International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V73I4P116

Volume 73 Issue 4, 114-119, April 2025

Original Article

Personalized Pricing Based on Behavioral Signals: Revenue Uplift and Fairness Tradeoffs in E-Commerce

Anmol Aggarwal

Haas School of Business, University of California, Berkeley, California, USA.

Corresponding Author : anmol_aggarwal@berkeley.edu

Received: 17 March 2025 Revised: 16 April 2025 Accepted: 23 April 2025 Published: 30 April 2025

Abstract - *Personalized pricing strategies increasingly rely on machine learning to tailor discounts based on user behavior. While these methods promise improved conversion and revenue, their fairness and targeting precision remain underexplored. In this study, we simulate two behavior-based pricing policies, one broad and one selective, using session-level e-commerce data and estimate their impact using uplift modeling techniques.We implement and compare the TwoModel (T-Learner) and Class Transformation frameworks to evaluate treatment effectiveness. Our results show that stricter treatment rules, though applied to fewer users, yield higher conversion uplift and revenue per treated session. However, they also exclude many high-opportunity users, raising fairness concerns.By integrating causal inference, pricing simulation, and behavioral fairness analysis, this paper highlights the tradeoffs between targeting precision, business value, and equitable incentive distribution. We propose uplift modeling as a robust foundation for building fair and profitable personalized pricing systems.*

Keywords - *Behavioral pricing, Conversion uplift, Machine learning, Personalized pricing, Revenue optimization.*

1. Introduction

In the age of algorithmic commerce, personalized pricing has emerged as a powerful strategy for maximizing revenue and conversion by offering individualized discounts. By analyzing user behavior in real-time, businesses adjust prices or promotions dynamically, often targeting users deemed at risk of not converting. While this machine learning-driven approach holds economic promise, it also raises critical concerns about effectiveness, transparency, and fairness. Many real-world pricing systems continue to rely on simplistic heuristics—such as whether a user is new or has a high bounce rate—without estimating how pricing interventions actually affect behavior. These signals may reflect urgency, disinterest, or loyalty, but their interpretation is often misaligned with business objectives. As a result, such systems risk either excluding high-intent users from discounts or misallocating incentives to low-probability converters. While personalized pricing is increasingly adopted in e-commerce, most implementations rely on non-causal, rule-based targeting rather than estimating the true incremental benefit of an offer. This introduces both revenue inefficiencies and fairness risks. Prior research focuses heavily on predictive response modeling [1, 3], yet few studies apply uplift modeling—a causal inference framework—to simulate and evaluate targeting strategies. This study addresses that gap.We apply uplift modeling to estimate the incremental effect of a price incentive on conversion rather than simply predicting purchase likelihood. Uplift models estimate the difference in outcome probability with vs. without treatment, making them well-suited for policy evaluation and targeting optimization.

This study makes three key contributions:

- We simulate two behavioral pricing policies using publicly available e-commerce data, reflecting real-world discount logic.
- We apply and compare two uplift modeling approaches—TwoModel and Class Transformation—to estimate conversion uplift and evaluate model robustness.
- We assess not only revenue outcomes but also fairness tradeoffs, asking: Are we targeting users who need support or just those who are easiest to influence?

By analyzing both the financial and ethical implications of pricing interventions, this study emphasizes the need for behavioral fairness in algorithmic pricing—ensuring that intent and loyalty are not penalized by targeting systems optimized only for click probability. Our findings offer practical guidance for data scientists and product managers building pricing systems that aim to be both profitable and justifiable.

2. Related Work

Personalized pricing has evolved significantly with the integration of machine learning techniques, enabling businesses to offer individualized prices based on real-time behavioral and contextual data. In their study on high-dimensional features and demand heterogeneity, Ban and Keskin [1] demonstrate that machine learning can model personalized elasticities and drive measurable improvements in pricing performance. Similarly, Baier and Stöcker [2] show that integrating uplift modeling into direct marketing campaigns leads to significant profit improvements by targeting customers based on estimated incremental gains rather than raw response rates. In e-commerce settings, El Youbi et al. [3] implement a machine learning-driven pricing strategy using Gradient Boosted Trees, showing notable predictive power when adjusting prices dynamically.

