Recommendation system as a business intelligence tool developed in the Jupyter notebook. Case: selection of productive personnel

Sistema de recomendaciones como herramienta de inteligencia de negocios desarrollado en el cuaderno de Jupyter. Caso: selección de personal productivo

ROMERO, Raúl†, VILLANUEVA, Guadalupe, MARTÍNEZ, Sergio* and LÁZARO, Pablo

Tecnológico de Estudios Superiores de Chalco, División de Ingeniería Informática.

ID 1st Author: *Raúl, Romero*

ID 1st Coauthor: *Guadalupe, Villanueva*

ID 2nd Coauthor: *Sergio, Martínez*

ID 3rd Coauthor: *Pablo, Lázaro*

Abstract

Recommendation Systems are a useful tool in an organization's decision-making process, because they base their recommendations on applied mathematical models. The organization's historical information-based Recommendation Systems make it possible to make even more accurate decisions because of the personal nature of the information that feeds them. Thanks to the various information technologies that exist, the creation of Customized Recommendation Systems according to the needs of the organization is of relative ease. The purpose of this article is to show the procedure for creating and operating a recommendation system developed in the Jupyter Notebook as a business intelligence tool that supports business decision-making applied to an organization-specific personnel selection case.

Business Intelligence, Jupyter Notebook, Automatic Learning, Decision Making, Recommendation System

Resumen

Los sistemas de recomendación son una herramienta útil en el proceso de toma de decisiones de una organización, ya que basan sus recomendaciones en modelos matemáticos aplicados. Los sistemas de recomendación históricos basados en la información de la organización permiten tomar decisiones aún más precisas debido a la naturaleza personal de la información que los alimenta. Gracias a las diversas tecnologías de la información que existen, la creación de sistemas de recomendación personalizados de acuerdo con las necesidades de la organización es de relativa facilidad. El propósito de este artículo es mostrar el procedimiento para crear y operar un sistema desarrollado en Jupyter Notebook como una herramienta de inteligencia empresarial que apoya la toma de decisiones empresariales aplicada a un caso de selección de personal específico de la organización.

Inteligencia de Negocios, Cuaderno de Jupyter, Aprendizaje Automático, Toma de Decisiones, Sistema de Recomendaciones

* Correspondence to Author (email: mdti.martinez@tescha-informatica.net)
† Researcher contributing first author.

Introduction

On May 26, 2016, the Technological of Higher Studies of Chalco signed a collaboration agreement with the Association of Industrialists and Entrepreneurs of the Volcanoes Region of the State of Mexico, ASIVE, whose purpose is to establish ties of cooperation and institutional participation with In order to establish collaboration and support agreements between the productive and educational sectors, this agreement seeks to combine technological and educational efforts in favor of regional development, allowing the link between the academic sector and the ASIVE member companies. The creation of the System of Recommendations developed in the Jupyter Notebook as a Business Intelligence tool that supports business decision making applied to a case of selection of specific personnel of an organization, is developed by TESCHA, educational institution of higher education , in order to support the decision-making process of a company belonging to ASIVE, which for reasons of information protection has been called: SAMEX CIA. This company provided the TESCHA group of researchers with a history of their information regarding the productive work performance of their former and current employees. The database provided by the company consists of 112 records, including circumstantial data of its employees as well as the level of productivity obtained by each employee.

The use of the Jupyter Notebook as a Business Intelligence tool represents for the organizations that employ it, the possibility of having a cluster of strategies and instruments that allow the extraction, transformation and analysis of their information in order to improve their internal processes of decision making.

The foregoing is mainly based on the fact that Jupyter's Notebook is a development environment that integrates libraries of scientific computation that allow entrepreneurs to extract and analyze information from various sources such as: social networks, transactional systems, government sources, websites, and own historical information among others. This information when analyzed and processed allows to support the decision-making process but this time based on the study of the data.

This article aims to show one of the many applications of the Jupyter Notebook in the business sector. In this case, the application shown is the creation of a recommendation system that supports the business decision-making process. The application case shown in this article is designed by the authors of the same who are part of the Research Line called: "Business Intelligence: Datawarehouse, Datamining and Knowledge Management"; which is a line of research belonging to the Computer Engineering degree at TESCHA. The case designed by the researchers is a business strategy that relies on a system of recommendations that seeks to be applied in the last stage of the personnel selection process that companies develop, with the distinction that said strategy focuses on the selection of personnel with some specific characteristic, which, in this case, is the characteristic of labor productivity.

The development of this work is contextualized in a business organization belonging to ASIVE, so it takes a model for a company dedicated to food packaging.

The topic of Personnel Selection is addressed in this work, because for the reference company (as for most companies), the human resource is one of its main assets, since it is the personnel, the transcendental resource that develops activities of creation and transformation that pay directly to the scope of business objectives; Therefore, if this resource is chosen in an appropriate manner, the company may have duly motivated collaborators willing to contribute their creativity, knowledge and skills to the overall development of the company.

