International
Journal of
**Human
Sciences
Research**

CURRENT TRENDS IN EDUCATIONAL PREFERENCES IN MEXICO: A COMPARATIVE ANALYSIS BETWEEN TECHNICAL AND UNIVERSITY EDUCATION

Yolanda Moyao Martínez
Benemérita Universidad
Autónoma de Puebla

Carmen Cerón Garnica
Benemérita Universidad
Autónoma de Puebla

Beatriz Beltrán Martínez
Benemérita Universidad
Autónoma de Puebla

Abstract: This article presents a regression model to predict preferences between two bachelor's degrees and four technical degrees for public high school students. The data analysis and prediction model aim to identify patterns and preferences among this group of students with the goal of understanding how their motivations or preferences for the next stage of education align with the profiles most in demand by Industry 4.0, with technological, digital, analytical, and multidisciplinary skills. For this reason, the analysis carried out using the prediction model focuses on bachelor's degrees and technical degrees related to computing. For this study, the necessary information was collected through questionnaires administered to a representative sample of students from thirty-two different high schools in the state of Puebla. The results reveal important trends and key points regarding educational preferences. This analysis using the prediction model becomes a useful tool for educators and institutions responsible for educational policies for the creation and improvement of academic programs.

Keywords: Data analysis, Demand profiles, Industry 4.0, Educational policies.

INTRODUCTION

Throughout their lives, high school students face one of the most difficult and important challenges, as they have to choose the bachelor's degree or technical career they wish to pursue in the coming years. This is a difficult process because it can be compromised by various factors such as the influence of other people, such as friends or acquaintances, personal interests, financial considerations, and the educational context in which they operate [García, 2020], [Aragón, 2020].

Therefore, this work, which seeks to identify patterns and trends through a prediction model of preferences for bachelor's degrees and technical careers related to computing among high school students, can be very significant as a tool to justify such choices [Lee, 2019].

There is a strong relationship between the choice of undergraduate or technical career and the current needs of companies and society in general, which is a key aspect of educational planning aimed at preparing professionals capable of meeting the demands of the labor market.

There are some results, such as those presented in [Real, 2024], which show that students are currently more focused on choosing university courses that give them the competitive tools and skills they need to face the challenges of society, which are becoming increasingly demanding and aligned with the need for the use of technology.

It is important to consider that, thanks to constant technological advances, the world is changing in terms of how we communicate, how we learn and how we teach. In other words, Industry 4.0, which is characterized by the use of emerging technologies such as Artificial Intelligence (AI), the Internet of Things (IoT), Big Data, and automation [González, 2018].

For this reason, it is necessary to constantly create or update educational programs, as Industry 4.0 demands professionals with increasingly specific profiles who can adapt to new needs through skills and abilities linked to technology and engineering [Bertram, 2019].

This study conducts a data analysis that aims to explore and understand various aspects related to the preferences between two bachelor's degrees and four technical degrees from a sample of 9,465 high school students in the state of Puebla. This study is based on data collected from a set of surveys conducted with high school students to identify patterns, trends, and relationships between different variables.

It also aims to contribute to the effort to adapt and create new educational programs, with the goal of contributing to decisions for the development of higher education that is aligned with the needs of students and will enable them to face challenges and contribute to meeting the demands of Society 4.0 or Industry 4.0 (Fourth Industrial Revolution).

METHODS

Below, we describe each of the steps necessary to perform data analysis, from data collection to the interpretation of results.

Data Collection

The data was collected through a questionnaire, as shown in Figure 1, and then stored in an Excel file. This includes information on age, gender, current semester, institution, student interests in two bachelor's degrees and four technical degrees, and an additional column for general comments. With this data set, different analyses could be performed.

Initially, the distribution of gender and age was graphed to help understand the demographic composition of the students. Subsequently, an exploratory analysis was performed using to represent preferences for the different bachelor's degrees, as well as a correlational analysis between the different variables.

Source: Own elaboration

Figure 1 Instrument for collecting data.

Data processing

The Excel file was cleaned so that all data, including age, gender, semester, institution, and preference for each degree program and technical career, was numerical, since the data must be standardized for analysis. In addition, some assignments were made to the names as-sociated with each of the twenty degrees, as can be seen in the Appendix: Notations. This was done because the names are too long to be shown in the graphs.

