ANALYTICAL DECISION LIST ALGORITHM FOR MANAGING CUSTOMER REACTION TO A MARKETING CAMPAIGN

Ushakova I., Hrabovskyi Y., Szymczyk K.*

Abstract: In this paper, the authors aim to develop a methodology for customer segmentation based on their response to marketing campaigns, considering customer value using predictive analytics methods and computer modeling tools. The scientific novelty of this article is the method of modeling and analyzing customer reactions to a marketing campaign. This method includes the following stages: questionnaire development and customer data collection; preliminary analysis of the received data; preparation of customer data in a formalized presentation; RFM analysis of customer value; building a model of customer feedback on a marketing campaign based on the solution list algorithm; analysis of the obtained results. The decision list algorithm was chosen to model customer response to marketing campaigns, which provides an inherent order to the rule set and a more accessible interpretation of the results. The IBM SPSS Modeler was used as a modeling tool. Customer information for the model was obtained through a survey conducted among customers of companies manufacturing packaging goods using a specially designed questionnaire. The practical value of the research lies in the application of the results of customer segmentation to create marketing strategies by a company that can consider the results of both models and group them to cover a wider range of customers.

Key words: predictive analytics, strategic foresight, marketing campaign, decision list algorithm, computer modelling

DOI: 10.17512/pjms.2023.28.1.21

Article history:
Received July 09, 2023; *Revised* August 21, 2023*; Accepted* September 20, 2023

Introduction

Wide access to information influences customer behavior, which quickly changes according to their preferences. Therefore, customer analytics has become an increasingly important tool for any business. If a company has information about customer purchasing habits and behavior, it can predict future customer behavior. The marketing campaign aims to increase profits, maintain a good reputation,

* **Iryna Ushakova**, PhD Assoc. Prof., Simon Kuznets Kharkiv National University of Economics, Ukraine;
✉ email: iryna.ushakova@hneu.net,
ORCID: 0000-0001-8315-0917
Yevhen Hrabovskyi, PhD Assoc. Prof., Simon Kuznets Kharkiv National University of Economics, Ukraine;
✉ email: yevhen.hrabovskyi@hneu.net,
ORCID 0000-0001-7799-7249
Katarzyna Szymczyk, PhD Asst. Prof., Czestochowa University of Technology, Poland;
✉ corresponding author katarzyna.szymczyk@pcz.pl,
ORCID: 0000-0003-0869-2454

increase brand awareness, attract new customers, and support existing customer loyalty (Wolniak, 2020; Kazmierczak-Piwko et al., 2022; Baryshnikova et al., 2021). To develop an effective marketing campaign strategy, marketers use various types of analysis such as SWOT analysis, PEST analysis, BCG matrix, and other tools (Czajkowska, 2016; Ho, 2014; Mohajan, 2017; Hambrick et al., 1982). This helps to identify what needs to be done to attract the maximum number of customers. The vital marketing campaign stages are customer feedback, effectiveness analysis, campaign control, and adjusting. A well-planned marketing campaign increases sales and promotes the popularity of the brand, services, and products. The authors of the paper assume that access to information affects customer behavior, and therefore, customer analytics might be an increasingly essential and helpful tool for any business. As mentioned above, the information about customer purchasing habits and behavior could be beneficial for the company in terms of strategically predicting and managing future customer behavior and developing marketing campaigns that will receive positive customer feedback (Karimi et al., 2015; Stancu and Meghisan, 2014; Knez and Obrecht, 2018). Thus, developing a methodology for customer segmentation based on their response to marketing campaigns, as well as considering customer value, using predictive analytics methods and computer modeling tools, can be effective for business policy. This allows for substantiating decision-making regarding the marketing strategy toward company customers. The authors present the research that includes modeling customer reactions to marketing campaigns. They discuss methods and tools for customer segmentation based on their response to marketing campaigns, considering customer value using computer modeling tools.

Literature Review

Recently, various researchers like Asniar and Surendro (2019), Liu et al. (2020), Brei (2020), Bahrami et al. (2020), Noori (2021), De Caigny et al. (2021), Kim and Lee (2021), Zhao and Keikhosrokiani (2022), Kasem et al. (2023) or Hou et al. (2023) dedicated their work to the issues of predictive customer behavior analytics. Customer information in marketing research is diverse, containing substantial data and structured and unstructured characteristics. Therefore, virtually all studies emphasize the need for modern information technologies to address a wide range of modeling tasks, relying on advanced classification algorithms, machine learning methods, and artificial intelligence (Dorohov et al., 2020; Gupta and Joshi, 2022; Hrabovskyi et al., 2022a; Tong and Saladrigues, 2023; Reshetnikova and Mikhaylov, 2023).

The impact of artificial intelligence and machine learning in the field of marketing is explored in the works of Brei (2020), Siau and Yang (2017), Sterne (2017), Kalicanin et al. (2019), Mauro et al. (2022). The authors discuss aligning significant types and strategies of machine learning and its algorithms with marketing research. However, the analysis also highlights that the existing toolkit to support machine learning may no longer be entirely up-to-date, primarily due to rapid developments