Due to the above, it is very important to select employees correctly, ensuring a job profile according to the requirements and needs of the organization; It is here that the use of a system of recommendations for the selection of personnel would allow, (analyzing the necessary data) to identify those elections that present a lower degree of uncertainty for the company.

In general, there are three internal benefits that this work seeks to provide: the first one seeks to establish collaboration and contribution ties between the academic sector and the productive sector of the region, which contributes to the objectives of the collaboration agreement concluded between TESCHA and ASIVE; the second; it is obtained by generating part of the scientific production of the research line of the educational institution; and the third benefit is that working with modern technologies demanded by companies today allows the educational institution to train professionals trained in the use of cutting-edge technologies.

The external benefit that this work provides is in the way in which the use of the Jupyter Notebook will impact the business decision-making processes, allowing organizations to adopt new and better business intelligence technologies to carry out the development of their management activities when discovering the generation of value through the finding or not of patterns in your information.

Finally, and in a general way, this project seeks to show the utility of a system of recommendations developed in the jupyter notebook within the business field, specifically as a business intelligence tool by offering support in the business decision-making process.

Methodology

Derived from the collaboration agreement between TESCHA and ASIVE, a brief investigation was carried out in each of the associated companies in order to know the level of use of technological tools in decision-making processes at the managerial level. The research information was obtained through the analysis of the answers given by the entrepreneurs to 19 questions contained in the questionnaire. The questions asked allowed, among other things, to identify the characteristics of the associated companies as well as their knowledge and use of technological tools for business intelligence in the development of their business activities.

Regarding the characteristics it is important to mention that there were 33 companies that provided information, of which 76% belong to the tertiary sector and 24% to the secondary sector.

The turns of the companies vary between food production, advertising, health sector, textile branch among others. With respect to the technological tools possessed, 84% have computer equipment, 60% have internet service, 33% use some specialized software, while 27% have computer servers.

However, the analysis of the information obtained allowed TESCHA to determine the following regarding the decision-making process: 1. When making decisions within your company, 54% of companies never use any of the methodologies for making of decisions. 2. That 42% of companies do not know any decision-making model. 3. That 90% of companies know Excel software while 0% know Jupyter's Notebook. 4. That 63% of entrepreneurs do not use specialized tools for decision-making mainly because they are unaware of their existence.

The answer to the previous question shows the main problem faced by entrepreneurs, which is the lack of knowledge of the existing technological tools to support the business decision-making process.

The offer for the resolution of this problem focuses on minimizing the ignorance of these technologies in entrepreneurs through the demonstration of the operation and application of a System and Recommendations as a Business Intelligence tool based on the Jupyter Notebook.

The resolution offer also seeks that through knowledge and subsequent use of the recommendation system, companies improve their internal processes while increasing their benefits.

The chapters that integrate this work are: Objectives, Theoretical Framework, Procedure, Results and Conclusions.

Objectives general

Show the creation and operation of a recommendation system as a business intelligence tool developed in Jupyter's notebook that supports business decision making applied to a case of personnel selection.

Specific

- Gather the Anaconda Software Ecosystem programs version 3-5.2.0-Windows-x86_64 for the installation of the development environment in the computer equipment destined for the realization of the project.

- Install and verify the correct functioning of the Anaconda Software Ecosystem version 3-5.2.0-Windows-x86_64 on the computer equipment destined to carry out the project.

- Run the "Jupyter Notebook" program, and import libraries for scientific computing and data analysis necessary for the operation of the Jupyter notebook.

- Apply the quality process to the data (remove headers, empty cells and unify values) that integrate the database for analysis.

- Make the connection to the database to extract the information for analysis.

- Define the variables to analyze, filter and export them for processing.

- Carry out the definition and functions of the algorithm to be used.

- Apply the supervised machine learning algorithm.

- Obtain the recommendations issued by the recommendation system.

- Interpret the recommendations thrown by the system.

Theoretical framework

This section aims to provide the theoretical concepts necessary to understand the development of the system of recommendations described in this article.

The concepts presented range from the fields of knowledge used to the specifications of the tools used.

Fields of knowledge

The creation of a recommendation system developed as a business intelligence tool, derives directly from knowledge contained in the field of Information Technology, as this is a knowledge generating information system, however, and because the recommendation system works with data, which are analyzed and worked through mathematical models, the field of Mathematics is also included; finally and derived from the fact that this system of recommendations seeks to be applied to a business model, it is included in the field of Administration. Therefore, the areas of knowledge of these three sciences involved in the development of the system of recommendations, as well as the correlation between such knowledge, are presented below.

Computer science.

