Identification of association rules between characteristics of female students enrolled in computer science programs at the Facultad Politécnica of the Universidad Nacional de Asunción: A first approach

Ellen L. Méndez Xavier[1,†], Christian von Lücken [2,*,†] and Rita Cantero [3,†]

[1,2,3] *Facultad Politécnica – Universidad Nacional de Asunción, San Lorenzo, Paraguay*

Abstract

The participation of women in science, technology, engineering, and mathematics (STEM) programmes, especially in computer science, remains a challenge and a relevant area of study. Understanding the characteristics, motivations, and backgrounds of female students admitted to these disciplines is important for designing support strategies and encouraging them to remain in these programmes. Therefore, this exploratory study aims to identify association rules among various characteristics of female students admitted (N=30) into informatics programs at the *Facultad Politécnica of the Universidad Nacional de Asunción (FP-UNA)* in the 2024 cohort. Data were collected using a 33-question survey instrument across six dimensions. Subsequently, the FP-Growth algorithm was applied for association rule mining. The findings reveal consistent patterns and direct associations. These preliminary results offer points for discussion on the diversity of incoming students' profiles and demonstrate the usefulness of association rules as a tool for better understanding this student group in the context of the FP-UNA.

Keywords

higher education, women, gender, computing, educational datamining, fpgrowth

1. Introduction

In the current context, where STEM disciplines (Science, Technology, Engineering and Mathematics), and particularly computer science and informatics, have a major impact and offer a wide range of career opportunities, it is essential to understand the background, environment and motivations of women who choose these areas. For them, admission to university in these fields represents both a personal commitment and a step towards greater equity and diversity in the technology sector.

In recent years, several studies have analyzed women's participation in STEM careers in different contexts. According to UNESCO [1], only 35% of students enrolled in STEM-related higher education degrees are women, and differences have been noted across these disciplines. For example, only 3% of female students in higher education choose to study information and communication technologies (ICT). In Paraguay, data from the Ministry of Education and Science [2] indicate that engineering and technology careers account for 7.74% of total university enrolment (men and women), representing 23,587 students. The proportion of women within this group is considerably lower.

[1*] Corresponding author.

[†] These authors contributed equally.

Proceedings XVII Congress of Latin American Women in Computing 2025, October 27–30, 2025, Valparaíso, Chile.

emendez@pol.una.py (E. Mendez); clucken@pol.una.py (C. von Lücken); rcantero@pol.una.py (R. Cantero);
0000-0001-6063-1769 (E. Mendez); 0000-0002-2198-1237 (C. von Lücken)

At the Facultad Politécnica of the Universidad Nacional de Asunción (FP-UNA), available data [3] indicate that between 2010 and 2021, the percentage of women who were admitted to informatics programs through entrance examinations reached a maximum of 31% in 2021, with an average of 24% over those eleven years. No clear trends of sustained increase in these indicators are observed.

Given the relevance of gender equality in technology education, this work aims to better understand the profiles of newly admitted students through a detailed analysis. Specifically, the objective is to identify patterns of association between various characteristics of female students admitted to informatics studies at FP-UNA in 2024, using data mining techniques.

Section II presents the theoretical framework of the study, focusing on educational data mining (EDM) and association rule techniques, as well as the most relevant algorithms. Section III describes the methodology used, including the design of the instrument and aspects of its application. Section IV analyses the results and discusses the main findings. Finally, section V summarizes the conclusions and suggests future research directions.

2. Theoretical framework

It is important to mention three specific points in the theoretical framework of this work. Below is an introduction to the concepts of educational data mining, association rules, and a discussion of their use in STEM work.

2.1. Educational data mining

According to [4], educational data mining (EDM) focuses on developing methods for analyzing data from educational environments. The information obtained through these methods is used to inform decisions aimed at improving these environments.

The choice of methodological approach depends on the type of problem to be solved, allowing multiple algorithms to be applied based on specific objectives. Various systematic reviews, such as those presented in [5] and [6], summarise the main lines of application of EDM. Among these, the models for predicting academic performance [7] and the techniques for grouping students according to their performance [8] stand out.

