Physical Education, Health and Social Sciences

https://journal-of-social-education.org

E-ISSN: 2958-5996

P-ISSN: 2958-5988

A Comparative Analysis of ROAS-Based, Multi-Metric, and Rule-Based Advertising Decisions: Evidence from Keyword-Level Campaign Data

Suhail Ahmed[1]*

[1] Independent Researcher, M.Sc.-Digital Marketing from University of Bedfordshire
Corresponding Author's Email: suhail@suhailahmed.co, suhail@hearthyfoods.com

DOI: https://doi.org/10.63163/jpehss.v3i4.1779

Abstract

This study aims to compare the performance of ads based on ROAS, multi-metric, and rule-based decision-making methods by testing them against 47 observations at the keyword level of an e-commerce advertising dataset. The ROAS analysis found that very high ROAS scores were frequently measured over very low ad spending amounts, while high revenue, high sales volume, and more were observed for moderate ROAS scores with higher ad spending. Other performance metrics, such as CTR, CPC, CPM, sales volume, and revenue, provided additional context on the performance that ROAS didn't. The classification based on the rule-based procedure had 65.96% agreement with the classification based on ROAS. The multi-metric assessment was descriptive instead of categorical, so no multi-metric agreement rate is reported. The results indicate that both ROAS and multi-metric views revealed different images of the same keywords when it comes to understanding the performance of advertising efforts, revealing a point where a single metric can tell only part of the story. The study does not determine whether one evaluation approach produces better commercial outcomes. Instead, it demonstrates that different performance metrics provide complementary information about keyword-level advertising performance.

Keywords: *ROAS, Multi-Metric, advertising performance, Rule-Based, Decision-making, Digital Advertising, Keyword-level Analysis, Campaign Optimisation*

Introduction

With the growth of e-commerce and digital advertising, it is more crucial than ever to adopt a data-driven approach to campaign planning, targeting, optimisation and measurement (Rainy, 2025). Digital advertising increasingly relies on performance metrics to support targeting, optimisation and budget-allocation decisions. The recent literature on AI advertising can be divided into the four main areas of AI-driven advertising, computational advertising, programmatic advertising, and advertising effectiveness (Ford et al., 2023). Recent studies show, for instance, that advertising targeting using AI can yield better results for campaigns than other targeting methods that do not rely on algorithms (Shan, 2024). These developments highlight the importance of selecting appropriate performance measures to evaluate digital advertising activities. Return on Advertising Spend (ROAS) is one of the key metrics of advertising effectiveness and is calculated as a ratio of the value of the advertising revenue to the amount of money spent on advertising. Advertising performance can be assessed through complementary indicators including clicks, impressions, and advertising expenditure, revenue, sales volume, and ROAS depending on the objectives of the evaluation.

Problem Statement

Although ROAS measures the revenue generated relative to advertising expenditure, it does not provide a complete assessment of keyword-level advertising performance. Some observations can lead to high ROAS with low sales or ad spend, while others can yield high

sales and revenue but moderate ROAS. Therefore, examining additional indicators, including CTR, sales volume, CPC, CPM, and advertising spending, to better understand the advertising performance. This study addresses the need to compare ROAS-based evaluation, descriptive multi-metric assessment, and predefined rule-based decisions using empirical advertising data.

Research Gap

This study is motivated by three areas of research gaps. First, previous studies on AI in advertising have largely been on individual metrics like ROAS, CTR, conversion, and ROI, and there is little research on how multiple metrics can be leveraged together to inform advertising decisions (Ford et al., 2023). Second, there is relatively little research that directly compares evaluation based on a single metric to an integrated multi-metric assessment based on engagement, conversion, sales, and cost metrics. From a research perspective of AI-augmented advertising, performance metrics must be tailored to the specific advertising goal, instead of a single metric (Almestarihi et al., 2024). Third, there is a need for empirical examination of how predefined rule-based advertising recommendations correspond with ROAS-based classification using keyword-level advertising data. This study addresses these limitations by comparing recommendations based on ROAS, recommendations based on multiple metrics, and rules-based recommendations.

Significance of the Study

The study contributes to Digital Advertising research and practice by examining how different performance measures and predefined decision rules inform keyword-level advertising evaluation. It plays a role in advertising performance measurement by looking at ROAS and several other indicators instead of relying exclusively on one to measure advertising performance. It contributes to advertising performance research by demonstrating the complementary information provided by ROAS engagement, sales, and cost indicators. The study offers a useful roadmap for e-commerce practitioners to measure the effectiveness of their campaigns by finance, engagement, conversion, sales, and cost metrics. The results could also enable more tactical application of AI as a decision-making aid when managing data-based ads.

Aim of the study

This study descriptively compares ROAS-based, multi-metric, and rule-based advertising decision rules at the keyword-observation level. It does not test an AI model; it examines rules that could inform AI-assisted decision support.

Research Questions

RQ1: How does advertising performance evaluation differ between ROAS-based and multi-metric approaches at the keyword-observation level?

RQ2: To what extent do rule-based advertising recommendations align with ROAS-based and multi-metric campaign evaluation?

Research Objectives

1. To compare keyword-level advertising observations using ROAS-based and multi-metric performance indicators.
2. To evaluate the extent to which rule-based advertising recommendations align with ROAS- based classifications.

