Journal of Agricultural Economics & Development

Homepage: https://jead.um.ac.ir

Research Article
Vol. 38, No. 4, Winter 2025, p. 351-370

A Data Mining Approach to Consumers' Choice of Retail Market: The Case of Urban Retail Markets in Iran

H. Hashemzadeh[1], N. Yousefian[2], S. Esfandiari Bahraseman[2], A. Karbasi[3], A. Firoozzare[4]*

1, 2, 3 and 4- M.Sc in Agricultural Economics, Ph.D. Student, Professor and Assistant Professor, Department of Agricultural Economics, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran, respectively.
(*- Corresponding Author Email: firooz@um.ac.ir)

Received: 20-05-2024 Revised: 06-07-2024 Accepted: 08-07-2024 Available Online: 08-07-2024	**How to cite this article:** Hashemzadeh, H., Yousefian, N., Esfandiari Bahraseman, S., Karbasi, A., & Firoozzare, A. (2025). A data mining approach to consumers' choice of retail market: The case of urban retail markets in Iran. *Journal of Agricultural Economics & Development, 38*(4), 351-370. https://doi.org/10.22067/jead.2024.88151.1269

Abstract

Urban retail markets are state-owned retail markets that were recently established in Iran to increase the welfare of consumers and producers. To achieve this goal and expand its presence in the Iranian retail sector, it is essential to gain a comprehensive understanding of consumer behavior within these markets. This study examines the various socio-economic factors influencing consumers' decisions in the retail market by using the C4.5 algorithm. The data were collected using a random sampling method through a survey of 189 consumers, focusing on the population of Mashhad, Iran, during 2019-2020. Results revealed that awareness of available discounts significantly drives consumer choices in urban retail markets. Despite existing discounts, awareness among consumers remains low, suggesting a need to review promotional strategies within the marketing mix. The study also identifies previous purchases from urban markets, household income, and education as influential factors. Findings offer valuable insights for policymakers, market strategists, and stakeholders seeking to enhance the effectiveness of local retail markets in Iran. By leveraging insights into consumer behavior and market dynamics, these markets can thrive, benefiting Iran's retail sector and overall economy. Following the study, recommendations such as enhanced promotional campaigns, education-oriented strategies, loyalty programs, collaborations with local producers, and inclusive marketing policies was made aim to improve access for all consumers to urban retail markets.

Keywords: Consumer behavior, Data mining, Decision tree, Machine learning

https://doi.org/10.22067/jead.2024.85975.1237

Introduction

By identifying the critical factors that influence buyers' purchasing behaviors, marketers can craft strategies that align with these determinants, thereby securing their competitive edge in the market (Shamsher, 2016). Consumer behavior has remained a focal point in the discourse on retail marketing and decision-making for more than 70 years (Rahim *et al.*, 2021; Tian *et al.*, 2021).

Understanding consumer behavior involves delving into the psychological, social, and cultural factors that influence how individuals make purchasing decisions. Initially emerging as a specialized sub-discipline within marketing, the study of consumer behavior has consistently shown that purchasing decisions are influenced by a complex interplay of sociological and psychological factors (Applebaum, 1951; Dominici *et al.*, 2021; Donoghue *et al.*, 2021). Furthermore, contemporary research in this field views consumer behavior as an interdisciplinary social science, integrating insights from psychology, sociology, anthropology, ethnography, marketing, and economics, particularly behavioral economics (Razu & Roy, 2019).

In recent years, Iran's retail industry has undergone a dramatic transformation, marked by the growing prominence of organized retail. This shift has brought profound changes within the sector (Shamsher, 2016; Gauri *et al.*, 2021). These changes include the proliferation of shopping malls, supermarkets, and hypermarkets, which have altered traditional shopping patterns and consumer expectations. The rise of organized retail in Iran has been accompanied by an increase in consumer access to a wider variety of products and brands, more competitive pricing, and enhanced shopping experiences. This evolution has necessitated that retailers in Iran and similar markets pay closer attention to consumer behavior to remain competitive.

Iran's retail sector has evolved into an oligopoly, dominated by a few corporations operating chains of hypermarkets that supply most grocery products (Jofreh, 2013). Recently, state-operated urban retail markets have entered the Iranian market. State-owned retail markets were common in many countries, such as China, where urban consumers were limited to shopping at state-run establishments until 1978 (Veeck & Veeck, 2000). Although these markets are often criticized for inefficiencies (Hingley *et al.*, 2009), they must improve their performance to compete in a competitive environment.

The establishment of state-operated urban retail markets in Iran aims to ensure a direct and timely supply of fruits, vegetables, meat, poultry, fish, detergents, and sanitizers in the domestic market; support consumer rights; establish suitable and effective pricing; organize and regulate food markets; and facilitate the preparation and distribution of products by creating a direct connection between producers and consumers (Golriz Ziaie *et al.*, 2015). Considering these factors, promoting shopping in these markets enhances consumer welfare. Increasing the market share of urban markets can also lead to greater profits for producers, as these markets provide a platform for producers to sell their goods and reduce marketing margins (Thakur *et al.*, 2023). In this regard, marketers of urban retail markets must gain comprehensive knowledge about the factors that determine consumers' choice of retail market and how to influence their purchasing decisions (Liyanage *et al.*, 2020; Thakur *et al.*, 2023).

By understanding consumer preferences, behaviors, and the socio-economic variables that impact shopping habits, marketers can develop strategies that attract more customers to these markets. Several studies have looked for factors that influence consumers' choice of retail markets. For instance, Laine (2014) and Bhatti *et al*. (2015) identified that a store's product selection, previous experience, and location are crucial factors affecting consumers' retail market choices. Zulqarnain *et al.* (2015) examined factors influencing consumers' choice of retail stores, focusing particularly on grocery stores. Pandey & Kaur (2018) conducted a study comparing retail marketing between rural and urban regions. Their findings underscored that urban as well as rural retail markets need to be addressed separately by any retail marketer because both regions offer several opportunities. Rasheed *et al.* (2018) found that if the store atmosphere is neat, the point of purchase display is attractive, promotional activities are influencing and payment facilities are provided then these will promote more impulse buying. Nguyen (2019) identified four factors influencing supermarket choice that includes location, perception of prices and products, employee attitudes, and references. Manuere (2023) investigated factors affecting customers' choice of supermarkets for grocery shopping in Chinhoyi town. The results of the study revealed that the location of the supermarket, children's play areas and parking areas are likely to influence the customers' choice of a supermarket.

The previous studies primarily focus on global contexts, overlooking specific insights crucial for the Iranian market. This gap hinders the development of customized strategies that align with local consumer preferences and socioeconomic dynamics. Moreover, the profound influence of culture on consumer behavior means that findings from other studies may not be applicable to Iran. This underscores the critical need for localized research to fill this gap and guide effective marketing strategies. The current study addresses these challenges by examining how socioeconomic variables impact consumer purchasing decisions in Iranian urban retail markets. Through the innovative application of machine learning algorithms for consumer classification, the study introduces a unique analytical framework tailored to predict and understand consumer behavior within Iranian urban retail markets. This approach aims not only to uncover the most influential socioeconomic factors shaping consumer choices but also to equip marketers with practical insights to enhance their ability to attract buyers to urban retail markets. In doing so, the study contributes to advancing our understanding of consumer behavior in a culturally specific context and introduces a novel method for consumer classification that can be adapted for similar studies.

Materials and Methods

Data Mining

Data mining is the investigation and examination of large amounts of data to find significant patterns and rules (Berry & Linoff, 2000). It tries to discover potentially useful, interesting, and previously unknown patterns from a large collection of data (Ram, 2022). Classification is one of the main goals associated with data mining. It is the characteristic knowledge that displays the shared properties of the same type of items and the shared characteristics of differences between different items (Vindigni *et al.*, 2022). It aims to retrieve rules from the database. The most popular multidimensional classification technique is decision tree (Moitra *et al.*, 2021; Quinlan, 2014), which is used in this survey.

Decision Tree

Decision trees (DT) are powerful tools that have been widely used to build classification models (Kotsiantis, 2013; Ooi *et al.*, 2017). The algorithm created by a DT can be turned into multiple IF-THEN rules that display the relationship between output and input traits. A DT is made up of one root node, a few internal nodes, and numerous leaf nodes. The DT begins with the root node. The internal nodes join the leaf and root nodes. Each leaf node is equipped with a class label. The routes leading from the root node to the leaf nodes show the classification rules (Dev & Eden, 2019; Meng *et al.*, 2020). The advantages of using DT compared with other methods are (Kotu & Deshpande, 2018; Nisbet *et al.*, 2018):

- DT is more accurate compared with other data mining methods.
- It can be used for a wide range of data types, including continuous and discrete data.
- Compared with other classification methods, it requires less time.
- Extracted rules are simple to understand

and interpret.

