



The impact of digital marketing strategies on customer's buying behavior in online shopping using the rough set theory

Ebrahim Forghani¹ · Reza Sheikh¹ · Seyed Mohammad Hassan Hosseini¹ ·
Shib Sankar Sana²

Received: 2 April 2021 / Revised: 24 May 2021 / Accepted: 24 August 2021

© The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2021

Abstract Digital Marketing strategies are sets of controllable e-marketing variables that organizations combine to achieve marketing goals and to meet customers' needs. These strategies are the most important factors that electronic-marketing managers pay attention to the best strategy in order to achieve sales and profitability. This study aims to investigate the effect of these strategies on the buying behavior of customers in online shopping stores in Tehran. For this purpose, five best-selling online stores in Tehran are selected and 79 samples are taken from each of them. For data collection, a 2-tuple fuzzy linguistic representation model is used in order to no lose the linguistic information obtained from customers. For data analyzing and extracting proper rules, two approaches of the rough set theory are used. Based on the results provided by Rosetta software, five rules governing customer behavior are identified as the most important factors affecting buying behavior in online shopping. To evaluate the result, a comparison is carried out between the extracted rules using the proposed rough set

technique and the tree diagram of the data obtained by Rapidminer software. Almost all provided rules are confirmed through this comparison along with the opinions of experts. Some of key results according to the obtained rules indicate that the most important digital marketing strategy is the search engine optimization. Moreover, the social media marketing and recommender engine play as second important issue of the marketing management.

Keywords Digital marketing · Rough set theory · 2-tuple fuzzy · Customer behavior

1 Introduction

In nowadays competitive environment, the ultimate goal of commercial companies is to increase their sales and market share. In order to sell more products, brands need to find the most successful marketing strategies. Marketing strategies are policies that organizations and marketers use to provide customers' expectations with appropriate products and services at low-cost. An efficient advertising strategy and interacting with distribution outlets play a vital role to increase demands for the company products and rising efficiency and profit. The reason behind it that marketing strategy is well known as a continuous process which must direct organizations where they want to business in the long term. So, a proper marketing strategy helps organization to manage the direction, and interaction among the marketing elements and the environmental elements in a specific situation.

On the other hand, revolution of internet and its expansion has led to a deep change of marketing both for organizations and the daily life of people. Due to the advent of

✉ Shib Sankar Sana
shib_sankar@yahoo.com

Ebrahim Forghani
forghani.ebrahm@gmail.com

Reza Sheikh
resheikh@shahroodut.ac.ir

Seyed Mohammad Hassan Hosseini
sh.hosseini@shahroodut.ac.ir

¹ Department of Industrial Engineering and Management, Shahrood University of Technology, Shahrood, Iran

² Department of Mathematics, Kishore Bharati Bhagini Nivedita College, Ramkrishna Sarani, Behala, Kolkata 700060, India

internet, moving towards electronic environment is inevitable for organizations. They can reach global marketplace just a mouse click away. Progress in technology have an effect on marketing especially online marketing that is based on web and internet. Nowadays, in Digital era, electronic media is helping organization in digital marketing for business and marketing. Organizations are trying to invest in digital marketing as well as offline marketing. Thus Digital marketing with exploring new facilities and technologies try to make new customers for organizations (Taherdoost, H. & Jalaliyoon, 2014; Sett et al. 2020; Sarkar et al. 2019).

According to the key role of digital marketing in successful internet-based sales and also companies' attempts to optimize their profit, they companies should consider different types of strategies in the modern competitive market (Sett et al. 2020). This study aims to provide relevant rules that demonstrate the effect of digital marketing strategies on customers' buying behavior in online shopping. In this way, we first determine the most important aspects of digital marketing. Then, the data collection is done. Afterwards, appropriate rules would be extracted using the rough set theory. Provided rules can accentuate the most determinative digital marketing indicators affecting customers' buying behavior in online buying.

Research findings would help online store managers to figure out the weaknesses and concentrate their efforts to increase their sales and market share. Moreover, when the managers have an accurate understanding of the factors affecting the buying behavior of customers, they can better benefit from the positive results and increase the profit and income of these centers as well as their competitiveness. All customers are somehow influenced by various elements of the digital marketing and their decisions are highly dependent on these elements. Also in the case of sales stores, the right use of digital marketing will be the condition of success and sales guarantee. In this research, 2-tuple fuzzy linguistic model is used. This Symbolic approach prevents losing some information in the approximation process. In the second part, this approach is fully explained and its mathematical relations are described.

The outline of the paper is as follows. Section 2 is devoted to present Preliminaries and the survey of studies related to this article. This section is categorized as digital marketing, customer behavior, rough set theory, decision tree, 2-Tuple Linguistic Term Sets, Picture Fuzzy Sets (PFSs), and Picture 2-Tuple Linguistic Sets (P2TLSs). The research methodology is described in Sect. 3. Section 4 presents the result of study and result analysis. And finally conclusions and suggestions of the research are provided in Sect. 5.

2 Preliminaries

2.1 Digital marketing

Digital marketing concept emanated from the internet and search engines but it is now more than just term internet marketing or E-marketing. It is known as a modern marketing activity which includes the use of different web-based media such as emails, web sites, blogs, or social networking. After bursting internet bubble, search engine optimization has been great grown for some companies. (Malik, 2017). Gopher was the first search engine started with a network protocol for enquiry and search in 1991. Thereafter, companies' goal was to achieve the maximum ranking on the web since the beginning of yahoo in 1994.

Rapid advent and expansion of information and communication technologies (ICTs) and digital media have been a considerable effect on people's communication and provide their socioeconomic and emotional needs (Dwivedi et al. 2020). The utilization of cellphone devices drastically growth the usage of internet and people around the world started connecting with each other more easily via social media. Already many people used digital media for a range of activities such as connecting with family and friends, services and places, reviewing products, selling and buying goods, etc. (Kulkarni et al. 2020).

With the rise of internet users, customers' buying procedure changed. They started searching and choosing products online instead of consulting salesperson, which make serious barriers for companies. These issues led companies and marketing managers found out the digital solution to develop their market. Digital's improvement changed marketers and businesses' method to use new technology in marketing. Digital platforms are increasingly incorporated into marketing and everyday life. Consequently, digital marketing was bring up in business when people used digital devices to meet their needs instead of physical visiting shops (Desai, 2019). Generally the methodologies and notions which refers to digital marketing would seem emergent in twentieth century specially in the last decade because the spread of internet and sales via electronic commerce would seem consolidated at that time (Saura et al. 2019). Digital world' characteristics and its use for marketing have promoted the increase of channels, forms and languages leading to strategies and devices which are in offline unthinkable. Nowadays Digital marketing; instead of just a topic of conventional marketing, has changed into a new phenomenon bringing customization and mass distribution together. To make the aims of marketing, the convergence of technology and the use of device have opened up the ways for man to think about internet marketing. They have developed a new notion of digital

marketing highlighting the feature of a user-centered, ubiquitous, more measurable and interactive activity (Piñeiro-Otero & Martínez-Rolán, 2016).

Digital marketing has become a comprehensive phrase which describe the deployment of digital technologies all over the whole marketing process for get and hold customers, building good image of organizations, and leading customer relationships. The development of digital marketing through the web and mobile phones, offers new access for company to advertising that is very influential (Sawhani & Susilo, 2020). Desai (2019) defined digital marketing as all activities which used internet or electronic apparatus. In this definition, businesses use digital channels, email, social media and all digital apparatus to help them connect with customers. This can also be named as 'web marketing', 'internet marketing' or 'online marketing'. It also can be defined as numerous digital strategies and ways to connect with people where they spend most of their time online. website, digital advertising, email marketing are widespread strategies get under the digital marketing term (Desai, 2019). Kannan and Li (2017) in their comprehensive view defined digital marketing as technology-based adaptive and process by companies to creating collaboration with all partners to achieve value for all stakeholders.