Literature Review

AI in Digital Advertising

Artificial Intelligence (AI) is now ubiquitous in digital advertising across a variety of areas, including targeting, campaign optimisation, media buying, performance prediction and automated decision making (Ziakis & Vlachopoulou, 2023). Machine-learning systems can examine huge volumes of advertising information and understand the patterns in audience

behaviour and forecast outcomes of advertising campaigns. AI can thus be used for tasks like audience targeting, bid optimisation, budget allocation, ad scheduling, and performance monitoring (Ford et al., 2023; Shan, 2024). Advertising systems can also be automated, so that they are able to continually analyse campaign signals and make a decision based on a set of objectives. Recent studies suggest that the use of AI in marketing can enhance the acquisition of customers and the performance of the firm, but the results are influenced by the context and the features of the firm (Mu & Zhang, 2025). With the advent of Generative AI, these features are further expanded by providing support in creating content, developing campaigns, and analysing marketing data (Kshetri et al., 2024; Grewal et al., 2025). Cillo and Rubera (2025) further map out the research opportunities for generative AI in innovation and marketing processes. However, the growing reliance on AI also raises concerns about ethical implementation, promoting calls for principles for advertising responsibility using generative AI (Sands et al., 2024). So, the use of AI is becoming more about actual automation, but also about using it as a tool to help marketers understand ad data and make suggestions. Its usefulness, however, is reliant on the information available, the goals set, and the degree of human judgment in the process of making advertising decisions.

ROAS as an advertising Performance Metric

Return on Advertising Spend (ROAS) is a commonly used advertising performance metric that measures the advertising revenue produced against the cost of advertising (Almestarihi et al., 2024). It could be defined as ROAS = Revenue/Advertising Spend. The primary benefit it offers marketers is its simplicity: they can easily understand if the money in the ads is getting revenue. ROAS also makes it easy to compare campaigns and keywords with a common metric of efficiency in ads. ROAS is beneficial in a data-driven advertising world to track campaign performance and make budget or optimisation choices (Ford et al., 2023). But there's a problem with ROAS: It offers only a partial perspective of campaign performance. It is important to note that this is a ratio of revenues to ad spend – it does not directly tell you about the quality of the customer engagement or the customer behaviour that follows the engagement. Ad campaigns can achieve the same ROAS with completely different CTR, conversion rates, sales volume, and/or ad spend(Vangelov, 2020). Similarly, a great ROAS isn't a sign of a successful campaign in terms of all goals. In the field of advertising analytics, one of the key points is that there is a need to link advertising performance metrics to the specific advertising goals, and not just one metric (Grewal et al., 2025). In this respect, ROAS is useful as an efficiency metric, but it might not be adequate for the complete evaluation of ads.

Multi-Metric Advertising Performance Evaluation

Multi-Metric Evaluation considers advertising performance across several complementary metrics, not just a single metric (Gogou, 2025). Click-through rate (CTR) is the ratio of people who click through to your site after seeing the ad, and it lets you know how the audience is reacting to your ad. Cost per click (CPC) is the cost incurred for every click, and cost per mille (CPM) is the cost of one thousand impressions. These indicators give an insight into engagement and ROI that ROAS doesn't. Other metrics such as hit rate or conversion rate, sales volume, revenue, or advertising expenditure can tell more about the outcomes of an ad campaign. For instance, a high CTR and low conversion rate could mean that engagement is not driving sales, or a high sales rate and high CPC or CPM could mean that the cost of advertising might be higher (Hu et al., 2010). In the context of AI advertising research, the concept of leveraging several performance indicators to measure advertising effectiveness is also emphasised (Shan, 2024). By aggregating the information on efficiency, engagement, conversion, sales, and costs, a multi-metric approach gives a wider picture of campaign performance. In this study, the ROAS is an important metric but is also combined with other metrics such as CTR, CPC, CPM, hit rate, sales volume, revenue, and advertising spend for a fuller picture of the metric.

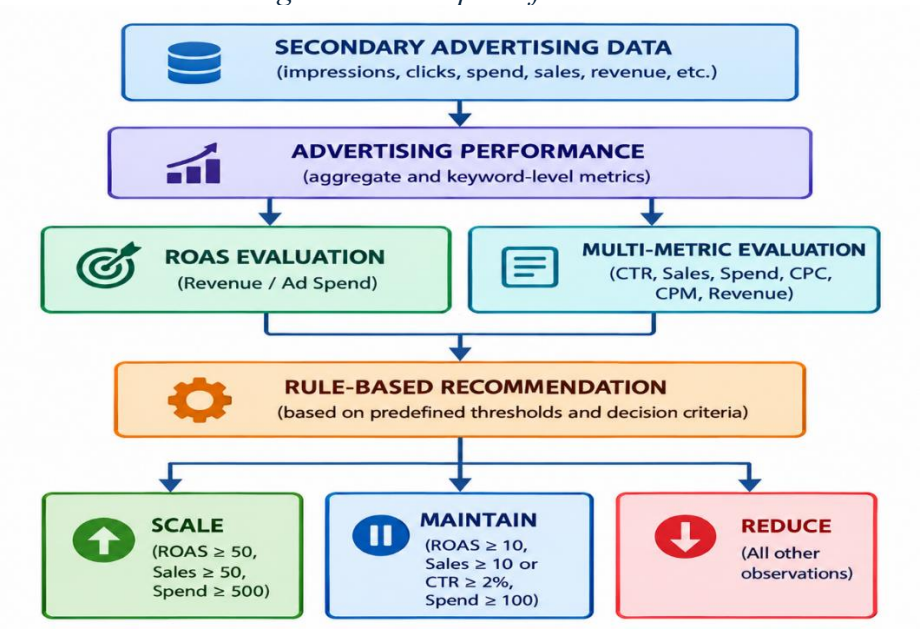
AI-Assisted Marketing Decision-Making

AI marketing decision-making refers to decision-making that relies on the use of AI systems to analyse marketing information, detect patterns, and make recommendations that aid managerial decisions (Ljepava, 2022). Such systems can make suggestions for targeting, campaign modifications, budget adjustments, as well as performance or campaign actions in advertising based on the available campaign data. AI's role in boosting marketing performance is revealed by research, which indicates that impacts vary based on organisational and contextual factors (Mu & Zhang, 2025). In the field of AI-driven advertising, too, studies show that results may differ based on targeting and the human-defined cut-off of ads (Häglund & Björklund, 2024). While AI recommendations offer valuable support, they are contingent upon the data, metrics, goals, and context fed into the system, making human oversight crucial. Human-AI collaboration in marketing has yielded mixed results across different tasks and the extent to which AI tools complement human skills (Ngo, 2025). In addition, human-AI collaboration also has the potential to impact managerial decision-making processes. So, AI recommendations are not necessarily right; they are suggestions to aid in making the right decision. Their value lies in the provision of adequate performance feedback and criteria that are consistent with the ad campaign objective.