Commonly utilized DT algorithms include the CART, Iterative Dichotomizer (ID3), and C4.5 algorithms (Barh, 2020). All the classification methods have a similar structure, to create a high-quality tree with a low error rate. Most of them use top-down greedy search to traverse the possible spaces of a DT. However, the majority of them have different branching and branch-cutting techniques (Han *et al.*, 2022). Developed by Quinlan (1996), C4.5 is an improved version of ID3 that uses the information gain ratio as the standard for selecting attributes rather than information gain. Like the CART algorithm, C4.5 considers each node as recursive and chooses the optimal part. This trend will continue until there are no optimal parts left. There are two main advantages associated with the C4.5 algorithm. First, it can handle continuous attributes better than the CART algorithm (not limited to binary attributes); second, the overfitting issue that exists with the ID3 technique can be resolved using the C4.5 approach (Zaki & Meira, 2020; Sun *et al.*, 2007). The process of forming a decision tree and selecting an attribute is described below (Sugumaran & Ramachandran, 2007; Kumar *et al.*, 2014):

1- A collection of available features is the input of the algorithm, and DT is the output of it.
2- Leaf nodes of the DT represent the class labels, and other nodes represent classified classes.
3- The tree's branches stand in for each predictive value of the feature node they originate from.
4- Starting at the tree's root and proceeding through it until one reaches a leaf node (which offers a categorization of the instance), the DT can be used to categorize feature vectors.
5- Using the right estimation criteria, one can choose the most beneficial feature for classification at each decision node in the DT. The ideas of entropy reduction and information gain—discussed in the following subsection—are applied in the criteria used to determine the optimal feature.

In the C4.5 algorithm, the decision tree is built from top to bottom. This technique uses information gain to determine which attribute should be utilized to categorize the current subset of data. Then, to choose the best alternative for the next node, the same process takes place for the data of each branch (Cherfi *et al.*, 2018). 'Information gain' is the anticipated loss of entropy owing to portioning the samples according to a specific attribute, while 'Entropy' describes the impurity of an arbitrary collection of examples. Increasing information lessens uncertainty. Entropies of the initial system and the system after information is added, are compared in information gain. Information gain (S, A) of a feature *A* relative to a collection of examples *S* is defined by the following formula (Reddy & Chittineni, 2021; Lintang *et al.*, 2022):

$$\text{Gain(S,A)} \equiv \text{Entropy}(S) - \sum v \in Value(A) \frac{|Sv|}{S} \text{Entropy } (S_v \quad (1)$$

Where *Values (A)* refers to the collection of all potential values for *A* attribute, and corresponding S_v designates the subset of *S* for which feature *A* has value *v*. *S* stands for the original collection's entropy in the equation above, and *A* is the expected value of entropy. The entropies belonging to each subset Sv are added together to get the expected entropy denoted by the word *A*. *Gain (S,A)* is the expected depletion in entropy as a result of the known value of attribute *A*. Entropy is:

$$Entropy(S) = \sum_{i=1}^{c} - p_i \log_2 p_i \quad (2)$$

Where P_i is the proportion of S belonging to class 'i' (Kumar *et al.*, 2014). The present study used the J48 algorithm (equal to the C4.5 algorithm in WEKA software) for learning the decision tree.

Goodness-of-fit of decision-tree models

The performance of the decision tree can be measured by indexes below:

- True Positive Rate
- False Positive Rate
- Precision
- Recall

- F amount

The True Positive Rate (TPR) criterion, also called the Hit Rate, is the ratio of buyers who are correctly classified in each group and is calculated as follows:

$$\text{TP Rate} \approx \frac{\text{Number of consumers correctly classified}}{\text{Total number of consumers}} \quad (3)$$

False Positive Rate represents the ratio of buyers who are classified incorrectly by DT and is calculated as follows:

$$\text{FP Rate} \approx \frac{\text{Number of consumers incorrectly classified}}{\text{Total number of consumers}} \quad (4)$$

The measures of Precision and Recall are also calculated with the following formulas:

$$\text{Precision} = \frac{\text{TP rate}}{\text{TP rate+FP rate}} \qquad \text{Recall} = \text{TP rate} \quad (5)$$

The index that combines FPR and TPR is the harmonic mean of Precision and Recall and is calculated as below (Walkinshaw, 2013):

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

Cohen's kappa (κ) statistic is a chance-corrected method for measuring agreement (rather than association) among raters. It answers the question of: What percentage of values not expected to be in agreement (by chance) are in agreement? The formula of Kappa is:

$$\widehat{K} = \frac{f_0 - f_E}{N - f_E} \quad (7)$$

Where f_0 is the relative observed agreement among raters, f_E is the hypothetical probability of chance agreement, and N is the total number of observations (Cohen, 1960). The Kappa result can be interpreted as follows: values ≤ 0 denote no agreement, 0.01-0.20 indicate no to slight agreement, 0.21-0.40 indicate fair agreement, 0.41- 0.60 indicate the moderate agreement, 0.61-0.80 indicate substantial agreement and 0.81-1.00 denote almost perfect agreement (McHugh, 2012).

Data

Examining urban retail markets in Mashhad is particularly significant due to the city's dual role as a major hub for both domestic residents and international tourists. With millions of visitors annually, Mashhad's retail sector experiences unique dynamics influenced by diverse consumer behaviors shaped by cultural, social, and economic factors specific to the region. In this context, a questionnaire was designed to identify and understand the socio-economic factors that impact consumers' choices in retail markets. Initially, pertinent literature was reviewed to select the most relevant socio-economic variables affecting consumer behavior. Following this, the questionnaire was refined through focus group discussions to comprehensively capture the nuances of these factors in Mashhad's retail markets. By conducting a pre-study, the sample size was determined, and necessary tests to ensure its validity and reliability were performed. In this regard, 30 questionnaires were filled out. It should be noted that the data were collected in 2019-۲۰۲۰ . The results of this sample showed that the variance of the studied attribute, which was calculated based on the probability of each of the dependent variable groups, is equal to 0.238. The current study employed a simple random sampling method to ensure equal opportunity for every individual in the studied population to be selected and to prevent bias in sample selection. Furthermore, the sample size of 189 observations was determined based on Equation 10.

$$n = \frac{Z^2 \delta_i^2}{d^2} \quad (8)$$

When Z is the amount of normal variable with $1-\alpha$ confidence level (in the present study 95% confidence level is considered), d is the margin of error and δ_i^2 stands for the variance of the studied attribute. Prior to conducting the survey, the Ferdowsi University of Mashhad granted ethics approval. All the participants in the survey were informed regarding the aim of this study and were asked to answer in an honest manner. They were assured regarding the confidentiality of their response, the survey's scientific intent, and their privacy. All the participants were responsible for the shopping in their families and were older than 18 years old.

Result and Discussion

Descriptive statistics

To discover effective parameters and factors affecting the selection of urban retail markets by consumers, this study used the C4.5 algorithm. Table 1 shows the variables used in this study.

Table 1- Description of variables in the study

Variables	Explanation
Priority in purchase (decision variable)	purchase from other shops=1; purchase from urban market=2
Gender	Female=0; Male=1
Age	Less than 40=1; More than 40=2
Job	Housekeeper, Unemployed, Retired=1; Self-employed=2; Employment in public and private sector jobs=3
Education	Less than university education =1; University education=2
Distance from urban market	Near=1; Far=2
Previous purchase from urban Markets	No=0; Yes=1
Existence of discounts and awareness of that in the urban markets	No=0; Yes=1
Household income per month	Less than 357 dollars =1; Between 357 and 655 dollars = 2; More than 655 dollars = 3

According to Table 1, effective socio-economic variables in the selection of urban retail markets include gender, age, job, education, distance from the urban markets, previous purchases from the urban markets, awareness of the discounts in the urban markets, and household income.

Fig. 1 provides an overall illustration of the sample. It is made up of nine different figures, each of which belongs to a different variable. On the left side of each picture are those who coded with the least number (0 or 1), and on the right side are those with higher codes. Those with blue colors have prioritized purchasing from other markets over urban markets.

