



# Intelligent Marketing Decision Model Based on Customer Behavior Using Integrated Possibility Theory and K-Means Algorithm



Hamed Fazlollahtabar\*

Department of Industrial Engineering, School of Engineering, Damghan University, Damghan, Iran

\* Correspondence: Hamed Fazlollahtabar ([h.fazl@du.ac.ir](mailto:h.fazl@du.ac.ir))

**Received:** 10-02-2022

**Revised:** 11-17-2022

**Accepted:** 11-23-2022

**Citation:** H. Fazlollahtabar, "Intelligent marketing decision model based on customer behavior using integrated possibility theory and K-means algorithm," *J. Intell. Manag. Decis.*, vol. 1, no. 2, pp. 88-96, 2022. <https://doi.org/10.56578/jimd010202>.



© 2022 by the authors. Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

**Abstract:** E-commerce is referred as any transaction in which the sale and purchase of goods or services takes place via the Internet and leads to the import or export of goods or services. Supply always needs demand. We need marketing to expand the demand in order to sell more services or products. Organizations need to have a good understanding of their customers and their desires in order to become succeed in business, and to get this understanding, they must use tools and techniques to measure the customer's interest. With using of data mining techniques and with the discovery of hidden and valuable knowledge of data, organizations don't miss the opportunity to sell more and provide better customer satisfaction. The customer segmentation is one of the methods of customer recognition. This method is used when we look for groups of similar data. Segmenting is one of the most important topics in reaching modern marketing and managing successful customer relationship management. The purpose of this paper is to design an electronic marketing model using the k-mean algorithm. First, customer's data is collected and after preparing and pre-processing data, using the k-mean algorithm, segmentation customers and future marketing strategies and recommendations are discussed and eventually using the theory of the possibility, the possibility and requirement for the proposal to be considered, and each of the recommendations or strategies are given numbers with the name of the possibility and necessity of the system output, and a more favorable proposal is obtained.

**Keywords:** Intelligent marketing; Customer behavior; Data mining; K-mean algorithm; Possibility theory

## 1. Introduction

E-commerce has been welcomed by many human societies over the last few years. Today, fewer people can be found that the term e-commerce is unfamiliar for them [1]. The development of information and communication technology has created an opportunity for institutions to carry out their business activities quickly. Today, most middle and large organizations in the world have a website. Websites are the tool for marketing [2]. The importance of marketing in today's world is very valuable that misleading marketing decisions may lead to the destruction or failure of the organization, especially in today's global marketplace, which is constantly evolving and rivals offer new products and services every day. On an e-commerce site, e-marketing should help to the customers for buying different goods, which requires customer information [3]. Marketers believe that, in the online shopping activity, customers disclose information that reflects their behavior such as the customer's email address and phone number, occupation, age, education, address, and customer orders, order history, methodologies for paying, favorites, and their requirements that can be used to understand customer behavior [4]. The reason for collecting information is we need to get a sense of customer behavior, how they decide on their needs, and what internet marketing methods are suitable and successful for them. In fact, the availability of customer data and transaction records provide a better understanding of customer behavior and their willingness to market [5]. By recognizing its customers, the marketer can expand relationships with customers and expand the various fields of work as well as improve the decision-making process in the marketing process. The volume of data gathered from the customer data is huge, and analysis of them is one of the problems, but one way to overcome this problem is to use the data mining technique. This tool uses from customer information to help marketers and customers find their products and

services more efficiently [6].

