REVIEW

Application of predictive models for territorial planning in Pasto, Colombia

Aplicación de modelos predictivos para la planificación territorial en Pasto, Colombia

Jheison Edilson Arteaga Quistial[1] ✉, Miguel Ángel Velásquez Bravo[1] ✉, Omar Alexander Revelo Zambrano[1]

[1]Universidad CESMAG, Facultad de Ingeniería, Ingeniería de sistemas. Pasto – Nariño, Colombia.

Corresponding Author: Jheison Edilson Arteaga Quistial ✉

ABSTRACT

The research analysed urban growth in the city of Pasto between 2000 and 2022. The need to properly manage large volumes of data using technological tools that would allow for the organisation, analysis and prediction of urban expansion behaviour was identified. To this end, a predictive model was implemented, supported by software developed with open-source technologies such as Python and Laravel, with an interface based on HTML and CSS. This software facilitated the identification of growth patterns through data mining and multi-criteria analysis techniques. The project received institutional support from the University of Nariño, as well as the participation of teachers and researchers from the fields of engineering and architecture. In addition, data was collected both digitally and through fieldwork in order to build a reliable repository to feed the information system. The study showed that urban growth is a multifactorial process influenced by social, economic, demographic and political variables. Finally, it was concluded that the predictive model made it possible to determine interactive scenarios of urban expansion and offered an effective tool for territorial planning. The results confirmed that it is possible to apply methodologies such as EMC and information systems to support strategic decisions in urban planning.

Keywords: Data Mining; Expansion; Prediction; Urban Planning; Planning.

RESUMEN

La investigación analizó el crecimiento urbano en la ciudad de Pasto entre los años 2000 y 2022. Se planteó la necesidad de gestionar adecuadamente grandes volúmenes de datos mediante herramientas tecnológicas que permitieran organizar, analizar y predecir el comportamiento de la expansión urbana. Para ello, se implementó un modelo predictivo respaldado por un producto software desarrollado con tecnologías libres, como Python y Laravel, con una interfaz basada en HTML y CSS. Este software facilitó la identificación de patrones de crecimiento mediante técnicas de minería de datos y análisis multicriterio. El proyecto contó con el apoyo institucional de la Universidad de Nariño, así como con la participación de docentes e investigadores de las áreas de Ingeniería y Arquitectura. Además, se recopilaron datos tanto digitales como mediante trabajo de campo, con el fin de construir un repositorio confiable que alimentara el sistema de información. El estudio evidenció que el crecimiento urbano es un proceso multifactorial influenciado por variables sociales, económicas, demográficas y políticas. Finalmente, se concluyó que el modelo predictivo permitió determinar escenarios interactivos de expansión urbana y ofreció una herramienta eficaz para la planificación territorial. Los resultados confirmaron que es posible aplicar metodologías como la EMC y sistemas de información para apoyar decisiones estratégicas en el ordenamiento urbano.

Palabras clave: Minería de Datos; Expansión; Predicción; Urbanismo; Planificación.

INTRODUCTION

Nowadays, information management has become an important element for decision making in different sectors of cities. The generation of large volumes of data is extremely relevant in the different studies that can be developed in many disciplines, always emphasizing that such information must be reliable and of value, therefore; information management is not only denoted in the organization and access to data, but also in the process of analyzing them in a safe, optimal and fast way.[1,2,3,4,5,6]

It is necessary to understand that large volumes of information are often difficult to manage in a safe, optimal and fast way since many times such information is not truthful, valid or consistent, therefore there is the need to apply techniques related to the fields of machine learning that allow organizing and classifying the information in a refined way in order to give a prediction of future results about the urban growth of the city of Pasto, thus allowing to visualize final results through analysis. Considering the analysis in the research process, data analytics and the importance of collecting information in a truthful and compact way are a great reference to have a good development and understanding of the situation between the periods 2000 and 2022. [7,8,9,10,11]

Information consists of a set of data related to an event, fact or phenomenon, which, when organized in a specific context, acquire a meaning. Its purpose is to reduce uncertainty or enrich the understanding of a subject.[12] In this sense, the aim is to develop a repository of urban data on the city of Pasto that facilitates the understanding of its growth in relation to urban development. This repository would provide a clear and structured vision of the urban growth of Pasto, thus allowing a deeper analysis of its evolution over time.

Certainly, information technology (IT) refers to the totality of aspects concerning the transformation, conservation, safeguarding, processing and transmission of data. It is used to encompass any technology that enables the management and communication of information;[12] therefore, information processes can be developed supported by computerized tools that provide real and feasible solutions to the changes and dynamics that today's city represents.

The analytics and implementation of a good repository of information helps to make a software product which allows to give a clear picture of what has been achieved in the process of research, collection and filtering of information, additionally it will allow to predict which areas are areas of potential space this will be possible through the implementation of neural networks, therefore we have "the research group of the Eafit University Research in Spatial Economics (RISE) developed two products: Urban Pixel software, an algorithm that forecasts the growth of cities and their population; and the Newton module, which takes growth data and uses it to calculate water demand in different geographical areas. "[13] It can be deduced that growth through technological tools is already a reality and therefore also a necessity, "We consider that these urban growth models are an important tool for public authorities to have a better idea of where their cities are going and to have more time to plan and design the necessary actions to anticipate that growth," said Juan Carlos Duque, professor in the Department of Mathematical Sciences at EAFIT and researcher at Rise[14], consequently, interactive scenarios can be determined from urban growth in the city of Pasto.

