Received: 5 July 2023 | Revised: 20 March 2024 | Accepted: 20 April 2024

DOI: 10.1002/cae.22751

RESEARCH ARTICLE

WILEY

Identifying beliefs about the gender gap in engineering professions among university students using community detection algorithms and statistical analysis

Noemi Merayo[1] | **Alba Ayuso**[2]

[1]Department of Signal and Communications Theory and Telematics Engineering, School of Telecommunications Engineering, Universidad de Valladolid, Valladolid, Spain

[2]Department of Pedagogy, Faculty of Medicine, Universidad de Valladolid, Valladolid, Spain

Correspondence
Noemi Merayo, Department of Signal and Communications Theory and Telematics Engineering, School of Telecommunications Engineering, Universidad de Valladolid, Valladolid 47011, Spain.
Email: noemer@uva.es

Funding information
Universidad de Valladolid

Abstract

Digital societies require professionals in the Technology and Engineering sectors, but their lack, particularly of women, requires a thorough understanding of this gender gap. This research analyzes the beliefs and opinions of university engineering students about the gender gap in their professional fields by means of a community detection algorithm to identify groups of students with similar belief patterns. This study leverages a community detection algorithm to analyze the beliefs of 590 engineering students regarding the gender gap in their field, together with a correlational and explanatory design using a quantitative paradigm. A validated questionnaire focusing on the professional dimension was used. The algorithm identified three student communities, two gender-sensitive and one gender-insensitive. The study uncovered a concerning lack of awareness regarding the gender gap among engineering students. Many participants did not recognize the importance of increasing the representation of professional women, maintained the belief that the gender gap affects only women, and assumed that men and women are equally paid. However, women show a higher level of awareness, while men perceive the gender gap as a passing trend, which is worrying. Students recognize the importance of integrating a gender perspective into university and engineering curricula. It is worrying that many students doubt the existence of the gender gap and that both genders lack knowledge about gender gap issues. Finally, community detection algorithms could efficiently and automatically analyze gender gap issues or other unrelated topics.

KEYWORDS

community detection algorithms, engineering education, engineering workforce, gender equality, professional dimension, STEM

1 | INTRODUCTION

The world is moving at a frenetic pace toward a digital society in which new technologies and the Internet will not only influence the economy but also our daily lives and social environments. This whole process implies a strong impact on the interaction and convergence between the physical and the digital world, incorporating new concepts and trends: Artificial Intelligence (AI), Big Data, Cybersecurity, Cloud Computing, Internet of Things (IoT), or the Fourth Industrial Revolution, so professionals are highly required in these areas fully connected with STEM (Science, Technology, Engineering, and Math) sectors [36]. In fact, the European Union forecasts a figure of more than 1.75 million new jobs in STEM sectors by 2030 [28]. The USA [84] estimates that employment in STEM will increase by 10.5% from 2020 to 2030, which means more than one million jobs. According to the Organisation for Economic Cooperation and Development [62], by 2030, 80% of jobs worldwide will be related to STEM sectors.

Despite the high demand, there is a scarcity of STEM professionals, particularly in engineering and technology, which is a major concern, especially for women. According to projections, the USA is expected to have 3.5 million STEM jobs by 2025, with two million of them remaining vacant [25]. By 2026, the scarcity of engineers is estimated at 1.2 million, with a loss of 545.000 software developers. Similarly, Germany, Poland, Hungary, and Croatia are also facing significant shortages, with Germany expecting a reduction of four million specialists by 2030. Furthermore, Switzerland will require over 117.900 STEM specialists, and in Australia, companies will need more than 520.000 specialists by 2026. [24].

