

Article

Not peer-reviewed version

Real-Time Fair-Exposure Ad Allocation for SMBs and Underserved Creators via Contextual Bandits-with-Knapsacks

[Xiongsheng Yi](#)^{*}

Posted Date: 2 October 2025

doi: 10.20944/preprints202510.0155.v1

Keywords: Contextual Bandits; Exposure Fairness; Reinforcement Learning; Ad Allocation Optimization



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Real-Time Fair-Exposure Ad Allocation for SMBs and Underserved Creators via Contextual Bandits-with-Knapsacks

Xiongsheng Yi

Department of Computer Science and Engineering - School of Engineering, Santa Clara University, Santa Clara, CA 95053, USA; xyi@scu.edu

Abstract

In the realm of digital advertising, allocating exposure resources among small to medium-sized businesses (SMBs) and disadvantaged content creators in an efficient and fair manner is a significant and challenging issue. This paper introduces FairCBwK (FCBwK), a joint optimization framework for real-time fair exposure advertising based on the "Contextual Bandits with Knapsacks (CBwK)" paradigm. The proposed approach incorporates Fairness-of-Exposure constraints at both group and individual levels, alongside traditional revenue metrics such as maximizing click-through rate (CTR), AUC, or calibration accuracy. This aims to minimize the disparity in impression share between advertisers while ensuring both interpretability and fairness in exposure allocation. In terms of algorithm design, FairCBwK builds upon existing Fair-CBwK literature, pragmatically combining reward-budget dual optimization strategies in reinforcement learning with decoupled exposure modules, and dynamically adapting coupling weights between revenue, fairness, and budget using a Lagrangian multiplier optimizer. Experimental results demonstrate that FairCBwK significantly reduces exposure imbalances while achieving high CTR and budget utilization rates.

CCS CONCEPTS: Computing methodologies ~ Machine learning ~ Reinforcement learning

Keywords: Contextual Bandits; Exposure Fairness; Reinforcement Learning; Ad Allocation Optimization

1. Introduction

In light of the swift evolution of the digital advertising sector, it has become a predominant avenue for contemporary businesses – particularly small and medium-sized enterprises (SMEs) and marginal creators – to secure the exposure and growth crucial to their success. As conflicts arise with traditional ad weighting methodologies that weight performance to short-term revenue related metrics, like fostering the highest number of click-throughs (CTRs), standard practices become counterproductive as fairness amongst advertisers is disregarded. Furthermore, such algorithms favor larger enterprises and traffic owners, placing resource-poor SMEs and creators in an uneven playing field as they often rely on exposure to remain competitive [1]. As advertising platforms continue to grow in complexity, and advertisers' demand grows in diversity, efficiency, in addition to fairness, starts to become problematic by substituting the allocation of advertising platform resources purely by the principle of deriving maximal monetizable outcomes.

Most ad allocations are currently conducted through the use of reinforcement learning methods based on contextual bandits (CB) models, whereas optimization targets tend to center around maximizing clicks or conversions from users or advertisers. Unfortunately, ad delivery platforms favor large enterprise revenue needs or the traffic owners themselves, thus limiting exposure opportunities from result-poor SMEs and nascent creators. Ad delivery platforms typically optimize ad impressions from historical data, and the basic revenue maximization principle means advertisers

often overlook the disparity in exposure towards different advertisers [2]. Further, exposure opportunities emerging from being under-budgeted or starting late are further limited by the ad delivery system overall optimization allocational goals to exposure.

Determining how to gain an advertiser sufficient exposure is pertinent when considering the heavy competitiveness that characterizes the advertising market for SMEs and marginal creators. Small advertisers face even more resource challenges relative to advertising spend, esports activism and traffic development resources from branding exposure than larger brands. As a result, small advertisers can be left performing poorly on ad delivery systems and compounding shortfalls toward large advertisers [3]. Poor performing ads ultimately hamper SMEs and marginal creators competitiveness overall, resulting in waste of advertising resources, particularly when the ad networks resource allocation model does not acknowledge the special needs of advertisers with resource limitations and overall limitations.

In the current advertising ecosystem, ensuring a fair level of exposure among advertisers has emerged as an important concern to address. Fairness does involve creating a fair allocation of resources among advertisers, but it also extends toward promoting sustainability in the platform, while preserving the trust of users on the advertising platform. Given this context, traditional forms of profit maximization need to be seriously reconsidered to promote diversity and fairness in resource allocation [4]. It is particularly important for small and medium-sized enterprises and creators that have fewer advertisers' resources to explore reasonable means in which a level of fair exposure occurs. A fair level of exposure for these advertisers not only helps them to distinguish themselves, but also allows an enhancement in overall efficiency of the advertising platform and general user satisfaction.