Feasibility
Operational

The research receives support in the field of human resources management from professors of the Systems Engineering program and the Architecture program.

Jointly, it has the institutional support and the collaboration of the research groups Alarife and Tecnofilia. This research originated from the project entitled "Analysis of urban growth in the expansion zones of Pasto 2000-2022: implications of the compact territorial model proposed by the land management plans in urban planning, by means of a predictive model"[15], which won first place in the call for research seedlings in 2022.

Technique

Data mining techniques and knowledge about Python were used to create a predictive model in order to extract, store relevant information and make future predictions. For the development of the software product, the Laravel framework was used for the management and structuring of the project. The user interface was designed using HTML, CSS and Bootstrap, which allowed not only to improve the visual appearance, but also to ensure a responsive and accessible user experience. The server deployment environment.

Economic

The collection of information for the research will be done for free; this is possible since certain data can be obtained from Google Maps, Google Earth or various pages that provide information necessary for the research process. Field work was carried out in certain areas where no information was available on the Internet, but the costs of this are minimal. In the development of the software product, free applications will be used, which have no cost, such as Python, CSS and HTML. Once the product itself has been realized, the cost that would be

incurred would be the cost of the web domain. It is worth mentioning that all the necessary resources will be financed by the researchers (Jheison Arteaga, Miguel Velázquez), budgeting in the same way the software and hardware tools that will be developed and implemented.

Delimitation

This research will be carried out in the city of Pasto, in the Colombian territory, focusing on the period from 2000 to 2022. In order to collect and analyze data strategically to ensure its veracity and validity. Focusing on different residential areas (buildings, complexes, condominiums, housing), using digital mapping tools and online sources to obtain detailed information. Similarly, surveys will be conducted to people with knowledge in the creation of housing, local residents and municipal officials in the specified period in order to compare the information and to be able to fill the matrix that is being made for the development of the project.

DEVELOPMENT

Background

International

In the research process about urban growth, it can be highlighted that it is not only a fact that develops in Colombia as such but in most of the countries of the world, for this it is necessary to understand the importance of knowing relevant and useful information about urban growth, in this sense it is possible to highlight some of the reasons that can solve the questions Why? and What for? To know about this topic is of great importance, because it can be noted that there is a reason delimited or synthesized in the local contextualization based on understanding that the implementation of urban growth strategies and policies of each country are different and not the same for all countries, this is crucial to understand the local context of each country in order to develop effective and appropriate approaches. A second reason would be the innovation and adaptation promoted by studying the urban growth of different countries in order to identify innovations and implement dynamic and creative solutions supported by technological tools capable of providing such solutions. Finally, a third reason can be based on knowing the different contexts present in various countries regarding urban growth, which is important in foreseeing errors, since it is possible to learn about the challenges that different countries have experienced with respect to their urban growth, thus providing information to prevent making such errors again in order to help save resources and minimize negative impacts.

Expansion

Firstly, there is a document researched by Lincoln Institute of Land Policy in 2016, which is a document adopted at the United Nations conference on Housing and sustainable urban development, exposed in Quito Ecuador in the year of 2016, said strategic document has been developing for many years through research processes which establishes a global vision for sustainable urban development worldwide providing guidelines to deal with the challenges and opportunities that cities live in the XXI century, detailing that if there is an adequate urban planning and a correct territorial management in a sustainable way, the challenges or challenges represented by urban growth as such could be addressed highlighting citizen participation, where different societies can be present in decision making and urban planning, understanding in this way that citizens have a fundamental role in the creation of sustainable cities.[16]

In the same way, a study conducted by Shlomo in 2012 explains a research process that tries to examine the phenomenon of urban growth in cities around the world addressing a total of 200 cities where it was possible to have a global vision of urban expansion, analyzing the urban patterns and trends that made possible the rapid urban growth of certain areas or regions, identifying the key factors that promoted such expansion, Additionally, the different urban planning policies and strategies used or implemented in different cities to manage expansion were analyzed, detailing the importance of proper planning together with the support of different sectors such as social, economic, political, among others; with which the urban challenges of urban growth can be addressed and dealt with.[17]

It is important to be clear that urban growth can have negative or positive impacts on an area as such, for this a series of studies have been conducted in many countries that allow identifying the above, an example of this is the study conducted by Andreas Haller in 2016 where he develops processes of physical, demographic and socio-cultural urbanization which have many impacts on the peri-urban area located between the city of Huancayo and the countryside, detonating a broad rural-urban relationship of the Andean intermediate cities and the potential for social inclusion and sustainable development of the peri-urban areas of Huancayo, where a series of data were obtained from the nearby areas of the city determining variables such as population, average annual growth and the projection in different years or periods of time obtaining in this way that north of the city center, on the left geographical side of the Shullcas River is developed a residential area for the emerging middle class and an important center for private educational institutions. The growing demand for lots has resulted in a price increase for land in the neighboring towns of Uñas and Vilcacoto, which in turn increases the sale of agricultural land to developers.[18]

EMC

In addition, in order to analyze the urban growth of an area, it is necessary to consider the following. For this reason, there must be a methodology that allows considering and evaluating these factors in order to make informed and/or equitable decisions that can generate a positive impact on the management of urban growth. The methodology discussed here is the EMC or multi-criteria evaluation, which is necessary and useful in the development of the project to evaluate multiple criteria that help identify the areas that will see urban growth.

