What can Romanian Supermarkets do in order to improve the Customer Experience? Recommendation Systems in Creating Personalized Offers

Cristina STOICESCU, Cristina STANCIU-MÂNDRULEANU
University of Economic Studies, Bucharest, Romania
kris_stoicesku@yahoo.com, cristina.stanciu@csie.ase.ro

Modern food retail stores have started to penetrate the Romanian market over 20 years ago and since then they have gained more than half of the total market share. More than 10 international brands have opened over 2500 stores creating a competitive environment where each company needs to develop new business strategies in order to achieve positive financial results. In order to gain new clients supermarkets are opening new stores every year, but in order to keep their actual clients they need to improve the customer experience. Introducing fidelity cards, loyalty points, instant awards, tasting sessions or moving in the online area are some of the actions taken by supermarkets in Romania, but none of them have yet developed a recommendation system that can propose personalized offers. This article describes a hybrid algorithmic model based on data mining and collaborative filtering techniques. The goal of the model is to recommend date specific products, similar products and cross-shopping cart products that fulfill the custom needs of each consumer. So at the end of the shopping session each client will receive an offer that includes specific products and customized discounts to his buying behavior. As already sustained by other markets, personalized offers will improve the customer experience and will increase the sales of supermarkets.
Keywords: *consumer behavior, food retailers, recommendation systems, customer experience*

1 Introduction

In the last decades, consumer behavior has changed drastically, becoming a focus point for all companies around the world. Consumers of the 21st century are more informed, the buying process being completed before a first interaction with sales, they are socially networked, more in-store purchases being abandoned due to negative online sentiment and less loyal, more customers being willing to try a new brand to get better customer service. The main causes for this phenomenon are the demographic, social and economic changes and the fast evolution of technology, which reshaped the consumer behavior. As a consequence, companies started to develop their own models in order to quantify consumers needs which have become more diverse and more custom, and to better predict their uncertain and changing behavior.

In Romania, the food retail market is becoming saturated, new stores being opened annually in order to satisfy the extended demands. In this context, in order to gain more market share, the food retail stores have realized the importance of gaining the loyalty of their customers.

The main goal of this paper is to describe a model that can anticipate the needs of consumers and generate personalized offers, which will increase the sales of the company. The model is a hybrid algorithm based on data mining and collaborative filtering techniques that in the end will create a personalized recommendation for each client. The offer includes three types of products: products that are specific to the client shopping cart, products bought by other clients with a similar shopping cart and specific day products. In this way, it is encouraged the purchase of more products from the same category, and the diversification of the current shopping cart. By increasing the number of the products sold, the financial results of the food retail companies will be improved.

Section 2 of the paper reviews some of the research literature concerning the evolution of consumers behaviour and the actual trends that shapes the food retail market in Romania. Section 3 presents the context needed for implementing the model that generates personalized recommendation, while section 4 describes in detail the main steps of model. Section 5 contains the conclusions and future research.

2 Literature survey

Consumer behavior is an emergent phenomenon that evolved along with human development. Thus, during prehistoric the human behavior occurred in a very limited way, people being grouped in small families with the only concern of surviving. According to Miller (2009), much later, people began to develop social skills that eventually led to the emergence of money, social status, wealth and ultimately to shaping consumer behavior. The study of consumer behavior is an interdisciplinary science that integrates information from psychology, sociology, social psychology, anthropology, artificial intelligence and economics, which currently has started to gain attention of all companies who want to survive in a competitive and evolving environment, including food retailers. [1]

Burke (2014) performed several studies regarding consumer behavior in supermarkets and he noted that 85% of the products purchased by a consumer don't change over time, economic agents having stable preferences for most of the bought products. [2] According to this, in order to increase its sales, a food retailer needs to primarily focus on quantity and not on variety. However, only a combination of both methods will determine in the end improved financial results. In Romania, supermarkets and hypermarkets are using bonus points and loyalty points programs to gain more market share, offering only general discounts or discount that are dedicated to an entire customer segment. Therefore, Romanian food retailers are focusing exclusively on the price strategy and not on loyalty programs based on personalized recommendations, programs already adopted and tested in other countries like USA or UK.