in this domain. Sarker (2022) presents a comprehensive view of modeling based on artificial intelligence, particularly emphasizing the complexity of tasks related to the dynamic nature of real-world problems and data. Noori (2021), Kim and Lee (2021), Liu et al. (2020), Dorokhov et al. (2020), Ushakova et al. (2021), and Hrabovskyi et al. (2022b) also pay attention to the quality and adequacy of the built models, and the accuracy of the obtained results. Many studies construct models as a means of justifying decision-making regarding further customer interactions and retaining the loyalty of valuable customers (De Caigny et al., 2021; Hou et al., 2023; Dorokhov et al., 2020), which the authors consider to be the objective of such research. Some studies are devoted explicitly to exploring customer churn issues (Tamaddoni Jahromi et al., 2014; De Caigny et al., 2021; Kim and Lee, 2021; Dorokhov et al., 2020), suggesting the use of models based on modern data analysis tools. Many works focus on the necessity of analyzing diverse information to determine customer loyalty, address retention issues, and attract new customers (Kim and Lee, 2021; Dorokhov et al., 2020; Gupta et al., 2022; Rauyruen and Miller, 2007). At the same time, traditional analysis methods like SWOT, PEST, BCG, or MCDA are commonly used for investigating customer response to marketing campaigns (Andersen et al., 2019; Bobocea et al., 2016; Grundey, 2010; Wątróbski et al., 2016; Oliveira and Dias, 2020; Castanho et al., 2023).
Nonetheless, noteworthy algorithms can be employed for analyzing customer reactions to various marketing initiatives [SPSS ModelerAlgorithms Guide, 2020; Hrabovskyi et al., 2022a; Giri and Paul, 2020; Alwis et al., 2018). Having explored the literature, the authors claim that conducting a comprehensive study of methods and tools for modeling and analyzing customer response to marketing campaigns, particularly customer segmentation based on their response and value, is highly relevant. The authors, therefore, hypothesize that modeling customer reactions to marketing campaigns through a decision list algorithm, using IBM SPSS Modeler as a tool, allows for an inherent ordering of the set of rules and a more straightforward interpretation of the results. Developing a methodology for segmenting customers based on their reactions to marketing campaigns and, importantly, considering customer value and using predictive analytics methods and computer modeling tools should allow the company to justify decisions regarding the marketing strategy toward its customers adequately. Therefore, using customer segmentation results to create marketing strategies has a practical dimension and value for business. If a company plans to improve its marketing strategy, it can take the results of both models and group them to cover a broader range of its customers.

Research Methodology

To develop a methodology for customer segmentation based on their response to marketing campaigns, considering customer value, using predictive analytics methods and computer modeling tools, the authors accomplished the following main tasks:

1) investigating the peculiarities of the decision list algorithm, understanding its essence and applicability for grouping objects with corresponding behavior, and defining metrics for evaluating the model as a whole and individual segments,
2) defining the stages of modeling customer response to marketing campaigns and the sequence of actions required to perform these stages,
3) developing a questionnaire for customer surveys,
4) conducting a preliminary analysis of the survey results,
5) preparing customer data in a formalized presentation,
6) performing RFM analysis of customer value,
7) building a computer model of customer response to marketing campaigns based on the decision list algorithm,
8) analyzing the obtained results.

First, it should be mentioned that decision lists are widespread in predictive analytics. Modeling is based on selecting the value of the target variable according to certain input variables and constructing a decision list where the internal nodes correspond to one of the input variables. Edges connect the nodes to their descendants for each possible value of the input variable. Each leaf represents the value of the target variable, determined by the values of the input variables from the root to the leaf.

In this work, the decision list is employed to build the predictive model, a specific case of a decision tree. The decision list, an ordered set of rules, is a one-sided tree in which each internal node has at least one leaf as a descendant. Each internal node has one terminal node and one internal node as its descendants, except for the bottom node, where the only descendant is a leaf node. Decision lists represent Boolean functions that can be quickly learned from examples. The decision list algorithm (DL) of length r has the following form:

$$
\begin{array}{l}
if\ f_1\ then \\
\quad output\ b_1 \\
else\ if\ f_2\ then \\
\quad output\ b_2 \\
\quad \ldots \\
else\ if\ f_r\ then \\
\quad output\ b_r
\end{array}
$$

Where fi is the $i-th$ formula, and bi is the $i-th$ Boolean value for $i\in\{1...r\}$.

The expression "$if{-}then{-}else$" in the algorithm means that the formula fr always equals $true$.

Depending on the number of terms contained in the formula fi, the authors distinguish:

the ($k{-}DL$) list, where all formulas contain at most k terms,

the ($1{-}DL$) list, where all formulas consist of a single variable or its negation.

In this paper, the objective of creating the model based on the decision list is to identify groups of objects, such as customers, with a clear behavioral pattern regarding the probability of purchasing a product. A typical set of decisions in the model consists of decision rules. A decision rule is a rule of the form "$if-then,$" which consists of two parts: the antecedent and the consequent. The antecedent part "if" is a Boolean expression of predictors, and the consequent "$then$" is the predicted value of the target variable when the predictor is true. The most straightforward construction of a decision rule is a segment based on a single predictor; for example, Company Size = "Medium" or Income ≥ 10000. An object is included in the rule if the antecedent rule is true. If any object falls under the action of one of the rules in the list of decisions, it is considered to fall under the action of that list. The order of rule traversal in the decision list is important; if an object falls under the rule, it will be assigned to that segment and subsequently ignored by the following rules. The last expression of the "*if-then-else*" algorithm means that the formula fr always evaluates to true. Depending on the number of terms contained in the formula fi, the authors distinguish the (k-DL) list, where all formulas contain at most k terms, and the (1-DL) list, where all formulas consist of either a single variable or its negation. The decision list defines subgroups or segments that show higher or lower likelihoods of binary data outcomes (yes or no) compared to the entire sample. For example, one could search for customers with a low probability of response or with a high probability of response to a specific offer or company. The principles of creating a decision list can be summarized as follows:
1. Potential rules can be found in the initial dataset.
2. The best rules are added to the decision list.
3. Data included in the decision list are removed from the database.
4. New rules can be found based on the reduced dataset.
5. The process is repeated until one or several stopping criteria are met.

Building the model includes creating the decision list, forming the rules, and generating the segments.

Creating the Decision List

1. Among all the elements used to form the set of M decision list models, select a subset of records - $\bar{X}_{M_i}$ that do not match the rules of the M_i decision list (1):

$$X_{\bar{M}_i} = X - X_{M_I}, \tag{1}$$

2. Create an alternative set of rules - R for the set of records - $\bar{X}_{M_i}$ by applying the decision rule algorithm. Build a new set of potential models by adding rules to R. Save the expanded list into the set of candidate decision list models - T.

3. Select the decision list models from the set of candidate decision list models - T: calculate the estimated response probability - $\hat{P}_{M_i}$ for each model in T:

$$\hat{P}_{M_i} = \frac{N(Y_n=1, X_n \in X_{M_i})}{N\ (X_n \in X_{M_i})} \tag{2}$$

- select the decision lists from the T set with the highest value of estimated response probability - $\hat{P}_{M_i}$.