Applying data for marketing as to make decisions intelligent was used profoundly [7]. For agricultural products where the lifecycle is significant it is necessary to make use of data driven approaches to increase the value chain. Recommender systems are useful tool in intelligent marketing based on customer suggestions. Li et al. [8] employed the attention allocation index concept to propose an intelligent prediction algorithm to join users of the same brand products. Martínez-López and Casillas [9] introduces knowledge discovery in database to marketing modeling as a decision support for consumer behavior. They hybridized artificial intelligence methods as a knowledge base marketing management approach. A new paradigm of industry 4.0 related products influenced marketing efforts and decisions. Khan et al. [10] investigated the impact of artificial intelligent and industry 4.0 on consumer behavior using a meta-analysis. They found that intention to buy, acceptance and need for recognition are significantly impact the consumer behavior. Business models are important in forming marketing models and decisions. Hence, Moradi and Dass [11] studied the artificial intelligence use in B2B marketing and listed several pros and cons. The study emphasized the role of machine learning methods throughout the customer life cycle. The analysis employed technology acceptance model to suggest recommendations for marketers and business policy makers. Considering user experience in developing marketing recommendation system is significant. Due to large amount of data, it is necessary to make use of user data into product design [12]. Rahim et al. [13] examined the purchasing behavior of Muslims towards Halal products. They collected a dataset and used structural equation modeling partial least square for statistical analysis. Outsourcing marketing decisions for different purposed were investigated in several researches [14, 15]. Eco-tourism infrastructure effectiveness, demand segmentation, personal relationships, knowledge sharing, and size of the company were the most emphasized factors justified outsourcing marketing decisions. Detecting frauds in market transactions was investigated in the study of Eshghi and Kargari [16] to analyze the trend and developing scenarios for marketers.

## 2. Proposed Problem

In this paper, we design an e-marketing model using the recommendation system. The customer's attitude in the online shopping process is not constant and is dynamic and variable based on gender, age, geographical location and tastes. So, we cannot definitely discuss the behavior of customers as a matter of fact. First, collect information from customers, and after preprocessing data and separating inappropriate data, the relevant data is being prepared for the proposed login. The output of the K-Mean algorithm includes strategies and recommendations for marketing, and finally, with helping of the theory, the possibility and the requirement of the proposal are calculated. The related tools: To use the marketing model, we need to introduce the tools and explain the reasons for using each one. In fact, the proposed system of the research, by analyzing the behavior of its users, proposes the most suitable items for business executives.

### 2.1 Possibility Theory

Among the mathematical theories in terms of ambiguity, the Possibility Theory can be considered the most appropriate and coherent theory in the analysis of the uncertainties of the decision environment. In summary, the content of this theory can be expressed in such a way that in the analysis of environmental events and conditions that we are not only looking for possible events, and in uncertain structures we are looking for all possible events that are introduced with the degree of possible events and the degree of possible contradictory events (see Table 1). In the theory of probability, we try to construe it with the probability of occurrence of the event with the probability that it does not occur, while in more general theory, the uncertainty of an event can be determined by two numbers [2, 6]. The degree of the possibility of self-occurrence and the degree of necessity of the event (the degree of necessity of the event = the degree of probability of contradiction-1).

**Table 1.** Formulas related to the possibility theory

Relationships for possible sizes	Relationship to the required sizes
$\Pi(A \cup B) = \max[\Pi(A), \Pi(B)]$	$\Pi(A \cap B) = \min[\Pi(A), \Pi(B)]$
$\Pi(A) = 1 - N(A^c)$	$N(A) = 1 - \Pi(A^c)$
$\max[\Pi(A), \Pi(A^c)] = 1$	$\min[N(A), N(A^c)] = 0$
$\Pi(A) + \Pi(A^c) \geq 1$	$N(A) + N(A^c) \leq 1$
$\Pi(A) < 1 \Rightarrow N(A) = 0$	$N(A) > 0 \Rightarrow \Pi(A) = 1$
$\Pi(A) \geq N(A)$	

## 2.2 Data Mining Technique

In this paper k-means technique is applied for clustering the customers based on their views and purchase behavior.

### The K-means algorithm:

Clustering is the process of grouping an item set into similar item classes. A popular clustering algorithm proposed by MacQueen is named K-means. The main idea of K-means is to discover k clusters that the records in each cluster are similar to each other and differ from other cluster records. The K-means algorithm is often used to preprocess data or to cluster because of its simplicity and efficiency [1].