Research Gaps and Conceptual Framework

The literature shows that AI is increasingly utilised to assist with advertising targeting, optimisation, prediction, and decision-making. Previous studies, however, tended to focus on a single performance measure or AI feature without directly contrasting ROAS-based evaluation with an integrated approach with multiple metrics. In this study, this gap is addressed by integrating ROAS and CTR, sales, advertising expenditure, CPM, and CPC to provide a more comprehensive assessment of advertising performance. The framework also considers the extent to which rule-based recommendations align with these other ways of measuring performance. Observations are evaluated using ROAS-based and multi-metric evaluation with secondary advertising data. The results of these assessments then lead to a rule-based recommendation for scaling, sustaining, or lowering observations. The framework therefore connects advertising performance measurement with rule-based decision support, and the fact that recommendations are based on the performance indicators used and the decision criteria.

Figure 1: Conceptual framework



Methodology

Research Design

In this study, the quantitative secondary data research design was used with a comparative decision analysis. This study was based on an existing advertising-performance dataset, and observations were compared using ROAS-based and multi-metric evaluation. The campaign information was then used in a rule-based decision procedure to generate advertising recommendations.No survey, interview, or primary data was used. Using structured numerical advertising data, the study was compared to ROAS-based evaluation, multi-metric evaluation, and rule-based recommendations.

Data Source

The ROAS dataset was publicly shared by Harun Gezici and is available via Mendeley Data. This dataset was published on 12 March 2025 (Gezici, 2025). The data were accessed for this analysis on 15 October 2025. The dataset contains targeted e-commerce advertising information, organised into 6 groups, each represented by a separate Excel file containing keyword-level advertising observations. Data points include keywords, advertising spend, impressions, clicks, hit rate, sales volume, revenue, ROAS, proposed CPM, actual CPM, and selected CPM. The data is licensed under the CC BY 4.0 terms (Gezici, 2025).

Dataset and Variables

The variables in the dataset include those related to advertising inputs, exposure, engagement, conversion, outcomes, efficiency, and cost. Advertising spend is expenditure, and keywords are the terms that are targeted at advertising. Impressions count the number of exposures; clicks count the number of engagements. Hit rate is calculated as clicks divided by impressions multiplied by 100, which is equivalent to CTR. Therefore, hit rate and CTR are treated as a single engagement measure rather than as separate conversion and engagement indicators. ROAS is a measure of the efficiency of ads. Proposed CPM, Actual CPM, and Selected CPM are recommended, realised, and selected CPM measures, respectively. The ROAS-based and multi-metric evaluations were based on these variables.

Table 1 Variables and operationalisation

Variable	Role	Measurement
Keyword	Unit identifier	Categorical
Advertising Spend	Advertising cost	USD
Impressions	Exposure	Count
Clicks	Engagement	Count
Hit Rate/CTR	Engagement/click-through	%
Sales Volume	Outcome	Count
Revenue	Financial outcome	USD
ROAS	Advertising efficiency	Ratio
Proposed CPM	Recommended cost	(USD)
Actual CPM	Realized cost	USD
Selected CPM	Selected cost	USD

Hit rate and CTR are treated as a single engagement measure because they are calculated identically (Clicks ÷ Impressions × 100). They are not separate conversion metrics.

Sales volume represents the total number of sales recorded for each keyword observation, as defined in the source dataset. The dataset documentation does not specify whether this measure represents individual units or completed orders.

Data Preparation

The six files in Excel provided in the dataset were downloaded and amalgamated into one analytical dataset. All the data were checked for missing data, duplicate observations, and differing variable formats. Appropriateness of analytical treatment of numerical and categorical

variables was checked. The available variables were used to derive advertising-performance measures. The final dataset was structured at the keyword-observation level, allowing for a consistent evaluation across ROAS-based evaluation, multi-metric evaluation, and rule-based recommendations.

Derived Measures

The variables were used to determine four advertising-performance measures. Click-through rate (CTR) was defined as "clicks/impressions x 100. Cost per click (CPC) was computed as advertising spend divided by clicks. CPM (Cost per Thousand Impressions) was determined by dividing the amount of money spent on the ads by the number of impressions times 1,000. ROAS was defined as revenue/ads spent. These measures complemented each other with respect to engagement, cost efficiency, exposure efficiency, and revenue generation. These derived measures were then used in the multi-metric evaluation and the rule-based decision-making process.

The formulas were:

$$CTR = (Clicks \div Impressions) \times 100$$

CPC = Advertising Spend / Clicks.

CPM = Advertising spend / Impressions x 1000

ROAS = Revenue / Spend on advertising.

Comparative Evaluation Procedure

There were two prespecified strategies for evaluating keyword-level observations.

First, a ROAS-based evaluation was performed, and each observation was assigned to one of three categories based on two pre-established ROAS thresholds: High ROAS (ROAS ≥ 50), Medium ROAS (10 ≤ ROAS < 50), and Low ROAS (ROAS < 10). These categories are aligned with the Scale, Maintain, and Reduce categories of the ROAS-based column of Table 2, and the agreement rate of 65.96% is calculated based on these categories. Secondly, a multi-metric descriptive analysis was carried out which included ROAS, CTR, sales volume, advertising spend, CPC, and CPM. Hit rate is not included since it repeats CTR. This assessment examined engagement, sales, and cost dimensions not captured by ROAS alone. Unlike the ROAS and rule-based approaches, the multi-metric assessment was descriptive and presented ROAS, CTR, sales volume, advertising spend, CPC, and CPM side by side without assigning decision categories.

