*Article*

# Using Machine Learning Tools to Classify Sustainability Levels in the Development of Urban Ecosystems

**Nidia Isabel Molina-Gómez** [1,2,*] , **Karen Rodríguez-Rojas** [1] , **Dayam Calderón-Rivera** [1] , **José Luis Díaz-Arévalo** [3] and **P. Amparo López-Jiménez** [2]

1   Department of Environmental Engineering, Universidad Santo Tomás, 110231 Bogota, Colombia; karen.rodriguezr@usantotomas.edu.co (K.R.-R.); dayamcalderon@usantotomas.edu.co (D.C.-R.)
2   Hydraulic and Environmental Engineering Department, Universitat Politècnica de València, 46022 Valencia, Spain; palopez@gmmf.upv.es
3   Department of Civil and Agricultural Engineering, Universidad Nacional de Colombia, 111321 Bogota, Colombia; jodiazar@unal.edu.co
*   Correspondence: nidiamolina@usantotomas.edu.co; Tel.: +57-300-552-6779

check for updates

**Abstract:** Different studies have been carried out to evaluate the progress made by countries and cities towards achieving sustainability to compare its evolution. However, the micro-territorial level, which encompasses a community perspective, has not been examined through a comprehensive forecasting method of sustainability categories with machine learning tools. This study aims to establish a method to forecast the sustainability levels of an urban ecosystem through supervised modeling. To this end, it was necessary to establish a set of indicators that characterize the dimensions of sustainable development, consistent with the Sustainable Development Goals. Using the data normalization technique to process the information and combining it in different dimensions made it possible to identify the sustainability level of the urban zone for each year from 2009 to 2017. The resulting information was the basis for the supervised classification. It was found that the sustainability level in the micro-territory has been improving from a low level in 2009, which increased to a medium level in the subsequent years. Forecasts of the sustainability levels of the zone were possible by using decision trees, neural networks, and support vector machines, in which 70% of the data were used to train the machine learning tools, with the remaining 30% used for validation. According to the performance metrics, decision trees outperformed the other two tools.

**Keywords:** urban sustainability; indicators; supervised classification; micro-territories

## 1. Introduction

For decades, sustainable development has been a significant challenge for nations, which is supported by, among other aspects, the environmental and socio-economic impacts associated with registered population growth. In 2018, 55% of the world's population lived in urban areas, which is expected to increase to 68% by 2050 [1]. The primary objective in addressing this challenge is to provide an orientation for a sustained improvement in the population's living conditions, which faces poverty, disease (associated with environmental and social determinants), and violence, among other situations. In this regard, the development and implementation of the Millennium Development Goals (MDGs) and the subsequent Sustainable Development Goals (SDGs) play an important role in determining the progress made towards achieving sustainable development.

The concept has been analyzed in different studies from different approaches [2–4], based on a broad spectrum of interpretations, primarily founded on the notion established in the report Our

Common Future, which states "development that meets the needs of the present while considering the needs of future generations" [5] (p. 16). Notwithstanding the global nature of the term [2,3], studies primarily focus on analyzing three fundamental pillars, environmental, social, and economic dimensions. Each dimension has its own specific challenges with respect to territorial conditions, in addition to being connected and integrated with one another, in order to make sustainable development achievable.

Evaluating sustainability establishes a degree of development for urban ecosystems in which natural and artificial structures interact and coexist. Ecosystem services, provided by natural systems, contribute to urban ecosystems' sustainability through the provision of goods and services. However, environmental conditions are altered (air emissions, waste, wastewater, among others) as the result of man-made structures and urban communities.

The pillars of sustainable development are looked at from a policy context, with a view towards an interaction between ecology and society, human ecology. The environmental dimension corresponds to natural resources and anthropogenic structures, while the biological community refers to the living components of ecosystems [6]. The social, economic, and institutional dimensions are part of the social system that is modified by technological infrastructure, knowledge, and social organization.

Social systems' influence on ecosystem services impacts the environmental dimension, not just at the resources level, but also in its biological community. In this manner, these interactions have been measured through indicators, whose objective is to establish conditions for the analyzed resource, in order to make decisions about its resilience.

Population growth and the ensuing pressure on natural systems through the use and exploitation of resources creates a need to understand these forms of pressure and possible measures that can be implemented to promote achieving the Sustainable Development Goals. Knowing the variation of sustainable development in a territory, based on its behavioral pattern, is an indispensable input for planning actions and measures. Natural ecosystems are basic to human life. As such, forecasting ecosystems' behavior, both natural and urban, can provide tools to protect human ecology.

Several studies have been developed to measure progress levels with respect to sustainability in countries and cities [2,4,7–11], in addition to other studies that have created inputs for forecasting nations' sustainability levels by using machine learning tools [12–15]. These studies have established procedures for the calculation, aggregation, and comparison of indicators in different settings, and have also proposed tools that can be useful in decision-making. However, these studies have been developed mainly from a global perspective, for a comparison between the behavior of countries and cities, leaving aside a more detailed territorial level approach, which is useful for the territorial synergy required to implement the SDGs. There is a need to integrate actors in an analyzed territory, in addition to structuring a comprehensive instrument to support the development of urban sustainability processes at the local level.

Machine learning tools have been used for decades in different settings to forecast the future behavior of input information. With the generation of large volumes of data, using these tools has become more useful in developing improvement strategies and analyzing sustainability from the smallest urban setting (organizations, households) to the territorial level. Therefore, it is essential to understand that achieving sustainable development is not only carried out through a national policy perspective but also in understanding the actions of territories that are part of cities and regions, urban micro-territories [16].

In this vein, this study seeks to establish a methodology for forecasting sustainability levels of an urban ecosystem through supervised modeling with machine learning tools. For the case study here described, the locality of Kennedy was selected, which is an urban territory in the city of Bogotá, the capital of Colombia. Kennedy has 1.2 million inhabitants with a rapidly growing population, 38% growth from 1993 to 2017. Additionally, 5.3% of this population lives in multidimensional poverty, among which, the health dimension (60%) is where most people are affected [17]. Kennedy is characterized by being one of the most polluted zones in Bogotá in terms of air quality, in addition to having high levels of insecurity. Several economic and service activities with contrasting environmental,

social, and economic behavior interact in this urban micro-territory. The analysis period for this study was 2009–2017.

Developing the aspects contained herein is innovative in that it applies machine learning tools to a territorial analysis approach. This study analyzed the dimensions of sustainable development in a more specific territorial scope that addresses aspects such as the difficulty in accessing information, a common characteristic in Latin America. This study is pioneering as it not only includes opinions from experts and community residents in the territory, but also an analysis of complaints and requests in the context of urban needs. The territorial scope established for sustainability analysis, in the field of human ecology, is a perspective that nations need to take into account in order to achieve better results related to sustainable development goals and targets. In general, there is a lack of machine learning models that forecast the sustainability behavior of urban territories, starting at the micro-territorial level, to support national and global perspectives for informed decision-making.

This study is structured as follows: Following this introduction, a description is given of the different steps undertaken for the supervised modeling of sustainability levels. These include collecting information by evaluating sustainability levels, applying machine learning tools, and an analysis of the same according to evaluation metrics. Afterwards, the results from applying this methodology in the case study are presented, in which the conditions of the micro-urban territory were identified, along with an indicator correlation within the framework of the sustainability dimensions. The territory's behavior over the years analyzed is presented through a categorization of sustainability levels, as well as the behavior of the machine learning models that were used. The study concludes with an analysis and discussion of the results, putting forth a suggested method to forecast sustainability levels in urban territories.

