sustainability

MDPI

Article

Characterisation of Youth Entrepreneurship in Medellín-Colombia Using Machine Learning

Adelaida Ojeda-Beltrán [1], Andrés Solano-Barliza [2], Wilson Arrubla-Hoyos [3], Danny Daniel Ortega [1], Dora Cama-Pinto [4,5,*], Juan Antonio Holgado-Terriza [6], Miguel Damas [4], Gilberto Toscano-Vanegas [7] and Alejandro Cama-Pinto [7,*]

1 Faculty of Economy, Universidad del Atlántico, Puerto Colombia 081007, Colombia; adelaidaojeda@mail.uniatlantico.edu.co (A.O.-B.); danny.ortega.23@gmail.com (D.D.O.)
2 Faculty of Engineering, Universidad de La Guajira, Riohacha 440002, Colombia; andresolano@uniguajira.edu.co
3 Faculty of Engineering, Universidad Nacional Abierta y a Distancia, Sincelejo 700002, Colombia; wilson.arrubla@unad.edu.co
4 Department of Computer Architecture and Technology, University of Granada, 18071 Granada, Spain; mdamas@ugr.es
5 Faculty of Industrial Engineering, Universidad Nacional Mayor de San Marcos, Lima 15081, Peru
6 Software Engineering Department, University of Granada, 18014 Granada, Spain; jholgado@ugr.es
7 Department of Computer Science and Electronics, Universidad de la Costa, Barranquilla 080002, Colombia; gtoscano1@cuc.edu.co
* Correspondence: dora.cama@unmsm.edu.pe (D.C.-P.); acama1@cuc.edu.co (A.C.-P.)

Abstract: The aim of this paper is to identify profiles of young Colombian entrepreneurs based on data from the "Youth Entrepreneurship" survey developed by the Colombian Youth Secretariat. Our research results show five profiles of entrepreneurs, mainly differentiated by age and entrepreneurial motives, as well as the identification of relevant skills, capacities, and capabilities for entrepreneurship, such as creativity, learning, and leadership. The sample consists of 633 young people aged between 14 and 28 years in Medellín. The data treatment was approached through cluster analysis using the K-means algorithm to obtain information about the underlying nature and structure of the data. These data analysis techniques provide valuable information that can help to better understand the behaviour of Colombian entrepreneurs. They also reveal hidden information in the data. Therefore, one of the advantages of using statistical and artificial intelligence techniques in this type of study is to extract valuable information that might otherwise go unnoticed. The clusters generated show correlations with profiles that can support the design of policies in Colombia to promote an entrepreneurial ecosystem and the creation and development of new businesses through business regulation.

Keywords: artificial intelligence; machine learning; data mining; K-mean; youth entrepreneurship

Citation: Ojeda-Beltrán, A.; Solano-Barliza, A.; Arrubla-Hoyos, W.; Ortega, D.D.; Cama-Pinto, D.; Holgado-Terriza, J.A.; Damas, M.; Toscano-Vanegas, G.; Cama-Pinto, A. Characterisation of Youth Entrepreneurship in Medellín-Colombia Using Machine Learning. *Sustainability* **2023**, *15*, 10297. https://doi.org/10.3390/su151310297

Academic Editors: Feng Liu, Peiwan Wang and Jiayin Qi

Received: 24 April 2023
Revised: 26 June 2023
Accepted: 26 June 2023
Published: 29 June 2023

1. Introduction

Entrepreneurial culture in Latin America is an increasingly important issue because of its potential to drive regional economic and social development. Some of the defining characteristics of entrepreneurial culture in Latin America include creativity and innovation, adaptability, resilience in the face of challenges, and a strong focus on collaboration and teamwork; however, much remains to be done [1].

An entrepreneurial culture is fundamental to the economic and social development of any country, and Colombia is no exception. Fostering an entrepreneurial culture in Colombia means promoting entrepreneurship and creativity, which can create new jobs and increase innovation in the country. Youth entrepreneurship in Colombia has some characteristics that distinguish it from other types of entrepreneurships. First, it is common

for young entrepreneurs in Colombia to create innovative enterprises, especially start-ups [2]. They also tend to be very focused on innovation and technology, using digital tools and technological solutions to improve their products and services [3]. It is also common for young entrepreneurs in Colombia to seek to solve social or environmental problems through their ventures [4], reflecting a concern for the country's problems and a desire to contribute to solving them as young entrepreneurs. Through its regulatory framework, Colombia is strengthening the country's entrepreneurial ecosystem and placing it at the centre of economic and social reactivation. This will strengthen Colombia's future and position it as an entrepreneurial nation and regional leader in innovation. According to the OECD, the macroeconomic environment, growth prospects, public institutional capacity, and policy and regulatory frameworks are essential elements to make this a reality [4].

Youth entrepreneurship has been studied from different perspectives; in this sense, advances in the fields of artificial intelligence (AI) and machine learning (ML) in recent years have provided researchers with new methodological tools to build models that can predict different human behaviours and provide accurate information about the predictive capacity of entrepreneurs. Studies by Schade and Schuhmacher [5] show how by using different supervised machine learning techniques, including decision tree, random forest, artificial neural network, k-nearest neighbour, extreme increasing degree tree ensemble, and naïve Bayesian, it is possible to obtain a baseline and estimate their relative values and achieve prediction from a model based on the Global Entrepreneurship Monitor dataset of 1,192,818 individuals from 99 countries, with an overall accuracy ranging from 70.1% to 91.2%. The results show that entrepreneurial self-efficacy and opportunity recognition are the most important characteristics for predicting opportunity-driven entrepreneurial activity.

Society demands knowledge of the primordial value of lifelong learning and research to generate knowledge that enables innovation in all sectors. Entrepreneurial culture is essential to this process, as it requires knowing and identifying cultural differences, creating knowledge, and promoting rapid economic and social growth [6].

Youth entrepreneurship perspectives are promoted in business models based on the gig economy (small order economy, on-demand). It has become an essential driver for the creation of new businesses [3], such as digital ones promoted by disruptive innovations [7], making entrepreneurship a viable tool to counter the effects of the pandemic [8]. It is not about specific actions at specific points in time but about a strategic and entrepreneurial orientation that is sustainable over time, as well as the assumption of responsibility and value-creating stakeholder networks [9]. A crisis such as COVID-19 poses challenges and provides new opportunities that increase people's intentions to start businesses [10].

The aim of this paper is to apply machine learning algorithms (K-means) to obtain a detailed analysis of the characteristics that define a dataset, which in this case was the "Youth Entrepreneurship" survey developed by the Secretariat of Youth of Colombia, to obtain some knowledge and insights about the qualities and skills for entrepreneurship among young people in Medellín.

2. Theoretical Framework

Entrepreneurship is a highly complex phenomenon that has been extensively studied in various disciplines, such as economics, psychology, and sociology. This section reviews the existing literature to gain a deeper understanding of the key concepts and factors associated with entrepreneurship. In this context, it is emphasised that entrepreneurship occurs when a person discovers an investment opportunity, i.e., when he or she manages to identify and exploit favourable situations [11]. Similarly, Gelderen et al. [12] point out that it implies the ability to identify and exploit opportunities. On the other hand, Lumpkin and Dess [13] define entrepreneurship as the ability to develop an entrepreneurial spirit.

Entrepreneurship refers to the personal characteristics and qualities that enable an individual to identify and seize business opportunities, take calculated risks, be innovative, and have a proactive mindset. According to Lumpkin and Dess [13], entrepreneurship

encompasses the ability to make independent decisions, the passion to achieve challenging goals, and the willingness to take responsibility. Baron [14] states that entrepreneurship involves the ability to see opportunities where others see problems, to take calculated risks, and to have the determination and tenacity to turn ideas into concrete realities. In other words, entrepreneurship implies the ability to identify investment opportunities and refers to the personal characteristics that enable someone to identify and seize those opportunities, take risks, be innovative, and have a proactive attitude in the pursuit of challenging goals.

2.1. Determinants of Youth Entrepreneurship

During childhood, children define entrepreneurial characteristics based on their creativity, risk-taking, and resilience, which are fostered by the parental and social environment [15]. Later, adolescents' entrepreneurship further depends on their social environment, motivation, resources, autonomy, risk-taking, and innovativeness [16–18]. They supplement their education with (a) courses, (b) workshops, (c) guest speakers, (d) simulations, and (e) entrepreneurship competitions [19–22]. At the university level, there is a link between students' entrepreneurial learning, knowledge acquisition/creation/use and their entrepreneurial mindset. Their competencies need to be reconfigured to strike a balance between exploring new roles and using existing ones to adapt to market demands [23]. Teacher–student support and student–student relationships are essential in creating the learning environment [24] which in some cases translates into the creation of incubators that play an essential role in the entrepreneurial ecosystem [25].

