Generation Of Decision Making For Processes In Health Services Smes

Hugo G. Hernández Palma[1] , William A. Niebles Núñez[2] , Claudia Rojas Martínez[2]

[1] Business Administration Program, Universidad del Atlántico, Colombia.

[2] Faculty of Economic and Administrative Sciences, Universidad de Sucre, Colombia.

ABSTRACT

Artificial intelligence is currently one of the latest revolutions being incorporated at the business level. For this reason, and with the aim of analyzing how it applies to business processes in small and medium enterprises, the results of a descriptive analysis are presented, supported by a documentary review, which includes the most significant aspects of the subject. The results show that data mining is a very valuable tool for the competitiveness of companies and the research consulted suggests its insertion for the promotion of more remarkable results in companies.

Keywords: Data Mining, Business Processes, Small and Medium Enterprises (SMEs), Decision Making.

1. INTRODUCTION

Since the appearance of new and better methods for data processing, which have been in constant evolution from the most important computational developments of the second half of the 20th century, management decision-making has benefited from the use of these techniques [1, 2]. Thus, in an economic and commercial environment as dynamic as the current one, information is the key to taking clear competitive advantages and achieving a better positioning in the marketplace [3].

In this order of ideas, data mining becomes a useful tool that can be exploited by organizations with the appropriate information architecture and training [4], for this reason, the purpose of this document is to make an analysis based on the concepts and theories on the most important elements of data mining and its relationship with business and business processes, presenting a brief description of the problem and the need to address this area of knowledge in higher education scenarios.

Organizations and their leaders face a variety of situations that warrant a critical and reflective decision-making process [5], in tune with the environment in which they operate and with greater emphasis on the facts, most of which are represented by data and information. Therefore, knowing, implementing, and mastering Decision Support System (DSS) analysis techniques, including data mining, could bring important benefits and advantages that substantially contribute to competitiveness and productivity, making operations more efficient, effective, and efficient in a comprehensive manner.

In some cases, especially in small businesses and with short life spans, decisions are based on intuition, judgement, experience, and other qualitative elements based on data, structured or unstructured [6] that create even greater problems for the organization, given that the decision-making process must be supported not only by these elements, but also by information from the processing of real data from the dynamics of the environment [7].

However, the implementation of data mining techniques to support business decisions implies the development of an adequate structure of information systems, where the commitment of management in terms of investment, training and adoption is key [8], this is because the use of these techniques must be mastered and understood by the organization to fully exploit their advantages and benefits, which generates greater pressure on future management professionals to master these issues.

2. METHODOLOGY

Based on the qualitative paradigm, it was decided to use the documentary research method to collect and systematize theories and approaches of authors. This method is based on literature to explore theoretical contributions on a specific topic and thus determine its most relevant characteristics [8]. To this end, some criteria were established to carry out the documentary review, such as scientific articles published in the main current journal in English or Spanish and with at least 5 to 8 years of validity for current affairs, without this implying the exclusion of classical theoretical references.

The flexibility of this type of method allows the researcher to reflect on the relevance and theoretical relevance of the topic addressed, given that he or she will be able to compare, analyze and compile the different perspectives of the authors, thus nurturing the process of critical analysis and the contributions that it can provide [9].

3. DATA MINING: EVOLUTION AND CONCEPT

At the end of the 1970s and beginning of the 1980s, the marketing area of organizations began to worry about market information, where the staff specialized in statistics understood that the applications and potentialities in the business world were beginning to be very well received [10]. Conceptually, data mining can be defined as a process in which a set of techniques are used to detect patterns, process, and extract useful information from large volumes of data, whether structured or unstructured, that cannot be easily found by traditional methods of analysis [11, 12].

Likewise, some authors such as Ezhilarasan and Prasad [13] indicate that data mining is part of a broader process called knowledge discovery in Database por (KDDD), which places a high value on the technique, given that, in the end, it is not only information, but also knowledge. On the other hand, according to the branch to which it belongs, some authors refer to data mining as a branch of computer science whose objective is to discover patterns in a high volume of data [14]; in addition, data mining opens the way to new processes related to the area of data analysis, such as business intelligence (Business Intelligence or BI for its, artificial intelligence (AI) [15].

