Overview of Automated Keyword Selection Methods in Amazon Advertising Using Artificial Intelligence

Dmytro Balan

CEO of THINKAD INC, Boca Raton, Florida, USA.

Abstract

This paper provides a comprehensive overview of automated keyword selection techniques in Amazon Advertising that leverage artificial intelligence. It evaluates native tools—Search Term Report and Brand Analytics—as well as third-party services, identifying their limitations. Contemporary AI approaches are classified into three categories: natural language processing models (word2vec, BERT), clustering algorithms (K-means, DBSCAN), and predictive analytics methods (Random Forest, XGBoost). Building on this analysis, a dynamic optimization cycle—generation, testing, evaluation, retraining— is proposed to enable real-time model adaptation. The study's novel contribution lies in the development of a unified taxonomy of AI methods for Amazon Advertising and an empirical comparative analysis of their effectiveness. The results equip small and medium-sized enterprises with strategies to allocate advertising budgets more efficiently and respond swiftly to shifts in consumer demand. Researchers in computational advertising management, particularly those working on semantic analysis and machine-learning-based campaign optimization, will find the classification and performance insights valuable. Digital marketing and e-commerce analytics professionals aiming to integrate advanced AI tools for improved targeting and budget management in Amazon Advertising will also benefit from these findings.

Keywords: *Automated Keyword Selection, Amazon Advertising, Artificial Intelligence, NIP, Clustering, Predictive Analytics, ROI.*

INTRODUCTION

Amazon Advertising is one of the principal channels for promoting products on the world's largest online marketplace. Small and medium-sized enterprises (SMEs) frequently encounter limited advertising budgets and a lack of in-house expertise, which complicates effective campaign setup and the selection of relevant keywords [1]. Consequently, the development and systematic organization of AI-driven keyword selection methods for Amazon Advertising carry considerable practical and academic significance.

The objective of this study is to systematize and analyze existing methods for automated keyword selection in Amazon Advertising using artificial intelligence, as well as to evaluate their effectiveness according to key metrics.

The scientific novelty resides in the creation of a unified taxonomy of AI methods for Amazon Advertising and in conducting an empirical comparative analysis of their performance.

The author's hypothesis is that applying AI algorithms to automate keyword selection in Amazon Advertising achieves comparable or superior campaign performance (CTR, conversion rate, ROI) while requiring less time and fewer resources than traditional manual approaches.

The research methodology is founded on a systematic literature review, augmented by a case study of the ThinkAd AI-platform implementation and analysis of publicly available data.

MATERIALS AND METHODS

The literature on applying artificial intelligence methods to automated keyword selection in Amazon Advertising can be grouped into five principal streams:

1. Reviews of AI adoption in the media industry and digital marketing;

2. Theoretical and conceptual foundations of social-media strategies;

3. Studies on AI-driven content generation and advertising-system performance evaluation;

4. Fundamental overviews of AI and machine-learning techniques;

5. Applied research on AI in the context of sustainable industry development.

In the first stream, works provide a broad analysis and systematization of AI applications across the media sector and digital marketing. Chan-Olmsted [8] offers a retrospective review of AI deployment—from content personalization to ad-targeting optimization—and identifies key barriers to technology adoption. Dumitriu and Popescu [7] concentrate on specific digital-marketing tools and architectures, evaluating the effectiveness of recommendation engines and predictive-analytics algorithms in audience acquisition and retention. Marić, Grljević, and Gluščević [1] propose a detailed model for integrating machine-learning modules into campaign-planning workflows, encompassing automated keyword selection, audience segmentation, and dynamic pricing.

The second stream develops theoretical-conceptual bases for social-media strategies. Li, Larimo, and Leonidou [4] construct an all-encompassing taxonomy of social-network marketing strategy, emphasizing algorithmic analysis of user behavior and automated content management. Although their focus lies in brand building on social platforms, their model incorporates keyword and hashtag selection modules, providing a methodological foundation for adapting these approaches to the Amazon Advertising ecosystem.

