Cruz-Eraso, Andrés-Felipe; González-Serrano, Carolina
Adaptive Model of Classification of Professions in Vocational Guidance Systems

Revista Facultad de Ingeniería
Journal Homepage: https://revistas.uptc.edu.co/index.php/ingenieria

Adaptive Model of Classification of Professions in Vocational Guidance Systems

Andrés-Felipe Cruz-Eraso[1]

Carolina González-Serrano[2]

[1] Universidad del Cauca (Popayán-Cauca, Colombia). andrescruz@unicauca.edu.co. ORCID: 0000-0002-1216-4876

[2] Ph. D. Universidad del Cauca (Popayán-Cauca, Colombia). cgonzals@unicauca.edu.co. ORCID: 0000-0002-8764-3177

Abstract

Vocational guidance is part of psychosocial development and is understood as a method that helps to determine the most appropriate profession according to the aptitudes and abilities of the student. The processes of vocational guidance are dynamic and focus on educating and favoring the decision-making process in the professional choice for a learning pathway throughout the student's life, which will benefit society in the long run. Most of the current solutions, both theoretical and applied, from Europe and North America differ when used in the Colombian context, mainly for adults, since the process of classifying professions is not accurate nor precise. In addition, there are various educational projects and evaluation systems in secondary education level institutions. At this level, the students have a changing vocational choice which implies taking into account specific characteristics of the context, also, the student profile vocational guidance determinants. The objective of this article is to describe the adaptive model of occupational classification integrated into the Intelligent Web Platform used in educational institutions in the Department of Cauca. The use of the CRISP-DM methodology allowed finding the Naive Bayes and Deep learning algorithms as those with the best performance in the classification of professions.

Keywords: adaptive intelligence; data mining: vocational guidance; web platform.

Modelo adaptativo de clasificación de profesiones en sistemas de orientación vocacional

Resumen

La orientación vocacional forma parte del desarrollo psicosocial y se entiende, como un método que ayuda a determinar la profesión más adecuada, en función de las aptitudes y capacidades de los estudiantes. Los procesos de orientación vocacional, son dinámicos y se enfocan en educar y favorecer el proceso de toma de decisiones en la elección profesional. proporcionándoles un camino de aprendizaje claro para seguir a lo largo de su vida y para que la sociedad se beneficie de su talento enfocado. La mayoría de soluciones actuales, teóricas y de aplicación, de Europa y

Norteamérica, difieren al utilizarse en el contexto colombiano, principalmente para adultos, las cuales en términos del proceso de clasificación de profesiones presentan bajos niveles de exactitud y precisión. Además, existen diversos proyectos educativos y sistemas de evaluación en los centros de enseñanza secundaria. En este nivel de educación, los estudiantes tienen una decisión vocacional cambiante que implica tener en cuenta características específicas del contexto, también, los determinantes su perfil de orientación vocacional. El objetivo del presente artículo es describir el modelo adaptativo de clasificación de profesiones, integrado a la Plataforma Web Inteligente, utilizada en instituciones educativas del Departamento del Cauca. El uso de la metodología CRISP-DM permitió encontrar los algoritmos de Naive Bayes y Deep learning como los de mejor desempeño en la clasificación de profesiones.

Palabras clave: inteligencia adaptativa; minería de datos; orientación vocacional; plataforma web.

Modelo adaptativo de classificação de profissões em sistemas de orientação profissional

Resumo

A orientação profissional faz parte do desenvolvimento psicossocial e é entendida como um método que ajuda a determinar a profissão mais adequada, com base nas aptidões e habilidades dos alunos. Os processos de orientação profissional são dinâmicos e têm como foco educar e favorecer a tomada de decisão na escolha profissional. proporcionando-lhes um caminho claro de aprendizado para seguir ao longo de sua vida e para que a sociedade se beneficie de seu talento focado. A maioria das soluções atuais, teóricas e aplicadas, da Europa e América do Norte, diferem quando utilizadas no contexto colombiano, principalmente para adultos, que em termos de processo de classificação de profissões apresentam baixos níveis de exatidão e precisão. Além disso, existem vários projetos educacionais e sistemas de avaliação nas escolas secundárias. Neste nível de ensino, os alunos têm uma decisão vocacional em mudança que implica ter em conta características específicas do contexto, bem como os determinantes do seu perfil de orientação

vocacional. O objetivo deste artigo é descrever o modelo adaptativo de classificação de profissões, integrado à Plataforma Web Inteligente, utilizado em instituições educacionais do Departamento de Cauca. A utilização da metodologia CRISP-DM permitiu encontrar os algoritmos Naive Bayes e Deep learning como os de melhor desempenho na classificação de profissões.

Palavras-chave: inteligência adaptativa; mineração de dados; orientação vocacional; plataforma Web.

I. INTRODUCTION

Vocational orientation is a process that allows a person to find the profession that will bring the greatest satisfaction and benefit, being the choice of career one of the most important decisions in life [1]. There are several determinants in a student's vocation, among which are: (a) personality, (b) interests, (c) self-concept, (d) cultural identity, (e) social connections, (f) role models, (g) economic resources, (h) parents' educational level, etc., which when recognized, become fundamental elements to guide the student in their learning and contribute to an appropriate career choice [2]. The process of vocational and career guidance has evolved in the world according to the different economic schemes specific to each era [3]. Globally many solutions have been proposed. The main objective is to facilitate the career choice for students [4]. In [5], the importance of having vocational guidance models that link different instruments for the collection of information, addressing various dimensions such as (i) self-knowledge, (ii) the world of work, and (iii) academic offerings, is highlighted. Additionally, it is considered essential to link artificial intelligence (AI) techniques to classify professions according to their profile and provide adequate feedback on career decisions [6].

