



# Machine Learning in E-commerce

Maria-Cristina ENACHE\*

## ARTICLE INFO

### Article history:

Accepted January 2019

Available online April 2019

JEL Classification

K22, M21

### Keywords:

IT, E-commerce

## ABSTRACT

With the explosive growth of data, it is one of the most important challenges for modern enterprises to develop data-based infrastructures. The scientific discovery of Artificial Intelligence (IA) is an open opening for a wide range of applications, which allows large quantities of real-time business data to be obtained. As sales of electronic electronics continue to grow, the question arises whether e-commerce will overcome or completely replace sales of physical stores. Although nature of technological advances and consumer behavior are increasingly unpredictable, there are some key trends that I believe will mark both the medium and long term growth of e-commerce.

© 2019 EAI. All rights reserved.

## 1. Introduction

Nowadays, most companies have adopted more realistic expectations and understand that AI will not solve all their problems but cannot be ignored.

An area that is greatly affected by this development of artificial intelligence is that of retail sales. The number of online purchases is steadily increasing, which allows companies to gather detailed data on the customers of the whole experience: what they are looking at, what order, how long, what day, what questions they may have, what they ultimately acquire aims to evaluate and review products. With AI, this date can then be used to improve this experience to make it easier, more efficient, more engaging, and more responsive to personal needs.

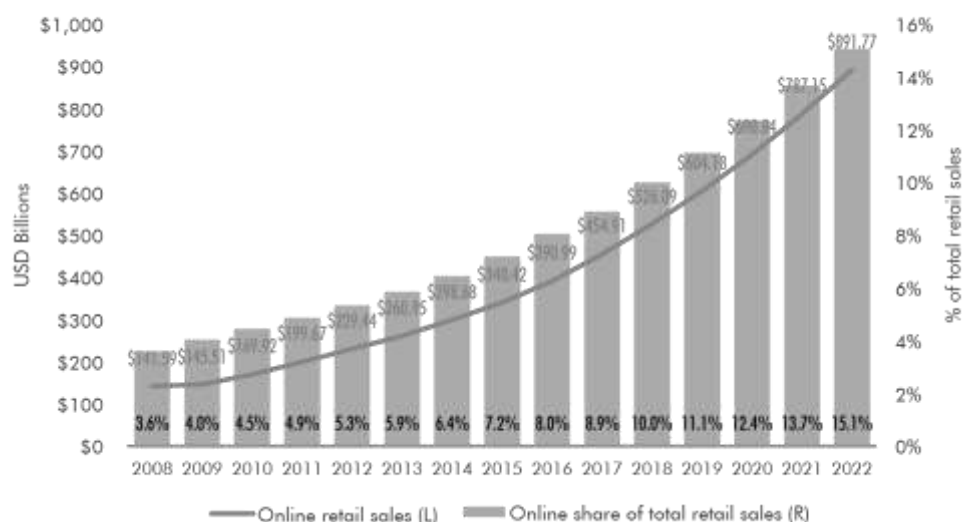


Fig 1. - Forecast E-commerce Sales

Source: eMarketer, 2018

As sales of electronic electronics continue to grow, the question arises whether e-commerce will overcome or completely replace sales of physical stores. Although nature of technological advances and consumer behavior are increasingly unpredictable, there are some key trends that I believe will mark both the medium and long term growth of e-commerce.

\*Dunarea de Jos University of Galati, Romania. E-mail address: [maria.podoleanu@ugal.ro](mailto:maria.podoleanu@ugal.ro)

Innovations in this sales area are based on the automatic learning method, which is the AI sub-field that develops autonomous learning algorithms. This article will outline some of the implications of using artificial intelligence in electronic commerce

The increase in e-commerce has varied considerably depending on the retail category and the retail price. By categories, goods such as clothing and digital goods have shown an increase in online sales higher than other categories such as health and beauty, which consumers still prefer to purchase in the store. Though technologies such as virtual reality and facial recognition can make online purchases easier in all categories, the growth of ecommerce will vary greatly by category.

At the same time, other categories will see lower e-commerce in cities than in suburbs, as urban consumers have easy access to physical store networks.