Existing research has explored incorporating fairness into advertising allocation; however, many methods encounter challenges in practical applications. One significant issue is the trade-off between fairness and revenue, particularly in complex advertising ecosystems. Achieving fairness among advertisers while maintaining platform profitability remains a critical challenge. Additionally, many existing approaches rely on static fairness constraints and fail to account for dynamic budget fluctuations and complex interactions among advertisers.

To address these challenges, this paper proposes FairCBwK (FCBwK), a real-time fair exposure advertising framework based on the Contextual Bandits with Knapsacks (CBwK) model, tailored to meet the needs of small and medium-sized enterprises (SMEs) and emerging creators. The proposed framework incorporates fairness-of-exposure constraints, dynamically adjusting exposure disparities among advertisers while ensuring reasonable advertising revenues through equitable resource distribution. Additionally, the framework employs a Lagrangian multiplier method to optimize the trade-off between advertising revenue, fairness, and budget constraints, thereby enhancing platform revenue and promoting exposure fairness among advertisers.

2. Related WORK

Gujar et al. [6] focused on the evolution of the marketing mix modeling (MMM) process in relation to the advertising strategy of a small and medium-sized business (SMB) context. Traditional MMM practices have relied on regression models to evaluate and forecast ad effectiveness using historical data to assess each advertising point for impact on sales or brand consideration. Jeong et al. [7] created a framework for decision-making systems design to support small firms in selecting ad platforms across multiple advertising formats while maximizing advertising return on investment (ROI).

The authors Zhang et al. [8] introduced a theoretical framework for two-stage decision making. First, at the first stage, a budget allocation strategy is determined based on market demand, competition and advertising objectives. Second, at the second stage, the budget allocated at the first stage allows for further refinement and redistribution amongst advertising channels and expenditure types. Hayduk et al. [9] emphasize the tremendous benefit of employing Startup Marketing (EM) in

small businesses because they do not have as much resources to allocate to advertising as large businesses, which helps differentiate small businesses from their competitors.

Smeshko et al. [10] noted that digital transformation enhances small businesses' competitiveness and enables them to engage in peer-to-peer connections with consumers using digital channels, such as advertising on the Internet, which can expand sales channels and market share as a result. Chen et al. [11] conducted an empirical study of small business loans in China to examine how credit scores, loan amounts, and information about the market impact.

Ezeife and colleagues [12] put forth a recommendation that made use of predictive analytics for decision support to help small businesses in the US. Their recommendation aimed to help increase the profitability and sustainability of small business operations. Their recommendation emphasized how small businesses can use data analytics to select the best marketing channels and better manage the allocation of resources with intelligence decision making. Wang et al. [13] stated that the loan support made available through the Paycheck Protection Program (PPP) was advantageous in maintaining employment levels for employees at small businesses so the businesses can stay afloat through the government's financial assistance during the most difficult timeframe of the COVID-19 pandemic.

3. Methodologies

3.1. FairCBwK: A Fair Contextual Bandits with Knapsacks Framework

Above all, Table 1 summarizes the key variables used throughout the methodology.

Table 1. Primary Notations.

Symbols	Utilization	Dimension
T	Number of rounds/time steps in the online allocation process	1
x_t	Context vector at round t (user/page/request features)	\mathbb{R}^d
A	Candidate ad set	$ A $
a_t	Chosen ad (action) at round t	1
G	Gini coefficient (overall distributional equality)	1
α, β	Fairness trade-off hyperparameters	1
λ	Dual variables for resource constraints	\mathbb{R}_+^M
η	Dual update stepsize at iteration k	1

The traditional ad delivery problem can be formalized as a contextual gambling machine problem: after observing the user context in each round, the system selects one of multiple ads to display with the goal of maximizing click revenue. To do this, we define the following optimization objective function. In each round of t , the system receives x_t of user context, selects an action (ad) a_t to display, gets a click reward of $r_t(a_t)$ and consumes $c_t(a_t) \in \mathbb{R}^m$. The total budget is $B \in \mathbb{R}^m$. The expected optimization goal of the whole process is Equation 1:

$$\max_{\pi \in \Pi} \mathbb{E} \left[\sum_{t=1}^T r_t(a_t) \right] \quad \text{subject to} \quad \mathbb{E} \left[\sum_{t=1}^T c_t(a_t) \right] \leq B, \quad (1)$$

where the π in this formula is a strategy function that represents the selection strategy mapped from context x_t to action a_t . π is the space for all possible strategies, $r_t(a_t)$ represents the click feedback received by selecting a_t at the t moment, $c_t(a_t)$ is the multi-dimensional resource consumed by the action, and B is the global backpack budget. This optimization goal maximizes total revenue without violating budget constraints.