The digital marketing ensures direct interaction between companies and customers. Digital platforms make adequate opportunities for companies interact with their customers and receive immediate feedback on how the customers feel about the products or services offered to improve their performance and thus increases market segment and profitability. Application of digital media has a larger audience compared to traditional media because it uses the various social media platforms to advertise services and products. Therefore, reaching a wide audience compared to other forms of media like television, a large number of the customers increase the demand of the products and services offered by the companies (Shah, 2018).

Digital marketing by creating platforms that allows the users to generate and share content changed the structure of marketing. Pages that open and sites created on virtual environment are not belong to special group. Rather they have become channels that address a wider user group. In other words, digital marketing can enable to organizations and individuals at different locations around the world to connect with each other in a faster way being interactive and digital are two essential features of digital marketing. Being interactive can be defined as giving information to potential customers and responding to questions and communicating with the customer by involving them in a two-way process. Being digital refer to quality of the new environment where marketing takes place matters (Cizmeci & Ercan, 2015).

Different other benefits can be obtained by using digital marketing such as ease of use, potential for increasing knowledge for customers (Tiago & Veríssimo, 2014), to make a great chance for building a brand image on the internet, presence and constant updates (Piñeiro-Otero & Martínez-Rolán, 2016). Furthermore, firms' access affordable promotion of in international markets, strengthen firms' relationships due to the ease of communication with customers, suppliers, or partners are the ways in which digital marketing makes to bossiness to access customers around the world. Also it can rises the capacity to control, optimize, and improve campaigns activities by preparing high flexibility and dynamism, enabling customized and precise segmentation (Carpio et al. 2020). Although digital marketing has various different advantages, some constraints and disadvantages can be found in the existing related literature. A summary of the digital marketing's advantages and disadvantages can be represented based on the related studies (GirchenkoTetiana & Ovsiannikova, 2016; Karimi & Naghibi, 2015; TODOR, 2016; Veleva & Tsvetanova, 2020) as follows:

- *Facilitation of communication and interplay with customers* Due to extant various digital systems such as web applications, websites, and social networks, customers can connect to each other and share their experiences about goods. On other hand companies and marketers can talk straightly with their potential customers to build trust and get feedback about the goods and services they offer.
- *Cost efficiency* Digital marketing compared to traditional marketing has much lower cost and in many cases websites can generate traffic even for free.
- *Personalization* In digital marketing, brands can offer programs that can be personalized based on customer behavior and their preferences
- *Unbounded time* Digital marketing offers convenient method of doing business transactions (not limited to specific hours). That is so because of absence of relation with tangible mediums.
- *Assessments* All performances, operations, and factors effects can be measured easily in a digital marketing

2.1.1 Disadvantages of digital marketing

- *Dependence on technology and technics* Digital marketing requires strong knowledge but this knowledge can make different technical faults because information and communication apparatuses used by it are not without errors.

- *Time consuming to get start* It's very time consuming for businesses, and it needs a perfect and experienced team to work at most of the time to make content that can react to customers' complaints.
- *Lack of customer trust* Multiplicity of cheats on online advertising has led to lack of customers' trust.
- Digital marketing strategies take some time to obtain measurable success.
- *Negative feedback* In digital platforms, all comments of customers about goods and services are accessible for all. So negative reactions and comments of some customers can lead to destroyed company's brand and customers outflow.

There are different strategies in digital marketing for organizations to use. Table 1 shows a summary of these strategies with their definition. Deployment of these strategies lead companies to have a set of capabilities and advantages that can be used for their improvement. Bianchi and Mathews (2016) referred to some key aspects of digital marketing including online sales, etc.

2.2 Customer behavior

Customer' behavior can help marketing managers to figure out how customers think, feel, and choose different products and brands, as well as how the environment influences their behavior. Customer buying behavior is considered according to various factors including cultural, social, personal and psychological factors. Most of these

Table 1 Digital marketing strategies

Strategy	Description
Search Engine Optimization (SEO)	SEO is known as all activities to achieve high visibility and improving rating of website in search engine outcome page. Generally, high rated sites on Search Engine Result Pages (SERP) appears mostly in search outcome list and take more traffic from search engine user. This traffic of web can convert visitors into potential customers (Terrance et al. 2018). Natalia and Liashenko proposed some following offer For effective SEO optimization: optimization of design of web-site structure, optimization of the search engine, collection and analysis of statistics of visitors' transfers to the website from the search engine, arrangement of the promotion strategy based on the analysis of the prepared statistics and its dynamics (Solodka & Liashenko, 2020)
Search Engine Marketing (SEM)	Companies use SEM in business as a general strategy to drive web traffic to their business. Some authors called it Paid Search Marketing. According to organization' structure, companies can choose some models such as pay-per-click (PPC), cost-per-click (CPC), or cost-per-thousand impressions (CPM) model. also Google Ad Words and Bing Ads are the most favorite system which companies can use as SEM (Bala & Verma, 2018)
Social Media Marketing (SMM)	Social media has converted into different connection platforms that people can share their moments from lives and some live in it. Social media is communication device to socialize. They use web-based technology to rapidly spread data and information to a large number of users in web space (Neti, 2015; Shen et al. 2020). SMM refers to social media that is designed to rapid spread data via social interactions. SMM empowers organization to advertise their products and services via online websites or channels and have connection with large community that may via traditional marketing may not have been accessible. Also SMM use some tools like Facebook, LinkedIn, etc. to access the potential customers. (Karimi & Naghibi, 2015)
Content Marketing	Digital content marketing (DCM) refers to creating and publishing of related and worth relevant to brand and business in digital platforms to make good mentality (Hollebeek & Macky, 2019). DCM tended to interplay with users of digital marketing by mobile device and online websites via making valuable video or a live program (Bu et al. 2020). All DCM's definitions highlight three points: (1) developing and distributing information is main purpose of content marketing. (2) Engaging customers online is one of the purposes of digital content marketing (3) the information and content shared must be related and valuable. (Wang & McCarthy, 2020)
Affiliate Marketing	Affiliate marketing is defined as a collection of different business' activities which should create a sustainable composition of strategies and other advertising methods (such as paid search engine marketing, SEO, e-mail marketing) to improve sales and promotion (Ivkovic & Milanov, 2010)
Pay-Per-Click (PPC)	PPC is defined as a strategy to make income for search engines (Kritzinger & Weideman, 2013). In this strategy, web traffic is driven to company' website and users pay publisher every time they clicked in company' ad. Facebook, promoted tweets and google ad are some of typical types of PPC
Recommender Engine (RE)	Recommender Engine (RE) is defined as a system that recommend related and likely item for buy to customers. RE use some platforms such as social media, emails or online buying web portals to service. In this procedure, according to user preferences of items such as movies, e-learning materials, travel destinations, applications, and services and products, used to find an algorithm to make prediction and recommendation (Behera et al. 2020)
Email Marketing	Sending mercantile messages to users via email is defined as email marketing (Bawm & Nath, 2014). Here we refer to some reasons for the popularity of email marketing. (1) email marketing is Very low cost (2) it need less time to prepare and execute (3) users typically respond to email more than other marketing' ways (4) it's easy to communicate with users (X. Zhang et al. 2017)

factors are uncontrollable and marketers do not have access to them; but to understand the complex behavior of customers, it is important to consider them (Stávková et al. 2008). These factors are summarized in the Table 2. Customer Behavior is defined as the behavior that customers meet in their search, acquisition, use, evaluation and use of the goods, services and ideas they expect to meet their needs. Individual or organizational customer is one who acquires goods or services for personal use or for the use of another person. The act of consumption is not only an economic behavior, but also a social behavior. Solomon (2010) defined customer behavior as follows: Customer behavior is the study of the process by which people choose to buy or use service products, ideas, or experiences to meet their needs and wants. This is a general statement that customer behavior refers to the first and most important thing in buying a particular product or service. Customer

behaviors studies in stores usually deal with identifying customers and their buying behavior patterns. The purpose of such studies is to determine who, when, where, how buys a product or service. In addition, such studies seek to gain information about customers' reactions to the organization's sales advertising system. The results of these studies are useful in solving a set of marketing problems and are of particular importance in marketing research (Applebaum, 1951). According to customers' use of the internet and online socializing tool, understanding customer behavior is important and is a key of success of digital marketing (Janathanan & Nizar, 2018). Marketers must know the importance of information derived from others in digital atmospheres. Various information and social elements received through digital marketing in the form of reviews or comments affect in the mind of customers (Sharma et al. 2021). If company can identify factors affecting online