The design of a knowledge base in [7] stores information about careers, experiences, and skills, and [8] uses fuzzy logic to select the most suitable professions according to the profile. Other solutions [7], [9] use recommender systems that link AI techniques and seek to offer careers based on the students' interests and skills.

In Colombia, the Ministry of National Education has published orientation methodological guides [10] as an "informative" process or guidance for students who evidence a conflictive situation regarding their professional choice; however, no results of the effectiveness of this strategy are shown. [11].

Other works have focused on neural networks using multiple intelligences as inputs and professions as network output [12], [13]. In addition to multiple intelligences, in [13], differential aptitudes and professional interests tests are used. [14] presents an expert system whose knowledge base uses interests, aptitudes, and skills; however, there is no evidence as results related to classifying professions. In [8], three

algorithms are used: 1. Decision trees, 2. K-nearest neighbor (KNN), and 3. Supported vector machine (SVM), where the SVM showed the best result of the classification of science, technology, engineering and mathematics STEM professions with a percentage of 72% for 142 records in the dataset. In [15], a dataset is outlined by the authors, taking into account academic aspects and using SVM and Artificial Neural Networks (ANN) as algorithms to classify professions. They achieved a performance of 84% of accuracy in the classification, the SVM algorithm offered the best results.

Despite many solutions developed to support vocational guidance processes, limitations or challenges are still evident, which should be addressed with new research projects. Such limitations and challenges include: 1. Scarce implementation of vocational guidance models in real contexts; 2. The absence of vocational guidance models, which consider the dynamic character of the student's profile; 3. High level of uncertainty in the vocational decision due to low levels of accuracy in the prediction of professions; and (iv) lack of vocational guidance solutions for the high school context.

Therefore, and considering the importance of guaranteeing a compelling orientation to students who are in the transition from high school to higher education, it is necessary to have counseling mechanisms and effective tools that allow recognition of competencies and skills to provide adequate guidance that contributes to a better career choice and, ultimately, to the personal and professional development of students.

Thus [16] presents in general terms the development of an intelligent web platform to support the process of vocational guidance in the context of secondary education. This article claims to describe the adaptive model of profession classification, developed and implemented to be used on an intelligent web platform to support vocational guidance processes in some educational institutions of the Department of Cauca.

The structure of this article is as follows: Section I presents some related works, highlighting mainly their strengths and weaknesses. Section II describes the methodological process for the development of the proposed solution. Section III

describes the results obtained following the stages of the selected methodology. Finally, conclusions and future work are presented.

In recent years, several computational tools have been developed to facilitate the vocational guidance process. From the mechanisms described in [17], [18], [19], and [20], playfulness is used to motivate the student to perform the vocational orientation process autonomously. However, the results delivered to the student are very broad, which does not contribute to the reduction of their uncertainty about their vocational decision.

In [21], speech and text recognition were used to facilitate interaction with the user of the vocational guidance system. However, the evaluation of the results shows an accuracy of 37% in predicting professions. [22] uses the internet of things (IoT) to propose the design of a guidance system in enterprises, which, with the use of radio frequency, allows students to see the compatibility of their skills with a specific job. This research does not show the levels of accuracy achieved during the guidance process.

In [6], a chatbot is used for virtual vocational guidance processes supported by DialogueFlow technology which applies a Google cloud application interface. It uses a dataset with attributes of activities, skills, occupations, and personality and proposes to give feedback to the student using Facebook chat. Similarly, in [23], Facebook and the internal Wit.ai application interface were used to convert voice responses and store them in the dataset. However, the two proposals do not evidence the effectiveness of the automated guidance process.

Other researches focus on the use of AI to support vocational guidance processes. In [24], fuzzy logic was used; however, the authors focus more on IT career classification. In [7], a vocational guidance prototype is presented, using cloud computing to integrate educational institutions and companies. In [25], ANNs are used to support the career decision-making process, and in [14], the recurrent neural network uses long short-term memory (LSTM) blocks to predict new jobs for adults. The use of ontologies and expert systems in [26] and [27] describe an expert system using fuzzy logic as in [28]. In these three works, low classification percentages are shown, with a maximum of 70% accuracy [29] and [30] using data mining and

distributed computing applications to analyze the career paths of adults based on their skills and requirements in the work context. Most of the above research is oriented toward the adult population.

In [9], [26], and [27], statistical and data mining models are used, with fixed architectures having specific characteristics of the vocational guidance process [26], [27] describes the use of Bayesian networks (Naïve Bayes NB), one rule Oner algorithm, and judgment repeated incremental pruning (JRIP), and case-based reasoning classifier (CBR). In [9], a mobile application that uses NB is proposed and validated in the middle school context. [31] suggests a profile evaluation related to college students' career choices using knowledge discovery in databases (KDD) to identify patterns in large amounts of data. This work evaluates the algorithm's performance using datasets with different numbers of records.

Finally, in [6], Multiclass Neural Networks (MNN) are used as the best classification algorithm with an accuracy of 90% for a data set of four attributes: activity, occupations, skills, and personality.

It is worth mentioning that online vocational guidance alternatives are a complement to face-to-face guidance [32]. Online vocational guidance communities help the student validate the results of face-to-face guidance, with the advantage of easy access and the diversity of knowledge received from other people. In [33], a web interface is presented for vocational guidance with different guidance alternatives and age ranges; however, the tool's accuracy in aspects of career classification is not shown.

In [34], a web platform is shown that using games motivates students to identify their skills and profiles through challenges that allow them to know themselves (weaknesses, strengths, and fears), project their career aspirations, and share and compare their results. This platform, as in the ludic web platforms [15], [17], [19], also does not show results in terms of the accuracy of the orientation of its users.