In addition to the above, the results presented in a study of youth entrepreneurship in low-income countries [26] found that discrimination associated with the condition of migrants encourages entrepreneurship. In the Latin American context, the phenomenon of migration is mainly produced by young people of Venezuelan nationality. In the case of Colombia, not only does the phenomenon of Venezuelan migration converge, but also the internal displacement of rural communities to cities caused by years of internal civil conflict. Despite their remarkable resilience, commitment, creativity, and personal resources in terms of skills and family support, most returnees find themselves in a situation of permanent insecurity. They face difficulties in meeting their basic, existential, and security needs due to several challenges [27].

On the other hand, Zulkifle and Aziz [28] propose four new independent variables: social awareness, prior entrepreneurial experience, cosmopolitanism, and cultural intelligence as determinants of entrepreneurship. Similarly, Rusu et al. [29] suggest that the determinants of entrepreneurial intention are related to aspects such as access to finances, knowledge, and education. Entrepreneurship education can significantly influence the determinants of entrepreneurship by providing the knowledge, skills, competencies, and networks necessary for successful entrepreneurship.

In that sense, entrepreneurship education can help young and potential entrepreneurs overcome barriers and develop the skills needed to start and run a successful business [30]. In recent years, the growing importance of entrepreneurship education has been recognised as central to promoting entrepreneurship and preparing individuals to meet the challenges of today's business environment. The literature consistently supports the need for sound entrepreneurship education that provides students with the skills and knowledge necessary to start and run their own businesses. In particular, the European Commission's report on entrepreneurship in education (2017) highlights the importance of entrepreneurship education in fostering an entrepreneurial mindset from an early age. Furthermore, research, such as that conducted by Fiet [31], highlights the need for a more holistic approach to entrepreneurship education, emphasising the importance of practice and practical experience rather than focusing solely on theory and concepts. Finally, a study by Neck et al. [32] highlights the importance of integrating the development of entrepreneurial skills into all levels of education, not only by offering specific entrepreneurship courses, but also by incorporating entrepreneurship elements into different disciplines and educational

programmes. This approach will enable students to acquire an entrepreneurial mindset and develop transferable skills that are relevant in today's business world.

A review of the literature allows us to conclude that several characteristics of the entrepreneur have been considered (described by Ramírez et al. [33] and summarised in Table 1).

Table 1. Characteristics of the Entrepreneur.

Reference	Characteristics
[34]	Risk-taking, tenacity, innovation, creativity, self-efficacy, industry knowledge, desire for money, optimism
[35]	Risk-taking, independence, leadership, knowledge of the industry, desire for money, hard work, integrity, teamwork
[36]	Independence, tenacity, self-efficacy, determination
[37]	Tenacity, innovation, self-efficacy, industry knowledge, ability to plan, passion, continuous learning, integrity, tolerance of frustration

Source: [33]

2.2. Challenges and Difficulties in Youth Entrepreneurship

The challenges and difficulties of entrepreneurship [16] suggest that a scarcity of financial resources is the main challenge to starting a new business, coupled with the difficulty of procedures that make potential entrepreneurs more reluctant to start their business legally. Consequently, it is helpful to implement mechanisms of access to credit and communication to minimise the costs and maximise the benefits for the entrepreneurial seed [38–40]. On the other hand, Kebede [41] also identifies deficits in entrepreneurship education, lack of effective administrative and regulatory frameworks, and lack of access to better technology systems and infrastructure, among others, for which fostering entrepreneurship through video games can be considered an innovative alternative [42].

In the context of collective entrepreneurship, Chatti and Hamrouni [43] identify the critical success factors and difficulties faced by entrepreneurs in developing business opportunities. The results of this study, which involved five entrepreneurs who had experienced collective entrepreneurship, show that diversity of knowledge and experience, as well as interpersonal conflicts, can be obstacles that hinder the development of business opportunities in the context of collective entrepreneurship.

In the Colombian case, the main difficulties focus on government support for entrepreneurship as a formal institution. However, given the lack of diversification through industrialisation and import substitution, as well as the violence that the country faces, entrepreneurship is a challenge [40]. The institutional framework for entrepreneurship in the country is not well developed [40]. A country's institutional framework plays a key role in fostering entrepreneurship, which drives economic growth. Institutional factors include political stability, government effectiveness, regulatory quality, strong rule of law, ease of starting a new business, and availability of credit [44].

2.3. Artificial Intelligence

Machine learning is the most common application of artificial intelligence and is mainly based on supervised learning in which AI is trained on human-provided eth-quotients. Deep learning, on the other hand, involves training on unlabelled data, allowing the machine to discover underlying patterns through its algorithms.

Artificial intelligence, as a tool applied to the study of entrepreneurship, and how they are combined and used, will determine their impact on humanity. Although researchers have made independent advances in understanding entrepreneurship and artificial intelligence, these two areas of research have largely progressed in parallel [45].

AI is a broad field that includes aspects of intelligence, perception, sensing, reasoning, and aspects of ML [46]. ML is a subfield of AI, defined as a set of methods that can automatically detect patterns in data to predict (future) or describe (present) [47]. Data

mining (DM) uses ML techniques to find patterns and correlations within large datasets to predict outcomes. Using techniques ranging from statistics to machine learning or artificial intelligence [48], we can conduct a comprehensive analysis of the Youth Entrepreneurship study and even predict future trends.

ML algorithms are classified into three paradigms: supervised learning, unsupervised learning, and reinforcement learning [47]. Supervised learning is used when the desired outputs are known, and a training process is carried out to obtain those desired outputs. When information about expected outputs is unavailable, grouping techniques that do not require supervision are often applied [46].

Given the current situation, this research aims to identify profiles of young Colombian entrepreneurs using data from the "Youth Entrepreneurship" survey developed by the Secretariat of Youth. We use a clustering method based on the K-means machine learning algorithm.

Clustering divides a dataset into groups so that the entities in each cluster are comparatively more like those in that cluster than those in the other clusters [49]. The purpose of clustering is to organise a dataset into groups so that the entities in each group are alike but different from those in the other groups. K-means and hierarchical clustering techniques are the most widely used clustering algorithms in the literature [50].

Clustering generates several subsets of a dataset, each of which contains data with a certain degree of similarity [51]. For numerical data, this similarity is given by the distance between them in the data space; the most similar are those with the minimum distance between them. As a reference point, the centre of each data group is the point at which, implicitly, the data close to this centre are also close. In this sense, the K-means clustering method [52] was developed to identify data groups that can be analysed for deeper purposes, such as classification, extraction of relevant information, and discovery of new information. K-means clustering is generally effective when data are concentrated around a central point, forming a spherical (or hyperspherical)-shaped subset in the data space [53].

The application of machine learning algorithms in population characterisation studies is of great importance, since by analysing large amounts of population data, it is possible to find hidden patterns and trends; machine learning algorithms can also help to segment the population into different groups with specific characteristics and needs.

2.4. Context of Entrepreneurship Research in Colombia

Colombia has several laws and regulations that promote entrepreneurship and business creation. Some of the most important laws related to entrepreneurship in the country are (a) The Law on Entrepreneurship in Colombia (Law 2069), enacted in 2021, which was approved with the aim of promoting the creation, growth, and consolidation of businesses. This law establishes benefits and incentives for entrepreneurs, such as the simplification of procedures and the promotion of an entrepreneurial culture in the country. (b) The Law on Financing: Although not exclusively focused on entrepreneurship, this law introduced several measures to encourage investment and entrepreneurship in Colombia. These include tax benefits for investors and support for the creation of start-ups. (c) The Law on Formalisation and Job Creation: This law aims to encourage the formalisation of companies and the creation of jobs in the country. It provides benefits to companies that formalise, such as reduced administrative burdens and access to employment support programmes. (d) The Law on Science, Technology, and Innovation: Although not exclusively focused on entrepreneurship, this law promotes research, technological development, and innovation in the country. It establishes funding and support mechanisms for science- and technology-based entrepreneurship projects.