In addition, some records that were inconsistent or had responses with values outside the permitted range were deleted. The age range is only from 14 to 20 years old, the value for gender can only be 1 "female," 2 "male," 3 "prefer not to say," the current semester ranges from 1 to 6, and the value for the choice of a bachelor's degree or technical degree can only be between 1 and 5, values represented on a Likert scale, for which the highest level of satisfaction is represented by the number five. For the analysis, responses with a value of 4 or 5 were considered.
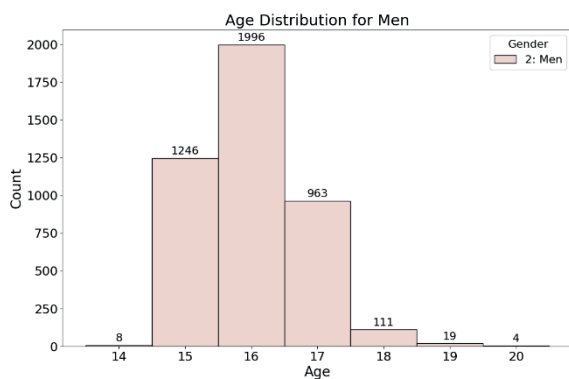
ANALYSIS AND GRAPHICAL REPRESENTATION OF THE DATA

DISTRIBUTION GRAPHS

Histograms were used to plot the age distributions for women and men. These histograms show how the ages of women and men are distributed ( ) and the age ranges in which the largest number of students are concentrated.
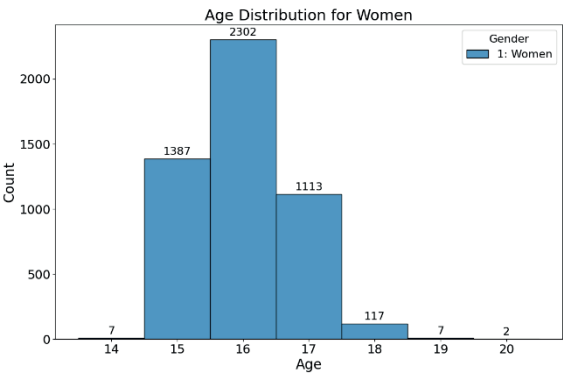
a) *Age distribution of women*

b) *Age distribution for men*

Figure 2 Frequency distributions.

Source: Python program

In Figure 2, sections a) and b), it can be seen, respectively, that the age range where the highest number of women and men are concentrated is 16 years old. It can also be seen that, in the population studied, there are a total of 4,935 women, which is greater than the 4,347 men, showing that at the preparatory or high school level, education is inclusive and egalitarian and governed by equal opportunities, as established by the UN, particularly in the "Sustainable Development Goals" (SDGs)4 [United Nations, 2015].

Sustainable Development Goals" (SDGs)4 [United Nations, 2015].

STATISTICAL SUMMARY

In Figure 3, sections a) and b), it can be seen, respectively, that the bar graphs show the level of interest that both women and men have in bachelor's degrees and technical careers, thus showing the preferences that best suit their academic needs for the next stage of their lives.

In the case of a population of 4,935 women in the " " (Women's University of the Pacific) program, 1,266 (25.7%) chose a TSU in Health Informatics, 1,214 (24.6%) chose a Bachelor's Degree in Nanotechnology, 1,192 (24.2%) chose a TSU in Environmental Te-

chnology Engineering and a r Sustainability, while 1,070, equivalent to 21.7%, chose TSU in Biotechnology, 1,027, equivalent to 20.8%, chose TSU in Digital Transformation, and 1,003, equivalent to 20.3%, chose the Bachelor's Degree in Computational Linguistics and Technology. This can be seen in Figure 4, subsection a).

From this, we can conclude, on the one hand, that TSU programs were chosen by a considerable number of women, which may mean that women prefer short programs with immediate practical application. On the other hand, women have a strong interest in educational programs focused on STEM areas that combine environmental biotechnology or computer science applied to biology and health, and they also show interest in emerging technologies such as nanotechnology.