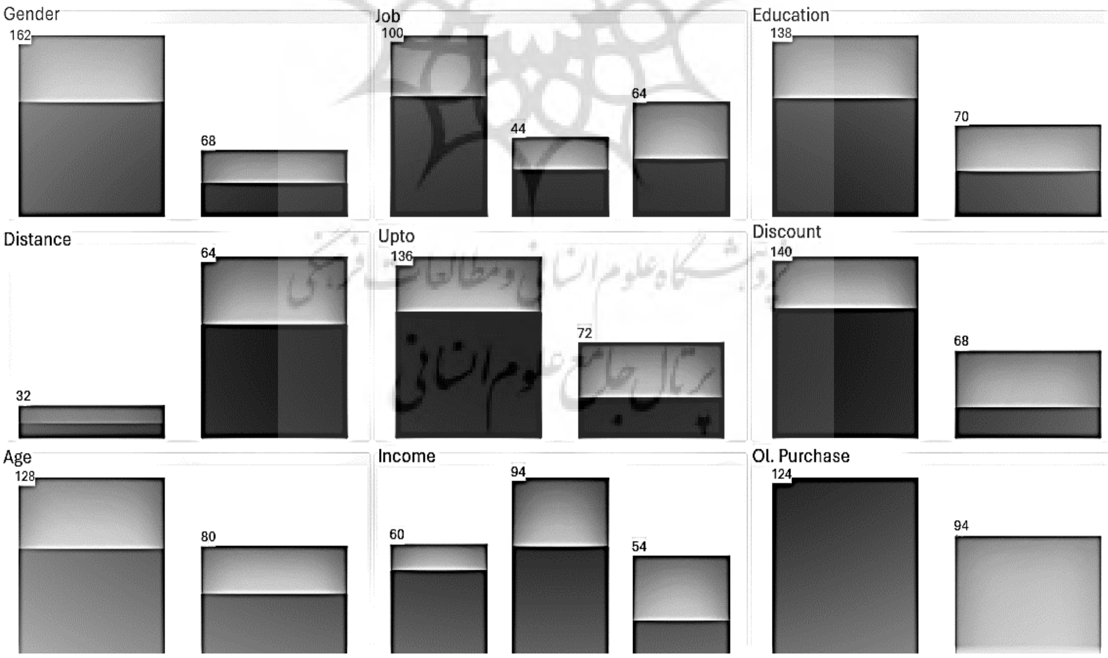

Figure 1- The relation between the independent variables and the purchasing from the urban markets or other shop

Table 2- Descriptive analysis of variables

Variable	Decision group variable	Group	Percentage
Gender	1	0	77.4
		1	22.6
	2	0	66.7
		1	33.3
	Total	0	73.1
		1	26.9
Age	1	1	62.9
		2	37.1
	2	1	59.5
		2	40.5
	Total	1	61.5
		2	38.5
Job	1	1	53.2
		2	21.0
		3	25.8
	2	1	40.5
		2	21.4
		3	38.1
	Total	1	48.1
		2	21.2
		3	30.8
Education	1	1	72.6
		2	27.4
	2	1	57.1
		2	42.9
	Total	1	66.3
		2	33.7
Distance from urban markets	1	1	11.3
		2	88.7
	2	1	21.4
		2	78.6
	Total	1	15.4
		2	84.6
Previous purchases from urban markets	1	0	75.8
		1	24.2
	2	0	50.0
		1	50.0
	Total	0	65.4
		1	34.6
Discount and being aware of it	1	0	80.6
		1	19.4
	2	0	47.6
		1	52.4
	Total	0	67.3
		1	32.7
Household Income Per Month	1	1	37.1
	2	2	46.8
	3	3	16.1
	1	1	16.7
	2	2	42.9
	3	3	40.5
	Total	1	28.8
		2	45.2
		3	26.0

Fig. 1 and Table 2 provide a detailed demographic overview of the sample. The study shows a predominant representation of female participants, particularly among shoppers from other markets compared to those from urban markets. Previous research in Iran has consistently observed a higher participation rate among women, a trend that is accentuated in this study. Awareness of discounts offered at urban markets is notably low, with a significant disparity between those who prioritize shopping elsewhere and others. A substantial portion of respondents had never made purchases from urban markets, indicating a potential barrier to market engagement. Interestingly, a large majority of participants perceive urban markets as distant from their residences, reflecting a common sentiment regardless of shopping preferences.

The results of the C4.5 algorithm

Using the C4.5 algorithm, the general behavior of consumers in their market selection is first investigated, and then the behavior of each consumer group is examined. Before analyzing the decision tree's algorithm in Fig. 1, it is required to investigate classification accuracy using various goodness of fit tests. Table 3 shows that the fitted tree in Table 1 has satisfactory classification accuracy, properly classifying more than 87 percent of first-group buyers, 83 percent of second-group buyers, and 86 percent of all purchasers.

Table 3- The general classification of consumers based on purchase priority from different markets using the C4.5 algorithm

Consumers	Number of observations	Results of decision tree C4.5 algorithm classification			
		Priority: purchase from other shops		Priority: purchase from the urban markets	
Priority: purchase from other shops (Group 1)	124	108 87.10%		16 12.91%	
Priority: purchase from the urban markets of the city (Group 2)	84	14 16.67%		70 83.33%	
Percentage of total correct classified observation		85.58			
Kappa Statistic		0.702			
Accuracy statistics based on class	TP rate	FP rate	Precision	Recall	Balanced F-score
Group 1	0.871	0.167	0.885	0.871	0.878
Group 2	0.833	0.129	0.814	0.833	0.824
Total	0.856	0.151	0.856	0.856	0.856

As shown in Fig. 1, the most effective socio-economic variable on buyers' choices is their awareness of the discounts in the urban markets, and after that, gender and household income are in the following order. The next degrees of significance are given to the factors of distance from the market, previous purchases from urban markets, and age.

In the DT with the C4.5 algorithm, the first number of each leaf node[1] represents consumer priority in choosing the market. It takes 1 if the consumers prefer buying from other markets, and it takes 2 if they prefer buying from the urban markets. The first number in parentheses in each leaf node represents how much data exists in the represented sample, and the second number illustrates the number of errors in the classification of buyers.

According to the classification that has been done based on the C4.5 algorithm of the DT, around 40% of buyers are women who were not aware of the urban markets' price discounts,

1- Each decision tree has two kinds of nodes: internal nodes and leaf nodes. Each internal node or non-leaf node is illustrated by an attribute in a way that represents a "test" on an attribute, and there are branches equal to the number of possible answers, and each of them shows the outcome of a test. Each leaf node also represents a class or group of answers.

were far from the urban markets, and had never purchased from these markets. These buyers will purchase from shops near their place of residence.

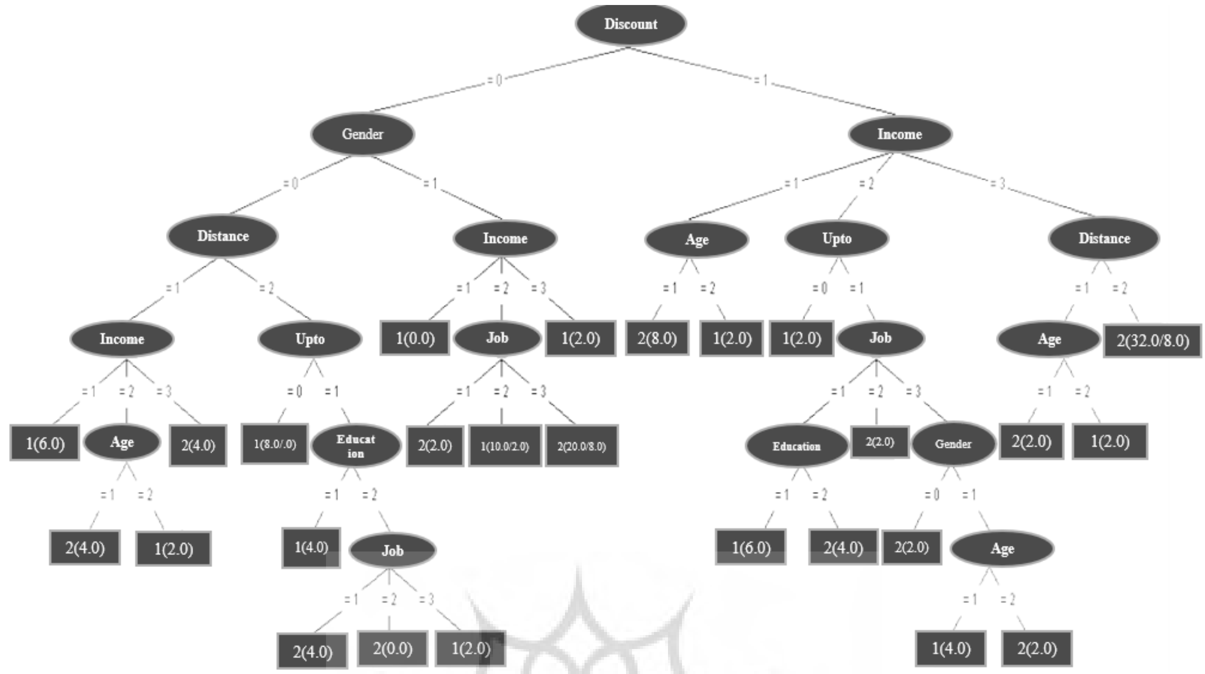

Figure 2- **C4.5 algorithm of buyers' general behavior in choosing among different markets**

According to Fig. 2, another group that makes up a high proportion of buyers is middle-income men who are unaware of discounts in the urban markets. In this group, those who work as employees choose the urban markets, while those with self-employed jobs choose other shops. Buyers who have an average income level and are aware of the discount in the urban markets do not buy from these markets if they do not have experience buying from them. However if they do have the experience, purchasing or not purchasing from the urban markets depends on other variables, such as job, gender, education, and age, which are explained in Fig. 1. Greenacre & Akbar (2019); Daoudi & Develi (2022) also stated that age and income play a significant role in influencing consumer buying behavior. The noteworthy point about this tree is how the middle-income group, as opposed to the low-income and high-income groups, follows a more complicated algorithm from all the branches that lead to household income. Planners and policymakers of urban marketplaces should consider this since this group makes up the majority of the population. Therefore, it is crucial to take additional factors such as age, gender, occupation, and education into account for the middle-income group. Investigating this tree also illustrates that people with a higher level of awareness regarding discounts are more likely to purchase from these markets. For a better understanding of consumer behavior, this behavior is discussed in different groups based on gender, education, and distance. Fig. 3 and 4 represent male and female consumers' behavior, respectively, and Table 4 represents the information accuracy of the DT represented in that figure. As shown in Table 4, the DT could accurately classify female consumers' purchasing patterns in 87 percent of cases, while male consumers' purchasing patterns were correctly classified in 86 percent of cases. Kol & Levy (2023); Chen *et al.* (2022); Savaşkan & Çatı (2021) stated that female customers are more interested in discounted products and gender difference affects the purchasing pattern. Furthermore,