A general review of several vocational guidance web platforms in [35] shows these platforms adopted features like text-based content, videos of professionals from various areas or occupations, online tests, social networks, vacancy search agents, and job recommendations. However, most of these publications do not have an

internal evaluation in which users can validate the platform's contribution to their vocational orientation.

In [36], some essential elements identified in the vocational orientation process are used in different research projects, which mainly include: 1. Professional interests, 2. Personality, and 3. Skills. In addition to the above characteristics, in this research, the following aspects are included: social, economic, academic, and family. These elements are adopted in the web platform proposal.

[37] suggests some general characteristics that should be included in a vocational guidance website, which are grouped into the quality of service and perceived quality of information. The mentioned factors are taken into account in the proposed platform, which were validated through a usability survey conducted on a sample of students.

II. METHODOLOGY

The aim is to describe the adaptive model aspects for occupational classification described in [16] and identify the best algorithms to include in the intelligent adaptive web platform for vocational guidance. Below are described the associated methodological processes.

a. Participants: 628 students belonging to an official secondary education institution in the city of Popayán whose sociodemographic characteristics are: females 48.7% n= 306, males with 51.3% n= 322; age range between 12 and 23 years with an average of 16.03 years and with a standard deviation of 1.94; socioeconomic stratification levels are1, 2, and 3.

b. Measures: criteria of accuracy, precision, and recall were used to compare six models to choose the best ones for classifying and predicting professions.

c. Design: the data mining methodology (Cross Industry Standard Process for Data Mining, CRISP-DM) was used [38], which considers the following stages: understanding the business, understanding the data, data preparation, modeling, evaluation, and deployment. Holland's validated theory of career environments RIASEC [39] was also used to identify students' career interests and formalize the

process of gathering and analyzing information related to students' vocational decisions.

The CRISP-DM methodology, in conjunction with the RIASEC theory, was used to collect information on students' professional profiles through a web interface and, subsequently, consolidate the information into a data set of 628 records, with which it was possible to identify the best classification models for the particular characteristics of the educational population.

III. RESULTS

A. Bussiness Understanding

The first phase, understanding the business, proposed in the CRISP-DM methodology, was achieved by systematically reviewing policies, standards, theories, methodologies and tools to establish the RIASEC theory as the most practical and understandable, as well as desirable characteristics to be included in a vocational guidance platform.

B. Understanding and Preparing the Data

With the use of a web survey and the design of a data set directly related to the identified characteristics of the vocational orientation process, we proceeded to the second phase of understanding and preparation. A data set was created in the beginning, according to the RIASEC theory, with variables for each professional and occupational environment in which a student can be classified: Realistic, Investigative, Artistic, Social, Entrepreneurial, and Conventional. In addition, variables of age, gender, and professional expectations of the students were included. Figure 1 shows the correlation degree of RIASEC variables, age, and sex, with the variable binomial classification (engineering profession).

Figure 1 shows that the variable with the highest correlation in the choice of the engineering profession is gender, and the environments of the RIASEC theory that contribute most to the professional decision are the Conventional, Realistic, and Investigative types.
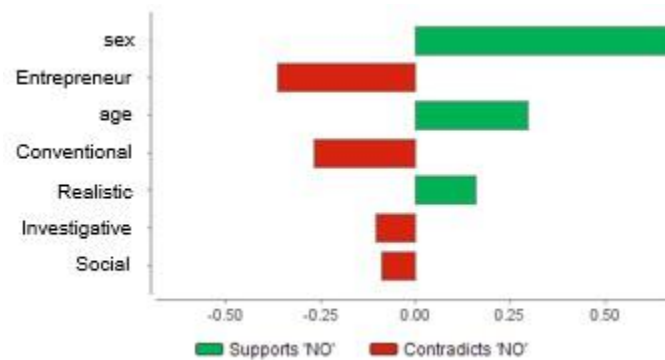
Fig. 1. Contribution of attributes in the classification of the engineering profession.

C. Modeling

Initially, a Naive Bayes (NB) algorithm was evaluated to create the first model, which showed low performance in accuracy and training time for multi-valued classification variable datasets. The results improved changing to a binomial variable. Later, in a comparison of up to five models (NB, linear logistic regression model (generalized linear model, logistic regression), large fast margin, and deep learning and decision tree), Naive Bayes was found again with higher accuracy and lower training time. In the comparison, two tools, Weka and Rapidminer, were used to determine the stability of the model in different datasets and sizes, starting from a stable model using the parameters of accuracy, precision, and completeness, as well as classification error and classification time. Datasets were projected to evaluate the algorithms' response according to the characteristics of the initial data.

D. Evaluation

In the evaluation phase, the first data set was analyzed using NB, Linear Regression, Logistic Regression, and Deep Learning models. The results are described in Table 1, with accuracy and training time shown in milliseconds for every 1,000 records for the generated models.

Table 1. Evaluation of algorithms with multi-valued classification variable.

Model	Accuracy	Training Time in ms for (1,000 Records)
NB	26%	4,179.9
Generalized Linear Model	27%	1,313.7
Logistic Regression	26%	1,503.2
Deep Learning	28%	2,781.8

In this initial result, the dataset used the classification variable "the profession" with ten different values of professions to address the expectations in professional choice on the part of the students. Due to the low accuracy obtained, as a result, this multi-valued variable was changed to a binomial value variable (YES/NO) specifically for the area of professions of engineering, architecture, and urban planning, and using the algorithms of NB, Generalized Linear Model, Logistic Regression, Deep Learning, and Decision tree. The results received are recorded in Table 2.

With this new data set, the results were analyzed in terms of accuracy, precision, and completeness, receiving a shorter training time and an acceptable performance for NB, which allowed the classification of professions based on the professional interests of the students (Table 2).

Table 2. Evaluation of algorithms with binomial classification variable.