## 2. Materials and Methods

Several variables influence a territory's sustainability level, and their interaction affects its population's quality of life. Machine learning tools such as decision trees (DT), support vector machines (SVM), and artificial neural networks (ANN), were used in developing this study. A model to classify the sustainability levels of an urban area was created by applying these tools, which is useful for decision-making. This study consists of three relevant procedural paths: Characterization of the study area with indicators, definition of the classification labels for the supervised learning model based on the calculation from the sustainable development index (SDI), and the development of machine learning models. The above made it possible not only by creating a method, but also a model for the SDI classification of an urban area at the micro-territorial level. These stages were developed in a sequential manner, as described below (see Figure 1).
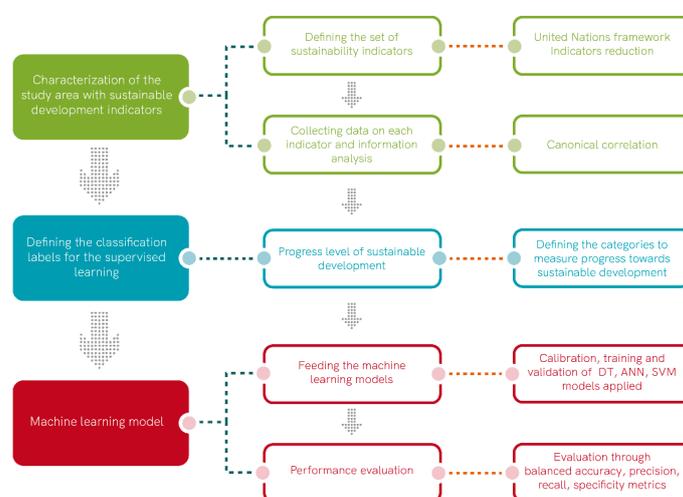


**Figure 1.** Methodological framework for the sustainability levels classified trough machine learning tools.

*2.1. Characterization of the Study Area with Sustainable Development Indicators*

The study area is the locality of Kennedy, an urban territory in Bogotá, the capital of Colombia. It is located at the coordinates 4°38'37'' N 74°09'12'' W (see Figure 2). This zone has a population density of 33,500 inhabitants/km$^2$, 36.7% greater than that of the city [17]. The locality is characterized by the presence of economic activities that include the provision of services, trade, and certain manufacturing activities. Additionally, 58.2% of the study area is residential, in addition to areas that are used for mixed purposes (services and trade). The zone has limited green space (6 m$^2$/inhabitant), with 3380 trees/km$^2$ [17]. Kennedy is made up of 12 zonal planning units in which different economic activities are developed along with housing areas.
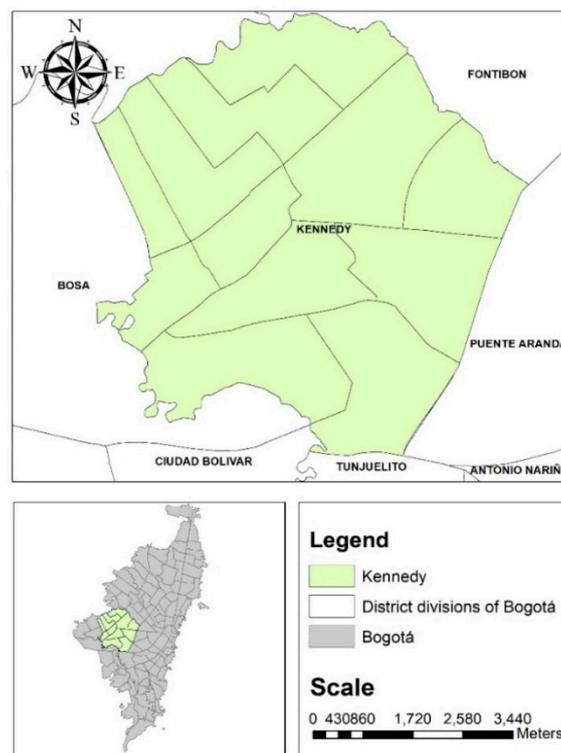


**Figure 2.** Locality of Kennedy as the study area in Bogota.

2.1.1. Defining the Set of Sustainability Indicators

To characterize this urban area, first, a set of environmental, social, economic, and institutional indicators was established based on the framework put forth by the United Nations [18,19]. This was followed by examining different studies that include analyses of sustainability dimension indicators [7,9,13,20–26]. The steps taken above made it possible to identify a set of indicators capable of rating the progress level, in terms of sustainable development, of the urban area from 2009 to 2017. In selecting the indicators, consideration was also given to whether these were part of the goals, targets and indicators of the Sustainable Development Goals/Millennium Development Goals (SDG/MDG) [18,27].

Subsequently, the indicators were reduced considering the following factors: (a) The opinions from community residents regarding different subjects of interest, by analyzing complaints filed with public sector entities, (b) the indicators' qualification characteristics, and (c) the importance of the indicators according to criteria established by technical experts and people with extensive knowledge of the territory.

The examination of the complaints filed by community members was carried out through systematic frequency analysis. With respect to the indicators' qualification characteristics, eight characteristics were

selected from the different studies analyzed [8,16,20–22,27–31]. These characteristics were: Access to information, analytical soundness, universality, policy relevant and usefulness to users, use of a multidimensional approach, measurable, unambiguous, and systematic. Each one was rated on a 1–10 scale, in which 10 is the highest value of the indicator characteristic. Characteristics with a total sum of less than 50 were discarded from the general base indicators.

In addition to analyzing the characteristics, a variation of the Delphi method was conducted [32], in which technical experts and people with extensive knowledge of the territory evaluated the established importance of the indicators. To this end, a web consultation was carried out using an electronic form addressed to technical experts in the specific areas of the indicators analyzed. Furthermore, two workshops were held with the experts in the study area to identify the importance of the indicators to residents in the territory.

The online expert consultation consisted of a series of closed-ended questions in which the participants stated their level of satisfaction with the eight characteristics for each indicator. A question was also included that inquired about the numerical importance of the indicator to achieve sustainable development in the study area.

With respect to the two workshops held with experts in the territory, a presentation was given on the project and how it was related to the SDGs, which was followed by an analysis of the indicators' importance in the territory. This evaluation was carried out through working groups and used a rating scale from 0 (low importance) to 5 (highly important).

It is important to note that considering the current participation spaces promoted by the local administration (i.e., local environment commission and economic observatory) 32 representatives from district entities, community leaders, and delegates from universities within the territory attended the workshops. This structured work developed so that different participants, with knowledge of the territory's priorities, could establish the importance of the indicators in the study area, making it possible to determine the set of indicators to evaluate the sustainability level of the urban zone.

2.1.2. Collecting Data on Each Indicator and Information Analysis

The next step consisted of collecting information on each indicator for the studied period 2009–2017. Written reports from twenty-seven district entities were consulted, as well as information from technical documents and annual reports on the study area created by these entities and other institutions [17,33–46].

A trend and behavioral analysis of the annual information for the period 2009–2017 was carried out for each indicator, which found incomplete information in some cases (8% of the total indicators used for the study period). Therefore, it was necessary to impute missing data in those cases in which specific annual information for the indicator was not available. The procedure followed in each case was to examine the indicator's behavior, and based on the same, the arithmetic average was taken by presenting an increasing or decreasing trend of the yearly information, or the moving average by presenting the variable's behavior from data with no apparent trend.

Once the set of annual indicators was established, a paired correlation analysis was performed via canonical correlation analysis. A comparison was made of the linear behavior between the variables representing the environmental, social, economic, and institutional dimensions. This procedure made it possible to determine the canonical variables and their correlation level.

It is important to note that the sustainable development of a territory implies the integration of the environmental, social, and economic pillars under the line of action defined through the institutional dimension. This study considered a characteristic parameter called "habitability", which is reflected in the relation between the environmental and social dimensions [7], in which indicators that describe the environmental health of a territory are relevant. An analysis was also carried out on the "viability" characteristic of the territory, which is based on the interaction between the environmental and economic dimensions [7], and also describes eco-efficiency indicators. Lastly, the analysis considered the importance of equitable development, based on the interaction between

the social and economic dimensions [7], described by the indicators' relation within the framework of social efficiency. The institutional pillar was analyzed from a global perspective that provides the basis to develop the individual pillars and their interactions.

## 2.2. Progress Level of Sustainable Development

In terms of planning, the term sustainable development has established a guideline from a global perspective, which aims to reduce inequities and improve conditions in the social, environmental, and economic dimensions with support from institutions. This study evaluated the study area's level of progress towards sustainability by considering the different indicators chosen for each pillar.