Their approach aligns with prescriptive strategies that not only estimate purchase likelihood but also consider optimal intervention thresholds. From a methodological perspective, uplift modeling has gained momentum as a causal technique for estimating individual treatment effects. Gutierrez and Gérardy [4] provide a foundational review of uplift modeling strategies such as the TwoModel and Class Transformation methods, which serve as the basis for our modeling framework. Lo and Pachamanova [6] build on this by developing prescriptive uplift analytics, integrating estimation uncertainty into pricing decision-making - a vital extension for real-world policy deployment where pricing errors carry direct costs.Fairness in algorithmic pricing has become an increasingly important research focus.

Kallus and Zhou [5] introduce a formal framework for evaluating equity and welfare in personalized pricing decisions, identifying tradeoffs between profit maximization and group-level fairness. Priester et al. [7] explore this issue from the consumer's viewpoint, finding that perceived fairness in dynamic pricing heavily depends on transparency and the framing of personalization. Zuiderveen Borgesius [9] critiques personalization through a regulatory lens, examining how algorithmic decision-making may violate anti-discrimination laws, particularly in the context of opaque targeting mechanisms. Complementing these theoretical contributions, open-source tools such as scikit-uplift [8] have made causal targeting techniques more accessible to practitioners. These tools are crucial for scaling experiments and evaluating personalization policies beyond correlation-based targeting.

Finally, Das et al. [12] present an engineering-focused study that applies machine learning to optimize real-time pricing in retail and e-commerce. While their work is largely predictive, it reinforces the practical momentum behind data-driven pricing systems. It signals a growing need for causal and fairness-aware pricing frameworks such as the one proposed in our study. Together, these works provide a strong foundation for our research, which contributes by simulating personalized discounting policies, estimating individual uplift, and analyzing behavioral fairness tradeoffs using real e-commerce data.

3. Methodology

3.1. Dataset and Preprocessing

We used the UCI Online Shoppers Purchasing Intention Dataset, which contains 12,330 user sessions. Each session includes user behavior metrics such as bounce rates, exit rates, page durations, number of product-related pages viewed, and session attributes like traffic source, month, and visitor type. Preprocessing steps included:

- One-hot encoding of Month and VisitorType
- Converting Revenue into binary (1 = purchase, 0 = no purchase)
- Standardizing boolean fields (Weekend) as integers
- Keeping bounded behavioral metrics like BounceRates in raw form.

The final dataset preserves behavioral signals necessary for realistic policy simulation.

3.2. Simulated Treatment Assignment

Since no real pricing interventions exist in the dataset, we simulated binary treatment T∈{0,1, where:

- T=1: user is treated (discount applied)
- T=0: the user is not treated

Policy 1 (Broad Targeting)

$$T_i^{(1)} = \begin{cases} 1 & \text{if VisitorType}_i = \text{New\_Visitor} \wedge \text{BounceRate}_i > Q_{0.75} \\ 0 & \text{otherwise} \end{cases}$$

Policy 2 (Selective Targeting)

$$T_i^{(2)} = \begin{cases} 1 & \text{if VisitorType}_i = \text{New\_Visitor} \wedge \text{BounceRate}_i > Q_{0.75} \wedge \text{ExitRate}_i > Q_{0.75} \\ 0 & \text{otherwise} \end{cases}$$

Where $Q_{0.75}$ is the 75th percentile of bounce/exit rates, these simulate targeting strategies based on disengagement, urgency, or perceived price sensitivity.

3.3. Uplift Modeling Approaches

We compare two widely used uplift modeling frameworks:

3.3.1. TwoModel (T-Learner)

- $f_1(X) = \mathbb{P}(Y = 1 \mid T = 1, X)$
- $f_0(X) = \mathbb{P}(Y = 1 \mid T = 0, X)$

The individual uplift score is computed as:

$$U(X) = f_1(X) - f_0(X)$$

This method allows high flexibility and has shown strong performance in heterogeneous treatment effect estimation (Gutierrez & Gérardy, 2017).