In the case of a population of 4,347 men, 1,418 (32.6%) chose the Bachelor's Degree in Nanotechnology, 1,305 (30.0%) chose the Bachelor's Degree in Computational Linguistics and Technology,
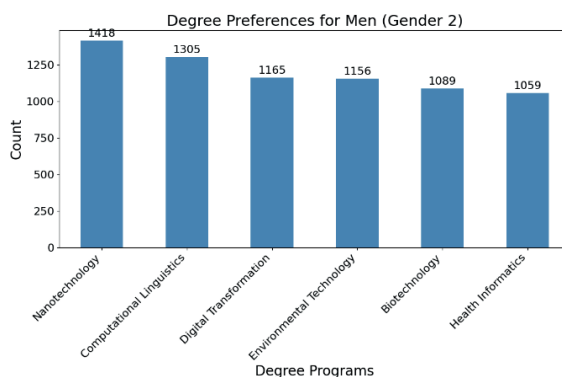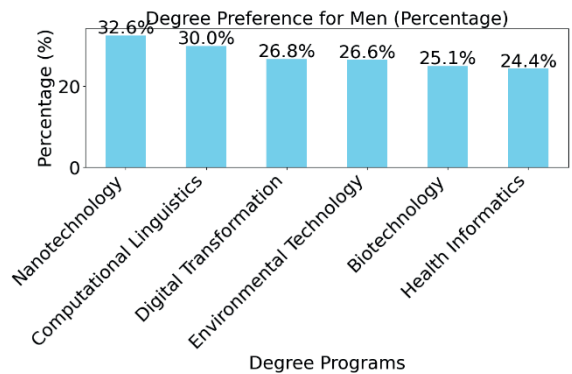

a) *Preference among women*

Degree Preferences for Women (Gender 1)
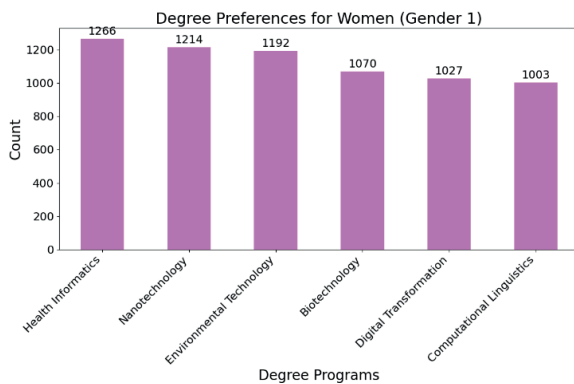
b) *Preference among men*

Degree Preference for Men (Percentage)

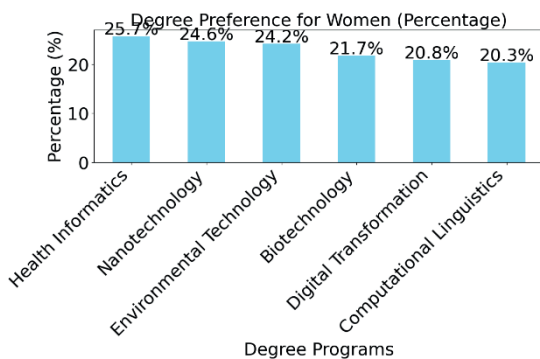b) *Male preference*

Figure 3 Preferences in bar chart.

Source: Python Program

Figure 4 Preference percentages.

Source: Python Program

1,165, equivalent to 26.8%, chose a TSU in Digital Transformation, while 1,156, equivalent to 26.6%, chose TSU in Environmental Technology and Sustainability Engineering, 1,089, equivalent to 25.1%, chose TSU in Biotechnology and r 1,059, equivalent to 24.4%, chose TSU in Health Informatics. This can be seen in Figure 4, section b).

Degree Preference for Women (Percentage)

a) *Percentage of female preference*

From the above, it can be concluded, on the one hand, that educational programs with theoretical training are just as important to men as educational programs that allow them to enter the labor market immediately, but with the acquisition of technical and practical skills.

On the other hand, based on men's preferences regarding bachelor's degrees, it can be concluded that they are more inclined toward technology, digitalization, and sustainability as aligned and essential elements for enhancing the skills that will enable them to adapt and meet the needs of a dynamic and constantly changing future workplace.