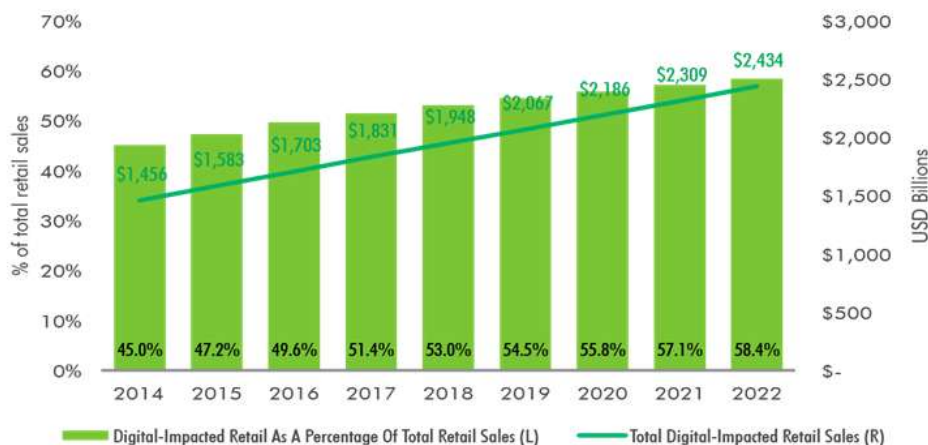


Fig 2 – Digital Sales  
Source: eMarketer, 2018

## 2. Machine learning in ecommerce

Product recommendation is usually the first thing people have in mind when thinking about machine learning. Features such as "If you love the product x, you'll probably also want the product" have proven to work remarkably well and can serve as a valuable tool to guide users through the growing masses of options they have at their disposal. Traditionally, recommendations have been added manually, based on the heavy-duty product categories, but this is extremely time-consuming, error-prone and fast-paced.

The modern recommendations algorithms are divided into two categories: collaborative filtering and content-based filtering. In collaborative filtering, the recommendations for a particular customer are based on what similar customers have chosen in the past. By simplifying somewhat, if the customer bought X and Y, and customer B only bought product X, Y could be recommended to customer B. On the other hand, customers (for example, gender, preference brand, or age) to product content (e.g., category, price, or color).

Recommendations system are automated learning systems or machine learning systems that help users discover new products and services. Every time you buy online an recommendations system guides you to the most likely product the user will buy. Today's recommendations system is an essential feature because users are often overwhelmed by the system of choice and recommendation, helping users quickly find the product they love. This leads to more sales, as well as making users happier. These systems make dollar companies recommend a product to a user to make a sale or to make the user think whether to buy a product or not. It's like a vendor who recommends a product, but a virtual one without a human cleansing, but a better cure based on your history.

In ecommerce systems, recommendations can be made from users or from products:

- The similarity of users is for checking the difference between two users. If two users have similar preferences for a product, we can assume that we have a similar interest. It's like a friend who recommends a product. The similarity of users has a great inconvenience that all user data must be stored and used in the product recommendation process. Recommendation based on the similarity of users requires prior user data. A new ecommerce site that needs to suggest a product starts with a handicap because it does not have a large number of clients. The product similarity does not come into this problem because it only requires product information. It is only the user's preference. Many of the ecommerce sites avoid this problem by asking users to prefer it while initiating a new account.
- The most useful or effective system for recommendations products is the one that relies on products. If the user navigates and searches for a particular product, similar products may be displayed. The similarity of products is especially useful when we do not know the user's profile, but we can recommend the product to the user based on his navigation history.