Considering a methodology to evaluate urban growth factors, we have the study conducted by Noelia Principi in 2021 where the application of a Multi-Criteria Evaluation (MCE) methodology is presented, analyzing different spatial criteria and land use suitability levels for urban expansion in the city of Luján. Likewise, Saaty's pairwise comparison method, known as Analytical Hierarchy Process (AHP), is used to perform the weighting of each of the criteria.

The results obtained show different levels of suitability for urban land use, highlighting the northwest sector of the city as an area with potential for urban development. The EMC supported by the use of Geographic Information Systems (GIS) allows the analysis of the geographic space for planning support purposes, generating important possibilities to improve the spatial decision making process.[19]

Continuing with the EMC implemented in the evaluation of multiple criteria and decision making in different contexts or social sectors is the study of Federico Benjamin in 2013 where a method is shown, which, supported by scientific proposals, helps in the planning of the sustainable use of the territory. This method is based on the evaluation models of the carrying capacity of the territory. It is based on the development of an evaluation methodology with a broad perspective of rural sustainability and contemplates the necessary enhancement of natural resources, the preservation of the environment for its environmental quality and the minimization of natural risks. It will be used to evaluate the implementation of buildings related to the process of diffuse urbanization in rural areas. For this purpose, the particular conditions of the rural environment are evaluated in order to better estimate its possibilities and its vulnerability to natural risks. As a result, an evaluation based on clearly defined value judgments and specific attributes of the territory has been carried out. Through this process, an information layer is obtained that shows a classification of the studied space with a valuation assigned to each part of the territory according to its capacity to accommodate the uses that have been evaluated.[20]

Predictive models

In order to understand how urban sprawl develops in a given area, predictive models that can forecast and predict the future behavior of urban growth as such can be used and implemented.

In the implementation of predictive models in the research process of urban expansion, an example is the study in 2007, where he explains the use of predictive models to generate future and dynamic scenarios of land use changes and the expansion or growth that this could present in a context of territorial planning or the evaluation that can generate the environmental impact, Thus, it is said that the implementation of predictive models and the use of future land growth patterns are useful when implemented with the use of sustainability indicators that can substantially help the challenge of sustainable development of nations and more specifically to the environmental management of urban growth.[21]

The study conducted in 2014 explains that in the face of the impact that urban expansion exerts on a certain area and its population, the use of simulation techniques that help to develop dynamic scenarios on urban expansion is necessary or crucial, therefore it is explained that the development that cellular and dynamic models (MBA) can function as tools to make decisions and implement them on

the object of study to simulate urban growth in the Corredor del Henares which is a community of Madrid in Spain, with this is achieved detailing that cellular models are suitable for spatial analysis with an urban approach, especially in the simulation of future scenarios identifying the main factors of expansion or in defining growth patterns.[22]

Information system

For the research process of urban growth it is important to find the variables that can alter or modify its behavior, in the course of the project a large amount of information can be obtained that must be recorded, so the project requires an information system that allows this process and that has been fed with information that will be debugged in a single repository by identifying the most significant variables of the research.

In this section we have the research in 2015 where he establishes the information system as a set of interrelated functions or components that form a whole, where information obtained during analysis or information gathering processes is obtained, processed, stored and distributed, which is useful to support decision making and control in an organization, supporting coordination, problem analysis, visualization of complex aspects.[23]

It can be detailed that the information system must be able to store large volumes of information in an optimal and secure way, therefore, we have the research in 2017 where she highlights the need to maintain

efficient information systems that facilitate decision making and the identification of useful and valuable information flows for an entity taking into account the security of information according to current standards.[24]

It can be added the fact of the implementation of an information system in decision making by already having refined information, with this we have the study in 2018 where she highlights that the information system can be useful in business operations, management and decision making.[24]

business operations, management and decision making, providing people with the information they need through the use of information technologies.[25]

In the process of supplying the debugged information in the information system, it is held that such information must be updated, timely and of value, to this can be considered the study in 2017 where he explains that, in environments such as technology, the competition and among others it is important to have updated and timely information for decision making, in order to create strategies that ensure the proper functioning of a given entity by supporting data processing.[26]

It is important to highlight that the information system has useful information, related to the urban growth of the city of Pasto and in turn serves as a provider for people who need such information, in this part we have the research in 2014 where the impact that can have the usefulness of information systems is studied highlighting that such information is of quality so that the systems function as information providers and in turn help in the automation of tasks detailing in this way the relationship between automation and the usefulness of information systems.[27]

Based on the above mentioned, the information system is capable of automating the analysis of the urban growth of the city of Pasto, since every information system has the characteristic of automation that is able to manage large volumes of information in an optimal and safe way. In this section we have the study in 2004 where she explains that information system is an organized combination of large amount of data that has a significant impact on the growth or increase of productivity within the organization.[28]

Data mining

In the development of the project it is sought to implement data mining for the process of finding patterns and correlations in the data that are being collected in the research process in order to predict interactive or dynamic scenarios of urban growth in the city of Pasto.