Table 4- Consumer classification based on gender and purchase priority using the C4.5 algorithm

Consumers	**Number of observations**		**Results of decision tree C4.5 algorithm classification**			
			Priority: purchasing from other markets		**Priority: purchasing from urban markets of city**	
	Female	Male	Female	Male	Female	Male
Priority: purchasing from other shops (Group 1)	96	28	88	24	8	4
			91.67%	85.71%	8.33%	14.28%
Priority: purchasing from urban markets of city (Group 2)	56	28	12	4	44	24
			21.42%	14.28%	78.57%	85.71%
	Female			Male		
Percentage of total correct classified observation	86.84%			85.71%		
Kappa Statistic	0.715			0.714		

Accuracy statistic based on class	TP rate		FP rate		Precision		Recall		Balanced F-score	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Group 1	0.917	0.857	0.214	0.143	0.880	0.857	0.917	0.857	0.898	0.857
Group 2	0.786	0.857	0.083	0.143	0.846	0.857	0.786	0.857	0.815	0.857
Total	0.868	0.857	0.166	0.143	0.868	0.857	0.868	0.857	0.867	0.857

The DT from the analysis of female consumer behavior is shown in Fig. 3. It demonstrates that for this group of consumers, awareness of the discounts is the most significant factor, and then the variables of distance from the urban markets and education are on the second level of effectiveness. The third most effective variable in influencing consumers' choices is income and previous purchases from the urban markets. In the case of the behavior analysis of male consumers, Fig. 4 illustrates that the job is the most significant variable, followed by income and previous purchases from the urban markets. Age, distance from the urban markets, and discount awareness are on the third level of effectiveness. Gomes (2018) emphasizes that education level, gender, and occupation profoundly impact individuals' shopping behavior patterns, playing pivotal roles in shaping consumer preferences, decision-making processes, and shopping habits across diverse demographic groups. Furthermore, Taylor *et al.* (2019) and Chandrakala *et al.* (2023) highlighted that discounts are a significant factor influencing consumer purchasing behavior. Šostar & Ristanović (2023) highlight that age, gender, and employment status are influential factors in shaping consumer behavior.

According to Fig. 3, if women are aware of the discount offered by the urban markets and have a high level of education, they purchase from the urban markets, but if they have a low level of education and a medium level of income, they will purchase from an alternative market. If planners want to draw in more consumers, they should take into account the significant impact of discount awareness on women's purchasing habits. Female consumers will not purchase from urban markets if they do not know about the discount, the distance between their place of residence and these markets is far, and they have not previously purchased from these markets. Planners must adopt and carry out remedial policies to address this problem.

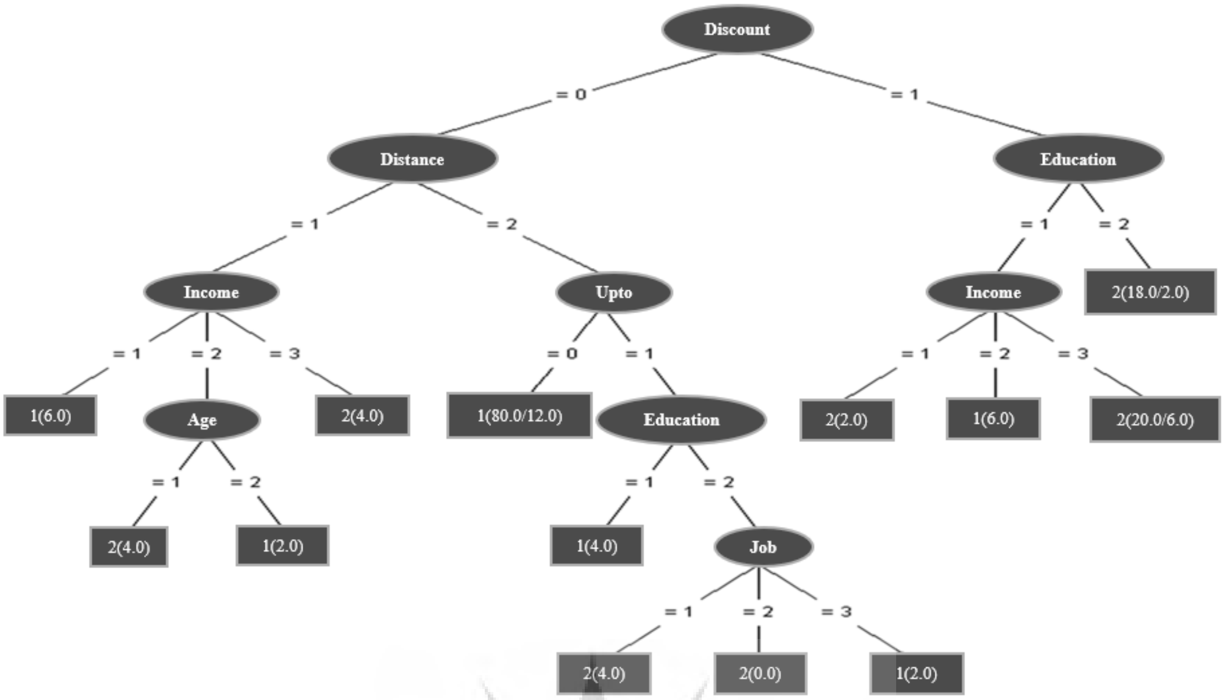

Figure 3- **C4.5 algorithm of female buyers' buying behavior in choosing between different alternatives of market**

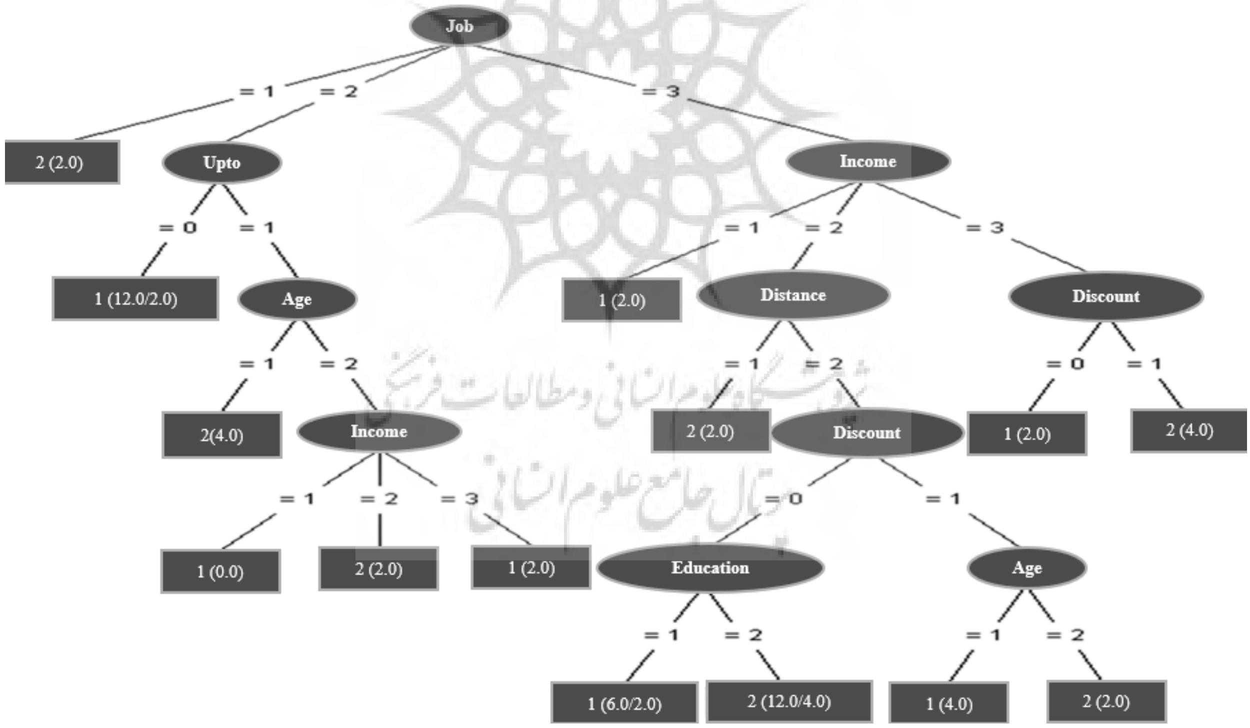

Figure 4- **C4.5 algorithm of male buyers' buying behavior in choosing between different market alternatives**

Regarding male consumers, Fig. 4 illustrates that those who are self-employed and have not previously shopped at urban markets tend to prefer other shops. However, employed men purchase from urban markets in two scenarios: first, if they have a high income and are aware of market discounts; and second, if they have a moderate income and reside near urban markets. Interestingly, they also shop at urban markets when they live far away, possess a high level of education, and are unaware of market discounts. In line with these findings, Roy *et al.* (2021) suggested that local retailers' spot discounts could enhance market potential and influence consumer preferences toward purchasing. Table 5 discusses the behavior of consumers with and without university degrees. It indicates that the C4.5 algorithm of the DT could effectively classify consumers based on socioeconomic factors and simulate consumer behavior in two groups, including those with and without a university degree, to a great extent. Fig. 5 and 6 display the classification's outcomes.