In Colombia, the promotion of the culture of entrepreneurship and the establishment of a regulatory framework to promote its growth and consolidation define the entrepreneur as "a person with the capacity to innovate, understood as the ability to produce goods and services in a creative, methodical, ethical, responsible and effective manner" and

entrepreneurship as "a way of thinking, reasoning and acting focused on opportunities, raised with a global vision and carried out through balanced leadership and calculated risk management, whose result is the creation of value that benefits the company, the economy and society" [54,55].

Regarding the dynamics of entrepreneurial activity in Colombia, the Global Entrepreneurship Monitor—GEM—presented in May 2022, in its report for the country, presents a panorama of great contrasts, evaluated to verify the conditions of the entrepreneurial ecosystem, highlighting financing, infrastructure, public policies, and government programmes. The GEM mentions that the national index of entrepreneurial ecosystem conditions (NECI) has improved in the productive and entrepreneurial characteristics of new businesses, surpassing those of Chile and Mexico [3].

Colombians' perceptions of entrepreneurship, measured by the variables of self-confidence, perception of risk, perception of opportunity, and evaluation of entrepreneurship as a life alternative, showed a cheerful disposition before the COVID-19 pandemic [56] and adopted a more demanding evaluation line afterwards. The country has achieved an economic reactivation of the productive system, demonstrating its capacity for resilience and renewal.

The studies carried out by Ramírez et al. [33] broaden and deepen the concept of the "entrepreneur" from a personal perspective. The information obtained from previous research is compiled, grouped, and graphically presented, and the psychological, sociological, and demographic dimensions are identified as perspectives for studying the entrepreneur, assigning to each of them the key factors of the individual and the environment, as well as the specific qualities and skills of the entrepreneur, thus creating an approximate profile of the entrepreneur. On the other hand, Quillas et al. [57] consider the relationship between the quality of life of individuals and their relationship with the society in which they live. Although quality of life is not limited to variables related to monetary wealth, the variables used, such as access to water and sanitation, information and communication, and advanced education, depend on the capacity of societies to generate wealth, making the role of entrepreneurs fundamental in the wealth creation process.

2.5. Review of the Literature

The results of a literature review in the field of entrepreneurship using a machine learning model show the results proposed by Lu et al. [58], who try to identify potential growth areas and improve the skills needed for entrepreneurship education among university students by integrating two powerful algorithms: random forest (RF) and logistic regression (LR). The main contributions of the work are the construction of a quality index for each topic of interest, using and ranking the indicators according to the quality index to assess strengths and weaknesses. In the same vein, Graham and Bonner [59] explore the determinants of entrepreneurship by applying decision tree algorithms to a large dataset from the Global Entrepreneurship Monitor. The results show considerable heterogeneity in the factors contributing to entrepreneurship, highlighting the need for academics and policy makers to consider the likelihood that there is no single set of motivating factors. Gosztonyi and Judit [60] examine the characteristics of entrepreneurs using the following algorithms: multivariate adaptive regression spline (MARS), support vector machine (SVM), random forest (RF), and AdaBoost.

On the other hand, Zhang et al. [61] propose an intelligent and reliable prediction model of entrepreneurial intentions to support and guide the positive entrepreneurial intentions of university students. The model mainly uses the improved crowd search algorithm (CSA) to optimise the kernel extreme learning machine (KELM) model with feature selection (FS), namely CSA-KELM-FS, to investigate entrepreneurial intention. The literature review, which corresponds to the search chain characteristics of entrepreneurship in the main databases Scopus/IEEE in an observation window of 10 years, is intended to expose the variables identified by the authors. Within the search are identified some of

the variables used in the construction of the questionnaire of questions asked by the entity responsible; this information is used in the results section.

3. Materials and Methods

3.1. Participant Data

According to the National Administrative Department of Statistics (DANE), Colombia is administratively divided into 32 departments, which are divided into municipalities, departmental districts, or districts. There is an intermediate territorial division between department and municipality, called provinces or subregions [62]. The research was carried out in the city of Medellín, the capital of the department of Antioquia. Within Latin America, Medellín stands out for having been a pioneer in implementing initiatives favouring entrepreneurship and innovation. It has a business base that has been forged over several historical cycles, leading to the emergence of a set of clusters.

The data contain 632 records of young people between 14 and 28 years old. Of these, 251 are male, 378 are female, and three respond as other. Regarding age, 52.2% are between 14 and 18 years old, 35.0% are between 19 and 23 years old, and 12.8% are between 24 and 28 years old.

Concerning the level of education, 0.5% say they have completed primary school, 22.3% have a bachelor's degree or a bachelor's degree in progress, 17.7% have a technical degree or technical degree in progress, 17.1% are technologists or technologists in progress, 38.8% are undergraduates or undergraduates in progress, and 3.6% have a postgraduate or postgraduate degree in progress. Regarding the characteristics of the population group, 78.3% say they do not belong to any differential group, 13.4% report being victims of the conflict, 4.9% belong to the Afro-descendant community, 1.6% have a disability, 1.1% are migrants, and 0.6% are indigenous.

Colombia has a sui generis system of classification of residential properties that arises as an approximation to the economic capacity of its residents [63]. The socioeconomic strata into which dwellings and properties can be classified are 6: 1. low–low 2. low 3. medium–low 4. medium 5. medium–high 6. high. Of these, strata 1, 2, and 3 correspond to low strata whose house users have fewer resources and who are beneficiaries of subsidies for public utilities. Strata 5 and 6 correspond to high strata whose house users have more significant economic resources and who must pay surcharges (contributions) on the value of public utilities. Stratum 4 is not a beneficiary of subsidies, nor does it have to pay surcharges; it pays precisely the value that the company defines as the cost of providing the service [64]. Table 2 shows that 47.3% belong to strata 1 and 2, 45.7% belong to strata 3 and 4, and 7% belong to strata 5 and 6.

Table 2. Social strata Colombia Percentage.

Social Strata Colombia	Percentage
1 and 2	47.3%
3 and 4	45.7%
5 and 6	7%

3.2. Variables

The results of the literature review indicate that there are several determinants that can influence entrepreneurship. For the development of this study, we used the variables considered in the data from an open data portal available online: http://medata.gov.co/dataset/emprendimiento-juvenil (accessed on 10 December 2022), results of the "Youth Entrepreneurship" survey, developed by the Colombian Youth Secretariat. This survey applies to young people in the city of Medellín and was carried out to determine the status of their enterprises to characterise the value chains in the youth economy based on the observation variables defined in Table 3.

Table 3. Observation Variables and Constructs.

Observational Variables	Constructs
Age	How old are you?
Commune	In which commune or corregimiento do you live?
Stratum	Stratum
Gender	What is your gender?
Population Group	Do you belong to any of the following population groups?
Educational level	What is the level of education you have achieved so far?
Income	I decided to become an entrepreneur because: (I needed to find an alternative way to generate income and support myself).
Livelihood	I decided to become an entrepreneur because: (I needed to find an alternative for income generation and livelihood for my family).
Talent	I decided to become an entrepreneur because: (I discovered a talent/skill in me that I could capitalise on through entrepreneurship).
Problem	I decided to become an entrepreneur because: (I identified a need/problem in my community to which I could provide an answer through my entrepreneurship).
Passion	I decided to become an entrepreneur because: (I am passionate about finding business opportunities and generating income from this activity).
Research	I decided to become an entrepreneur because: (My project was born as part of a process of research and academic training).
Learning	With 0 being the lowest performance and 7 being the highest performance level. (Active and strategic learning).
Creativity	With 0 being the lowest performance and 7 being the highest performance level. (Creativity, originality, and initiative).
High Thinking	With 0 being the lowest performance and 7 being the highest performance level. (Critical thinking).
Problem Solving	With 0 being the lowest performance and 7 being the highest performance level. (Problem Solving).
Leadership	With 0 being the lowest performance and 7 being the highest performance level. (Leadership)
Time Management	With 0 being the lowest performance and 7 being the highest performance level. (Time management and coordination).
Emotional Intelligence	With 0 being the lowest performance and 7 being the highest performance level. (Emotional intelligence).
Technological Design	With 0 being the lowest performance and 7 being the highest performance level. (Technological design and programming).
Manual Skills	With 0 being the lowest performance and 7 being the highest performance level. (Manual dexterity, endurance, and precision).
Management in Finance	With 0 being the lowest performance and 7 being the highest performance level. (Management of finances).
Analytical Thinking	With 0 being the lowest performance and 7 being the highest performance level. (Analytical thinking]
Visual Skills	With 0 being the lowest performance and 7 being the highest performance level. (Visual and discursive skills).
Starting Year	Year of starting activity
First Entrepreneurship	Is this your first time as an entrepreneur?