## 2.3 Architecture of the Conceptual Model of Marketing

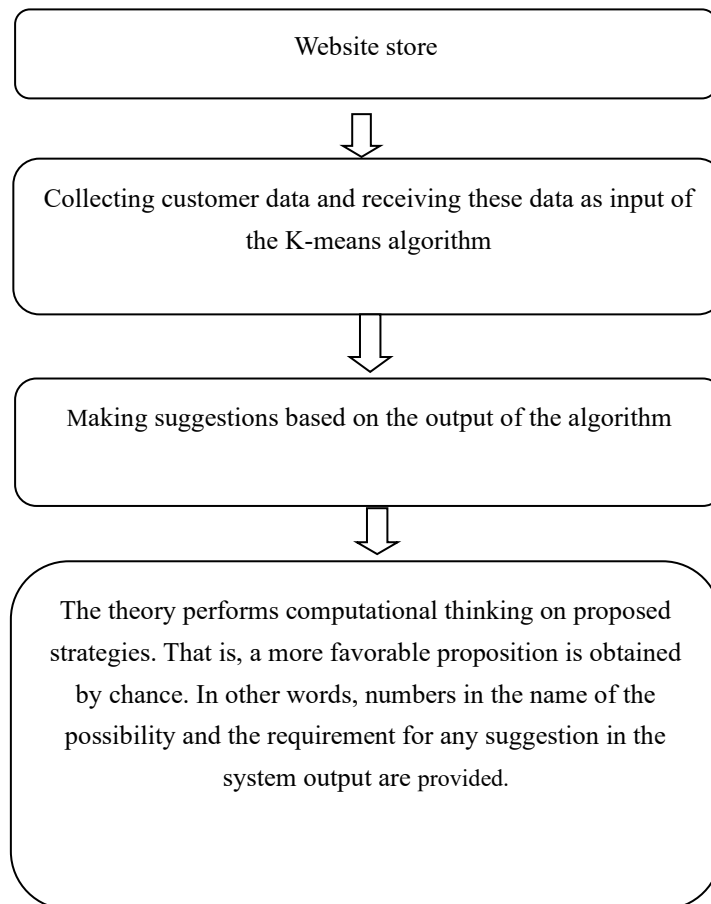
The architecture of our proposed marketing model consists of several steps as shown in Figure 1:

Step 1: Data Collection: Customer data is collected after customers have made their online shopping.

Step 2: Classify Customers: The K-Mean algorithm receives customers' data as inputs, and classifies customers.

Step 3: The marketer provides advice and recommendations.

Step 4: the numbers in the name of the possibility and the requirement for any recommendation in the system output are provided.



**Figure 1.** Conceptual marketing model structure

## 3. Implementation of the Proposed Marketing Model

In this section, we will implement the system for an online store. From the virtual stores, we consider the DigiKala store. DigiKala store is the largest online store in the country, with a range of diverse products for all segments of society and in every age group. The range of goods of this site is in the range of the following goods: digital goods, home appliances, beauty and health, culture and art, sports and entertainment. The steps of

implementing the marketing Model with the K-Mean Algorithm:

**Step 1: Get the necessary information from actual customers:**

The customer purchases his/her product from the DigiKala website, the following information is received from the customer: email address, mobile phone number, the code of product, date of birth, gender, the place of residence, purchased goods. Sample data collected from customers from DigiKala site is as follows (see Table 2):