Table 5- Consumer classification based on education and purchase priority using the C4.5 algorithm

Consumers	Number of observations		Results of decision tree C4.5 algorithm classification			
			priority: purchase from other shops		priority: purchase from urban markets	
	High[1]	Low	High	Low	High	Low
priority: purchase from other shops (Group 1)	32	90	26	86	6	4
			81.25%	95.56%	18.75%	4.44%
priority: purchase from urban markets (Group 2)	36	48	4	12	32	36
			11.11%	25.00%	88.89%	75.00%
	High			Low		
Percentage of total correct classified observation	85.71%			88.41%		
Kappa Statistic	0.714			0.734		

Accuracy statistics based on class	TP rate		FP rate		Precision		Recall		Balanced F-score	
	High	Low	High	Low	High	Low	High	Low	High	Low
Group 1	0.824	0.956	0.111	0.250	0.875	0.878	0.824	0.956	0.848	0.915
Group 2	0.889	0.750	0.176	0.044	0.842	0.900	0.889	0.750	0.865	0.818
Total	0.857	0.884	0.145	0.179	0.858	0.885	0.857	0.884	0.857	0.881

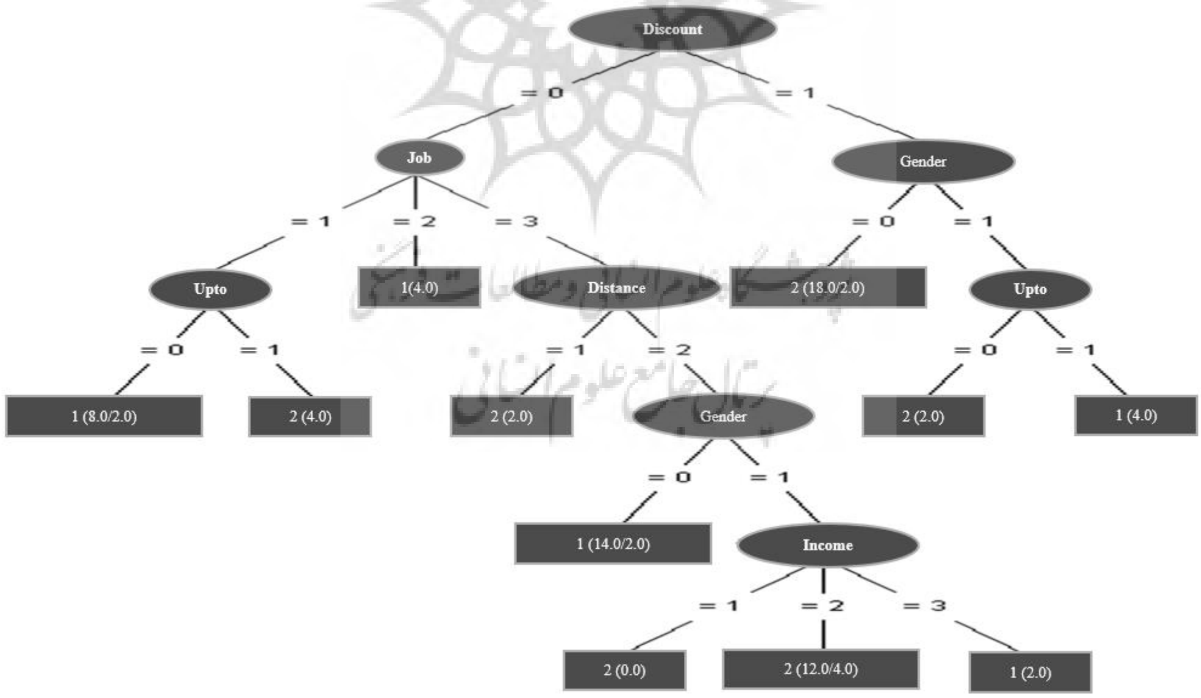

Figure 5- C4.5 algorithm of buyers' behavior in choosing different alternatives of the market who do not have a university degree

1- High means people with university degree and low means people with low degree

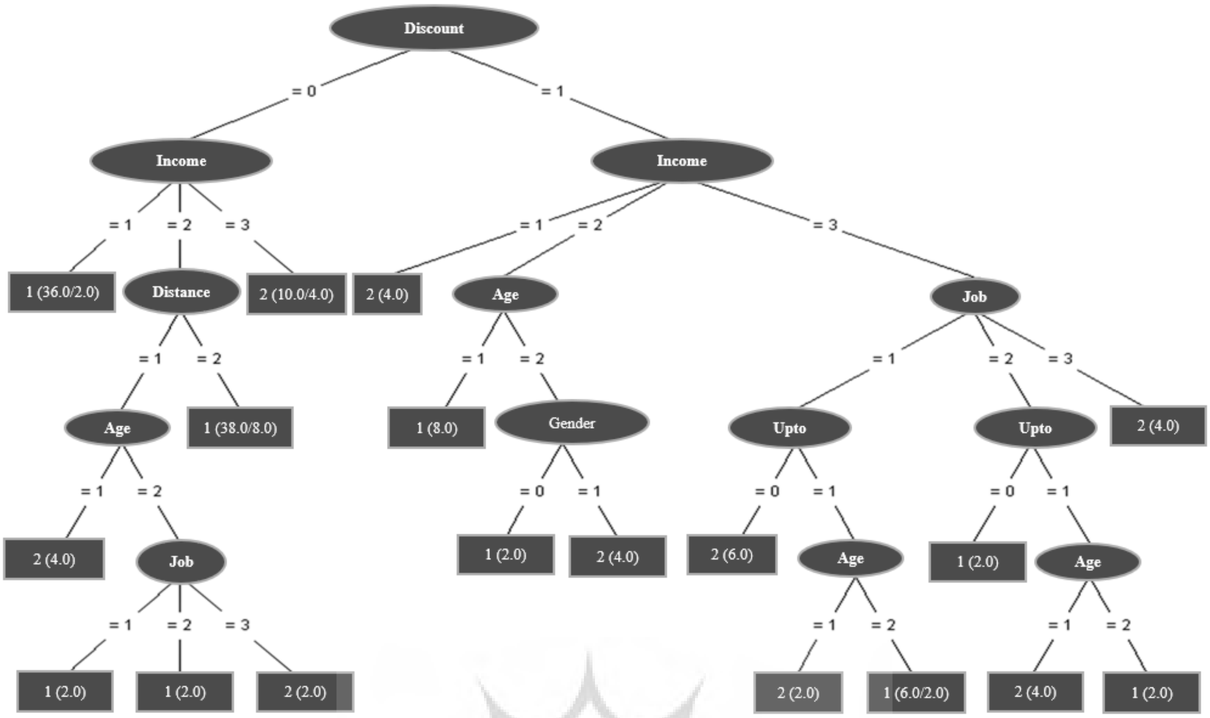

Figure 6- **C4.5 algorithm of buyers' behavior in choosing different alternatives of the market who have a university degree**

As it is shown in Fig. 5 and 6, the awareness of the urban market's discount is the most important component that influences consumers' choice among market options in both groups with and without a university degree.

In the group of consumers without university degrees, the second most essential criterion for both consumers with and without awareness of urban markets' discounts is their income. They are less likely to purchase from the urban markets if they are unaware of discounts and have low incomes. This is also true for consumers with a medium level of income who live far from the urban markets. Finally, since the distance from the urban markets is one of the important factors influencing consumer behavior, planners in this area can affect consumers' choices by selecting the right site.

Fig. 7, illustrates socio-economic variables that are effective on the behavior of consumers at a far and close distance from the urban markets. Prior to that, Table 6 provides information regarding different statistics that must be used to assess the accuracy of fitted trees. The DT correctly classified 82% of the consumers. Accuracy statistics based on class also illustrate the appropriate accuracy of the represented DT. Zulqarnain *et al.* (2015) also identified distance as a critical factor influencing consumers' choice of retail stores for grocery shopping. Terano *et al.* (2015), Yildirim *et al.* (2015) stated that significant differences in customer preferences based on age, gender, and education level.

As shown in Fig. 7 and 8, the consumer's decision to shop at the urban markets first depends on the discount offered at these markets and the consumers' awareness of that. One of the most crucial paths that planners should take into account is the path of women who are not aware of urban markets' discounts, live far away from the urban markets, and have never made purchases there before. These characteristics have prevented these individuals from choosing the urban markets over other markets. However, males living a far distance from the urban markets with a medium level of income who are not aware of discounts will buy from the urban markets if they have university degrees; otherwise, they will purchase from other shops. Therefore, the implementation of a specific marketing strategy for these groups can lead to an increase in their purchases from these markets.