3.3. Data Collection

The Google Colaboratory/Jupyter notebook tool was selected to run Python code to process the data obtained from the government portal, which is available online: http://medata.gov.co/dataset/emprendimiento-juvenil (accessed on 10 December 2022). Table 4 shows the data information using the pandas library to process the 632 rows and 19 columns that make up the dataset.

Table 4. General information of the dataset.

#	Column	Non-Null Count	Dtype
0	Age	632	Int64
1	Gender	632	object
2	Income	632	Int64
3	Livelihood	632	Int64
4	Talent	632	Int64
5	Problem	632	Int64
6	Passion	632	Int64
7	Investigative	632	Int64
8	Active_Learning	632	Int64
9	Creativity_Originality	632	Int64
10	High Thinking	632	Int64
11	Problem Solving	632	Int64
12	Leadership	632	Int64
13	Time Management	632	Int64
14	Emotional Intelligence	632	Int64
15	Technological Design	632	Int64
16	Manual Skills	632	Int64
17	Financial Management	632	Int64
18	Analytical Thinking	632	Int64
19	Visual Skills	632	Int64

In the case of the age variable, the set of desired ranges was constructed as 3 ranks, (14–18), (19–23), and (24–28), to express the possible categories. In machine learning techniques with Python, it is usually used to create dummies for categorical variables to represent them as numbers 0 or 1. Then, we can determine binary categorical variables, defining only two states or two categories. In our case, three binary categorical variables were set.

3.4. Data Processing for Analysis

Clustering allows a dataset to be divided into groups so that elements of each cluster are comparatively more like the elements within it compared to those of any other cluster [49]. Clustering aims to organise a similar dataset significantly different from the other clusters. It is worth noting that K-means and hierarchical clustering techniques are the most used clustering algorithms in the literature [50].

The application of clustering requires the calculation of a K-value to classify objects or cases into relatively homogeneous groups [65]. As this information is often unavailable, researchers usually perform trial and error exercises [65]. However, the elbow method can determine the optimal K-value or clusters. Figure 1 shows the results of the elbow diagram, plotted using the sum of squares within the cluster for a range of K-values. The optimal number of groups (K-value) is determined by choosing the "elbow" value of K, i.e., the point at which the WCSS starts to decrease linearly [53]. Therefore, we assume the number of clusters is 5 for the dataset. Figure 1 shows the results of the K-means clustering algorithm.

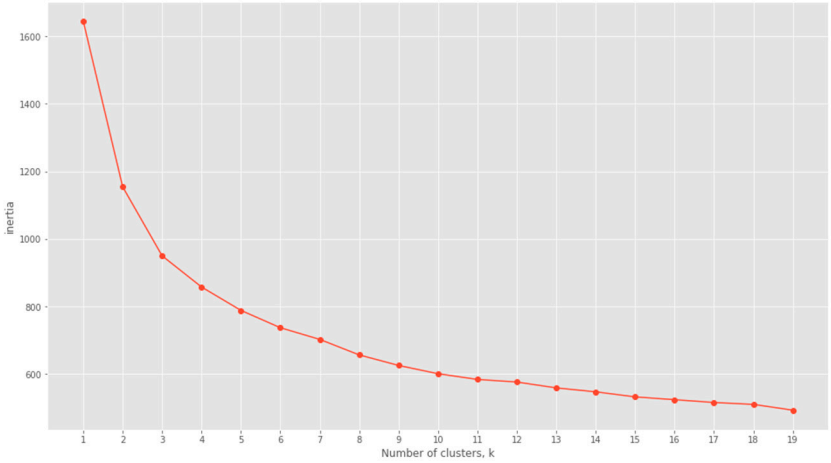

Figure 1. Result of the K-means clustering algorithm (elbow diagram). The red line shows the shape of the elbow to define the number of clusters.

3.5. To Validate the Clustering Model, the following Metrics Can Be Applied

Inertia: It calculates the sum of all points within a cluster from the centroid of that cluster; for all clusters, the final inertia value is the sum of all these distances. The distance within clusters is known as the intracluster distance. The following equation was used for the calculation of inertia:

$$\sum_{i=1}^{N}(x_i - C_k)^2 \quad (1)$$

N = the total number of points in the cluster.
x_i = a specific point within the cluster.
C_k = the centroid of the cluster for the point x_i.

The formula calculates the squared Euclidean distance between each point x_i and the centroid C_k of the cluster to which it belongs. These distances are then squared for all points within the cluster.

Silhouette: The silhouette coefficient (see Equation (2)) indicates the quality of the clustering in the data and is calculated using the following equation:

$$s(x) = \frac{b(x) - a(x)}{max[a(x), b(x)]} \quad (2)$$

where $a(x)$, called cohesion, is the average distance from x to all other points in the same cluster, and $b(x)$, called separation, is the average distance of x to all other points in the nearest cluster.

In the case of the analysed data, we find that all of them exceed the minimum values allowed within the silhouettes, as illustrated in Figure 2. In Figure 2, we show the values of the error plot for the clusters. This internal measure is mainly used to evaluate the cohesion of the clusters that the clustering algorithm generated, where k is the number of clusters, x is a point of the cluster C_i, and m_i is the centroid of the cluster C_i (see Equation (3)).

Equation (3) is the mathematical formula for calculating the sum of squares within (*SSE*) error in a K-means clustering model. In the equation, i represents the number of clusters, k is the total number of clusters, x represents an observation or point in cluster C_i, and m_i represents the centroid of cluster C_i. For all clusters, the *SSE* is calculated as the sum of the squared distances from each point x in cluster C_i to its centroid m_i. The *SSE* is a measure of cluster cohesion, i.e., how close the points of a cluster are to each other. Generally, one seeks to minimise the *SSE* to obtain more compact and homogeneous clusters.

$$SSE = \sum_{i=1}^{k}\sum_{x \in C_i} dist^2(m_i, x) \quad (3)$$

k = the total number of clusters.
i = the number of clusters under consideration.
x = an observation or point within the cluster C_i.
m_i = the centroid of the cluster C_i.
$dist(m_i, x)$ = the distance between the point x and the centroid m_i of cluster C_i.

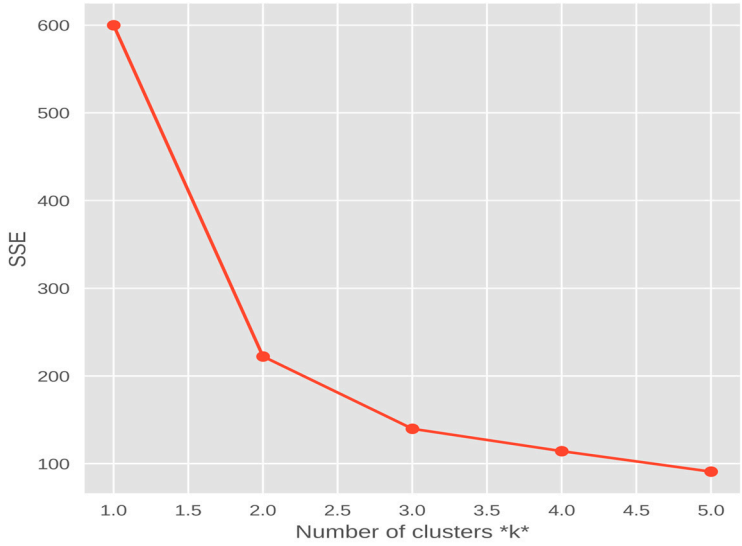

Figure 2. SSE error about. The red line shows the shape of the elbow to define the number of clusters.

The formula calculates the sum of the squared distances from each point x in the cluster C_i to its centroid m_i. These distances are then squared for all points within each cluster. The result SSE (sum squared error) is an accuracy measure where the errors are squared and then added, which quantifies the error of the dispersion within the clusters.

Figure 2 shows the results of applying the equation for each of the clusters as follows, ranging from cluster 1 to 5, respectively: [600.0000000000001, 222.36170496502302, 139.82049635974982, 114.30449485021941, and 90.92751382392049].