According to a study by Precima (2017) traditional marketing techniques are still dominating the food retail markets around the world, although companies are not very pleased with their results, 22% generating zero benefits for the companies questioned in 2017. Traditional techniques aim to identify the customer segments that will positively react to an already existing general promotion, while current trends suggest the necessity to create personalized recommendations based on the needs and behavior of each consumer. [3]

During LEAD Marketing Conference, McVie (2017) spoke about the importance of personalized recommendations in food retail stores and how to set discounts according to their impact on business performance. There is a directly proportional relationship between discounts, sales volume and profit, its intensity becoming negative once the discounts reach a 40% threshold, the optimal discount level being between 10% and 20% when all financial indicators are recording maximum rates. Discounts up to 20% influence loyal customers and generate high sales volume, while discounts over 30% primarily affect customers who follow exclusively offers and generate small earnings. Discounts over 40% are generating loss. [4]

In conclusion, personalized discounts should fall within a range of 5% to 20%, focusing primarily on loyal customers. In this regard, Romanian food retailers should introduce personalized offers to gain customer loyalty and competitive advantages, by integrating recommendation systems in their business strategy. According to Isinkaye et al. (2015), recommendation systems assume data collection, data processing and adaptation of data filtering techniques, finally generating a recommendation, a feedback loop being present between these stages. Recommender systems may use different approaches like collaborative filtering, content-based filtering or hybrid techniques. [5]

According to Leskovec et al. (2014) content-based filtering techniques have the advantage of being independent of other consumer choices, and if a consumer's preferences change the recommendation system can be adapted in a very short time. The main disadvantage is that products must have similar content because the data are collected for the same types of variables and in the absence of data the results can be disruptive. [6] According to Herlocker et al. (2004), collaborative filtering techniques are mainly used when heterogeneous products with different characteristics need to be selected. Collaborative filtering techniques start from the assumption that users with similar shopping carts have common preferences, so recommendations will be made on the basis of similarities between consumers, not products. [7]

In 2006, Netflix launched a contest to improve its recommendation system named Cinematch. At that time, the recommendation system relied on content-based filtering techniques, using as input the ranking results and reviews of users. Recently, in order to generate personalized recommendations for each consumer, user profile analysis techniques based on historical activity have been also integrated into the system (Gomez-Uribe and Hunt, 2016). [8]

The large volume of information has also become a problem for social networks like Facebook, selecting the right information that might caught consumers attention becoming a real challenge. Baatarjav et al. (2008) proposed a recommendation system for group socialization on Facebook by integrating hierarchical cluster analysis techniques and decision trees. According to their results, the group recommendation system has an accuracy of 73%, using as main data the consumers profile. [9]

Recommendation systems are migrating from e-commerce to brick-and-mortar industries, becoming a solution for companies that sell a large range products and want to obtain a stable market share and gain customer loyalty.

3 Hypotheses and Prerequisites

The goal of the model is to generate personalized offers that will satisfy the specific needs of each consumer. In order to define and describe the model, the following points were taken into consideration:

- Who will receive the personalized offer – new or already existing customers;
- How to select and filter the recommended products – how many products are included in the recommendation and how are they prioritized within the recommendation;
- How to customize the additional information – the discounts, the validity or the newsletters;
- How to distribute the personalized offer – electronic or printed and when should it be delivered.

3.1. Customer selection

The model is based on a prescriptive system which analysis the historical data of consumers purchases in order to generate the personalized recommendations. Therefore, the first prerequisite is the presence of an instrument used to collect information regarding consumer behavior. In this sense, loyalty cards is the best solution. This marketing strategy already exists on the Romanian food retail market, but is used only for isolated situations.

Consumers that hold a loyalty card will be considered already existing customers for which a history of purchase can be analysed and recommendations can be generated, while consumers who do not have such a card will be considered new customers, although they might possibly do regular shopping. The offers generated by the model will be fully personalized and will not include general offers. In order to eliminate the risk that a general offer would be more attractive than the personalized one, negatively affecting the recommendation system, the company may cumulate the discounts or may introduce the general offers as variable in the model increasing significantly its complexity. According to the business strategy, the company can also decide to create personalized offers for new customers. This might happen if the company is launching on a new market, if it wants to build a stable database or seeks to gain a larger market share.

Food retail market in Romania is becoming saturate and the already existing stores have stable positions on the market so their main focus should be on gaining customer loyalty. The model will generate full personalized offers only for consumers who have a fidelity card, while general offers will primarily appeal to attracting new customers. The customers will be able to cumulate the discounts.