This approach transforms the target variable using treatment and outcome:

$$Z = \begin{cases} 1 & \text{if } T = 1 \wedge Y = 1 \\ 1 & \text{if } T = 0 \wedge Y = 0 \\ 0 & \text{otherwise} \end{cases}$$

A standard classifier is then trained on this transformed label. The predicted probability is interpreted as an uplift proxy. While this technique reduces model complexity, it may introduce noise if the transformation is not well aligned with the underlying causal structure.

3.4. Training Details

- All models implemented using scikit-learn and scikit-uplift
- Random Forest (100 estimators, balanced subsampling) used as base learner
- Training/test split: 70/30 stratified by treatment label
- Features used: behavioral and temporal (e.g., PageValues, BounceRates, Month_*, etc.)

We trained models separately under Policy 1 and Policy 2 to isolate the impact of treatment design.

3.5. Evaluation Metrics

We use three categories of metrics:

3.5.1. Qini AUC Score

A standard evaluation metric in uplift modeling that captures cumulative gain relative to a random baseline. Higher values indicate better identification of persuadable users.

3.5.2. Revenue Proxy

We simulate revenue with synthetic pricing:

$$\text{Price}_i = \begin{cases} 90 & \text{if treated} \\ 100 & \text{otherwise} \end{cases} \quad \Rightarrow \quad \text{RevenueProxy}_i = \text{Price}_i \cdot Y_i$$

This allows us to compute ROI tradeoffs across policies.

3.5.3. Fairness Metrics

To assess behavioral fairness, we compute the following:

- The top 20% uplift-scored users who were treated
- Average uplift score among treated users
- Treatment rates stratified by VisitorType

These provide insight into whether a policy favors disengaged users at the expense of loyal ones.

4. Results

We evaluated two behavioral treatment policies, Policy 1 (broad) and Policy 2 (selective), using TwoModel and Class Transformation uplift estimators. Each policy's effectiveness was measured using Qini AUC, revenue proxy, and fairness metrics.

4.1. Uplift Modeling Performance

To assess each model's ability to identify users most influenced by price incentives, we measured Qini AUC, a ranking-based metric widely used in uplift modeling.

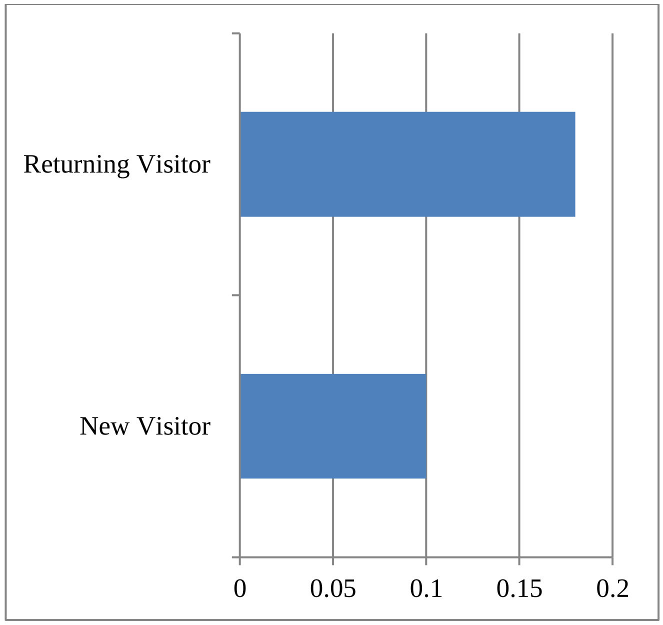

Fig. 1 Average Uplift by Visitor Type: Returning visitors had a significantly higher average uplift than new visitors, suggesting they were more likely to convert if treated—yet many were excluded under behavior-based targeting policies.