In the field of STEM and gender, studies such as [9] use text mining and syntactic and semantic analysis methods to highlight differences in motivation, interests, and training applications among students by gender within an engineering faculty in North America.

This paper focuses on identifying relationships among different student characteristics using association rule algorithms.

2.2. Association rules

Association rule methods enable the discovery of frequent patterns between attributes in transactional datasets. These techniques facilitate the identification of relationships relevant to decision-making across various fields, including education.

In [10], an association rule is formally expressed as: $X \rightarrow Y$, where X (antecedent) and Y (consequent) are disjoint sets of items. The rule suggests that the presence of items in X implies a higher probability of finding items in Y, quantified by the confidence of the rule. This allows relationships between variables to be established.

The strength of an association rule can be measured in terms of its support and confidence. Support determines how often a rule applies to a given dataset, while confidence determines how often the elements of Y appear in transactions containing X. The formal definitions of these metrics are [10]:

$$Support,\ \ s(X\to Y) = \frac{\sigma(X\cup Y)}{N}$$

$$Confidence,\ \ c(X\to Y) = \frac{\sigma(X\cup Y)}{\sigma\ (X)}$$

In addition to support and confidence, another important metric is Lift, which is calculated as:

$$Lift\ (X\to Y)\ =\ \frac{c(X\to Y)}{s(Y)}$$

$Lift >$ 1indicates a positive association, i.e., the occurrence of X increases the probability of Y occurring.

The process of searching for association rules consists of two main steps:

1. Detecting frequent item sets, whose occurrence exceeds a minimum number of transactions (minimum support).
2. Obtaining the association rules associated with these frequent item sets, which have a confidence level above the required threshold.

The **Apriori** algorithm identifies frequent item sets in transaction databases and generates association rules. Its primary disadvantage is its high computational cost, as it requires multiple database scans. The **FP-Growth** algorithm is an efficient alternative to traditional association rule mining algorithms by compressing the dataset into a structure known as a Frequent Pattern Tree (FP), which allows the mining of frequent item sets without generating candidate sets

In education, various studies use rule mining to analyze student satisfaction, make predictions in combination with other techniques, or identify common characteristics among students [11, 12, 13]. Due to its greater efficiency, the FP-Growth algorithm has been selected for this work.

3. Methodology

For data collection, a structured survey instrument was designed, consisting of 33 questions distributed across six thematic dimensions. Table I presents each dimension along with the corresponding questions, their numbering within the instrument, and the type of response associated with each item.

To validate the instrument, the expert judgement technique was used [14, 15], with a panel composed of three professionals: two specialists in statistics and one in computer education, who evaluated the items in terms of clarity, relevance, consistency, and significance.

Table 1

Dimensions, questions, and possible responses of the instrument.

Dimension	Questions		Answer options
Identification	1	Can you tell us your age range?	Under 18, 18–20, 20–24, Over 25
	2	What is your gender?	Male, Female
	3	What degree are you studying?	Informatics Engineering, Bachelor of Science in Informatics
	4	Do you consider that you have any specific learning needs or conditions that may influence your educational process?	Yes, No, I prefer not to answer.
Academic background	5	Select your year of admission	2024