On the other hand, the latest EU report indicates a persistent gender gap in digital skills among professionals [27], with only 19% of technology professionals and approximately one-third of STEM graduates being women. Little progress has been observed as these figures have remained stable in recent years. In Europe, men in the technology sector earn 19% more than women and account for more than 80% of these specialists on average. Currently, women account for 29.4% of junior (fresh out of university) positions, but this figure drops sharply to between 12.4% and 17.8% in senior management positions. UK statistics from 2022 show that women represent only 16.5% of engineers, a mere 6 percentage point increase since 2010 [89]. Similar trends can be seen in the United States, where female engineers account for only 14%, and female computer and mathematical engineers make up around 26% [54]. Australia also mirrors these figures, with approximately 28% of women working in STEM and a gender pay gap of around 18% [6]. The World Economic Forum's Global Gender Gap Report [92] reveals that it will take 132 years to achieve full parity. Technology and engineering fields show the most significant gender disparities, yet there is a substantial economic impact of excluding women from STEM, particularly in the digital and technology sectors. Closing the gender gap in STEM is estimated to increase GDP per capita by 2.2% to 3.0% by 2050 in the EU European Institute for Gender Equality [29]. Consequently, countries are implementing strategies to promote the integration of women professionals in STEM, especially in engineering and technology. Therefore, women's participation in STEM is indispensable for a future diverse, inclusive, fair, and sustainable digital economy and society. Thus, there is a need to guess the causes and the time in school when girls lose motivation for STEM disciplines, as this will generate a huge waste of human resources that will have a long-term economic and social impact. Indeed, this decline of professionals and students in STEM sectors also persists in Spain, with increasingly worrying data [47, 51, 62]. Under this global scenario, schools have a vital role to play, due to the great need to foster motivation in STEM, especially among girls. Many studies have focused for many years on examining the factors and causes contributing to the decrease in girls' and young women's interest in STEM studies during secondary school [50, 8, 35, 43, 48, 55, 63, 72, 78, 85]. At the university level, there is also a significant number of research studies [4, 5, 59, 13, 17, 32, 38, 57, 64, 70, 86, 90], but few focus on describing the beliefs and opinions of university students regarding the gender gap in STEM sectors [65]. But despite all these studies, the existence of a gender gap in STEM university studies remains a problem that much research around the world is still trying to address and mitigate today.

On the other hand, the integration of AI methodologies, like community detection algorithms, to analyze the gender gap has been relatively unexplored. These algorithms offer advantages by identifying homogeneous groups with similar tastes or beliefs and extracting groups based on variables. Utilizing these technologies to automatically analyze students' beliefs, motivations, and opinions regarding the gender gap in STEM Higher Education represents an innovative approach that holds promise for other related subjects. Taking all of the above into account, the following research questions are posed:

- Are there communities of students who share similar beliefs regarding the gender gap in their professional fields?
- What do students think about the gender gap in their professional fields?

To this end, the following hypotheses are being pursued in this study:

- There exists a correlation among STEM students' perceptions of the gender gap in their professional fields.
- STEM degree students' perspectives on the gender gap in their professional fields vary based on gender.
- Various communities within the STEM student population hold distinct views on the gender gap in their professional fields.

Consequently, the main objectives of this study are:

- To conduct an analysis of STEM undergraduates' beliefs regarding the gender gap in their professional fields, with particular emphasis on gender-specific perspectives.
- To model and integrate community detection algorithms to identify patterns in the beliefs of STEM undergraduates concerning the gender gap in their professional fields.

The results of this research are partial findings of a broader study, which covers both the professional and educational dimensions, but in this paper, we focus on the professional domain. This paper is organized as follows. Section 2 shows the background. Section 3 explains the methodology and sample. Section 3 describes the community detection algorithm. Section 4 shows the main results and Section 5 the discussions. Section 6 provides the conclusions.