This is conducted through a trial-and-error exercise in which updating the centroids involves modifying the position of the centroid of each group by replacing it with the average of the positions of the objects belonging to that group. This process is repeated iteratively until the centroids either stop moving or their movement is less than a threshold distance at each iteration.

4. Results and Discussion

The number of clusters formed corresponds to 5. The clusters are composed of men and women of different ages, motives for entrepreneurship, and skills. The results obtained identify sociodemographic factors that affect entrepreneurship, such as the level of education attained by the parents, marital status, number of dependents, and socioeconomic stratum, as defined in Arias Vargas et al. [66]. Therefore, universities must stimulate entrepreneurial thinking by developing curricula and mechanisms that foster entrepreneurship [67]. The results of this clustering are shown in detail in Figure 3.

The results obtained from the application of the K-means algorithm are detailed in Table 5. It is constructed from the cluster assignment to each record in the dataset.

The results show that entrepreneurs are people who take action and engage in vigorous and persistent efforts to turn their ideas and visions into profitable operating enterprises, and they identify and explain the variables that characterise the level of entrepreneurial activity (entrepreneurs, businesses, and environment), which serves to set the tone for public policy proposals that increase the quality and quantity of entrepreneurial activity [3]. Under this approach, the K-means clustering method is used to identify the critical determinants of youth entrepreneurship in Medellín. The clustering results show that creativity and

learning ability are the essential characteristics influencing entrepreneurial development, as highlighted in [68].

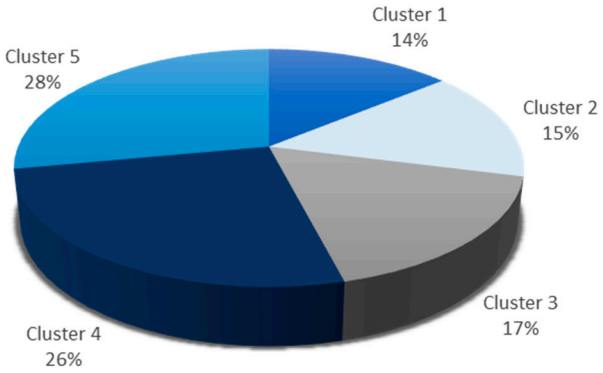

Figure 3. Amount of data in each cluster.

Table 5. Clustering results based on the K-means algorithm.

Cluster	Gender	Age	Motivation for Entrepreneurship	Skills
Cluster 1	Men Women	(14–17)	Livelihood Passion	Creativity High Thinking Emotional Intelligence Manual Skills Visual Skills
Cluster 2	Men Women	(17–20)	Problematic Talent Livelihood Income Passion	Visual Skills Analytical Thinking Financial Management High Thinking Leadership Investigative Problem Solving Creativity Visual Skills Time Management Learning
Cluster 3	Men Women	(20–23)	Research Issues Passion	Visual Skills Analytical Thinking Financial Management High Thinking Leadership Investigative Problem Solving Leadership Time Management Learning Technology Design Learning

Table 5. *Cont.*

Cluster	Gender	Age	Motivation for Entrepreneurship	Skills
Cluster 4	Men Women	(23–26)	Problematic Support Income	Financial Management High Thinking Leadership Investigative Problem Solving Creativity Learning
Cluster 5	Men Women	(26–29)	Problematic Support Income	Financial Management High Thinking Leadership Investigative Problem Solving Leadership Time Management Learning

In our study, Table 6 shows the results of the silhouette coefficient values by clusters. The value of s(x) can vary between −1 and 1, where −1 is poor clustering, 0 is indifferent, and 1 is good.

Table 6. Silhouette coefficient value by clusters.

Clusters	Silhouette Coefficient Value
Cluster 1	0.45833333
Cluster 2	0.5
Cluster 3	0.05484746
Cluster 4	0.05779661
Cluster 5	0.66666667

The accepted clusters have an average value of the silhouette of 0.34, an acceptable value in the grouping of the data given that the value is between −1 to 1 and the closeness to 1 indicates a better grouping of data in a cluster and vice versa.

Creativity is identified in relation to the skills associated with cluster 1, which significantly influences different family scenarios by fostering entrepreneurial imagination [68]. High thinking enables intellectual development to develop new habits, recognise inconsistencies, develop intellectual empathy to understand others, and confront ideas; this skill is encouraged by business education [69]. This should then be fostered by the different pedagogical initiatives proposed in universities, which include collaborative work and managerial simulations that foster the competencies of entrepreneurs [70,71]. Regarding emotional intelligence, there is growing evidence of the importance of identifying and managing emotions to understand entrepreneurial behaviour; people with high emotional intelligence show more capacity and higher levels of self-efficacy and confidence to take on these challenges [72].

Cluster 2 highlights skills such as leadership as positively impacting entrepreneurship, and it can therefore be considered an essential factor in improving productivity and effectiveness [73]. It also identifies real world-based problem-solving skills that boost students' creativity in entrepreneurship's novelty, utility, aesthetics, and authenticity dimensions [74].

Cluster 3 identifies skills related to technological design, which undoubtedly associates research and entrepreneurship from a design science approach that brings considerable benefits to the field [75]. Since design thinking is a human-centred innovation process that emphasises observation, collaboration, rapid learning, visualisation of ideas, rapid

prototyping of concepts, and simultaneous business analysis, it influences entrepreneurship strategy [76].

Cluster 4 highlights skills such as learning and financial management, while cluster 5 associates efficient time management skills, enabling improved productivity, successful time organisation, and achieving short- and long-term goals [77].

From these descriptive variables, it is identified that age is a determining factor in developing new entrepreneurs because young people play a new force in innovation and entrepreneurship [78]. Likewise, objective, and straightforward strategies that enable the implementation of entrepreneurship education and the development of unique skills, such as creativity, leadership, and learning, could contribute to a better-prepared and adapted society. For their part, government interventions can help young people overcome obstacles to productive entrepreneurship [79,80].

The clusters generated show correlations with profiles that can support the design of policies in Colombia; policies affecting entrepreneurship can have a significant impact on the entrepreneurial ecosystem and the creation and development of new businesses through business regulation and improved access to finance; education policies can promote training in entrepreneurial skills and foster entrepreneurship; and intellectual property protection policies can be important to promote innovation and the creation of knowledge-based businesses. The results of this research are intended to inform policy development with possible implications for (a) the design of support policies by providing valuable information on the characteristics, skills, and needs of entrepreneurs; (b) the promotion of innovation and creativity by showing that certain entrepreneurial profiles are associated with these characteristics; (c) the promotion of science, technology, engineering, and mathematics (STEM) education and entrepreneurial environments that encourage experimentation and collaboration. This could include promoting science, technology, engineering, and mathematics (STEM) education and fostering entrepreneurial environments that encourage experimentation and collaboration; (d) Inclusion and equity policies: Research on entrepreneurial profiles can shed light on existing disparities in entrepreneurship in terms of gender, race, educational attainment, or other factors and lead to policies aimed at promoting inclusion and equity in entrepreneurship, removing barriers, and providing equal opportunities for all demographic groups; (e) Resource allocation: Understanding the profiles of entrepreneurs can help policy makers to allocate available resources more effectively to support entrepreneurship. For example, if research shows that certain profiles are more likely to succeed, governments can target their investments towards these specific groups; and (f) Regulatory framework: The results of research on entrepreneurial profiles can influence the formulation of regulatory policies related to entrepreneurship [81].

The results obtained in the background review allow us to identify from different perspectives the existence of a gap between the entrepreneurial capacity of men and women [82,83]. The most significant barriers to female entrepreneurship identified by [84,85] in the literature include low levels of entrepreneurial human capital, limited networks, low legitimacy in entrepreneurship, limited ecosystems, and access to capital. This suggests the need for policies to support women's entrepreneurship [86]. Traditionally and historically, the figure of the entrepreneur has been equated with that of men and, therefore, with their skills [87–89].

5. Conclusions

Applying the K-means clustering technique gives a deeper and more detailed understanding of the characteristics and needs of young entrepreneurs in Medellín. The detailed analysis of the determining characteristics of the dataset has allowed us to identify risk factors and establish variables, such as age, as determining factors in the development of new entrepreneurs. Furthermore, it has been observed that gender is indifferent as a determinant factor, indicating that the general discourse that women are not entrepreneurs is invalid. The main motives for entrepreneurship are income generation, livelihood, and

the solution to problems identified by young entrepreneurs. More excellent value has been placed on creativity, high thinking, and emotional intelligence.