According to Chiavenato 1999, computer science is the science that manages information through the use of digital tools in order to transform and convert it into knowledge. Considering the above, it is interesting to know the various fields in which information technology is divided. According to the model of (Heckhausen, 1975.) There are 5 fields that integrate computer science according to its material domain, which are: 1. Data / Information / Knowledge. 2. Systems. 3. Algorithms. 4. Information / Knowledge Systems and 5. Information and Communication Technologies. The previous classification allows to identify the branch to which this project belongs, which is branch 4. Information / Knowledge Systems. The following figure is intended to show this classification. Information / knowledge systems. An Information / Knowledge System "is a set of interrelated components that collect, process, store and distribute data and information and provide a feedback mechanism to meet an objective. This mechanism is what helps organizations achieve their goals "(Stair & Reynolds, 2000. p. 52). It should be noted that one of the main objectives of organizations is to correctly apply the decision-making process, since the above allows the organization to generate a climate of certainty based on information, in addition to generating competitive advantage and therefore better yields One of the main examples of information / knowledge systems focused on business decision making are the Recommendation Systems.

Recommendation Systems.

"The recommended systems consist of software tools and techniques used to provide suggestions in an automated way, about topics (items) that may be useful, to users of certain computer environments." (Cingolani, 2014, p. 11). To this, it can also be added that the recommendation systems aim to offer useful information to the interested party from the analysis of a large amount of information, which, due to its volume, could not be analyzed effectively by the user. The recommendation systems provide the user with information with a high degree of certainty because their answers are based on processes of machine learning.

Machine learning.

Machine learning is understood as the development of various techniques through which computers are instructed so that they are capable of performing tasks similar to those performed by human beings, including language comprehension, image recognition , object classification and decision making, among others. Therefore, machine learning is considered:… "as a branch of artificial intelligence whose objective is to develop techniques that allow computers to learn, that is, to generalize behaviors and knowledge based on information provided as an example" (Gago, 2017, p. 23). It is important to note that machine learning has two major divisions: supervised and unsupervised learning; the first one is named this way because the nature of the result to be obtained is known, since said result has its origin from a set of previous known information; the second, unsupervised learning is called that because it only provides general information without emphasizing the labels or specifications of the information, which is why the results cannot be known or intuited in advance. In the construction of this system of recommendations supervised learning is applied, because the data provided for the analysis are known by the organization.

Supervised learning.

"It is called that because it trains with a set of data for which the result is known, and the algorithm can use this output value to optimize its operation, and they return a function that will predict the correct result from that moment with greater or less precision "(Luque, p. 23).

The above definition allows both the analysis data and the algorithm to be used to be identified. According to its own definition, an algorithm is composed of a set of logical and articulated instructions that allow certain activities to be carried out. There are different types of algorithms, however, one of its main classifications is according to their functionality, which allows them to be classified as follows: a) sorting algorithms. b) search algorithms and c) routing algorithms. Due to the nature of this recommendation system, the algorithm used for its operation is the K-NN algorithm itself, which belongs to the classification of sorting algorithms.

K-NN algorithm.

"It is based on the assumption that the closest prototypes have a similar a posteriori probability. If Ki (X) is the number of class samples present in the k neighbors closest to X. "(García y Gómez, 2012. p. 4). Therefore, the K-NN (K-Neighbors Neighbors) or Nearest Neighbors algorithm is a supervised classification algorithm that is based on the analysis of a set of training data. In the case of application of the system of recommendations shown in this article, the training data that serve as the basis for the analysis of the algorithm are data provided by the company used as a model; Due to the fact that said company seeks to know the level of labor productivity in future cases of personnel selection, the information shown corresponds to the historical data contained by the Human Resources department in accordance with the tags of: SEX, AGE, LEVEL_ACADEMIC , STATE_CIVIL, TRANSFER_TIME_IN_MINUTES, DISTANCE_EN_KMS, AUTO_PROPIO, NUMBER_ OF_MARIMAL_WAYS as well as that of PRODUCTIVITY_LEVEL, which is the label that provides the recommendation. As a reminder, the fact that the System of Recommendations shown in this work belongs to the field of knowledge 4. Information / Knowledge Systems according to the Heckhausen model is highlighted, since it is in this field that its foundation is theoretical, however for its development process, the System of recommendations as a Business Intelligence tool developed in the Jupyter Notebook. Case: Productive Personnel Selection, is located in field 5. Information and Communication Technologies of the same model.

Information and communication technologies.

As can be seen in the procedure section, for the implementation of the system of recommendations, it is necessary to use various elements of information and communication technologies. According to the above, the first thing that needs to be determined is the use of the Software Ecosystem to be used. "A software ecosystem is a work space in which a series of tools coexist that, together with good practices, allow a development team to model a work methodology" (Recena, 2019). A Software ecosystem is composed among other elements, mainly of Programming Languages, libraries and Development Environments.

Anaconda.

The ecosystem used for the development of the recommendation system is Anaconda in its version 3.5. Anaconda is a business data science platform employed by leading business information technology professionals or by data scientists. Data science finds this platform as basic because it has several libraries focused on the development of machine learning. The following 3 Anaconda elements were used for the development of this project: 1. Python Programming Language. 2. Numpy and Pandas bookstores. 3. Jupyter Notebook Development Environment (Jupyter Notebook).