Dimension	Questions		Answer options
	6	Can you indicate your type of admission?	Admission test, Direct admission, Transfer from another degree program, Transfer from another campus, Admission by Agreement.
	7	If you entered the program through transfer, can you indicate your original program?	--
	8	What type of school did you come from?	Public, private, subsidised
	9	What type of secondary education did you receive?	Social Sciences, Technical
	10	If you studied a technical secondary education program, was it in Informatics?	Yes, Other
	11	How much time passed between finishing secondary school and starting this degree program?	Less than 1 year, 1-2 years, 3-4 years, 5 or more years.
	12	Did you take any programming courses before entering university?	Yes, at the educational institution; Yes, self-taught; No
	13	Have you participated in events and/or competitions related to technology, mathematics, or science?	Yes. Specify___ No.
Sociodemográficos	14	Can you tell us your area of origin?	Urban, Rural
	15	Can you tell us your city and department of origin?	--
	16	Can you tell us the city and county where you currently live?	--
	17	Can you tell us the educational level of the father/guardian	No education, Primary Education, Secondary Education, Technical Secondary Education, University, Postgraduate, I do not have a father /guardian.
	18	Can you tell us the profession or postgraduate studies of your father/guardian?	--
	19	Can you tell us the educational level of your mother/guardian?	No education, Primary Education, Secondary Education, Technical Secondary Education, University, Postgraduate, I do not have a mother/guardian.
	20	Can you tell us your mother/guardian's profession or postgraduate studies?	--
	21	Do you have siblings who are professionals or students in the same field as you or in similar fields?	Yes, No
	22	Do you have significant family responsibilities that may require some of your time?	Yes, No, I prefer not to answer
Employment status	23	Are you currently employed?	Full-time, Part-time, No
	24	If you are working, is your job related to your career?	Yes, No
Motivation and expectations	25	What was the main reason you chose to study Computer Science?	I am passionate about technology and computing, Good job and salary opportunities, Recommendation from family or friends, It is a career with a future, Other: ___
	26	How long have you been interested in computer science or technology?	Since childhood, Since secondary school, Recently, I'm not really interested
	27	What are your expectations regarding your career choice?	To acquire advanced technical knowledge, To find a well-paid job, To start my own business or venture, To continue with postgraduate studies.
	28	Please indicate your level of interest in these areas.	Very high, high, I don't know, low, not interested.
	29	Is there another area not mentioned that interests you and that you would like to indicate due to its relevance?	--
	30	Did anyone influence your decision to study this degree?	Yes, family members; Yes, friends; Yes, teachers; No, it was my own decision
	31	Did you have any previous experience in programming or technology before entering university?	Yes, at school or college; Yes, self-taught (courses, tutorials, etc.); No, this is my first time
Access to resources	32	Do you have access to a personal computer?	Yes, No
	33	Do you have access to the internet at home?	Yes, No

The instrument was administered to new students in the 2024 cohort of the Informatics Engineering (IIN) and Bachelor of Science in Informatics (LCIk) programs, regardless of the admission method (test, transfer, agreement, or direct admission). Only data from female students was used for the analysis in this study.

Data collection was carried out using an electronic form sent by institutional email, accompanied by clear instructions. The form included information about the study's objectives, voluntary and anonymous participation, and the confidential use of data. Submission of the completed form was considered informed consent. In addition, first-year teachers collaborated in disseminating the instrument in their classrooms. The data were anonymized before analysis to protect the participants' identities.

Figure 1 shows the total female population of 54 students, 19 of whom are enrolled in the IIN program, while 35 are enrolled in the LCIk program. A total of 30 surveys were completed, 16 from the IIN program and 14 from the LCIk program. This represents an overall response rate of 55.56% of the total female intake population.

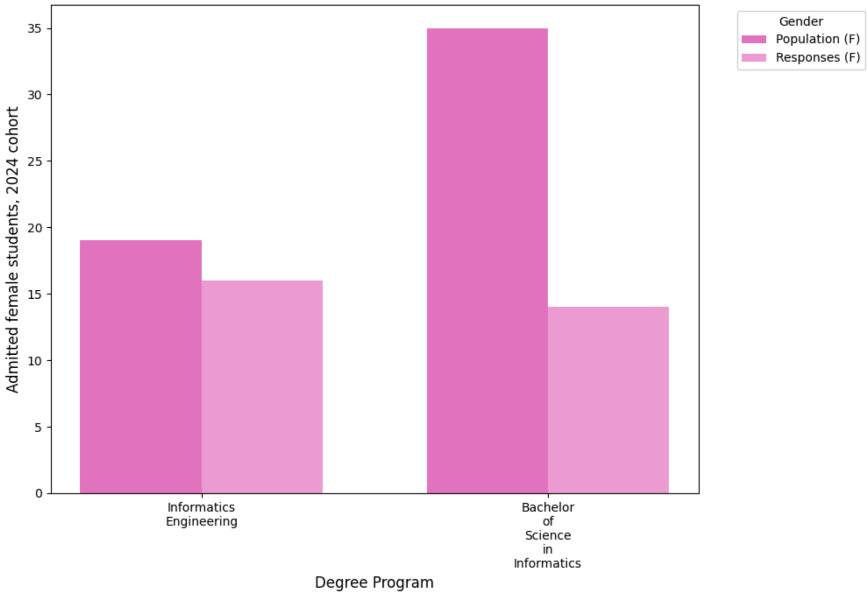

Figure 1: Distribution of the incoming female population (N=54) and valid surveys received (n=30) in the 2024 cohort, by degree program.

