

RESEARCH ARTICLE

Decision-Making by Data, An Application of Data Science with Cognitive Ergonomics

Thomas Jeferson Nogueira dos Santos¹, Claudia Lysia de O Araujo² Henrique Martins Galvao²

- ¹ Graduate Program Stricto Sensu in Design, Technology and Innovation - Professional Master's Degree from the University Center Teresa D'Ávila - UNIFATEA.
- ² Stricto Sensu Postgraduate Program in Design, Technology and Innovation Professional Master's Degree from the University Center Teresa D'Ávila - UNIFATEA.

Abstract

With the growth of computing and technology in general in the last few decades, there has also been an abundance of available data, information previously unimaginable, which now cause companies and organizations to encounter needs they have not previously had. Such a need comes in the face of its ability to make decisions in the increasingly shorter timeframe, the fact that it poses major challenges for data collection, transport, processing and performance. In this context, a new denomination known as Data Science, capable of handling data from several areas of knowledge, is a great capacity to display quick and objective information, which in a Business context, support decision making, a very important factor To remain competitive in today's commercial environment. This article aims to study this new denomination, Data Science, and related topics, such as Data Mining, Cognitive Ergonomics and fundamentals for decision making.

Keywords: *Data Science, Decision Making, Data Mining, Cognitive Ergonomics.*

Introduction

In the intense rhythm with which information is emerging and impacting directly on the economic and corporate conjunctures of organizations, making quick decisions can often mean commercial survival and be a powerful factor in business competitiveness. In this environment, the visual perception, in the parallel processing of the data, or simply in the intuition, are related to the development of the applied strategy as a way to obtain fast results from the decision making.

According to Simon [1] the essence of creativity and intuition lies in the organization of knowledge for a quick identification of the component elements of an event. Following this reasoning [2], point out that probability, risk and uncertainty contribute differently to decision-making processes. From this logic, the reasoning of the decision is related to the intuition generated after analyzing the information acquired. Offering, thus, support the knowledge about the action and subsequent

application of the chosen strategy. In the business environment, this decision-making task requires tools and methods that help decision-making. In a knowledge-based economy [3], data acquisition, analysis, and access to support decision making have become indispensable to business success or survival. With the emergence of this new scenario, companies employ statistics, quantitative analysis and predictive modeling as the primary competitive elements [4].

In this context, the emergence of a new science, called Data Science, begins to gain notoriety. Its definition continues to be developed, but, in short, it is a combination of several areas of knowledge and distinct abilities, aiming at collecting, preparing, analyzing, visualizing, managing and preserving large amounts of information. The theme is of great relevance because it deals with factors that currently impact on the success or survival of Companies and the lack of related works make this theme important

for its development. The research will be based on bibliographical works such as articles and theses. The methodology used will be qualitative, because it is a bibliographical research and the type of analytical / explanatory research in order to deepen the knowledge of reality, analyzing and interpreting the factors that led to such foundations.

Theoretical Grounding

Data Science

The most common terms used to define Data Science have been used over the past few years, are linked to other areas of knowledge such as Data Analysis, Data Processing, Statistics, Knowledge Discovery in Database (KDD), Data Mining, Big Data, among others.

Data Science is an interdisciplinary field that divides definitions and areas of action with other fields that support and guide the extraction of information and knowledge from structured and unstructured data. Data Science [5] involves principles, processes, and techniques for understanding phenomena through automated data analysis. It can be said that, Data Science is the scientific domain that is dedicated to discovering Knowledge through data analysis. It can be said that, Data Science is the scientific domain that is dedicated to discovering Knowledge through data analysis.

Figure 1 shows the "Venn Data Science Diagram" [6], which illustrates the interdisciplinary involved in Data Science.



Figure 1: Venn Diagram of Data Science

Source: (Conway, 2010)

The application, together, of the skills described in the Venn Diagram-computational skills, statistical and mathematical knowledge, and domain/subject expertise - define, in practice, Data Science [6].