3.2. Product selection

The model splits the products into three categories, date specific products, specific shopping cart products and products different from the consumers shopping cart. For each category a different selection and filtering logic will be defined. The opportunity cost increases along with the number of products a consumer can choose from, which can make the buying decision very difficult. So, the total number of products included in the personalized offer must be limited, a larger discount for fewer products having a higher impact than a lower discount for multiple products.

The model will recommend every time 6 products for each consumer, the number of products specific to the shopping cart being directly proportional to the total number of products, thus having the highest weight. Only one product date specific and one product different from the shopping cart will be included in the offer. Products will be prioritized according to the category they belong to, so the first products will be those date specific and different from consumer shopping cart, while the other products will be those specific to the shopping cart.

3.3. Additional information selection

Discounts on birthdays or name days is one of the traditional marketing techniques used by food retailers. The model will generate a general discount of 15% on the customer's birthday and 5% on his name day. The discount will be applied only once for a full shopping cart within three days, including the celebration day. This discount will be included in the offer as a header separate section before all other discounts and will be announced in advance so it may be included in several successive recommendations. This discount will not affect the total number of products included in the offer which is a default configuration. The discount will be automatically granted once the customer is identified in the system and will be cumulated with all the other offers. If birthday and name day are in the same period, only the birthday discount will be generated, as its value is significantly higher.

The value of discounts for the other products is influenced by customer loyalty, being directly proportional to the frequency of purchases and purchasing power of the consumer. The discount can be customized for each product, influencing the complexity of the model, or can be applied as a general discount for the entire personalized offer. The model will determine a general discount for all products included in the recommendation.

The discounts will be applied on a time frame established by default or it can be also customized according to the consumer behavior. In Romanian, shopping in supermarkets is a weekly activity for the majority of consumers, so the model will generate for the recommendation a one week validity period that will start the day after the current shopping session.

Depending on the company's strategy, the offer can be used only once in that time frame or whenever the customer comes to shopping during that time.
In order to make recommendations even more attractive, they will also include a personalized Newsletter, based on nutritional information, diets, or recipes for a food product included in the offer or based on up-to-date information for a non-food product included in the offer. This will also be a separate section at the end of the offer or on the back of it, and the information will not be repetitive.
For customers who have a low shopping frequency, the quantity of the products sold needs to be increased, so selecting one of the specific shopping cart products included in the offer would be the best solution for generating the Newsletter. For customers who have a high shopping frequency, they aim is to increase the diversity of their shopping carts, so the Newsletter will be generated for products different from the shopping cart or even for random products that have not been purchased or recommended before.

3.4. Distribution channel

The personalized recommendations will be delivered once a day only for the first shopping session in order not to stimulate splitting the products into several purchases. Usually offers have the greatest impact when consumer are still focused on the shopping session, so the recommendations will be distributed at the end of the shopping session in order to increase the view rate. The consumer will be able to receive a printed offer along with the shopping bill or by e-mail. The consumer will be able to choose the distribution mode when requesting the fidelity card and will then be able to change this option according to his needs. No matter how the recommendation is distributed, the consumer must present it at the end of the shopping session in order to apply the discounts.

4 Input variables and main steps of the model

The shopping frequency is represented by the number of shopping sessions made in a recent time unit. The time unit can be predefined by the company or it can be represented by the entire period a consumer has become a loyal client by purchasing a loyalty card. As consumer behavior may change over time, affecting also the frequency of purchases, the time unit is represented by the last six months. The shopping frequency is expressed in days and is calculated using the following formula:

$$SF = \frac{TU}{SS}$$

SF – shopping frequency
TU – time unit
SS – number of shopping sessions in the defined time unit

Customers with a shopping frequency less than 7 days will be considered loyal customers, while customers with a shopping frequency greater than one month will be considered casual customers.
There are many social, psychological and economic factors that influence consumer behavior, such as gender, age, level of training, socio-professional category, residence environment, the main factor being the income. For most consumers the products price is a decisive component in the purchasing process, products from the same category being sold at different prices depending on their quality or supplier. The purchasing power of a consumer will be determined by the price category of the purchased products. The model attaches to each product a price category by calculating the ratio between the price of the product and the average price of all products in the same category. Thus, each product belongs to one of the following categories:

- Category 1 – budget product
- Category 2 – mass product
- Category 3 – premium product
- Category 4 – luxury product