Table 2: Decision Rules

Decision Category	ROAS-Based Classification	Rule-Based Procedure
Scale (High ROAS)	ROAS ≥ 50	ROAS ≥ 50 AND Sales ≥ 50 AND Spend ≥ 500
Maintain (Medium ROAS)	10 ≤ ROAS < 50	ROAS ≥ 10 AND (Sales ≥ 10 OR CTR ≥ 2%) AND Spend ≥ 100
Reduce (Low ROAS)	ROAS < 10	All remaining observations

Rule-Based Decision Procedure

The 47 keyword-level observations were subjected to a rule-based procedure previously defined. The metrics tracked for every observation were advertising spend, impressions, clicks, CTR, sales volume, revenue, ROAS, CPC, and CPM. Observations were then classified for each as Scale, Maintain, or Reduce, based on the rule-based thresholds listed in Table 2 (Section 3.6). The rules were applied as follows: Scale: ROAS ≥ 50, sales volume ≥ 50, advertising spend ≥ $500; Maintain: ROAS ≥ 10, sales volume ≥ 10 or CTR ≥ 2%, advertising spend ≥ $100; Reduce: others. Before comparing them with the ROAS-based and multi-metric classifications, the rule-based classifications were recorded. There was no evidence that a

person was more consistent by agreeing with the decision rules, only that it was a sign of agreement.

Data Analysis

Descriptive statistical results were summarised by calculating the mean, median, minimum, maximum and standard deviation of the variables of advertising performance. Decision-category frequencies were calculated for the ROAS-based and rule-based approaches. The multi-metric assessment was summarised descriptively because it did not assign observations to discrete decision categories. The recommendations based on rules were then compared to the classifications based on ROAS and multi-metric. Agreement rates were expressed as a percentage of the total number of observations (47), giving the number of classifications that were agreed on. The results are given in descriptive tables and comparative figures. Agreement was taken as a sign of agreement with the pre-defined decision rules and not that one method yields better real-world advertising outcomes.

Results

Overall Advertising Performance

There are no missing values in the master dataset; it has 47 observations of keywords by group and 13 different keywords. In aggregate, the observations represent $70,792.36 in advertising spend, $3,308,998.47 in revenue, and 26,619 sales. The aggregate ROAS was 46.74, with an overall CTR of 4.25%, CPC of $2.36, and CPM of $100.28. There were significantly different efficiencies and scales of group performance. Group 5 recorded the highest advertising spend ($43,933.77), revenue ($1,943,380.17), and sales volume (14,808), with an aggregate ROAS of 44.23. Group 3 had the highest aggregate ROAS (88.35), the lowest aggregate CPC ($1.31), and the highest CTR (7.12%), generating $1, 228,841.19, in revenue and 10833 sales from $13,908.04 in spend. The lowest aggregate ROAS was achieved with Group 6 (7.06). Group 2 did not spend as much on advertising ($860.64) and had an average ROAS of 23.15. Group 1 and Group 4 were 12.37 and 11.58, respectively. The results also showed that there was no correlation between the efficiency of the ads and the scale of the ad spend: Group 3, with the highest ROAS, did not have the highest spend or revenue, and the group that had the highest revenue and sales (Group 5) had a much lower ROAS.

Table 3: Dataset Characteristics

Characteristic	Value
Total observations	47
Number of groups	6
Number of keywords	13
Missing values	0
Range of advertising spend	$0.25–$17,665.35
Range of revenue	$105.61–$1,811,070.41
Range of ROAS	0.26–5,622.52

ROAS-Based Evaluation

There were considerable differences in ROAS for the 47 observations. The highest ROAS was recorded for Organic in Group 2 (5,622.52), followed by Baby butter in Group 3 (473.83), Biscuit in Group 1 (317.14), Baby semolina in Group 6 (231.84), and Baby curd in Group 3 (198.02). These were extremely high ROAS levels; however, with very low advertising spend, Organic in Group 2 had only $0.25 in spend and achieved 5,622.52, whereas Baby butter in Group 3 had $2.11 in spend and achieved 473.83. In contrast, Makarna lutfen in Group 3 achieved a ROAS of 117.30 from $10,331.77 in spend, generating $1,211,873.38 in revenue and 10,703 sales. Makarna lutfen in Group 5 generated $1,811,070.41 in revenue and 13,790 sales from $17,665.35 in spend, with a ROAS of 102.52. These observations included high ROAS as well as the ability to see significant scale and sales volume in ads. The descriptive

statistics show a mean advertising spend of $1,506.22 (median $183.78; range $0.25–$17,665.35), a mean sales volume of 566.36 (median 10; range 1–13,790), and a mean revenue of $70,404.22 (median $1,538.10; range $105.61–$1,811,070.41). ROAS had a mean of 167.62 (median 13.90; range 0.26–5,622.52). Mean CTR was 2.34% (median 1.13%), mean CPC was $6.51 (median $4.88), and mean CPM was $68.07 (median $61.13). The mean ROAS of 167.62 was substantially higher than the median of 13.90, reflecting the influence of extreme observations. Organic in Group 2 recorded a ROAS of 5,622.52 from only $0.25 in advertising spend. Therefore, the median provides additional context for interpreting the typical ROAS across the 47 observations.