2 | BACKGROUND

2.1 | Analysis of gender gap in STEM university education

The existence of a gender gap in STEM university studies continues to be a problem that many research studies around the world have been trying to address and mitigate for many years. Studies show that young women lack confidence, face gender bias, and have limited identification with engineering degrees. Family and societal expectations play a role in their career choices. Moreover, integrating a gender perspective into university curricula is necessary. In this way, retaining minority students, like black women in engineering, is a specific focus [70]. Furthermore, gender differences in engineering students' self-efficacy reveal disparities despite similar grades [57] and gender bias persists even when women outperform men [2, 17]. Indeed, studies report a gender gap in engineering self-efficacy, with men reporting higher means [13, 64]. Recommendations include the incorporation of humanistic and nontechnical content in engineering curricula [64], the integration of soft robotics to improve self-efficacy and mitigate gender differences [38], or the integration of flipped classroom models [18], which can help improve the grades of engineering students, especially female students. Thus, more attention is paid to gender equality policy interventions and mainstreaming in curricula [86]. Besides, female students show less affinity toward professional engineering practices [34, 64] and detect a gender gap in entrepreneurial attitudes [4]. In this way, teacher support for women in STEM degrees is relatively low [38], while family, personality, and expectations strongly influence their career decisions [55, 81]. Finally, in the systematic review conducted by Msanbwa and colleagues (165 scholarly publications) [53], they explained that women were reluctant to participate in STEM due to various factors. About 10% attributed it to personal factors, 61% to environmental factors, and 29% to behavioral factors. The latter included negative attitudes, lack of career expectations, diminished interest, low self-concept, low self-efficacy, and low motivation, all of which were identified as factors influencing girls' low participation in STEM. In fact, other studies suggest that the inequality between men and women in terms of pay or employment status in these professions is significant from the time of entry into the wage market [93].

According to the last World Economic Forum's Global Gender Gap report 2023 [92] (146 countries covered), the gender gap is most pronounced in the engineering and technology sectors. In fact, the percentage of women graduates in technological degrees is 1.7%, compared to 8.2% of men. In Engineering, the same figures are given, with 24.6% of men and 6.6% of women. In fact, based on statistics regarding STEM education in 2023 [75], women represent less than 30% of researchers, indicating an underrepresentation of women in STEM fields worldwide and across Europe. Gender segregation is prevalent in the technology and engineering sectors, where approximately 82% of students are male. This trend is reinforced by the DESI 2022 report [26], which highlights that only one in three STEM graduates and a mere 19% of ICT (Information and Communication Technology) specialists are female. As a result, men comprise over 80% of these specialists on average in Europe. However, after graduating in STEM fields, there is a noticeable drop in women's presence within the workforce, with representation declining significantly within 1 year. Currently, women represent 29.4% of entry-level positions, but this figure decreases considerably to between 12.4% and 17.8% for more senior roles. In this regard, the statistics in Spain are equally concerning, as indicated by the latest report on women scientists published by the Spanish Government [51]. This report highlights significant gender imbalances in engineering studies, with only 25% of graduates being women. In fact, the most recent comparative data from the Spanish Ministry's “White Paper on Women in Technology” [47] and the recent report on the gender gap in STEAM by the Spanish Ministry of Education [18] demonstrate a decline in the number of women pursuing STEM studies. The sharpest decline is in Computer Science, with a loss of around 20% of women between 2002/2003 (around 31% female students) and 2019/2020 (12% female students). Another significant drop is in Mathematics, with a loss of 14% of women between 2004/2005 (50%) and 2019/2020 (around 36%). Moreover, the percentage of women also decreased in Telecommunications Engineering by around 3% between 2003/2004 (25%) and 2019/2020 (22%) and around 1% in Aeronautical Engineering degrees in the same period. However, it is important to point out that although the problem is currently very relevant, it has been persistent in Spain for years, as corroborated by figures from different research studies. In fact, there are studies from the 1980s, such as Braizan and colleagues' research [9], where the authors concluded that women in the computer science and engineering sector working in private companies felt discriminated against in terms of economic disadvantage, level of responsibility, and promotion prospects compared to their male counterparts. In this way, Pérez-Artieda et al. [66] have analyzed the enrollment of women in engineering degrees at the Public University of Navarra over time. The study revealed that since 1996, the overall percentage of women in all engineering programs at this university has remained below 25%. Besides, the qualitative research conducted among their students demonstrated the strong influence of the family on academic decision-making and suggested the inclusion of role models of female engineers to inspire young women. Another research at the University of the Basque Country for the decade 1995–2005 revealed that the percentage of females in Computer Engineering has decreased [60], and they suggest integrating actions that involve providing more information to high school students, explaining the role of engineers in society, and adopting a proactive and informative attitude towards these disciplines within society. In this way, in the data collected by the Gender Equality Units of the universities of Galicia (between 1997 and 2008), gender segregation in the presence of women in engineering degrees was observed; this figure has remained constant over time [3]. Indeed, a specific survey on electrical engineering degrees (in Vigo) highlighted the absence of promotion policies in the Spanish education system and recommended actions such as improving the image of the profession, introducing promotion schemes in preuniversity studies, providing training for educators and disseminating information on the history of women in engineering. However, this decrease in female representation does not imply a lower academic performance, as a study carried out in aerospace engineering at the University of Valencia showed that there was no gender gap in academic performance [52]. Moreover, Ramirez and colleagues' study [68] demonstrated that in Western countries, including Spain, the number of women enrolled in STEM fields in higher education increased by 20% between 1970 and 2000 but remained steady from 2000 to 2010. Considering this previous analysis, it can be inferred that gender disparity persists in STEM fields, particularly pronounced in the field of engineering. Therefore, continuing research in these domains to mitigate their impacts in an increasingly digitized society is of crucial importance for the progress of these societies.