3.1 Data Mining Methodologies

There are several methodologies in the literature for performing data mining processes, however, there are two major methods for achieving this purpose. On the one hand, the SAS Institute proposes the SEMMAS (Sample, explore, model, modify and assess) [16] methodology, while on the other hand, the CRISP-DM (Cross-Industry Standard Processor Data Mining) methodology has been widely used and is organized in six fundamental stages similar to the one formulated by SAS, namely: understanding the problem, understanding the data, data preparation, modelling, evaluation and exploitation [17].

3.1.1 SEMMAS Methodology

Created by the information analysis company SAS in coordination with the data mining software developed by the company, it allows an organized exploration of the data to understand and exploit the patterns identified in large volumes of data (Dutta and Bose, 2015). This methodology is shorter and simpler and covers the following stages:

- Phase 1. Sample: The purpose of this first phase is to work with a reduced set of data that comply with the representativeness of the population. With this, it is possible to reduce the time and cost of processing the data with which it will be worked and thus make the process more efficient, however, care must be taken that the sample is statistically representative of the data universe.
- Phase 2. Exploration: Once the set of data with which the analysis will be carried out has been determined, it is necessary to evaluate them to detect anomalous data, errors, inconsistencies, and other factors that may alter the subsequent results. To this end, various statistical techniques can be used to corroborate the integrity of the data, as well as visualization techniques to support data analysis and behavior.
- Phase 3. Modification of the data: Once the structure and purification of the data has been defined, the necessary transformations of the variables will be carried out at this stage, according to the models and techniques that will be used later. In other words, depending on the model used, the data and variables must have a specific structure which must be constructed and adapted.
- Phase 4. Modeling: It is in this phase that data mining techniques are most relevant, as predictive models are built. This is a process that involves all the data in the modified set to achieve the expected results, hence the importance of phases 2 and 3 of the methodology.

- Phase 5. Evaluation: The models produce results that must be validated according to the requirements of the project and the success criteria, i.e., if the outputs of the model respond to the problems raised in the initial phase. This validation is performed on the sample being worked on, which, if successful, can be implemented for the total data population. Figure 1 shows in summary the sequence of the phases of the SEMMAS methodology.

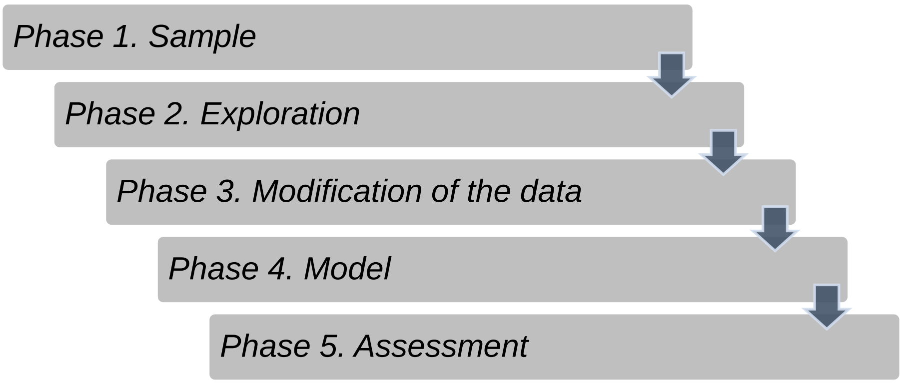

Figure 1. Phases of the methodology SEMMAS

3.1.2 CRISP-DM Methodology

This methodology was created by SPSS, NCR and Daimer Chrysler in early 2000 and, according to statistics, is the most widely used reference for the development of data mining projects (Shafique and Qaiser, 2014). A brief description of each of the phases of the methodology will follow:

- Phase 1. Understand the business or problem: It corresponds to one of the key phases of the whole process, as it outlines and defines the objectives and requirements of the project and stakeholders, and then converts them into technical objectives, with a logical sequence of activities for the project plan. It is at this stage that the problem should be studied as rigorously as possible, observing the possible results, involvement, limitations, resources, among other aspects, so that the data needed to solve it can be collected efficiently. Without a correct definition and delimitation of the problem, the collection, processing, and analysis of data will most likely fail or fail to meet customer expectations.

