DEVELOPING A PEDAGOGICAL METHOD TO DESIGN INTERACTIVE LEARNING OBJECTS FOR TEACHING DATA MINING

Diana Montoya, Instituto Tecnológico Metropolitano
Jorge Eduardo Hincapie Vargas, Politécnico Innova—ITS Colombia
Jonathan Sanchez Giraldo, Politécnico Innova—ITS Colombia
Natalia Correa Hincapié, Instituto Tecnológico Metropolitano

ABSTRACT

This article presents the design of a pedagogical method to produce interactive learning objects that allow users to learn about new statistical process control techniques, such as Data Mining, in order to address the curricular needs of college students enrolled in academic programs related to quality engineering. Acquiring skills in this type of tools enables students to have a more competitive profile to work at different organizations.

Keywords Teaching object, quality engineer, data mining, statistical data, interactive learning object.

INTRODUCTION

Technology-based education can support pedagogical processes in the fields of humanities and exact sciences, where several learning objects are needed to teach very specific topics. Furthermore, it has become increasingly necessary to learn about computation time, the treatment of great volumes of information, the complexity of calculations, the speed of processes and answers, and the accuracy of results of data treatment by means of machine processing.

In Colombia, four higher education institutions offer undergraduate programs in quality engineering: Politécnico Colombiano Jaime Isaza Cadavid, Productivity and Quality Engineering; Universidad Pontificia Bolivariana, Quality Engineering; Fundación Universitaria Panamericana, Quality Management; and Instituto Tecnológico Metropolitano, Quality Engineering. They are available in Medellín and Bogotá, capital cities home to a significant percentage of the manufacturing and service industries in the country (Ramirez & De Aguas, 2015).

The graduate profiles of all the programs above are oriented toward process improvement and project development and evaluation. Their curricula are mainly focused on statistical systems; nevertheless, no data mining course is taught to evaluate processes and projects based on the data that companies generate and their use of computer technologies.

In order to evaluate the progress of the study and the use of data, it is important to review the history of philosophy and mathematics and, more specifically, the early conceptions of knowledge and learning (see Table 1).

Table 1. Early Conceptions of Knowledge and Learning

Author	Idea(s)
Descartes (1596–1650) and Leibniz (1646–1716)	The mind is connected to the physical world.
John Locke (1690)	In the beginning there was the mind.
Hume (1779)	Knowledge is a product of perception, acquired by experience (induction), and represented by logical theories.
Darwin (1859)	The theory of evolution by natural selection.

However, such notions changed later in time, when mathematical thinking was philosophically based on formal rules (see Table 2).

Table 2 Formal Rules for Mathematical Thinking

Author	Idea(s)
Boole (1815–1864) and Frege (1848–1925)	Fundamentals of mathematical logic.
Gödel (1906–1978) and Turing (1912–1954)	Limits to computability (incompleteness theorem).
Fermat (1601–1665), Bernoulli (1700–1782), and Bayes (1702–1761)	Probability and probabilistic reasoning.

Providing aspiring engineers with an education that sparks their creativity, dynamism, and skills acquisition to face and lead technological changes is essential for higher education institutions. Universities should motivate engineers to work for continuous improvement and resource optimization in a globalized and competitive environment and achieve maximum competitiveness supported by technologies and permanent knowledge transfer. The competences acquired by engineering students should contribute to solve problems based on their being, knowing, and doing, as well as solid training in basic sciences and the integration of contemporary curricular contents (Tres Palacios Ortiz, et al., 2008).

The current competitive and globalized environment generates large amounts of information that is subsequently used to create new knowledge. However, an analysis of the graduate profiles of quality engineering programs at higher education institutions in Colombia suggests that future engineers should orient their learning process toward techniques, methods, and tools to treat such information and support and improve engineering practices. Artificial Intelligence has opened up countless subfields such as data mining, which is responsible for treating massive amounts of information and generating new queries to make decisions. Additionally, it enables users to automatically forecast and diagnose information, which represents a great advantage for companies.