Python programming language.

Python is a simple syntax programming language with the ability to run on a large number of platforms. "Python is a very powerful and flexible general purpose programming language, while simple and easy to learn. It is a high-level language, which allows you to easily process all types of data structures, both numerical and text. "(Caballero, 2019). The simplicity feature of Python is one of the main features that allowed the implementation of the recommendation system.

Bookstores numpy and pandas.

The libraries are code files with different functionalities already established, which can be added to the programs to enrich its activities.

In the case of the development of this project, the libraries of Numpy and Pandas were used whose specific characteristics and functions are the following: Numpy: "Fundamental package for scientific computing with Python that is characterized by containing an N-dimensional matrix object for linear algebra and integrates with a wide variety of databases "(Numpy.org, 2019.) Pandas:" It is used for operations and manipulations of structured data… applied to the phase of purification and preparation of the data. It is a recently added library, its great utility has boosted the use of Python in the scientific community".

Jupyter Notebook.

The Jupyter Notebook is a tool that is used to carry out operations and manipulations of structured data, focused on debugging and preparing them. One of its main features is to have open source that allows you to create and share documents that contain live code, equations, visualizations and narrative text. "Its uses include: cleaning and data transformation, numerical simulation, statistical modeling, data visualization, machine learning." (Project Jupyter, 2019). The above concepts are a general and specific part of the computer field. The concepts used for the development of the recommendation system in the field of mathematics are presented below.

Mathematics.

According to the more generalized definitions of mathematics as a discipline of study, it can be determined that mathematics is an exact science that, through logical reasoning, studies the properties and quantitative relationships between abstract entities, these relationships are studied through their various branches. The classification contemplates arithmetic, algebra, trigonometry, probability and statistics, and calculus integrate as the main branches of mathematics. The mathematical foundation of the present work is in the field of algebra, mainly linear algebra.

Linear algebra.

ROMERO, Raúl, VILLANUEVA, Guadalupe, MARTÍNEZ, Sergio and LÁZARO, Pablo. Recommendation system as a business intelligence tool developed in the Jupyter notebook. Case: selection of productive personnel. RINOE Journal-Financial Economy. 2019

Prior to the concept of linear algebra, the concept of algebra defined by Aurelio Baldor stands out as: "branch of Mathematics that studies the quantity in the most general way possible" (Baldor, 2017. p.5.).

Among the main areas of study of algebra are linear algebra, abstract algebra, vector algebra, tensorial algebra, multilateral algebra, homological algebra, commutative algebra, differential algebra, Boolean algebra and elementary algebra. Linear algebra is a branch of mathematics that studies concepts such as vectors, matrices, systems of linear equations among others. Because the recommendation system data is worked on a matrix basis, it is necessary to identify the concepts of Matrices and Vectors, which are applied in the development of the system.

Matrices.

According to Howard Anton, a matrix is: "an array of numbers. The numbers in the array are called the matrix elements".

This concept is cited because, as is known, the function of a matrix is to locate the data in a two-dimensional manner, a model in which the data analyzed by the recommendation system.

There is a concept of utmost importance for the implementation of the recommendation system, this concept is even the one that directly links the field of mathematics with the field of computer science.

The concept is algorithm. In the understanding that an algorithm is a logical, systematized and orderly sequence of instructions that when applied mean a solution to a given problem, it is important to mention that in the creation of the recommendation system, the algorithm is the sequence of activities Once applied, they allow the classification of the data, which, when analyzed, generate knowledge. It is important to remember that in the case of this work, the algorithm used in the creation of the recommendation system is the K-NN algorithm, which is addressed in section 1.1.4 of this section.

Administration

The administration is the third field of knowledge applied in this work, because the system of recommendations made focuses on the resolution of a need presented by an organization. According to Hurtado, "The administration is a process through which resources are used that are scarce in order to achieve proposed objectives." (Hurtado, 2008. p.41). The Administration is a social science that has presented several changes in a relatively short time, these changes are best appreciated by identifying the various approaches from which it has been studied. The following image shows an outline of the various approaches to administration. As the previous figure shows, the administration study has been carried out based on at least 8 different approaches, including: classical approach, humanistic approach, neoclassical approach, structural approach, management behavior approach, approach systemic, situational approach, as well as the most current approach called: New approaches to management. It is important to mention that the realization of this project is located in the last approach.

The new management approaches are integrated by a compendium of innovative actions focused on the specific resolution of an administrative topic. As part of these new approaches is Knowledge Management.

Knowledge Management.

"Knowledge management is a logical, organized and systematic process to produce, transfer and apply in concrete situations a harmonious combination of knowledge, experiences, values, contextual information and expert insights that provide a framework for its evaluation and incorporation of new experiences and information. "(Davenport and Prusak, 2001. 77). Another definition is the following:" Process of generating value from the intangible assets of an organization "(Fontalvo, Quejada and Puello, 2011, p. 81). The foregoing is important to highlight that knowledge management implies the generation of value for the organization, a value that intrinsically relates to the nature of the business or to what is considered as Business Intelligence.