For processing the FP-Growth algorithm, we used the implementation of Python's mlxtend module[2], which provides the function for calculating frequent item sets and the function for association rule mining. The survey data was pre-processed to generate a matrix format using One-Hot encoding, a technique that transforms categorical variables into binary format. In this matrix, each row represents a respondent, and each column corresponds to a specific combination of question and answer option; the value is True if that combination is present in the response.

In converting the instrument, each response option to categorical questions was transformed into an individual item following the format 'QuestionID_ResponseOption'. Open-ended questions were manually categorized, generating new categories for the most frequent or significant responses, and question 28, which corresponded to a Likert scale, was converted into a separate item. In the cases of questions 4 and 22, the option 'I prefer not to answer' was considered a valid category.

Given the sample size (N=30), the following parameters and criteria were established for the extraction and selection of rules:

[2] MLXtend: https://rasbt.github.io/mlxtend/

- **Minimum Support** (min_support): This was set at 0.2, which means that a set of items must appear in at least six responses to be considered frequent. This value was defined empirically after preliminary tests, seeking to balance the number of patterns obtained with the manageability of the results in this exploratory study with a small sample.
- **Minimum Confidence** (min_confidence): This was set at 1.0, to identify only perfect implication relationships. This decision, appropriate to the exploratory approach and the small sample size (N=30), was made after iterative parameter adjustment to obtain a manageable and meaningful set of rules.
- **Lift metric**: The lift metric was calculated for all generated rules. Rules with a lift greater than 1.2 are considered, as this indicates a positive association and that the items co-occur more than would be expected by chance.
- **Maximum item length:** The length of the item sets was limited to two elements (max_len = 2) to facilitate the interpretation of the rules and ensure the manageability of the results. This decision responds to the exploratory approach of the study and seeks to ensure greater clarity in the analysis.

4. Results and Discussion

This section presents a general descriptive analysis of the responses obtained.

4.1. General descriptive analysis

Table 2 presents the most relevant responses regarding the general characteristics of new students in general and by degree program. The predominant profile of new students is that of a student between the ages of 18 and 20, coming from an urban area. In terms of secondary education, the majority graduated from technical secondary schools in informatics, with a similar distribution between public and private schools.

Table 2
Frequency of most relevant responses in general and by degree program

Question	Response	General (n=30)	IIN (n=16)	LCIk (n=14)
1 Can you tell us your age range?	18 to 20	24	14	10
4 Do you consider that you have any specific learning needs or conditions that may influence your educational process?	No	22	10	12
8 What types of school do you come from?	Public	16	6	10
9 What types of secondary education did you receive?	Technical	22	11	11
11 How much time has passed between finishing secondary school and starting this degree program?	Less than 1 year	13	7	6
	1 to 2 years	13	8	5
12 Did you take any programming courses before entering university?	No	15	8	7
	Yes	15	8	7
13 Have you participated in events and/or competitions related to technology, mathematics or science?	No	23	12	11
14 Can you tell us your area of origin?	Urban	27	16	11
17 You can tell us the educational level of the father/guardian	Primary Education	1	0	1
	Secondary Education	3	1	2