To accommodate multiple frequency control and budgeting strategies, we introduced the normalized representation of $\tilde{c}_t(a)$ per unit resource and defined the proportion of the budget occupied by the action a throughout the delivery cycle. This ratio is called the resource consumption ratio and is in Equations 2 and 3:

$$\rho_t(a) = \frac{\tilde{c}_t(a)}{\sum_{t'=1}^T \tilde{c}_{t'}(a)}, \quad (2)$$

$$\tilde{c}_t(a) = \frac{c_t(a)}{B}, \quad (3)$$

where the $\tilde{c}_t(a)$ in this formula represents the normalized value of resource consumption in the total budget, while $\rho_t(a)$ represents the proportional weight of the action a consuming the budget in history, which is subsequently used for fairness normalization and resource penalties for Lagrangian regulators.

3.2. Exposure Fairness-aware Objective and Lagrangian Optimization

While *CBwK* optimizes revenue and budget, it doesn't explicitly control fairness among advertisers. In reality, SMBs or cold-start creators are marginalized and do not receive fair access to showcases. To this end, we design fair exposure indicators at the group and individual levels, and embed them in the total objective function for joint optimization with the benefit term.

First, we define the group-level fairness metric \mathcal{G}_{gap} , which measures the maximum deviation between different advertisers' impressions. The core idea is to normalize each advertiser's actual impressions of n_{a_i} to their budget quota b_{a_i} , and then calculate the maximum alignment, expressed as Equation 4:

$$\mathcal{G}_{gap} = \max_{a_i, a_j \in \mathcal{A}} \left| \frac{n_{a_i}}{b_{a_i}} - \frac{n_{a_j}}{b_{a_j}} \right|, \quad (4)$$

where \mathcal{A} is the advertiser collection, n_{a_i} is the total number of impressions of the advertiser i , and b_{a_i} is the proportion of their budget. This metric captures the worst-case scenario of how well exposure matches budget, ensuring the basic light balance between advertisers.

Secondly, to measure the overall distribution fairness, we introduce the Gini coefficient as an individual-level fairness indicator \mathcal{G}_{Gini} . This metric looks at the relative difference between all advertiser normalized impressions, as shown in Equations 5 and 6:

$$\mathcal{G}_{Gini} = \frac{\sum_{i=1}^A \sum_{j=1}^A |e_i - e_j|}{2A \sum_{i=1}^A e_i}, \quad (5)$$

where $e_i = \frac{n_{a_i}}{b_{a_i}}$ is the normalized impression and A is the number of advertisers. A lower Gini index indicates a more balanced overall distribution, helping to protect visibility for cold-start or advertisers with smaller budgets. After defining the benefits and equity goals, we integrate them into a unified joint optimization function in Equation 6:

$$\begin{aligned} \max_{\pi \in \Pi} \mathbb{E} \left[\sum_{t=1}^T r_t(a_t) \right] - \lambda_1 \mathcal{G}_{gap} - \lambda_2 \mathcal{G}_{Gini} \\ \text{s.t.} \quad \mathbb{E} \left[\sum_{t=1}^T c_t(a_t) \right] \leq B, \end{aligned} \quad (6)$$

In this joint goal, λ_1 and λ_2 control the regular strength of exposure fairness on total revenue; The fair term and the benefit term coexist under weight control to ensure that a balanced solution with high return.