Table 4: Descriptive Statistics

Variable	Mean	Median	Min	Max	SD
Advertising Spend	1,506.22	183.78	0.25	17,665.35	3,324.61
Impressions	15,019.49	5,280	110	114,206	24,672.34
Clicks	638.17	36	3	13,709	2,372.44
Sales Volume	566.36	10	1	13,790	2,412.58
Revenue	70,404.22	1,538.10	105.61	1,811,070.41	289,126.73
ROAS	167.62	13.90	0.26	5,622.52	809.50
CTR (%)	2.34	1.13	0.23	14.23	3.02
CPC ($)	6.51	4.88	0.08	37.19	7.54
CPM ($)	68.07	61.13	1.89	261.97	50.79

Table 5 ROAS- Based Classification

keyword	Group	ROAS	ROAS category	Decision
Makarna lutfen	1	14.57	Medium	Maintain
Baby biscuit	1	2.76	Low	Reduce
Biscuit	1	317.14	High	Scale
Organic	1	2.69	Low	Reduce
Pasta	1	0.26	Low	Reduce
Toddler food	1	44.59	Medium	Maintain
Baby semolina	1	27.93	Medium	Maintain
Baby tarhana	1	8.36	Low	Reduce
Gluten-free pasta	1	13.95	Medium	Maintain
Makarna lutfen	2	30.95	Medium	Maintain
Baby semolina	2	13.02	Medium	Maintain
Pasta	2	6.25	Low	Reduce
Baby tarhana	2	14.73	Medium	Maintain
Gluten-free pasta	2	16.07	Medium	Maintain
Pudding	2	194.79	High	Scale
Toddler food	2	6.69	Low	Reduce
Organic	2	5622.52	High	Scale
Pudding	3	1.03	Low	Reduce
Makarna lutfen	3	117.30	High	Scale
Organic	3	4.55	Low	Reduce
Pasta	3	0.73	Low	Reduce
Baby curd	3	198.02	High	Scale
Baby butter	3	473.83	High	Scale
Toddler food	3	8.04	Low	Reduce
Makarna lutfen	4	30.38	Medium	Maintain
Toddler food	4	4.84	Low	Reduce
Pasta	4	0.79	Low	Reduce

Organic	4	79.31	High	Scale
Biscuit	4	1.30	Low	Reduce
Baby biscuit	4	30.70	Medium	Maintain
Baby semolina	4	13.90	Medium	Maintain
Makarna lutfen	5	102.52	High	Scale
Pasta	5	7.15	Low	Reduce
Toddler food	5	8.02	Low	Reduce
Baby semolina	5	7.45	Low	Reduce
Baby tarhana	5	6.10	Low	Reduce
Organic	5	0.77	Low	Reduce
Baby butter	5	0.79	Low	Reduce
Rice flour	5	16.45	Medium	Maintain
Makarna lutfen	6	31.68	Medium	Maintain
Rice flour	6	4.68	Low	Reduce
Organic	6	41.24	Medium	Maintain
Pasta	6	71.76	High	Scale
Toddler food	6	4.94	Low	Reduce
Baby semolina	6	231.84	High	Scale
Baby butter	6	3.11	Low	Reduce
Baby curd	6	37.46	Medium	Maintain

Multi-Metric Advertising Performance

Examining the metrics together allowed a distinction between high ROAS generated from a meaningful amount of spend versus high ROAS generated from a very small amount of spend at the observation level. Pasta in Group 6 had a ROAS of 71.76, but received only six sales with $12.98 spent. By contrast, Toddler food in Group 6 had a ROAS of 4.94 but generated 139 sales and $16,264.60 in revenue from $3,292.64 in spend. Biscuit in Group 1 had an ROAS of 317.14 from only $16.56 in spend, with 32 sales and a CTR of 0.31%. The trends displayed here indicate that having a strong ROAS doesn't always mean having strong sales, engagement, or advertising scale. ROAS was used to measure revenue efficiency, and CTR, sales, CPC, and CPM measured engagement, output, and cost efficiencies.

Table 6 Multi-Metric Evaluation

Keyword	Group	ROAS	CTR (%)	Sales	CPC ($)	CPM ($)
Makarna lutfen	1	14.57	1.64	130	8.33	136.33
Baby biscuit	1	2.76	0.71	8	7.98	56.69
Biscuit	1	317.14	0.31	32	1.18	3.66
Organic	1	2.69	0.98	3	6.84	66.80
Pasta	1	0.26	0.28	1	22.34	63.05
Toddler food	1	44.59	4.46	13	1.01	45.06
Baby semolina	1	27.93	2.80	21	2.85	79.61
Baby tarhana	1	8.36	1.08	3	5.78	62.48
Gluten-free pasta	1	13.95	0.75	6	4.88	36.38
Makarna lutfen	2	30.95	1.11	92	5.34	59.09
Baby semolina	2	13.02	2.36	9	2.62	62.00
Pasta	2	6.25	0.36	5	12.40	44.36
Baby tarhana	2	14.73	4.34	6	1.41	61.13
Gluten-free pasta	2	16.07	2.34	5	1.00	23.51
Pudding	2	194.79	0.32	3	0.79	2.53
Toddler food	2	6.69	1.34	2	4.63	61.95
Organic	2	5622.52	2.27	10	0.08	1.89
Pudding	3	1.03	0.53	4	7.42	39.69
Makarna lutfen	3	117.30	10.70	10703	1.03	109.79
Organic	3	4.55	1.72	6	1.51	26.01

Pasta	3	0.73	0.43	10	29.25	124.45
Baby curd	3	198.02	3.72	13	1.10	40.74
Baby butter	3	473.83	8.18	4	0.23	19.18
Toddler food	3	8.04	1.68	93	5.75	96.60
Makarna lutfen	4	30.38	0.85	138	7.98	68.08
Toddler food	4	4.84	0.65	8	11.46	75.02
Pasta	4	0.79	0.42	1	18.41	76.69
Organic	4	79.31	2.00	3	1.71	34.15
Biscuit	4	1.30	0.23	8	37.19	85.00
Baby biscuit	4	30.70	1.13	4	4.62	52.20
Baby semolina	4	13.90	2.65	189	3.08	81.72
Makarna lutfen	5	102.52	12.00	13790	1.29	154.68
Pasta	5	7.15	1.12	371	11.75	131.52
Toddler food	5	8.02	2.71	111	5.11	138.72
Baby semolina	5	7.45	2.54	342	5.44	138.03
Baby tarhana	5	6.10	2.32	131	7.30	169.13
Organic	5	0.77	1.96	26	13.39	261.97
Baby butter	5	0.79	0.51	27	14.25	72.00
Rice flour	5	16.45	0.87	10	4.49	38.91
Makarna lutfen	6	31.68	0.76	54	5.67	43.23
Rice flour	6	4.68	0.56	5	5.11	28.36
Organic	6	41.24	3.39	8	1.36	46.16
Pasta	6	71.76	0.92	6	1.08	9.90
Toddler food	6	4.94	1.85	139	5.31	98.04
Baby semolina	6	231.84	14.23	20	0.27	38.85
Baby butter	6	3.11	0.73	28	2.84	20.70
Baby curd	6	37.46	1.07	18	1.24	13.28

Note: The multi-metric results are reported descriptively. No discrete multi-metric decision categories were defined, so no multi-metric agreement rate is presented.