- Phase 2. Understanding the data: Once the problem has been well defined, it is necessary to start with the data collection phase, based on the following questions: What data are required to solve the problem? is the data structured in an organized database? what are the different sources of the data? In this way it is then possible to become familiar with the structure and quality of the data, identifying limitations and future constraints and other elements that may pose a risk to the analysis process.

- Phase 3. Preparing the data: This phase is key to adapting the data based on the techniques to be used in data mining. The fundamental support of this stage is, in the first place, the objectives set in the initial phase, as they will guide the techniques that best fit the analysis process. In addition, validation and cleaning processes should be carried out to ensure that the data structure is as reliable as possible.

- Phase 4. Modeling: Depending on the objectives, the relevant modeling techniques are proposed according to the following criteria: agreement with the problem (phase 1), adequate data for analysis (phase 2), knowledge of the modeling technique in terms of results and interpretation.

- Phase 5. Evaluation: Once the model has been implemented, its performance is evaluated according to the objectives. In this sense, the criteria for success are given by the answers obtained, that is, whether the model answered the initial questions and met the objectives set out. In addition, it is important to evaluate the reliability of the model and its goodness of fit to ensure that the outputs are reliable and can be used as a basis for decision making.

- Phase 6. Implement: With the analysis model evaluated, it is then implemented in the organization, where the documentation of the process to create the model is fundamental to make future adjustments. The implementation involves technical sub processes and training for personnel using this type of solution, so that it can be used and manipulated. Sequentially, the phases of the CRISP-DM methodology can be seen in Figure 2.

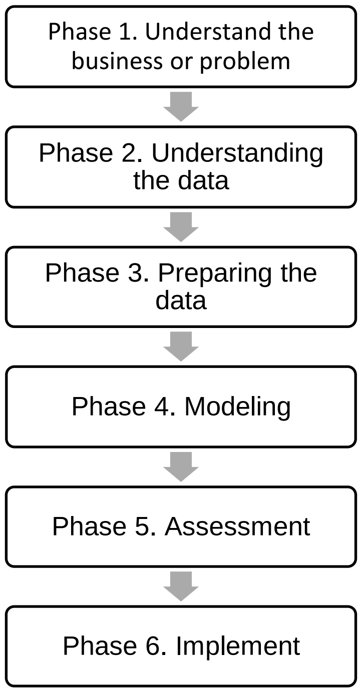

Figure 2. Phases of the methodology CRISP – DM

3.1.3 Data mining techniques

Regardless of the data mining methodology used, the techniques provide a set of analysis tools for each case or problem being addressed. The review of the literature indicates that there are numerous techniques that can be used in this regard, however, they are basically defined in three main groups: predictive, descriptive, and auxiliary techniques [18].

Predictive and descriptive techniques are widely developed in the statistical literature, and are basically composed of regression analysis, time series, discriminant analysis, neural networks, among others, while descriptive techniques form the group of techniques that allow the most important characteristics and features of the data set to be highlighted [19]. On the other hand, the auxiliary techniques correspond to methods of verification and extraction of information, such as analytical processes of transactions (OLAP), queries and reports, among others. To illustrate this classification, a brief outline of the most common techniques based on the literature review is presented below.

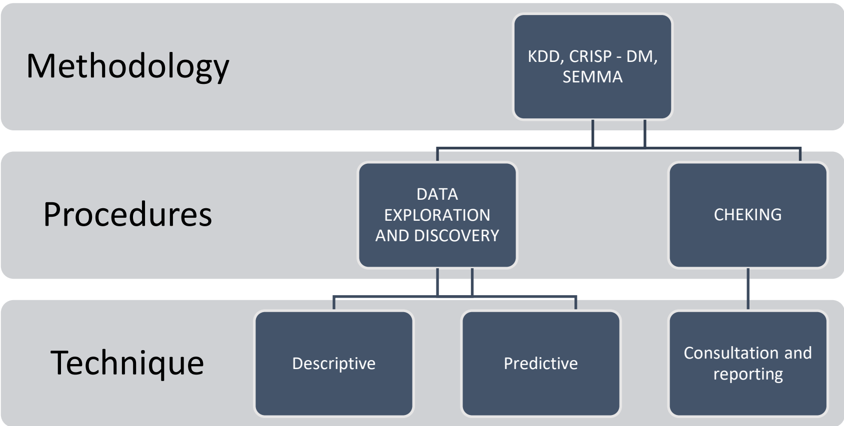

Figure 3. Classification of data mining techniques

In general, there are many data mining techniques, however, what is important is their usefulness and the degree of coherence between them and the objectives of the analysis [20]; it is important to clearly define the problem being studied and based on this information, to define the most suitable technique.