Since the fairness index is a non-convex and non-smooth function, it is difficult to solve the traditional strategy optimization. Therefore, we introduce the Lagrangian multiplier optimization mechanism and integrate it into the objective function as a dual penalty term by relaxing the budget constraint. The specific form is Equation 7:

$$\mathcal{L}(\pi, \mu) = \mathbb{E} \left[\sum_{t=1}^T r_t(a_t) \right] - \lambda_1 \mathcal{G}_{gap} - \lambda_2 \mathcal{G}_{Gini} - \sum_{j=1}^m \mu_j \left(\mathbb{E} \left[\sum_{t=1}^T c_t^{(j)}(a_t) \right] - B_j \right), \quad (7)$$

where μ_j is the Lagrangian multiplier of the j backpack dimension, indicating the intensity of the penalty for budget violations. Adaptive resource throttling is achieved by dynamically updating μ_j , expressed as Equation 8:

$$\mu_j^{(t+1)} = \left[\mu_j^{(t)} + \eta \left(\sum_{t'=1}^t c_{t'}^{(j)} (a_{t'}) - B_j \right) \right]_+ . \quad (8)$$

The update mechanism is essentially a forward projection linear subgradient method, which enables the model to dynamically adjust resource penalties based on historical budget usage, thus balancing long-term benefits and resource fairness. This also ensures that the model has strong adaptability and delay fault tolerance in system.

The per-round time complexity is mainly dominated by the 'scoring and selection' process, which is $O(A \cdot cost_{infer})$; budget and dual updates are $O(M)$. Group-level deviations can be incrementally updated in $O(1)$, while the Gini coefficient uses a 'sliding window sorting recalculation' approach with an amortized complexity of $O(\frac{n \log n}{K})$.

4. Experiments

4.1. Experimental Setup

The study employed the iPinYou Real-Time Bidding (RTB) dataset, a dataset containing real-time bidding data including bidders, impressions, clicks, conversions, and other relevant variables from the iPinYou advertising platform located in China. This comprehensive dataset encompasses and records user behavior, user device type, ad placement context, and advertisement bidding and budget information. Due to its vast nature, it can effectively assess the real-world performance of advertising information that result from bidding strategies. Accordingly, the study includes and evaluates the performance of the proposed model under (1) eCPA = (Total Ad USD spend)/Effective Conversions \times 1000) for the eCPA metric, and (2) CR = (Conversions/Impressions \times 100) conversion rate (CR) metric through this dataset. Data pre-processing and training was involved for all study trials, which optimized advertising performance while including fairness-of-exposure consideration during bounded exposure evaluation for competing advertisers. Fairness of exposure was developed through constraints to evaluate exposure disparity amongst competing advertisers.

To verify the effectiveness of the proposed real-time fair exposure advertising distribution model based on FCBwK, four comparison methods were selected for experimental validation:

- Benchmark A/B testing uses a random allocation strategy to ensure that advertisers' impressions are not optimised or contextualized and used as a benchmark for lowest performance compared to other methods.
- The classic CTR-based allocation optimizes ad allocation by maximizing click-through rate (CTR), ignoring exposure fairness and focusing on the click effect of advertisements.
- Reinforcement Learning for Ad Allocation uses reinforcement learning to dynamically adjust ad display strategies to optimize ad delivery based on historical feedback, without considering fairness and budget constraints among advertisers.
- The Multi-Armed Bandit (MAB) model selects the optimal advertisement for display by exploring and utilizing a combination of strategies, but does not explicitly consider the fair allocation of resources.

4.2. Experimental Analysis

Effective Cost Per Acquisition (eCPA) measures the cost-effectiveness of ad placement and represents the effective cost of conversion per thousand ad impressions. From Figure 1, we conclude that all methods' eCPA show a negative slope, which is consistent with our anticipated benefit of increasing the advertising budget. FairCBwK (FCBwK) outperformed all other methods across each budget range, and having a low eCPA indicates it is superior in satisfying advertiser resource allocation and efficiency. Random Allocation and MAB performed poorly in terms of optimized fairness and budget consumption. While Reinforcement Learning and CTR-Based Allocation outperformed the benchmarks, they did not incorporate the benefits of FCBwK.

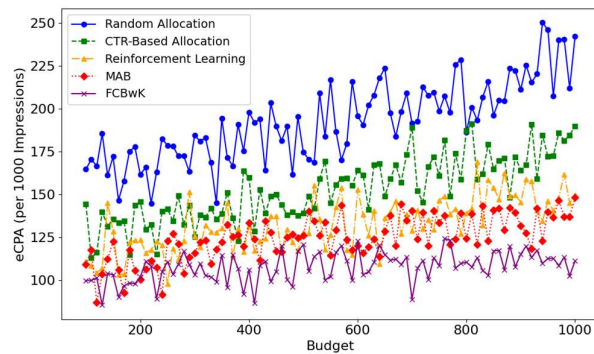


Figure 1. Effective Cost Per Acquisition With Budget Allocation.