Comparative interpretation of ROAS and Multi-Metric Results

The difference between ROAS and multi-metric evaluations was more extreme in the cases where ROAS was extremely high, but spend was extremely low. Organic in Group 2 had a ROAS of 5,622.52 on $0.25 in spend, whereas Makarna lutfen in Group 3 had a ROAS of 117.30 with $1,211,873.38 in revenue and 10,703 sales from $10,331.77 in spend. The observation with very low advertising spend recorded a substantially higher ROAS than the larger-scale observation because its advertising expenditure was only $0.25, resulting in an exceptionally high revenue-to-spend ratio. The highest ROAS at the group level was for Group 3 (88.35), and the highest revenue and sales volume was for Group 5. There was no efficiency / commercial scale relationship. This multi-metric view was clearly visible: It was possible to differentiate between high ROAS that came from a meaningful campaign versus high ROAS that came from minimal spend.

Results relevant to Rule-Based Decision-Making

The following rule-set was used to categorize each observation:

- Scale: ROAS ≥ 50 AND sales ≥ 50 AND spend ≥ 500
- Maintain: ROAS ≥ 10 AND (sales ≥ 10 OR CTR ≥ 2%) AND spend ≥ 100

Keep in mind that all other observations will be reduced.

There were 2 Scale, 7 Maintain, and 38 Reduce observations in the procedure.

These two recommendations from Scale were:

The two Scale recommendations were linked to Makarna lutfen. But in both of these observations, recorded sales volume was more than the corresponding click volume. For Group 3, the sales volume was 10,703 while there were approximately 10,031 clicks, and for Group 5, the sales volume was 13,790 while there were approximately 13,694 clicks. The differences

will be reported as per the source dataset and are subject to interpretation based on the definition of Sales Volume in the source dataset.

The seven recommendations that were made in Maintain were:

- Makarna lutfen in Groups 1, 2, 4, and 6
- Baby semolina (Groups 2 and 4)
- Rice flour, Group 5

Table 7 Rule-Based Recommendations by Group

Group	Scale	Maintain	Reduce	total
Group 1	0	1	8	9
Group 2	0	2	6	8
Group 3	1	0	6	7
Group 4	0	2	5	7
Group 5	1	1	6	8
Group 6	0	1	7	8
Total	2	7	38	47

Table 8 Rule-Based Recommendations (selected Observations)

Keywords	Groups	ROAS Decision	Rule-Based Recommendations
Makarna luften	1	Maintain	Maintain
Baby biscuit	1	Reduce	Reduce
Biscuit	1	Scale	Reduce
Organic	1	Reduce	Reduce
Pasta	1	Reduce	Reduce
Toddler food	1	Maintain	Reduce
Baby semolina	1	Maintain	Reduce
Baby tarhana	1	Reduce	Reduce
Gluten-free pasta	1	Maintain	Reduce
Makarna lutfen	2	Maintain	Maintain
Baby semolina	2	Maintain	Maintain
Pasta	2	Reduce	Reduce
Baby tarhana	2	Maintain	Reduce
Gluten-free pasta	2	Maintain	Reduce
Pudding	2	Scale	Reduce
Toddler food	2	Reduce	Reduce
Organic	2	Scale	Reduce
Pudding	3	Reduce	Reduce
Makarna lutfen	3	Scale	Scale
Organic	3	Reduce	Reduce
Pasta	3	Reduce	Reduce
Baby curd	3	Scale	Reduce
Baby butter	3	Scale	Reduce
Toddler food	3	Reduce	Reduce
Makarna luften	4	Maintain	Maintain
Toddler food	4	Reduce	Reduce
Pasta	4	Reduce	Reduce
Organic	4	Scale	Reduce
Biscuit	4	Reduce	Reduce
Baby biscuit	4	Maintain	Reduce
Baby semolina	4	Maintain	Maintain
Makarna luften	5	Scale	Scale
Pasta	5	Reduce	Reduce

Toddler food	5	Reduce	Reduce
Baby semolina	5	Reduce	Reduce
Baby tarhana	5	Reduce	Reduce
Organic	5	Reduce	Reduce
Baby butter	5	Reduce	Reduce
Rice flour	5	Maintain	Maintain
Makarna luften	6	Maintain	Maintain
Rice flour	6	Reduce	Reduce
Organic	6	Maintain	Reduce
Pasta	6	Scale	Reduce
Toddler food	6	Reduce	Reduce
Baby semolina	6	Scale	Reduce
Baby butter	6	Reduce	Reduce
Baby curd	6	Maintain	Reduce

Agreement Analysis

This classification process is also referred to as agreement between rule-based and ROAS-based classification.