Business Intelligence.

"BI Business Intelligence (Business Intelligence) is a tool under which different types of organizations can support decision-making based on accurate and timely information; guaranteeing the generation of the necessary knowledge that allows choosing the alternative that is most convenient for the success of the company. "(Rosado y Rico, 2010. p. 321). Considering the above, the relevance of the concept is observed for the implementation of the recommendations system, since it is through this that the organization seeks to provide relevant information considered as knowledge. A concept that is directly related to Business Intelligence is decision making.

Decision Making.

Decision-making is understood as the application of a systematic process through which the organization chooses an alternative among several others that are characterized by containing a lower degree of uncertainty and a greater degree of viability. However, it is worth mentioning that in most organizations this process has very little specialization to the extent that in many cases organizations do not have an established and specialized method for decision making; This translates into various negative aspects that impact the organization, either in terms of productivity, positioning or obtaining returns. Currently, the condition described above has presented a series of advances, both theoretical (such as the birth of methodologies for decision-making) and practical (such as the creation of recommendation systems.

Recommendation systems.

A recommendation system is considered as an intelligent system that issues information through suggestions (recommendations) on a given topic. The main feature of these systems is that the recommendations they issue are made based on the processing of their own information, that is, personal recommendations are issued in accordance with the analysis of the specific information.

The specification of the thematic concepts allows to have a theoretical overview of the basic concepts used in the construction System of recommendations as a Business Intelligence tool developed in the Jupyter Notebook.

It remains only to clarify that the system of recommendations developed in this article focuses on providing recommendations to the company that was used as a model, specifically in the Productive Personnel Selection process; process that is developed in the area of Human Resources of the company.

Process.

This section will show the procedure used to develop the recommendation system. The steps and activities that were carried out for its development are shown progressively according to the process of creating the recommendation system.

Gather the anaconda software ecosystem programs version 3-5.2.0-windows-x86_64 for the installation of the development environment in the computer equipment destined for the realization of the project.

To gather the necessary programs it is necessary to download the Anaconda software 3-5.2.0-Windows-x86_64, which must be downloaded from the official Anaconda website https://www.anaconda.com/distribution/ of the download section, which is available for Microsoft Windows, MacOS and Linux platforms. In this work, the Anaconda version 3-5.2.0-Windows-x86_64 available for Microsoft Windows versions has been used. The following image shows the download section of the Anaconda software. (See figure 1).
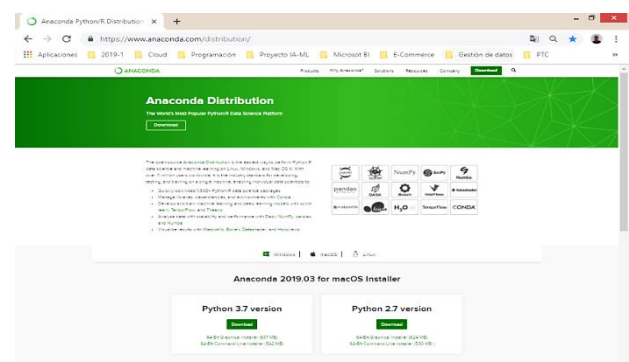
Figure 1 Anaconda Software Download
Source: Anaconda official site

Install and verify the correct functioning of the anaconda software ecosystem version 3-5.2.0-windows-x86_64 on the computer equipment destined to carry out the project.

Once the Anaconda program version 3-5.2.0-windows-x86_64 has been installed, you must run the Anaconda Navigator file that has been located as a shortcut on the Windows desktop, then the following window should appear. (See Figure 2).
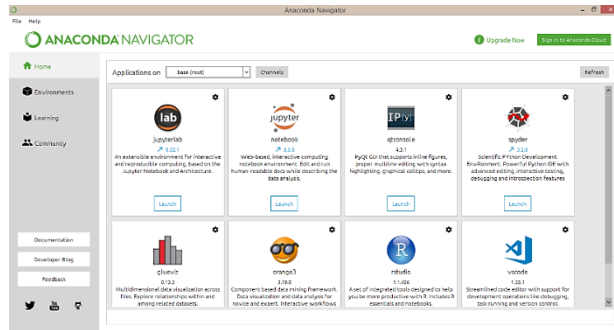
Figure 2 Anaconda Software Execution
Source: Anaconda official site

Execute the "Jupyter Notebook" program and import libraries for scientific computing and data analysis necessary for the operation of the Jupyter Notebook.