Table 2 Classification of factors affecting customer behavior

Main factors	subcategory	Description
Cultural factors	Culture	Items that refer to the core values, needs and behaviors learned by community members from family and other institutions
	sub-culture	Each culture includes subcultures that include nationality, religion, racial groups, and geographic areas
Social factors	Social class	Members who share common behaviors and interests. Social classes are usually recognized according to the combination of income, education, occupation, and wealth
	Groups	Refers to two or more people interacting to achieve mutual goals. An individual's behavior is determined by small groups or reference groups. These groups can include family, religious groups, circle of friends, and so on
	Family	Family members can shape buying behavior. Marketers look at different roles in the family to shape their buying behavior
	Positions and roles	The position of the individual in each group is determined by his role
Individual factors	Age and life cycle stage	People change their behavior over time. This behavior also affects the buying behavior
	occupation	Occupation affects the purchase of goods and services
	The economic situation	An individual's economic level affects the choice of products. Marketers need to focus on people's income, wealth and interest rates
Psychological factors	Motivation	When the customer feels a need in himself, he subconsciously tries to meet that need. Motivation makes a person take action
	Learning	People are trained when they act, and this training can change their behavior
	Perception	Perception is the process by which people to organization, identification, and interpretation of sensory meaningful picture of the world
Marketing mix	product	The product includes either goods or services that delivers to the customer by creating added value
	Price	It is the amount the customer pays to buy the product
	Place	Includes delivering the product to customers
	promotion	All the operations needed to communicate with the customers and promote the product or service in the target market
	people	All the people who play a constructive role in presenting the product to the customers
	Process	A process is the way in which a product or service is delivered to an end customer and refers to a system that helps the organization deliver the service to the customer
	Physical evidence	This element refers to the environment in which services and goods are presented

customer behavior in digital marketing, it can help to design appropriate technology for this emerging class of customers and according to the size of their cohort, it can become a powerful buying group (Kim & Ammeter, 2018).

2.3 Rough set theory

Rough set theory is using for handling vague, inconsistent and uncertain data and is a new soft computing technique proposed by Z. Pawlak (Pawlak, 1998). Similar to probability theory and fuzzy set theory, rough set theory is another mathematical tool to deal with inaccurate and incomplete data. Rough set and fuzzy set theories are similar methods used to deal inaccurate data. However, in contrast to the concept of partial membership in fuzzy set theory, rough set theory expresses uncertainty, and imprecision by a boundary region of a set. The rough set refers to the process of topological operations, also known as approximations (Xie et al. 2020). Due to novel thinking, ease of employment, and a vast array of fields of application, the rough set theory has made its way to become one of the most practical tools to deal with ambiguity and incomplete data set. Giudice et al. (2017) offered an application of Rough Set Theory (RST) to the real estate field, in order to highlight its operational potentialities for mass appraisal goals. This research showed operational potentialities of RST as a valuation methodology (Del Giudice et al. 2017). Hamoudav et al. (2018) presented a new and robust breast cancer prediction and diagnosis system based on the Rough Set (RS). They introduced the robust classification process based on some new and most effective attributes and incomplete databases (Moheimani et al. 2021).

So, this technique has attracted the attention of many researchers around the world in recent years. It is a strong method in the ground of artificial intelligence, and used as a part of hybrid method in data mining and machine learning (Wei & Liang, 2019). In rough set theory, data tables are analyzed. Data tables in this theory can be obtained by experts or sampling. Typical data can be showed in the form of a decision table with rows including objects and columns including criteria or attributes. RST can used decision tables to extract rules via inductive process (Chiaselotti et al. 2019). We use 4-tuple method as Eq. (1) to represent the rough set theory as information system:

$$S = (U, R = C \cup D, V = \{V_a | a \in R\}, f = \{f_a | a \in R\}) \quad (1)$$

where, U is the universe of discourse with a finite number of objects and a is a finite nonempty set of attributes. The subset C is called condition set and the subset D is called decision set (Das et al. 2018). In addition, V is the nonempty set of the values where V_a represents the possible

values of the attribute a . Moreover, f is an information function, and $f_a(x)$ represents the value of the objects $x \in U$ on the attribute a . The information system can be simply written as $S = (U, R, V, f)$ (Jia et al. 2020).

In next, some basic concepts of the rough set theory, such as indiscernible relation, approximation that involved in this paper will be reviewed. In rough set theory, for a subset of attributes $A \subseteq R$, the indiscernible relation denoted by $IND(A)$ and defined in Eq. (2):

$$IND(B) = \{((x, y) | (x, y) \in U^2, \forall a \in A (f_a(x) = f_a(y)))\} \quad (2)$$

Obviously, $IND(A)$ is an equivalence relation, which is reflexive, symmetric and transitive. The family of all equivalence classes of $IND(A)$ will be denoted by $U/IND(A)$, or simply U/B ; an equivalence class of $IND(A)$ containing x will be denoted by $[x]_{IND(A)}$ or simply $[x]$ (Han et al. 1999). By considering the relationship of equivalence classes, the lower approximation and the upper approximation of X with respect to the set of attributes A are defined. The scope that certainly belong to the desired collection in defined as low approximation and the scope that can't certainly belong to the desired collection is defined as upper approximation. Equations (3) and (4) showed lower and upper approximations (Sajjadian et al. 2018):

$$\underline{apr}(X) = \{x \in U | [x] \subseteq X\} \quad (3)$$

$$\overline{apr}(X) = \{x \in U | [x] \cap X \neq \emptyset\} \quad (4)$$

The difference between upper and lower approximations which is shown in Fig. (1) defined as Boundary region approximations. The universe U can be partitioned into three disjoint area, the Boundary, positive, and negative area which are provided in Eq. (5) to (7) (Yao, 2008).

$$BND(X) = \overline{apr}(x) - \underline{apr}(x) \quad (5)$$

$$POS(X) = \underline{apr}(x) \quad (6)$$

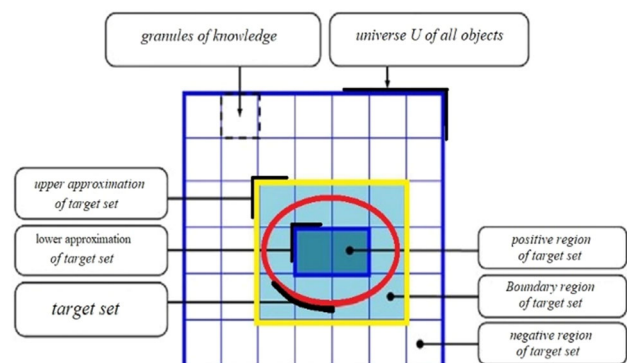


Fig. 1 Upper and lower approximations (Othman et al. 2009)

$$\text{NEG}(X) = U - \overline{\text{apr}}(x) \tag{7}$$

If an object $x \in \text{POS}(x)$, then it certainly belongs to target set X . If an object $x \in \text{BND}(x)$, then it does not certainly belong to target set X . If an object $x \in \text{NEG}(x)$, then it cannot be deduced whether the object x belongs to target set X or not. If $X \sqsubseteq U$ be an optional target set of events comprising a single class of a particular status, Fig. 2 shows schematics diagram of the concepts of rough set theory such as upper and lower approximations, positive, negative and boundary regions.