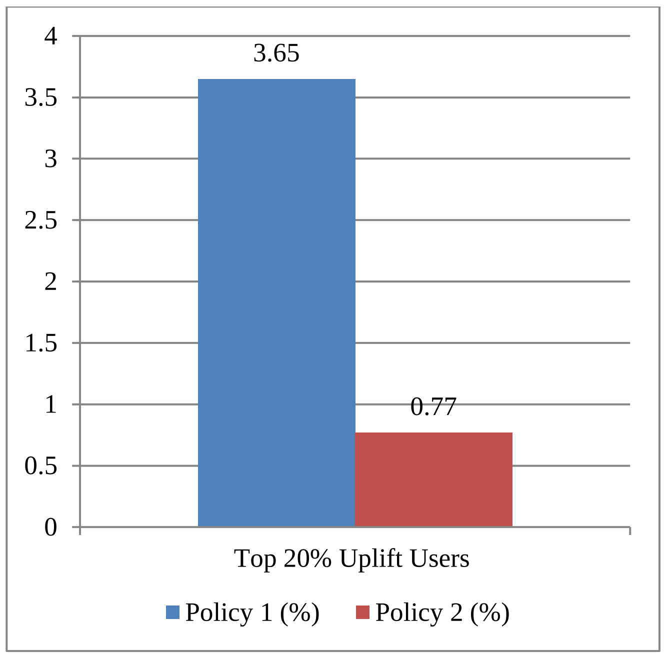

Fig. 2 Percentage of high-opportunity users (top 20% uplift) who received treatment under each policy. Policy 1 reached 3.65% of high-uplift users, while Policy 2 reached only 0.77%, highlighting a tradeoff between precision and coverage in targeting.

Table 1. Qini AUC Scores by Model and Policy

Model	Policy	Qini AUC
TwoModel	Policy 1	0.3784
TwoModel	Policy 2	**0.3817**
Class Transformation	Policy 2	0.3671

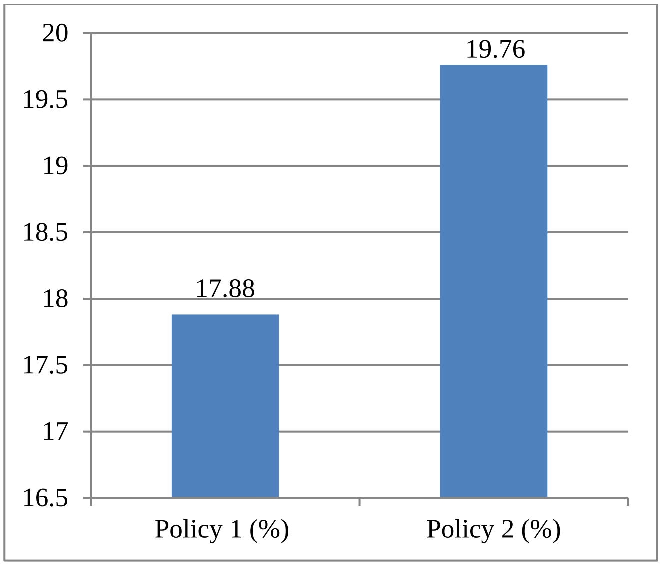

Fig. 3 Revenue Proxy by Treatment Group: Average revenue generated from treated sessions under each policy. Policy 2, though more selective, led to higher per-session revenue ($19.76 vs $17.88), supporting the effectiveness of precision-based targeting.