I Defined Some Concepts Related to Data Science, Which Are [7]

- Data Science is the study of generalizable knowledge extraction from data.
- A common epistemic requirement to assess whether new knowledge is likely to be a resource for decision making is its predictive power, not just its ability to explain the past.
- A data scientist requires an integrated set of skills that encompasses mathematics, machine learning, artificial intelligence, statistics, databases and optimization, along with a deep understanding of the craft of formulating a problem to design effective solutions.

As previously stated, the concept of Data Science has been elaborated over the last few years. Nowadays, companies of the most diverse segments are seen in the middle of a vast amount of data, available in diverse forms. They focus their attention on the exploitation of this data in order to achieve competitive advantage or maintain its survival. The convergence of this phenomenon has given rise, more and more, the widespread application in the Data Science business ⁽⁵⁾.

Discovery of Knowledge in Database

The Knowledge Discovery in Databases (KDD) is focused on developing methods and techniques to make sense of the data collected. It is a process used to identify valid standards in the analysis of large volumes and data sets, and can discover important information that can help in the strategic decision making of the Company, for example in the direction of marketing, in increasing the profitability of a given commerce or

business, among others. Mapping low-level data (which is typically too bulky to easily understand and synthesize) is a problem in the KDD process, as in other forms that may be more compact, such as a simple, more abstract report, a descriptive or more useful approximation, as a predictive model to estimate the value of future cases for example. At the heart of this process is the application of specific data mining methods

for pattern discovery and extraction [8]. Data Mining can be defined as a process with defined phases. Extracting information from structured and unstructured datasets and transforming them into an understandable structure for later use. In addition to this main process, Data Mining, there are also five other processes that make up the KDD, as shown in Figure 2 below:



Figure 2: Steps of the KDD process

Source: (Fayyad et al., 1996)

The Stages of the KDD Process Are Detailed Below

- Selection - The selection phase is the first phase of the KDD. It is a very important phase because it is the decision phase which sets of data will be relevant to obtain results with useful information;
- Preprocessing - In the preprocessing phase, data cleansing and attribute selection takes place. At this stage, missing, erroneous or inconsistent information in the databases must be corrected so as not to compromise the quality of the knowledge models to be extracted at the end of the KDD process;
- Transformation - At this stage the transformation of the data happens, ie, the important data that were taken out in the previous process are modified so that the next step can be performed. The transformation is nothing more than analyzing the data and reorganizing them in a specific way and will be interpreted by a data mining software;
- Data mining - In the mining phase is where everything happens, the data after preprocessing will be read and interpreted. Mining makes mere data to be transformed into information, such information is indicated through rules that can only be interpreted through

brute force, that is, reading rule by rule and interpreting them;

- Interpretation / Evaluation - In this last phase is where the rules indicated by the previous process will be interpreted and evaluated. After the interpretation may arise patterns, relationships and discovery of new facts, which can be used for research, optimization and others.

In terms of methods and techniques, the KDD area is commonly confused with what has been defined as Data Science. This process can be used on any type of database as long as the data is cleaned in advance so that it is only the most important and necessary.

Data Mining

Data Mining or Data Mining is the technique of exploring large volumes and data sets for the purpose of establishing relationships, associations and patterns, transforming raw data into high value information. For this to be possible, learning algorithms are used that classify the data based on statistical information and neural network crossings.

The results are usually presented in the form of rules, hypotheses, decision trees and graphs. Data Mining is divided into two parts, the task of specifying what we are looking for in the data, what kind of regularities or category of patterns we are interested in finding, or what type consists in

specifying methods that guarantee us how to find the patterns that interest us. Among the main techniques used in data mining are statistical techniques, machine learning techniques and techniques based on information validation, which are: decision trees, neural networks, and Bayesian classification, among others.

Cognitive Ergonomics

Cognitive Ergonomics [9] is characterized by aspects related to the issues of comprehension, logic, compatibility of repertoires and information, meaning of messages, complexity of the task, among other aspects that result in disturbances to the selection of information for cognitive strategies and compromise their autonomy in problem solving and decision making, such as the difficulties of decoding, learning and memorization, in the face of logical inconsistencies and navigation of the communicational and dialogical subsystems.