To calculate the sustainability level or sustainability development index (SDI), Equation (1) (see Table 1) was used, in which the SDI is evaluated as the average behavior of the sub-indices for the environmental, social, economic and institutional pillars. Different indicators were established to be used as inputs for the process. These will be described in Section 3.

**Table 1.** Equations used to calculate the sustainability level.

| | Equation | Variables | Studies Consulted |
|---|---|---|---|
| (1) | $\text{SDI} = \left(\frac{1}{4}\right)\sum_{1}^{4}\text{DI}$ | SDI = Sustainable development index<br>DI = Indexes of each dimension | [10,32] |
| (2) | $\text{I} = \sum_{i=x}^{n}(w_i * x_i)$ | I = Indicator<br>$w_i$ = Relative weight of the indicator<br>$x_i$ = Normalized value of each indicator | [10] |
| (3) | $y_t^i = \frac{x_t^i - \min(x_t^i)}{\max(x_t^i) - \min(x_t^i)} \in (0,1)$ | $y_t^i$ = Normalized value<br>$x_t^i$ = recorded data value for period t<br>$\min(x_t^i)$ = minimum data value of the indicator<br>$\max(x_t^i)$ = maximum data value of the indicator | [10,47] |
| (4) | $\text{DI} = \sum_{1}^{n}(\text{I})$ | DI = Index by dimension<br>n = number of indicators of the dimension<br>I = Indicator | [32] |

Each sub-index is calculated from the sum of the normalized indicators for each dimension by considering the relative weight of each within the dimensional index (see Equations (2)–(4) in Table 1).

The relative weight, $w_i$, in Equation (2) was calculated by using the analytic hierarchy process (AHP). This process is based on a paired comparison of variables, considering the Saaty Rating Scale, 1987 [48], the eight defined characteristics, and the values of importance established in the participatory work developed with the technical experts and the people with extensive knowledge of the territory. It is noteworthy that the indicators' level of importance, which was stated by the people with extensive knowledge of the territory, was one of the characteristics included in the AHP assessment.

In parallel, the min-max scaling method was used to normalize the indicators (see Equation (3) in Table 1). This method uses the distance between the maximum and minimum values of the analyzed indicators, considering the data of each indicator in the analysis period (2009–2017). Consequently, the indicator values were set to values in the 0–1 range, in which 0 represents the worst indicator performance, and 1 reflects the best performance [10,47,49].

Lastly, by conjugating the variables in Equation (1), the SDI was calculated for each analysis period. This same procedure was applied to the regular values or pre-established permissible levels for each indicator, either at the national level or based on international guidelines. This was done in order to compare the results for the study area with values established at the national and/or international levels that are deemed desirable for each indicator.

The calculated SDI values were the basis for the classification labels chosen in the supervised learning models that were applied in this study. Three categories for the sustainability levels were considered: Low (0.0–0.33), medium (0.34–0.66), and high (0.67–1.0).

### 2.3. Machine Learning Model

This study used three different supervised machine learning tools to classify sustainable development levels: Decision trees (C.5.0Tree), artificial neural networks (perceptron algorithm), and support vector machines (SVMradial) were used.

Decision trees (DTs) are a hierarchical predictive model of decisions and their consequences. They consist of nodes, branches, and leaves that characterize the model, and also establish the complexity of the decision tree. Complexity characteristics include the depth of the decision tree and the number of attributes used. The more complex the decision tree is, the more complexity there will be with respect to the accuracy of the results. Induction rules are applied when developing decision trees [50]. Different algorithms for decision trees have been developed, including the C.5.0tree, which evolved from C4.5. The C5.0tree algorithm is characterized by using entropy to measure the purity of tree divisions. This algorithm includes or removes predictors (in this case, indicators) based on their relationship with the labels established for supervised learning. In this manner, the model that is created includes only the most important predictors, taking into consideration that the error rate is reduced. In the event that the error rate is higher, due to not having included all the predictors in the classification model, they are left as predictors for the model [51].

For their part, artificial neural networks (ANNs) are mathematical models inspired by the biological functioning of neurons [51]. As with decision trees, this model is composed of nodes. In this case, they act as input, output, or intermediate processors connected to each other through links. They are characterized by their use of adaptive learning and self-organizing algorithms, and they process information in a non-linear manner. The node receives an input that has an associated weight, which is modified in the learning process. Basis and activation functions are necessary for the network to function.

Lastly, as a classification tool, support-vector machines (SVMs) use proximity to classify samples in a vector space. The maximum distance in the hyperplane is measured by the points closest to it. In this manner, the categories will have a distance from each side of the hyperplane, serving as a classification space. The representation by the mean of Kernel functions provides a solution to this problem, projecting the information to a larger characteristic space, which increases the computational capacity of the linear learning machine [52].

2.3.1. Information Required to Feed the Models

The information used to feed the models corresponds to two important inputs: Indicators according to dimension and supervised classification parameters.

(a) Indicators that describe the behavior of the study area according to the sustainable development dimension: Environmental, social, economic, and institutional. This study used machine learning tools on the indicators that were normalized through Equation (3) in Table 1, with information on yearly (81 indicators) and monthly (16 indicators) scales. An annualized basis of indicators was used, taking into account reporting characteristics in the study area. However, given the nature of how DTs, SVMs, and ANNs function, the results were derived from monthly information.

This study aims to establish a forecasting method for sustainability levels by using machine learning tools. Therefore, examples of variable data are required for the process to train and validate the models. Consequently, in the cases in which it was not possible to complete the monthly information, the indicator was discarded from the information base that would feed the model. Furthermore, in cases in which invariable information behavior was observed, these indicators were not included in the learning model with monthly information. That is, indicators, such as the drinking water supply, which during the year does not vary significantly, but which over the years has a degree of variation, as well as wastewater treatment. for example, were eliminated from the information set to be included in the model. In each case, it was verified that the sustainability pillars were represented in the indicators in order to develop the learning and classification process.

(b) Regarding the selection of classification parameters for supervised modeling, the results from the evaluation of the study area's sustainability level were used to establish the supervised classification labels. Three sustainability level categories were established: High (0.67–1.0), medium (0.34–0.66), and low (0.0–0.33). It is worth noting that given the characteristics of the results from the index calculation, scenarios were created to allow training data to be entered into the model, specifically for the low and high sustainability labels. These scenarios were generated by considering each indicator's threshold value, ensuring that the models had enough training examples in the data set and for validation, in accordance with the proposed scenarios. A 108-data point set was available for monthly reporting purposes, 70% of which was used for training and 30% for validation in the classification process. The same ratio for training and validating was applied to the yearly data set.

### 2.3.2. Performance Evaluation of Machine Learning Models

The metrics used in each model to measure its performance correspond to balanced accuracy, precision, recall, and specificity, or true negative rate, as determined by the confusion matrix. The matrix is a $3 \times 3$ table with different combinations of predicted and actual values regarding the classification labels (in this case, a high, medium, and low sustainability level). The balance accuracy metric prevents inflated performance estimates in unbalanced data sets. The metric determined the accuracy of the classifier to forecast each sustainability category: High, medium and low. In this vein, if the complete set of labels predicted for a sample strictly coincides with the real set of labels, the accuracy of the subset is 1.0. For its part, the precision metric made it possible to know the capacity of the classifier to not classify a result in a sustainability category or level that belongs to another category. The best results from this metric are 1.0, falling in an average close to 0.0. The recall metric refers to the classifier's capability to find all samples belonging to the sustainability category being evaluated, with a value of 1.0 referring to the best results for the metric.

Furthermore, the level of importance of the input variables was established by using the Gini index in the implementation of the supervised learning models.

To develop machine learning models, the open-source R software was used along with the caret package library, specifically for the following models: Decision tree (method: C5.0Tree) [53], artificial neural networks (method and package: nnet) [54], and the function of the package e1071 for the support vector machine [55].