A study conducted by in 2015 where he explains that data mining has become a tool used in the analysis of data and large volumes of information in order to identify and show patterns of behavior on the data and achieve the generation of hidden knowledge about the information that previously was difficult to share and produce due to the complexity of the information, an example of the above can be highlighted in business intelligence strategies and advanced prediction models. Additionally, there is the study in 2015 where she denotes that data mining involves a large selection process[29] and the exploration and modeling of large volumes of data are fundamental to discover patterns choosing the classification technique to generate such models presenting a tree with the different attributes giving the best rules of the interactions of the students of the national polytechnic university of Ecuador.[30]

National

Expansion

In a study by Bazant et al. it has been observed that cities expand or grow daily; areas or zones that were considered and are considered uninhabitable are expected to be inhabited, divided, lotified and subdivided in the future, therefore it is deduced that for the growth of cities there are no limits, there is no area, no matter how dangerous or essential for the environmental balance, that can stop such expansion.[31]

EMC

Continuing with the study of Bazant et al., it allows to explain methods of multi-criteria decision analysis (MCDA). Multicriteria decision analysis (MCDA) methods based on the useful multicriteria evaluation for the prioritization of technologies from a governmental perspective where a systematic search of decision criteria was carried out, which allowed the identification, classification and evaluation of their importance defining the components that frame the proposed model in future decision making[25] which for these studies help the health sector, but the multicriteria evaluation methodology of decision criteria based on the study by Calderón et al. can be used in the study.[25]

by Calderón et al. can be used for the development of the project when prioritizing the criteria and/or variables involved in the urban growth of the city of Pasto.[32] The study by Vargas et al. can also be considered, which highlights the importance of incorporating multiple criteria evaluation processes in the analysis of a problem using various dimensions: socio-cultural, economic, ecological-environmental, or others.[33]

Predictive models

A research conducted denotes land cover and land use changes where predictive models are established as an efficient tool in the analysis that facilitates decision making to anticipate, prevent and mitigate the effects generated by such changes.[34] There is also the study where it is stated that urban expansion produces an unprecedented transformation in land use, therefore, by not recognizing its dynamics and incidence in sustainable planning, a predictive model is needed to help make future decisions that allow for territorial planning to be sustainable when urban growth occurs.[35]

Information system

Considering the study the existence of Geographic Information Systems (GIS), capable of managing large volumes of spatial data derived from different sources, is reported.[36] In addition, another study reports that the application of technology in the registration, management and processing of data is always necessary in the automation of research processes.[37] Finally, the research of Peña et al. tells us that the information held in the repositories must be filtered before being supplied to the information systems so that such information is truthful.[38]

Data mining

Finally, there is the study which shows that data mining is useful in the process of pattern recognition in large amounts of data, having in the first place the "Understanding of the business", where the information supplied is analyzed; in second place there is the "Understanding of the data", where a descriptive statistical analysis of the data is made, and in third place there is the "Understanding of the data", where a descriptive statistical analysis of the data is made, and in third place there is the "Understanding of the business", where the information supplied is analyzed. Descriptive statistical analysis of the data; thirdly, the "Data preparation", where the variables are profiled, duplicate records are eliminated, null values are managed and outliers are eliminated; fourth is "Modeling", where 3 analytical models are designed and applied for the study of the data corresponding to a clustering analysis, a selection of factors and a prediction; fifth is "Evaluation" where the degree of reliability and certainty of the models is measured; finally there is "Implementation" where the results obtained with the analytical models are concluded, validating the results obtained as support for decision making.[39]

Regional

Expansion

A study conducted by Ceballos et al. in the region of Nariño, has identified a socio-spatial evolution of the urban area of Sotomayor understanding the environment where the population has developed identifying that the territories have different characteristics and therefore do not manifest similar dynamics of urban growth.[40]

EMC

In the region of Nariño there have been multiple investigations where they exemplify the zoning by susceptibility to flooding phenomena in the area of urban influence and urban growth, where the Multicriteria Evaluation technique was implemented to visualize the areas involved, selecting the necessary criteria and assigning them a weight, based on identifying which has the greatest weight in percentage to the slope variable.[41] Another study carried out in the region of Nariño, explains that the evaluation and hierarchization of variables are important and necessary for future decision making where the methods, criteria and data sets that should be taken into account in order to solve the problem posed in the selection and location of possible potential coffee distribution zones are described.[42]

Information Systems

An investigation carried out in Nariño shows that offering more confidence to develop, analyze and apply information in decision making is useful as long as it is used and quality information is available.[43] A following research is that of Pineda et al. where he details the importance of confidentiality of the information provided to information systems, where this information must be kept secure so that it is visible to users who require it.[44]

Data mining

Finally, there is the study where data mining is implemented by consulting all published and available sources of information that contain relevant information about the object of study to be investigated.[45] The last study explains the extraction of patterns from the information obtained in collection phases or found in different repositories using data mining techniques to be subsequently recorded in an information system.[46]

Theoretical research assumptions

For the development of the project, the formulation of solid theoretical assumptions has been highlighted to guide the research process of urban growth in the city of Pasto, for this reason these assumptions have been grouped according to the context where they are based, having the contexts of the object of study which is urban growth, the data obtained during the research, the predictive model that will be implemented and the software that will support or support the management of information; once the contexts are identified, the assumptions as such have been identified; having the following:

Urban growth assumptions

Assuming that urban growth is a diverse and complex process influenced by a series of key assumptions, such as infrastructure, economy, population, government policy, it can be assumed that it is possible to devise different future scenarios of urban growth that would be affected by changes in the key assumptions mentioned above.