Question	Response	General (n=30)	IIN (n=16)	LCIk (n=14)
	Technical Secondary Education	7	4	3
	University	13	7	6
	Postgraduate	2	1	1
	I do not have a father /guardian	4	3	1
19 Can you tell us the educational level of your mother/guardian?	Primary Education	2	0	2
	Secondary Education	3	3	0
	Technical Secondary Education	5	4	1
	University	16	8	8
	Postgraduate	3	1	2
	I do not have a mother /guardian	1	0	1
22 Do you have significant family responsibilities that may require some of your time?	No	18	10	8
	I might have	12	6	6
23 Are you currently employed?	No	26	14	12
25 What was the main reason you chose to study Computer Science?	Good job opportunities and salaries	10	8	4
	I am passionate about technology and computing	10	5	5
26 How long have you been interested in computing or technology?	Since secondary school	19	11	8
27 What are your future expectations regarding your career choice?	To find a well-paid job	14	5	9
	Acquire advanced technical knowledge	9	7	2
28 Did anyone influence your decision to study this degree?	No, it was my own decision	13	6	7
	Yes, family members.	11	6	5
29 Did you have any previous experience in programming or technology before entering university?	Yes, at school or college	16	7	9

There is a very similar proportion between those who have taken a programming course and those who have not. Most do not participate in STEM events and/or competitions and are not working when they start their careers.

Interest in computer science usually arises during secondary school or shortly before entering university. The main motivations for choosing this career are associated with the job opportunities and salaries offered by the sector.

All female students were admitted to their degree programs through an admission test. In general, their parents have a university education, and no influence from siblings was identified in their choice of degree program. All female students reported having access to a computer and the internet.

Figure 2 shows the main areas of interest. Among the most popular are Software Development, Video Games, Cybersecurity, Artificial Intelligence, and Machine Learning. The least interesting areas are Infrastructure and Networks, and IT Management. In the management area, there is a higher level of unfamiliarity than in other areas.

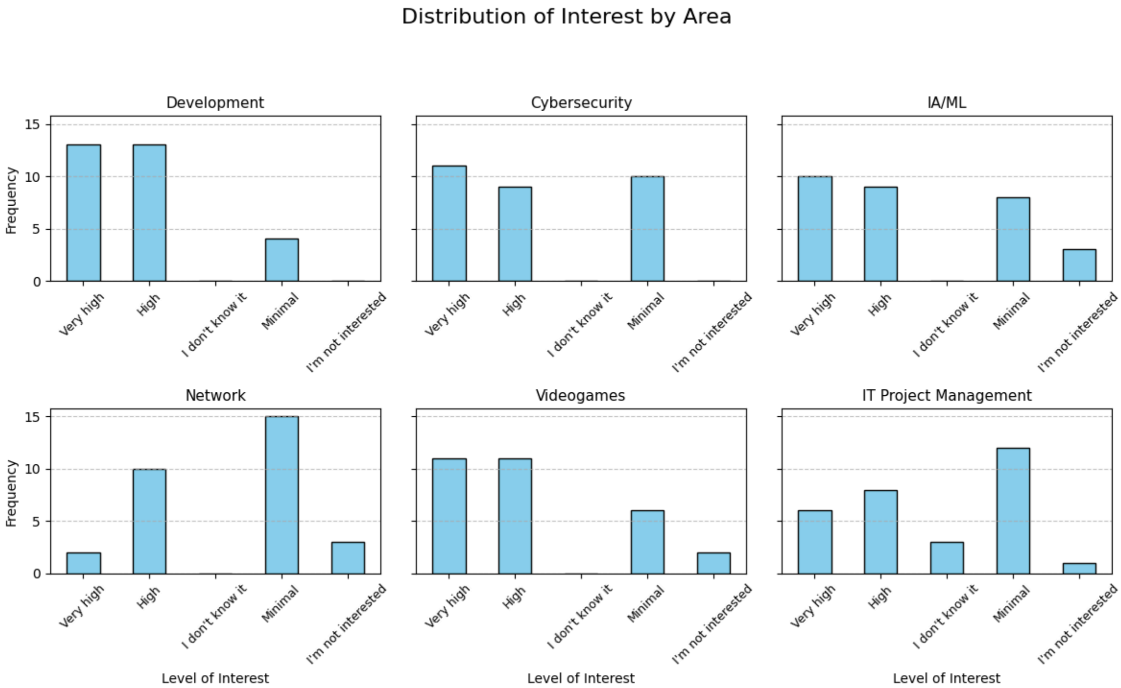

Figure 2: Frequency distribution of interest by area.