**Table 2.** Sample data collected from customers on the DigiKala website after the purchase process

Purchased item	Email	Phone number	Location	Age	Gender
Memory card	X1@gmail.com	PhoneNumber1	Mazandaran	27	Male
Mobile phone	X2@gmail.com	PhoneNumber2	Mazandaran	25	Male
Mobile phone	X3@gmail.com	PhoneNumber3	Tehran	26	Male
Shoes	X4@gmail.com	PhoneNumber4	Tehran	30	Female
Electronic mixer	X5@gmail.com	PhoneNumber5	Mashhad	30	Female
Perfume	X6@gmail.com	PhoneNumber6	Tehran	30	Female
Laptop & mobile phone	X7@gmail.com	PhoneNumber7	Mazandaran	30	Male
Hard	X8@gamil.com	PhoneNumber8	Ahvaz	27	Male
Headphone	X9@gmail.com	PhoneNumber9	Semnan	20	Male

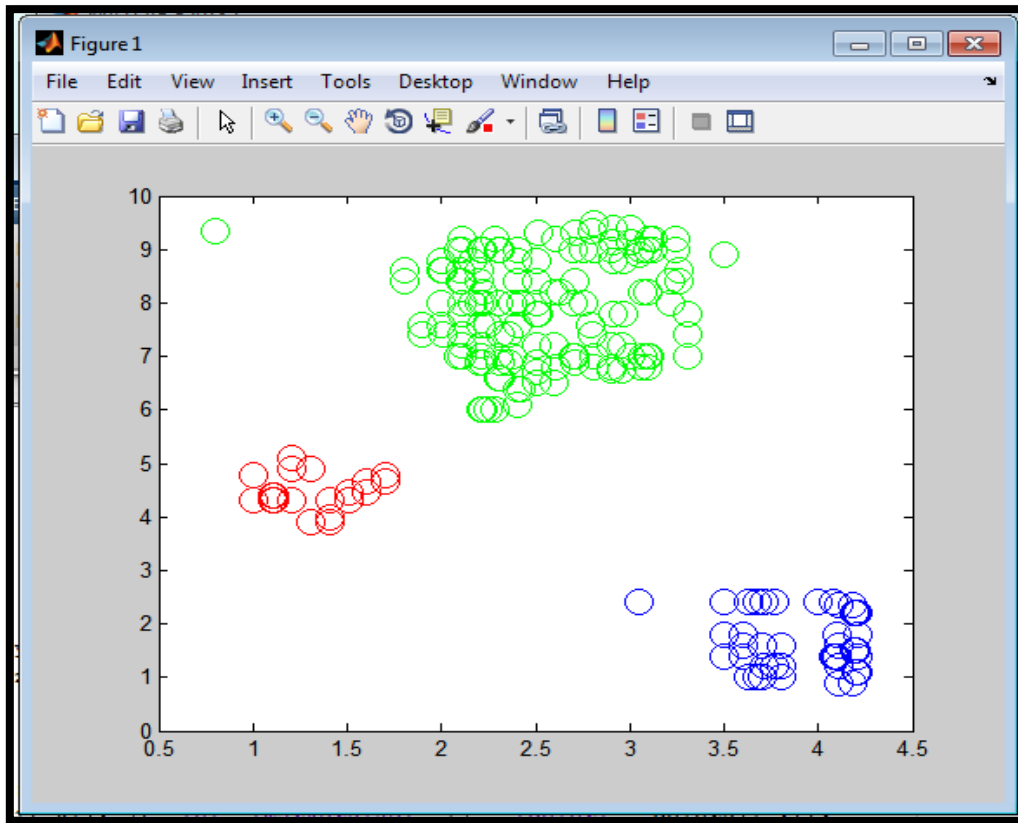
**Step 2: Data clustering using the K-Means algorithm:**

At this stage, the data are categorized using the K-Means clustering algorithm. Output of the K-Means algorithm for the two characteristics of the age and goods (see Figure 2 and Figure 3):

```

1 - clc;
2 - clear;
3 - close all;
4
5 - data=load('mydata13');
6 - y=data.y;
7
8 - k=3;
9 - z=linkage(y);
10 - I=cluster(z, 'maxclust' , k);
11
12 - colors=hsv(k);
13 - for i=1:k
14 -     xi=y(I==i,:);
15 -     plot(xi(:,1),xi(:,2), 'o', 'markersize',12 , 'color' , colors(i,:));
16 -     hold on;
17 - end
    
```

**Figure 2.** The k-Means algorithm used for the two characteristics of age and goods



**Figure 3.** Data clustering based on K-Means algorithm

**Suggested strategy for K-Means algorithm output:**

Vertical Line: desired product: Household appliances are in range from 1 to 3. Cultural and sports goods are in range from 3 to 6. Digital goods are in range from 6 to 10.