It deals with how people conceptualize and process information absorbed in situations arising from their work, in the mental aspect of perception, attention, storage and memory retrieval, etc. Cognition [10] is what happens in the human mind when daily activities are performed. Such activities involve cognitive processes such as thinking, remembering, learning, fantasizing, making decisions, seeing, reading, writing and speaking.

"Human cognition stems from the capacity developed by men and women for the creation or composition of mental representations and imaginative processes, starting from the memory of feelings, feelings and ideas. These creations or compositions are caused by internal disturbances which in part derive directly from the stimuli received from the environment in which humans are inserted [11]. "Therefore, the Cognitive Ergonomics approach is, in fact, a contribution to the decision-making issues, whether they are relevant and that they run logically.

Decision Making

Decision-making is a natural occurrence that happens constantly within organizations, to a greater or lesser extent, and directly affects the organizations' daily lives. For decision-making to occur, the people involved in the process should be provided with as much information as possible.

However, as indicated, much of your decision will take into account the intrinsic knowledge you have or even your intuition. Organizational [12-13] decisions can be classified according to the administrative activity belonging to it,

According to Three Levels

- Operational level: Signifying the effective and efficient use of existing facilities and all resources to perform operations. The operational level decision is a process by which it ensures that operational activities are well developed. For this, the quality of the information and the speed that it is made available are fundamental for a correct decision, especially when time is a preponderant factor for success.
- Tactical level: Including generic acquisition of resources and tactics for acquisition, localization of projects and new products, decisions at the tactical level are usually related to administrative control and are used to decide on the operations of: Formulate new decision rules ; Variation from a planned operation; Analysis of decision possibilities in the course of actions. These decisions do not require so much speed to be taken, but if a very long time to make a decision occurs, this can cause losses and possible losses.
- Strategic level: It encompasses the definition of objectives, policies and general criteria to plan the course of the organization.
- In relation to decisions at strategic levels, it is to develop strategies so that the organization is able to reach its macro-objectives. These decisions are most often marked by historical information and the urgency to make a strategic decision need not be so great as to take an operational decision. Decision-making at all three levels should follow criteria of optimization and timeliness of information to address possible asymmetries.

Clarifies, Decision Making is a Concern of the Individuals of our Society for a Long Time [14]

[...] decision-making is, of course, part of a stream of thought begun in times when man, in the face of uncertainty, sought guidance from the stars. From then on, the search for new decision-making tools, from the Arabic-Arabic numerical system to the

systematic empiricism of Aristotle, from the advances in Occam's logic to Francis Bacon's inductive reasoning to the application of the scientific method by Descartes, has never ceased. The increasing sophistication of risk management, understanding the variations in human behavior, and the technological advancement that supports and simulates cognitive processes, and in many situations has improved decision making.

The decision-making process can be exemplified as how people act at a particular moment in which a decision-making action is necessary. Speaks of naturalistic decision-making, in which action occurs in a natural environment using the most diverse sources of power [15].

Decision Making, in the Naturalistic View, can be Defined as Any Action Involving the Following Characteristics

- Time Limitation;
- High risk and of any sort;
- Expert decision-makers or strategists;
- Presence of inadequate, ambiguous or wrong information;
- Presence of ill-defined objectives and / or procedures,
- Existence of learning by suggestion,
- Stresses,
- Presence of coordinated teams

In this way, it is important to bring together the choice of an option so that it is satisfactory and that it involves the best solution. To arrive at a solution level of a problem it is necessary to have the largest number of variables analyzed and to have reliable information of the problem under various angles and forms. For coherent decision making, it is necessary to have in its domain all the necessary ways to exclude any doubt, together with the experience and the appropriate preparation. These elements will guide you in a correct and safe way when making decisions and have a great chance of success within the organization.