The purchasing power is calculated by selecting all products purchased in the predefined time unit and applying the following weighted arithmetic mean:

$$PP = \begin{cases} \left[\frac{1*x_1+2*x_2+3*x_3+4*x_4}{x_1+x_2+x_3+x_4}\right], \\ \left\{\frac{1*x_1+2*x_2+3*x_3+4*x_4}{x_1+x_2+x_3+x_4}\right\} \le 0{,}5 \\ \left[\frac{1*x_1+2*x_2+3*x_3+4*x_4}{x_1+x_2+x_3+x_4}\right]+1, \\ \left\{\frac{1*x_1+2*x_2+3*x_3+4*x_4}{x_1+x_2+x_3+x_4}\right\} > 0{,}5 \end{cases}$$

x_i – number of products from Category i
PP – purchasing power
The purchasing power can be an essential factor in dividing customers into clusters and thus reducing the complexity of the customer based filtering algorithm. Also, the purchasing power of the customer is used to accurately identify the exact product that will be recommended to a customer, always selecting products whose price category does not exceed the customer's purchasing power.
The discount calculated by the model will be between 5% and 15%, being directly proportional to the shopping frequency. Therefore, the discount will be granted on the basis of the following distribution function:

$$D = \begin{cases} 5\%, & FC > 14\, zile \\ 10\% * \frac{7}{SF}, & 5 < FC < 14 \\ 15\%, & FC < 5\, zile \end{cases}$$

D – personalized discount
SF – shopping frequency

A product should not be included in two consecutive offers, thus a *recommendation range* is defined, expressed in number of days.

$$R = SF * c$$

R – recommendation range
SF – shopping frequency
c – constant, defined by the company

The constant c is influenced by the diversity of the products and the company's business strategy. The model will consider c = 4, so a product will appear in no more than one recommendation per month.

The selection and filtering algorithm consists of three parts, one for each product category included in the recommendation. The model selects one date specific product, one product different from the shopping cart and N-2 shopping cart specific products. Taking into account the impact of opportunity cost, the number of products included in the final personalized recommendation is N = 6. The algorithm is executed only for one consumer at a time. The main steps of the model are summarized in figure 1.

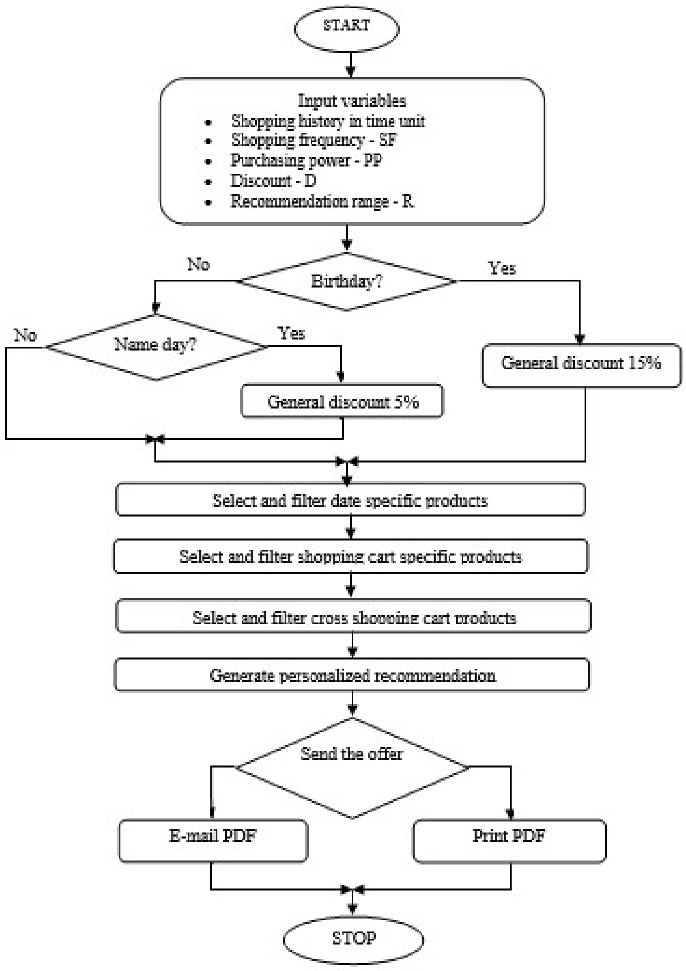

Fig. 1. The Logical schema of the recommendation model

4.1. Select and filter date specific products

STEP 1: check if the current date is included in a special period, like Christmas, Easter or St. Valentines. A custom table in the database maintains the special periods.