Agreement Between Rule-Based and ROAS-Based Classification

In 31 of the 47 (65.96%) cases, the rule-based classification was identical to the ROAS-based classification. Agreement was highest in Group 5 (8/8, 100%) and lowest in Groups 2 and 6 (4/8, 50% each). Group-level agreement was:

Table 9 Rule-Based VS ROAS agreement rates by dataset group (n=47)

Group	Agreements	Observations	Agreement Rate
Group 1	5	9	55.6%
Group 2	4	8	50.0%
Group 3	5	7	71.4%
Group 4	5	7	71.4%
Group 5	8	8	100.0%
Group 6	4	8	50.0%
Total	31	47	65.96%

Most disagreements involved observations with very high ROAS and very low spend. In this instance, Biscuit, spending $16.56 in Group 1, had an ROAS of 317.14; and Organic in Group 2 had an ROAS of 5,622.52, spending $0.25. Both were deemed Scale by ROAS but Reduce by the rule-based procedure due to the spend and sales criteria. Likewise, Baby curd and Baby butter in Group 3 had ROAS over 100 and spend below $10 were considered Reduce. The interpretation of this agreement should be done with great care. The Scale and Maintain tiers are based on rules and reuse ROAS thresholds of 50 and 10, as well as spend filters. The 65.96% value is thus an indication of consistency of rules, rather than accuracy or independent validation.

Multi-Metric Agreement

A multi-metric agreement rate against the rule-based procedure is not reported because the multi-metric assessment was descriptive and did not produce a discrete decision category. Agreement can only be calculated between two categorical classifications; the multi-metric assessment reported metrics without assigning categories.

Overall Agreement Summary

Table 10 Agreement Summary

Comparsion	Agreements	Observations	Agreement Rate

Rule-based vs ROAS-based	31	47	65.96%
Rule-based vs multi-metric	Not applicable	—	—

This rule-based procedure was compared to the classification based on ROAS, and a level of agreement of 65.96% was obtained. There was no agreement rate calculated for the multi-metric assessment since it did not provide performance indicators side by side and did not give categorical decisions. The agreement outcome is therefore the agreement that is most consistent between the consistency of the rule-based procedure and the ROAS-based classification, and not the correctness or commercial success of the recommendations.

Discussion

RQ1: ROAS vs Multi -Metric Evaluation

The results of this study show that the ROAS and multi-metric approaches yield significantly different impressions of the advertising performance, especially in situations where the advertising expenditures are different. The results show that ROAS is useful for assessing revenue efficiency, but it does not independently capture advertising scale, engagement, or sales volume. The side-by-side assessment of additional indicators provides complementary information for interpreting keyword-level advertising performance. This supports the contention of Almestarihi et al. (2024) that measures need to be adjusted to specific advertising goals and cannot be based on a single measure. ROAS is useful for comparing campaigns, but it's not enough to understand the impact of advertising, as pointed out by Ford et al. (2023). The multi-metric analysis has shown that Group 3 had the highest aggregate ROAS (88.35), Group 5 had the highest revenue ($1,943,380.17), and Group 5 had the highest sales volume (14,808). This contrast bolsters the argument of Grewal et al. (2025) to make advertising performance metrics specific to a goal, rather than interchangeable. Vangelov (2020) also noted that the same ROAS can be obtained by having a completely different CTR, conversion rate, and number of sales. The current study demonstrates that ROAS and the additional performance indicators provide different but complementary information about keyword-level advertising performance.

RQ2: Rule-Based Recommendations

The rule-based procedure matched the classification based on ROAS in 31 out of 47 observations (65.96%). An agreement rate with the multi-metric assessment was not calculated because this assessment was descriptive and did not assign categorical decisions. The rule-based procedure included the same ROAS thresholds as the ROAS-based classification, as well as supplemental sales, CTR, and advertising spend thresholds. Therefore, the agreement observed is not evidence of improved alignment or decision quality, but an agreement that is consistent with the structure of the predefined rules. Nevertheless, the differences show a significant shortcoming. There were a few observations that were classified as Reduce in the rule-based procedure and Medium or High in the multi-metric assessment, which had high sales volume but low to moderate ROAS. In Group 6, for instance, Toddler food had 139 sales and revenue of $16,264.60, but its ROAS (4.94) and spend ($3,292.64) were below the Maintain thresholds, so the channel was recommended for Reduce. The findings illustrate how predefined thresholds influence rule-based advertising recommendations, particularly when observations have substantial sales volume but do not satisfy the ROAS requirement. The findings demonstrate that ROAS alone does not capture all dimensions of keyword-level advertising performance, particularly when observations generate different levels of sales and advertising expenditure.

Comparison with previous studies

This is consistent with Ziakis and Vlachopoulou (2023), who analysed Artificial Intelligence in digital marketing and found that there are several complementary metrics needed to measure performance. Another finding of the study is that Gogou (2025) is right when he recommended the use of marketing metrics together rather than separately. The results highlight the

importance of selecting appropriate performance indicators when evaluating digital advertising. The findings also demonstrate that the recommendations made by rules are subject to the predefined thresholds, as observations can be given different recommendations based on different predefinitions of ROAS, sales, CTR, and advertising-spend requirements. These results corroborate the overall message of this study that performance measures of ads should go beyond a single metric (Almestarihi et al., 2024; Ford et al., 2023).

Theory and Practical Interpretation

The most important theoretical finding of the present research is the multidimensional nature of advertising performance evaluation and the fact that the usefulness of any decision procedure will rely on the metrics provided as well as the thresholds and decision criteria used. An ROAS that's high with little spend isn't the same as an ROAS that's moderate with significant spend. For online advertisers, these findings have implications for budget decisions: ROAS alone should not drive budget cuts, as advertising observations with relatively moderate ROAS may still generate substantial revenue and sales volume.

Limitations and Future Research

Limitations

- **Secondary data:** The study uses a secondary data set that was used for commercial purposes and has less control over the data quality and variable definitions.
- **The dataset lacks some context:** it is not explained what was being done, who the target audience was, and what the market conditions were.
- **Limited generalisation to broader advertising strategies** – keyword-level analysis, as the unit of analysis is individual keywords, not complete campaigns.
- **No profit margin data:** ROAS is not a true measure of profitability because there are no costs or profit margins included in the calculation.
- The recommendations are based on rule-based thresholds; different thresholds may lead to different classification.
- **No causal inference:** The study is not able to demonstrate a cause-and-effect relationship between some advertising decision and some business result.
- **Restriction on Generalisability:** The data set might not be representative of all the e-commerce businesses, markets, or advertising situations.