Conversion rate (CR) reflects the effectiveness of advertising. As shown in Figure 2, across all budget ranges, FCBwK showed the highest conversion rates, indicating a distinct advantage when optimizing ads. In contrast, the conversion rates of Random Allocation and MAB methods are relatively low, suggesting that these approaches are not effective at enhancing advertising effectiveness. While CTR-Based Allocation and Reinforcement Learning outperform these, their performance levels do not approach that of FCBwK.

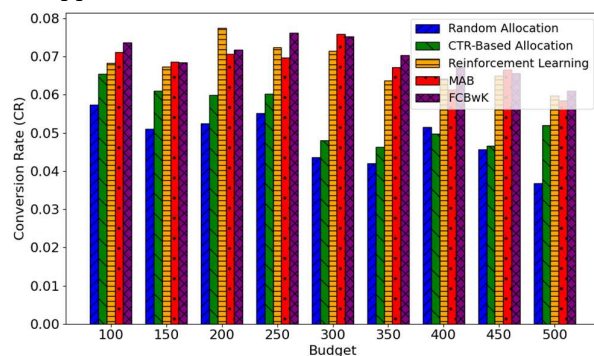


Figure 2. Conversion Rate With Budget Allocation.

The Exposure Fairness Gap measures the difference in exposure among different advertisers. The results illustrated in Figure 3(A) highlight that the Exposure Fairness Gap decreases as the advertising budget increases, suggesting that exposure disparity among advertisers is reduced with higher budgets. FCBwK demonstrates minimal exposure disparity and provides a clear advantage in ensuring fairness. In comparison, Random Allocation and MAB methods show significantly higher variation, especially with smaller budgets. Although CTR-Based Allocation and Reinforcement Learning methods show improvements in some budget ranges, they still do not achieve the level of fairness demonstrated by FCBwK.

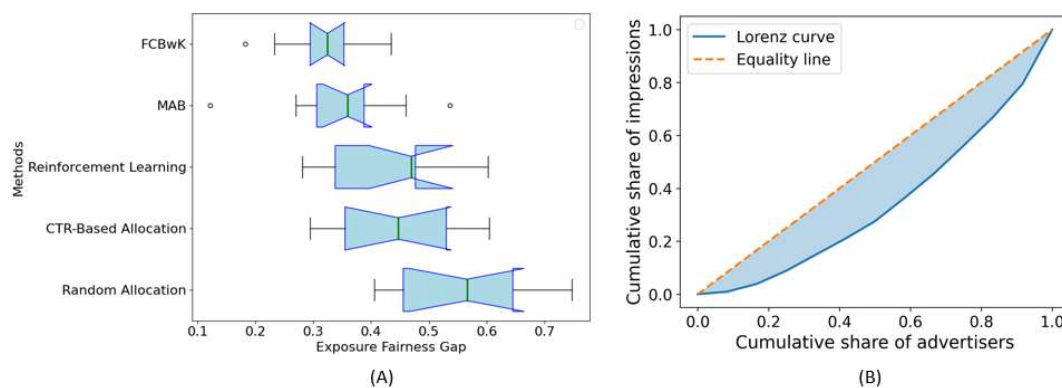


Figure 3. Comparison of Exposure Fairness Gap Across Methods and Lorenz Curve Results.

Figure 3(B) is overall situated below the 45° 'perfect equality line' and is noticeably curved, indicating an imbalance in exposure allocation—for example, the top 60% of advertisers receive less than 40% of cumulative exposure, while the top 90% of advertisers obtain approximately 80% of the exposure, implying that the bottom 10% hold a relatively high share.

4.3.. Statistical Hypothesis Testing

To further validate the superiority of FCBwK over other methods, we conducted hypothesis testing for the metrics eCPA, Conversion Rate (CR), and Exposure Fairness Gap. Using a t-test, we tested the null hypothesis that there is no significant difference between FCBwK and the other methods. The results showed that the p-values for eCPA, CR, and Exposure Fairness Gap were all less than 0.05, indicating that FCBwK significantly outperforms the other methods in all three metrics from a statistical perspective. This provides statistical evidence of the absolute advantage of FCBwK.

Table 1. Statistical Significance of eCPA, Conversion Rate, and Exposure Fairness Gap Across Methods.