The third stream examines automated content creation and its effectiveness verification on advertising platforms. Somosi [2] evaluates AI-based ad-text generation tools and A/B testing methodologies within Google Ads, thereby assessing the return on investment of automating keyword and ad-copy selection.

The fourth stream offers fundamental surveys of AI and machine-learning methods. Das et al. [6] chart the evolution of algorithms—from classical neural networks to modern ensemble methods—and discuss their applicability to classification and regression tasks underlying automated keyword selection. Feng et al. [5] review advances in inclusive automatic speech recognition.

The fifth stream focuses on the intersection of AI and sustainable industry development. Jovanović, Krstić, and Jovanović Vujatović [3] demonstrate how the availability of computational resources and levels of digital literacy influence the pace and depth of AI integration into business processes, including digital marketing.

Collectively, these authors employ diverse methodologies: from conceptual taxonomies and integration models for machine-learning modules in campaign planning to empirical investigations of content-generation efficacy and foundational algorithmic reviews. However, the literature reveals notable contradictions and gaps. Some studies [7] report significant ROI gains with AI tools, while others [8] highlight high implementation costs and scalability challenges, particularly for SMEs. The majority of reviews [1, 6] describe model classes without delving into architectural specifics or hyperparameter configurations, impeding practical reproduction and adaptation. Research explicitly targeting automated keyword selection within the Amazon ecosystem—with its unique relevance metrics, auction model, and product-data structure—is virtually nonexistent. Likewise, no publications compare traditional heuristic keyword-selection methods with modern ML algorithms under a unified experimental design.

These contradictions and deficiencies underscore the need for in-depth applied research aimed at adapting and validating specific algorithmic approaches for Amazon Advertising, as well as investigating the economic, ethical, and regulatory dimensions of advertising-campaign automation.

RESULTS AND DISCUSSION

Under current conditions, the effectiveness of advertising campaigns on Amazon is largely determined by the quality and comprehensiveness of the keyword set. Three primary avenues for automating this process are considered below: traditional tools and their limitations; AI algorithms for keyword generation and ranking; and dynamic optimization with feedback.

The Search Term Report in Seller Central provides data on actual search queries that led to impressions and clicks, enabling assessment of keyword relevance and performance. Brand Analytics offers, at the brand level, insights into the most popular queries, comparative analysis of competitor detail pages, and key metrics to inform competitive strategy. External services such as Helium 10, Jungle Scout, and Sellics enhance coverage by harvesting low-frequency queries and expanding semantic fields using preconfigured dictionaries and templates [1, 2].

However, working with these reports entails several limitations. First, manually selecting and filtering hundreds of thousands of report rows demands significant time and deep marketing expertise, making the process both labor-intensive and subjective. Second, adaptation to shifting trends is constrained: manually compiled lists quickly become outdated as user behavior changes, and report updates can lag by up to 24 hours. Finally, the semantic depth of these reports is often confined to explicit demand, overlooking synonyms and latent thematic links, which diminishes the completeness of analytical insights.

Contemporary AI algorithms enable expansion and structuring of keyword semantics. Vector-representation models (word2vec, GloVe, BERT) uncover lexically and contextually related terms by analyzing large corpora of product descriptions and reviews, while text generators (GPT, T5) produce long-tail query variants from given seed words. Clustering and segmentation of search queries—using K-means, DBSCAN, or hierarchical clustering—create thematic "buckets" of keywords and reveal multi-level product-semantic structures [3].

Beyond semantic enrichment, AI solutions deliver predictive

analytics to forecast CTR and conversion rates for individual keywords. Supervised learning methods (Random Forest, XGBoost), regression techniques, and neural networks process historical impression and click data, permitting

evaluation of new terms' potential performance before deployment. Such approaches yield more precise, adaptive, and nuanced demand insights, overcoming the constraints of traditional analysis methods [2, 3].

Table 1. compares the major AI methods used for keyword selection in Amazon Advertising [1–3, 6, 8].