Formation of Decision Rules

Using decision rules involves searching for segments to increase the number of occurrences of the target outcome. Decision rules can be designed to identify segments with either higher or lower probabilities. To create decision rules, the following steps are necessary:

1. Select, from the set of records - X, a subset of records - X_{R_i} covered by the R_i rule with A_j attributes.
2. Create a new segment in the set of segments - S according to the R_i rule.
3. Calculate the estimated response probability - $\hat{P}_{R_i}$ for the created rule.

$$\hat{P}_{R_i} = \frac{N(Y_n=1, X_n \in X_{R_i})}{N\ (X_n \in X_{R_i})} R_i. \tag{3}$$

Select the rules with the highest $\hat{P}_{R_i}$ and add them to the R rules.

Segment Generation

The division of decision rules generates segments with a high response rate from a single attribute (field). Segments are created based on records and their attributes. The algorithm applies to all attributes that have unambiguously ordered values. The segments generated by the algorithm can be used to expand an n-dimensional rule into an (n + 1)-dimensional rule.

Stages of Generation

1. Calculate the observed response probability - P_i for each $i-th$ category (attribute) - C_i from the set of all attributes - C, a sorted list of attribute values.

$$P_i = \frac{N(Y_n=1, X_{n,c} \in C_i)}{N\ (X_{n,c} \in C_i)}, P_0 = P_{(M+1)} = 0. \tag{4}$$

2. Find local maxima - P_i to create a set of segments. Select the segment with the highest response probability.

The following metrics are used to evaluate the model as a whole and individual segments:

1. Confidence interval.

The confidence interval (p^-, p^+) for the response probability - $\hat{p}$ is calculated using the following formulas:

$$\mathrm{p}^- = \begin{cases} \dfrac{\mathrm{x}}{\mathrm{x}+(\mathrm{n}-\mathrm{x}+1)\mathrm{F}_{2(\mathrm{n}-\mathrm{x}+1),2\mathrm{x};1-{}^{\alpha}/_{2}}}, \mathrm{x} \neq 0 \\ 0,\ \mathrm{x} = 0 \end{cases} \tag{5}$$

$$p^+ = \begin{cases} n - \dfrac{(x+1)F_{2(x+1),2(n-x);1-{}^{\alpha}/_{2}}}{x+(x+1)F_{2(x+1),2(n-x);1-{}^{\alpha}/_{2}}}, x \neq n \\ 1,\ x = n \end{cases} \tag{6}$$

where n is the coverage of the rule or list,
x is the frequency of responses for the rule or list,
α is the desired level of confidence,
$F_{a,b,c}$ is the inverse cumulative distribution function for F with degrees of freedom - a and b for the percentile.

2. The following metrics are used for each segment:

2.1. Coverage.
The number of records in the segment:

$$N(S) \tag{7}$$

2.2. Frequency.
The number of records in the segment for which the response is true:

$$N(Y_n = 1, X_n \in S) \tag{8}$$

2.3. Probability.
The proportion of records in the segment (Frequency/Coverage) for which the response is true:

$$(N(Y\_n = 1, X\_n \in S))/N(S) \tag{9}$$

The decision list algorithm comprises a set of individual classification rules that collectively form a classifier and has several advantages:
-unlike an unordered set of rules, decision lists have built-in order, making classification relatively straightforward,
-the input space is fully open (i.e., there are no limitations on the number and content of input rules),
-the input space is divided based on unlimited dimensions, allowing all possible logical propositions,
-the input space is divided arbitrarily, not by automatically detected conditions, and does not optimize the loss function of specific criteria.
As shown above, modeling customer response to marketing campaigns is complex and involves specific stages. As mentioned at the beginning, the following stage involves creating a questionnaire and conducting a survey. Developing a model for predicting customer response to marketing campaigns for promoting new products is considered in a company specializing in producing various types of tubes, including laminated, polyethylene, and aluminum ones. The company's marketing system is an integral part of the effective management of production and sales activities, thus requiring a well-established information exchange system between these processes. By leveraging modern information technologies, data accumulated during sales, production, and customer interactions can be used for predictive analytics to fine-tune marketing campaigns. The online survey based on the questionnaire was conducted in Ukraine in 2023, in the second quarter of the year, during the break between blackouts and active phases of hostilities. It involved 512 respondents from various types of companies, with 39.9% of small enterprises, 43.2% of medium-sized enterprises, and 17% of large enterprises. The companies were divided based on their production type, with 29.3% – engaged in food production, 24.5% – in the pharmaceutical industry, 24.1% – in household chemicals production, and 0.8% – in other sectors.
Having created the survey questionnaire and conducting the study, the authors conducted RFM analysis in the next stage. The final stages included building a computer model of customer response to marketing campaigns based on the decision list algorithm and discussing the obtained results.

Research Results

In the survey, 34.10% showed interest in laminate tubes, 19.8% in polyethylene tubes, 33% in aluminum tubes, and 14.7% were interested in tube caps. Among all the surveyed companies, almost 60% had already bought products, while more than 40% had yet to make any purchases. The survey was to indicate whether the respondents were satisfied with the company's pricing policy to check the customer feedback. The answers were almost evenly divided. 51.2% were content with the pricing, and 48.8% expressed dissatisfaction. As for the product quality, only about a quarter of the respondents (27.6%) were dissatisfied with the company's products.
In the next stage, the data were extracted from the obtained survey results and customer database, including information about contracts, such as the number of previous purchases by each company, the date of the last purchase, and other relevant details. However, to construct the model, these data need to be presented in a more structured and convenient format for processing. Preparing the input information for model building involves formalizing the data into a more structured presentation suitable for further processing.
The following step involves conducting RFM analysis. An important indicator to consider when modeling customer response to a marketing campaign is customer value, which can be determined based on RFM analysis, prioritizing each customer's importance to the business. To assess this value, each customer is examined based on the following attributes:
Recency - refers to the time since the last purchase.
Frequency - indicates the purchase frequency.
Monetary - represents the monetary value of purchases.
The information for determining these attributes is gathered from the survey results and/or existing databases of the company's information system:
Recency - calculated as the time elapsed since the last purchase date.
Frequency - extracted from the company's product sales database.
Monetary - computed as the total amount spent by the customer.
The first step of the analysis is to categorize customers based on each attribute. The authors identify five categories for each attribute. To determine the categories, it is necessary to find the minimum and maximum values for each attribute:
Recency: min - the most recent purchase, max - the earliest purchase.
Frequency: min - the lowest number of orders; max - the highest number.
Monetary: min - the smallest order amount, max - the largest order amount.
The determination of the step for categorization is done using the formula:

$$L_g = max_g - min_g, \tag{10}$$

$s_g = \frac{L_g}{n}$,
where L_g – The range of values for g attribute (Recency, Frequency, Monetary accordingly);
s_g, – the category step size for the g attribute;

max_g, mix_g – the maximum and minimum value for the g attribute.
n – the number of categories, in our case 5.

Subsequently, the boundaries for each category of the attributes are determined. The first category ranges from mix_g to point - $x1_g$, the second category ranges from point $x1_g$ to point $x2_g$ (Figure 1). The category boundaries are rounded in consideration of analysis. Based on these points, the matrix of customer categories is created (Table 1).

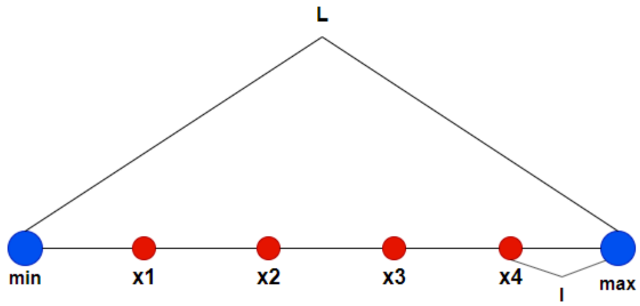

Figure 1: Division of customers into groups
Source: Own elaboration

Table 1. Matrix of client categories by RFM

Category	Recency, days		Frequency, times		Monetary, thousand UAH	
1	Long ago	> 600	Very rarely	≥ 1 ≤6	Very little	<2500
2	Very long ago	> 450 ≤600	Rarely	> 6 ≤12	Little	≥2500<5000
3	Moderately long ago	>300≤450	Average	>12≤18	Average	≥5000<7500
4	Recently	>150≤300	Often	>18 ≤24	Many	≥10000<12500
5	Very recently	>1≤150	Very often	> 24	A lot	≥12500

Source: Own elaboration

Each customer is assigned a numerical value for the corresponding attribute category. For example, if the Recency value falls within the range from min to x1, it is assigned the value 5, which corresponds to the category "very recently." The overall RFM score is obtained as the sum of the categories.

$$RFM = \sum_{i=1}^{3} k_i, \quad (11)$$

where RFM- the overall customer value indicator,
k_i – the value of the customer category by the corresponding RFM attribute.

Thus, all the customers are divided into 15 categories in the segment. An example of the results of customer segmentation into categories according to RFM analysis is shown in Table 2.

Table 2. Customer segmentation by RFM

№	Recency	R	Frequency	F	Monetary	M	RFM
1	370	3	2	1	67800	1	5
2	256	4	1	1	56500	1	6
3	344	3	6	1	918000	1	5
4	252	4	24	4	1101600	1	9
5	96	5	12	3	370800	1	9
6	40	5	9	2	143100	1	8
7	293	4	5	1	249000	1	6

Source: Own elaboration

Customer segmentation by RFM allows one to focus on more valuable customers for the business and orient the modeling based on this criterion. The decision tree algorithm is utilized to build a predictive model for customer response to a marketing campaign. The algorithm generates rules determining the probability of a binary outcome, "yes" or "no." The visual modeling tool - IBM SPSS Modeler is employed to construct the model. This data exploration toolkit allows for rapidly developing predictive models that leverage business knowledge and experience. This enables the implementation of these models into business operations to enhance decision-making processes. The constructed model is presented as a flow consisting of several sequential nodes (Figure 2).

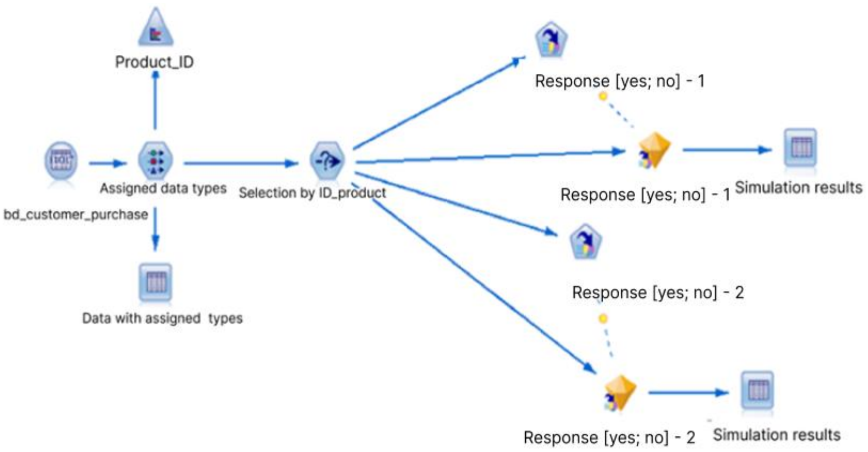

Figure 2: Customer Response Modelling
Source: Own elaboration

The first node defines the data source. This node attaches the input data file in Excel format. IBM SPSS Modeler reads the data values from this file to build the model

based on them. The next node is the "Data Type" node, which determines the data type for each column in the table to facilitate further processing. At this stage, one must select the target field, "Response" - the customer's response to the company. This field can only take two values: 1 - the customer responded to the company, or 0 - the customer did not respond to the company. Above the "Data Type" node, there is a node called "Diagrams," which is connected to the "Data Type" node to show the distribution of purchase percentages by product type visually. The diagram is constructed using the "product_id" field (Figure 3).
Additionally, the number of responses for each product type is determined on the chart using the target field "Response." The diagram shows eight variants of product purchases: the first four correspond to one product type (3; 1; 2; 4), and the remaining four represent a combination of purchasing two product types simultaneously (1,4; 2,1; 2,4; 3,1), each with a lower percentage volume. Hence, these combined records can be ignored in the model-building process.