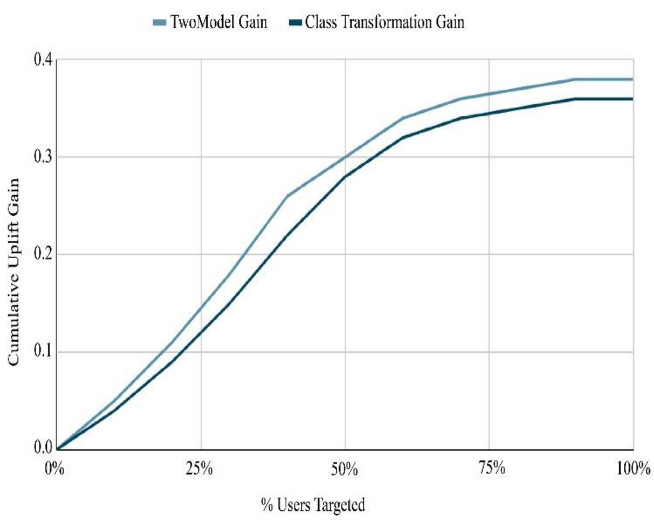

Fig. 4 Qini Curve – TwoModel vs Class Transformation: Qini curves comparing the uplift ranking performance of TwoModel and Class Transformation approaches. TwoModel achieved a marginally better uplift ranking throughout the targeting range.

Interpretation

Policy 2 achieved a higher Qini AUC than Policy 1 under both models despite treating fewer users. This indicates that more targeted pricing (using bounce and exit rates) better isolated persuadable users. TwoModel slightly outperformed Class Transformation, but both supported the same directional conclusions.

4.2. Revenue Impact of Pricing Policies

We simulated session-level revenue by assigning:

- $90 to treated users (received discount)
- $100 to untreated users (paid full price)

Table 2. Average Revenue by Policy and Treatment Group

	Avg Revenue (Treated)	Avg Revenue (Untreated)
Policy 1	$17.88	$15.56
Policy 2	**$19.76**	$15.65

Interpretation

Policy 2, though applied to only ~0.4% of sessions, generated significantly more revenue per treated user. This suggests that precision in discounting may be more profitable than widespread offers, particularly when behavioral filters isolate users most at risk of abandonment.

4.3. Fairness Metrics: Coverage vs Precision

To evaluate fairness, we measured:

- How many high-uplift users were treated
- The average uplift among those treated
- The treatment rate by VisitorType

Table 3. Fairness Metrics by Policy

	% Top 20% Uplift Treated	Avg Uplift (Treated Users)
Policy 1	3.65%	0.0296
Policy 2	**0.77%**	**0.0498**

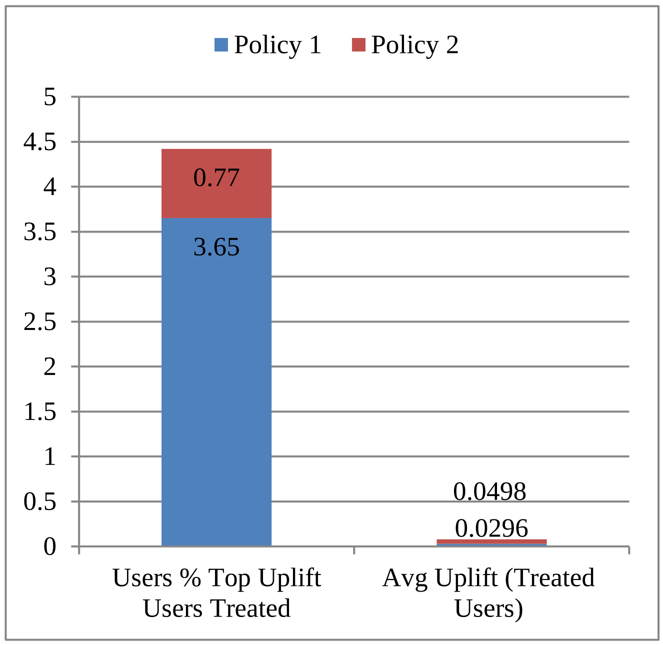

Fig. 5 Fairness Tradeoff – Coverage vs Precision: Comparison of two fairness metrics: the percentage of top uplift users treated (coverage) and the average uplift among those treated (precision). Policy 1 favored coverage; Policy 2 prioritized precision.

4.3.1. Interpretation

Policy 1 was more inclusive, and it treated more of the top 20% uplift users. However, Policy 2 had better uplift efficiency, treating fewer users with a much higher chance of converting. This reveals a coverage vs. precision tradeoff that has both ethical and business implications.