### 3. Recommendations system

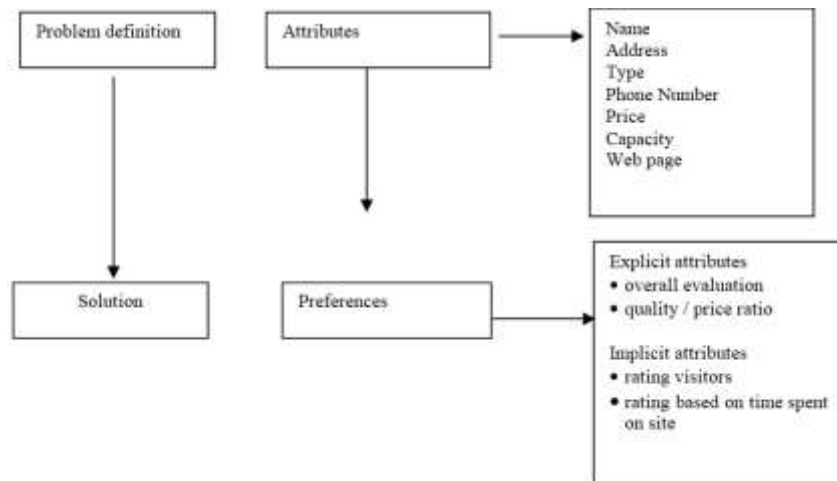
The recommendation function in e-commerce works to suggest users to buy the products they need and takes over the user information and predicts the user's rating for the product. Forecasting user rating right before user rating makes the product strong. Suggesting users only products they would miss very much and not the products they do not care.

- The user-product relationship - Some users have an affinity or preference for the specific products they need. For example: Women are a preference for articles related to clothing or cosmetics so that the e-commerce site will build a product relationship to the female user type.
- The product-product relationship - When we look at products, some products are similar in nature. Hence aspect or similarity is an important factor. In e-commerce, this is worthwhile, for example, by categorizing: same-sex books, same-sex music, dishes from the same kitchen, news items from a particular event.
- User-User Relationship - Some customers have similar tastes about a particular product or service.
- User behavior data - Useful user engagement information on the product can be collected from ratings, clicks, historical acquisitions.
- Demographics of users - personal user information such as age, education, income, location are found in contact forms and are important elements in the system's recommendations.
- Product Characteristics - Product identifier information is added by the store manager as other vital information in making the recommendation.

### 4. Example

The most useful or effective system for suggesting or recommendations products is the one that relies on products. If the user navigates and searches for a particular product, similar products may be displayed. Users often expect to quickly find the products they want and move on if they have difficulty finding the relevant product. When the user chooses to select a product or read more about it, it can receive information about another similar product, or if the user buys the product, he can receive advertisements or coupons via e-mail to purchase a similar product. The similarity of products is especially useful when we do not know the user's profile, but we can recommend the product to the user based on his navigation history. The content-based recommendation system uses the attributes of each product to recommend new products. Users are suggested products based on attributes. The content-based recommendation works well when content descriptive information is provided in advance. "Similarity" is measured by attributes of the product.

The proposed example is based on restaurant activity. So we took into account some of the most important or significant attributes of his products, as well as consumer preferences that are starting points in the realization of the consumer profile.



We will consider a set of products  $P = \{p_1, p_2, \dots, p_n\}$

Each product is characterized by a set of attributes  $p_i = \{at_{i1}, at_{i2}, \dots, at_{in}\}$

The attributes considered are not very complex, they consist of significant adjectives to describe them, sometimes nouns or values.

In this example, products and preferences/ratings are represented as such

$At = \{\text{name, address, phone number, type, price, capacity, web page}\}$

$Int^e = \{\text{quality/price ratio, quantity, serving}\}$

$Int^d = \{\text{visited web pages, time needed for recommendation, time earned}\}$

An experience of a visitor at the Res Galati restaurant could be:

$E = \langle \{\text{Restaurant ResGalati }, \text{"St. nr 3 Street, "0236/345345", "traditional", "55 ron", 100, "www.resgalati.ro"}\}, \{0.83.0.76.0.91\} \{2,21,12\} 0.83 \rangle$

Thus, for a restaurant with a capacity of between 120 and 150, traditional, creative, national, the price from 55 to 100.

The consumer profile is given based on a history of the heights through  $\text{Profile} = \langle E_i, C_i \rangle$ , where

- $E_i$  is the set of past experiences. The set of past experiences that make up the user's profile is marked with  $E_i$  and is included in  $E$ , that is, the set of possible experiences.
- If we denote  $A = \{a_1, a_2, \dots, a_r\}$  a training set then the coordinating agent will have  $C_i = \{(a_{i1}, t_{i1}), \dots, (a_{ir}, t_{ir})\}$  where  $a_{ij}$  is an element of  $A$  and  $t$  is a numerical value between  $[0,1]$ , which represents an assigned trust value.