Table 6-Consumer classification based on distance from the urban markets and purchase priority using C4.5 algorithm

Consumers	Number of observations		Results of decision tree C4.5 algorithm classification			
			Priority: purchase from other shops		Priority: purchase from urban markets	
	Close	Far	Close	Far	Close	Far
Priority: purchase from other markets (Group 1)	14	110	12	98	2	12
			85.71%	89.09%	14.28%	10.90%
Priority: purchase from urban markets (Group 2)	18	66	4	18	14	48
			22.23%	27.27%	77.77%	72.73%
			Close		Far	
Percentage of total correct classified observation			81.25%		82.95%	
Kappa Statistic			0.625		0.630	

Accuracy statistic based on class	TP rate		FP rate		Precision		Recall		Balanced F-score	
	Close	Far	Close	Far	Close	Far	Close	Far	Close	Far
Group 1	0.857	0.891	0.222	0.273	0.750	0.845	0.857	0.891	0.800	0.867
Group 2	0.778	0.727	0.143	0.109	0.875	0.800	0.778	0.727	0.824	0.762
Total	0.813	0.830	0.178	0.211	0.820	0.828	0.813	0.830	0.813	0.828

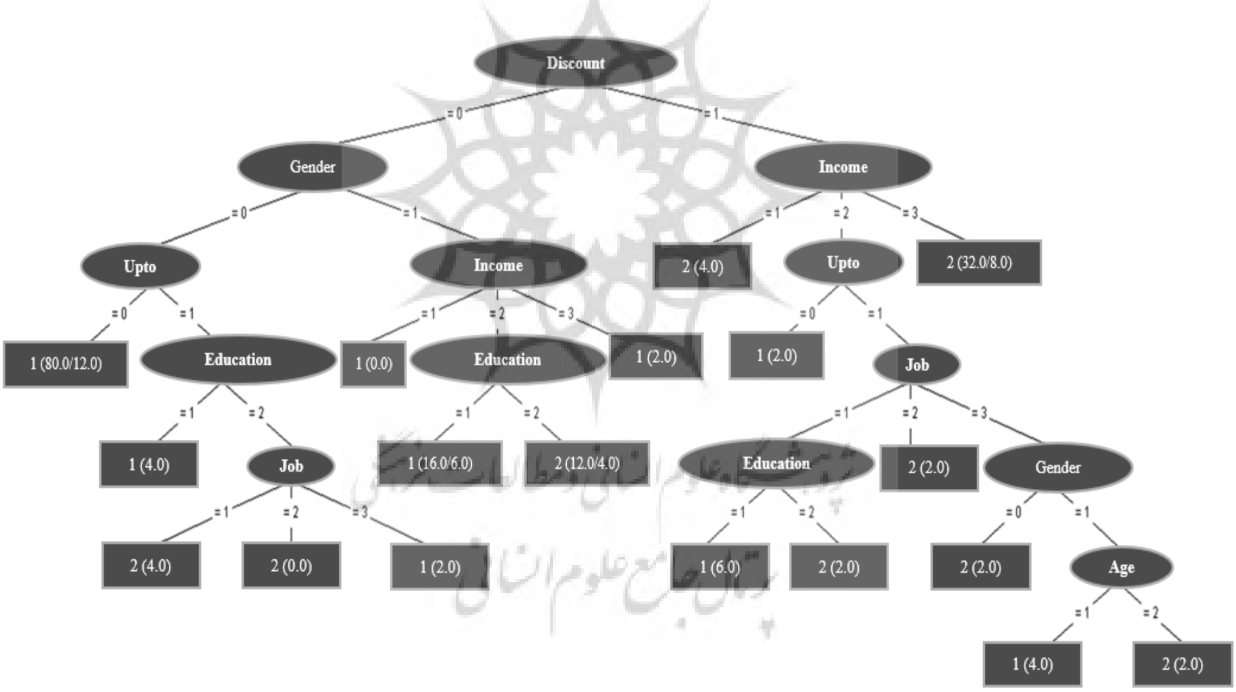

Figure 7- C4.5 algorithm of buyers' behavior with far distance from urban markets

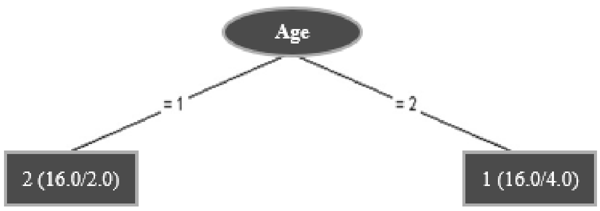

Figure 8- C4.5 algorithm of buyers' behavior at close distance from urban markets

Another considerable point about the distance from the urban markets is the fact that for those with a close distance from the urban markets, their choice between the urban markets and other shops only depends on age. Those who are less than 40 years old prefer urban markets, while others do not. Makgosa & Sangodoyin (2017) highlighted age as a critical determinant of consumers' store preferences. They categorized younger shoppers as recreational quality seekers and older shoppers as novelty–quality seekers.

Conclusion and Recommendation

This study used the C4.5 algorithm, one of the most accurate decision tree algorithms, to survey consumers and model their decision processes. In accordance with previous studies, this study demonstrates that C4.5 is a helpful tool for creating a hierarchical decision support model. It provides a simple schematic diagram of the consumers' decision-making process. This approach allows for the statistical testing of model validation. This enables the model's reliability to be taken into account. This study's decision tree has acceptable reliability. The findings highlight that consumer behavior is predominantly influenced by awareness of discounts, aligning closely with Alsini *et al.*'s (2023) findings on the impact of promotional campaigns on grocery shopping in Saudi Arabia. Discounts play a significant role in retail dynamics by boosting sales and enhancing customer satisfaction but also present challenges such as reduced profit margins and increased price sensitivity. Less than a third of consumers are aware of available discounts, suggesting deficiencies in current marketing strategies and the need for reassessment. Hecht *et al*. (2020) emphasized the importance of regularly reviewing marketing tactics to inform consumers about discounts effectively. In addition, Noor (2020) emphasized the significant impact of discounts on consumer buying behavior. Büyükdağ *et al.* (2020) also investigated the varying effects of discounts on men and women, highlighting that gender differences can influence how consumers respond to promotional offers.

The study also revealed that the majority of consumers in these markets have advanced degrees, likely due to policies targeting educated individuals or a lack of advertising that has inadvertently attracted this demographic. Conversely, those with lower education levels are less inclined to shop at these markets, likely due to factors such as income disparity, limited exposure to marketing, and less familiarity with urban retail markets. Hanaysha (2018) confirms that income is a key factor influencing purchase decisions. Kumar (2018) emphasized that comprehensive product offerings, convenience, and parking facilities contribute to the preference for urban retail markets. Current policies favoring high-income consumers highlight the need for equitable policy reviews to broaden market accessibility. Customer retention strategies, such as periodic discounts or conditional sales, can significantly influence subsequent consumer behavior and retention. In conclusion, this study provides valuable insights into consumer behavior within urban retail markets but acknowledges its limitations. Future research should explore comprehensive datasets, additional socio-demographic variables, and the long-term impact of marketing interventions. Based on these findings, the following suggestions are presented for policy-making and implementation:

1. Enhanced Promotional Campaigns: Enhanced promotional campaigns can effectively utilize social media and online advertising to raise awareness about ongoing discounts and special offers.
2. Education-oriented strategies: Collaborate with community centers and educational institutions to host workshops informing consumers about the benefits of urban retail markets. Furthermore, simplify access to discount details and product offerings through user-friendly apps or strategically placed informative brochures across the city.
3. Loyalty Programs: Introduce a loyalty

program that offers points or discounts to repeat customers, along with exclusive promotions, to attract more consumers to enroll and actively engage in the market.

4. Collaborations with Local Producers: Strengthen partnerships with local producers to ensure a consistent supply of high-quality products. Highlight these partnerships in marketing campaigns to emphasize the support for local businesses and fresh, local produce.
5. Inclusive Marketing Policies: Implement pricing strategies that cater to diverse income levels. Consider introducing tiered pricing or discount days tailored for low-income families to enhance affordability. Additionally, collaborate with community leaders and organizations to gain insights into the preferences of lower-income and less-educated consumers. Use this feedback to inform and adjust marketing and operational strategies accordingly.

By adopting these strategies, urban retail markets in Mashhad can enhance their appeal, attract a more diverse consumer base, and ensure sustainable growth. These efforts will not only boost consumer welfare but also support local producers and contribute to the overall economic development of the city.