2.2 | Analysis of community detection algorithms in STEM education

Community detection algorithms in STEM have not been applied to a large extent, but offer advantages when working with large data, enabling the discovery of communities and extracting groups based on different factors. They aid in analyzing and visualizing complex data, uncovering patterns difficult to identify otherwise. However, they can be computationally demanding for very large networks, and dynamic communities pose challenges [73, 97]. In higher education, some community detection algorithms have focused on predicting student performance or drop-out rates. The work of Adraoui et al. [1] detected at-risk learning communities, aiding the development of educational resource recommendation systems, while Wang and Wang [88] proposed a community detection algorithm to predict student dropout and cooperative learning interactions. Iam-On and Boongoen [37] used K-means clustering to identify dropout patterns in learning networks based on grading profiles. Regarding recommended systems in higher education, Khaled et al. [39] developed recommender systems to tailor learning content and resources to individual learners' needs. Besides, Mahnane [42] proposed data mining algorithms for recommending learning activities based on learning styles. Senthil-Kumaran et al. [40] utilized similarity-based clustering techniques to suggest courses to students based on skills and interactions. On the other hand, other studies are focused on analyzing learning processes in Massive Open Online Courses (MOOCs). Wise et al. propose community detection to understand learning through discussion forums in MOOCs by examining social interactions and relationships [91]. Besides, Sun and Bin [79] propose an interaction model to solve the problem of the learners' capability gap. The algorithm effectively classifies learners' abilities by analyzing information about their behavior on MOOC platforms. Other researchers [94] investigate student engagement with a video learning repository and employ community detection algorithms to identify behavioral patterns among students. The findings illustrate that algorithms can successfully identify and assess learning communities that may not be directly tied to a specific topic. Finally, on the assessment of student behavior, Maldonado-Mahauad et al. [45] applied clustering to group students into profiles based on their behavior. Wang and Zhang [87] implemented K-means clustering strategies to detect groups of students with different behavioral characteristics in blended learning courses. In the same context, Mengoni et al. [49] apply an algorithm to find communities of learners based on their interactions in virtual environments. In terms of targeted research on community detection in STEM, Yuen and Pickering [95] put forth a system designed to analyze the characteristics of STEM education communities on Twitter. The community detection algorithm offers potential benefits for STEM teachers, enabling them to enhance their interactions on Twitter and fostering a more robust and collaborative STEM education community among teachers and students. As can be seen from the previous analysis, it can be concluded that there are few proposals that apply community detection algorithms in higher education. Indeed, to the best of our knowledge, no community detection algorithms have been applied to analyze the gender gap among STEM university students, so our approach is particularly novel and powerful and can be applied to many different topics.

3 | METHODOLOGY AND SAMPLE CHARACTERIZATION

A descriptive correlational and explanatory design is carried out from a quantitative paradigm [14].