Research Method

As for the method, this is an analytical / explanatory study with the literature review as scientific articles, sites related to the area

of Data Science, concentrating the studies on the topics of Decision Making, Knowledge Discovery and Cognitive Ergonomics. The study's research is intended to address the concepts of Data Science and relate to Decision Making as a competitive and survival advantage for Business Organizations in order to discuss foundations of knowledge discovery that can be used in their models business, deepening the knowledge on the subject, analyzing and interpreting the bibliographical information obtained.

Final Thoughts

Speaking in decision making without first checking how the information that is received and handled, and how to act within this process would not be possible without the approach on Data Science. To ensure how this information is processed we have to study and understand the human cognitive process and how it is in its decision making process. In this context it is verified that the decision-making process is not simple, since the individual searches for relationships and combinations of ideas, concepts, facts and information, in this large amount of information, seeking to transform them into knowledge and, in this way, knowledge in the intended process.

It is not a simple process, because the great amount of information, when not organized and treated, can undermine the process of decision making. Discovering useful patterns in data that can be used in decision making has gained a variety of names, such as: Data Mining, Knowledge Extraction, Information Discovery, Information Collection, among others. However, the term Data Mining has been used by statisticians, data analysts, and communities of information systems managers to help with decision making.

The proposed study is relevant because it aimed to generate a brief relationship between the current needs of decision making by the Companies and the current processes of Knowledge Discovery. Another factor is the lack of knowledge about the topic of Data Science and its applications, ranging from its denomination to its processes involved. It is concluded that the decision-making process can be defined as the way that individuals act when a decision-making action is demanded, in which the analysis and judgment of the available information must

be taken into account. Thus, using processes and methods to assist in this decision are critical to business survival. Decision-making

with support for data science implies the search for risk reduction and greater success of the proposed actions.

References

1. Simon HA (1987) "Making Management Decisions: The Role of Intuition and Emotion." *Academy of Management Executives*. 5: 7-64. In: Mintzberg, H. Sternberg, R., 2000, *Psicologia Cognitiva*. Porto Alegre: Artmed.
2. Rushworth MFS, Behrens TEJ (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience* [S.I.], 11,(4):389-397.
3. Godin B (2006) The Knowledge-Based Economy: Conceptual Framework or Buzzword? *Journal of Technology Transfer*, 31,(1):17-30,
4. Davenport TH, Cohen D, Jacobson A (2006) "Competing on Analytics" In: *Harvard Business Review*, 98-107.
5. Provost F, Fawcett T (2013) "Data Science and its relationship to big data and data-driven decision making" In: *Big Data Journal*, 1, 51-59.
6. Conway D (2010) "The Data Science Venn Diagram". Disponível em <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>.
7. Porter M, Millar V (1985) "How information gives you competitive advantage." *Harvard Business Review*.
8. Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) "From Data Mining to Knowledge Discovery in Databases", In: *Advances in Knowledge Discovery and Data Mining*, 1-34.
9. Moraes A, Mont'alvão CR (2003) *Ergonomia: conceitos e aplicações Metodologia Ergonômica*. Rio de Janeiro : iUsEr 139. 16.
10. Preece J, Rogers, Y, Sharp H (2005) *Design de Interação – Além da interação homem-computador*. Porto Alegre: Bookman.
11. Santos FMR, Sousa RPLO (2010) *conhecimento no campo de engenharia e gestão do conhecimento. Perspectivas em Ciência da Informação*.
12. Drucker PF (2001) *O melhor de Peter Drucker – A Administração*. São Paulo. Livraria Nobel AS.
13. Santos LP, Wagner R (2008) *Processo decisório e tomada de decisão: um dualismo*. In: SEGET - Simpósio De Excelência Em Gestão E Tecnologia.
14. Passos A (2008) *Como Tomar Uma Decisão?* Portal da Administração.
15. Klein GA (1998) *Fontes de poder: o modo como as pessoas tomam decisões*. Lisboa: Instituto Piaget.