To produce a certain learning outcome, a set of problem situations, solution criteria, and activities are necessary. Moreover, according to Irigoyen Jiménez, & Acuña (2011), different types of problems are related to the development of specific skills. Conceptual problems require activities to analyze and explain the phenomena being studied. Methodological problems deal with the instrumentation of procedures and measurement to study relevant variables and their systematic operation. Technological and axiological situations highlight the adaptation and application of scientific knowledge to evaluate and solve problems in social contexts.

This article describes a method to create Interactive Learning Objects (ILOs) in order to teach quality engineering enrollees and professionals in related fields how to solve problems by applying data mining (DM) techniques. These ILOs combine the acquisition of necessary decision-making skills with technological scenarios adapted to computational environments and information and communications technologies that introduce important changes to the academic activity (Horruitiner Silva, 2012). Furthermore, such digital resources (ILOs) include several components: content, activities, contextual elements, and an external (metadata) structure that makes them easy to store, identify, and retrieve. The set of digital self-contained entities addressed in this article will be designed to fit into an actual study plan.

In Colombia, the use of data mining is largely ignored by college students in programs other than software engineering. Nevertheless, a well-rounded education should consider connections and complements between different fields of knowledge that strengthen learning processes by combining different tools. Consequently, this study aims to raise awareness about the advantages of DM in multiple fields.

PROCESS QUALITY CONTROL

One of the main topics of study in quality engineering and related fields is process quality control, which encompasses the measurement and data acquisition of processes and activities in order to control defective products and operating ranges. Statistical process control emerged as a science that studies the behavior of a process based on data to provide information that determines objective criteria, thus controlling quality and meeting product and process specifications.

There is a clear connection between statistical process control and DM, which refers to the application of specific algorithms to extract patterns (models) from the data. DM involves a series of steps: data preparation, selection and cleaning, incorporation of previous knowledge, and interpretation of results.

According to Villanueva Manjarres, Moreno Sandoval, and Salinas Suárez (2018), data mining is an "emerging discipline that seeks to develop methods to explore vast amounts of data in order to better understand behaviors, interests, and results" (p. 235). In turn, Mitra and Acharya (2003) argue that

> *data mining tasks can be descriptive, (i.e., discovering interesting patterns or rela-tionships describing the data) and predictive (i.e., predicting or classifying the be-havior of the model based on available data). In other words, the objective of this interdisciplinary field is to predict outcomes and reveal relationships between data. It uses automated tools and sophisticated algorithms to discover hidden patterns, associations, anomalies, and/or structures in large amounts of data stored in data warehouses or other repositories in order to subsequently filter the necessary in-formation from this big dataset (p. 107).*

Within this theoretical framework and in order to review the available literature on DM teaching, the authors searched online databases using terms such as statistical process control, DM, quality, virtual learning, and different synonyms. This process retrieved few documents related to the topic under study, which suggests there are benefits in creating an object to teach DM to quality engineers and professionals in similar fields.

QUALITY CONTROL AND DATA MINING

According to Rendón (2013), statistical quality control is the inspection of processes to determine the quality of products from each one of the operations that are conducted to manufacture them; at the same time, data are monitored and captured to determine variations and thus establish the necessary corrective and improvement actions.

Different statistical tools that allow users to establish the phenomena and behavior of each process are employed to determine these controls. Specifically, data mining offers a solution to define equations for specific company processes and take the corresponding corrective and improvement actions. ILOs are essential tools that adopt cognitivism to have students master a new technology and learn how to apply it to the solution of real problems at product- or service-based companies.

Specifically, the method proposed in this work to create ILOs follows the trend of cloud computing because higher education institutions can provide not only software but also flexible and scalable technological capacity to all users through web technologies. Cloud computing allows students to understand the importance of distance, mobility, and risk and reduces tuition cost; moreover, such services are generally standardized and promote dynamic interaction in machine processing.

TEACHING METHODS

Cognitivism defines knowledge acquisition as a mental activity that involves the internal coding and structure of the student's mind. Furthermore, cognitive theories conceptualize student learning processes and examine the way information is received, organized, stored, and retrieved, thus turning the student into an active participant (Ertmer, & Newby, 1993).

