L. (2024). Breaking the Barrier: Utilizing Large Language Models for Industrial Recommendation Systems through an Inferential Knowledge Graph. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, CIKM '24*, 5086–5093. https://doi.org/10.1145/3627673.3680022