Horizontal Line: Customer's age range: The customer's age is divided into 10.

Number of clusters:  $k = 3$ .

**For more detail information, see Table 3.**

**Table 3.** The proposed strategy for the two characteristics of the age and goods

Cluster	Proposed strategy
Green Cluster	This cluster, which is the largest market segment, includes 318 out of 456 customers, accounting approximately 69.7%. This segment of the market has the most sales of digital goods compared to other market segments, which range of age is from 18 to 38. Therefore, it is suggested to the marketer that, if new products enter the market, they will start advertising from people aged 18 to 38 years old. Also, because of the large volume of buyers in this group, are used both E-mail and SMS advertising. For this category of customers, special discounts will be considered for future purchases.
Blue Cluster	This cluster consists 96 out of 456 customers, which is 21.05% of the sample size. The most popular home appliances are in the age range of 37 to 45 years. Therefore, it is suggested to the marketer that if new goods arrive at a better quality in the market place, they be sent to customers in the age range of 37 to 45 years and are used SMS advertising. Also, for this category of customers, in the future purchases, will take place free delivery goods.
Red Cluster	This cluster, which is the smallest market segment, includes 44 out of 456 customers, accounting for 9.25% of the sample size. This part of the market includes the most game of software and magazines and sports shoes that age range from 11 to 18 years. Therefore, it is suggested to the marketer that if new game packs and magazines and sports shoes are introduced, they be sent to customers in this age range with SMS ads. Also, for this bunch of customers, the guaranty of product will be included in future purchases.

Output of K-Means algorithm for gender and goods are shown in Figure 4 and Table 4.

The image shows a MATLAB Editor window titled "Editor - C:\jj\app2.m". The window contains a script for performing K-Means clustering. The script is as follows:

```
1 - clc;
2 - clear;
3 - close all;
4
5 - data=load('mydata18');
6 - y=data.y;
7
8 - k=3;
9 - z=linkage(y);
10 - I=cluster(z, 'maxclust', k);
11
12 - colors=hsb(k);
13 - for i=1:k
14 -     xi=y(I==i,:);
15 -     plot(xi(:,2),xi(:,3), 'o', 'markersize',12, 'color', colors(i,:));
16 -     hold on;
17 - end
```

The status bar at the bottom right indicates "Ln 15 Col 12".