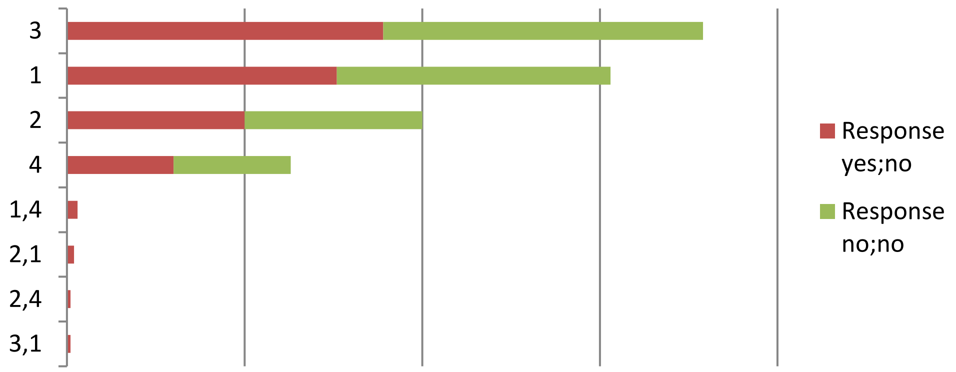

Figure 3: Distribution of responses broken down by products
Source: Own elaboration

The next node is the "Selection" node (Figure 2). This node defines the conditions for selecting data based on the "product_id" field. Although our data, as shown in the diagram (Figure 3), include information about four different product types and four possible combinations of products, only single purchases are selected for model building under the following conditions:

$$product\_id = '1' \; or$$
$$product\_id = '2' \; or$$
$$product\_id = '3' \; or$$
$$product\_id = '4'.$$

Subsequently, the actual process of model creation begins. In this case, two models are constructed with different configurations to achieve better results and use the results of both models to cover a more extensive customer base. Subsequently, the "Decision List" node is added to the flow, which directly creates the model for

customer response to the marketing campaign using the decision list algorithm. To build the model, one needs to define its parameters: the value for the target field - "Response"; the maximum number of customer segments - 5; the minimum segment size, in this case, for model_1 – 20, for model_2 - 30; the confidence interval for model_1 – 55, for model_2 – 50.
The following fields were selected for modeling customer feedback: an overall indicator of client usefulness (RFM), a sign of future cooperation, duration of the company's existence, company size, duration of cooperation, purchase sign, income from the last purchase, number of previous purchases, satisfaction with the price.
After configuring both variants of the model nodes, the model is launched for execution. A corresponding node is added to see the fields selected during the simulation. The node adds three new fields, including a score (1 meaning the record is included, or $null$ for excluded records), the probability (hit rate) of the segment the record falls into, and the ID number for the segment.

Discussion

In the preliminary modeling results, several alternative options can be chosen even before the selection of segments. At the same time, all records fall into the "residual". Of all the entries, 249 responses are selected, which is a 50.30% probability. Since the creation of three alternative options was chosen during the setup process, the system issues exactly three options (Table 3). Among the three options, "Alternative 1" has the highest hit probability (68.2%), so the authors leave this option for further modeling.

Table 3. Alternative options

Name	Appointment	Number of segments	Coating	Frequency	Probability
Alternative 1	1.0	5	29	88	68.22%
Alternative 2	1.0	4	103	70	67.96%
Alternative 3	1.0	4	134	89	66.42%

Source: Own elaboration

Using the rules based on predictors, among which there were selected revenue from the customer's last purchase, number of previous orders, size of the customer's company, and RFM, the model identifies five segments (Figure 4) with a response rate more significant than the overall sample. The finished model covers 129 records from the entire sample. And the number of hits to coverage is 88.

ID	Segment rules	Grade	Coating	Frequency	Probability
			495	249	50.30%
1	Income, Number_prev_purch Income>159153.500 и Income<=288591.000 и Number_prev_purch>2.000 и Number_prev_purch<=14.000	1.0	22	17	77.27%
2	RFM, Income RFM>5.000 и Rfm<=7.000 и Income>14331.000 и Income<=56463.000	1.0	32	22	68.75%
3	RFM, Income RFM>5.000 и Rfm<=7.000 и Income<=548817.000	1.0	22	15	68.16%
4	Income, Company_size Income>288591.000 и Income<=414857.500 и Company_size="M"	1.0	27	19	70.37%
5	Number_prev_purch Number_prev_purch>0.000 и Number_prev_purch<=1.000	1.0	26	15	57.69%
	Remainder		366	161	43.99%

Figure 4: Segmentation results, model 1
Source: Own elaboration

Among all the customers analyzed in the model, those who meet the three conditions are included in the first segment: responded positively to future marketing offers, placed an order for UAH 159.153 (approx. EUR 4.06) up to UAH 288.59 (approx. EUR 7.36), placed from 2 to 14 orders. The customers are included in the second segment of the list based on such criteria as RFM. The amount of the last purchase, namely, the customer value to the enterprise (RFM), is between 5 and 7 units (this is lower than the average, as the actual range is between 3 and 15); the amount of the last order ranges from UAH 14.331 to UAH 56.463 (approx. EUR 0.37 to EUR 1.44). The selection criteria for another segment are company size and revenue. The companies with the feedback meet the following conditions: medium size (M – medium), order UAH 288.59 up to UAH 414.86 (approx. EUR 7.36 EUR to EUR 10.58).

Subsequently, the authors analyzed the results using the gain chart (Figure 5). To construct this diagram, the data in Table 4 were used.