4 CHALLENGES OF DATA MINING IN COLOMBIAN SMES

The continuous changes in the economic environment mean that SMEs are increasingly adopting innovative measures to overcome the difficulties they face, including data mining [21]. There is now a growing interest in using and employing these techniques, supported without a doubt by the data that is continuously generated in day-to-day operations [22]. However, to take advantage of these analytical processes, it is not enough to have a large volume of information, given that if it is not related or processed, it will hardly open the way to new knowledge for decision making, and therefore would represent a waste of opportunities for business growth [23].

The areas in which data mining can make a significant contribution are particularly those in which large amounts of data are generated. In trade and banking, predictive analysis of customer preferences and needs can facilitate the design of new products or improve existing ones [24], and credit risk models also provide relevant information on the consumer profile to predict, with some degree of confidence, whether it is or not feasible to extend credit [25].

In more specialized sectors such as medicine, biochemistry, industry and pharmaceuticals, the applications are numerous, since based on the data generated it is possible to create better and innovative procedures for continuous improvement and innovation. Ultimately, there are many areas where useful information can be extracted to enhance the quality of life, processes, and decision-making [26]. In this way, in the Colombian sphere, great opportunities deriving from this type of orientation can be appreciated, however, the use of these guidelines will depend to a large extent on the decision to accept these mechanisms, since if no decisions are taken in this regard and plans are made in the short and medium term, it will be difficult for SMEs to meet the demands of the issues reviewed.

5 CONCLUSION

Undoubtedly, data mining techniques represent an opportunity for the growth and development of SMEs, not only in economic terms, but also in terms of the knowledge management generated in the processes, helping leaders to better understand their organization and make better decisions [20]. However, the implementation and adoption of these techniques requires commitment, training, and investment in the structure of information systems to ensure efficient, quality and reliability of data.

One of the challenges for the SME that wants to implement these solutions is to integrate the different sources of data generation, to have clarity in the organizational objectives that allow to formulate the correct questions and to have the qualified personnel to plan, design, implement and evaluate the analytical models; therefore, in this last aspect, the responsibility of the training in this area falls fundamentally on the educational institutions, who are called to create spaces to face the new demands of the market.

REFERENCES

[1] Gandomi, A., Haider, M.: Beyond the hype: big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144 (2015).

[2] Hamidi, H.: Decision-Making Processes Used by Malaysian Managers of Small and Medium-Sized Enterprises When Deciding Their Participation in Government Sponsored Training Programs. Global Business and Management Research: An International Journal, 13(2) (2021).

[3] Palma, H. H., Sierra, D. M., Arbeláez, D. C.: Enfoque basado en procesos como estrategia de dirección para las empresas de transformación. Saber, Ciencia y Libertad, 11(1), 141-150 (2016).

[4] Zhong, R. Y., Huang, G. Q., Lan, S., Dai, Q. Y., Chen, X., Zhang, T.: A big data approach for logistics trajectory discovery from RFID-enabled production data. International Journal of Production Economics, 165, 260-272 (2015).

[5] Niebles-Núñez, W. A., Hernández-Palma, H. G., Cardona-Arbeláez, D.: Gestión tecnológica del conocimiento: herramienta moderna para la gerencia de instituciones educativas. Revista de Investigación, Desarrollo e Innovación, 7(1), 25-36 (2016).

[6] Martínez, D. L. L. R., Gómez, C. E. P.: Metodología de Estudio del Rendimiento Académico Mediante la da Minería de Datos. Campus Virtuales, 3(1), 56-73 (2015).

[7] Estrada, M. R.: Manejo de problemas y toma de decisiones (Vol. 8). Editorial El Manual Moderno (2015).