Figure 4. K-Means algorithm used

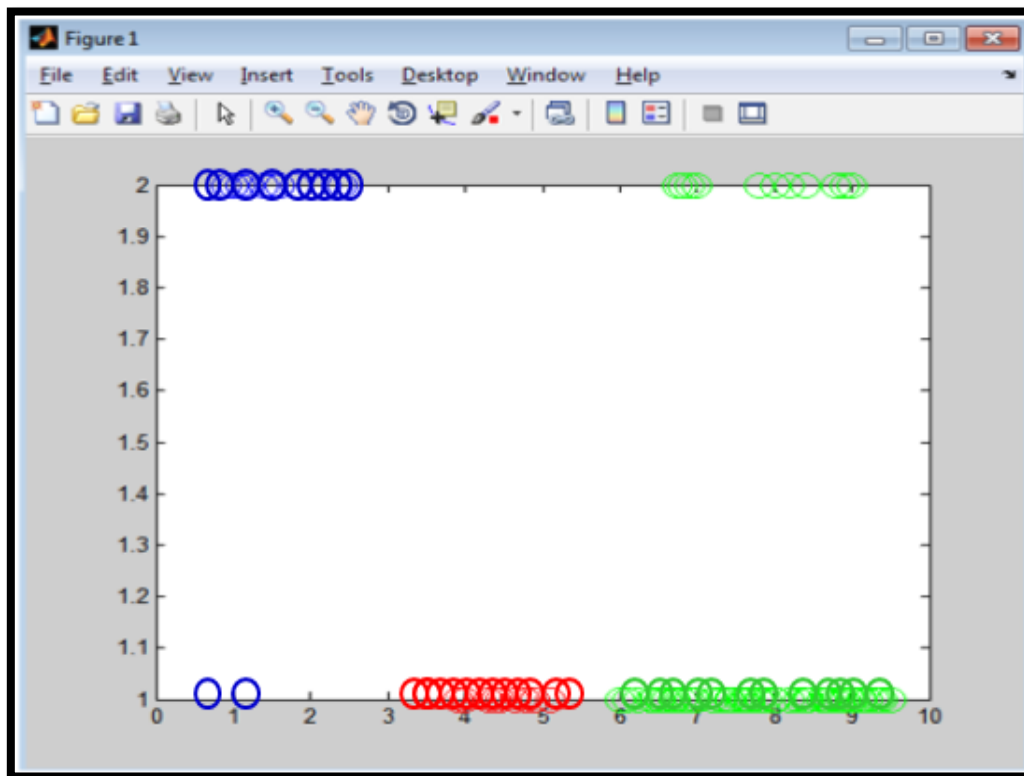


Figure 5. Data clustering based on K-Means algorithm

**Suggested strategy for K-Means algorithm output:**

Vertical Line: desired product: Household appliances are in range from 1 to 3. Cultural and sports goods are in range from 3 to 6. Digital goods are in range from 6 to 10.

Horizontal Line: Customer gender: male (1), female (2).

Number of clusters:  $k = 3$

**For more detail information, see Table 4.**

**Table 4.** The proposed strategy for gender and the goods

Cluster	Proposed strategy
Green Cluster	This cluster, which is the largest market segment, includes 238 out of 318 consumers who are purchasers of digital goods, with 75% of the sample size. This segment of the market has the most purchases of digital goods compared to other parts of the market, which is also done by men. Therefore, it is suggested to the marketer that, if new goods are introduced to the market, they start their advertising to men and consider discount for future purchases. Also, because of the large volume of buyers in this group, they use both email and SMS advertising for this group.
Blue Cluster	The cluster consists of 96 customers that 93 out of 96 are women, with more than 95% of household appliances being sold by women. Therefore, it is suggested to the marketer that, if the new product arrives in the market, it will be sent to this category by SMS advertising, and also for this category of customers consider the free delivery of goods in future purchases.
Red Cluster	The cluster, which is the smallest market segment, includes 44 customers which 39 are male, representing 88.6% of the sample size. This part of the market includes the most popular packages of software game and magazines and sports shoes. Therefore, it is suggested to the marketer that if the new product is introduced to the market, they will insert on their personal page on the DigiKala website and also consider the goods guaranty for this category of customers in future purchases.

From the two tables above one can extract the strategy suggested in Table 5:

**Table 5.** The proposed strategy based on three characteristics: Age and gender and goods purchased from the website

Suggested strategy
Most of the digital sales are in the range of 18 to 38 years old, and most of the shoppers are male. Therefore, it is suggested to the marketer that if new products are introduced into the market, they start advertising male at the range age 18 - 38, and for this category of customers will be considered special discounts in future.
Most of the home appliance sales are in the age range of 37 to 45 years, and most of the buyers are women. Therefore, it is suggested to the marketer that advertisements should be sent to female customers in this age range, and also for this group timely delivery of goods in the future purchases of free delivery.
Approximately the largest purchase of software packages and sports magazines and sports shoes are done by men at the age range of 11 to 18 years. Therefore, it is suggested to the marketer that, if new goods are introduced to the market, will be sent SMS in this age range for men and also will be considered guaranty for future purchases.