Table 4. Coverage and Profit by Segment, Model 1

№ of segment	Coating, %	Win, %
1	4.4	6.8
4	5.45	7.6
2	6.46	8.8
3	4.4	6
5	5.25	6
All	25.96	35.2

Source: Own elaboration

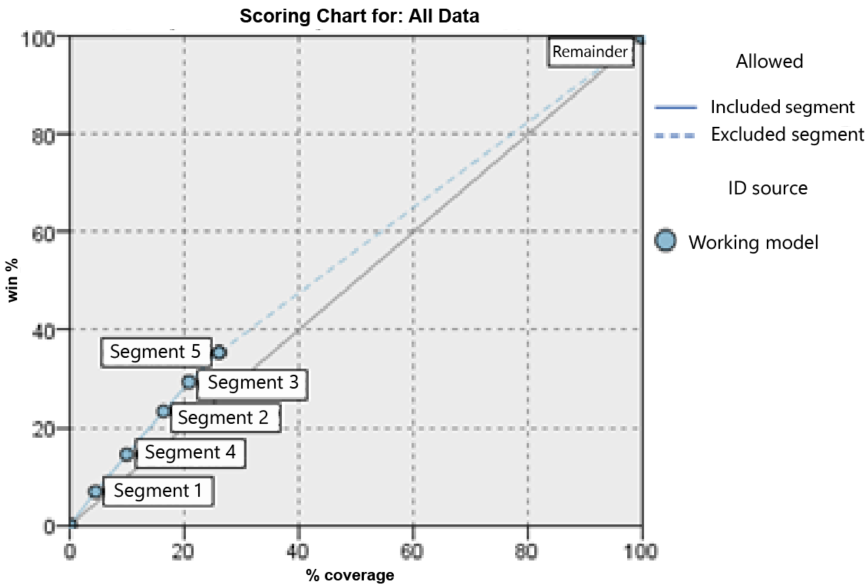

Figure 5: Gain diagram, model 1
Source: Authors' elaboration

The win chart shows the percentage in relation to the total number of hits. The gain is defined as the ratio of the number of hits on each division of the scale to the total number of hits in the tree according to the following formula:

$$G_i = \frac{g_i}{\sum_i g_i} \cdot 100\%, \tag{12}$$

where – g_i the number of hits in the segment,
$\sum_i g_i$ – total number of hits.
The gain diagram clearly illustrates how wide the tree needs to be searched to get a given percentage of all hits. The diagonal line shows the expected response for the entire sample without using the model. In this case, the response rate would be a constant since the probability of response for one customer is the same as for another.

The curved line indicates how much the feedback can be increased by including only those customers who are ranked in the highest percentages based on winnings.
The analysis of the results of the first model shows that the constructed model gives positive results. For comparison, the authors analyzed the results of the second model. Among the three issued alternatives, the authors chose the first one (Table 5). Although the first alternative covers fewer records, it has a higher probability of being hit.

Table 5. Alternative options, model 2

Name	Appointment	Number of segments	Coating	Frequency	Probability
Alternative 1	1.0	4	161	100	62.11%
Alternative 2	1.0	4	162	101	61.96%
Alternative 3	1.0	4	163	101	61.96%

Source: Own elaboration

The results of the second model show four constructed segments, which cover 161 records with a hit frequency of 100. Four fields are involved in the segments of this model: RFM, $Income$, $Number\_prev\_purch$, $Collaboration\_time$.
The coverage and profit by segments and the profit diagram for the second model are presented in Table 6 and Figure 6, respectively.

Table 6. Coverage and profit by segment, model 2

№ of segment	Coating, %	Win, %
1	8.2	11.64
2	6.26	8
3	6.26	8
4	11.7	12.4
All	32.42	40.04

Source: Own elaboration

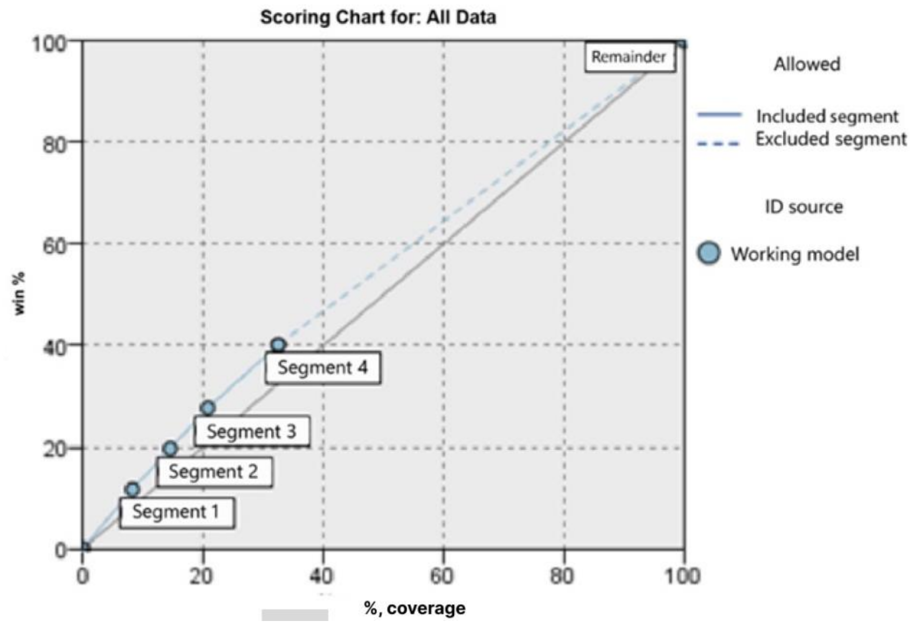

Figure 6: Gain diagram, model 2
Source: Own elaboration

Having analyzed the results of both models, one may conclude that the coverage of records in the second model is more extensive (the first model - 129, the second model - 161), and the frequency is also wider in the 2nd model (the first model - 88, the second model - 100). The profit percentage in the second model is 5.02% higher than in the first; one may expect a more significant customer response using the second model. Such results were obtained by reducing the probability of a hit: the probability of the first model is 68.22%, and the second – 62.11%. In this way, the gain and coverage of the model can be increased by slightly reducing the hit probability, which is relatively high for both models. Both models are suitable for practical use in the enterprise and can be used individually to build future marketing proposals based on grouped results. The main result is precisely the indicators (predictors) and their values determining customer segments. This information is essential for making decisions about the company's marketing strategy.