With the results obtained above, it can be inferred that both women and men follow a common guideline in their interests, as both have a strong tendency towards areas of technology for sustainable development, despite the fact that their preferences in general are from a slightly different point of view. This can be seen in the histogram in Figure 5.
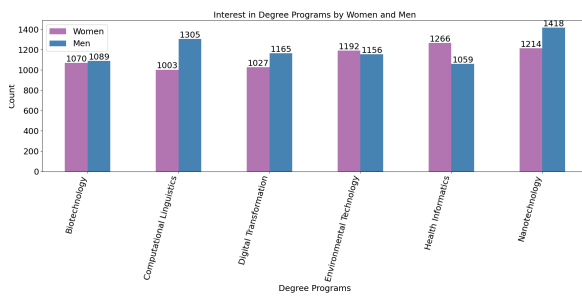
Figure 5 Comparison of preferences between women and men.

Source: Python program

EXPLORATION OF CORRELATIONS

Figure 6, sections a) and b), show the correlation matrices between age and preference for different degrees, where the correlations range from -0.06 to -0.01. Similarly, the correlations between gender and preference for degrees range from -0.1 to -0.02. Thus, it can be concluded that the correlation between the variables is negative or close to zero and, furthermore, very weak, meaning that there is no significant relationship between demographic variables and preferences for different majors.
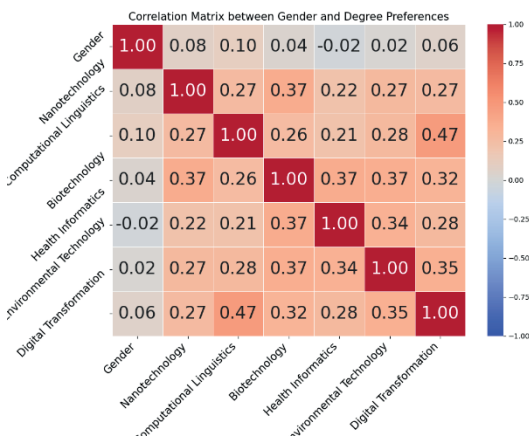
a) Between age and degree programs
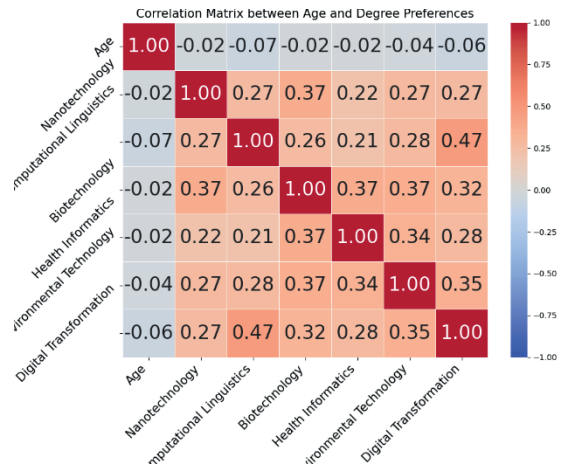
b) Between gender and degree programs

Figure 6 Correlations between age and gender variables.

Source: Python program

MACHINE LEARNING MODEL.

MODEL SELECTION

A logistic regression model is used. This type of model is useful when binary classification is required, i.e., when only two different classes need to be predicted: class 1, where a student chooses the degree that best suits their interests, or class 0, where a student does not choose a degree. It is a widely used and fairly simple standard model for data analysis.

MODEL TRAINING

Once the model has been selected, 80% of the data is used to train the model so that it can learn and generalize to new data. The remaining 20% is used to perform the necessary tests to evaluate its performance, robustness, and validity, thus ensuring a well-trained model with the most accurate evaluation possible.

INTERPRETATION OF RESULTS

Figure 6 shows the different evaluations, the classification report, the confusion matrix, and the accuracy of the model. Some important metrics such as probability distribution and the relationship between precision and recall are also displayed ( ).

The model has an accuracy of:

- 72.87% of the time it predicted Class 0: "did not choose Nanotechnology" as correct.
- 73.89% of the time it predicted Class 1: "chose Nanotechnology" as correct.

The results for recall are also shown, i.e., the number of cases that the model detected correctly for each of the classes.