Future Research

- **Company-level datasets:** Future studies should utilise company-level datasets that include more detailed campaign and business data.
- **Profitability measures:** More research is needed to add actual profits, contribution margin, and product-cost data to assess commercial outcomes in addition to revenue efficiency.
- **Alternative decision models:** Multiple decision procedures and AI models should be tested to see if recommendations remain the same when using different classification.
- **Longitudinal Analysis:** Advertising performance and the quality of decisions should be compared with advertising data over a period of time.
- **Cross-platform comparison:** Future research should examine how advertising is performing on each platform (Amazon, Google, and Meta) to determine if there are platform-specific patterns.
- **Human–AI comparison:** Benchmarks against human marketing experts could be created for evaluating AI-aided advertising choices.
- **Experimental testing:** Experimental testing of the rule-based and AI-generated recommendations will help check the effectiveness of the actual advertising and commercial results in future campaigns.

Conclusion

This study aimed to compare the advertising decisions made using ROAS-based evaluation and the advertising decisions made using a multi-metric evaluation, and to discuss the relationship between the recommendation with the ROAS-based and multi-metric evaluations. The ROAS-based analysis revealed a significant amount of variation across the observations: ROAS values were often very high, which was due to very little investment in advertising. The multi-metric analysis provided valuable additional information: showing that several high-ROAS observations were associated with low sales and low advertising spend, while some moderate-ROAS observations generated high sales volumes and revenue. The rule-based procedure classified the same category for 31 of 47 observations (65.96%), matching the ROAS-based classification, which is partly due to the ROAS thresholds being the same. The mapping from multi-metric labels to rule-based decisions was not pre-defined; hence, there is no multi-metric agreement rate. The study concludes that keyword-level advertising performance cannot be fully understood through ROAS alone. Considering complementary indicators such as CTR, sales volume, advertising expenditure, CPC, and CPM provides additional context for interpreting performance.

Conflict of interest: The author declares no conflict of interest.
AI-Use Statement: No AI was used in this study.

Data-Availability: The dataset analyzed during this study is publicly available in the Mendeley Data repository. The data were released by Gezici (2025) and can be accessed at Mendeley Data. https://doi.org/10.17632/jpz9kcnrhd.1. The dataset consists of 6 Excel files containing keyword-level e-commerce advertising data, which were amalgamated into a single dataset for this analysis.

References

Almestarihi, R., Bani Ahmad, A. Y. A., Frangieh, R. H., Abu AlSondos, I. A., Nser, K. K., & Ziani, A. (2024). Measuring the ROI of paid advertising campaigns in digital marketing and its effect on business profitability. *Uncertain Supply Chain Management, 12*, 1275–1284. https://doi.org/10.5267/j.uscm.2023.11.009

Cillo, P., & Rubera, G. (2025). Generative AI in innovation and marketing processes: A roadmap of research opportunities. *Journal of the Academy of Marketing Science, 53*, 684–701. https://doi.org/10.1007/s11747-024-01044-7

Ford, J. B., Jain, V., Wadhwani, K., & Gupta, D. G. (2023). AI advertising: An overview and guidelines. *Journal of Business Research, 166*, 114124. https://doi.org/10.1016/j.jbusres.2023.114124

Gezici, H. (2025). *ROAS* [Data set]. Mendeley Data. https://doi.org/10.17632/jpz9kcnrhd.1

Gogou, E. (2025). *Marketing metrics and analytics in consumer behaviour* [Doctoral dissertation, Athens University of Economics and Business, School of Business, Department of Business Administration].

Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2025). How generative AI is shaping the future of marketing. *Journal of the Academy of Marketing Science, 53*(3), 702–722. https://doi.org/10.1007/s11747-024-01064-3

Hu, Y., Shin, J., & Tang, Z. (2010). Pricing of online advertising: Cost-per-click-through vs. cost-per-action. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–9). IEEE. https://doi.org/10.1109/HICSS.2010.470

Kshetri, N., Dwivedi, Y. K., Davenport, T. H., & Panteli, N. (2024). Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda. *International Journal of Information Management, 75*, 102716. https://doi.org/10.1016/j.ijinfomgt.2023.102716

Ljepava, N. (2022). AI-enabled marketing solutions in marketing decision making: AI application in different stages of marketing process. *TEM Journal, 11*(3), 1308–1315.

Ngo, V. M. (2025). Human–AI collaboration for marketing capabilities: A meta-analysis. *Marketing Letters, 36*(4), 839–855. https://doi.org/10.1007/s11002-025-09783-5

Rainy, T. A. (2025). AI-driven marketing analytics for retail strategy: A systematic review of data-backed campaign optimisation. *International Journal of Scientific Interdisciplinary Research, 6*(1), 28–59. https://doi.org/10.63125/0k4k5585

Sands, S., Campbell, C., Ferraro, C., Demsar, V., Rosengren, S., & Farrell, J. (2024). Principles for advertising responsibly using generative AI. *Organizational Dynamics, 53*(2), 101042. https://www.sciencedirect.com/science/article/pii/S0090261624000159

Shan, L. (2024). Computing advertising, intelligent computing, and push based on artificial intelligence in the big data era. *Heliyon, 10*(17), e37252. https://doi.org/10.1016/j.heliyon.2024.e37252

Vangelov, N. (2020). Managing key performance indicators for successful online advertising campaigns. In *Communication management: Theory and practice in the 21st century* (pp. 219–227). Faculty of Journalism and Mass Communication, Sofia University "St. Kliment Ohridski."

Ziakis, C., & Vlachopoulou, M. (2023). Artificial intelligence in digital marketing: Insights from a comprehensive review. *Information, 14*(12), 664. https://doi.org/10.3390/info14120664