Infrastructure affects the urban growth or expansion of cities due to future planning by public policies, thus denoting that morphological changes in the structure of cities induce economic growth, improving the labor market and the welfare of societies by focusing on infrastructure management in territorial planning considering the physical facilities and different systems that help or benefit the functioning of a city; such as bridges, roads or health care and education services, generating significant impacts on such societies. In this way focusing on transportation infrastructure, achieving a correct planning of such infrastructure results favorable for people mobilization within the city essential for the connection of workplaces, housing or services allowing the growth of urban areas.[2]

The economic processes that develop in societies such as economic globalization generate profound changes in the use of the territory by different social groups, thus experiencing that economic, technological advances, and productive systems produce significant transformations triggering new forms of metropolitan expansion, achieving a greater diversification of the activities that can be developed in a given city, thus including expansion in sectors such as education, technology, or territory that can lead to a more balanced and diverse growth that is bidirectional because it influences the impacts between economy and urban growth or vice versa.[7]

Focusing on urban planning, which is directly involved with the urban expansion of societies, it is important to point out that effective and active citizen participation is required, capable of promoting and analyzing the participation of the population in social organizations, thus understanding that the population is not simply part of society and is the object of study of urban growth, but also the population can be directly involved with issues or social aspects that generate impacts on the expansionism of cities or societies, thus jointly understanding that the density and size of the population radically influence the urban planning of cities so that strategies must be developed to help with the management of these societies, for this reason; It is common for the development of high-rise buildings or residences to be used or adapted to areas of lower territorial density, causing urban development towards larger areas, thus understanding that the relationship between population and urban expansion is intrinsic, which means that the changes that can occur in a population can have large-scale impacts on the way a city expands or grows.

Public or governmental policies are considered as a plan to achieve an objective of public interest that benefits a society, implementing techniques and tools of the traditional forms of state planning, constituting a solid and efficient public policy that incites sustainable urban development or expansion by identifying its basic components, and in this way achieve the approach of general proposals that relate to the territory, urban development and housing, therefore; governmental decisions and regulations play a fundamental role in the development of urban growth because they can influence many aspects of such development, from land use planning to resource management, understanding that governmental decisions can have lasting impacts on the way cities grow and evolve over time.[20]

Assumptions of the data collected

It is assumed that there is a set of reliable and complete historical data related to the urban growth of the city of Pasto, including relevant and quality information that allows identifying the areas of Pasto that may experience urban expansion in the future through the use of such data, thus achieving an adequate and sustained identification of the problems, opportunities and challenges that the city could experience with respect to urban growth, evaluating possible solutions to the problems and challenges mentioned above. Additionally, it is assumed that the application of data analytics on the data collected will be useful to analyze or study such data effectively and obtain valuable information regarding the urban growth of the city of Pasto allowing cities to optimize the management processes and allocation of resources by identifying and determining the areas where investment could have or generate significant impacts, in order to ensure that the resources available in a given city are used efficiently and appropriately, using them in multiple processes involving different sectors of societies as such.[35]

Every research process requires relevant data or information that helps or develops in an adequate way the object of study, which for this project is urban growth, having said this, it is important to consider the process of collecting detailed, complete, reliable data and also the constant improvement of the quality of the information, through direct or indirect evaluations with the purpose of correcting basic data that allow detecting changes in different sectors such as health, economy, housing, or territory, economy, housing, or territory, thus managing to study and understand the causes and consequences produced by urban growth or expansion in a given society by applying historical data analysis techniques, thus demonstrating that the data collected provide valuable information that is useful for the planning and management of urban areas, thus influencing decision making and policies related to urban development.

It is important to highlight that the collected data when not taken to analysis and management processes become irrelevant and worthless information due to the lack of certainty about the reliability of such data, which leads to the application of data analytics involved with an exhaustive inquiry process, demonstrating the challenges that mean the access and manipulation of data by implementing different scientific methods involved with such process, important for making future decisions that involve in any sector of society, understanding that the implementation of data analytics in the process of studying urban growth is favorable due to the fact that it can help in identifying the best way to use the available land of cities by using algorithms and optimization of these that can find and work factors that can exemplify the most efficient allocation of land use of a certain area as such.[8]

Assumptions of the predictive model

It is assumed that it is possible to develop a predictive model capable of identifying patterns in the collected data that are based on urban growth or expansion, together with the assumption that such a model can be scaled and validated with different historical data, assuming that such a model will have a sufficient degree of accuracy to provide useful and truthful information for decision making regarding urban expansion.