4.2. Association rules

For the application of the FP-Growth algorithm, tests were carried out based on the combination of algorithm parameters indicated in the previous section. Twenty significant rules were identified. Table 3 shows the combination of questions and answers resulting in the rules found by the algorithm. For a better layout of the results, this data has been coded.

Table 3
Legend of Item Codes for Association Rules

Code	Question and answer option
I1	If you have participated, can you specify which events and/or competitions?_response_nan
I2	If you studied at a technical secondary school, was it in Computer Science?_answer_YES
I3	If I studied for a technical secondary school, was it in Computer Science?_answer_nan
I4	Do you consider that you have any specific learning needs or conditions that may influence your educational process?_response_NO
I5	What are your future expectations regarding your career choice?_response_TO ACQUIRE ADVANCED TECHNICAL KNOWLEDGE
I6	What type of school did you attend?_response_PUBLIC
I7	How long have you been interested in computing or technology?_answer_SINCE CHILDHOOD
I8	How long have you been interested in computing or technology?_answer_SINCE SECONDARY SCHOOL
I9	Have you participated in events and/or competitions related to technology, mathematics, or science? E.g., Omapa events, science clubs, etc._answer_NO
I10	Have you participated in events and/or competitions related to technology, mathematics, or science? E.g., Omapa events, science clubs, etc._answer_YES
I11	Have you taken any programming courses before entering university?_answer_NO
I12	Have you taken any programming courses before entering university?_answer_YES, SELF-TAUGHT

Code	Question and answer option
I13	Have you taken any programming courses before entering university?_answer_YES, AT AN EDUCATIONAL INSTITUTION
I14	Did anyone influence your decision to study this degree?_answer_NO, IT WAS MY OWN DECISION
I15	Can you tell us the educational level of your parent/guardian?_answer_UNIVERSITY
I16	Can you tell us your level of interest in these areas [IT Project Management]_answer_VERY HIGH
I17	Can you tell us your level of interest in these areas [Artificial Intelligence and Machine Learning]_response_LOW
I18	Can you tell us how interested you are in these areas [Video Games]_response_NOT VERY INTERESTED
I19	Can you tell us your age range?_response_18 TO 20 YEARS OLD
I20	What type of secondary education did you complete?_response_SOCIAL SCIENCES
I21	What type of secondary education did you complete?_response_TECHNICAL
I22	Do you have significant family responsibilities (caring for children, adults, etc.) that may require some of your time?_answer_NO
I23	Did you have any previous experience in programming or technology before entering university?_answer_NO, THIS IS MY FIRST TIME
I24	Did you have any previous experience in programming or technology before entering university?_response_YES, AT SCHOOL OR COLLEGE

Table 4 presents the rules resulting from the application of the FP-Growth algorithm with the previously defined parameters. Twenty significant association rules were identified. In this context, it is important to clarify the behavior of rules (marked in color) that show consistency with the logic of the instrument used:

- **I1 (event specification = nan):** Given rule I1 ⟶ I9 (Rule No. 1 with conf=1.0), I1 is equivalent to I9 (Did not participate in events).
- **I2 (if Technical Secondary School was in Computer Science = YES):** Given rule I2 ⟶ I21 (Rule No. 6 with conf=1.0), I2 is equivalent to I21 (Technical Secondary School).
- **I3 (if Technical Secondary School was in Computer Science = nan):** Given rule I20 ↔ I3 (Rule No. 3 and No. 4 with conf=1.0), I3 is equivalent to I20 (Secondary School in Social Sciences), or more precisely. The question about the specialization of the technical secondary school did not apply because the secondary school was in Social Sciences.

Table 4
Resulting rules with their metrics.