## 3. Results

### 3.1. Characterization of the Study Area

A set of 81 indicators was established to be used as inputs for the process. The table presented in the Supplementary Material puts forth a description of the indicator set according to the dimension to which it belongs, the intersection if the indicator is part of an intersection (livable, equitable, viable), as well as the related sustainable development goal and target. Each indicator has an identification code, a combination of a letter and a number. The E letter identifies indicators belonging to the environmental dimension, the S letter identifies indicators belonging to the social dimension, the letters EC identify indicators belonging to the economic dimension and, the letter I identifies indicators of the institutional dimension. Table 2 presents an outline of the indicator set, displaying the number of indicators according to the characteristics established for each cell.

With regard to the environmental dimension, over the analysis period, the study zone has improved in terms of its indicators on air quality, waste collection, and areas allocated for green spaces. However, domestic wastewater generated in the locality is discharged into water sources without any type of treatment. On the other hand, while some indicators behave in a relatively constant manner, the importance of their improvement is noteworthy, specifically km$^2$ of green areas and recreational spaces.

**Table 2.** Set of indicators for the analysis of urban sustainability in the micro-territory.

| Dimension | Environmental = 13 | Social = 47 | Economic = 17 | Institutional = 4 |
|---|---|---|---|---|
| **Intersection** | Livable = 20, Equitable = 24, Viable = 1, Sustainable = 14 | | | |
| **Subject** | Air = 4<br>Water = 4<br>Waste = 2<br>Green spaces = 3 | Health = 21<br>Education = 6<br>Demography = 3<br>Security = 4<br>Coverage of public services = 6<br>Transportation = 2 | Economic structure = 7<br>Poverty = 3<br>Consumption and production = 7<br>Income and expenditure = 4<br>Employment = 4 | Government = 2<br>Social community services = 2 |
| **Related SDGs** | ▪ Clean water and sanitation<br>▪ Sustainable cities and communities<br>▪ Life of terrestrial ecosystems | ▪ End of poverty<br>▪ Zero hunger<br>▪ Health and well-being<br>▪ Quality education<br>▪ Gender equality<br>▪ Clean water and sanitation<br>▪ Industry, innovation, and infrastructure<br>▪ Sustainable cities and communities<br>▪ Peace, justice and solid institutions | ▪ End of poverty<br>▪ Health and well-being<br>▪ Affordable and no polluting energy<br>▪ Decent work and economic growth<br>▪ Sustainable cities and communities | ▪ End of poverty<br>▪ Partnerships to achieve goals |

With respect to the social dimension, a substantial number of indicators (25%) are related to the subject of health, given the influence exercised by socio-environmental determinants. These indicators' behavior does not reflect a marked upward or downward trend but responds specifically to the health determinant conditions present each year in the study area. Despite the variability, improvements are seen in indicators such as the child malnutrition rate, under-five mortality rate, all-cause infant mortality rate, and maternal mortality ratio.

Regarding the education indicators, gross education coverage decreased in 2016 and 2017 in the study area. However, the indicator behavior improved for areas such as years of schooling completed, illiteracy rate, population with middle and high school level education, and school attendance rate during the analysis period. Furthermore, with respect to population, the number of inhabitants per square kilometer has seen an upward trend, but the number of square kilometers with informal settlements has decreased, while coverage of the storm drainage system and the number of passengers transported by the mass transportation system have increased.

The study area is noted for having many security concerns, shown in indicators such as theft, aggravated robbery, and reports of domestic, family and child abuse, indicators which had a negative behavior trend during the study period.

Concerning its economic structure, the locality has high levels of its population living under the poverty line, with its highest recorded value in 2015, with 183,966 inhabitants in this condition. In the final two years of the study period, this indicator decreased by nearly 10%, in which there was a higher risk of water shortages (on average, 171 people ± 42). However, there was an improvement in indicators such as access to electricity (a yearly increase of nearly 2%), per capita household income, and improvements to the road network in the urban area.

Lastly, the institutional dimension is supported by policies and actions from the institutional sphere to meet the needs of the other pillars. The indicators that comprise this dimension had stable behavior during the analysis period.

As shown by the indicators, these characteristics are consistent with the frequency analysis of complaints filed by community members, which had high values concerning safety (15% of the 46,800 written complaints analyzed). This is in addition to the situation of the canonical correlation that enabled the indicators to be conjugated, which is described below.

Canonical Correlation

In the correlation analysis of the 81 indicators with an annual frequency in the period 2009–2017, the comparison between environmental protection and economic growth (see Figure 3) found a relation between indicators such as $PM_{10}$, $PM_{2.5}$, access to public services and the unemployment rate. The upper right-hand margin of Figure 3 shows an important grouping of economic indicators. All have positive behavior, in the sense of increased per capita household income (EC5), an increase in energy consumption (EC12), and growth of the employed population (EC3), for example. In this grouping, there are environmental indicators such as the average annual concentration of $PM_{10}$ (E1), the number of trees per hectare (E13), and the water quality of the Tunjuelito River (E10). Furthermore, the same quadrant includes indicators regarding $PM_{2.5}$ (E2) and the road network in good condition (EC15), both with improving trends.
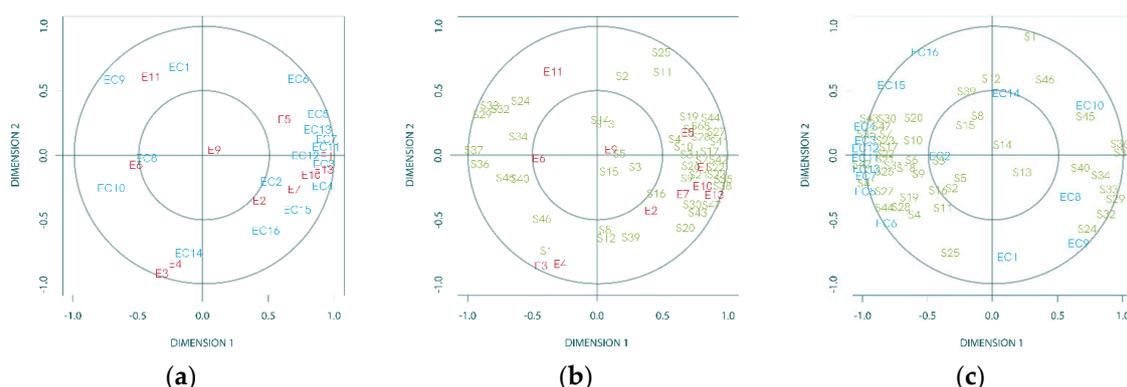
**Figure 3.** Canonical correlation between indicators of the (**a**) environmental and economic pillars, (**b**) environmental and social pillars, (**c**) social-economic pillars. Indicators with an E code refer to environmental indicators, indicators with an EC code refer to economic indicators, and indicators with an S code are social indicators. In (**a**–**c**), Dimensions 1 and 2 are the canonical variables that make the best representation of the total variation of indicators interactions. These canonical variables maximize the discrimination between groups of indicators. Supplementary Material shows the environmental, social, and economic indicators, represented in this figure through letters E, EC, or S, followed by an identification number.

The second chart (Figure 3b) shows an initial grouping of indicators that measure mortality rates: All-cause infant mortality (S6), under-five mortality from pneumonia (S4), under-five mortality (S10), perinatal mortality (S18), and life expectancy at birth (S28). The air quality index (E5) is included within this set of indicators in Figure 3b. There is also a set of health indicators such as acute malnutrition in children under five (S7) and the infant death rate (S21), indicators that characterize the physical conditions of the study area such as km$^2$ of areas susceptible to flooding (S38), as well as service indicators, which include the number of passengers who commute via the mass transportation system (S35) and households with access to natural gas service (S42). Furthermore, there are education indicators such as school attendance rate (S23), average years of schooling completed (S22), and population with a middle and high school education (S26). Another social indicator in this grouping corresponds to deaths due to firearms (S31). In addition to this set, there is the average annual concentration of $PM_{10}$ (E1) and closely related indicators such as the water quality of the Tunjuelito River (E10) and the number of trees per hectare (E13). This same chart shows the closeness of indicators that report excesses of $PM_{10}$ (E3) and $PM_{2.5}$ (E4), as well as the indicator that corresponds to the mortality rate due to cardiopulmonary disease, pulmonary circulation diseases and other forms of heart disease (S1).