Once Anaconda Navigator is open, the Jupyter Notebook program will be found in the main panel, which must be executed by clicking on the launch button. Once the Jupyter Notebook is opened, the New option will be selected and from the drop-down list that will appear, the Python3 option must be selected, then the corresponding code must be written to import the Numpy and Pandas libraries. The following image shows the code needed to import the mentioned libraries. (See Figure 3).
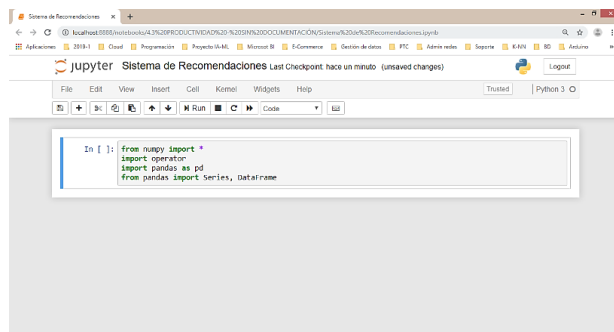
Figure 3 Execution of the Jupyter Notebook
Source: System of recommendations developed in the Jupyter Notebook

Apply the quality process to the data (remove headers, empty cells and unify values) that make up the database for the analysis.

The process of applying quality to the data lies in reducing inconsistencies at the time of performing the data analysis, this process carries out activities such as elimination of headings, elimination of empty cells and unification of criteria regarding the capture of the data; This is because sometimes the data comes from different sources. In the case of the database obtained from the company, model of this article, the elimination of the head of the main sheet, the title in which the columns are classified and the unification of the values of the level of productivity were applied. The following image shows part of the data for the analysis to which the quality process has been applied. (See Figure 4).

Figure 4 Application of the Data Quality Process
Source: Database

Make the connection to the database to extract the information for analysis. Anaconda version 3-5.2.0-windows-x86_64 is compatible with various platforms because it integrates multiple libraries that connect to most existing databases. In the present work the following instruction was used:

bd = pd.read_excel ('productivity.xlsx', skiprows = [1]) to connect the Jupyter Notebook with the database of the model company, which was in an Excel software spreadsheet. The following image shows the connection of the Jupyter Notebook with the spreadsheet and the data obtained. (See Figure 5).
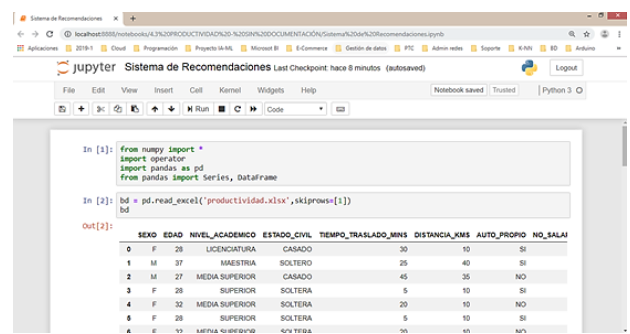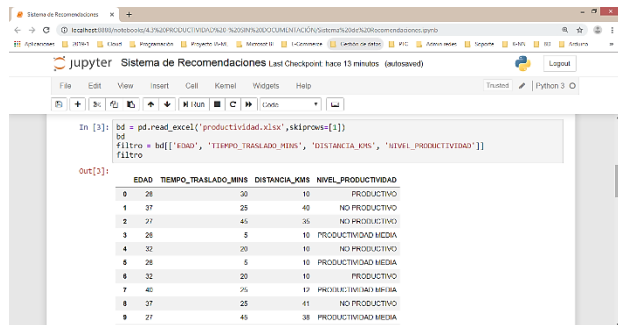
Figure 5 Database connection
Source: System of recommendations developed in the Jupyter Notebook

Define the variables to analyze, filter and export them for processing. In this section, the fields of the database that will feed the algorithm of the closest neighbors (K-NN) must be identified, which should be divided into two parts, dataSet and labels. The dataSet, represent the characteristics of the data set and the labels the possible results of the algorithm, then the result must be exported to a text file in csv format that is the format that supports the algorithm. (See Figure 6).
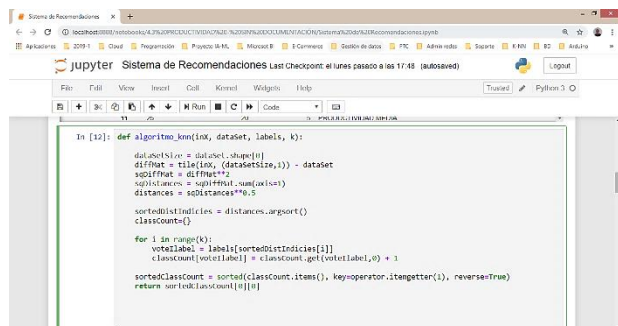
Figure 6 Variable Definition
Source: System of recommendations developed in the Jupyter Notebook

Make the definition and functions of the algorithm to be used. This section refers to the definition of the functions of the recommendation system which consists of:

a) Function of the K-NN algorithm

b) Conversion function

c) Standardization function.