According to cognitivism, teaching practices are aimed at students acquiring knowledge. In particular, Davini (2008) indicates that, regardless of the contents, teaching processes try to simultaneously favor the assimilation of knowledge and the development of thinking skills, and different approaches are adopted for that purpose. They can be divided into four big groups: inductive, instructional, and cognitive flexibility and conceptual change. Such general frameworks can be analyzed independent of the context and concrete users and, in the field of education, they can be studied, reconstructed, and combined with each other for specific purposes according to the field of study or application being taught. Cognitivism, in addition, supports the practice of analyzing a task and breaking it down into manageable parts and setting objectives and then measuring performance based on such objectives (Fatemi Aqda, Hamidi, & Ghorbandordinejad, 2011).

This work combines two approaches: inductive and instructional. The first is used to introduce the concepts of DM and statistical quality control by means of data observation, management, organization, and utilization. The second focuses on the assimilation of knowledge in order to apply procedures of DM to statistical process control.

Both approaches use virtual spaces as learning environments and ILOs as mediators in the teaching process.

García, Portillo, Romo, and Benito (2007) maintain that current pedagogical models require student-centered learning environments to meet the criteria of adaptability, dynamism, and orientation. In those scenarios, professors are mediators who make a significant contribution to the acquisition of the skills the students need for their professional development. Additionally, instructors should present the processes in an organized manner so that knowledge of the topic is acquired, autonomous learning is promoted, and new teaching tools are implemented. The method proposed in this work fulfills those conditions, and it constitutes an effective framework that enables students to learn the concepts of DM and solve problems in that field.

INTERACTIVE LEARNING OBJECTS

The Ministry of National Education of Colombia defines ILOs as digital resources that can be reused in different educational contexts; among other types, they are courses, images, pictures, movies, videos, and documents that have clear educational objectives (Colombia Aprende, 2012).

As previously mentioned in the words of another author, ILOs are entities (digital or otherwise) that can be used, reused, or referenced for technology-supported learning, and they offer flexibility, customization, modularity, adaptability, reutilization, and durability. To guarantee student learning, ILOs must exhibit several characteristics: accomplishment of educational goals, ease-of-use, quality of the audiovisual environment, interaction with contents, and quality content (Callejas Cuervo, Hernández Niño, & Pinzón Villamil, 2011).

Several standards are available to create ILOs. In particular, our pedagogical method adopts the Sharable Content Object Reference Model (SCORM), which fulfils several requirements:

> *It provides access through web-based technologies, it is adaptable to the needs of individual users and organizations, it is permanent regardless of the evolution of technology, and it can be used in different types of platforms (interoperability) and reused in several applications and contexts. The objective of said standard is to establish a reference model to create objects with structured teaching contents, thus facilitating their exchange in different educational systems (Callejas Cuervo et al., 2011, p. 180).*

ILOs are, in turn, complementary tools that support inclusion processes and enable different people to access teaching and learning resources. Hence, they work as essential tools to promote the improvement of educational processes, equality, and the diversification of education (Alquati Bisol, Valentini, & Rech Braun, 2015). Moreover, ILOs should be flexible enough to reduce the duplication of work and costs among institutions, which is essential in planning and developing distance learning courses (Falcão, De Moraes, & Rossato, 2016).

METHOD

The process for developing the pedagogical method proposed in this study was divided into three stages. Each one of them comprised specific activities and outputs in order to enable the creation of ILOs to teach data mining to students of quality engineering programs and similar fields. During the first stage, the learning need was identified in the curricula of undergraduate quality engineering programs in the country. The second stage was the application of cognitivism to guide the design of ILOs, focusing on competency-based learning and combining inductive and instructional methods to teach DM. Additionally, a general scheme of the ILOs was created at this point. Finally, the method was applied based on ideograms used as instructions to create ILOs.

CREATING INTERACTIVE LEARNING OBJECTS TO TEACH DATA MINING TO QUALITY ENGINEERING STUDENTS

First, the authors searched for studies in the field of quality and summarized the results in Figure 1. The occurrence of the terms units of measurement and quality control can be observed from 2004 to the present. These expressions were mostly used in 2016, when they were mentioned in more than 20 publications in the literature. These data suggest the importance of incorporating DM and information technologies into the curricula of quality engineering programs.