Lastly, the third graph (see Figure 3c) shows a comparison between social inclusion and economic growth in which there is a correlation between indicators such as access to public services, the economically active population, and education level.

*3.2. Progress Level of Sustainable Development*

Applying Equations (1) to (4) (see Table 1), the sustainability categories were calculated for each analysis year in Kennedy. The locality has had low to medium sustainability levels (see Figure 4). However, the behavior in 2016 and 2017 surpassed the medium sustainability level (0.33–0.66). Moreover, the biogram presented in Figure 5 shows the behavior of the environmental, social, economic, and institutional sub-indices for the study area.
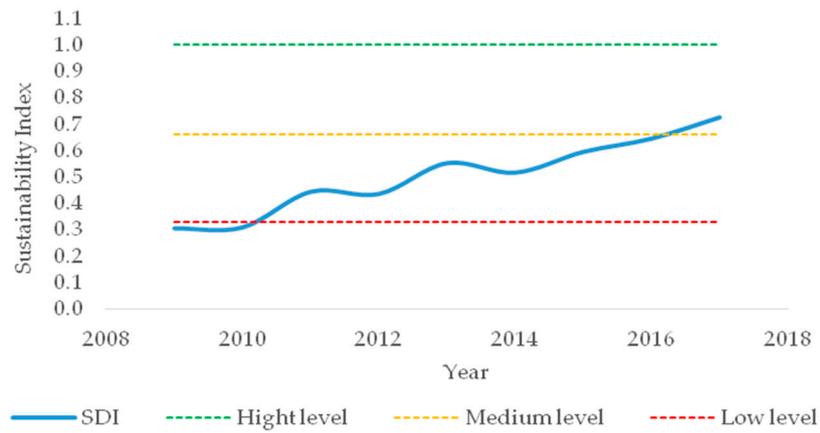
**Figure 4.** 2009–2017 sustainability index for the locality of Kennedy.
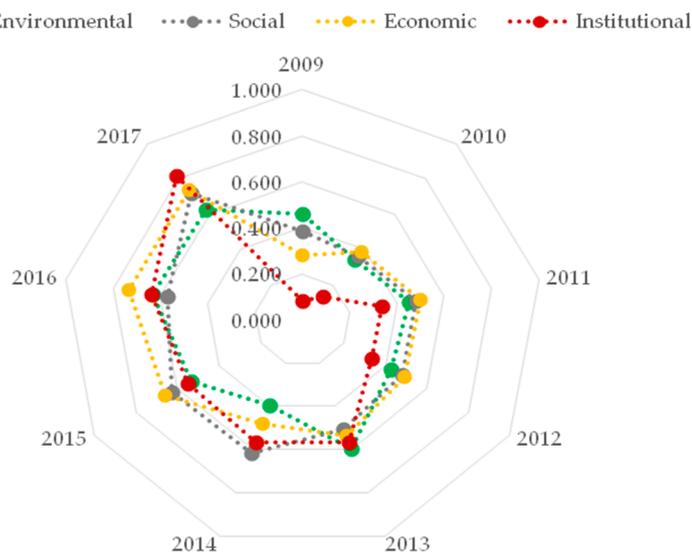


**Figure 5.** Biogram of the influence of the sustainable development dimensions for each year of the study period in the Kennedy urban area.

Figure 5 shows the influence of the institutional and economic dimensions, with a lag seen in the environmental pillar when compared with the other dimensions. In general, the behavior related to the SDI has improved for each dimension from 2015 to 2017.

*3.3. Machine Learning Model*

As mentioned in the methodological description, yearly and monthly information was used to develop the models. Each model was calibrated based on specific parameters for each machine learning tool, following the selection criteria provided by the kappa and accuracy measurements, as presented in Table 3.

**Table 3.** Calibration parameters for the machine learning tools to define the classification model of the Sustainable Development Index for the urban micro-territory.

| Tool | Calibration Parameters |
|---|---|
| Decision trees (C.5.0 Tree) | Iterations: 1<br>Kappa = 0.28 ± 0.17, Accuracy: 0.72 ± 0.11 |
| Artificial neural networks | Size: 1<br>Weight decay: 0.1<br>Kappa = 0.42 ± 0.12, Accuracy: 0.81 ± 0.07 |
| Support vector machine (SVMradial) | Sigma: 0.04<br>c: 1<br>Kappa = 0.4 ± 0.22, Accuracy: 0.83 ± 0.06 |

By applying the models, we found that due to the limited number of observations (nine data points for each indicator), models based on yearly information turn out to be inconclusive. Given the low volume of observations entered, it was not possible to forecast sustainability levels. However, using a monthly scale increased the number of observations, which enabled a greater volume of information to be available to train and validate the models. Table 4 presents the results for the three models developed. The labels high, medium and low correspond to the classification categories of the sustainability level assigned to the model for training and subsequent forecasting. Values with results in the 0.67–1 range belong to the high sustainability category, values with results in the 0.34–0.66 range correspond to the medium category, and values with results ranging from 0 to 0.33 belong to the low category.

**Table 4.** Metrics generated by the machine learning models in the classification of sustainability levels in the micro-territory.

| Model | Balanced Accuracy | | | Precision | | | Recall | | | Specificity | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Medium | Low | High | Medium | Low | High | Medium | Low | High | Medium | Low |
| Decision tree—C.5.0 Tree | 0.95 | 0.81 | 0.60 | 0.75 | 0.90 | 0.33 | 1.00 | 0.82 | 0.33 | 0.91 | 0.80 | 0.86 |
| Neural networks—Nnet | 0.96 | 0.80 | 0.50 | 0.75 | 0.85 | - | 1.00 | 1.00 | 0.00 | 0.92 | 0.60 | 1.00 |
| Support Vector Machine—SVMradial | 0.79 | 0.70 | 0.50 | 0.67 | 0.79 | - | 0.67 | 1.00 | 0.00 | 0.92 | 0.40 | 1.00 |

As this is a multi-class model as a whole, the decision tree model yields the best metrics (see Table 4). Decision trees and neural networks were 95% and 96% accurate, respectively. The high and medium territory sustainability categories were 81% and 80% accurate, respectively. While the support vector machine was not as accurate, it performed well in the classification, with values of 79% for the high category and 70% for the medium category.

The accuracy of the low classification category indicates that neural networks and the support vector machine classify the information for this category in a random manner. Only decision trees were 60% accurate in the low classification category.

These values are consistent with the results established by the precision metric, in which the decision tree and neural network models correctly predicted 75% of the labels in the high category. According to the recall metric, 100% of the labels for this category were forecasted. With respect to the medium sustainability category, the precision metric shows that 90% of the forecasted labels were correct in the decision tree model, and according to the recall metric, 82% of the category was forecasted.