There are several factors by which the sufficient and necessity of the extracted rules could be identified. In the following, these factors are briefly introduced.

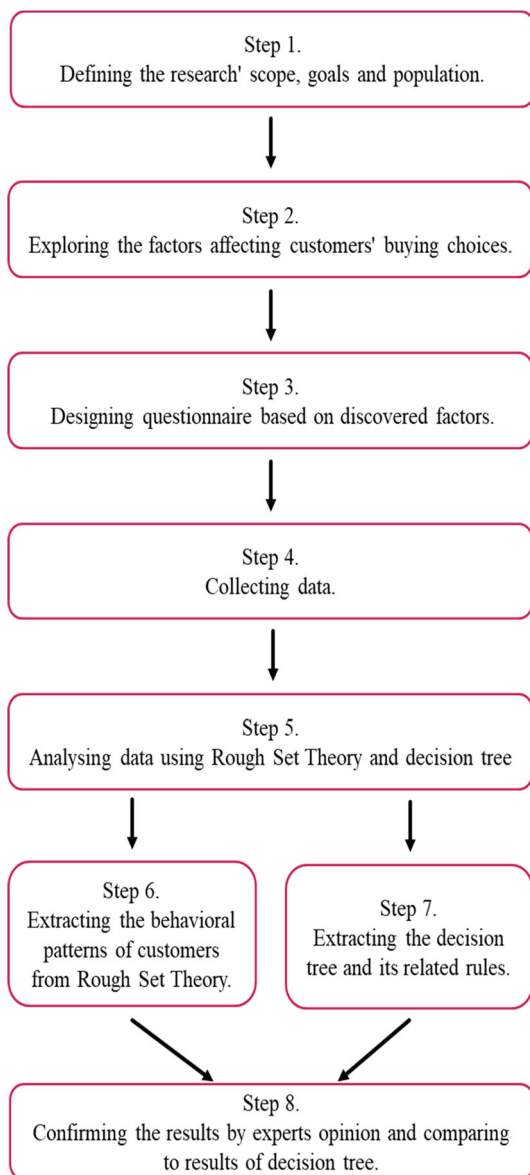


Fig. 2 Process of the proposed research

Support factor is the number of instances which have been covered by a decision rule and is specified in Eq. (8).

$$\text{Support}_x(C, D) = |C(x) \cap D(x)| \tag{8}$$

Certainty factor is the conditional probability for a decision rule given the conditional patterns and is defined as Eq. (9).

$$\text{Cer}_x(C, D) = \frac{\text{Support}_x(C, D)}{|[X]_C|} = \frac{\sigma_x(C, D)}{\pi([X]_C)} \tag{9}$$

$$\pi([X]_C) = \frac{|[X]_C|}{|U|} \tag{10}$$

Based on the Eq. (10), the certainty factor could be considered as the conditional probability of y belonging to $[X]_D$ given y belongs to $[X]_C$.

Strength factor is the probability for a decision rule to occur and is defined in Eq. (11).

$$\text{Strength}_x(C, D) = \sigma_x(C, D) = \frac{\text{Support}_x(C, D)}{|U|} \tag{11}$$

Coverage factor is the inverse probability and is specified in Eq. (12).

$$\text{Cov}_x(C, D) = \frac{\text{Support}_x(C, D)}{|[X]_C|} = \frac{\sigma_x(C, D)}{\pi([X]_D)} \tag{12}$$

$$\pi([X]_D) = \frac{|[X]_D|}{|U|} \tag{13}$$

In fact, the coverage factor is the certainty factor of the inverse decision rule and can also be computed by dividing the support of a decision rule by the number of objects that match the decision rule consequence.

Accuracy factor of a decision rule can be calculated as its support divided by the number of objects that match its rule antecedent. In other words, accuracy is simply the fraction of test observations classified to the correct class. The two factors coverage and accuracy are used more commonly and are defined based on the two relations RHS and LHS (i.e. the right-hand side and the left-hand side of the rules).

- *LHS Support*: Number of observations matching the IF-part.
- *RHS Support*: Number of observations matching the IF-part and the THEN-part.
- *RHS Accuracy* : $\frac{\text{RHS support}}{\text{LHS support}}$ (Accuracy is 1.0 unless the THEN-part contains several decisions)
- *LHS Coverage* : $\frac{\text{LHS support}}{\text{Number of all observation in the set}}$
- *RHS Coverage* : $\frac{\text{RHS support}}{\text{Number of all observation in the set}}$

- *LHS Length*: Number of attributes in the IF-part of the rule.
- *RHS Length*: Number of attributes in the THEN-part of the rule.

2.4 Decision tree

Data mining defined as a method of discovering hidden patterns in huge data to process of decision making. Classification is one of decision making techniques which use to extract pattern and also Classification algorithms are too many. One of important these techniques is decision tree which by breaking down data cause to build a tree structure (Fakir et al. 2020). In decision tree some remarked concepts are: node, branch, leaf and pruning. Node refer to an attribute, branch refer to value of the attribute and finally the leaf refers to class. Scientists use pruning to optimize decision tree and abridge the extracted rules which typically has two ways: pre pruning and post-pruning (Qowidho et al. 2019). It can be extracted a set of rules from each branch in the form of IF–THEN (Fitrani et al. 2019).

2.5 2-Tuple linguistic term sets

By means of the following elements, the linguistic information can be modelled in the 2-tuple fuzzy linguistic representation (Martínez et al. 2015):

Suppose that $s_i \in S = \{S_0, \dots, S_g\}$ is a linguistic term like the fuzzy in which the grammar is chosen regarding the choices provided by the fuzzy linguistic approach and semantics is presented by a fuzzy membership function. For example, S can be defined as:

$S = \{s_0 = \text{extremely poor}, s_1 = \text{very poor}, s_2 = \text{poor}, s_3 = \text{medium}, s_4 = \text{good}, s_5 = \text{very good}, s_6 = \text{extremely good}\}$.

The term α is a numerical value of Symbolic Translation that indicates the translation of the fuzzy membership function which represents the closest term $s_i \in S = \{s_0, \dots, s_g\}$, if s_i does not match exactly the computed linguistic information. The value of α is then defined by Eq. (14):

$$\alpha = \begin{cases} [-0.5, 0.5) & \text{if } s_i \in \{s_1, s_2, \dots, s_{g-1}\} \\ [0, 0.5) & \text{if } s_i = s_0 \\ [-0.5, 0) & \text{if } s_i = s_g \end{cases} \quad (14)$$

Now, the linguistic data shown by these elements is represented as the 2-tuple linguistic value. Notice that a symbolic computation of linguistic items in S gets a value $B \in [0, g]$ and will be changed into a similar 2-tuple linguistic value (s_i, α) through Δ_s function which is described as follows (Herrera & Martínez, 2000):

Let $S = \{s_0, \dots, s_g\}$ be a set of linguistic terms and \bar{S} the 2-tuple set associated with S defined as $\bar{S} = S \times [-0.5, 0.5)$. The function $\Delta_s : \bar{S}[0, g] \rightarrow \bar{S}$ is given by Eq. (15).

$$\Delta_s(\beta) = (s_i, \alpha), \text{ with } \begin{cases} i = \text{round}(\beta) \\ \alpha = \beta - i \end{cases} \quad (15)$$

With round the function that assigns to β the closest integer number $i \in S = \{0, 1, \dots, g\}$ to β .

The followings are some introductory 2-tuple fundamental operations:

1. Arrangement of 2-tuple linguistic data

The arrangement of 2-tuple linguistic data shown in Eq. (16) by 2-tuple linguistic values is done as stated by ordinary lexicographic arrangement.

Suppose (S_K, α_1) as well as (S_L, α_2) are two 2-tuple linguistic information:

When $k < l$ so $(S_K, \alpha_1) < (S_L, \alpha_2)$.