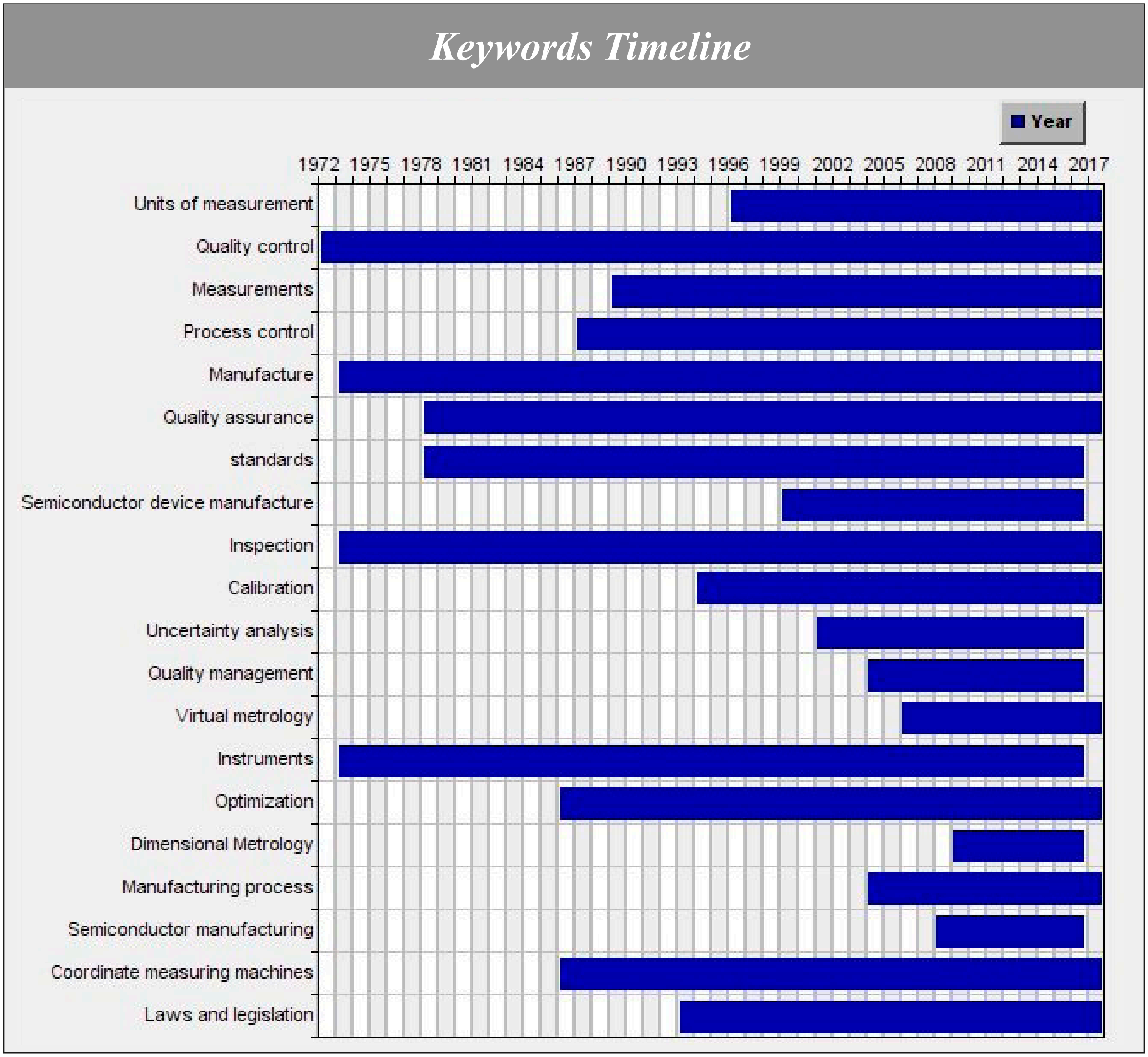

Figure 1. Occurrence of different terms in quality engineering studies.