Model	Classification error	Accuracy	Precisión	Recall	Training time (1,000 Records)
Naive Bayes	20%	80.3%	80%	80%	0.02
Decision Tree	19%	82.5%	50.20%	8.4%	0.11
Generalized Linear Model	20%	79.40%	0%	0%	0.24
Logistic Regression	20%	77%	74%	80%	0.42
Deep Learning	20%	72.20%	38.1%	64%	2.25

Subsequently, and taking into account the characteristics of data independence, which the NB model presupposes, it was decided to generate a data set with a more significant number of unique records (38,399) among the variables used in the data set for the classification process. These data were generated taking into account the restrictions in score values for each of the environments (from 0 to 30), as well as the values of age, ranges found in the data of the surveyed students (aged 12 to 23

years) in the initial data set. The results described in Table 3 show that the values of accuracy, precision, and completeness decreased.

Table 3. Evaluation of algorithms, and data set using random data.

Model	Classification error	Accuracy	Precisión	Recall	Training time (1,000 rows)
Naive Bayes	49.6%	50.4%	50.3%	95.9%	15.3
Decision Tree	49.8%	50.2%	50.2%	100.0%	100.3
Deep Learning	49.8%	50.2%	50.2%	100.0%	346.5
Fast Large Margin	50.3%	49.7%	49.0%	6.1%	148.6
Logistic Regression	50.7%	49.3%	49.8%	22.6%	25.2
Generalized Linear Model	50.8%	49.2%	49.7%	23.1%	29.1

A new data set was created to improve these results by attending to the resulting more stable model given in Table 2, for which, in addition to using the restrictions on the data boundaries, the production model was attended to, and the influence of each of the variables in the data set on the classification variable (the profession) can be seen in Table 4.

Table 4. Correlation of variables with class variable "profession" of binomial classification.

Variable	Correlation
Sex	9.77
Age	0.52
Realistic	0.65
Investigative	0.02
Artistic	0.02
Social	0.64
Entreprenur	0.20
Conventional	0.39

The strongest correlations were taken into account, in addition to the continuity in the highest values for groups of three contiguous environments according to the RIASEC hexagon [40]. With these new model considerations, a new dataset was created with 1,048,561 records, and the results are shown in Table 5.

A new dataset created with 1,048,561 records following the characteristics of the model was re-evaluated, increasing its classification percentage to 83% accuracy.

Other variables were added to this model to determine its stability by increasing the attributes of the dataset (personality, socioeconomic) to verify the impact on the classification.

Table 5. Evaluation of algorithms, data set using production model, correlation and RIASEC of three environments.

Modelo	Classification error	Accuracy	Precision	Recall	Training time (1,000 rows)
Naive Bayes	23%	77%	63%	39%	7.2
Decision Tree	20%	80%	59%	84%	9.3
Logistic Regression	23%	77%	62%	38%	20.8
Generalized Linear Model	23%	77%	63%	37%	24
Fast Large Margin	22%	78%	59%	62%	157.1
Deep Learning	17%	83%	65%	80%	373.6

As described in the values in Table 5, it is found that Naive Bayes is still a stable algorithm with a shorter training time. It is also visualized that, for a larger volume of data, Deep Learning is projected as a better alternative for classification processes of the data corresponding to students of the specific context. This indicates that the intelligent web platform used the best-performing algorithm allowing more efficient results during the adaptation process, considering the number of records.

E. Deployment

In the deployment phase, the evaluated classification model was enriched by integrating two models, NB and Deep Learning, linked to the web platform. Figure 2 abstracts its operation for understanding, and Figure 3 describes the components that integrate it. This configuration ensured the platform's adaptability in an intelligent way by allowing the automatic update of the classification model, taking into account the previous choices of the students and, with increases in the size of the dataset, ensuring optimal levels of accuracy, precision, and completeness.

Finally, the web application was developed taking into account features found in other tools proposed for vocational guidance, both in the functional part and in usability aspects, and in the developed web application, the adaptability

functionalities that make use of the mining module and the administration module were added.
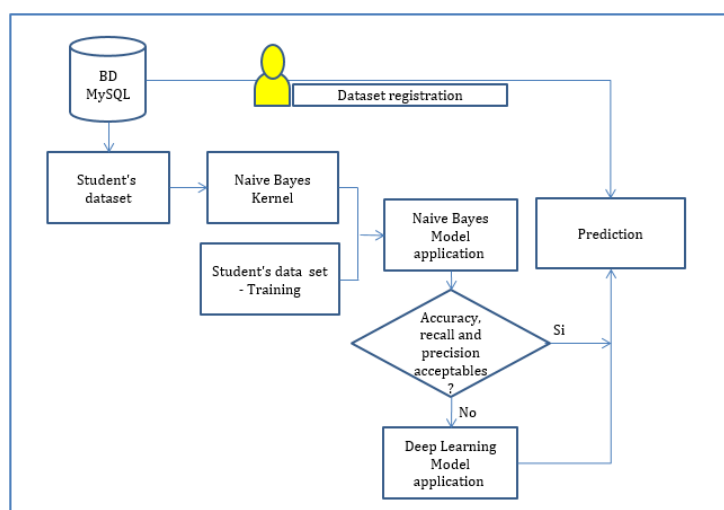