Once the information or data collected is available, it is important or necessary to recognize and analyze patterns in such data, examining the new areas of urban expansion that could emerge in a given area, it should be noted that to find such patterns it is essential to consider different factors that allow understanding such expansionism, it is worth mentioning that the process of urban diffusion can be evaluated by analyzing patterns of change with respect to land use, space or areas of expansion and the time that delimits the study about urban expansion by implementing analysis techniques that are useful for making predictions of future events involved with the behavior based on historical data and patterns identified in such data.

The implementation of a valid and efficient model in different analysis processes is convenient for decision making, however, it is required that the model at the time of its implementation has been previously evaluated to understand its level of accuracy, therefore it must have enough training data that are essential for the model to learn and acquire the ability to make future predictions or make decisions based or based on new data and thus as a result generate data for model validation using machine learning algorithms; It is also important to consider the time that the learning process can take for the model, therefore the implementation of a model in the research process can provide information of great relevance for the long-term strategic planning of a given city including different sectors.

Software product assumptions

It is assumed that users will interact with the results of the predictive model in a visual, dynamic and optimal way, which leads to the creation of a user-friendly and easy to manipulate navigation interface. Additionally, it is assumed that the software requirements will be precise and complete, which will allow to break down in an optimal way the behavior that the system should have together with the quality characteristics with which the aforementioned software to be developed should have. It could also be assumed the fact in the prioritization of the requirements of the software product where it is assumed that said requirements can be categorized and hierarchized in such a way that the priority requirements are based or focused on the precise understanding of the needs or priorities of the users that will interact with the software, and that the changes in the prioritization of the identified requirements would be made in a reasonable and justified way.[11]

The development of an efficient product that is capable of displaying the data analyzed post-implementation of the predictive model is necessary for users, because in this way they can interact properly with the relevant information about the urban growth of the city, for this in the development process could apply the type of applied-technological research where you can carry out multiple software processes such as descriptive studies, collection of information on the topics that should contain the software, therefore could be broken down in the development of software, phases such as planning; where the needs to be solved through the implementation of the software product are defined, the next phase is the analysis, where the characteristics of the users that will interact with the application are determined, the next phase is the design, necessary to determine the topics and contents of the software, adapting the interfaces so that they are intuitive and easy to understand

at the time of navigation by the users, another phase supported by the implementation, where an initial version of the software is developed, and the last phase is the testing and launching of the final product.[16]

Focusing on the context of the identification of the software requirements, it is important to emphasize that for the product to be optimal and useful in the research process, it is necessary to have a correct planning process for each and every one of the requirements of the final product, for this reason; it is necessary to establish in a precise, clear, detailed and organized way the set of requirements that must be satisfied by the software to be developed in order to determine in a total and conscious way the requirements of the application, then the analysis of the resulting outputs, the decomposition or breakdown of the data and the processing of these, in order to obtain a software product with quality standards that is efficient and meets the needs of users who will interact with this.[3]

For the approach of the requirements of the software product it is relevant to consider the final functionalities with which it should have, in this way; highlighting the approach of the requirements it is possible to denote the final functionalities that allow to show how the system should behave; therefore the prioritization of the functional and non-functional requirements could arise applying the MoSCoW technique useful in the project to prioritize, and categorize the requirements of the final product according to its level of importance within the development of this, where 4 degrees of prioritization can be established for the requirements, which are:

- Must (must): priority grade 1 where the requirements for this prioritization are mandatory to be implemented in the final product to be developed.
- Should: priority level 2 where the requirements for this prioritization are necessary to be implemented in the final product to be developed.
- Could: priority grade 3 where the requirements for this prioritization are important to be implemented in the final product to be developed.
- Won't (No): grade 4 of priority where the requirements for this prioritization are relegated from being implemented in the final product to be developed but could be worked on in the future for the development of the product as such.

The previous degrees of priority represent the importance of each requirement with respect to its functionality within the application, thus highlighting which are the requirements that should be worked primarily and those that can be worked in a complementary way, MoSCoW is used in multiple areas or multidisciplinary, such as in the areas of Business Analysis and Software Development in order to achieve a correct and efficient common understanding of the stakeholders with respect to the requirements.

common understanding of the interested parties regarding the importance of each requirement that will be delivered as part of the final product, which achieves effectiveness in the prioritization of the requirements that the application will have, thus achieving an optimal and efficient product.

Study variables

Dependent Variable: in this research, the predictive model is considered the dependent variable, since its performance and accuracy depend on a series of factors that affect it. This model is designed to forecast urban growth in the city of Pasto, adjusting itself according to the influences of several independent variables that determine its effectiveness in projecting results.[4]

Independent Variables: the independent variables that influence the predictive model are called urban variables and include: Project Name, Housing Type, Risk or Hazard Zone, Principal Ecological Structure, Housing Quantity, Housing Type (Morphology), Housing Type (Subsidy), Housing Strata, Housing Size (m^2), Housing Price, Start Year, Population Density (hab/Hec), Amenities, and Residence Location (Longitude and Latitude). These variables are fundamental to understand the factors that impact urban growth and to optimize the model's predictions.