3.1 | Participants

Inclusion criteria for participant selection encompassed undergraduate students enrolled in engineering programs at the specific university where the study was conducted, with no upper age limit. Participation was strictly voluntary, ensuring that all participants provided informed consent. Exclusion criteria involved students from nonengineering fields and those who did not provide informed consent. The intentional sample consisted of 590 students ($x = 20.26$; $\sigma = 2.014$) from the University of X, aged 18–35. Participants were selected from specific engineering degrees: 32.7% from Industrial engineering, 27.7% from Telecommunications engineering, and 39.6% from Computer Science. Students voluntarily participated in the questionnaire. Year-wise distribution: 14% first year, 34.8% second year, 24.3% third year, 25.2% fourth year, and 1.7% master's degrees. Gender distribution: 71.2% men, 25% women, and 3.7% nonbinary or preferred not to answer.

3.2 | Data collection and analysis procedure

These results are part of a broader study focusing on the professional sphere, while the questionnaire used was adapted from García-Holgado et al. [33] for generic STEM degrees. The instrument's reliability was measured using Cronbach's α coefficient, resulting in a value of 0.778, indicating good reliability (fairly high according to studies indicating good values between 0.76 and 0.95) [11, 22, 58, 80]. The questionnaire was administered anonymously via Microsoft Forms.

The professional dimension consisted of 8 Likert-type questions (rated from 1 to 5; 1 representing total disagreement and 5 representing total agreement):

- Q1. Gender equality is an important issue that needs to be tackled at all levels (family, education, social, and professional).
- Q2. Women have more problems than men in technical tasks in engineering/technology.
- Q3. Men are better prepared than women for jobs in Engineering/Technology.
- Q4. Women have more problems than men in finding jobs in Engineering/Technology.
- Q5. Men and women are paid the same for similar positions in Engineering/Technology.
- Q6. More women professionals are needed in Engineering/Technology.
- Q7. The gender gap is a passing fad.
- Q8. Engineering/Technology professionals must help to reduce the gender gap in their sector.
- Q9. The gender gap is a problem that only affects women.

The above questions in the professional field have been correlated with others that belong to the educational dimension and are of particular interest because of their close connection. Specifically, these questions are:

- Q10. Gender equality must be part of university curricula.
- Q11. Gender gap is not a problem that needs to be addressed specifically in engineering studies.

- Q12. Gender influences the completion of engineering studies.

All survey responses have been collected and analyzed with the statistical package SPSS 28.0 for Windows. In all tests, a confidence level of 95% was established. First, an analysis was carried out with the entire population of this study. Frequencies, means, and percentages were established for each variable. Apart from conducting descriptive statistics, the performed tests included Pearson's correlation to assess the strength correlations among the Likert-type variables. Furthermore, the population was divided into female and male student groups to explore gender-specific differences more comprehensively. A descriptive analysis of means and standard deviation was conducted for each group. Subsequently, the Kolmogorov–Smirnov normality test was employed to determine the normal distribution of the sample. Table 1 shows the results of the Kolmogorov-Smirnov, which indicate significant differences between the variables and their distribution.

Taking into account the results from Table 1, non-parametric tests such as the Mann–Whitney test were chosen. This test was applied to identify any disparities between the medians of the variables in the two groups. Additionally, the Louvain community detection algorithm was utilized to understand the relationships and links between university students within large data sets expressed in graphs. This algorithm is an iterative method to identify community structures in complex networks, maximizing modularity (how strongly a network can be divided into robust groups). Its main advantages include its computational efficiency and the ability to handle large data sets, as well as its capacity to detect hierarchical and overlapping communities. These sophisticated, machine-learning-oriented techniques help define and cluster communities of data that would not be able to detect with the naked eye.

TABLE 1 Kolmogorov–Smirnov normality test.

	Kolmogorov–Smirnov	
	Statistic	**Sig.**
Q1	0.381	0.000
Q2	0.408	0.000
Q3	0.375	0.000
Q4	0.211	0.000
Q5	0.192	0.000
Q6	0.214	0.000
Q7	0.195	0.000
Q8	0.245	0.000
Q9	0.221	0.000

To better describe each community of students and to deepen the results of the algorithm, a descriptive statistical analysis was performed with the mean and standard deviation of each community. Finally, the Bonferroni test with post hoc analysis was used to contrast the communities.