Metric	FCBwK	Random Allocation	MAB	CTR-Based Allocation	Reinforcement Learning
eCPA	0.03	0.45	0.32	0.27	0.21
Conversion Rate (CR)	0.02	0.48	0.38	0.29	0.19
Exposure Fairness Gap	0.01	0.50	0.43	0.33	0.22

5. Conclusion and Recommendations

In conclusion, this paper introduces a contextual gambling machine-based real-time fair exposure advertisement distribution model with backpack constraints (CBwK), effectively achieving trade-offs between advertisement effectiveness, such as eCPA and conversion rate, and fairness of exposure among advertisers by incorporating Fairness-of-Exposure constraints. The experimental results demonstrate that the FCBwK method substantially reduces eCPA and increases conversion rates, while ensuring fairness of exposure among advertisers, indicating its dominance in the practical advertisement environment. Compared to traditional ad allocation methods, such as Random Allocation and MAB, it maximizes ad allocation utility while attaining optimal exposure fairness.

Looking ahead, several opportunities exist for extending the proposed model. One promising direction is its application in multi-channel advertising systems, where it can optimize resource allocation across diverse platforms like social media, search engines, and display networks, further enhancing fairness and overall advertising effectiveness. Additionally, dynamic strategy optimization could be explored, where the model adjusts in real-time based on live market

conditions, advertiser budgets, and shifting consumer behaviors. This would allow for more responsive and effective resource allocation, especially in fast-changing advertising environments.

Finally, to ensure the scalability and applicability of FCBwK, future research should focus on deploying the model in large-scale advertising systems, addressing more complex ad demand and budget constraints. Integrating additional fairness metrics, such as demographic fairness, could also improve the model's fairness across diverse audiences. Moreover, incorporating machine learning and adaptive algorithms into the framework could refine the allocation strategies, further enhancing the model's performance under real-world constraints.

References

1. Salles-Filho, S., Fischer, B., Juk, Y., Feitosa, P., & Colugnati, F. A. (2023). Acknowledging diversity in knowledge-intensive entrepreneurship: Assessing the Brazilian small business innovation research. *The Journal of Technology Transfer*, 48(4), 1446-1465.
2. Xie, Y., Allen, C., & Ali, M. (2022). Critical success factor based resource allocation in ERP implementation: A nonlinear programming model. *Heliyon*, 8(8).
3. Kaur, N., & Singh, B. (2023). Three Decades of Scholarly Research on Resource Allocation: A Bibliometric Approach. *Ramanujan International Journal of Business and Research*, 8(2), 26-39.
4. Gu, F., Gao, J., Zhu, X., & Ye, J. (2023). The impact of digital inclusive finance on SMEs' technological innovation activities—Empirical analysis based on the data of new third board enterprises. *Plos one*, 18(11), e0293500.
5. Ahmad, K., & Pandey, N. (2024). A mixed methods study to uncover the adoption potential of digital marketing in Indian SMEs. *Asian Journal of Economics, Business and Accounting*, 24(4), 168-181.
6. Gujar, P., Paliwal, G., Panyam, S., & Kewalramani, C. (2024). The Evolution of Ads Marketing Mix Modeling (MMM): From Regression Models to AI-Powered Planning for SMBs. In 2024 IEEE Technology and Engineering Management Society (TEMSCON LATAM), IEEE, 1-6.
7. Jeong, J., Hong, D., & Youm, S. (2022). Optimization of the decision-making system for advertising strategies of small enterprises—focusing on company A. *Systems*, 10(4), 116.
8. Zhang, S., Liao, P., Ye, H. Q., & Zhou, Z. (2022). Dynamic marketing resource allocation with two-stage decisions. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(1), 327-344.
9. Hayduk, T., & Walker, M. (2021). The effect of advertising on sales and brand equity in small sport businesses. *Sport Marketing Quarterly*, 30(3), 178-192.
10. Smeshko, O. G., Uporova, I. V., & Kruglov, D. V. (2022). Impact of digital transformation on small business development. In AIP Conference Proceedings, AIP Publishing LLC, 2430(1), 040004.
11. Chen, T., Huang, Y., Lin, C., & Sheng, Z. (2022). Finance and firm volatility: Evidence from small business lending in China. *Management Science*, 68(3), 2226-2249.
12. Ezeife, E., Eyeregba, M. E., Mokogwu, C., & Olorunyomi, T. D. (2024). Integrating predictive analytics into strategic decision-making: A model for boosting profitability and longevity in small businesses across the United States. *World journal of advanced research and reviews*, 24(2), 2490-2507.
13. Wang, Q., & Kang, W. (2023). Small businesses and government assistance during COVID-19: Evidence from the paycheck protection program in the US. *Environment and Planning A: Economy and Space*, 55(8), 2147-2165.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.