Variable Importance Based on the Gini Index

For the decision tree model, the variables with the greatest importance were: Population with access to health services (S47), residential per capita water consumption (EC16), and excess $PM_{10}$ (E3) (see Figure 6). For the neural network model, the variables with the greatest importance were: Reports

of violence and domestic abuse (S32), excess PM$_{10}$ (E3), theft and aggravated robbery (S33), mortality rate due to pneumonia in adults older than 64 years of age (S3), and average annual concentration of PM$_{2.5}$ (E2) (see Figure 6). With respect to the support vector model, the most influential variables that exceeded 60% importance were: Population with access to health services (S47), passengers who commute via the public mass transportation system (S35), reports of violence and domestic abuse (S32), energy consumption (EC13), average annual concentration of PM$_{2.5}$ (E2), excess PM$_{10}$ (E3), and residential per capita water consumption (EC16). The above can be seen in Figure 6a–c, related to each forecasted level of sustainable development.
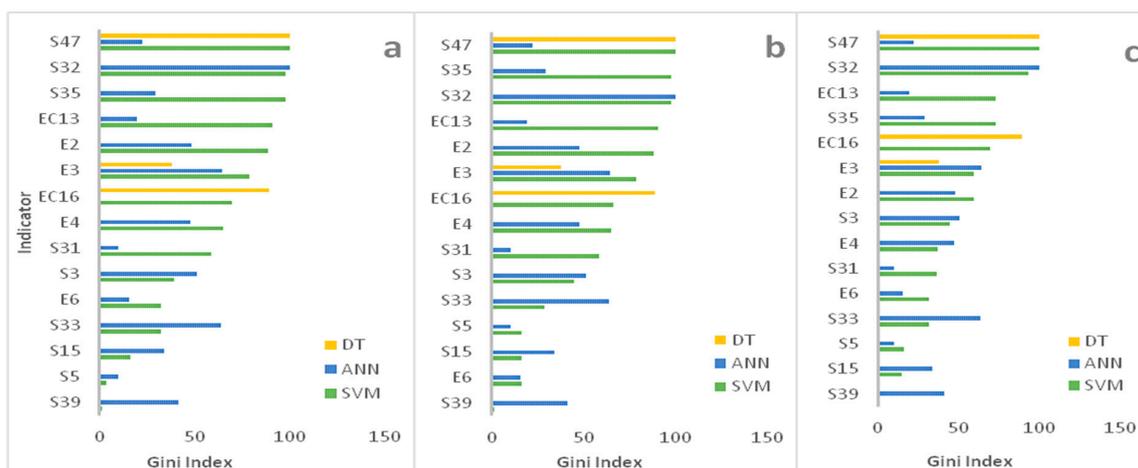


**Figure 6.** Variable importance according to the classification of the sustainable development index: (**a**) High sustainable development index score, (**b**) medium sustainable development index score, (**c**) low sustainable development index score. DT: Decision tree, ANN: Artificial neural network, SVM: Support vector machine. Supplementary Material includes the description of the environmental (E code), social (S code), and economic (EC code) indicators.

When comparing the most influential variables in the models, the excess of PM$_{10}$ variable (E3) is present in the three applied models, with similar levels of importance: 64% for ANN, 78.4% for SVM, and 37.8% for DT, for the high and medium sustainability categories (see Figure 6a,b). Additionally, its importance drops by 19 percentage points in the low category for the SVM model (see Figure 6c). While the population with access to health services variable (S47) is the most important variable in the DT and SVM models, it scores less than 30% in the ANN model. The role of the social dimension's variables, related to security, stands out, given its influence on the classification of sustainability levels of the urban area.

## 4. Discussion

The canonical correlation analysis found that the behavior described by the indicators shows that the urban area has different needs regarding the sustainability pillars and residents' quality of life. This is reflected in the interactions between indicators that seemingly do not show a direct relationship, yet describe specific determinants of the micro-territory's reality in the habitable and equitable interactions in the urban area [10].

There is an interaction between indicators such as the employed population between 12 and 64 years old (EC3), the economically active population (EC2), and indicators related to the habitable interaction, such as water quality of the Tunjuelo River (E13) and trees per hectare (E10). In addition to the analysis, there is a connection between indicators regarding economic issues and those that address social characteristics in the area, in terms of education and security (theft and violence). The grouping with the canonical correlation reflects behavior as described by Tanguay (2017) [10], for each of the pillars' interactions. Furthermore, the grouping of sustainability indicators, such as passengers

transported (S35), aging rate (S30), households with access to water (S42), energy consumption (EC12) and acute malnutrition of children (S7), which, despite the classification of specific issues, result in the interaction of sustainability dimensions in the territory. With respect to these interactions, it is important to note that the priorities in evaluating and measuring urban sustainability are determined by the territorial characteristics themselves [2]. That said, it is necessary to establish a comparison line in order to identify territories' evolution. To this end, the Sustainable Development Goals and its targets are an appropriate platform that brings together common goals.

Previous studies on the city of Bogotá have determined that the most relevant variables in the sustainable development index are poverty, crime, and unemployment [4], in which the index was calculated by applying a sustainability assessment by fuzzy evaluation. These variables are consistent with the results from this study in the complaints analysis as an input to prioritize indicators and calculate the Sustainable Development Index. However, it is considered that they should not be the only factor of interest as sustainable development is achievable only to the extent that interactions are addressed and balanced, such as the livable, viable and equitable dimensions [7,11], as shown by the canonical correlation analysis.

These indicators' behavior establishes that the population increase in the urban area and its resulting impacts, substantiate the need to advance a process of continuous feedback in order to support improving the conditions of the environmental, social, economic and institutional dimensions in territories. These are the results obtained from evaluating the Sustainable Development Index.

Kennedy is the second most populated territory in Bogotá. According to the SDI evaluation, the SDI of the urban area has moved from the low to the medium category over the period 2009–2015, with values that surpassed the medium sustainability category in 2016 and 2017 (See Figure 4). Prior studies have determined that Bogotá has reached a medium sustainability level (0.55, on a 0–1 scale), ranking 88 among 106 European, African, Asian, and Latin American cities [4]. Another study that applied multivariate statistical techniques [8] identified a medium sustainability level for Kennedy. Despite the difference in the methods applied to evaluate sustainability, these studies were consistent with the results presented in this paper. Furthermore, the variation in the numerical values recorded is limited, which is counterbalanced by studies that analyzed the variation in results with respect to the methodological variation in calculating sustainability, which yielded similar results even with different methodologies applied [10]. That said, it is important to note the importance of indicator selection for a relevant evaluation of sustainability.

Furthermore, a comparison of the influence of a micro-territory with better socio-economic behavior than Kennedy found that the results obtained through the SDI evaluation for Kennedy in this study are consistent with results from prior studies [8]. Teusaquillo is another micro-territory in Bogota, which, unlike Kennedy, is characterized by having greater purchasing power, more employed people, as well as having better educational, financial, cultural, and recreational services. In this vein, according to Carrillo and Toca (2013) [8], Teusaquillo achieved a high sustainable level in the evaluation. These are aspects that, despite the difference in methodologies, influence territories' progress towards sustainability.

Moreover, it has been noted that the development and implementation of a machine learning model require enough observations to ensure adequate training and validation of its behavior. This project faced limitations associated with not having enough information. Some of the available information corresponds to specific data concerning the city of Bogotá, primarily corresponding to the periods in which surveys, reports on the implementation of government plans, or the gathering of information for specific purposes were carried out. Planning and territorial evaluation processes do not consider creating range indicators for urban sustainability dimensions at the micro-urban territory level. In the face of these limitations, the following three specific aspects stand out:

(1) Benchmarking was used to select the indicators for this study, which was carried out by examining many existing studies on these types of indicators, in addition to reviewing the framework of the SDGs to achieve congruity amongst the indicators. The analyses presented herein are consistent

with those presented by L.-Y. Shen et al. (2011), Shen et al. (2013), and Verma et al. (2018), regarding the need to have valid objectives and targets for each territory as a clear support mechanism to evaluate progress made towards sustainability [2,3,26]. The indicators are matters of governance, but not issued by the government [8]. As such, it is necessary to develop a collection of historical data on territorial behavior, as this provides evidence of territories' evolution and support for sustainable development processes. Furthermore, given that population is an essential component of urban activities [2], participation from interest groups and including their needs to determine the set of indicators is necessary.

(2) The evolution of territories, as a goal of sustainable development in which human beings are the central axis of governments, requires coherence and coordination to identify, collect, and process information. Several studies use national statistics that have been published on various platforms for years prior to the implementation of the Millennium Development Goals as the basis for their information sources. Unfortunately, a clear example of the need to prioritize indicators can be seen in Latin American territories, where a greater impulse is required in information management, as demonstrated in the micro-territory analyzed in this study. It is also a mitigating circumstance for the capital city's position in the ranking of cities with the lowest sustainability levels, according to the results from Phillis et al. (2017) [4].