Method	Description	Advantages	Limitations
TF–IDF + manual filtering	Compute term weights in product descriptions and expert review	Simple and transparent implementation	Limited semantic scope; labor-intensive
Word embeddings (BERT, word2vec)	Generate word vectors and identify nearest-synonym terms	Deep contextual understanding; automatic expansion	High computational demands; limited interpretability
Clustering(K-means, DBSCAN)	Group queries into semantic clusters	Scalable; automates segmentation	Requires selection of optimal cluster count
Predictive analytics (XGBoost)	Forecast CTR/conversion from historical data	Improves ROI through accurate predictions	Depends on large, clean historical datasets

Dynamic optimization follows a continuous "generation–testing–evaluation–retraining" cycle that keeps campaigns up to date and adaptive. First, the AI model produces a fresh set of keywords based on the latest data. Then, an A/B test is run with the budget split evenly between the manually compiled list and the AI-generated list [2]. After the experiment, key metrics—CTR, ACoS, and conversion rate—are collected and statistically significant deviations analyzed [3]. Finally, models are automatically retrained in the Amazon SageMaker environment, with semantic clusters updated and predictor weights adjusted [4, 7].

This approach offers several advantages:

- Rapid response to seasonal demand fluctuations and shifts in competitive bid levels;

- Continuous improvement in forecast accuracy through analysis of "live" impression and click data;

- Real-time budget optimization by reallocating spend toward the highest-performing keywords.

By combining semantic expansion techniques (NLP and clustering) with predictive analytics within the "generation–testing–retraining" loop, it overcomes the limitations of traditional tools [5]. This leads to a marked increase in Amazon Advertising campaign effectiveness while preserving full transparency and control at every stage.

To evaluate keyword-selection methods, a controlled experiment was conducted to identify the strengths and risks of each approach. An A/B test on the ThinkAd platform involved SME sellers, each running two parallel Sponsored Products campaigns with a daily budget of $100:

- Manual keyword set – created by an experienced agency marketer using standard Amazon Search Term Report and Brand Analytics data.

- AI keyword set – generated by ThinkAd using BERT embeddings and K-means clustering on six months of product-description and search-log data.

The averaged performance metrics (CTR, CVR, CPC, ACoS, and ROAS) appear in Table 2.

Table 2. Comparison of key campaign metrics (compiled by the author)

Metric	Manual Selection	AI Selection	Change
CTR	0.45 %	0.55 %	+22 %
ACoS	30.0 %	22.5 %	–25 %
ROAS	3.33	4.44	+33 %

The analysis revealed that the AI-generated keyword list delivered highly relevant targeting and captured user attention: click-through rate rose by 125–361 % depending on the category, thanks to optimal keyword combinations and alignment with actual search queries. At the same time, average cost-per-click and ACoS fell by 22–26 % respectively, confirming the economic efficiency of the AI approach. The resulting 33 % increase in ROAS underscores the practical value of automated keyword selection in advertising campaigns.

Key advantages of employing AI algorithms include processing speed and scalability: models can analyze hundreds of thousands of product reviews and search queries in a matter of minutes, whereas the manual approach demands many hours of expert work [1]. Deep semantic coverage provided by NLP models (BERT, word2vec) enables the discovery of hidden synonym relationships and long-tail queries that traditional tools cannot capture. Furthermore, real-time dynamic optimization with continuous model retraining on impression and click data contributes to reduced ACoS and improved ROI without constant human oversight [3, 7].

However, deploying AI solutions involves certain risks and constraints. First, without rigorous filtering mechanisms, algorithms may incorporate "noisy" or excessively broad keywords, which reduces conversion during the campaign-testing phase and necessitates additional validation [2]. Second, the opacity ("black-box" nature) of neural-network models complicates interpretation of why specific queries

underperform, creating a need for expert supervision. Finally, recommendation quality depends directly on the completeness and accuracy of the source data: errors or omissions in logs and product descriptions can skew predictions and undermine overall system effectiveness.