Provided that $k = l$ therefore.

- a) When α_1 equal to α_2 ; $(S_K, \alpha_1), (S_L, \alpha_2)$ indicates similar data.
- b) When α_1 is less than α_2 ; $(S_K, \alpha_1) < (S_L, \alpha_2)$
- c) When α_1 is more than α_2 ; $(S_K, \alpha_1) > (S_L, \alpha_2)$

2. Process of negation of a 2-tuple linguistic information

The concept of negation of a 2-tuple linguistics, knowledge expression, system of logic, etc., is vital for handling the 2-tuple linguistic model. The definition is an extent of the classic negation operator for linguistic data in the fuzzy linguistic approach (Yager, 1988):

Suppose that (s_i, α) be a 2-tuple linguistic data and $S = \{s_0, \dots, s_g\}$, the negation operator is defined in Eq. (17):

$$\text{Neg}((S_i, \alpha)) = (g - (\Delta(S_i, \alpha))) \text{ With } g + 1 \text{ granularity of uncertainty of } S \quad (17)$$

2.6 Picture fuzzy sets (PFSs)

Definition 1: A picture fuzzy set A on a universe X is an object of the Eq. 18 (Cường 2015):

$$A = \{ \langle x, \mu_A(x), \eta_A(x), \nu_A(x) \rangle | x \in X \} \quad (18)$$

Where $\mu_A(x) \in [0, 1]$ is called the Degree of positive membership of x in A; $\eta_A(x) \in [0, 1]$ is called the Degree of neutral membership of x in A and $\nu_A(x) \in [0, 1]$ is called the degree of negative membership of x in A wherein, $\mu_A, \eta_A,$ and ν_A satisfy the following condition:

$$0 \leq \mu_A(x) + \eta_A(x) + \nu_A(x) \leq 1 \forall x \in X \text{ and } \pi_{A(x)} = 1 - \mu_A(x) - \eta_A(x) - \nu_A(x), \forall x \in X \text{ degree of refusal membership of } x \text{ in } A.$$

2.7 Picture 2-tuple linguistic sets (P2TLsS)

Definition 2: A picture 2-tuple linguistic set on the universe X is an object of the Eq. 19(Wei, 2017):

$$A = \{(S_\theta(x), \rho), (\mu_A(x), \eta_A(x), \nu_A(x)), x \in X\} \tag{19}$$

where $(S_\theta(x), \rho) \in S, \rho \in [-0.5, 0.5], \mu_A(x) \in [0, 1], \eta_A(x) \in [0, 1], \nu_A(x) \in [0, 1]$ with the condition $0 \leq \mu_A(x) + \eta_A(x) + \nu_A(x) \leq 1 \forall x \in X, S_\theta(x) \in S, \rho \in [-0.5, 0.5]$

The numbers $\mu_A(x), \eta_A(x), \nu_A(x)$ represent the degree of positive membership, degree of negative membership, and degree of negative membership of the element x to linguistic variable, $(S_\theta(x), \rho)$ respectively. Then for $x, \pi_{A(x)} = 1 - \mu_A(x) + \eta_A(x) + \nu_A(x), \forall x \in X$ could be called the degree of refusal membership of the element x to linguistic variable $(S_\theta(x), \rho)$.

3 Research method

3.1 Questionnaire design

The questionnaire of this study includes two parts; the first part contains demographic characteristics of the customers. The second part shows some information about digital marketing elements and their subsets. A 5-point Likert-type scale including five linguistic terms is used which is ranging from “strongly disagree” to “strongly agree” to access expectations and perceptions of the customers in using products. In this study, four of the most important digital marketing strategies were selected with the help of experts who specialize in digital marketing. These factors include SEO, Recommender Engine, Content Marketing and Social Media Marketing. Also, to cover the purchase process by customers, the marketing mix factor was selected as the fifth conditional variable. This factor includes 4p of marketing mix which are product, price, place and promotion.

3.2 Sampling and data collection

The statistical population of this study includes all customers who have referred to five best-selling online shopping stores in Tehran to prepare products. The ranking of the stores site has been used to determine the best-selling stores. Figure 2 summarizes the process of the proposed research.

3.3 Methodology

After distributing the questionnaires, customers' answers are collected in the form of language variables. First, the answers of individuals are converted to 2-tuple fuzzy numbers through Table 3.

The fuzzy numbers selected in this study are defined as follows (Zhang et al. 2019):

$$A = \{(S_\theta(x), \rho), (\mu_A(x), \eta_A(x), \nu_A(x)), x \in X\}$$

, Where set S are defined as follows:

$S = \{s_0$: strongly disagree, s_1 : disagree, s_2 : no comment, s_3 : agree, s_4 : strongly agree}.

After that, according to the attached questionnaire, each dimension is determined based on a number of questions. Therefore, for each customer, the average of fuzzy binary numbers determines the number questions related to each dimension. For example, suppose one of the dimensions is determined by four questions and one of the customers gives the following answers to the four questions: Answer to the first question:

$$[(s_1, 0), (0.15, 0.35, 0.4)].$$

Answer to the Second question:

$$[(s_3, 0), (0.45, 0.15, 0.1)].$$

Answer to the third question:

$$[(s_4, 0), (0.5, 0.5, 0.05)].$$

Answer to the fourth question:

$$[(s_4, 0), (0.5, 0.1, 0.05)].$$

Using the following formula (20) and maple software, the average of these answers is calculated as follows:

$$\begin{aligned} Av &= [AV_j]_{1 \times n} = \left[\frac{\sum_{i=1}^m r_{ij}}{m} \right]_{1 \times n} \\ &= \left[\Delta \left(\sum_{j=1}^n \frac{1}{m} \Delta^{-1}(\delta_j, \rho_j) \right), \left(1 - \prod_{j=1}^n (1 - \mu_j)^{\frac{1}{m}} \right), \right. \\ &\quad \left. \left(\prod_{j=1}^n (\eta_j)^{\frac{1}{m}} \right) \left(\prod_{j=1}^n (\nu_j + \mu_j)^{\frac{1}{m}} - \prod_{j=1}^n (\eta_j)^{\frac{1}{m}} \right) \right] \end{aligned} \tag{20}$$

Using Eq. (20), we have

$$\begin{aligned} &\Delta \left(\frac{1}{4} (\Delta^{-1}(s_1, 0) + \Delta^{-1}(s_3, 0) + \Delta^{-1}(s_4, 0) + \Delta^{-1}(s_4, 0)) \right) \\ &= \Delta \left(\frac{12}{4} \right) = (s_3, 0) \end{aligned}$$

With

$$\mu_A(x) = 0.4153, \eta_A(x) = 0.1513 \text{ and } \nu_A(x) = 0.3986$$

Table 3 Linguistic terms

Linguistic terms	P2TLNs
Strongly disagree	$[(s_0, 0), (0.1, 0.4, 0.45)]$
Disagree	$[(s_1, 0), (0.15, 0.35, 0.4)]$
No comment	$[(s_2, 0), (0.4, 0.2, 0.15)]$
Agree	$[(s_3, 0), (0.45, 0.15, 0.1)]$
Strongly agree	$[(s_4, 0), (0.5, 0.1, 0.05)]$

Where $\mu_A(x)$ is called the degree of positive membership of x in A , $\eta_A(x)$ is called the degree of neutral membership of x in A and $\nu_A(x) \in [0, 1]$ is called the degree of negative membership of x in A .

Finally, the following average number is obtained for these answers:

$$[(s_3, 0), (0.41, 0.15, 0.39)].$$

It should be noted that decimal numbers are rounded to the nearest number. Therefore, 5 conditional dimensions, 1 decision dimension and 395 customers constitute the information system shown in Table 4. To collect information in this research, five best-selling online stores in Tehran have been selected and 79 questionnaires were collected for each of these stores. In the table above, the information is expressed as A_m^n , where n represents the store and m represents the customer. For example, A_{52}^2 represents store number 2 and customer number 52 who bought from this store. It should be noted that we have changed the Iranian currency to dollar in this research. The characteristic part of the decision is defined as follows:

A^+ =Customer purchase is more than one hundred dollars.