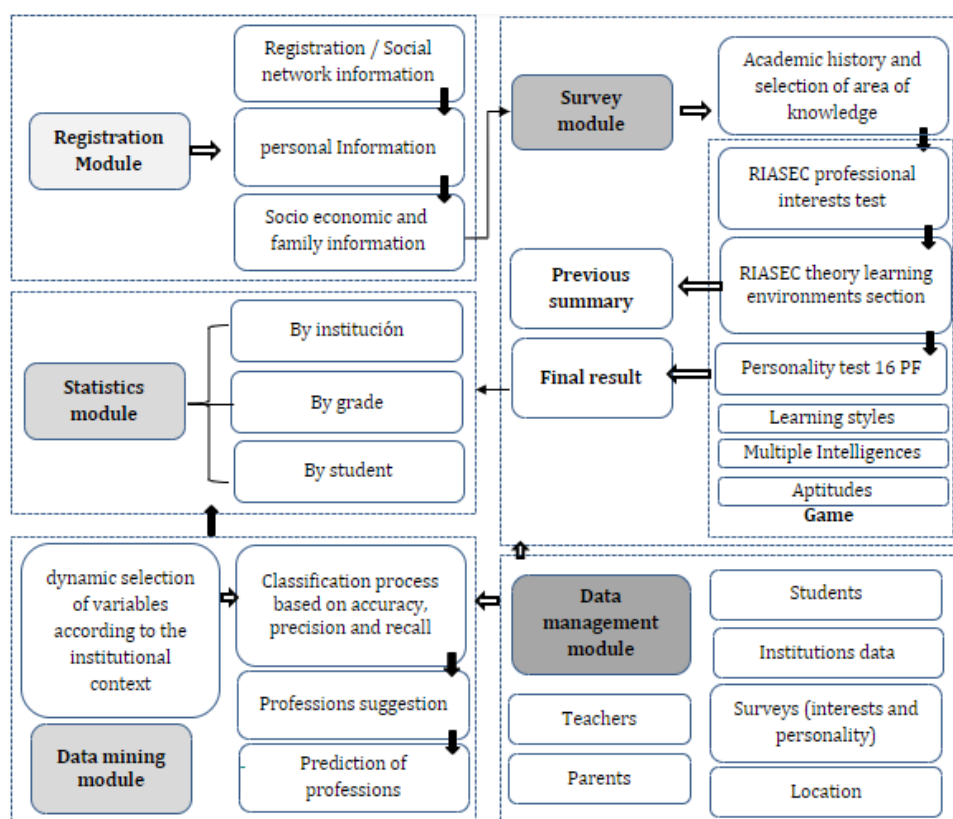**Fig. 2**. Occupational classification model.

Fig. 3. Description of the web platform.

In terms of functionality, the web platform allows users to register a new account by entering their data or using their account on Facebook. Moreover, users can be created directly by the system administrator. Once the profile is completed by entering family, demographic and economic information, it is possible to access the survey module where professional interests, personality, learning styles, multiple intelligences, and aptitudes (as the determinants of vocation) are evaluated.

Then, the mining module performs the adaptive classification process and generates the associated prediction according to the assessed profile. The integrated statistics module into the platform allows the system user and the administrator to view results and changes over time, allowing for comparisons and sequential and online monitoring within the institution (Figure 4).

The adaptive model considers the characteristics of the student profile related to the determinants of their vocation, including 1. Sociodemographic characteristics, 2. Interests, 3. Personality, 4. Learning styles, 5. Multiple intelligences, and 6. Aptitudes, and, additionally, the previous study expectations and preferences by subject, for which Figure 4 shows the statistical results of one of the preferred subjects per institution.

Fig. 4. Statistics by institution of taste by subjects.

This reading of the context, and the student's profile, must also consider changes in professions related to the academic world, which vary according to supply and demand statistics of the labor context. With this, educational institutions may require an adaptive model that allows them to select some characteristics according to the perception of their context. The classification algorithm is adapted according to the number of records, generating a new classification model that allows greater accuracy (Figure 5).
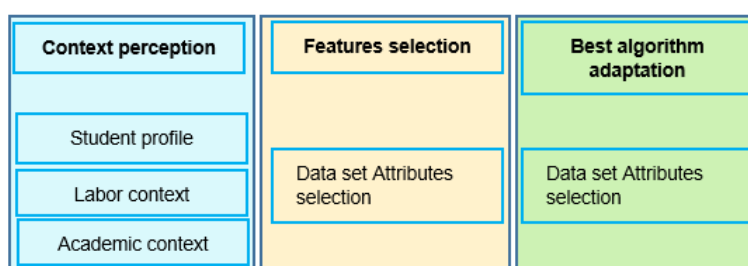

Fig. 5. Abstraction of the implemented adaptive model.

Regarding the prediction of professions, the professional branches suggested by the Ministry of National Education, according to [41], are taken into account; they are 1. Engineering architecture urbanism and related, 2. Agronomy veterinary and related, 3. Fine arts, health sciences, 4. Social and human sciences, 5. Economics administration accounting and related, 6. Mathematics and natural sciences, and 7. Education sciences, among others.

The model includes two algorithms, Naive Bayes and Deep Learning, which adapt automatically according to the number of records. The model uses supervised learning to provide a better classification variable whose categories are the professions to offer a better classification percentage.
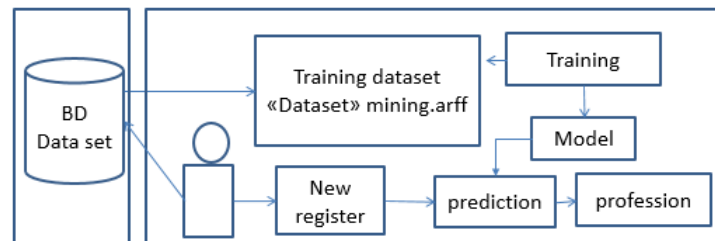

Fig. 6. Description of mining module functionality.

The classification of professions follows the binomial variable and performs the respective prediction. A student obtains the evaluation score of the knowledge acquired about the RIASEC environments and the classification for the suggested professional branch.