Conclusion

The authors of the paper carried out the study of existing approaches and methods for analyzing customer reactions to marketing companies, identified the main problems in this area, and proposed some methods to solve them based on the application of a comprehensive approach to the analysis and forecasting of customer reactions to companies, considering their value. The method based on the solution list algorithm was chosen for customer classification. This method is a type of decision tree that uses an ordered set of rules. The advantage of this method is the possibility of easier interpretation of the obtained results, which is important in marketing research. The authors of the work identified the main stages of modeling and analysis of customer reactions to a marketing campaign, which include developing a questionnaire and collecting data about customers; preliminary analysis of the received data; preparation of customer data in a formalized presentation; RFM analysis of customer value; building a model of customer feedback on a marketing campaign based on the solution list algorithm using the IBM SPSS Modeler computer modeling tool; analysis of the obtained results. The questionnaire was developed during an experiment, on the basis of which 512 customers were interviewed. The preliminary analysis of the survey results found that most customers are medium-sized enterprises, and more than 60% of customers responded to the marketing campaign. To facilitate further analysis, client data were converted to a digital format. The next stage was conducting RFM analysis of customer value, an essential indicator in modeling customer response. All customers were segmented into categories 3 through 15, and the category value was added to the customer database. In the visual environment of SPSS Modeler, models are built based on the solution list algorithm. When building the models, the client's response to the marketing campaign was chosen as the target field: yes or no. Two models were built with different parameters regarding the maximum number of segments, the minimum segment size, and the confidence interval to find the best results. Both models showed positive results, as confirmed by the metrics such as record coverage and win percentage. In each model, the most likely alternative options were selected. In the first model, five segments were identified, and in the second - four. Although the winning percentage in the second model is slightly higher, both models can be used and implemented in the process of creating marketing activities. To improve the company's marketing strategies, it is possible to group the results of both models, instead of concentrating on one solution, and get more coverage among its customers.

References

Alwis, P. K. D. N. M., Kumara, B. T. G. S. and Hapuarachchi, H. A. C. S., (2018). Customer Churn Analysis and Prediction in Telecommunication for Decision Making. *2018 International Conference on Business Innovation (ICOBI)*, NSBM, Colombo, Sri Lanka.

Andersen, P., Weisstein, F. L. and Song, L., (2019). Consumer response to marketing channels: A demand-based approach. *Journal of Marketing Channels*, *26*(1), 43–59.

Asniar, Surendro, K., (2019). Predictive Analytics for Predicting Customer Behavior. *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT)*.

Bahrami, M., Bozkaya, B. and Balcisoy, S., (2020). Using Behavioral Analytics to Predict Customer Invoice Payment. *Big Data*, *8*(1), 25-37.

Baryshnikova, N., Kiriliuk, O. and Klimecka-Tatar, D., (2021). Enterprises' strategies transformation in the real sector of the economy in the context of the COVID-19 pandemic. *Production Engineering Archives*, *27*(1), 8–15.

Bobocea, L., Spiridon, S., Petrescu, L., Gheorghe, C. and Purcarea, V., (2016). The management of external marketing communication instruments in health care services. *Journal of Medicine and Life*, *9*(2), 137–140.

Brei, V. A., (2020). Machine Learning in Marketing: Overview, Learning Strategies, Applications, and Future Developments. *Foundations and Trends® in Marketing*, *14*(3), 173–236.

Castanho, R. A., Vulevic, A., Gómez, J. M. N., Cabezas, J., Fernández-Pozo, L., Loures, L., ... & Kurowska-Pysz, J. (2023). Assessing the impact of marketing and advertising as strategic approaches to Eurocities development: An Iberian case study approach. *European Journal of International Management*, *19*(1), 58-91.

Czajkowska, A., (2016). SWOT analysis application for indications of the strategy action chosen enterprise in the construction sector. *Production Engineering Archives*, *10*(1), 33–37.

De Caigny, A., Coussement, K., Verbeke, W., Idbenjra, K. and Phan, M., (2021). Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach. *Industrial Marketing Management*, *99*, 28–39.

Dorokhov, O., Dorokhova, L., Malyarets, L. and Ushakova, I., (2020). Customer churn predictive modeling by classification methods. *Series III - Matematics, Informatics, Physics*, *13(62)*(1), 347–362.

Giri, A., Paul, P., (2020). *Applied marketing analytics using SPSS*. PHI Learning Pvt. Ltd.

Grundey, D., (2010). Planning for Sales Promotion at Lithuanian Supermarkets. *Economics and Sociology*, *3*(2).

Gupta, S., Joshi, S., (2022). *Predictive Analytic Techniques for enhancing marketing performance and Personalized Customer Experience*. IEEE Xplore.

Hambrick, D. C., MacMillan, I. C. and Day, D. L., (1982). Strategic Attributes and Performance in the BCG Matrix—A PIMS-Based Analysis of Industrial Product Businesses. *Academy of Management Journal*, *25*(3), 510–531.

Ho, J. K. K., (2014). Formulation of a systemic PEST analysis for strategic analysis. *European academic research*, *2*(5), 6478–6492.

Hou, R., Ye, X., Zaki, H. B. O. and Omar, N. A. B., (2023). Marketing Decision Support System Based on Data Mining Technology. *Applied Sciences*, *13*(7), 4315.

Hrabovskyi, Y., Kots, H. and Szymczyk, K., (2022a). Justification of the innovative strategy of information technology implementation for the implementation of multimedia publishing business projects. *Proceedings on Engineering Sciences*, *4*(4), 467–480.

Hrabovskyi, Y., Minukhin, S. and Brynza, N., (2022b). Development of an information support methodology for quality assessment of the prepress process. *Eastern-European Journal of Enterprise Technologies*, *6*(2 (120)), 30–40.

Kalicanin, K., Colovic, M., Njegus, A. and Mitic, V., (2019). Benefits of Artificial Intelligence and Machine Learning in Marketing. *Proceedings of the International Scientific Conference - Sinteza 2019*.