3.3 | Case study: Modeling and application of the Louvain community detection algorithm to a gender questionnaire

Community detection is a hot topic in modern network science due to the rise of large network data sets and the impact of networks on our lives [12]. These algorithms play an essential role in the graphical analysis of data, as graphs allow different types of data to be stored within a structure to understand their relationships across multiple vertices. Thus, communities are clusters of vertices that are more likely to be connected to each other than to members of other clusters. Hence, the detection of communities is closely linked to visualization, as it assists users in data analysis and enhances the identification of patterns. Visual or qualitative analysis of communities in data networks effectively supplements statistical (quantitative) analysis, as users interact with visual representations and can identify patterns that might be challenging to discern using other techniques. Contrary to clustering, which classifies data sets based on similarities and dissimilarities between their data, community detection algorithms are used to identify subcommunities within a graph of links. Thus, clustering algorithms, such as K-means or Hierarchical clustering [15] or clustering analysis performed by programs such as the SPSS statistical package, attempt to group together the objects that share the same characteristics, while community detection tries to find communities of closely connected and less densely interconnected nodes [46]. In this way, community detection is more suitable for understanding and analyzing the structure of large and complex networks, which depend on a single attribute type called edges. In contrast, clustering algorithms have a tendency to separate single peripheral nodes from the communities to which they should belong and, in addition, require some specific parameters to be specified in advance, such as the number of clusters [67]. Consequently, community detection uses the properties of edges in graphs or networks and is therefore more suitable for network analysis than a clustering approach. Therefore, as this research aims to identify the structure of communities in a very complex network and without prefixing some characteristics, it was decided to use community detection techniques.

There are several types of community detection techniques such as divisive, agglomeration, or optimization algorithms [12]. The quality of community partitions resulting from these methods is usually measured by the modularity of the partition. Modularity measures the density of links within the community compared to links between communities, that is, it measures how strongly a network can be divided into groups, clusters, or disjoint communities. A division into communities is strong when the nodes in each module are well connected to each other. One of the first highly efficient modularity-based algorithms was Newman's method, later improved by Clauset [19] who made changes to the way modularity between groups was calculated. Nevertheless, Blondel's algorithm, known as Louvain's algorithm, performs better than Clauset's algorithm in terms of maximum modularity. This algorithm is one of the most efficient, fast, and accurate algorithms for community detection [82].

The Louvain algorithm is a modularity optimization algorithm able to extract the community structure of a network and its target is to maximize the modularity of the data. It is a greedy, hierarchical, and agglomerative algorithm, so at the beginning, each node is considered an independent community. To find the best partition of the network, the algorithm heuristically groups these nodes into communities so that modularity is maximized [21]. Modularity is defined as a value that measures the density of links within the community compared to links between communities (Equation 1).

$$Q \quad (P) = \frac{1}{2m}\sum\left(A_{ij} - \frac{d_i d_j}{2m}\right)\sigma(C_i, C_j), \tag{1}$$

where:

- A_{ij}: This is the adjacency matrix of the network, the weight of the connecting edge between nodes i and j,
- d_i: degree of node I,
- d_j: degree of node j,
- m: Number of edges or links,
- Function $\sigma(C_i, C_j)$ is 1 if nodes i and j are in the same community ($C_i = C_j$) and 0 otherwise.

The Louvain algorithm works in two main phases, the first one forms the partitions, and the second one processes the obtained communities to identify their hierarchical relationships. The Louvain algorithm works in the following phases [20, 58]:

1. Phase 1. Initially, each node is assigned a unique community, so that the total number of nodes is equal to the total number of unique communities.
2. Phase 2. Iteratively, each node "i" is assigned to the community of its neighboring node "j" and the modularity of the network is recalculated. If the modularity improves compared to when node "i" was not in the community of node "j", node "i" will be assigned to the same community to which node "j" belongs, otherwise not. In the first case, node "j" shall be a neighbor of node "i."