To ensure the student's atention in the surveys, the platform integrates playful challenges, presenting images associated with the test questions of the underlying theory for selection and taking into account the average time of students who answer each section of questions in a concentrated manner, the student acquires a positive point when selecting the images or loses if it takes more than the average time. In addition, it encourages their attention and motivation, making the process more objective. Figure 7 shows the results after the classification process has been carried out. In the example, a positive prediction process is evidenced for the branch of engineering, architecture, and related fields, with an accuracy percentage of over 80%.
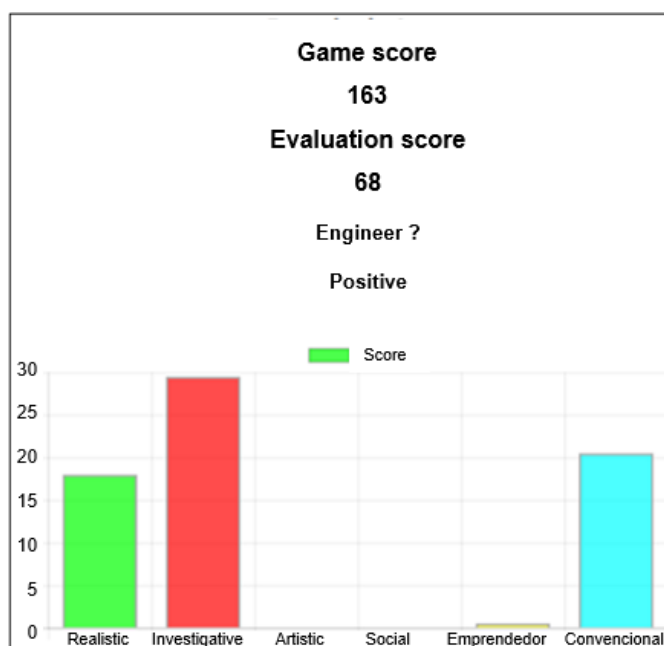
Fig 7. Mining module functionality description

Finally, Figure 8 shows the results of the surveys on 1. Professional interests, 2. Personality, 3. Learning styles, 4. Multiple intelligences, and 5. Aptitude tests.
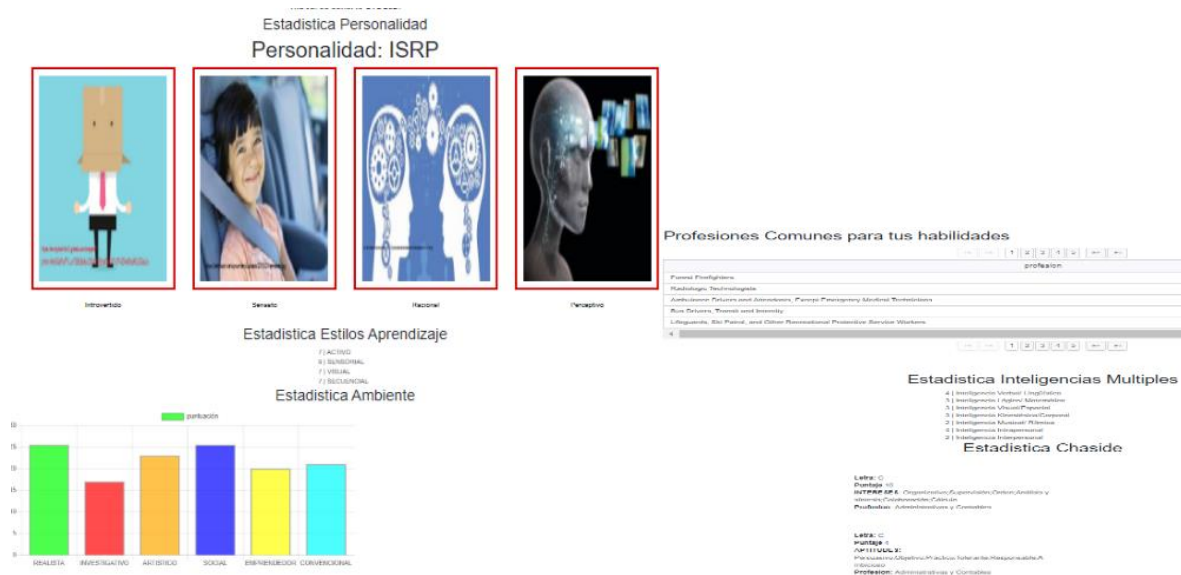

Fig. 8. Results of self-knowledge and career suggestion tests.

IV. CONCLUSIONS

The evaluation of data sets with a varying number of records, using Naive Bayes, Decision Tree, Logistic Regression, Generalized Linear Model, Large Fast Margin, and Deep Learning algorithms, allowed identifying the NB algorithm as a stable algorithm with the accuracy of classification of the profession binomial variable of up to 80.02%.

The model obtained describes the context of an educational institution in the city of Popayán using the theory of six environments (realistic, investigative, artistic, social, entrepreneurial, and conventional) (Holland's RIASEC). Additionally, the attributes of sex, age, profession, and sociodemographic context data were used to form a data set that allowed verifying the classification stability of the NB model and projecting larger data sets.

The best-performing algorithm In a dataset of 1,048,561 records was Deep Learning, with a classification profession accuracy percentage of 83%, thus recognizing these two algorithms. NB is recognized for small amounts and Deep Learning for large quantities of records, which were integrated into the web platform to respond adaptively, looking for the best classification percentage and learning from the previous model.

The statistics related to students' subject preferences show artistic profiles, specifically for one of the institutions with the data evaluated over two consecutive years, which allows decisions to be made regarding the strengthening of these skills in the curricular integration of area plans.

The model's adaptability provides the student with a more optimal career that is for them according to their profile characteristics. On the other hand, it allows the educational institution to evaluate which attributes of the profile contribute more to the vocation of its students.

For future work, it is expected to conduct a more significant number of tests, with a more considerable amount of data, linking students from different educational institutions in Popayán city and the Department of Cauca.