**Step 3: Using theory of possibility:**

At this stage, data analysis is done using the theory of Possibility. In this way, after classifying and analyzing information, we must examine the degree of probability and necessity of the event. The following table, based on the latest reports and customer information from the DigiKala website in the last 6 months, estimates the average amount of purchases of goods based on customers' age ranges as possible in Table 6.

**Table 6.** Possible for customers' purchased items over the period of 6 months

		A	B	C
Age of customers per year	Possibility	18-11	38-18	45-38
Digital goods	$\prod_1$	0.3	1	0.09
Artistic goods	$\prod_2$	0.06	0.02	0.01
Household appliances	$\prod_3$	0	0.04	0.19

Now we would like to answer the following questions:

What is the chance for customers to buy digital goods from 11 to 38 years old? What is the Possibility and necessity of this event?

Customers are between the age range of 11 and 38: customers are from 11 to 18 and 18 to 38 years old, is equivalent to the following:

$$\prod_1(A \cup B) = \max[\prod_1(A) \cup \prod_2(B)] = \max[0.3, 1] = 1 \quad (1)$$

The Necessity of the above-mentioned event is calculated according to the following formula:

$$N(A) = 1 - \Pi(A^c) \quad (2)$$

$$N_1(A \cup B) = 1 - \prod_1[(A \cup B)^c] = 1 - \prod_1[C] = 1 - 0.01 = 0.99 \quad (3)$$

Therefore, based on the necessity and possibility of purchasing digital goods at the age range of 11 to 38 years, we can conclude about 99% of digital sales are in this age range, so new entering digital goods will be sent to people in this age range either by email or SMS. The purpose of this paper is to design a marketing model that covers the disadvantages of past models. In order to improve the marketing planning process, a marketing model has been provided that benefits from the previous models and covers its disadvantages. In this proposed model, the method of computer model marketing and group decision making have been used in combination.

#### 4. Conclusions

In this paper, an electronic marketing system was proposed by using the recommendation system and the possibility theory based on the customer's random behavior. From the perspective of technology, the proposed marketing system has been designed to provide the necessary synchronization to operate with existing technologies. The purpose of this system is the development of e-mail marketing, which is needed in today's market, customer analysis and reporting marketers for marketing. This marketing system will create a competition that will be effective in the marketplace. In this system, the marketer can find out, at any time, the number of sales of different types of goods in different areas, based on the gender and age of the customers and provide marketing recommendations. Then, using the theory, it is possible to examine the possibility and necessity of recommendations and to assure the marketer of the recommendations.

#### Data Availability

The data used to support the research findings are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### References

- [1] Y. Zhen, S. Yain-Whar, Z. Defu, X. X. Zeng, S. C. H. Leung, and T. Li, "A decision-making framework for precision marketing," *Expert Syst. Appl.*, vol. 42, no. 4, pp. 3357-3367, 2015. <https://doi.org/10.1016/j.eswa.2014.12.022>.
- [2] L. Y. Chen and T. C. Wang, "Optimizing partners' choice in IS/IT outsourcing projects: The strategic decision of fuzzy VIKOR," *Int J. Prod. Econ.*, vol. 120, no. 1, pp. 233-242, 2009. <https://doi.org/10.1016/j.ijpe.2008.07.022>.
- [3] R. F. Saen, "Developing a new data envelopment analysis methodology for supplier selection in the presence of both undesirable outputs and imprecise data," *Int J. Adv Manuf Tech.*, vol. 51, no. 9-12, pp. 1243-1250, 2010. <http://dx.doi.org/10.1007/s00170-010-2694-3>.
- [4] B. M. Hsu, C. Y. Chiang, and M. H. Shu, "Supplier selection using fuzzy quality data and their applications to touch screen," *Expert Syst. Appl.*, vol. 37, no. 9, pp. 6192-6200, 2010. <https://doi.org/10.1016/j.eswa.2010.02.106>.
- [5] B. Chang and H. F. Hung, "A study of using RST to create the supplier selection model and decision-making rules," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8284-8295, 2010. <https://doi.org/10.1016/j.eswa.2010.05.056>.