Cognitivism is based on the idea that human beings are information processors and that the human mind and the workings of a computer share similarities. Both of them, mind and computer, process data in a knowledge or learning society. Furthermore, Ertmer and Newby (1993) concluded that

> *cognitive theories are devoted to the conceptualization of student learning processes and the way information is received, organized, stored, and retrieved. Learning is related not so much to what students do but to what they know and how they acquired that knowledge (p. 9).*

For that reason, and in order to develop a pedagogical method, we analyzed and selected specific inductive and instructional approaches that embrace cognitivism and enable students enrolled in quality engineering and similar programs to acquire DM skills. Figure 2 presents the relationships between the methods that were integrated for that purpose.

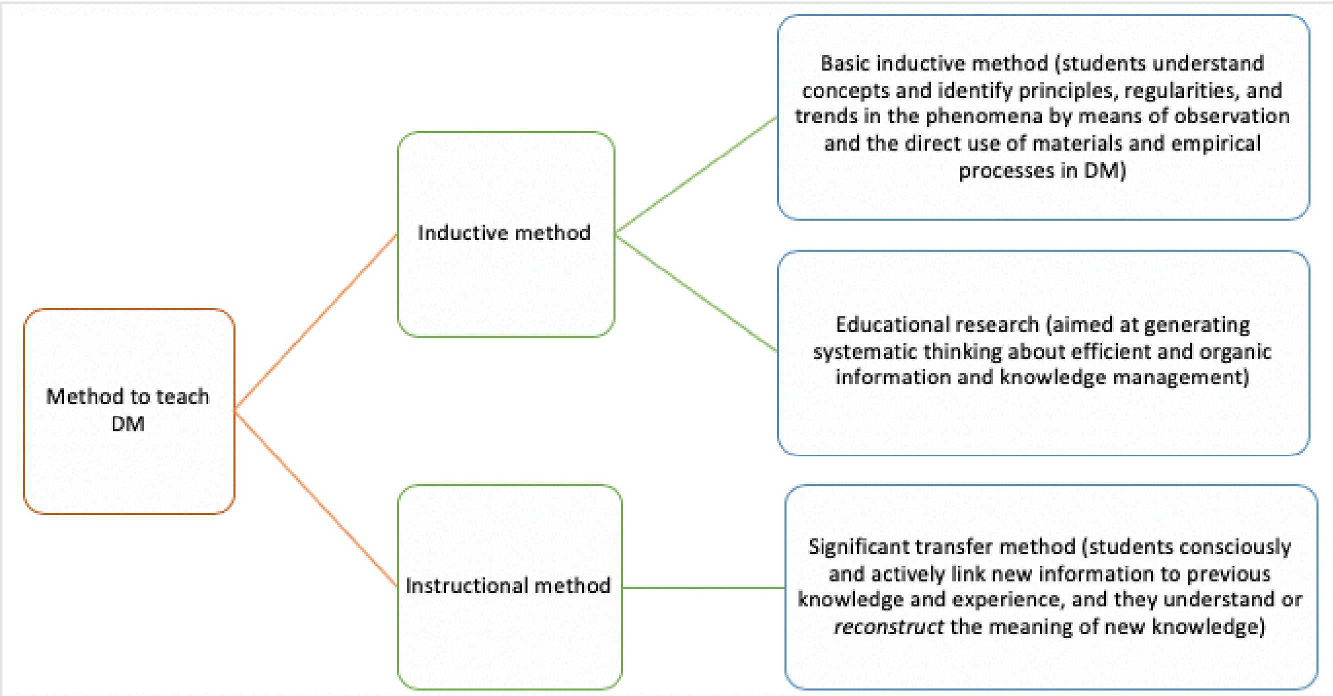

Figure 2. Integration of methods to teach DM.