The importance of entrepreneurship education and government interventions to help young people overcome barriers to productive entrepreneurship has also been highlighted. The research also highlights the relevance of technological design and design thinking as innovation processes that can influence entrepreneurship strategies. It is essential to mention that entrepreneurship is an economic driver that can contribute to solving social and environmental problems. In this sense, promoting social and sustainable entrepreneurship is an urgent need today.

In summary, it is concluded that entrepreneurship is a challenging path. However, with the proper education, valuable skills, and supportive policies, achieving entrepreneurial success and contributing to developing a more just and sustainable society is possible.

In Colombia, policies affecting entrepreneurship can have a significant impact on the entrepreneurial ecosystem and the creation and development of new firms, through business regulation and improved access to finance. Education policies can promote training in entrepreneurial skills and foster entrepreneurship. Intellectual property protection policies can be important to promote innovation and the creation of knowledge-based firms.

6. Limitations of the Research and Future Lines of Research

The main limitations identified in this study are related to sample size; future lines of research on youth entrepreneurship in Colombia using machine learning models can address issues related to sample expansion, representativeness, comparisons, geographical diversity, and long-term impact assessment.

Author Contributions: Conceptualization, A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; methodology, A.O.-B., A.S.-B., D.C.-P., A.C.-P. and D.D.O.; validation, A.O.-B., A.S.-B., A.C.-P., J.A.H.-T., M.D. and W.A.-H.; and A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; formal analysis, A.O.-B., A.S.-B., A.C.-P., J.A.H.-T., G.T.-V., M.D. and W.A.-H.; investigation, A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; data curation, A.O.-B., A.S.-B., D.D.O., A.C.-P., J.A.H.-T. and W.A.-H.; writing—original draft preparation, A.O.-B., A.S.-B., D.D.O., G.T.-V., A.C.-P. and W.A.-H.; writing—review and editing, A.O.-B., A.S.-B., D.C.-P., M.D., A.C.-P. and W.A.-H.; visualization, A.O.-B., A.S.-B., D.C.-P., A.C.-P., G.T.-V. and W.A.-H., J.A.H.-T.; supervision, A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; project administration, A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; resources, A.O.-B., A.S.-B., A.C.-P. and W.A.-H.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received support from the AUIP (Iberoamerican University Association for Postgraduate Studies).

Data Availability Statement: The data is available in http://medata.gov.co/dataset/emprendimiento-juvenil (accessed on 10 December 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lederman, D.; Messina, J.; Pienknagura, S.; Rigolini, J. El Emprendimiento en América Latina: Muchas Empresas y Poca Innovación. World Bank. 2014, pp. 61–95. Available online: https://www.worldbank.org/content/dam/Worldbank/document/LAC/EmprendimientoAmericaLatina_resumen.pdf (accessed on 15 April 2023).
2. Innpulsa Colombia. (s.f.). Boletín Analítica. Available online: https://www.innpulsacolombia.com/sites/default/files/documentos-recursos-pdf/Boletin%20Analitica.pdf (accessed on 15 April 2023).
3. Pereira Laverde, F.; Osorio-Tinoco, F.; Pinzón, N. *Nuestro Reto: Impactar la DINAMICA Emprendedora Colombiana: GEM Colombia 2021–2022*; Ediciones: Bogotá, Colombia, 2022.
4. Velasco Chávez, R.; Ordoñez Arias, C.; Marion, R.S.; Juan, C.E. Ley de Emprendimiento en Colombia. In Innpulsa, Colombia. Publications. 2021. Available online: https://www.innpulsacolombia.com/sites/default/files/documentos-recursos-pdf/Boletin%20Analitica.pdf (accessed on 15 April 2023).
5. Schade, P.; Schuhmacher, M.C. Predicting entrepreneurial activity using machine learning. *J. Bus. Ventur. Insights* **2023**, *19*, e00357. [CrossRef]
6. Buitrago Nova, J.A. Emprendimiento en Colombia. *Adm. Desarro.* **2014**, *43*, 7–21. [CrossRef]

7. Berger, E.S.; Von Briel, F.; Davidsson, P.; Kuckertz, A. Digital or not—The future of entrepreneurship and innovation: Introduction to the special issue. *J. Bus. Res.* **2021**, *125*, 436–442. [CrossRef]
8. Cruz, P.D.A.; Covarrubias, L.P.C.; Pérez, P.A. The Effect of Resilience on Entrepreneurial Intention in Higher Education Students in a Post-Covid Stage 19. *J. High. Educ. Theory Pract.* **2022**, *22*, 6–20. [CrossRef]
9. Bosse, D.A.; Harrison, J.S.; Pollack, J.M.; Schrempf-Stirling, J. Entrepreneurial Opportunities as Responsibility. *Entrep. Theory Pract.* **2022**, *47*, 3–16. [CrossRef]
10. Çera, G.; Ndoka, M.; Dika, I.; Çera, E. Examining the Impact of COVID-19 on Entrepreneurial Intention through a Stimulus–Organism–Response Perspective. *Adm. Sci.* **2022**, *12*, 184. [CrossRef]
11. Shane, S.; Venkataraman, S. The Promise of Entrepreneurship as a Field of Research. *Acad. Manag. Rev.* **2000**, *25*, 217–226. [CrossRef]
12. Van Gelderen, M.; Kautonen, T.; Fink, M. From entrepreneurial intentions to actions: Self-control and action-related doubt, fear, and aversion. *J. Bus. Ventur.* **2018**, *33*, 780–794. [CrossRef]
13. Lumpkin, G.T.; Dess, G.G. Clarifying the entrepreneurial orientation construct and linking it to performance. *Acad. Manag. Rev.* **1996**, *21*, 135–172. [CrossRef]
14. Baron, R.A.; Shane, S.A. Entrepreneurship: A process perspective. In *The Psychology of Entrepreneurship*, 2nd ed.; Psychology Press: Mason, OH, USA, 2007.
15. Segal, S.; Mikulincer, M.; Hershkovitz, L.; Meir, Y.; Nagar, T.; Maaravi, Y. A Secure Base for Entrepreneurship: Attachment Orientations and Entrepreneurial Tendencies. *Behav. Sci.* **2023**, *13*, 61. [CrossRef]
16. Elkayaly, D.; Fahim, I.S. Challenges and Motivations for Youth Entrepreneurship Start-Ups: Empirical study from Egypt. In Proceedings of the 2021 3rd International Sustainability and Resilience Conference: Climate Change, Virtual, 15–17 November 2021; pp. 339–341. [CrossRef]
17. Postigo, Á.; Cuesta, M.; García-Cueto, E.; Prieto-Díez, F.; Muñiz, J. General versus specific personality traits for predicting entrepreneurship. *Pers. Individ. Differ.* **2021**, *182*, 111094. [CrossRef]
18. Lien, T.T.H.; Hoang, N.D. The impact of young employees' perceptions of current paid jobs on the entrepreneurial intention with the mediator of job satisfaction: The case of Vietnam. *Entrep. Bus. Econ. Rev.* **2022**, *10*, 55–71. [CrossRef]
19. Onjewu, A.K.E.; Haddoud, M.Y.; Nowiński, W. The effect of entrepreneurship education on nascent entrepreneurship. *Ind. High. Educ.* **2021**, *35*, 419–431. [CrossRef]
20. Wu, L.; Jiang, S.; Wang, X.; Yu, L.; Wang, Y.; Pan, H. Entrepreneurship Education and Entrepreneurial Intentions of College Students: The Mediating Role of Entrepreneurial Self-Efficacy and the Moderating Role of Entrepreneurial Competition Experience. *Front. Psychol.* **2022**, *12*, 727826. [CrossRef]
21. Lyu, J.; Shepherd, D.; Lee, K. From intentional to nascent student entrepreneurs: The moderating role of university entrepreneurial offerings. *J. Innov. Knowl.* **2023**, *8*, 100305. [CrossRef]
22. Ndofirepi, T.M. Predicting the Sustainability-Oriented Entrepreneurship Intentions of Business School Students: The Role of Individualistic Values. *Soc. Sci.* **2023**, *12*, 13. [CrossRef]
23. Thomas, E.; Pugh, R.; Soetanto, D.; Jack, S.L. Beyond ambidexterity: Universities and their changing roles in driving regional development in challenging times. *J. Technol. Transf.* **2023**, 1–20. [CrossRef]
24. Liu, M.; Gorgievski, M.J.; Zwaga, J.; Paas, F. How entrepreneurship program characteristics foster students' study engagement and entrepreneurial career intentions: A longitudinal study. *Learn. Individ. Differ.* **2023**, *101*, 102249. [CrossRef]
25. Lechner, C.; Delanoë-Gueguen, S.; Gueguen, G. Entrepreneurial ecosystems and actor legitimacy. *Int. J. Entrep. Behav. Res.* **2022**. *ahead-of-print*. [CrossRef]
26. Banha, F.; Coelho, L.S.; Flores, A. Entrepreneurship Education: A Systematic Literature Review and Identification of an Existing Gap in the Field. *Educ. Sci.* **2022**, *12*, 336. [CrossRef]
27. Riaño, Y. Migrant Entrepreneurs as Agents of Development? Geopolitical Context and Transmobility Strategies of Colombian Migrants Returning from Venezuela. *J. Int. Migr. Integr.* **2022**, *24*, 539–562. [CrossRef] [PubMed]
28. Zulkifle, A.M.; Aziz, K.A. Determinants of Social Entrepreneurship Intention: A Longitudinal Study among Youth in Higher Learning Institutions. *Soc. Sci.* **2023**, *12*, 124. [CrossRef]
29. Rusu, V.D.; Roman, A.; Tudose, M.B. Determinants of Entrepreneurial Intentions of Youth: The Role of Access to Finance. *Eng. Econ.* **2022**, *33*, 86–102. [CrossRef]
30. Soomro, B.A.; Shah, N. Entrepreneurship education, entrepreneurial self-efficacy, need for achievement and entrepreneurial intention among commerce students in Pakistan. *Educ.+ Train.* **2022**, *64*, 107–125. [CrossRef]
31. Fiet, J.O. The pedagogical side of entrepreneurship theory. *J. Small Bus. Manag.* **2017**, *55*, 126–141.
32. Neck, H.M.; Greene, P.G.; Brush, C.G. *Teaching Entrepreneurship: A Practice-Based Approach*; Edward Elgar Publishing: Cheltenham, UK, 2019.
33. Ramírez, K.; Serrano, L.; Pineda, J. Categorization the entrepreneur's characteristics from the perspective of the person. In Proceedings of the 2017 Congreso Internacional de Innovacion y Tendencias en Ingenieria (CONIITI), Bogotá, Colombia, 4–6 October 2017.
34. Chamorro, E.T.; Erazo, J.A.Z.; Montenegro, E.A. Knight y sus aportes a la teoría del emprendedor. In *Estudios Gerenciales*; Elsevier: Amsterdam, The Netherlands, 2008; Volume 24, pp. 83–98.