K-NN algorithm function. The following image shows the corresponding code. (See Figure 7).
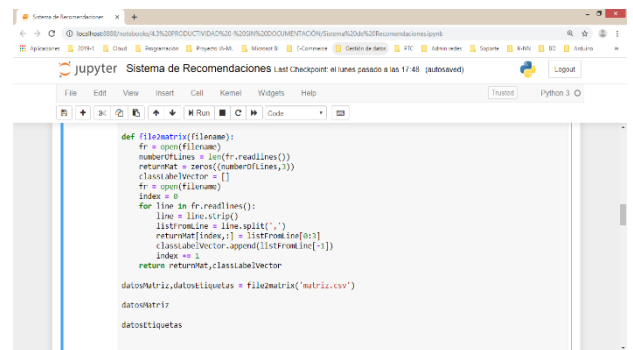
Figure 7 K-NN algorithm definition
Source: System of recommendations developed in the Jupyter Notebook

The code: "takes four inputs: the input vector to classify called inX, the complete array of training examples called dataSet, a vector of labels called labels, and, finally, k, the number of closest neighbors to use." (Harrington, 2012, p. 23).
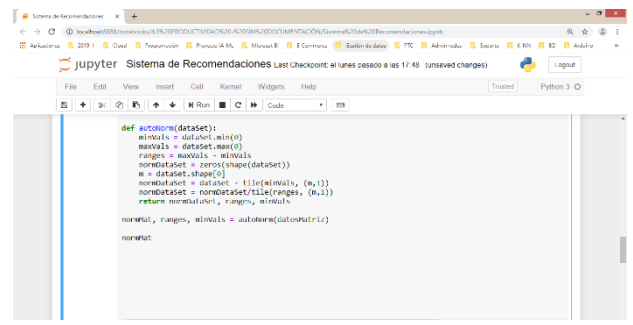
b) Conversion function. The following image shows the code of the conversion function that converts the output file generated in step 6 to the format that the K-NN algorithm classifier accepts. (See Figure 8).

Figure 8 Conversion function
Source: System of recommendations developed in the Jupyter Notebook

c) The function "takes a string from the file and generates two things: an array of training examples and a vector of class tags." (Harrington, 2012, p. 25).

d) Normalization function. The following image shows the code that allows to normalize the values of the generated matrix to 0 and 1, when these are in different ranges: (See Figure 9).
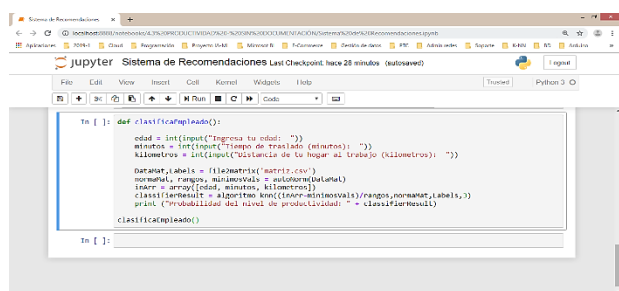
Figure 9 Normalization Function
Source: System of recommendations developed in the Jupyter Notebook

The following describes the function where: "get the minimum values of each column and place this in minVals; Similarly, get the maximum values.

The 0 in dataSet.min (0) allows you to take the minimums of the columns, not the rows. Next, calculate the range of possible values seen in the data set and then create a new matrix to return it. To obtain the normalized values, subtract the minimum values and then divide them by the range". (Harrington, 2012, p. 30).

Apply the supervised machine learning algorithm. This section introduces the new values (employee characteristics) that will be evaluated by the algorithm and will then be classified according to the number of closest neighbors with similar characteristics. The following image shows the code where the values of: age, travel time in minutes and distance in kilometers are entered and the probability of the level of productivity is calculated. (See Figure 10).

Figure 10 Algorithm Application
Source: System of recommendations developed in the Jupyter Notebook

Obtain the recommendations issued by the recommendation system. Once the algorithm has been applied to the new data, the result (recommendation) will be displayed on the screen, which will indicate the level of productivity according to the following:

Recommendation
a) Productive.

Recommendation.
b) Average Productivity.

Recommendation.

c) Non-productive.

The following Figure shows the recommendation options based on the traffic light model. (See Figure 11).

RANGOS DE LA RECOMENDACIÓN:
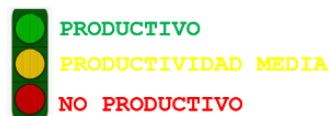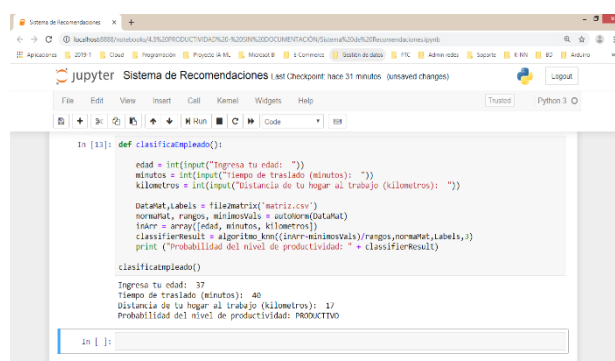
Figure 11 Recommendation ranges
Source: own elaboration

Interpret the recommendations thrown by the system. The following image shows the entry of the new value or new data to be classified, and also shows the result in a recommendation format given by the recommendation system. (See Figure 12).