Cognitivism and evaluation are closely linked by competences because cognitivism examines the way students acquire knowledge through basic inductive method and educational research and how they apply their knowledge and skills (Fernández March, 2010). According to cognitivism, skills are transferred and developed when students understand how to apply the knowledge they acquired in different contexts through the significant transfer method. This process is supported by ILOs as they use specific cases and students' solutions. Other authors support the idea that intelligence in the "knowledge society" is modeled by electronic machines that process and store data to simulate routine activities of the human brain (Runing, Schraw, & Norby, 2004)

In line with the theories above, several components were selected to develop the pedagogical model in this work: identification of the teaching method, the target population, the educational objective, content writing, pedagogical strategies, learning exercises and activities, score evaluation, pedagogical and content requirements for multimedia specialists, and test grading, all based on knowing, doing, and being. As indicated by the teaching methods in Figure 2, the ILOs that implement our method should describe real cases experienced at organizations with concrete and comparable results between DM applications and traditional statistical control techniques. Therefore,

Figure 3. Academic components of the interactive learning object.

such ILOs will apply two of Piaget's principles: learning by doing and learning from example.

The authors used the skill-based pedagogical model adopted by the Metropolitan Institute of Technology in Medellín (Cadavid Alzate & Urrego Giraldo, 2005) as a basis to develop their own method. Figure 3defines the academic components in the interactive learning object and details the elements that compose the method proposed to develop and design ILOs. They are classified into the three groups defined in the knowing-doing-being framework.

Knowing

In this cognitive field, the method is oriented toward mental constructs that enable understanding and subsequent learning. Multiple pedagogical strategies were applied to design the method, and its components can be planned, controlled, and defined according to the target population. In the proposed method, elements such as data, facts, information, opinions, and knowledge are contextualized as follows:

- **Method:** Identify the teaching method.
- **Target population:** Quality Engineering students.
- **Educational objective:** The Educational objective is to develop students' skills in the use of cutting-edge DM technologies that enable them to obtain, filter, and use their results for decision making regarding quality processes at an organization.
- **Content writing:** The main topics selected for students are introduction to Data Mining, algorithms, structures, models, test and validation, queries, solutions, architecture, and tools. Each one of these elements will be strictly focused on quality engineering processes.
- **Online learning strategies:** The flexibility of the proposed pedagogical method has effects on several aspects, such as accessibility, location, space, pacing, class hours, access to an ever-growing number of technologies, and the importance of tuition loans to pay for course credits.

Doing

Technological knowledge acquisition in the method above is conceived as a process of knowledge retention and development supported by ILOs at different levels of complexity and depth. The objective is that students acquire DM skills, abilities, and techniques to apply and translate their knowledge into actions inside organizations in dissimilar fields. For that purpose, the following elements are necessary:

- **Learning exercises and activities:** These are designed using digital texts that guide students as they practice the techniques they have learned, and they consciously use deductive strategies for the given examples. Such learning activities aim at more cognitive strategies, as defined by Oxford (1990), that use meaning inference techniques in each application to a particular case. This type of activities has students see and understand the usefulness of adopting their own learning strategies.
- **Evaluation of virtual test results:** The results of the strategies that are implemented should be assessed to determine if they have successfully promoted learning. For that purpose, ILOs should provide evaluation tools to validate the stages of awareness, acquisition, and knowledge transfer, as well as students' understanding of all the concepts, methods, and theoretical ideas covered in each learning unit.
- **Pedagogical and content requirements for multimedia specialists:** The technological tools necessary to implement the ILO are selected considering the available software, computers, servers, video cameras, and repositories of digital resources. In addition, the environments, learning activities, tests, and avatars to be designed are defined.
- **Test grading:** After the evaluation is submitted, the system should enable students to automatically see their scores and corrections as well as their mistakes or achievements. The instructor will be responsible for managing and configuring the grading tool as well as creating options that specify the rules to interact with students. Likewise, the tool should describe the steps the faculty should follow to configure questions in the test.

Being

The virtual environments of ILOs created by applying their method should motivate students to develop autonomously their full intellectual potential. Additionally, such environments will be transformed and improved to enable learning while integrating inclusive educational practices that result in learning opportunities regardless of personal background. Moreover, this pedagogical method integrates concepts of the theory of computational linguistics proposed by Berwick & Chomsky (2016) that discusses the representation of knowledge and language grammar.

- **Communication with the interactive learning object**: The topic is introduced using different digital resources so that learners assume responsibilities regarding their interaction with the object.