REFERENCES

[1] I. Gati, N. Levin, S. Landman-Tal, "Decision-Making Models and Career Guidance," in *International Handbook of Career Guidance*, 2019, pp. 115–145. https://doi.org/10.1007/978-3-030-25153-6_6

[2] S. Whiston, N.Mitts, Y. Li, "*Evaluation of Career Guidance Programs*," in *International Handbook of Career Guidance*, J. A. Athanasou and H. N. Perera, Eds. Cham: Springer International Publishing, 2019, pp. 815–834. https://doi.org/10.1007/978-3-030-25153-6_38

[3] S. Niles, R. Vuorinen, A. Siwiec, "Training Career Practitioners for the Current Context*,*" in *International Handbook of Career Guidance*, J. A. Athanasou and H. N. Perera, Eds. Cham: Springer International Publishing, 2019, pp. 529–553. https://doi.org/10.1007/978-3-030-25153-6_25.

[4] Y. Zambrano, *Desarrollo de una aplicación web para la orientación vocacional y promoción de carreras STEM implementando técnicas de Data Mining,* Grade Thesis, Universidad del Norte, Colombia, 2021. http://hdl.handle.net/10584/9857

[5] F. W. Vondracek, E. J. Porfeli, D. H. Ford, *Living Systems Theory: Using a Person-in-Context Behaviour Episode Unit of Analysis in Career Guidance Research and Practice*. 2019. https://doi.org/10.1007/978-3-030-25153-6_23

[6] O. Zahour, E. H. Benlahmar, A. Eddaouim, O. Hourrane, "A comparative study of machine learning methods for automatic classification of academic and vocational guidance questions," *International Journal of Interactive Mobile Technologies*, vol. 14, no. 8, pp. 43–60, 2020. https://doi.org/10.3991/ijim.v14i08.13005%0AOmar

[7] H. El-Sofany, S. El-Seoud, "A cloud-based educational and career guidance model using fuzzy logic concepts," *ACM International Conference Proceeding Series*, pp. 167–172, 2019. https://doi.org/10.1145/3328833.3328846

[8]   H. Youssef, A. Moumen, D. Gretete, "Fuzzy Logic for Academic Orientation and Its Impact on Success: Content Analysis," in *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning - BML*, pp. 546–551, 2022. https://doi.org/10.5220/0010742200003101

[9]   R. M. Acerado, R. C. Morco, J. R. Santos, J. J. Carpio, H. A. Isanan, "Predict: A mobile application for predicting the students' career using naïve bayes algorithm," in *Proceedings of the 2nd International Conference on Software Engineering and Information Management*, 2019, pp. 119–123. https://doi.org/10.1145/3305160.3305169

[10]   H. Botello Peñaloza, "Incidencia de los programas de orientación vocacional en Colombia," *Horizontes Pedagógicos*, vol. 1, no. 16, pp. 89–97, 2014. https://doi.org/10.33881/0123-8264.16108.

[11]   C. M. Gonzalez Diaz, L. Ortegón, A. M. Diaz, *Rutas de Vida. Manual de acompañamiento en orientación socio ocupacional*, Bogotá D.C., 2015. https://www.mineducacion.gov.co/1759/w3-article-356514.html?_noredirect=1

[12]   T. Pérez, J. Contreras, A. Puentes, *Orientación vocacional aplicando sistemas basados en conocimiento*, 2017.

[13]   D. Marcela, P. Morales, A. Celmira, G. Gualteros, *Red neuronal artificial para orientación profesional* UDProfession, 2009.

[14]   F. Téllez, E. Pineda, T. Meneses, J. Medina, "Sistemas expertos y orientación vocacional en educación a distancia virtualmente mediada," *Aibi*, vol. 8, no. S1, pp. 186-195, 2021. https://doi.org/10.15649/2346030X.2424

[15]   Y. Nieto, V. García-Díaz, C. Montenegro, R. G. Crespo, "Supporting academic decision making at higher educational institutions using machine learning-based algorithms," *Soft Computing*, vol. 23, no. 12, pp. 4145–4153, 2019. https://doi.org/10.1007/s00500-018-3064-6

[16]   A. Cruz, L. Orozco, C. Gonzales, "*Intelligent web platform for vocational guidance*," in *Proceedings International Conference on Virtual Reality and Visualization*, pp. 205–207, 2019. https://doi.org/10.1109/ICVRV47840.2019.00049

[17]   A. Shipepe, A. Peters, "The hard struggle: A co-designed interactive career guidance online game," in *ACM International Conference Proceeding Series*, 2018, pp. 314–317. https://doi.org/10.1145/3283458.3283512.

[18]   E. Guardiola, S. Natkin, "A Game Design Methodology for Generating a Psychological Profile of Players," in *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*, C. S. Loh, Y. Sheng, and D. Ifenthaler, Eds. Cham: Springer International Publishing, 2015, pp. 363–380. https://doi.org/10.1007/978-3-319-05834-4_16

[19]   I. Dunwell, P. Lameras, S. de Freitas, P. Petridis, K. Star, M. Hendrix, S. Arnab, "Metycoon: A game-based approach to career guidance," in *5th International Conference on Games and Virtual Worlds for Serious Applications*, 2013, pp. 1-6. https://doi.org/10.1109/VS-GAMES.2013.6624237

[20]   O. Garcia, J. Serra, J. Membrives, J. Juarez, "Waypass: A Gamified Self-Knowledge Quest for Teenagers," in *8th International Conference on Games and Virtual Worlds for Serious Applications*, 2016, pp. 1–4. https://doi.org/10.1109/VS-GAMES.2016.7590380