STEP 2: select all date specific products. Each product will be marked if it is specific to a period and each product will not be associated with more than one special period.

STEP 3: select all products previously purchased by the customer, regardless of the predefined time unit. This category of products is sold with an annual frequency so the time unit restriction is eliminated.

STEP 4: identify all date specific products purchased by the customer over time. If there is more than one product, the model will select the one purchased most often. If such a product could not be identified, then the date specific product is randomly selected.

4.2. Select and filter shopping cart specific products

Shopping cart specific products are also divided into three categories, identical products, similar products and complementary products, the algorithm having the same main logic for each subcategory.

The identical products are the products that were previously purchased by the customer. They are identified based on the purchasing history in the predefined time unit. For this category the price is not a filtering condition.

STEP 1: select all products previously purchased in the predefined time unit and remove one time purchased products which are considered spontaneous products.

STEP 2: remove the products that are no longer sold by the food retail store.

STEP 3: check if the products can be included in the recommendation according to the consumption period and the recommendation range.

A product can't be included in two consecutive recommendations and it can't be recommended before it is purchased and consumed. If the date of the last product recommendation summed up with the recommendation range exceeds the current date, or if the date of the last purchase cumulated with the consumption period exceeds the current date, then the product is removed from the selection.

The similar products are the products included in the same category as those previously purchased by the consumer. They could be identified by using an item based filtering techniques, but due to the diversity of the products the model will be slow and inefficient. So, the similar products are identified by CPV code. The Common Procurement Vocabulary (2008) establishes a single classification system for public procurement aimed at standardising the references used by contracting authorities and entities to describe procurement contracts. CPV codes are based on a branched structure of up to nine digits.

STEP 1: select all products similar to those already purchased in the predefined time unit. Select those products for which the first six digits of the CPV code is the same and the price category is the same as the consumer's purchasing power.

STEP 2: remove the products that are previously purchased because their selection has already been made using the identical algorithm.

STEP 3: remove the products that are no longer sold by the food retail store.

STEP 4: check if the products can be included in the recommendation. If the date of the last product recommendation summed up with the recommendation range exceeds the current date, the product is removed from the selection.

The complementary products are those products that are sold separately, but can be consumed together with other products, being in a direct association relationship.

STEP 1: select all products that are complementary to those already purchased in the predefined time unit, taking into account the price category. Select those products for which the price category is the same as the consumer's purchasing power.

STEP 2: remove the products that were previously purchased or the similar products because their selection has already been made using the above described algorithms.

STEP 3: remove the products that are no longer sold by the food retail store.

STEP 4: check if the products can be included in the recommendation. If the date of the last product recommendation summed up with the recommendation range exceeds the current date, the product is removed from the selection.

The model selects X1 identical products, X2 similar products and X3 complementary products, all being products specific to the consumer's shopping cart. A product will be uniquely found in one of three categories. The model selects randomly N-2 products to be included in the recommendation.

4.3. Select and filter cross shopping cart products

In order to select the products that are different from the shopping cart a customer based filtering technique is used by the model. To reduce the complexity of the model the customers are split in consumer clusters using as main rule of differentiation the purchasing power. The algorithm is applied only for customers from the same cluster.

STEP 1: select all customers who are part of the same cluster as the consumer for whom the model is running.

STEP 2: select all items purchased by each customer in the predefined the time unit. Determine how many times each product was purchased. Identify the common products purchased, taking into account only the products from the current shopping cart.

STEP 3: identify the customers who have purchased the most common products and select the different product that was purchased most often and is not already included in the offer.

STEP 4: remove the products that are no longer sold by the food retail store and check if the products can be included in the recommendation.

For example, the current bill contains products A, B, C, D, E. There are 4 consumers belonging to the same cluster and having the same purchasing power as the consumer for which the recommendation is generated. The consumer based filtering technique is illustrated in table 1.

Table 1. Consumer based filtering technique implemented in the model

Consumers	**C_1**	**C_2**	**C_3**	**C_4**
Products purchased in the predefined time unit	A_2 G_7 F_6 D_3	B_3 D_7 M_5 O_3 N_1	D_1 F_6 G_2 N_9	A_4 C_6 H_2 N_9
Common products	2 (A_2,D_3)	2 (B_3,D_7)	1 (D_1)	2 (A_4,C_6)
Degree of similarity	5	10	-	10
Cross shopping cart products	N_{10} M_5 O_3 H_2			

The model selects product N because it has been purchased most often by the consumers with the highest degree of similarity. The degree of similarity between two consumers is calculated as the number of common products purchased by the two consumers. If product N is specific to a period, is specific to the shopping cart, it is no longer sold or if it has been included in a recent offer, then another different product that meets the filtering conditions is recommended.