Figure 4 presents the process and the skills students gain when learning about DM. Users are guided step-by-step to achieve the learning outcome, starting with understanding a topic and setting the objectives. Subsequently, students are taught how to create a dataset according to the objective by selecting and integrating target data from multiple and heterogeneous sources. Afterward, they learn how to clean and process data, selecting and using skills to manage the missing information in the datasets because so-called dirty data can commonly deceive the mining process and lead to incorrect or unreliable results. Finally, students learn how to apply algorithms, which enable them to process a dataset to obtain new information about it.

CONCLUSIONS

New educational alternatives that address the needs and interests of university programs should be explored to satisfy the training requirements of organizations. As a result, the development of ILOs is expected to promote and define pedagogical strategies to teach DM. Fortunately, the latter includes flexible processes that can be adapted to different educational stages, cases, and situations.

Turing, along with Newell and Simon, claimed that any cognitive agent, such as a person or machine able to acquire knowledge, has information input units, a memory, a processor, and output units. In opposition to Descartes' ideas, knowledge is no longer conceived as being restricted to humans, and computers, in conjunction with ILOs, can hence be considered cognitive agents.

Cognitive theories are devoted to the conceptualization of student learning processes and the way information is received, organized, stored, and retrieved. Instructional explanations, demonstrations, examples, and counterexamples are tools that should guide student learning. Consequently, the ILOs designed adopting the pedagogical method proposed in this work to teach DM should meet the educational needs of students by applying cognitivism and integrating multiple learning tools. As a result, users will be able to acquire and identify important DM concepts and principles and transfer significant knowledge, thus demonstrating they understand, i.e., reconstruct, the meaning of the new knowledge.

This pedagogical method also satisfies the criteria of new teaching models in its adaptability, dynamism, use of new environments, development-centered pedagogy, student management, and support of the role of professors as mediators.

An ILO designed implementing the method discussed in this work is currently undergoing an experimental stage. The objective is to integrate such an ILO into an actual course with online educational content, tools, and virtual environments. After several months of testing, the results will

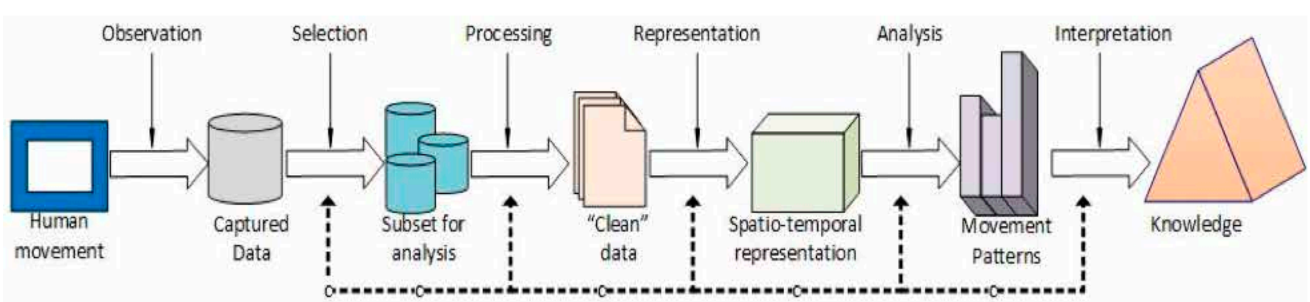

Figure 4. Process for teaching data mining.

be evaluated to establish the effectiveness of this method.

Data evaluation and treatment offer a wide range of possibilities. DM, a subfield of AI, is currently revolutionizing the world and supporting the development and maximization of corporate results and processes. Quality engineering graduates should possess extensive knowledge of this subfield because they will need to implement its methods, techniques, and tools in their work in order to find specific parameters for the activities of different organizations to support their continuous improvement.

REFERENCES

Albaiges Olivart, J. M. (2015). El poder de la memoria: Metodo infalible para mejorar la memoria. Barcelona: El Aleph.

Alquati Bisol, C., Valentini, C. B., & Rech Braun, K. C. (2015). Teacher education for inclusion: Can a virtual learning object help? Computers & Education, 85, 203–210. doi:10.1016/j.compedu.2015.02.017

Berwick, R. C., & Chomsky, N. (2016) Why only us: Language and evolution. Cambridge, MA: MIT Press.