35. Aparicio, S.; Audretsch, D.; Urbano, D. Why is export-oriented entrepreneurship more prevalent in some countries than others? Contextual antecedents and economic consequences. *J. World Bus.* **2021**, *56*, 101177. [CrossRef]
36. Gutiérrez, J.G.; Baquero, J.E.G. New cross-proposal entrepreneurship and innovation in educational programs in third level (tertiary) education. *Contaduría Y Adm.* **2017**, *62*, 239–261. [CrossRef]
37. Teixeira, R.; Ducci, N.; Sarrassini, N.; Munhê, V.; Ducci, L. Empreendedorismo jovem e a influência da família: A história de vida de uma empreendedora de sucesso. *Rev. Gestão* **2011**, *18*, 3–18. [CrossRef]
38. Guerrero, M.; Espinoza-Benavides, J. Does entrepreneurship ecosystem influence business re-entries after failure? *Int. Entrep. Manag. J.* **2021**, *17*, 211–227. [CrossRef]
39. Arnold, G. Does Entrepreneurship Work? Understanding What Policy Entrepreneurs Do and Whether It Matters. *Policy Stud. J.* **2021**, *49*, 968–991. [CrossRef]
40. Aparicio, S.; Urbano, D.; Gómez, D. The role of innovative entrepreneurship within Colombian business cycle scenarios: A system dynamics approach. *Futures* **2016**, *81*, 130–147. [CrossRef]
41. Kebede, G.F. Entrepreneurship and the Promises of Inclusive Urban Development in Ethiopia. *Urban Forum* **2022**, *34*, 1–30. [CrossRef]
42. Pascucci, T.; Sanchez, B.H.; Carlos, J.; Garcia, S. Video Games Entrepreneurship: A Confrontation between Strong Multinational and Indie Software House. *Int. J. Adv. Technol.* **2023**, *14*, 1000229.
43. Chatti, H.; Hamrouni, A.D. Identifying success factors and difficulties in setting up business opportunities in the context of collective entrepreneurship. *Acad. Entrep. J.* **2021**, *27*, 1–15.
44. Sendra-Pons, P.; Comeig, I.; Mas-Tur, A. Institutional factors affecting entrepreneurship: A QCA analysis. *Eur. Res. Manag. Bus. Econ.* **2022**, *28*, 100187. [CrossRef]
45. Shepherd, D.A.; Majchrzak, A. Machines augmenting entrepreneurs: Opportunities (and threats) at the Nexus of artificial intelligence and entrepreneurship. *J. Bus. Ventur.* **2022**, *37*, 106227. [CrossRef]
46. Jan, Z.; Ahamed, F.; Mayer, W.; Patel, N.; Grossmann, G.; Stumptner, M.; Kuusk, A. Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities. *Expert Syst. Appl.* **2023**, *216*, 119456. [CrossRef]
47. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
48. Pinto, J.F.; Cabral, M. Statistical Methods with Applications in Data Mining: A Review of the Most Recent Works. *Mathematics* **2022**, *10*, 1–22.
49. Virupakshappa, K.; Oruklu, E. Unsupervised Machine Learning for Ultrasonic Flaw Detection using Gaussian Mixture Modeling, K-Means Clustering and Mean Shift Clustering. In Proceedings of the 2019 IEEE International Ultrasonics Symposium (IUS), Glasgow, UK, 6–9 October 2019; pp. 647–649. [CrossRef]
50. Ezugwu, A.E.; Ikotun, A.M.; Oyelade, O.O.; Abualigah, L.; Agushaka, J.O.; Eke, C.I.; Akinyelu, A.A. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Eng. Appl. Artif. Intell.* **2022**, *110*, 104743. [CrossRef]
51. Celebi, M.E.; Kingravi, H.A. Linear, Deterministic, and Order-Invariant Initialization Methods for the K-Means Clustering Algorithm. In *Partitional Clustering Algorithms*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 79–98. [CrossRef]
52. Hoda, M.N.; Chauhan, N.; Quadri, S.M.K.; Srivastava, P.R. *Software Engineering: Proceedings of CSI 2015*; Springer Nature: Singapore, 2019.
53. Pham, T.T.; Lobos, G.A. Innovación en Minería de Datos para el Tratamiento de Imágenes: Agrupamiento K-media para Conjuntos de Datos de Forma Alargada y su Aplicación en la Agroindustria—Innovation in Data Mining for the Image Processing: K-means Clustering for Data Sets of El. *Inf. Tecnol.* **2019**, *30*, 135–142. [CrossRef]
54. Goverment of Colombia. Ley 2069 de 2020 Por Medio del Cual se Impulsa el Emprendimiento en Colombia. In Dapre.presidencia.gov.co. Available online: https://dapre.presidencia.gov.co/normativa/normativa/LEY-2069-del-31-dediciembre-de-2020.pdf (accessed on 20 May 2023).
55. Arrubla-Hoyos, W.; Ojeda-Beltrán, A.; Solano-Barliza, A.; Rambauth-Ibarra, G.; Barrios-Ulloa, A.; Cama-Pinto, D.; Arrabal-Campos, F.M.; Martínez-Lao, J.A.; Cama-Pinto, A.; Manzano-Agugliaro, F. Precision Agriculture and Sensor Systems Applications in Colombia through 5G Networks. *Sensors* **2022**, *22*, 7295. [CrossRef] [PubMed]
56. Torres, M.M.; Ramos, R.J.L.; Galvis, M.M.M.; Ramos, C.J.L.; Biswell, J.J.E. Determinantes del emprendimiento juvenil en Colombia: Un análisis desde la nueva economía institucional. *Rev. Métodos Cuant. Econ. Empresa* **2021**, *32*, 300–323. [CrossRef]
57. Quillas, C.I.L.; Rodríguez, H.F.R.; Rodríguez, A.H. Instituciones informales, emprendimiento y progreso social: Un estudio comparativo y correlacional. *Rev. Guillermo Ockham* **2022**, *21*, 113–129. [CrossRef]
58. Lu, Q.; Chai, Y.; Ren, L.; Ren, P.; Zhou, J.; Lin, C. Research on quality evaluation of innovation and entrepreneurship education for college students based on random forest algorithm and logistic regression model. *PeerJ Comput. Sci.* **2023**, *9*, e1329. [CrossRef]
59. Graham, B.; Bonner, K. One size fit all? Using machine learning to study heterogeneity and dominance in the determinants of early-stage entrepreneurship. *J. Bus. Res.* **2022**, *152*, 42–59. [CrossRef]
60. Gosztonyi, M.; Judit, C.F. Profiling (Non-)Nascent Entrepreneurs in Hungary Based on Machine Learning Approaches. *Sustainability* **2022**, *14*, 3571. [CrossRef]