For the predictive model, the variables would be related to the parameters and characteristics used to make future predictions, such as expansion rates, regression variables, as well as for the model's accuracies, variables such as mean square errors, error rates, could be identified to see if the model is efficient or not efficient.[18]

With the software as such, it is also important to consider or take into account some variables related to the user experience, such as ease of use, software adoption, to this part can be added the evaluation made from the quality characteristics of a software product through ISO 25010.

Nominal definition of variables

It should be considered that nominal variables are used to label or categorize elements in groups without any existing order or hierarchy where no mathematical operations can be performed with these variables. With the above we have that the nominal variables of the project are:[21]

- Name of the Project.

- Type of Housing.
- Risk or Hazard Zone.
- Main Ecological Structure.
- Type of housing (morphology).
- Type of housing (subsidy).
- Experience: rated in the categories of good, average, bad experience by the user when using the developed software.
- Ease: rated in the categories easy, intermediate, difficult with respect to ease when a user uses the software product.
- Adoption: rated according to the categories easy to understand, medium, difficult to understand the functionality of the software.

Operational definition of variables

It is important to consider that the operational variables represent properties or characteristics of interest in a study and that they can be measured or quantified numerically, therefore mathematical operations can be performed, with the above, the nominal variables of the project are:[33]

- Housing stratum.
- Size of dwelling.
- Price of the house.
- Year of start.
- Population Density (inhab/Hec).
- Facilities (m).
- Length.
- Latitude.
- Expansion rates: important to generate the change in the predictions made by the model.
- Regression variables: important for detailing the accuracy of the model.
- Mean square errors: needed in the model to measure the amount of error between the data sets to be worked with.
- Error rates: denoting an error rate by the implemented model.

Formulation of hypotheses
Research hypothesis
Hypothesis Statement

"It is possible to determine the interactive scenarios of urban growth in the city of Pasto by means of a predictive model, supported by a software product."

Null hypothesis
Hypothesis Statement

"It is not possible to determine the interactive scenarios of urban growth in the city of Pasto by means of a predictive model, supported by a software product."

Alternate hypothesis
Hypothesis Statement.

"The research provides tools for the understanding and planning of land use planning in relation to urban growth in the city of Pasto through a predictive model, supported by a software product."

CONCLUSIONS

From the exhaustive research process presented, it is concluded that urban growth in the city of Pasto between 2000 and 2022 has been a complex phenomenon, influenced by multiple social, economic, territorial and technological factors. The lack of efficient management tools and the outdatedness of available data have hindered adequate territorial planning, which highlights the need to implement predictive models and reliable information systems to support urban planning decision making.

The development of a software product, based on data mining techniques, advanced analytics and methodologies such as Multicriteria Evaluation (MCA), has proven to be a viable alternative to address the challenges posed by urban sprawl. This approach allows not only to visualize past growth patterns, but also to predict future scenarios, thus facilitating strategic and sustainable territorial planning.

Furthermore, it is evident that the integration of relevant urban variables, combined with a predictive model supported by free and accessible technologies, represents an operational, technically and economically feasible solution. This model, applied specifically to the local context of Pasto, becomes a highly valuable

tool for public entities, urban planners and decision makers, by offering updated, accurate and structured information.

Consequently, the research confirms the main hypothesis: it is possible to determine interactive urban growth scenarios by means of a predictive model supported by a software product, which contributes significantly to land use planning and to the understanding of urbanization processes in the city of Pasto. This work lays the foundations for future research and development on intelligent urban territorial management, both at the regional and national levels.

BIBLIOGRAPHIC REFERENCES

1. Apablaza M, Henriquez C. URBAN PLANNING AND GROWTH: dislocations and URBAN SUSTAINABILITY CHALLENGES OF REGIONAL SANTIAGO, METROPOLITAN. Rev electron geogr cienc soc. 2010;14.

2. Ulloa-Espíndola R, Pérez-Albert Y. Validación de un modelo de predicción del crecimiento urbano en Quito (Ecuador). EURE (Santiago). 2022;48(144):1–27. https://www.scielo.cl/scielo.php?pid=S0250-71612022000200006&script=sci_arttext

3. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022. Convocatoria Universidad CESMAG; 2022.

4. Antonio L. Un modelo de crecimiento urbano vertical con factores característicos basado en inteligencia artificial. Uaemex.mx; 2021. http://ri.uaemex.mx/handle/20.500.11799/112219

5. Santos D. 20 ejemplos de páginas interactivas excepcionales. Hubspot.es; 2022. Disponible en: https://blog.hubspot.es/website/ejemplos-paginas-interactivas

6. Ley 388 de 1997. Disponible en: https://www.minambiente.gov.co/wp-content/uploads/2021/08/ley-388-1997.pdf

7. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022. Convocatoria Universidad CESMAG; 2022.

8. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022b. Convocatoria Universidad CESMAG; 2022.

9. Cañas EA, Díaz IC. Estrategias para la gestión de grandes volúmenes de datos por medio de Big Data. MVM Ingeniería de Software.

10. El Tiempo. Ahora se puede predecir el crecimiento urbano de las ciudades. 2021. Disponible en: https://www.eltiempo.com/colombia/medellin/innovacion-paisa-permite-predecir-el-crecimiento-urbano-de-las-ciudades-593757

11. Universidad EAFIT. Investigadores de EAFIT pronostican el crecimiento urbano y la demanda de agua en las ciudades. Disponible en: https://www.eafit.edu.co

12. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022. Convocatoria Universidad CESMAG; 2022.