- For class 0 (did not choose the degree), it correctly detected 73.62%.
- For class 1 (did choose the degree program), it correctly detected 73.14%.

	precision	recall	f1-score	support
0	0.728669	0.736207	0.732419	580
1	0.738908	0.731419	0.735144	592
accuracy	0.733788	0.733788	0.733788	0.733788

	Predicted No	Predicted Yes
No	427	153
Yes	159	433

Figure 6 Classification report and confusion matrix for the Bachelor's Degree in Nanotechnology.

Source: Python Program

Regarding the confusion matrix, which shows the model's performance in terms of correct and incorrect predictions, the following can be noted:

- Not class 0: 427 correct (correct prediction that they did not choose the bachelor's degree).
- Class 1: 153 incorrect (false negatives: students who actually choose the bachelor's degree, but the prediction was that they would not choose it).
- Not class 0: 159 incorrect (false positives: students who do not choose the bachelor's degree, but the prediction was that they would).
- Yes class 1: 433 correct (correct prediction that they will choose the degree program).

Considering that the results obtained for the other five degrees proposed for this study are very similar to those obtained for the Bachelor's Degree in Nanotechnology, it can be concluded that the model works well, with consistent metrics ( ) between precision and recall for both classes and an overall accuracy of 73.38% of all predictions.

RESULTS

Based on the analysis shown above, the results obtained in this study are listed below:

- Based on age and gender distribution, most women are in the 16-year-old age range, with a total of 4,935. Most men are also in the 16-year-old age range, with a total of 4,347.
- Based on the percentages obtained for the proposed careers, women have a strong interest in educational programs that combine areas such as environmental biotechnology or computer science applied to biology and health. The percentages of choice are distributed as follows: TSU in Health Informatics (25.7%), Bachelor's Degree in Nanotechnology (24.6%), TSU in Environmental Technology and Sustainability Engineering (24.2%), TSU in Biotechnology (21.7%), Digital Transformation (20.8%), and Bachelor's Degree in Computational Linguistics and Technology (20.3%).
- Based on the percentages obtained for the preferred degree programs, educational programs with theoretical training are equally important for men, as are educational programs that allow them to enter the workforce immediately, but with the acquisition of practical skills. The percentages of choice are distributed as follows: Bachelor's Degree in Nanotechnology (32.6%), Bachelor's Degree in Computational Linguistics

and Technology (30.0%), TSU in Digital Transformation (26.8%), TSU in Environmental Technology and Sustainability Engineering (26.6%), TSU in Biotechnology (25.1%), and TSU in Health Informatics (24.4%).

- It can be observed that the trends in choice are geared towards educational programs to train professionals who can adapt to the new needs of Industry 4.0.

- Regarding the correlations between the variables, age and gender and preference for different careers, it can be concluded that the correlation between the variables is negative or close to zero and, furthermore, very weak, meaning that there is no significant relationship between them.

- The model has an overall accuracy and recall of around 73% and a fairly balanced performance between the two classes (0 and 1), which can be considered acceptable performance.

DISCUSSION

This study, applied to BUAP high school and baccalaureate students to determine their preferences regarding university majors in Puebla, reveals results with significant patterns in preferences classified by gender, which are also strongly related to technology and sustainability.

- The results obtained in the analysis clearly show a high level of interest among women in educational programs that tend to integrate variables from areas such as environmental biotechnology and computer science applied to biology and health. This interest breaks historical gender stereotypes and can be used as a basis for strengthening policies for female inclusion in STEM (Science, Technology, Engineering,

and Mathematics) [Lee, 2019].

- In fact, [Smith, 2018], [Wang, 2017] present some results that show a strong interest among women in STEM disciplines, and this is particularly reflected in their preference for areas that have a strong impact on health and the environment.

- The results obtained in the analysis clearly show that men have a high interest in educational programs that lean more towards technology, digitalization, and sustainability as aligned and essential elements for enhancing the " " capabilities that will allow them to adapt and face the needs of a constantly changing future workplace [Jones, 2020].

- The results of this study clearly show the urgent need to create or adapt educational programs to ensure that they are aligned with the demands of Industry 4.0.

- This is necessary because the new skills and knowledge that professionals need on a daily basis must be integrated. Although this is a major challenge, it is essential in order to adapt and make an impact in this Industry 4.0 revolution.