This will be repeated until no further gain in modularity is observed when moving any node to its neighbor's community and we have reached a maximum of modularity.

The case study in which the Louvain algorithm is performed is based on the proposed questionnaire, that is, questions described in Section 3.2, to automatically infer communities that follow similar patterns of thinking within a network of STEM university students. Louvain is an unsupervised algorithm, so it is not necessary to specify the number of communities or the size of the communities. The algorithm will build a network with nodes and edges. A network consists of nodes that represent individuals, people, or things, and edges represent the connection or relationship between nodes. In this case, nodes represent students and edges represent the relationship between these students, so edges join two nodes to indicate a relationship. Thus, Louvain groups similar nodes based on edge weights. In our case study, the criteria used to calculate these edge weights is defined as the number of questions where two students fully agree, that is, where they have given the same Likert score for a question. Thus, each time both agree on a question, a value of 1 (in scalar) is added to the edge weight. For example, if two students give the same score to three questions, that is, they agree on three questions, the associated joint edge weight for those students is 3. All questions in the questionnaire are equivalent to this calculation, that is, they all have the same weight for the computation. Then, Louvain randomly orders all the students to a graph in the modularity optimization phase. It then removes and inserts each student into a different community until there is no significant increase in modularity. Networks with high modularity have dense connections between nodes within each community, but sparse connections between nodes in different communities. Modularity is then understood as the fraction of edges connecting students of the same community, that is, the level of agreement of all students belonging to the same community. Finally, our Louvain model was implemented in Python using the Community [56] and the Networkx libraries [20].

4 | RESULTS

This section is divided into two subsections: one analyzing STEM students' opinions on the gender gap in professional environments, and the other focusing on developing a community detection algorithm to identify behavioral and belief patterns regarding the gender gap.

4.1 | Results of the Louvain community detection algorithm

The proposed Louvain algorithm allows us to visualize different groups of students in the data set of this study. In fact, Figure 1 shows the existence of three distinct communities, obtained from the optimal, that is, highest, modularity. When comparing modularity optimization methods, the most important quality measures are the speed and the modularity value [21, 41, 96]. Greater speed indicates higher efficiency of a method compared to others, making it preferable. Similarly, a higher modularity value is desirable as it signifies well-defined communities within a network. Modularity measures the extent to which a network can be partitioned into distinct and cohesive communities. A strong division into modules is characterized by well-connected nodes within each module and fewer connections with nodes outside the module. Modularity values range from −1 to 1, with values around 0.3 or higher indicating strong modularity and the presence of strong communities [21, 96].

In this way, the resulting modularity value in our algorithm is 0.29, so the performance of our proposal is quite good. In terms of speed, our proposal can find communities in around 1 min, so it can be said that the algorithm is fast and efficient. In fact, Louvain obtains good results in terms of computational efficiency and time complexity compared to other similar methods such as Girvan-Newman [21, 82, 96], proving to be very fast in providing optimal partitions in large networks (its algorithmic order is $O(n*\log(n))$ [21]. Furthermore, the communities are quite homogeneous in terms of the number of students. In fact, Community A has 209 students (51.67% men, 48.33% women), Community B 211 students (78.67% men, 21.33% women), and Community C 170 students (89.41% men, 10.59% women), a slightly smaller number of students. Moreover, the communities have also been analyzed according to several variables, specifically gender, type of degree, and year of study. As for the type of engineering (Telecommunications, Industrial, Computer Science), there are no differences in the percentage of students belonging to each community (around 30% in each community). The same occurs with the year to which the students belong, since the percentage of students of each year inside each community is homogeneous. Finally, the most female students are found in Community A, followed by Community B, with Community C being the community with the lowest number of female students. Once the final graph with the communities has been obtained, a descriptive and comparative statistical analysis of each community has been carried out to gain an in-depth understanding of the features of each community (intercommunity analysis). Moreover, the Bonferroni test has also been carried out to find out the correlations between communities (intracommunity analysis). These two analyses can provide the quality of the algorithm in terms of intercommunity and intracommunity quality measurements.