13. Blei AM, Angel S, Civco DL. Urban Expansion in a Global Sample of Cities, 1990–2014. Working Paper WP18AB2. Lincoln Institute; 2018.

14. Haller A. Los impactos del crecimiento urbano en los campesinos andinos. Espacio y Desarrollo. 2023.

15. Principi N. Evaluación multicriterio aplicada a la expansión urbana en Luján. UNLu. 2021.

16. Bravo Cobeña CM, Valdivieso Guerra PEA, Arregui Pozo R. Propuesta de modelos predictivos en la planificación territorial. Rev Electr Geogr Cienc Soc. 2007;XI(245).

17. Gómez Delgado CM. Diseño de un modelo basado en agentes para simular el crecimiento urbano. Boletín A G E. 2016;(70).

18. Abrego Almazán D, Medina Quintero JM, Sánchez Limón ML. Los sistemas de información en el desempeño organizacional. Investig Adm. 2015;(115):7-23.

19. Castillo G. Los sistemas de información gestionando grandes volúmenes de datos. Scielo.org.co. 2023.

20. Moreno-Cevallos JR, Dueñas-Holguín BL. Sistemas de información empresarial: la información como recurso estratégico. Dominio Cienc. 2018.

21. Orellana B, Ormeño S, Rodón Módol J. Impacto de la automatización sobre el desempeño. Esan.edu.pe. 2014.

22. Prieto A, Martínez M. Sistemas de información en las organizaciones. Rev Cienc Soc (Ve). 2004;X(2):322-37.

23. Jaramillo Valbuena S, Cardona SA, Fernández A. Minería de datos sobre streams de redes sociales. Inst Investig Bibliotecológicas. 2015;(33):63-74.

24. Jaramillo A. Aplicación de técnicas de minería de datos para determinar interacciones estudiantiles. Espol.edu.ec. 2023.

25. Bazant J. Procesos de expansión y consolidación urbana de bajos ingresos. Bitácora Urbano Territorial. 2008;13(2):117-32.

26. Lucia O. La evaluación multicriterio social y su aporte a la conservación de bosques. Rev Fac Nac Agron Medellín. 2023;58(1):2665-83.

27. Marcela L, Stevens G. Aplicación de un modelo predictivo para el cambio de cobertura urbana. Udistrital. edu.co. 2022.

28. Enrique J. Acerca de los sistemas de información geográfica. Univ del Valle. 2013.

29. Manuel J, Gitierrez P. Desarrollo de un sistema de información para gestión de proyectos. UCatólica Colombia.

30. Lucía M, Elvirita L. Calidad de la gestión del sistema de información de una IPS. Cuader Lat Admin. 2016;9(17):58-71.

31. Oviedo Carrascal AI, Jiménez Giraldo J. Minería de datos educativos: análisis del desempeño estudiantil. Rev Politécnica. 2019;15(29):128-40.

32. Michele K, Yurani T. Evolución socio-espacial del área urbana de Sotomayor. Udenar.edu.co. 2021.

33. Benitez K, Gomez D. Zonificación por susceptibilidad a fenómenos de inundación. Udenar.edu.co. 2015.

34. Mauricio D, Jaramillo Molina C, Alberto C. Evaluación espacial de centros de distribución de café. Rev EIA. 2022;19(38):10.

35. Lucía M, Elvirita L. Calidad de la gestión del sistema de información de una IPS. Cuader Lat Admin. 2016;9(17):58-71.

36. Carolina L, Sánchez JJ. Mapa de fallas de los volcanes Chiles-Cerro Negro. Bol Geol. 2017;39(3):71-86.

37. Pereira T, Clara. Minería de datos en supervivencia de cáncer cervical. Univ Salud. 2023;14(2):117-29.

38. López GA. ¿Las infraestructuras como proyecto de ciudad? Ciudades (Valladolid). 2023;(11):105-32.

39. Pérez Bustamante L, Salinas Varela E. Crecimiento urbano y globalización en Concepción, Chile. 2007.

40. Herrmann MG, Van Klaveren A. Disminución de participación ciudadana y planificación urbana. EURE. 2016.

41. Vista de Hacia una política pública de desarrollo urbano sostenible en el Perú. Paideia.: http://45.231.72.143

42. De Población En et al. Repositorio UC. Disponible en: https://repositorio.uc.cl

43. Rodríguez P, Palomino N, Mondaca J. El uso de datos masivos para políticas públicas. BID. https://publications.iadb.org

44. Modelando el crecimiento de ciudades medias. Google Books. 2014.

45. Carlos J. Análisis de modelos predictivos basados en visión computacional. Unam.edu.pe. 2019.

46. Fernando L, Carlos J. Desarrollo de software educativo sobre Ubuntu 13.10. Unach.edu.ec. 2016.

FINANCING

CONFLICT OF INTEREST
Authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Conceptualization: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.
Data curation: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.
Formal analysis: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.
Drafting - original draft: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.
Writing - proofreading and editing: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.