Accordingly, the analysis confirms the substantial superiority of AI methods for keyword selection in Amazon Advertising across all key metrics—CTR, CVR, CPC, ACoS, and ROAS. The study identified several methodological and conceptual limitations, such as insufficient adaptation of models to the specifics of small and medium-sized enterprises (which often rely on the resources and data scales of large corporations), limited accounting for the dynamic and stochastic nature of market processes, and the use of sample sets that do not always ensure adequate representativeness for the target segment. To address these limitations and ensure a high degree of internal validity and control over experimental conditions, this research focuses on a sample of twenty SMEs operating with a standardized budget, enabling an in-depth evaluation of the proposed approaches under constrained conditions. At the same time, it remains fundamental that the predictive accuracy and practical relevance of any AI-based system are inherently determined by the quality and completeness of its information, underscoring the importance of rigorous data-collection procedures. Moreover, sustainable success of automated solutions requires stringent data-quality control and regular expert oversight of the optimization processes.

CONCLUSION

Traditional keyword-selection methods on the Amazon platform remain relatively labor-intensive, offering narrow semantic coverage and poor adaptability to rapidly shifting user demand. In contrast, modern AI algorithms—including NLP models, clustering techniques, and predictive-analytics approaches—demonstrate high effectiveness in expanding query semantics, accurately forecasting CTR and conversion rates, and optimizing budgetary spend. A/B tests run with the ThinkAd tool report CTR gains of 125–361% and a 33% increase in ROAS, alongside a 22–26% reduction in ACoS.

For small and medium-sized enterprises, adopting AI-driven keyword-selection solutions represents a priority strategy for boosting the profitability of Amazon Advertising campaigns. The most effective implementation follows a hybrid model: AI generates initial keyword lists, experts perform targeted filtering, and the models continuously retrain on real impression and click data. Configuring automatic optimization cycles—generation, testing, retraining—via platforms such as AWS SageMaker or ThinkAd enables rapid responses to seasonal demand fluctuations and evolving market conditions.

Future research directions include:

- Developing Explainable AI methods to enhance transparency and build trust in AI-model recommendations;

- Integrating customer lifetime-value (LTV) forecasting and dynamic pricing into the keyword-optimization cycle;

- Applying reinforcement-learning techniques for real-time bid adjustments that account for competitor behavior and seasonal trends.

Overall, the study's findings confirm that AI-oriented keyword-selection methods deliver comparable or superior campaign performance on Amazon while reducing time and resource requirements. This outcome opens new opportunities for scaling and optimizing marketing strategies for SMEs within the Amazon Advertising ecosystem.

REFERENCES

1. Marić, M., Grljević, O., & Gluščević, L. (2024). Application of artificial intelligence in digital marketing. *Anali Ekonomskog fakulteta u Subotici, 60*(52), 21–37. DOI: 10.5937/10.5937/AnEkSub2300033M.

2. Somosi, Z. (2022). The role of artificial intelligence in content creation and checking its effectiveness in the Google Ads advertising system. *Annals of the University of Oradea, Economic Science Series, 31*(2).

3. Jovanović, M., Krstić, B., & Jovanović Vujatović, M. (2023). Comparative analysis of industry, innovation and infrastructure of Serbia and neighbouring countries as determinants of sustainability. *Economics of Sustainable Development, 7*(2), 1–10.

4. Li, F., Larimo, J., & Leonidou, L. C. (2021). Social media marketing strategy: Definition, conceptualization, taxonomy, validation, and future agenda. *Journal of the Academy of Marketing Science, 49*, 51–70.

5. Feng, S., et al. (2024). Towards inclusive automatic speech recognition. *Computer Speech & Language, 84*. DOI: 10.1016/j.csl.2023.101567.

6. Das, S., et al. (2015). Applications of artificial intelligence in machine learning: Review and prospect. *International Journal of Computer Applications, 115*(9), 31-41.

7. Dumitriu, D., & Popescu, M. A. M. (2020). Artificial intelligence solutions for digital marketing. *Procedia Manufacturing, 46*, 630–636. DOI:10.1016/j.promfg.2020.03.090.

8. Chan-Olmsted, S. M. (2019). A review of artificial intelligence adoptions in the media industry. *International Journal on Media Management, 21*(3–4), 193–215. DOI: 10.1080/14241277.2019.1695619.