A =Customer purchase is between sixty dollars to one hundred dollars.

B =Customer purchase is between forty dollars to sixty dollars.

C =Customer purchase is between twenty dollars to forty dollars.

D =Customer purchase is less twenty dollars.

In the next step, in order to make it possible to discover the rules, it is necessary to convert the obtained numbers into linguistic variables according to Table 5. The new information system is shown in Table 6.

4 Results

4.1 Profile of respondents

A questionnaires are planned and sent to the customers with the experience of buying in five online shopping stores based on the abovementioned dimensions. Finally, 395 questionnaires having correct data are received. Demographic variables of respondents are shown in Table 7. They are based on the first part of the questionnaire. The following characteristics of participants are presented in the demographic information: age, level of education, gender and sex. Figure 3 shows the percentage of customers and their buying behavior. For example, the percentage of customers that buying behavior are (A^+) is 16. From the 395 respondents in this study, 256 (64.8%) are female and 139 (35.2%) are male, 206 (52.1%) are single and 189 (47.9%) are married. In terms of age groups, the majority of the customers are 21–30 years old (34.6%). Most of the respondents' level of education is in the bachelor level (42.5%).

After evaluation of customers' viewpoints; based on RST, a logical connection between the collected data has been established. First, deduct and core are calculated based on the existing algorithm. The attributes that belongs to the Core include the most original and influential factors. Due to the large number of research data, 95 logical rules have been extracted with the help of RST. Tables 8, 9, 10, 11, 12 show the first three rules extracted from each of the buying behaviors.

Many of extracted rules have weak support factors. From these, 5 rules with the most repetition and according to other validation rules have been selected, which are shown in Table 13.

Table 4 Average value of data collected

Respondent	Marketing mix (4p)	Social media marketing	Search engine optimization	Content marketing	Recommender engine	Buying behavior
A_1^1	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_0, 0), (0.1, 0.4, 0.45)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	B
A_2^1	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_4, 0), (0.5, 0.1, 0.05)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_4, 0), (0.5, 0.1, 0.05)]$	A^+
A_{79}^1	$[(s_1, 0), (0.15, 0.35, 0.4)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_1, 0), (0.15, 0.35, 0.4)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_1, 0), (0.15, 0.35, 0.4)]$	B
A_{77}^5	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_1, 0), (0.15, 0.35, 0.4)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_1, 0), (0.15, 0.35, 0.4)]$	D
A_{78}^5	$[(s_1, 0), (0.15, 0.35, 0.4)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_1, 0), (0.15, 0.35, 0.4)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	C
A_{79}^5	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_2, 0), (0.4, 0.2, 0.15)]$	$[(s_3, 0), (0.45, 0.15, 0.1)]$	$[(s_4, 0), (0.5, 0.1, 0.05)]$	A

Table 5 Normalize fuzzy numbers

2-tuple	Equivalent
(s ₀ , 0)	Weak
(s ₁ , 0)	Medium
(s ₂ , 0)	Good
(s ₃ , 0)	Very good
(s ₄ , 0)	Excellent

According to the above obtained rules, the search engine optimization, the social media marketing and the recommender engine have been identified as the most important marketing strategies for online shopping respectively. It seems that if these factors appeared more intensely and other factors are in a relatively good condition, the store would have high sales. In the case of stores with relatively good sales, despite the favorable condition of the SEO social media marketing and recommender engine, the impact of the recommender engine is gradually diminishing. In other words, the social media marketing and good SEO of the website alone cannot bring good sales for stores. In general, it can be concluded that the importance of store SEO, recommender engine and social media marketing in customer behavior is more important than others. Of course, it should be noted that although these three factors have the most importance, other factors also have their special place and, in case of failure, it may cause inappropriate results in customer behavior.

4.2 Compared RST method with decision tree method

To check the accuracy of the obtained rules, the decision tree algorithm is used to check the accuracy of the rules. After entering the data into the Rapid Miner software, the conditional variables as well as the decision variable are specified, and by running the software, the decision tree is obtained (RapidMiner Inc., 2020).

Table 6 Average value of data collected after convert to linguistic term

IF						Then
Respondents	Marketing mix (4p)	Social media marketing	Search engine optimization	Recommender engine	Content marketing	Buying behavior
A ₁ ¹	Good	Good	Very good	Weak	Good	C
A ₂ ¹	Very good	Excellent	Good	Very good	Excellent	A+
A ₃ ¹	Medium	Good	Medium	Good	Medium	B
A ₇₃ ⁵	Good	Good	Medium	Good	Medium	D
A ₇₄ ⁵	Medium	Good	Very good	Medium	Good	C
A ₇₅ ⁵	Very good	Very good	Good	Very good	Excellent	A

Table 7 Demographic characteristics of study participants (N=395)

	Variables	N	(%)
<i>Gender</i>	Male	139	35.2
	female	256	64.8
<i>Age</i>	0–20	76	19.2
	21–30	137	34.6
	31–40	126	31.9
	41–50	41	10.6
	51–60	15	3.7
<i>Married status</i>	Single	206	52.1
	Married	189	47.9
<i>Level of education</i>	High school and less	90	22.8
	Bachelor	168	42.5
	Master	119	30.1
	Ph.D. and more	18	4.6

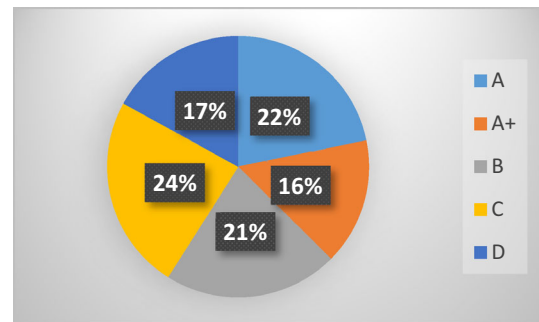


Fig. 3 Distribution of customer types

Table 8 The first three rules of buying behavior A+

IF					THEN	Certainty factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
Excellent	–	Excellent	Very good	–	A+	47	47	1	0.1189	0.7580	3	1.0	
–	Excellent	–	Very good	–		47	47	1	0.1189	0.7580	2	1.0	
–	Excellent	–	–	Excellent		36	36	1	0.0916	0.5806	2	1.0	

Table 9 The first three rules of buying behavior A

IF					THEN	Certainty factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
Good	Good	Very good	Good	–	A	81	81	1	0.2050	0.9418	4	1.0	
Very good	Good	–	Good	Good		3	2	0.66	0.0076	0.0232	4	1.0	
–	–	Excellent	Good	–		1	1	1	0.0025	0.0116	2	1.0	

Table 10 The first three rules of buying behavior B

IF					THEN	Certainty factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
–	Good	Medium	Good	–	B	64	64	1	0.1620	0.7529	3	1.0	
–	Good	–	Good	Medium		63	63	1	0.1594	0.7411	3	1.0	
–	Good	–	Medium	Good		5	5	1	0.0126	0.0588	3	1.0	

Table 11 The first three rules of buying behavior C

IF					THEN	Certainty factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
–	Medium	Medium	–	Medium	C	69	69	1	0.1746	0.7263	3	1.0	
Good	Medium	–	–	Medium		67	67	1	0.1692	0.7056	3	1.0	
Good	Medium	Medium	Good	–		64	64	1	0.1620	0.6736	4	1.0	

After obtaining the decision tree, the rules derived from the decision tree are also extracted to be compared with the rules obtained from Rough Set theory. Due to the large number of rules, only the rules with the most repetition are displayed. Table 14 shows both sets of rules obtained by the Rough Set Theory and the Decision Tree technique. As it

can be seen in this table, most of rules derived from Rough Set Theory are the same as rules derived from Decision Tree technique or in some occasion is a subset of rules derived from decision tree.