4.4. Summary of Key Tradeoffs

We simulated session-level revenue by assigning:

Metric	Policy 1 (Broad)	Policy 2 (Selective)
Treatment Coverage	Higher (~1.3%)	Lower (~0.4%)
Qini AUC (TwoModel)	0.3784	**0.3817**
Revenue per Treated Session	$17.88	**$19.76**
% Top Uplift Treated	3.65%	0.77%
Avg Uplift (Treated)	0.0296	**0.0498**
Business Efficiency	Moderate	**High**
Fairness Coverage	Higher	**Lower (Risky)**

4.4.1. Interpretation

Policy 2 provides better business outcomes and more targeted uplift but may exclude valuable users. Policy 1 is more equitable in access but less efficient in ROI per treated user. These tradeoffs inform how businesses must weigh profit optimization against inclusive access to benefits.

5. Discussion

Our experiments demonstrate that the design of personalized pricing policies, particularly the behavioral signals used to determine eligibility, can significantly influence both business outcomes and perceived fairness. Policy 2, which used a stricter targeting heuristic (BounceRate + ExitRate), consistently outperformed Policy 1 across multiple dimensions: it achieved higher uplift (Qini AUC of 0.3817 vs. 0.3784), increased revenue per treated user, and higher average uplift among treated sessions. Interestingly, despite this improved targeting precision, Policy 2 treated a smaller proportion of top uplift users (0.77%) than Policy 1 (3.65%). This points to a key tradeoff between coverage and precision. Broader policies may reach more users, including some who might benefit marginally from incentives, but they risk diluting effectiveness and driving up costs.

Conversely, more selective strategies may leave value on the table by under-treating potentially persuadable users. From a modeling perspective, the comparison between TwoModel and Class Transformation uplift estimators confirms the robustness of our findings. While the Class Transformation model yielded a slightly lower Qini AUC (0.3671), it still ranked Policy 2 as superior. This cross-model consistency suggests that our insights are not model-dependent and that behavioral pricing can be studied reliably using standard uplift frameworks. Ethically, this study contributes to the emerging conversation on behavioral fairness in machine learning for personalization. Whereas most fairness frameworks focus on group-level equity (e.g., by gender or race), personalized pricing raises new challenges. Should we offer discounts only to those who are least likely to buy, or should loyalty, intent, and prior behavior also factor into pricing decisions? Our analysis shows that systems optimized for conversion may implicitly penalize loyal users by withholding incentives, creating algorithmic asymmetries not based on identity but behavior. By bridging uplift modeling, pricing, and fairness evaluation, this study provides a foundation for future work on fair incentive allocation where both business performance and equitable access to benefits are optimized together.

6. Limitations

While this study provides practical insights into personalized pricing strategies using uplift modeling, it is important to acknowledge certain limitations. First, our analysis is based on a publicly available clickstream dataset that lacks granular pricing, loyalty, and demographic variables. As a result, we simulate treatment effects using behavioral proxies like bounce and exit rates, which may not fully capture price sensitivity or willingness to pay. In contrast, studies such as Ban and Keskin [1] utilize detailed transaction-level data, enabling real-world causal evaluation of personalized pricing strategies. Second, our uplift modeling framework estimates the incremental effect of treatment but does not incorporate uncertainty-aware or Bayesian optimization methods that could better capture pricing volatility or estimation risk.

This stands in contrast to prescriptive frameworks like that of Lo and Pachamanova [6], who explicitly account for estimation errors in treatment selection. Third, we do not include multi-treatment uplift or multi-objective optimization. Our modeling focuses on binary discount treatment (treated vs. not treated), whereas more advanced personalization strategies might assign varying price levels or use reinforcement learning–based optimization, as suggested in broader dynamic pricing literature. Fourth, while we analyze fairness in treatment allocation (e.g., % of high-uplift users treated), we do not evaluate post-treatment user outcomes or downstream impact on customer loyalty or trust, areas explored in greater depth in works such as Priester et al. [7] and Kallus and Zhou [5]. Lastly, generalizability is limited. Our findings are based on a single e-commerce dataset, and results may vary significantly across industries or user contexts, particularly in regulated sectors like finance or healthcare, where pricing personalization is subject to legal constraints, as discussed by Zuiderveen Borgesius [9].