No.	Antecedent	Consequence	Support	Confidence	Lift
1	I1	I9	0,666667	1	1,304348
2	I23	I11	0,333333	1	2
3	I20	I3	0,266667	1	3,75
4	I3	I20	0,266667	1	3,75
5	I7	I14	0,2	1	2,307692
6	I2	I21	0,566667	1	1,363636
7	I24	I21	0,533333	1	1,363636
8	I15	I19	0,433333	1	1,25
9	I10	I19	0,233333	1	1,25
10	I12	I6	0,233333	1	1,875
11	I17	I4	0,266667	1	1,363636
12	I13	I2	0,266667	1	1,764706

No.	Antecedent	Consequence	Support	Confidence	Lift
13	I13	I22	0,266667	1	1,666667
14	I13	I21	0,266667	1	1,363636
15	I5	I1	0,3	1	1,5
16	I5	I9	0,3	1	1,304348
17	I5	I19	0,3	1	1,25
18	I18	I8	0,2	1	1,578947
19	I16	I1	0,2	1	1,5
20	I16	I9	0,2	1	1,304348

A general analysis of the rules reveals the following pattern:

- Rule No. 2: Students who indicated that this is their first exposure to programming or technology (I23) always come from Asunción, the capital (I11). This pattern requires validation with future cohorts, as it could reflect a contextual relationship or sample bias.
- Rule No. 5: Students who expressed an interest in computer science or technology since childhood (I7) always indicated that they decided to study the degree on their own initiative, without external influences (I14). The high Lift suggests that a very early interest in the area encourages greater autonomy in choosing a degree.
- Rule No 7, 12, 14: Those who had previous experience in programming or technology at their school (I24) took a technical secondary school (I21), and students who took a programming course at an educational institution before university (I13) also indicated that their technical secondary school was in informatics (I2). This indicates that programming experience acquired formally within secondary educational institutions is linked to technical secondary schools, mainly in informatics.
- Rule No. 8: Students whose father/guardian has a university education (I15) are, in all cases, in the 18-20 age range (I19). In other words, the youngest group of entrants tends to have parents with a university education. This could be related to a family culture that promotes early access to higher education, although further exploration is needed to confirm this relationship.
- Rule No. 10: Students who taught themselves programming before university (I12) always come from public schools (I6). This pattern could be related to the search for ways to supplement limited formal resources.
- Rule No 15 and 16: Students whose main expectation is to acquire advanced technical knowledge (I5) always indicated that they had not participated in technological/scientific events before university (I9, and consequently I1). Those with a strong orientation towards advanced technical learning did not necessarily have a history of participation in events.
- Rule No 18: This rule outlines a subgroup of students whose interest in technology was consolidated during their secondary school years (I8), but which does not extend strongly to the field of video games (I18). This suggests that interest in computer science is not necessarily linked to recreational use of technology, and that it can develop from more academic, professional or practical motivations.
- Rule No 19 and 20: Students who expressed a 'Very High' interest in IT Project Management (I16) always indicated that they did NOT participate in events and/or competitions related to technology, mathematics or science (I9) before entering university. This association could serve to identify a student profile with a penchant for organization,

strategic planning and leadership in technological contexts, rather than competitive or technical-practical activities.

4.3. Limitations of the results

The main limitation is the small sample size (N=30). Although this sample allowed for a detailed exploratory analysis and the identification of 'perfect' associations within this specific group, the results are not statistically generalizable to the total universe of female entrants to the FP-UNA or other institutions.

The associations found, although they meet the strict metric criteria established (confidence=1.0, lift>1.2), should be considered as initial hypotheses that reflect strong patterns within the 2024 cohort.

Likewise, the sample size directly influenced the choice of algorithm parameters, requiring a restrictive approach (e.g., max_len=2, min_confidence=1.0) to obtain a manageable set of results, which may have left valid rules out of the analysis.

5. Conclusions and future work

This paper constitutes a first exploratory approach to identifying association rules between various characteristics of female students who enrolled in 2024 in informatics programs at the FP-UNA. Based on the analysis carried out, the following findings stand out:

- Regarding academic background: patterns linked to academic background were identified; for example, a perfect association between having completed a technical secondary school and previous experience in programming or technology, and a relationship between a secondary school in social sciences and non-participation in STEM events before university.
- Regarding previous experiences and motivations: it was observed that certain previous experiences, such as self-taught programming, are associated in the sample with coming from public schools. Likewise, an early interest in technology (from childhood) was consistently linked to an autonomous decision to choose a career.
- Regarding specific profiles: profiles were identified, such as female students who were interacting with technology for the first time but had developed an interest in secondary school, or those with a very high interest in IT Project Management who had not participated in previous STEM events.