5 Conclusions and future research

Due to the exponential growth of population and the rapid evolution of technology, consumers needs have become in the last centuries very diverse. Given this radical change of the economic environment where the bargaining power has shifted from sellers to buyers, companies need to adapt their business and marketing strategies in order to survive in a competitive environment.

The food retail market in Romania is becoming saturated, so the stores are realizing that keeping their actual clients is less expensive than attracting new ones and that gaining consumers loyalty have a big impact on their financial results. So, more and more companies are starting to use the newest technologies and software to increase customer satisfaction. This strategy can also be adopted by food retail stores.

The model described in this paper is based on the principles of recommendation systems originally developed for e-commerce companies. The hybrid algorithmic model based on data mining and collaborative filtering techniques generates real time personalized offers. The model can be used by the large food retail stores in Romania, as it analyses a wide range of products that have different characteristics. The complexity of the model may increase exponentially depending on the number of variables and techniques integrated, the ultimate goal being the recommendation of a limited number of products per consumer. The products are selected and filtered according to the consumer's specific buying behavior, and the final offer is personalized to the smallest details.

In the future, the model will be implemented in the SAP ERP program and tested on a real-life data base. The generated recommendations will be given to the trial participants in order to quantify the results of the model and to make the necessary adjustments.

References

[1] G. Miller, *Spent: Sex, Evolution and Consumer Behavior*, Penguin Books, 2009.

[2] R. R. Burke, *Innovative Uses of Video Data Analysis in Marketing: Measuring Shopper Attention, Engagement, and Purchase*, AMA-ECMI-EMAC Marketing & Innovation Symposium, Erasmus University, Rotterdam, 2014.

[3] Precima, *How important is price and promotion personalization to grocers and shoppers?,* article retrieved March 15, 2018, from https://www.precima.com/content/how-important-is-price-and-promotion-personalization-to-grocers-and-shoppers

[4] G. McVie, *Are personalized offers the remedy for ineffective trade promotions?,* LEAD Virtual Marketing Conference, 2017.

[5] F. O. Isinkaye, Y. O. Folajimi, B. A. Ojokoh, *Recommendation systems: Principles, methods and evaluation*, Egyptian Informatics Journal, 16(3), 261-273, 2015.

[6] J. Leskovec, A. Rajaraman, J. D. Ullman, *Mining of Massive Datasets,* 2nd Edition, Cambridge University Press, 2014.

[7] J. L. Erlocker, J. A. Konstan, L. G. Terveen, J. T. Rield, *Evaluating collaborative filtering recommender systems,* ACM Trans Inform Syst, 22(1), 5-53, 2004.

[8] C. A. Gomez-Uribe, N. Hunt, *The Netflix Recommender System: Algorithms, Business Value, and Innovation,* Transactions on

Management Information Systems, 6(4), no.13, 2016.
[9] E. A. Baatarjav, S. Phithakkitnukoon, R. Dantu, *Group Recommendation System for Facebook,* OTM: On the Move to Meaningful Internet Systems, 211-219, 2008.


Cristina STOICESCU graduated the Faculty of Economic Cybernetics, Statistics and Informatics, with a bachelor degree in Economic Cybernetics in 2012. In 2014 she got her masters degree from the same faculty of the Bucharest University of Economic Studies, specialization in Cybernetics and Quantitative Economics. Currently she is taking her PhD degree in Economic Cybernetics and Statistics at Bucharest University of Economic Studies, coordinated by Professor Crişan Albu. Her PhD thesis is “Experimental Economics in shaping consumer behavior”.


Cristina STANCIU-MÂNDRULEANU obtained her bachelor's degree as an economist after graduating the Faculty of Economics Sciences, at the University of Craiova in 2001. In 2009, she completed her masters degree at the Bucharest University of Economic Studies - Faculty of Master Administration. Currently, she is taking her PhD degree in Economic Cybernetics and Statistics at Bucharest University of Economic Studies, coordinated by Professor Crişan Albu. Her PhD thesis is “Means and methods of analyzing the behavior of a company with application on the medical oxygen market”.