References

1. Alsini, N., Kutbi, H.A., Hakim, N., Mosli, R., Eid, N., & Mulla, Z. (2023). Factors influencing grocery shopping choices and the prevalence of food label use among Saudi mothers: a cross-sectional pilot study. *Nutrition & Food Science, 53*(2), 432-444. https://doi.org/10.1108/NFS-11-2021-0345
2. Applebaum, W. (1951). Studying customer behavior in retail stores. *Journal Market,* 172–178. https://doi.org/10.1177/002224295101600204
3. Barh, D. (2020). *Artificial Intelligence in Precision Health: From Concept to Applications*. Academic Press. https://doi.org/10.1016/B978-0-12-817133-2.09988-2
4. Berry, M.A., & Linoff, G.S. (2000). Mastering data mining: The art and science of customer relationship management. *Industrial Management & Data Systems*. https://doi.org/10.1108/imds.2000.100.5.245.2
5. Bhatti, K.L., Latif, S., & Latif, R. (2015). Factors affecting consumer's store choice behavior. *International Journal of Application or Innovation in Engineering & Management (IJAIEM), 4*(9), 71-77.
6. Büyükdağ, N., Soysal, A.N., & Kitapci, O. (2020). The effect of specific discount pattern in terms of price promotions on perceived price attractiveness and purchase intention: An experimental research. *Journal of Retailing and Consumer Services, 55*, 102112. https://doi.org/10.1016/j.jretconser.2020.102112
7. Chandrakala, V.G., Sowmya, C.U., & Nagesha, H.G. (2023). A study on factors influencing the consumer buying behavior with reference to organized apparel retail outlets. *Journal of Innovations in Business and Industry, 1*(2), 85-92. https://doi.org/10.61552/JIBI.2023.02.005
8. Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y. (2022). The impact of online reviews on consumers' purchasing decisions: Evidence from an Eye-Tracking study. *Frontiers in Psychology, 13*, 865702. https://doi.org/10.3389/fpsyg.2022.865702
9. Cherfi, A., Nouira, K., & Ferchichi, A. (2018). Very fast C4.5 decision tree algorithm. *Applied Artificial Intelligence, 32*(2), 119-137. https://doi.org/10.1080/08839514.2018.1447479
10. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104
11. Daoudi, H., & İldem Develi, E. (2023). Impact of gender and monthly income on consumer buying behavior. *Journal of International Trade, Logistics and Law, 9*(1), 241-250.
12. Dev, V.A., & Eden, M.R. (2019). Gradient boosted decision trees for lithology classification. *In Computer Aided Chemical Engineering, 47*, 113-118. https://doi.org/10.1016/B978-0-12-818597-

1.50019-9

13. Dominici, A., Boncinelli, F., Gerini, F., & Marone, E. (2021). Determinants of online food purchasing: The impact of socio-demographic and situational factors. *Journal of Retailing and Consumer Services*, *60*. https://doi.org/10.1016/j.jretconser.2021.102473
14. Donoghue, S., Wilken-Jonker, I., Steffens, F.E., & Kirsten, J.F. (2021). South African consumers' willingness to pay a premium for Karoo Lamb: The influence of subjective and objective knowledge, label information and demographics. *Journal of Retailing and Consumer Services*, *63*. https://doi.org/10.1016/j.jretconser.2021.102664
15. Gauri, D.K., Jindal, R.P., Ratchford, B., Fox, E., Bhatnagar, A., Pandey, A., & Howerton, E. (2021). Evolution of retail formats: Past, present, and future. *Journal of Retailing*, *97*(1), 42-61. ISSN 0022-4359. https://doi.org/10.1016/j.jretai.2020.11.002
16. Golriz Ziaie, Z., Moghaddasi, R., & Yazdani, S. (2015). Estimation of customer satisfaction index of food markets, Case study: Mashhad Urbanity's Hypermarkets. *Journal of Agricultural Economics and Development*, *29*(2), 181-191. ISSN 2008-4722. https://doi.org/10.22067/JEAD2.V0I0.40763
17. Gomes, A. (2018). Influencing factors of consumer behavior in retail shops available at SSRN: https://ssrn.com/abstract=3151879 or http://dx.doi.org/10.2139/ssrn.3151879
18. Greenacre, L., & Akbar, S. (2019). The impact of payment method on shopping behavior among low income consumers. *Journal of Retailing and Consumer Services*, *47*, 87-93. https://doi.org/10.1016/j.jretconser.2018.11.004
19. Han, J., Kamber, M., & Pei, J. (2022). Data mining: concepts and techniques, Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-381479-1.00020-4
20. Hanaysha, J. (2018). An examination of the factors affecting consumer's purchase decision in the Malaysian retail market. *PSU Research Review*, *2*(1), 7-23. https://doi.org/10.1108/PRR-08-2017-0034
21. Hecht, A.A., Perez, C.L., Polascek, M., Thorndike, A.N., Franckle, R.L., & Moran, A.J. (2020). Influence of food and beverage companies on retailer marketing strategies and consumer behavior. *International Journal of Environmental Research and Public Health*, *17*(20), 7381. https://doi.org/10.3390/ijerph17207381
22. Hingley, M., Lindgreen, A., & Chen, L. (2009). Development of the grocery retail market in China: A qualitative study of how foreign and domestic retailers seek to increase market share. *British Food Journal*, *111*(1), 44-55. https://doi.org/10.1108/00070700910924227
23. Jofreh, M. (2013). An investigation of business activities in Iran retailing industry. *Research Journal of Applied Sciences, Engineering and Technology* *6*(5), 858-861.
24. Kol, O., & Levy, S. (2023). Men on a mission, women on a journey - Gender differences in consumer information search behavior via SNS: The perceived value perspective. *Journal of Retailing and Consumer Services*, *75*, 103476. https://doi.org/10.1016/j.jretconser.2023.103476
25. Kotsiantis, S.B. (2013). Decision trees: a recent overview. *Artificial Intelligence Review*, *39*(4), 261-283. https://doi.org/10.1007/s10462-011-9272-4
26. Kotu, V., & Deshpande, B. (2018), Chapter 4 – Classification, Data science: concepts and practice (Second Edition), Morgan Kaufmann, 65-163.
27. Kumar, D. (2018). Factors influencing consumers' choice of retail store format in Assam, India. *Journal of Marketing Management*, *1*(17), 22-37.
28. Kumar, M.G., Hemanth, K., Gangadhar, N., Kumar, H., & Krishna, P. (2014). Fault diagnosis of welded joints through vibration signals using Naïve Bayes algorithm. *Procedia Materials Science*, *5*, 1922-1928. https://doi.org/10.1016/j.mspro.2014.07.514
29. Laine, K. (2014). The factors influencing the choice of grocery store among finnish consumers. Business. ttps://urn.fi/URN:NBN:fi:amk-201405096933
30. Lintang, M., Pandiangan, N., & Hyronimus, D. (2022). Use of the C4.5 Algorithm to Analyze Student Interest in Continuing to College. SHS Web of Conferences.

31. Liyanage, L., PLGSD, P., & Rathnayake, T. (2020). Determinants of consumers' selection of supermarkets for grocery shopping; Empirical Evidence from Western Province, Sri Lanka. In International Conference on Marketing Management.
32. Makgosa, R., & Sangodoyin, O. (2017). Retail market segmentation: the use of consumer decision-making styles, overall satisfaction and demographics. *The International Review of Retail, Distribution and Consumer Research, 28*(1), 64–91. https://doi.org/10.1080/09593969.2017.1334690
33. Manuere, F. (2023). Factors affecting customers' choice of supermarkets for grocery shopping in Chinhoyi Town. *International Journal of Academic Research in Public Policy and Governance, 9*(1), 10–20
34. McHugh, M.L. (2012), Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276-282.
35. Meng, X., Zhang, P., Xu, Y., & Xie, H. (2020). Construction of decision tree based on C4.5 algorithms for online voltage stability assessment. *International Journal of Electrical Power & Energy Systems, 118,* 105793. https://doi.org/10.1016/j.ijepes.2019.105793
36. Moitra, A.K., Bhattacharya, J., Kayal, J.R., Mukerji, B., & Das, A.K. (2021). Innovative exploration methods for minerals, oil, gas, and groundwater for sustainable development. https://doi.org/10.1016/C2020-0-00590-6
37. Nguyen, T.D.E. (2019). Factors affecting customer loyalty of different strategic groups in the Vietnamese supermarket sector (Doctoral dissertation, University of hull).
38. Nisbet, R., Miner, G.D., & Yale, K. (2018). Chapter 9 – Classification, Nisbet, R., Miner, G. D., & Yale, K., Handbook of statistical analysis and data mining applications. 169-186 in Academic press.
39. Noor, Z. (2020). The effect of price discount and in-store display on impulse buying. *Journal Ilmu-ilmu Sosial dan Humaniora. 22*(2), 133-139. https://doi.org/10.24198/sosiohumaniora.v22i2.26720
40. Ooi, M.P., Sok, H.K., Kuang, Y.C., & Demidenko, S. (2017). Alternating decision trees. In P. Samui, S. S. Roy, & V. E. Balas (Eds.), Handbook of Neural Computation. PP 345-371 in Academic Press.
41. Pandey, A., & Kaur, D.A. (2018). A comprehensive study on evolution, present scenario and future prospects of retailing. *International Journal of Current Research in Life Sciences, 7*(2), 1158-1162.
42. Quinlan, J.R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Research, 4,* 77–90. https://doi.org/10.48550/arXiv.cs/9603103
43. Quinlan, J.R. (2014). C4.5: Programs for Machine Learning. 58-60. https://books.google.com/books/about/C4_5.html?id=b3ujBQAAQBAJ
44. Rahim, M.A., Mushafiq, M., Khan, S., & Arain, Z.A. (2021). RFM-based repurchase behavior for customer classification and segmentation. *Journal of Retailing and Consumer Services, 61*. https://doi.org/10.1016/j.jretconser.2021.102566
45. Ram, S. (2022). Data Mining. Computer Sciences. from Encyclopedia.com: https://www.encyclopedia.com/computing/news-wires-white-papers-and-books/data-mining
46. Rasheed, A., Shahid Yaqub, R., & Baig, F. (2018). Factors affecting impulse buying behaviors in shopping malls: Evidence from Bahawalpur Region, Pakistan. *Journal of Marketing and Consumer Research. 39.*
47. Razu, M.A., & Roy, D. (2019). Prominent factors influencing consumers' choice of retail store. *Journal of Business and Management, 21*(6), 62-66. https://doi.org/10.9790/487X-2106026266
48. Reddy, G.S., & Chittineni, S. (2021). Entropy based C4.5-SHO algorithm with information gain optimization in data mining. *Peer Journal Computer Science, 7*. https://doi.org/10.7717/peerj-cs.424