[8] Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., Lichtendahl Jr, K. C.: Data Mining for Business Analytics: Concepts, Techniques, and Applications in R. John Wiley & Sons (2017).

[9] S. Carrasco, G. Chinguel, M. Cubas, R. Cieza.: El estudio y la investigación documental: Estrategias metodológicas y herramientas TIC. Gerardo Chunga Chinguel, (2017).

[10] Bhowmik, S., Chattopadhyay, R., Chatterjee, U.: A Review on Security Measures in Data Mining. Imperial Journal of Interdisciplinary Research, 2(6) (2016).

[11] Ezhilarasan, M., Prasad, A. M.: Data Mining in Human Resource Information Systems (HRIS). JARDCS, (7) 301-308 (2017).

[12] Wuang. Y, Kung. L, Ting. C, Byrd. T.: Beyond a Technical Perspective: Understanding Big Data Capabilities in Health Care. 48th Hawaii International Conference on System Sciences, (2015).

[13] Ristoski, P., Paulheim, H.: Semantic Web in data mining and knowledge discovery: A comprehensive survey. Web semantics: science, services and agent son the World Wide Web, 36, 1-22 (2016).

[14] Kasemsap, K.: The role of data mining for business intelligence in knowledge management. Integration of data mining in business intelligence systems, 12-33 (2015).

[15] Martínez, A.: Acercamiento al proceso de toma de decisiones en las PYMES del sector industrial de la ciudad de Duitama, Boyacá, Colombia. FACE: Revista de la Facultad de Ciencias Económicas y Empresariales, 16, 48-60 (2016).

[16] Cardona, T., G. Brand, L. Marín.: Visualizaciones analíticas para la toma de decisiones en pequeñas y medianas empresas utilizando Data Mining. Cuaderno Activa, 8(8) 1-40 (2016).

[17] Dutta, D., I. Bose.: Managing a big data project: the case of ramcocements limited. International Journal of Production Economics, 165, 293-306 (2015).

[18] Shafique, U., H. Qaiser.: A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). International Journal of Innovation and Scientific Research, 12(1) 217-222 (2014).

[19] Gutiérrez, J., B. Molina.: Identificación de técnicas de minería de datos para apoyar la toma de decisiones en la solución de problemas empresariales. Revista Ontare, 3(2) 33-51, (2016).

[20] Levine, D., T. Krehbiel, M. Berenson.: Estadística para administración. Pearson Educación, (2014).

[21] Coleman, S.: Data-mining opportunities for small and medium enterprises with official statistics in the UK. Journal of Official Statistics, 32(4) 849-865 (2016).

[22] Niebles-Núñez, W. A., Hernández-Palma, H. G., Cardona-Arbeláez, D.: Gestión tecnológica del conocimiento: herramienta moderna para la gerencia de instituciones educativas. Revista de Investigación, Desarrollo e Innovación, 7(1) 25-36, (2016).

[23] Parody A. et al.: Aplicación de un diseño central compuesto de superficie de respuesta para la determinación de la planitud en las láminas de acero de un acero colombiano. En: Tan Y., Shi Y., Tang Q. (eds.) Minería de datos y Big Data. DMBD 2018. Notas de la conferencia en Ciencias de la computación, vol 10943. Springer, Cham (2018).

[24] Troncoso-Palacio A., Neira-Rodado D., Ortíz-Barrios M., Jiménez-Delgado G., Hernández-Palma H.: Using Discrete-Event-Simulation for Improving Operational Efficiency in Laboratories: A Case Study in Pharmaceutical Industry. In: Tan Y., Shi Y., Tang Q. (eds) Advances in Swarm Intelligence. ICSI 2018. Lecture Notes in Computer Science, vol 10942. Springer, Cham (2018).

[25] Joshi, A., Pandey, N., Chawla, R., Patil, P.: Use of data mining techniques to improve the effectiveness of sales and marketing. International Journal of Computer Science and Mobile Computing (IJCSMC), 4, 81-87, (2015).

[26] Olson, D. L., Wu, D. D.: Data Mining Models and Enterprise Risk Management. In Enterprise Risk Management Models (pp. 119-132). Springer, Berlin, Heidelberg, (2017).