Karimi, S., Papamichail, K. N. and Holland, C. P., (2015). The effect of prior knowledge and decision-making style on the online purchase decision-making process: A typology of consumer shopping behaviour. *Decision Support Systems*, *77*(1), 137–147.

Kasem, M. S., Hamada, M. and Taj-Eddin, I., (2023). Customer Profiling, Segmentation, and Sales Prediction using AI in Direct Marketing. *ArXiv (Cornell University)*.

Kazmierczak-Piwko, L., Kułyk, P., Dybikowska, A., Dubicki, P. and Binek, Z., (2022). Sustainable consumption among children and adolescents. *Production Engineering Archives*, *28*(3), 257–267.

Kim, S., Lee, H., (2022). Customer Churn Prediction in Influencer Commerce: An Application of Decision Trees. *Procedia Computer Science*, *199*, 1332–1339.

Knez, M., Obrecht, M., (2018). How can people be convinced to buy electric cars? – case of Slovenia. *Production Engineering Archives*, *21*(21), 24–27.

Liu, C.-J., Huang, T.-S., Ho, P.-T., Huang, J.-C. and Hsieh, C.-T., (2020). Machine learning-based e-commerce platform repurchase customer prediction model. *Plos One*, *15*(12), e0243105.

Mauro, A. D., Sestino, A. and Bacconi, A., (2022). Machine learning and artificial intelligence use in marketing: a general taxonomy. *Italian Journal of Marketing*, *2022*(4), 439-457.

Mohajan, H., (2017). An Analysis on BCG Growth Sharing Matrix. *Noble International Journal of Business and Management Research*, *2*(1), 1–6.

Noori, B., (2021). Classification of Customer Reviews Using Machine Learning Algorithms. *Applied Artificial Intelligence*, *35*(8), 567–588.

Oliveira, G. D., Dias, L. C., (2020). The potential learning effect of a MCDA approach on consumer preferences for alternative fuel vehicles. *Annals of Operations Research*, *293*(2), 767–787.

Rauyruen, P., Miller, K. E., (2007). Relationship quality as a predictor of B2B customer loyalty. *Journal of Business Research*, *60*(1), 21–31.

Reshetnikova, M. S., Mikhaylov, I. A., (2023). Artificial Intelligence Development: Implications for China, *Montenegrin Journal of Economics*, *19*(1), 139-152.

Sarker, I. H., (2022). AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. *SN Computer Science*, *3*(2). springer.

Siau, K., Yang, Y., (2017). Impact of Artificial Intelligence, Robotics, and Machine Learning on Sales and Marketing. *MWAIS 2017 Proceedings*, *48*.

SPSS ModelerAlgorithms Guide (2020). IBM Corporation. 804 p.

Stancu, I., Meghisan, G.-M., (2014). Strategic Planning Based on Consumers' Decision Making Process towards Mobile Telecommunications Operators. *Procedia Economics and Finance*, *15*, 1528–1534.

Sterne, J., (2017). *Artificial intelligence for marketing: practical applications*. Wiley. Copyright.

Tamaddoni Jahromi, A., Stakhovych, S. and Ewing, M., (2014). Managing B2B customer churn, retention and profitability. *Industrial Marketing Management*, *43*(7), 1258–1268.

Tong, Y., Saladrigues, R., (2023). The Influence of Intellectual Capital on the Financial Performance of Spanish New Firms, *Montenegrin Journal of Economics*, *19*(2), 179-188.

Ushakova, I., Skorin, Y., Shcherbakov, A. and Kharkiv, S., (2021). *Methods of quality assurance of software development based on a systems approach*. III International Scientific and Practical Conference "Information Security and Information Technologies," Odesa, Ukraine.

Wątróbski, J., Jankowski, J. and Ziemba, P., (2016). Multistage performance modelling in digital marketing management. *Economics and Sociology*, *9*(2), 101–125.

Wolniak, R., (2020). Main Functions of Operation Management. *Production Engineering Archives*, *26*(1), 11–14.

Zhao, X., Keikhosrokiani, P., (2022). Sales Prediction and Product Recommendation Model Through User Behavior Analytics. *Computers, Materials and Continua*, *70*(2), 3855–3874.

ALGORYTM ANALITYCZNEJ LISTY DECYZYJNEJ DO ZARZĄDZANIA REAKCJĄ KLIENTA NA KAMPANIĘ MARKETINGOWĄ

Streszczenie: W artykule autorzy stawiają sobie za cel opracowanie metodologii segmentacji klientów na podstawie ich reakcji na kampanie marketingowe, z uwzględnieniem wartości klienta z wykorzystaniem metod analityki predykcyjnej i narzędzi modelowania komputerowego. Nowością naukową artykułu jest metoda modelowania i analizy reakcji klientów na kampanię marketingową. Metoda ta obejmuje następujące etapy: opracowanie kwestionariusza i zebranie danych o klientach; wstępną analizę otrzymanych danych; przygotowanie danych klienta w sformalizowanej prezentacji; analizę RFM wartości klienta; zbudowanie modelu opinii klientów o kampanii marketingowej w oparciu o algorytm listy rozwiązań; analizę uzyskanych wyników. Algorytm listy decyzyjnej został wybrany do modelowania reakcji klientów na kampanie marketingowe, co zapewnia nieodłączne uporządkowanie zestawu reguł i bardziej przystępną interpretację wyników. Jako narzędzie modelowania wykorzystano program IBM SPSS Modeler. Informacje o kliencie dotyczące modelu uzyskano poprzez ankietę przeprowadzoną wśród klientów firm produkujących towary opakowaniowe za pomocą specjalnie zaprojektowanej ankiety. Praktyczna wartość badania polega na zastosowaniu wyników segmentacji klientów do tworzenia strategii marketingowych przez firmę, która może uwzględnić wyniki obu modeli i pogrupować je w celu objęcia szerszego grona klientów.

Słowa kluczowe: analityka predykcyjna, foresight strategiczny, kampania marketingowa, algorytm listy decyzyjnej, modelowanie komputerowe