61. Zhang, L.; Fu, Y.; Wei, Y.; Chen, H.; Xia, C.; Cai, Z. Predicting Entrepreneurial Intention of Students: Kernel Extreme Learning Machine with Boosted Crow Search Algorithm. *Appl. Sci.* **2022**, *12*, 6907. [CrossRef]
62. Ramírez, J.C.; de Aguas, J.M. *Configuración Territorial de las Provincias de Colombia Ruralidad y Redes*; Comisión Económica para América Latina y el Caribe (CEPAL): Bogotá, Colombia, 2016.
63. Marcela, D.; García, A.; García, A.; Marcela, D. Crecer en sociedades desiguales: Interpretaciones de niñas y niños colombianos sobre la estratificación socioeconómica. *Tempo E Argum.* **2022**, *14*, e0203.
64. Departamento Nacional de Estadística (DANE). Mesa de Expertos de Estratificación Socioeconómica (Informe Final). 2021. Available online: https://www.dnp.gov.co/programas/vivienda-agua-y-desarrollo-urbano/Paginas/Mesa-de-expertos-de-Estatificacion-Socioeconomica.aspx (accessed on 20 June 2023).
65. Morissette, L.; Chartier, S. The k-means clustering technique: General considerations and implementation in Mathematica. *Tutor. Quant. Methods Psychol.* **2013**, *9*, 15–24. [CrossRef]
66. Arias Vargas, F.J.; Ribes-Giner, G.; Arango-Botero, D.; Garcés Giraldo, L.F. Factores sociodemográficos que inciden en el emprendimiento rural de jóvenes en Antioquia, Colombia. *Rev. Venez. Gerenc.* **2021**, *26*, 1218–1240. [CrossRef]
67. Svotwa, T.D.; Jaiyeoba, O.; Roberts-Lombard, M.; Makanyeza, C. Perceived Access to Finance, Entrepreneurial Self-Efficacy, Attitude Toward Entrepreneurship, Entrepreneurial Ability, and Entrepreneurial Intentions: A Botswana Youth Perspective. *SAGE Open* **2022**, *12*, 21582440221096437. [CrossRef]
68. Meng, X.; Huang, W. Analysis of Key Factors Affecting Undergraduate Entrepreneurship Ability from a Big Data Perspective. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 2198948. [CrossRef]
69. Zhan, Z.; Fong, P.S.W.; Lin, K.-Y.; Zhong, B.; Yang, H.H. Editorial: Creativity, innovation, and entrepreneurship: The learning science toward higher order abilities. *Front. Psychol.* **2022**, *13*, 1063370. [CrossRef] [PubMed]
70. Ojeda, A.D.; Solano-Barliza, A.D.; Ortega, D.D.; Cañavera, A.M. Análisis cuantitativo de un proceso de enseñanza soportado en una estrategia pedagógica de gamificación. *Form. Univ.* **2022**, *15*, 83–92. [CrossRef]
71. Vargas, J.D.; Arregocés, I.C.; Solano, A.D.; Peña, K.K. Aprendizaje basado en proyectos soportado en un diseño tecno-pedagógico para la enseñanza de la estadística descriptiva. *Form. Univ.* **2021**, *14*, 77–86. [CrossRef]
72. Solano, A.D.; Aarón, M.A. Enseñanza en ingeniería de manera colaborativa a partir de un diseño tecnopedagógico, usando SMILE. *Form. Univ.* **2020**, *13*, 201–210. [CrossRef]
73. López-Núñez, M.I.; Rubio-Valdehita, S.; Díaz-Ramiro, E.M. The role of individual variables as antecedents of entrepreneurship processes: Emotional intelligence and self-efficacy. *Front. Psychol.* **2022**, *13*, 978313. [CrossRef]
74. Thuy, D.T.T.; Viet, T.Q.; Van Phuc, V.; Pham, T.-H.; Lan, N.T.N.; Ho, H. Impact of Leadership Behavior on Entrepreneurship in State-Owned Enterprises: Evidence from Civil Servant Management Aimed at Improving Accountability. *Economies* **2022**, *10*, 245. [CrossRef]
75. Weng, X.; Chiu, T.K.; Tsang, C.C. Promoting student creativity and entrepreneurship through real-world problem-based maker education. *Think. Ski. Creat.* **2022**, *45*, 101046. [CrossRef]
76. Seckler, C.; Mauer, R.; Brocke, J.V. Design science in entrepreneurship: Conceptual foundations and guiding principles. *J. Bus. Ventur. Des.* **2021**, *1*, 100004. [CrossRef]
77. Chou, D.C. Applying design thinking method to social entrepreneurship project. *Comput. Stand. Interfaces* **2018**, *55*, 1339–1351. [CrossRef]
78. Dmytryshyn, M.; Goran, T. Proposal of an Effective Time Management System. *Management* **2022**, *27*, 283–298. [CrossRef]
79. Zhiguo, L.; Yumei, W.; Shaoqin, L.; Beibei, L. The Efficiency Evaluation Study of "Creative Space" for Promoting the Development of Youth Innovation and Entrepreneurship Projects. In Proceedings of the 31st Chinese Control and Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 3797–3802. [CrossRef]
80. Brixiová, Z.; Ncube, M.; Bicaba, Z. Skills and Youth Entrepreneurship in Africa: Analysis with Evidence from Swaziland. *World Dev.* **2015**, *67*, 11–26. [CrossRef]
81. Campo-Ternera, L.; Amar-Sepúlveda, P.; Olivero-Vega, E. Interaction of potential and effective entrepreneurial capabilities in adolescents: Modeling youth entrepreneurship structure using structural equation modeling. *J. Innov. Entrep.* **2022**, *11*, 13. [CrossRef]
82. Kaminskiene, L.; Horlenko, K.; Chu, L.Y. Applying Eye-Tracking Technology in the Field of Entrepreneurship Education. FGF Studies in Small Business and Entrepreneurship. In *Artificiality and Sustainability in Entrepreneurship*; Springer: Cham, Switzerland, 2023; pp. 163–187. [CrossRef]
83. Laouiti, R.; Haddoud, M.Y.; Nakara, W.A.; Onjewu, A.-K.E. A gender-based approach to the influence of personality traits on entrepreneurial intention. *J. Bus. Res.* **2022**, *142*, 819–829. [CrossRef]
84. Oladipo, O.; Platt, K.; Shim, H.S. Female entrepreneurs managing from home. *Small Bus. Econ.* **2023**, 1–18. [CrossRef]
85. Avnimelech, G.; Rechter, E. How and why accelerators enhance female entrepreneurship. *Res. Policy* **2023**, *52*, 104669. [CrossRef]
86. Georgieva, S. Women's entrepreneurship—Segmentation and management dimensions. *Pol. J. Manag. Stud.* **2022**, *26*, 144–161. [CrossRef]
87. Kille, T.; Wiesner, R.; Lee, S.Y.; Johnson Morgan, M.; Summers, J.; Davoodian, D. Capital Factors Influencing Rural, Regional and Remote Women's Entrepreneurship Development: An Australian Perspective. *Sustainability* **2022**, *14*, 16442. [CrossRef]

88. Manzanera-Román, S.; Brändle, G. Abilities and skills as factors explaining the differences in women entrepreneurship. *Suma Neg.* **2016**, *7*, 38–46. [CrossRef]
89. Gieure, C.; del Mar Benavides-Espinosa, M.; Roig-Dobón, S. The entrepreneurial process: The link between intentions and behavior. *J. Bus. Res.* **2020**, *112*, 541–548. [CrossRef]