Each user has a different degree of interest or preference for any product  $p$ . Such preference can be expressed by the consumer through explicit or automatically captured attributes as a result of customer interaction (expressed by implicit attributes).

Explicit preferences are usually more important, or higher than implicit, but not always accessible. Implicit preferences are useful when deciding products that would interest the customer. as both categories of preferences have their importance, they will be noted

- explicit preferences set  $\text{Int}^e = \{\text{int}_1^e, \text{int}_2^e, \dots, \text{int}_m^e\}$
- implicit preferences set  $\text{Int}^d = \{\text{int}_1^d, \text{int}_2^d, \dots, \text{int}_m^d\}$  -
- $\text{int}_j^e, \text{int}_j^d$  have values between  $[0,1]$

Consumer history can be defined according to preference sets like this

$$E_i = \langle p_i, \text{Int}^e_i, \text{Int}^d_i, \beta_i \rangle \text{ where}$$

- $p$  - product
- $\text{Int}^e_i$  is explicit preferences set
- $\text{Int}^d_i$  is implicit preferences set
- $\beta_i$  is a temporary parameter that indicates the relevance of the experience and whose value is initially set to 1, but can be found in  $[0,1]$ . This value will change with the user profile.

In order to be able to find consumer preferences and to generate recommendations, the first step is to achieve the history, in other words, the introduction of training data.

Thus the consumer is asked to indicate or select some products for which he will select some of his favorite attributes, by which the system will train to make recommendations.

A. In the first step you will calculate the preference for a product,

$$V_i = \beta_i * g(f^e(\text{Int}^e_i), f^d(\text{Int}^d_i)) \text{ where}$$

- $f^d$  is a function that combines implicit preferences
- $f^e$  is a function that combines explicit preferences
- $g$  is a function that combines the results of function  $f^e$  and  $f^d$

B. In the second step, the preference value for the product  $r$  will be calculated from the similarity of the products  $I_r$

$$I_r = \frac{\sum_{i=1}^x (\text{Sim}(r, i) * V_i)}{\sum_{i=1}^x \text{Sim}(r, i)}$$

- $x$  is the number of similar products,
- $\text{Sim}(r, i)$  is the similarity between the products  $i$  and  $r$ ,
- $V_i$  is the preference value for the product  $i$ .

Starting from the example of the ResGalati restaurant, if the weights for the explicit attributes selected by the customer are

<i>Explicit preferences</i>	<i>Weight</i>
quality / price ratio	0.6
product weight	0.1
service efficiency	0.3

then the function  $f_{e_i} = 0.6 \cdot 0.91 + 0.1 \cdot 0.83 + 0.3 \cdot 0.76 = 0.86$

## 5. Conclusions

Machine learning does not always require a lot of data to work. The most useful or effective application for suggesting products to users The recommendations systems are everywhere. Users used it without knowing it. Consumers have infinite choices, but the recommendation system plays an important role in finding consumers the content they need without having to spend on things they do not like. Different users have different preferences and requirements. However, it is necessary to know the mechanisms of their functioning and to be aware of their implications in the decisions we take every day, as recent history shows, the recommendations are not always ethical or moral, depending on the policy of the company using these systems.

## Acknowledgement

This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0800 / 86PCCDI/2018 - FutureWeb, within PNCDI III.

## References

1. P. Funk and P.A. González Calero - *Hybrid Recommender Systems with Case-Based Components*, ECCBR 2004, LNAI 3155, pp. 91–105, 2004.
2. Ben Schafer, Joseph A. Konstan - *E-Commerce recommendation Application - Data Mining and Knowledge Discovery*, 115–153, 2001
3. A.Y. Tawfik and S.D. Goodwin - *Evaluating a Smart Recommender for an Evolving E-learning System: A Simulation-Based Study - Canadian AI 2004*, pp. 439–443, 2004.
4. <http://attend.it.uts.edu.au/icita05/> International Conference on InforTechnology for Application (ICITA 2004)
5. <http://ltf.ieee.org/icalt2004/ppt/66.pdf>
6. <https://www.emarketer.com/articles/>
7. <https://towardsdatascience.com/>