49. Roy, G., Debnath, R., & Mitra, P.S. (2021). Analytical study of low-income consumers' purchase behaviour for developing marketing strategy. *International Journal Syst Assur Engineering Management*, *12*, 895–909. https://doi.org/10.1007/s13198-021-01143-6
50. Savaşkan, A., & Çatı, K. (2021). Investigation of consumer behavior in market shopping in the gender context. *Elektronik Sosyal Bilimler Dergisi*, *20*(77), 255-272. https://doi.org/10.17755/esosder.767017
51. Shamsher, R. (2016). Store image and its impact on consumer behavior. *Elk Asia Pacific Journal of Marketing and Retail Management*, *7*, (2), 1-27. ISSN 0976-7193. https://www.researchgate.net/project/STORE-IMAGE-AND-ITS-IMPACT-ON-CONSUMER-BEHAVIOR
52. Šostar, M., & Ristanović, V. (2022). Assessment of influencing factors on consumer behavior using the AHP model. *Sustainability*, *15*(13), 10341. https://doi.org/10.3390/su151310341
53. Sugumaran, V., & Ramachandran, K.I. (2007). Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical Systems and Signal Processing*, *21*, 2237–2247. https://doi.org/10.1016/j.ymssp.2006.09.007
54. Sun, K., Likhate, S., Vittal, V., Kolluri, V.S., & Mandal, S. (2007). An online dynamic security assessment scheme using phasor measurements and decision trees. *IEEE Transactions on Power Systems*, *22*(4), 1935-1943. https://doi.org/10.1109/TPWRS.2007.908476
55. Taylor, M.J., Kwasnica, V., Reilly, D., & Ravindran, S. (2019). Game theory modelling of retail marketing discount strategies. *Marketing Intelligence and Planning*, *37*(5), 555-566. ISSN 0263-4503.
56. Terano, R., Yahya, R., Mohamed, Z., & Saimin S. (2015). Factor influencing consumer choice between modern and traditional retailers in Malaysia. *International Journal of Social Science and Humanity*, *5*(6). https://doi.org/10.7763/IJSSH.2015.V5.509
57. Thakur, P., Mehta, P., Devi, C., Sharma, P., Singh, K.K., Yadav, S., Lal, P., Raghav, Y.S., Kapoor, P., & Mishra, P. (2023). Marketing performance and factors influencing farmers choice for agricultural output marketing channels: The case of garden pea (*Pisum sativum*) in India. *Frontiers in Sustainable Food Systems*, *7*, 1270121. https://doi.org/10.3389/fsufs.2023.1270121
58. Tian, X., Cao, S., & Song, Y. (2021). The impact of weather on consumer behavior and retail performance: Evidence from a convenience store chain in China. *Journal of Retailing and Consumer Services*, *62*. https://doi.org/10.1016/j.jretconser.2021.102583
59. Veeck, A., & Veeck, G. (2000). Consumer segmentation and changing food purchase patterns in Nanjing", PRC. *World Development*, *28*(3), 457-471. https://doi.org/10.1016/S0305-750X(99)00135-7
60. Vindigni, G., Peri, I., Consentino, F., Selvaggi, R., & Spina, D. (2022). Exploring consumers attitudes towards food products derived by new plant breeding techniques. *Sustainability*, *14*(10). https://doi.org/10.3390/su14105995
61. Walkinshaw, N. (2013). Reverse-engineering software behavior, Memon, A., Advances in Computers PP 1-58 in Elsevier.
62. Yildirim, K., Cagatay, K., & Hidayetoğlu, M.L. (2015). The effect of age, gender and education level on customer evaluations of retail furniture store atmospheric attributes. *International Journal of Retail & Distribution Management*, *43*(8), 712-726. https://doi.org/10.1108/IJRDM-01-2013-0034
63. Zaki, M.J., & Meira, W. (2020). *Data mining and machine learning: Fundamental concepts and algorithms*. Cambridge University Press.
64. Zulqarnain, H., Zafar, A.U., & Shahzad, M. (2015). Factors that affect the choice of consumers in selecting retail store, for grocery shopping. *International Journal of Multidisciplinary and Current Research*, *3*(1), 1167-1172S. http://ijmcr.com

Journal of Agricultural Economics & Development
https://jead.um.ac.ir

مقاله پژوهشی
جلد ۳۸، شماره ۴، زمستان، ۱۴۰۳، ص. ۳۷۰-۳۵۱

رویکرد داده کاوی برای انتخاب بازار خرده‌فروشی مصرف‌کنندگان: مطالعه موردی بازارهای خرده‌فروشی شهری در ایران

هومن هاشم‌زاده[۱]- نازنین یوسفیان[۲]- ساسان اسفندیاری بهرآسمان[۲]- علی‌رضا کرباسی[۳]- علی فیروززارع[۴*]

تاریخ دریافت: ۱۴۰۳/۰۲/۳۱
تاریخ پذیرش: ۱۴۰۳/۰۴/۱۸

چکیده

بازارهای خرده‌فروشی شهری، بازارهای خرده‌فروشی دولتی هستند که اخیراً در ایران با هدف افزایش رفاه مصرف‌کنندگان و تولیدکنندگان تأسیس شده‌اند. برای دستیابی به این هدف و گسترش حضور در بخش خرده‌فروشی شهری، درک جامعی از رفتار مصرف‌کننده در این بازارها ضروری است. این مطالعه با استفاده از الگوریتم C4.5 به بررسی عوامل مختلف اجتماعی-اقتصادی مؤثر بر تصمیم‌گیری مصرف‌کنندگان در بازارهای خرده‌فروشی پرداخته است. داده‌ها از طریق نظرسنجی از ۱۸۹ مصرف‌کننده با استفاده از روش نمونه‌گیری تصادفی در شهر مشهد در سال‌های ۱۳۹۸ و ۱۳۹۹ جمع‌آوری شد. نتایج نشان داد که آگاهی از تخفیف‌های موجود به‌طور قابل توجهی انتخاب مصرف‌کننده را در بازارهای خرده‌فروشی شهری تحت تأثیر قرار می‌دهد. با این حال، علی‌رغم تخفیف‌های موجود، آگاهی در میان مصرف‌کنندگان پایین است که نیاز به بازنگری در استراتژی‌های تبلیغاتی می‌باشد. همچنین نتایج نشان داد تجربه خرید از بازارهای شهری، درآمد خانوار و تحصیلات از جمله عوامل تأثیرگذار بر انتخاب مصرف‌کننده می‌باشند. یافته‌های این مطالعه می‌تواند بینش‌های ارزشمندی برای سیاست‌گذاران و سهامداران فراهم آورد که در پی افزایش اثربخشی بازارهای خرده‌فروشی محلی در ایران هستند. علاوه بر این، با بهره‌گیری از این بینش‌ها در زمینه رفتار مصرف‌کننده و پویایی بازار، این بازارها می‌توانند رونق گرفته و نقش بسزایی در بهبود بخش خرده‌فروشی و اقتصاد ایران ایفا کنند. در این راستا، توصیه‌هایی مانند کمپین‌های تبلیغاتی، استراتژی‌های آموزش‌محور، همکاری با تولیدکنندگان محلی و سیاست‌های بازاریابی فراگیر با هدف بهبود دسترسی همه مصرف‌کنندگان به بازارهای خرده‌فروشی شهری ارائه گردید.

واژه‌های کلیدی: داده کاوی، درخت تصمیم، رفتار مصرف کننده، یادگیری ماشین

۱، ۲، ۳ و ۴- به‌ترتیب دانشجوی کارشناسی ارشد دانشجوی دکتری استاد و استادیار، گروه اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران
(*- نویسنده مسئول: Email: firooz@um.ac.ir)

https://doi.org/10.22067/jead.2024.88151.1269