- [6] J. Y. Chai, J. N. K. Liu, and E. W. T. Ngai, "Application of decision-making techniques in supplier selection: A systematic review of literature," *Expert Syst. Appl.*, vol. 40, no. 10, pp. 3872-3885, 2013. <https://doi.org/10.1016/j.eswa.2012.12.040>.
- [7] X. Liu, "Agricultural products intelligent marketing technology innovation in big data era," *Procedia Comput. Sci.*, vol. 183, pp. 648-654, 2021. <https://doi.org/10.1016/j.procs.2021.02.110>.
- [8] S. G. Li, X. W. Song, H. Y. Lu, L. Y. Zeng, M. J. Shi, and F. Liu, "Friend recommendation for cross marketing in online brand community based on intelligent attention allocation link prediction algorithm," *Expert Syst. Appl.*, vol. 139, Article ID: 112839, 2020. <https://doi.org/10.1016/j.eswa.2019.112839>.
- [9] F. J. Martínez-López and J. Casillas, "Marketing intelligent systems for consumer behaviour modelling by a descriptive induction approach based on Genetic Fuzzy Systems," *Ind Market Manag.*, vol. 38, no. 7, pp. 714-731, 2009. <https://doi.org/10.1016/j.indmarman.2008.02.003>.
- [10] S. Khan, S. Tomar, M. Fatima, and M. Khan, "Impact of artificial intelligent and industry 4.0 based products on consumer behaviour characteristics: A meta-analysis-based review," *Sustain Oper. Comput.*, vol. 3, pp. 218-225, 2022. <https://doi.org/10.1016/j.susoc.2022.01.009>.
- [11] M. Moradi and M. Dass, "Applications of artificial intelligence in B2B marketing: Challenges and future directions," *Ind Market Manag.*, vol. 107, pp. 300-314, 2022. <https://doi.org/10.1016/j.indmarman.2022.10.016>.
- [12] B. Yang, Y. Liu, and C. Chen, "A twin data-driven approach for user-experience based design innovation," *Int J. Inform. Manage.*, vol. 68, Article ID: 102595, 2023. <https://doi.org/10.1016/j.ijinfomgt.2022.102595>.
- [13] H. Rahim, H. Irpan, and M. S. A. Rasool, "Consumers Attitude toward Halal Food Products in Malaysia: Empirical Evidence from Malaysian Millennial Muslims," *Asian Proc. Soc Sci.*, vol. 9, no. 1, 2022. <https://doi.org/10.31580/apss.v9i1.2252>.
- [14] L. Bezuhla, I. Koshkalda, I. Perevozova, S. Kasian, and N. Hrechanyk, "Marketing management of the regional ecotourism infrastructure," *Int J. Ind Eng., Prod. Res.*, vol. 33, no. 1, pp. 1-11, 2022. <http://dx.doi.org/10.22068/ijiepr.33.1.3>.
- [15] Y. Hosseini, H. Fazlollahtabar, and M. T. Ashoori, "Outsourcing marketing plans for small and medium enterprises using knowledge sharing process case study: Tehran wood and furniture companies," *Int J. Ind Eng., Prod. Res.*, vol. 32, no. 2, pp. 1-18, 2021. <https://doi.org/10.22068/IJIEPR.1161>.
- [16] A. Eshghi and M. Kargari, "Detecting frauds using customer behavior trend analysis and known scenarios," *Int J. Ind Eng., Prod. Res.*, vol. 29, no. 1, pp. 91-101, 2018. <https://doi.org/10.22068/IJIEPR.29.1.91>.