Figure 12 Recommendation Result
Source: System of recommendations developed in the Jupyter Notebook

Results

The following section shows the results obtained from the process of creating the Recommendations System as a Business Intelligence tool developed in the Jupyter Notebook. Case: Selection of Productive Personnel. It should be noted that this section of results is shown through images that show the development of the system directly from the application, that is directly from the Jupyter Notebook. It is equally important to highlight that because the Jupyter Notebook is a versatile and complete tool, this one within its same structure allows to be commented, which will allow the total follow-up of the creation system of the recommendations system until the results process.

ROMERO, Raúl, VILLANUEVA, Guadalupe, MARTÍNEZ, Sergio and LÁZARO, Pablo. Recommendation system as a business intelligence tool developed in the Jupyter notebook. Case: selection of productive personnel. RINOE Journal-Financial Economy. 2019

Figure 13

Figure 14

Figure 15

Figure 16

Figure 17

Figure 18

Figure 19

Figure 20

Figure 21

Figure 22

Conclusions

Companies are complex entities that derived from their own activities, are immersed in the development of various business activities.

Decision-making based on a systematic, scientific and grounded process on their own experiences and characteristics allows these organizations to generate value, mainly because it reduces times, efforts and levels of uncertainty. The above characteristics allow the organization to be more competitive over the others. However, the use of technological tools that allow and facilitate the decision-making process, based on own information, are not yet as common among the business circle, which minimizes the possibility that companies are at the forefront. Most of the ASIVE Associated companies do not use technological tools applied and specialized to the decision-making process, the main reason is the ignorance that the business field has of these tools; However, it is also attributed to the fact that if businessmen know these tools, they think they are complex. The advantages of making decisions based on the company's own elements are diverse, without this meaning that the external business environment does not matter, however when the decision making process is focused on the analysis of the company's own information Decisions made are full of greater certainty that is based on the experience itself. Such is the case as exemplified in this work. The company that served as a model provided a database of 112 records that contained current and own information on the characteristics of its employees because its main problem was the low productivity of its staff, some general and particular characteristics were taken into account in order to determine the previous probability of productivity in potential employees.

Referenes

Chiavenato, I, (2007) "Administración de Recursos Humanos", 8va Edición, Cd. de México: McGraw-Hill/Interoamericana editores, S. A. de C. V. p 169

Heckhausen, Heinz, (1979), "Disciplina e interdisciplinariedad", en Apostel, Leo et al., Interdisciplinariedad, Problemas de la enseñanza y de la investigación en las universidades, ANUIES, México.

Barchini, G, Sosa, M, Herrera, S, (2012) "La informática como disciplina científica. Ensayo de mapeo disciplinar" Argentina, Universidad Nacional de Santiago del Estero.

Stair, R., & Reynolds, G. (2000). Sistemas de información. Enfoque administrativo, México. Ed CENGAGE Learning p. 52

Cingolani, E, Evaluación de los Sistemas Recomendadores de Contenidos Educativos a través de Estudios de Usuarios. (Magister en Tecnología Informática) Facultad de Tecnología Informática Universidad Abierta Interamericana 2014, p. 75

Gago, R, Uso de algoritmos de aprendizaje automático aplicados a bases de datos genéticos, Tesis (Máster universitario en Bioinformática y Bioestadística), Universitat Oberta de Catalunya, 2017

Luque, A, Análisis de sistemas complejos usando Machine Learning. Tesis (Grado en Matemáticas), Universidad de Sevilla. 2018 p. 23

García, C, Gómez, I, Algoritmos de aprendizaje: K-NN & KMEANS Inteligencia en Redes de Telecomunicación, Universidad Carlos III de Madrid p. 4.

Project Jupyter, Jupyter (Apr 12, 2019), recuperado de: https://jupyter.org/

Baldor A, Álgebra, 2017, México, Grupo Editorial Patria, p5

Hurtado, D, Principios de administración, Medellín, Colombia, Instituto Tecnológico Metropolitano, 2008, p.41

Davenport, T. ; Prusak, L. (2001). Conocimiento en acción: como las organizaciones mejoran lo que saben. Buenos Aires: Pearson Editores. P. 77

Fontalvo, T., Quejada, R. y Puello, P., La gestión del conocimiento y los procesos de mejoramiento, Dimensión empresarial, Vol. 9(1), 2011, p. 81.

Rosado, A., Rico, D, Scientia et Technica Año XVI, No 44, Abril de 2010. Universidad Tecnológica de Pereira, p.321

Harrington, P., 2012, Machine Learning in Action, Manning Publications Co., p.p. 23, 25, 30.