[21]   D. Calvo, L. Quesada, G. López, L. Guerrero, "Multiplatform Career Guidance System Using IBM Watson, Google Home and Telegram," in *International Conference on Ubiquitous Computing and Ambient Intelligence*, 2017, pp. 689–700. https://doi.org/10.1007/978-3-319-67585-5_67

[22] H. Mrabet, A. Moussa, "Smart school guidance and vocational guidance system through the internet of things," in *ACM International Conference Proceeding Series*, 2019. https://doi.org/10.1145/3320326.3320404

[23] N. Suresh, N. Mukabe, V. Hashiyana, A. Limbo, A. Hauwanga, "Career Counseling Chatbot on Facebook Messenger using AI," in *ACM International Conference Proceeding Series*, pp. 65–73, 2021. https://doi.org/10.1145/3484824.3484875

[24] T. Razak, M. Hashim, N. Noor, I. Halim, N. Shamsul, "Career path recommendation system for UiTM Perlis students using fuzzy logic," in *5th International Conference on Intelligent and Advanced Systems*, pp. 1–5, 2014. https://doi.org/10.1109/ICIAS.2014.6869553

[25] A. S. Rao, B. S. Kamath, R. R, Shreya Chowdhury, S. A Pattan, R. K. Kundar, "Use of Artificial Neural Network in Developing a Personality Prediction Model for Career Guidance: A Boon for Career Counselors," *International Journal of Control and Automation*, vol. 13, no. 4, pp. 391–400, 2020.

[26] H. I. Bülbül, Ö. Ünsal, "Comparison of classification techniques used in machine learning as applied on vocational guidance data," in *Proceedings - 10th International Conference on Machine Learning and Applications*, 2011, pp. 298–301. https://doi.org/10.1109/ICMLA.2011.49

[27] M. Nawaz, A. Adnan, U. Tariq, F. Salman, R. Asjad, M. Tamoor, "Automated Career Counseling System for Students using CBR and J48," *Journal of Applied Environmental and Biological Sciences*, vol. 4, no. 7S, pp. 113–120, 2015.

[28] A. Mundra, D. S. Chauhan, A. Soni, S. K. Sharma, P. Kumar, "Decision support system for determining right education career choice," in *ICC 2014-Computer Networks and Security*, pp. 8–17, 2014.

[29] C. Zhenyu, "The Application of Big Data in Higher Vocational Education Based on Holland Vocational Interest Theory," in *Proceedings - 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration*, 2017, pp. 37–40. https://doi.org/10.1109/ICIICII.2017.59

[30] W. C. Lu, "Future work-self salience and proactive career behavior among college student-athletes in Taiwan: A career construction model of adaptation," *Journal of Hospitality, Leisure, Sport & Tourism Education*, vol. 27, e100259, 2020. https://doi.org/10.1016/j.jhlste.2020.100259

[31] L. I. Jimenez-Raygoza, A. S. Medina-Vazquez, G. Perez-Torres, "Proposal of a computer system for vocational guidance with data mining," in *IEEE International Conference on Engineering Veracruz*, 2019. https://doi.org/10.1109/ICEV.2019.8920523

[32] J. A. Cerrito, J. Trusty, R. J. Behun, "Comparing Web-Based and Traditional Career Interventions With Elementary Students: An Experimental Study," *Career Development Quarterly*, vol. 66, no. 4, pp. 286–299, 2018. https://doi.org/10.1002/cdq.12151

[33] Y. Turganbayev, G. Adilgazinov, Y. Barabanova, A. Zhakupov, G. Zhukibayeva, "Information System for Vocational Guidance, Employment and the Forecasting of Labor Demand: The Case of Kazakhstan," in *IEEE International Conference on Smart Information Systems and Technologies*, pp. 28–30, 2021. https://doi.org/10.1109/SIST50301.2021.9465984

[34] A. Afolabi, R. Ojelabi, L. Amusan, F. Adefarati, "Development of a web-based building profession career portal as a guidance information system for secondary school students," in *Proceedings of the IEEE International Conference on Computing, Networking and Informatics*, pp. 1–10, 2017. https://doi.org/10.1109/ICCNI.2017.8123771

[35] U. Paukstadt, K. Bergener, J. Becker, V. Dahl, C. Denz, I. Zeisberg, "Design recommendations for web-based career guidance platforms – Let young women experience IT careers!," in *Proceedings of the Annual Hawaii International Conference on System Sciences,* pp. 5116–5125, 2018. https://doi.org/10.24251/hicss.2018.639

[36] J. C. G. Vargas, J. L. P. Laguna, A. Carrillo-Ramos, *"*KunaySoft: Adaptive system of vocational guidance*,"* in *10th Computing Colombian Conference*, pp. 448–455, 2015. https://doi.org/10.1109/ColumbianCC.2015.7333460

[37] N. Cobelli, A. Bonfanti, S. Cubico, G. Favretto, "Quality and perceived value in career guidance e-services," *International Journal of Quality and Service Sciences*, vol. 11, no. 1, pp. 53–68, 2019. https://doi.org/10.1108/IJQSS-12-2017-0114

[38] F. Martinez-Plumed, L. Contreras-Ochando, C. Ferri, J. Hernández-Orallo, M. Kull, N. Lachiche, M. J. Ramírez-Quintana, P. Flach, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, 2021. https://doi.org/10.1109/TKDE.2019.2962680

[39] J. L. Holland, *Making vocational choices: A theory of vocational personalities and work environments*, US: Psychological Assessment Resources, 1997.

[40] A. Erdogan, *Holland's theory of careers and vocational choice,* 2022.

[41] J. F. Martin Calvo, "Calidad educativa en la educación superior colombiana: una aproximación teórica," *Sophia*, vol. 14, pp. 4–14, 2018. https://doi.org/10.18634/sophiaj.14v.2i.799