7. Conclusion

This study evaluated personalized pricing strategies using uplift modeling techniques applied to session-level e-commerce data. We simulated two treatment policies based on bounce and exit rates and compared their performance across multiple machine-learning models. Our results show that stricter behavioral filters, despite treating fewer users, can improve uplift accuracy, increase per-session revenue, and enhance targeting efficiency. However, they also introduce fairness concerns by excluding users with a high potential to respond. Modeling comparisons further showed that uplift insights are stable across architectures, with TwoModel performing slightly better than Class Transformation in our setting. These findings suggest that effective personalized pricing requires not only strong models but also responsible policy design. Fairness-aware uplift modeling, constrained optimization, and long-term customer equity metrics are promising directions for developing pricing systems that maximize both profit and trust.

References

[1] Gah-Yi Ban, and N. Bora Keskin, "Personalized Dynamic Pricing with Machine Learning: High-Dimensional Features and Heterogeneous Elasticity," *Management Science*, vol. 67, no. 9, pp. 5549-5568, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Daniel Baier, and Björn Stöcker, "Profit Uplift Modeling for Direct Marketing Campaigns: Approaches and Applications for Online Shops," *Journal of Business Economics*, vol. 92, pp. 645-673, 2022. [CrossRef] [Google Scholar] [Publ isher Link]

[3] Raouya El Youbi, Fayçal Messaoudi, and Manal Loukili, "Machine Learning-Driven Dynamic Pricing Strategies in E-Commerce," *14th International Conference on Information and Communication Systems*, Irbid, Jordan, pp. 1-5, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Pierre Gutierrez, and Jean-Yves Gérardy, "Causal Inference and Uplift Modelling: A Review of the Literature," *International Conference on Predictive Applications and APIs*, vol. 67, pp. 1-13, 2017. [Google Scholar] [Publisher Link]

[5] Nathan Kallus, and Angela Zhou, "Fairness, Welfare, and Equity in Personalized Pricing," *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and* Transparency, Virtual Event Canada, pp. 296-314, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Victor S.Y. Lo, and Dessislava A. Pachamanova, "From Predictive Uplift Modeling to Prescriptive Uplift Analytics: A Practical Approach to Treatment Optimization While Accounting for Estimation Risk," *Journal of Marketing Analytics*, vol. 3, pp. 79-95, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[7] Anna Priester, Thomas Robbert, and Stefan Roth, "A Special Price Just for You: Effects of Personalized Dynamic Pricing on Consumer Fairness Perceptions," *Journal of Revenue and Pricing Management*, vol. 19, pp. 99-112, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] P. Skolov, Scikit-Uplift: A Python Package for Uplift Modeling, 2021. [Online]. Available: https://github.com/maks-sh/scikit-uplift

[9] F.J. Zuiderveen Borgesius, "Discrimination, Artificial Intelligence, and Algorithmic Decision-Making," *European Business Law Review*, vol. 31, no. 6, pp. 821-838, 2020. [Google Scholar] [Publisher Link]

[10] Brent Daniel Mittelstadt, and Luciano Floridi, *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*, The Ethics of Biomedical Big Data, Springer Cham, pp. 445-480, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[11] Piotr Rzepakowski, and Szymon Jaroszewicz, "Decision Trees for Uplift Modeling With Single and Multiple Treatments," *Knowledge and Information Systems*, vol. 32, pp. 303-327, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[12] Pritom Das et al., "Optimizing Real-Time Dynamic Pricing Strategies in Retail and E-Commerce Using Machine Learning Models," *The American Journal of Engineering and Technology*, vol. 6, no. 12, pp. 163-177, 2024. [CrossRef] [Google Scholar] [Publisher Link]