As an initial exploratory study, these findings constitute a first approximation of the interrelationships between characteristics within this specific group of students. They provide guidance for future; more in-depth and far-reaching research focused on the characteristics and needs of female students in computer science programs at FP-UNA. It is hoped that the instrument will be applied to new cohorts of admissions, to increase the number of surveys. This would allow the rules found in this initial approach to be validated and, possibly, less restrictive support and confidence thresholds to be applied. Similarly, it would be important to include male students in the analysis to conduct comparative studies of profiles and gender-specific patterns of association.

Declaration on Generative AI

During the preparation of this work, the authors used Gemini and Grammarly in order to: Grammar and spelling check. After using these tools/services, the authors reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] United Nations Educational, Scientific and Cultural Organization. (2019). Descifrar el código: la educación de las niñas y las mujeres en ciencias, tecnología, ingeniería y matemáticas (STEM).

[2] ABC Color. (6 de diciembre de 2024). Estas son las carreras con más universitarios en Paraguay. ABC Color. https://www.abc.com.py/nacionales/2024/12/06/estas-son-las-carreras-con-mas-universitarios-en-paraguay/. Consultado el 2025/05/11

[3] Xavier, E. L. M., von Lücken, C., & Cantero, R. (2022). Applications, Admissions and Graduations of Women in Computer Science Carrers for the Universidad Nacional de Asuncion. In LAWCC@ CLEI (pp. 12-19).

[4] Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.). (2010). Handbook of educational data mining. CRC press.

[5] Dutt, A., Ismail, M. A., & Herawan, T. (2017). A systematic review on educational data mining. *Ieee Access*, *5*, 15991-16005.

[6] Afif, F. E., Bouyahia, F., & Rafouk, L. (2025). Educational Big Data and Data Mining: A Systematic Literature Review. Industrial and Logistics Systems Design and Efficient Operation: Cyber Physical Production and Logistics Systems, 1332, 191.

[7] Dol, S. M., & Jawandhiya, P. M. (2024). Systematic review and analysis of EDM for predicting the academic performance of students. *Journal of The Institution of Engineers (India): Series B*, *105*(4), 1021-1071.

[8] Sultan Alalawi, S. J., Mohd Shaharanee, I. N., & Mohd Jamil, J. (2023). Clustering student performance data using k-means algorithms. *Journal of Computational Innovation and Analytics (JCIA)*, *2*(1), 41-55.

[9] Chopra, S., Gautreau, H., Khan, A., Mirsafian, M., & Golab, L. (2018, July). Gender differences in undergraduate engineering applicants: A text mining approach. International Educational Data Mining Society. In International Conference on Educational Data Mining.

[10] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, et al. *Introduction to data mining*, volume 1. Pearson Addison Wesley Boston, 2006.

[11] Ruiz, S., & Herrera, M. (2019). Variables asociadas a la satisfacción de alumnos mediante técnicas de minería de reglas de asociación. In *XIV Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2019), (Universidad Nacional de San Luis, 1 y 2 de julio de 2019).*

[12] Alangari, N., & Alturki, R. (2020). Association rule mining in higher education: A case study of computer science students. *Smart Infrastructure and Applications: Foundations for Smarter Cities and Societies*, 311-328.

[13] Dol, S. M., & Jawandhiya, P. M. (2023). Classification technique and its combination with clustering and association rule mining in educational data mining—A survey. *Engineering applications of artificial intelligence*, *122*, 106071.

[14] Lynn, M. R. (1986). Determination and quantification of content validity. Nursing research, 35(6), 382-386.

[15] Polit, D. F., & Beck, C. T. (2008). Nursing research: Generating and assessing evidence for nursing practice. Lippincott Williams & Wilkins.