Table 12 The first three rules of buying behavior D

IF					THEN	Certainty factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
–	Weak	Medium	Medium	Medium	D	60	60	1	0.1518	0.8955	4	1.0	
Weak	Weak	–	–	–		7	7	1	0.0177	0.1044	2	1.0	
Weak	–	–	–	Medium		7	7	1	0.0177	0.1044	2	1.0	

Table 13 The generated rules

IF					THEN	Certainty Factors							
Marketing mix (4p)	SMM	SEO	RE	Content marketing	Buying behavior	LHS support	RHS support	RHS accuracy	LHS coverage	RHS coverage	LHS length	RHS length	
Excellent	–	Excellent	very good	–	A+	47	47	1	0.1195	0.7580	3	1.0	
Good	Good	Very good	Good	–	A	81	81	1	0.2061	0.9418	4	1.0	
–	Good	Medium	Good	–	B	64	64	1	0.1628	0.7529	3	1.0	
–	Medium	Medium	–	Medium	C	69	69	1	0.1755	0.7263	3	1.0	
–	Weak	Medium	Medium	Medium	D	58	58	1	0.1475	0.8923	4	1.0	

Table 14 Extract rules from Rough Set Theory and Decision Tree

Extract rules from Rough Set Theory			Extract rules from Decision Tree		
IF	THEN		IF	THEN	
If (Marketing mix=excellent) and (SEO=excellent) and (recommender engine=very good)	A+		If (Marketing mix=excellent) and (SEO=excellent) and (recommender engine=very good)	A+	
If (Marketing mix=good) and (SMM=good) and (SEO=very good) and (RE=good)	A		If (Marketing mix=good) and (SMM=good) and (SEO=very good) and (Content Marketing=good) and (RE=good)	A	
If (SMM=good) and (SEO=medium) and (RE=good)	B		If (SMM=good) and (SEO=medium) and (RE=good)	B	
If (SMM=medium) and (SEO=medium) and (Content Marketing=medium)	C		If (SMM=medium) and (SEO=medium) and (Content Marketing=medium)	C	
If (SMM=weak) and (SEO=medium) and (RE=medium) and (Content Marketing=medium)	D		If (SMM=weak) and (SEO=medium) and (RE=medium) and (Content Marketing=medium)	D	

5 Managerial insights

Due to the competitive market for online shopping stores, managers are offered the following suggestions to improve service and ultimately increase sales.

- a) Since the SEO factor is one of the most important elements from the customers' attitude and it can be seen in all extracted rules, SEO managers need to be conscious about different SEO techniques which guarantee the most related rich content website at the top of the result pages searched in internet.
- b) Choosing the best SOE' strategy will help marketing managers to focus on website optimization that can also reduce search engine marketing budget for organization.
- c) Managers should also identify effective keywords and increase their effectiveness' presence in the website content.
- d) Marketing managers must pay the most attention to engage costumers with diverse brand-related process and activities via different social media in order to create a relationship that is permanent and stable.

- e) Managers should benefit by focusing on how products customers use social media when making products purchasing decisions.
- f) Senior technology managers of companies should choose the best RE's strategy solution. They are also encouraged to use this strategy for the customers regarding the customers' buying pattern, interests, observations, and perchanche rather than assumptions or guesswork.
- g) In order to improve the marketing mix element, recommended to increase the variety of colors, materials, and design related to products. Also, they should use eminent brands in offering products to customers have a great importance that managers should be noticed. In addition, the result of the effective factors in the element of price and promotion emphasizes that discounts and similar appropriate policies can help store managers to attract potential customers to their stores. Pay attention to impact of web design, the appearance of the website, the fonts, and other related factors also have a significant impact on customer's behavior in online shopping.

6 Conclusions

The main purpose of this research is to present efficient rules for online shopping stores owners to recognize the factors affecting customer's buying behavior. Moreover, findings of this paper help marketing managers to provide services tailored to customer's requirements to more welcome which leads to increase the profitability of the store. It is observed in the obtained rules, SEO is seen in all effective rules and has the most important in terms of rank among all rules. Moreover, two factors RE and SMM seen in four rules and both are the second priority.

Like to each research, this work faces to some limitations that should be considered by management decision makers in marketing area. Digital marketing strategies are very complex and changeable. However, this study has been performed by considering the most important factors that have the greatest influence on customers' online shopping behavior. Moreover, we have relaxed some digital marketing methods such as Pay-Per-Click (PPC), Search Engine Marketing (SEM), E-mail, and other similar approach in the present article. These factors may be included in the further extension of the study. Therefore, the related managers and decision makers are advised to use findings of this study after review and customizing the effect factors.

It also should be noted that the research scope of the present study is limited to digital marketing methods and behavior of customers in online shopping stores. The result

may not be appropriate for other marketing policies or for other purchases.

Future researchers can evaluate the impact of customer's buying behavior by examining other data mining methods. Also, using SAS (statistical analysis systems) and Python data mining software to obtain association rules and decision tree can help future researchers to achieve accurate performance in the optimal result.

Funding The author do not have any financial interests/personal relationships which may be considered as potential competing interests.

Declarations

Conflict of interest The author declared that there is no conflict of interest.

References

- Applebaum W (1951) Studying customer behavior in retail stores. *J Mark* 16(2):172–178
- Bala M, Verma D (2018) A critical review of digital marketing. *Int J Manag* 8(10):321–339
- Bawm ZL, Nath RPD (2014). A Conceptual Model for effective email marketing. In 2014 17th International Conference on Computer and Information Technology (ICCIT) (pp. 250–256). IEEE.
- Behera RK, Gunasekaran A, Gupta S, Kamboj S, Bala PK (2020). Personalized digital marketing recommender engine. *Journal of Retailing and Consumer Services*, 53, 101799.
- Bianchi C, Mathews S (2016) Internet marketing and export market growth in Chile. *J Bus Res* 69(2):426–434
- Bu Y, Parkinson J, Thaichon P (2020) Digital content marketing as a catalyst for e-WOM in food tourism. *Australasian Marketing Journal (AMJ)*.
- Carpio K V, Arce S, Enjolras M, Camargo M (2020) How Institutions Promote Digital Marketing in Small and Medium International Companies: a Comparison Between Costa Rica and France. *Technology Innovation Management Review*, 10(4).
- Chiaselotti G, Gentile T, Infusino F (2019) Decision systems in rough set theory: a set operatorial perspective. *J Alg Appl* 18 (01):1950004
- Çizmeçi F, Ercan T (2015) The effect of digital marketing communication tools in the creation brand awareness by housing companies. *Megaron*, 10(2).
- Cường BC (2015) Picture fuzzy sets. *J Comput Sci Cybernet*. <https://doi.org/10.15625/1813-9663/30/4/5032>
- Das AK, Sengupta S, Bhattacharyya S (2018) A group incremental feature selection for classification using rough set theory based genetic algorithm. *Appl Soft Comput* 65:400–411
- Desai V (2019) Digital marketing a review. *Int J Trend Sci Res Develop*. <https://doi.org/10.1016/j.jretconser.2019.101909>
- Dwivedi YK, Rana NP, Slade EL, Singh N, Kizgin H (2020). Editorial introduction: Advances in theory and practice of digital marketing.
- Fakir Y, Azalmad M, Elaychi R (2020) Study of the ID3 and C4. 5 learning algorithms. *J Med Inf Dec Mak* 1(2), 29.
- Fitrani AS, Rosid MA, Findawati Y, Rahmawati Y, Anam AK (2019). Implementation of ID3 algorithm classification using web-based weka. In *Journal of Physics: Conference Series* (Vol. 1381, No. 1, p. 012036). IOP Publishing.