(3) At the international level, proposals for forecasting sustainable development in different cities and countries have been developed using indicators with a yearly scale [12,13,15]. However, the present study was not able to yield conclusive results for this time scale. In applying DTs, as one of the simplest tools for this type of classification problem, and the SVM and ANNs as robust tools, nine observations were not enough to properly train the model and validate its results. As stated above, 70% of the data was used for training, and 30% for behavioral validation. Therefore, using these types of tools requires large amounts of information, which prevents generalization problems and ensures the information's quality to support decision-making. In this vein, the model for this study reduced the working scale to monthly indicators, finding that the decision trees had the best behavior, with neural networks having the potential for improvement.

Lastly, the method applied and structured through this study established a logical procedure that begins with identifying the most influential parameters in an urban territory and concludes with forecasting their behavior in terms of sustainable development (see Figure 7). This procedure collected experiences developed in various studies that combine community participation in the territory, the technical expertise of professionals in areas of sustainable development, and the robustness offered by machine learning tools such as decision trees, neural networks, and support vector machines. This study was innovative in that it took a methodological step forward by integrating the community who are affected by their government's decisions, while including experiences from different studies, and the vision of the SDGs. It also integrated different tools for decision making, to be used for annual and statistical collection plans, as well as to manage the different resources that characterize the sustainability pillars.

Future studies should focus on the importance of having spatialized information, which enables the identification of the behavior of habitability interactions and the viability of sustainable development in different territories. This information can be used to forecast sustainability categories with machine learning tools as additional support for decision-making. Similarly, it can resolve difficulties in accessing information [2], even at the level of an urban micro-territory analysis, which was chosen for this study.
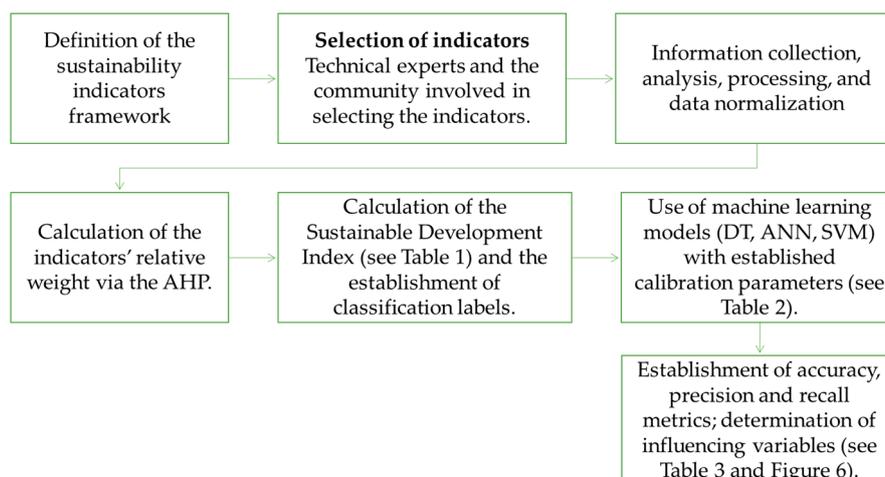
**Figure 7.** Suggested methodological process to classify the sustainable development index in urban micro-territories.

## 5. Conclusions

As shown in the present research, urban ecosystems include a combination of diverse micro-ecosystems, whose interaction supports economic development, yet leads to environmental damage and the deterioration or improvement of the population's quality of life. In this manner, the continuous evaluation and forecasting of this behavior contribute to developing strategies to improve the habitability, viability, and equity of urban territories with a view towards meeting the targets established by the SDGs.

While some studies have been developed to forecast sustainable development, these have focused either on specific sustainability dimensions or on understanding countries' evolution regarding the same. The latter are analyzed from a global perspective based on behavior in different territories. Along these lines, this study, which includes coordinating a series of procedures, contributes to the advancement of sustainability at the urban micro-territory scale. Its comprehensive method contributes to the academic and public arenas in the sense that it puts forth a tool that forecasts the category level of future sustainability in a micro-territory, such as Kennedy. It provides an opportunity to develop information-gathering strategies and action plans, as well as monitor their implementation.

This instrument stands out in the sense that it reduces the territorial and temporal scope of information, in order to have a better territorial observation and to make use of systematized tools to analyze the portfolio of governmental proposals as techniques in different fields of sustainability, thus contributing to habitability, viability, and equity interactions.

The micro-territory analyzed as a case study in this research study is representative of different environmental, social, and economic conditions in Bogota. Kennedy is one of the most populated areas of the city, is one of the most polluted zones in Bogota in terms of air quality, in addition to having high levels of insecurity. It also represents an important economically active population of the city. The results from this study show consistent progress in implementing several policies and show the value of using statistical and machine learning tools to identify behavioral patterns of variables that influence the performance of micro-territories in the city, which is useful for decision-makers. Currently, decision-makers need to understand future situations regarding the implementation of current measures. Knowing of indicators that influence sustainable development enables leaders to make more informed decisions.

Concerning the results of the statistical analysis and the important variables through the Gini index in machine learning models, it is important to note that the later reinforces results from traditional methods.

This study found limitations on information availability for indicators that describe the behavior of sustainability dimensions in the micro territory. It is necessary to have a significant amount of information either for an appropriate characterization of each sustainability dimension, or to feed the machine learning models. Therefore, the information gathering phase required the most time and resources of this study.

Further research studies will be able to apply the methodology developed herein, in conjunction with machine learning models for each micro-territory in Bogota. The studies contemplate an analysis of micro-territories and how sustainable dimensions and their interactions are influenced by socio-economic aspects. This will enable a comparative analysis of the behavior of micro-territories, taking into account indicators on the environmental, social, and economic dimensions, as useful tools for decision-making related to resource prioritization and allocation. Additionally, conducting research that considers spatialized information will identify the behavior of habitability interactions and the viability of sustainable development in different territories.

## References

1. United Nations, Department of Economic and Social Affairs, Population Division. *World Urbanization Prospects The 2018 Revision (ST/ESA/SER.A/420)*; United Nations: New York, NY, USA, 2019.
2. Shen, L.; Kyllo, J.; Guo, X. An Integrated Model Based on a Hierarchical Indices System for Monitoring and Evaluating Urban Sustainability. *Sustainability* **2013**, *5*, 524–559. [CrossRef]
3. Verma, P.; Raghubanshi, A.S. Urban sustainability indicators: Challenges and opportunities. *Ecol. Indic.* **2018**, *93*, 282–291. [CrossRef]
4. Phillis, Y.A.; Kouikoglou, V.S.; Verdugo, C. Urban sustainability assessment and ranking of cities. *Comput. Environ. Urban Syst.* **2017**, *64*, 254–265. [CrossRef]
5. United Nations. *Report of the World Commission on Environment and Development: Our Common Future*; Available online: http://www.un-documents.net/our-common-future.pdf (accessed on 18 January 2019).
6. Gerry Marten, Human Ecology: Basic Concepts for Sustainable Development—Populations and Feedback Systems. Available online: http://gerrymarten.com/ecologia-humana/capitulo02.html (accessed on 28 January 2020).
7. Tanguay, G.A.; Rajaonson, J.; Lanoie, P. Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecol. Indic.* **2010**, *10*, 407–418. [CrossRef]
8. Carrillo-Rodríguez, J.; Toca, C.E. Sustainable performance in Bogota: Building an indicator from local performance. *EURE* **2013**, *39*, 165–190.
9. Mapar, M.; Jafari, M.J.; Mansouri, N.; Arjmandi, R.; Azizinejad, R.; Ramos, B. Sustainability indicators for municipalities of megacities: Integrating health, safety and environmental performance. *Ecol. Indic.* **2017**, *83*, 271–291. [CrossRef]
10. Rajaonson, J.; Tanguay, G.A. A sensitivity analysis to methodological variation in indicator-based urban sustainability assessment: A Quebec case study. *Ecol. Indic.* **2017**, *83*, 122–131. [CrossRef]
11. Toumi, O.; Le Gallo, J.; Ben Rejeb, J. Assessment of Latin American sustainability. *Renew. Sustain. Energy Rev.* **2017**, *78*, 878–885. [CrossRef]

12. Li, Y.; Wu, Y.-X.; Zeng, Z.-X.; Guo, L. Research on forecast model for sustainable development of Economy-Environment system based on PCA and SVM. In Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, China, 13–16 August 2006; pp. 3590–3593.