- The correlation matrices show that the relationship between demographic variables is negative or close to zero and, moreover, very weak, suggesting that there is no significant relationship between them and preferences for different careers.

- Therefore, the results suggest that both individual and contextual variables, as well as demographic variables, influence career choice, as shown by the results of [García, 2020].

CONCLUSIONS AND FUTURE WORK

With the results obtained through the prediction model and shown above, it can be inferred that both women and men follow a common guideline in their interests, as both have a strong tendency towards areas of technology for sustainable development, despite the fact that their preferences in general are from a slightly different point of view.

It is important to mention that, based on the patterns identified in this analysis through the prediction model, there is a strong demand for the updating or creation of new educational plans that meet the demands of Industry 4.0.

Therefore, educational organizations and institutions need to incorporate skills into their programs that will enable professionals to be prepared for major challenges, such as automating industrial methods through the effici y use of emerging technologies, and thus be ready for new job opportunities.

Although patterns and some results that favor decision-making have been obtained through this analysis and the prediction model, for future research, it would be beneficial to have some other variables that could have a significant impact on the choice of one of the proposed careers.

With the results obtained from this analysis through the prediction model, we propose to promote the integration of the STEM approach in upper secondary education. In other words, to integrate the disciplines of science, technology, engineering, arts, and mathematics across the curriculum to develop skills in students, such as critical thinking, creativity, and problem-solving that require more complex solutions.

Through the use of digital technologies, the aim is to train students to face the challenges of an increasingly interconnected and dynamic world.

Although the predictive regression model performs reasonably well at around 73% for both classes, it can still be improved by training it with more features or adjusting hyperparameters.

The model could be further refined by adjusting hyperparameters, using other classification algorithms, or adding more relevant variables. This is because there may be additional factors that are not being fully captured by the current variables, especially in the field of education.

REFERENCES

1. Aragón Macías, L., Arras Vota, A. M. de G., Guzmán Ibarra, I. Realidad actual de la elección de carrera profesional desde la perspectiva de género. Revista de la Educación Superior, No. 49, Vol. 195, 35-54. https://doi.org/10.36857/resu.2020.195.1250, 2020.

2. Bertram, M. Future Skills: How to Prepare Students for the Digital Economy. Routledge, 2019.

3. García, M., Sánchez, L. Factores determinantes en la elección de carrera universitaria entre estudiantes de preparatoria. Revista de Educación. No. 35, Vol. 2, 145-160, 2020.

4. González, M. E., Cuadrado, F. La educación en la era digital: Perfiles y competencias para la sociedad 4.0. Editorial Universitaria, 2018.

5. Jones, A., Brown, L. The impact of digitalization on career preferences among high school students. Journal of Educational Technology, No. 45, Vol 1, 67-80, 2020.

6. Lee, J., Kim, H., Park, S. Predicting career choices of high school students using machine learning algorithms. International Journal of Artificial Intelligence in Education, No. 29, Vol. 3, 323-345, 2019.

7. Real Roby, R. A., Mora Herrera, E. Y., Contreras Moscol, D. F. Hacia un futuro sostenible: El impacto transformador de la tecnología educativa en la educación superior. Revista de Tecnología Educativa, No. 4, Vol 2, 55-72. https://doi.org/10.37549/rte.2024.v4n2a3, 2024.

8. Smith, J., Williams, K., Johnson, L. Women in STEM: Interest and participation trends. Gender Studies Journal, No. 27, Vol. 4, 89-105, 2018.

9. United Nations. Transforming our world: The 2030 Agenda for Sustainable Development. United Nations. Recuperado de https://sdgs.un.org/2030agendaS, 2015.

10. Wang, M. T., Degol, J. L. Gender gaps in STEM career interest: The role of science identity. International Journal of STEM Education, No. 4, Vol.1, 1-16, 2017.

ANEXO: NOTACIONES

Licenciaturas

LNanotecnología: Lic. En Nanotecnología

LLingCompuTecno: Lic. en Lingüística Computacional y Tecnología

Tbiotecnología: TSU en Biotecnología

TInformáticaSalud: TSU en Informática para la Salud

Titecnológicaambient: TSU en Ingeniería tecnológica ambiental y sustentabilidad

TTransformaciónDigi: TSU en Transformación Digital