- Girchenko T, Ovsianikova Y (2016) Digital marketing and its role in the modern business processes. *Eur Cooper* 11(18):24–33
- Han Z, Zhang Q, Wen F (1999) A survey on rough set theory and its application. *Kongzhi Lilun Yu Yingyong/control Theory Appl* 16(2):323–333. <https://doi.org/10.1016/j.trit.2016.11.001>
- Herrera F, Martínez L (2000) A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Trans Fuzzy Syst* 8(6):746–752
- Hollebeck LD, Macky K (2019) Digital content marketing's role in fostering consumer engagement, trust, and value: Framework, fundamental propositions, and implications. *J Interact Mark* 45:27–41
- RapidMiner Inc. (2020). *RapidMiner _ Best Data Science & Machine Learning Platform*. <https://rapidminer.com/>
- Ivkovic M, Milanov D (2010) Affiliate internet marketing: Concept and application analysis. In 2010 International Conference on Education and Management Technology (pp. 319–323). IEEE.
- Jia X, Rao Y, Shang L, Li T (2020) Similarity-based attribute reduction in rough set theory: a clustering perspective. *Int J Mach Learn Cybernet* 11:1047–1060. <https://doi.org/10.1007/s13042-019-00959-w>
- Kannan PK (2017) Digital marketing: a framework, review and research agenda. *Int J Res Mark* 34(1):22–45
- Karimi S, Naghibi HS (2015) Social media marketing (SMM) strategies for small to medium enterprises (SMEs). *Int J Inf Bus Manag* 7(4):86
- Kim D, Ammeter AP (2018) Shifts in online consumer behavior: a preliminary investigation of the net generation. *J Theor Appl Electron Commer Res* 13(1):1–25
- Kritzinger WT, Weideman M (2013) Search engine optimization and pay-per-click marketing strategies. *J Organ Comput Electron Commer* 23(3):273–286
- Kulkarni KK, Kalro AD, Sharma D, Sharma P (2020) A typology of viral ad sharers using sentiment analysis. *J Retail Consum Serv*. <https://doi.org/10.1016/j.jretconser.2019.01.008>
- Malik R (2017) An Emp Study Dig Market Elem 8(3):715–718
- Martínez L, Rodríguez R M, Herrera F (2015) 2-tuple linguistic model. In *The 2-tuple Linguistic Model* (pp. 23–42). Springer, Cham.
- Moheimani A, Sheikh R, Hosseini SMH, Sana SS (2021) Assessing the preparedness of hospitals facing disasters using the rough set theory: guidelines for more preparedness to cope with the COVID-19. *Int J Syst Sci Oper Log*, 1–16.
- Neti S (2015) Social media and its role in marketing. *Bus Econ J*. <https://doi.org/10.4172/2151-6219.1000203>
- Nizar NA, Janathanan C (2018). Impact of digital marketing on consumer purchase behaviour. In *APIIT Business, Law & Technology Conference*.
- Othman ML, Aris I, Abdullah SM, Bakar MLAMA, Othman MR (2009) Discovering decision algorithm of numerical distance relay using rough-set-theory-based data mining. *Eur J Sci Res* 33(1):30–56
- Pawlak Z (1998) Rough set theory and its applications to data analysis. *Cybern Syst* 29(7):661–688
- Piñero-Otero T, Martínez-Rolán X (2016). *Understand Digit Market Bas Actions*. https://doi.org/10.1007/978-3-319-28281-7_2
- Qowidho T, Zarlis M, Nababan E B, Agusnady A, Sembiring BS (2019). Analysis Of c4. 5 and ID3 Methods in Determining Student Graduation. In *Journal of Physics: Conference Series* (Vol. 1255, No. 1, p. 012025). IOP Publishing.
- Sajjadian F, Sheikh R, Sourí ME, Sana SS (2018) Application of rough set and netnography in tourism marketing analysis. *J Modell Manag*. <https://doi.org/10.1108/JM2-02-2018-0029>
- Sarkar B, Ullah M, Choi SB (2019) Joint inventory and pricing policy for an online to offline closed-loop supply chain model with random defective rate and returnable transport items. *Mathematics* 7:497–551
- Saura JR, Palos-Sanchez PR, Correia MB (2019) Digital marketing strategies based on the e-business model: Literature review and future directions. *Org Transf Manag Innov Fourth Ind Revol*. <https://doi.org/10.4018/978-1-5225-7074-5.ch005>
- Sawłani DK, Susilo D (2020) How digital marketing helps higher education institution branding. *Int J Multidis Edu Res* 5(3):45–55
- Sett BK, Dey BK, Sarkar B (2020) The effect of O2O retail service quality in supply chain management. *Mathematics* 8(10):1743
- Shah J (2018) Impact of digital marketing in business and politics. *Int J Sci Eng Res* 9(2):10–14299
- Sharma S, Singh S, Kujur F, Das G (2021) Social media activities and its influence on customer-brand relationship: an empirical study of apparel retailers' activity in India. *J Theor Appl Electron Commer Res* 16(4):602–617
- Shen CW, Luong TH, Ho JT, Djailani I (2020) Social media marketing of IT service companies: analysis using a concept-linking mining approach. *Ind Mark Manage* 90:593–604
- Solodka N, Liashenko O (2020) SEO web-resource strategy development. In *International scientific and technical conference Information technologies in metallurgy and machine building* (pp. 424–427).
- Solomon, M. R. (2010). *Consumer behaviour: A European perspective*. Pearson education.
- Stávková J, Stejskal L, Toufarová Z (2008) Factors influencing consumer behaviour. *ZEMEDELSKA EKONOMIKA-PRAHA* 54(6):276
- Taherdoost H, Jalaliyoon N (2014). Marketing vs E-Marketing. *International J Acad Res Manag (IJARM)*, 3(4), 335–340. <http://elvedit.com/journals/IJARM/wp-content/uploads/2015/01/marketing-e-marketing.pdf>
- Terrance AR, Shrivastava S, Kumari A (2018) Importance of search engine marketing in the digital world. In: *ICITKM*, vol 14, pp 155–158
- Tiago MTPMB, Veríssimo JMC (2014) Digital marketing and social media: Why bother? *Bus Horiz* 57(6):703–708
- Todor RD (2016). Blending traditional and digital marketing. *Bulletin of the Transilvania University of Brasov. Economic Sciences. Series V*, 9(1), 51.
- Veleva S S, Tsvetanova AI (2020). Characteristics of the digital marketing advantages and disadvantages. In *IOP Conference Series: Materials Science and Engineering* (Vol. 940, No. 1, p. 012065). IOP Publishing.
- Wang P, McCarthy B (2020) What do people “like” on Facebook? Content marketing strategies used by retail bank brands in Australia and Singapore. *Australasian Marketing Journal (AMJ)*.
- Wei G (2017) Picture 2-tuple linguistic Bonferroni mean operators and their application to multiple attribute decision making. *Int J Fuzzy Syst* 19(4):997–1010
- Wei W, Liang J (2019) Information fusion in rough set theory: An overview. *Inform Fusion* 48:107–118
- Xie Z, Du L, Lv X, Wang Q, Huang J, Fu T, Li S (2020) Evaluation and analysis of battery technologies applied to grid-level energy storage systems based on rough set theory. *Trans Tianjin Univ*. <https://doi.org/10.1007/s12209-020-00237-9>
- Yager RR (1988) On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans Syst Man Cybern* 18(1):183–190

- Yao Y (2008) Probabilistic rough set approximations. *Int J Approximate Reasoning* 49(2):255–271
- Zhang X, Kumar V, Cosguner K (2017) Dynamically managing a profitable email marketing program. *J Mark Res* 54(6):851–866
- Zhang S, Gao H, Wei G, Wei Y, Wei C (2019) Evaluation based on distance from average solution method for multiple criteria group

decision making under picture 2-tuple linguistic environment. *Mathematics* 7(3):243

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.