13. Zhang, Y.; Huan, Q. Research on the evaluation of sustainable development in Cangzhou city based on neural-network-AHP. In Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, China, 13–16 August 2006; pp. 3144–3147.

14. Pérez-Ortíz, M.; de La Paz-Marín, M.; Gutiérrez, P.A.; Hervás-Martínez, C. Classification of EU countries' progress towards sustainable development based on ordinal regression techniques. *Knowl.-Based Syst.* **2014**, *66*, 178–189. [CrossRef]

15. Zhang, Y.; Shang, W.; Wu, Y. Research on sustainable development based on neural network. In Proceedings of the 2009 Chinese Control and Decision Conference, Guilin, China, 17–19 June 2009; pp. 3273–3276.

16. Dizdaroglu, D. Developing micro-level urban ecosystem indicators for sustainability assessment. *Environ. Impact Assess. Rev.* **2015**, *54*, 119–124. [CrossRef]

17. Distrital Secretariat of Planning. *Monograph 2017 Diagnosis of the Main Territorial, Infrastructure, Demographic and Socio-Economic Aspects Kennedy Locality 08*; Bogota City Hall: Bogotá, Colombia, 2018.

18. United Nations. *Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development*; United Nations: New York, NY, USA, 2017.

19. United Nations. *Indicators of Sustainable Development: Guidelines and Methodologies*, 3rd ed.; No. October; United Nations: New York, NY, USA, 2007.

20. Niemeijer, D.; de Groot, R.S. A conceptual framework for selecting environmental indicator sets. *Ecol. Indic.* **2008**, *8*, 14–25. [CrossRef]

21. Quiroga Martínez, R.; Stockins, P.; Holloway, M.; Taboulchanas, K.; Sanchez, A. *Methodological Guide for Developing Environmental and Sustainable Development Indicators in Latin American and Caribbean Countries*; United Nations: New York, NY, USA; Economic Commission for Latin America and the Caribbean: Santiago de Chile, Chile, 2009.

22. Scipioni, A.; Mazzi, A.; Mason, M.; Manzardo, A. The Dashboard of Sustainability to measure the local urban sustainable development: The case study of Padua Municipality. *Ecol Ind.* **2009**, *9*, 364–380. [CrossRef]

23. Alpopi, C.; Manole, C.; Colesca, S.E. Assessment of the sustainable urban development level through the use of indicators of sustainability. *Theor. Empir. Res. Urban Manag.* **2011**, *6*, 78–87.

24. Cecchini, S. *Social Indicators in Latin America and the Caribbean*; United Nations: New York, NY, USA; Economic Commission for Latin America and the Caribbean: Santiago de Chile, Chile, 2005.

25. Klopp, J.M.; Petretta, D.L. The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities* **2017**, *63*, 92–97. [CrossRef]

26. Shen, Y.J.; Ochoa, J.; Shah, M.N.; Zhang, X. The application of urban sustainability indicators—A comparison between various practices. *Habitat Int.* **2011**, *35*, 17–29. [CrossRef]

27. Hák, T.; Janoušková, S.; Moldan, B. Sustainable Development Goals: A need for relevant indicators. *Ecol. Indic.* **2016**, *60*, 565–573. [CrossRef]

28. Escobar, L. Synthetic indicators of environmental quality: A general model for large urban areas. *Rev. EURE* **2006**, *32*, 73–98.

29. Sotelo, J.A.; Tolón, A.; Lastra, X. Indicators for and by sustainable development, a case study. *Estud. Geográficos.* **2012**, *72*, 611–654. [CrossRef]

30. Feleki, E.; Vlachokostas, C.; Moussiopoulos, N. Characterisation of sustainability in urban areas: An analysis of assessment tools with emphasis on European cities. *Sustain. Cities Soc.* **2018**, *2018. 43*, 563–577. [CrossRef]

31. Ocampo, L.; Ebisa, J.A.; Ombe, J.; Geen Escoto, M. Sustainable ecotourism indicators with fuzzy Delphi method—A Philippine perspective. *Ecol. Indic.* **2018**, *93*, 874–888. [CrossRef]

32. Torres-Delgado, A.; López Palomeque, F. The ISOST index: A tool for studying sustainable tourism. *J. Destin. Mark. Manag.* **2018**, *8*, 281–289. [CrossRef]

33. Distrital Secretariat of Planning. *"Knowing Kennedy: Diagnosis of Physical, Demographic and Socioeconomic Aspects"*; Bogota City Hall: Bogota, Colombia, 2009.

34. SALUDATA—Health Observatory of Bogota 2019. Available online: http://saludata.saludcapital.gov.co/osb/index.php/datos-de-salud/salud-ambiental/consultaurgencias14anios/ (accessed on 28 February 2019).

35. Hospital del Sur. *E.S.E. Local Diagnosis with Social Participation 2014 Locality of Kennedy*; Bogota City Hall: Bogota, Colombia, 2014.

36. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2009*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_2009_RMCAB.pdf (accessed on 20 August 2019).
37. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2012*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_2012_RMCAB.pdf (accessed on 20 August 2019).
38. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2015*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_RMCAB_2015.pdf (accessed on 20 August 2019).
39. Local Mayor of Kennedy. *Local Risk Management and Climate Change Council General Characterization of Risk Scenarios*; Bogota City Hall: Bogota, Colombia, 2018.
40. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2014*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_2014.pdf (accessed on 20 August 2019).
41. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2013*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_anual_2013_CalidadAire-RMCAB_V3.pdf (accessed on 20 August 2019).
42. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2011*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_2011.pdf (accessed on 20 August 2019).
43. Distrital Secretariat of the Environment. *Bogota Annual Air Quality Report 2010*. Available online: http://rmcab.ambientebogota.gov.co/Pagesfiles/Informe_Anual_2010.pdf (accessed on 20 August 2019).
44. Portal Geoestadistico 2019. Available online: http://www.sdp.gov.co/gestion-estudios-estrategicos/informacion-cartografia-y-estadistica/portal-geoestadistico (accessed on 20 November 2019).
45. Bogota City Hall. *Local Diagnosis with Social Participation 2009–2010 Locality of Kennedy*; Bogota City Hall: Bogota, Colombia, 2010.
46. Bogota City Hall. *Local Environmental Plan Kennedy Better for All Locality Example for All 2017–2020*; Bogota City Hall: Bogotá, Colombia, 2016.
47. Cui, X.; Fang, C.; Liu, H.; Liu, X. Assessing sustainability of urbanization by a coordinated development index for an Urbanization-Resources-Environment complex system: A case study of Jing-Jin-Ji region, China. *Ecol. Indic.* **2019**, *96*, 383–391. [CrossRef]
48. Saaty, R.W. The analytic hierarchy process-what it is and how it is used. *Math. Model.* **1987**, *9*, 161–176. [CrossRef]
49. Schuschny, A.; Soto, H. *Methodological Guide for Composite Indicators Design for Sustainable Development*; United Nations: New York, NY, USA; Economic Commission for Latin America and the Caribbean: Santiago de Chile, Chile, 2009.
50. Rokach, L.; Maimon, O. *Data Mining with Decision Trees: Theory and Applications*, 2nd ed.; World Scientific Publishing Co. Pte. Ltd.: Singapore, 2015; p. 5.
51. Kuhn, M. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.
52. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
53. Caret: Classification and Regression Training. R package version 6.0-72. Available online: https://CRAN.R-project.org/package=caret (accessed on 10 February 2020).
54. nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models. Available online: https://cran.r-project.org/web/packages/nnet/index.html (accessed on 29 November 2019).
55. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. Available online: https://cran.r-project.org/web/packages/e1071/index.html (accessed on 29 November 2019).