Cadavid Alzate, G., & Urrego Giraldo, M. I. (2005). Cuaderno de Pedagogía. Medellín: ITM.

Callejas Cuervo, M., Hernández Niño, E., & Pinzón Villamil, J. N. (2011). Objetos de Aprendizaje, un estado del arte. Entramado, 7(1), 176–189. Retrieved from http://www.scielo.org.co/pdf/entra/v7n1/v7n1a12.pdf

Carmignato, S. (2012). Accuracy of industrial computed tomography measurements: Experimental results from an international comparison. CIRP Annals, 61(1), 491–494. doi:10.1016/j.cirp.2012.03.021

Colombia Aprende. (2012). Colombia Aprende. Retrieved from http://aprende.colombiaaprende.edu.co/sites/default/files/naspublic/articles-318264_recurso_reda.pdf

Davini, M. C. (2008). Métodos de enseñanza. Didáctica general para maestros y profesores. Buenos Aires: Santillana.

Ertmer, P. A., & Newby, T. J. (1993). Conductismo, cognitivismo y constructivismo: Una comparación de los aspectos críticos cos desde la perspective del diseño de insutrucción. Performance Improvement Quarterly, 6(4), 50–72. Retrieved from https://www.galileo.edu/faced/files/2011/05/1.-ConductismoCognositivismo-y-Constructivismo.pdf

Falcão, E., De Moraes, M., & Rossato, J. (2016). Evaluation of virtual objects: Contributions for the learning process. International Review of Research in Open and Distributed Learning, 17(6), 62–74. doi:10.19173/irrodl.v17i6.2620

Fatemi Aqda, M., Hamidi, F., & Ghorbandordinejad, F. (2011). The impact of constructivist and cognitive distance instructional design on the learner's creativity. Procedia Computer Science, 3, 260–265. doi:10.1016/j.procs.2010.12.044

Fernández March, A. (2010). La Evaluación orientada al aprendizaje en un modelo de formación por competencieas en la educación universitaria. Revista de Docencia Universitaria, 8(1), 11–34. Retrieved from http://red-u.net/redu/files/journals/1/articles/144/public/144-130-2-PB.pdf

García, F., Portillo, J., Romo, J., & Benito, M. (2007). Nativos digitales y modelos de aprendizaje. Universidad del País Vasco / Euskal Herriko Unibertsitatea

Horruitiner Silva, P. (2012). La Universidad Cubana: El modelo de formación. La Habana: Editorial Universitaria del Ministerio de Educación Superior.

Irigoyen, J. J., Jiménez, M. Y., & Acuña, K. F. (2011). Competencias Y Educación Superior. Revista Mexicana de Investigación Educativa, 16(48), 243–266. Retrieved from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662011000100011

Mitra, S., & Acharya, T. (2003). Data mining: Multimedia, soft computing and bioinformatics. Hoboken, NJ: John Wiley & Sons.

Ramirez, J. C., & De Aguas, J. M. (2015). Escalafón de la competitividad de los depar-tamentos de Colombia, 2015. (CEPAL, Ed.) Cepal—Serie Estudios y Perspectivas, 91.

Rendón C., H. D. (2013). Control estadístico de calidad. Medellin: Universidad Nacional de Colombia.

Runing, R. H., Schraw, G. J., & Norby, M. N. (2004). Cognitive psychology and instruction (4th ed.). Upper Saddle River, NJ: Pearson.

Tres Palacios Ortiz, C. A., Rada Espinosa, M. E., Hernández Flores, M. M., Guerrero Hernánde, z. T., Loóez, L., & Martha. (2008). Curriculum y docente: Encuentro de significados. Educación Médica Superior, 22(2), 1–6. Retrieved from http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412008000200005

Vanniarajan, S. (1990). Language Learning Strategies: What Every Teacher Should Know by Rebecca L. Oxford. Issues in Applied Linguistics

Villanueva Manjarres, A., Moreno Sandoval, L. G., & Salinas Suárez, M. J. (2018). Data mining techniques applied in educational environments. Digital Education Review, 33, 235–266.