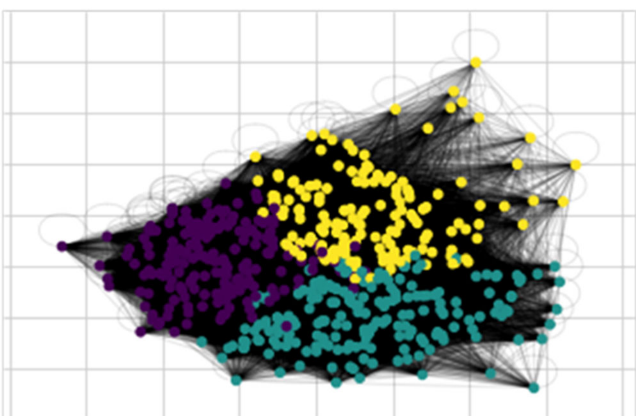

FIGURE 1 Results of the Louvain algorithm for STEM undergraduate students.

Therefore, Figure 2 shows the mean value and standard deviation inside each community (Community A, B, and C). It is observed that the standard deviation is not very high for any of the variables, indicating that there is not much dispersion in the distribution of the data within each community. But there are also differences between communities and also common patterns within them. To determine these common features, the three variables with the best and worst scores in each community were analyzed. Community A and B are quite similar in the three highest scoring variables, but the main difference is that Community A scores all variables with higher values, namely: "Q1. Gender equality is an important issue that needs to be tackled at all levels" (Community A = 4.94; Community B = 4.41), "Q8. Engineering/Technology professionals must help to reduce the gender gap in their sector" (Community A = 4.58; Community B = 3.68), and "Q6. More women professionals are needed in Engineering/Technology" (Community A = 4.44; Community B = 3.39). In contrast, Community C is very

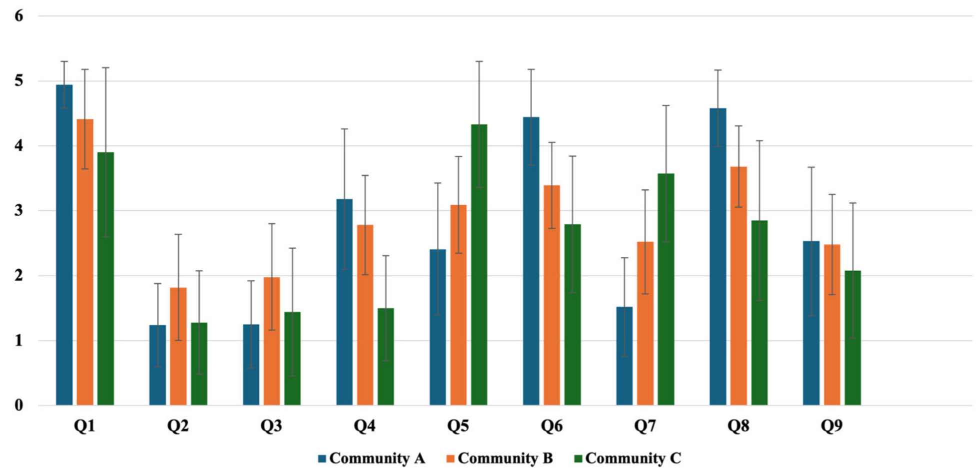

FIGURE 2 Descriptive statistics by each community.

different from the other two communities as the three variables with the highest scores are: "Q5. men and women are paid the same in Engineering/Technology" (4.33), "Q1. Gender equality is an important issue that needs to be tackled at all levels" (3.90), and "Q7. The gender gap is a passing fad" (3.57). Regarding the worst-scoring variables, a common pattern can be observed in the three communities, as the two variables with the worst scores are the same, that is: "Q2. Women have more problems than men in technical tasks in Engineering/Technology" and "Q3. Men are better prepared than women for jobs in Engineering/Technology", although Community A scores the lowest, followed by Community B and finally Community C. In contrast, the third worst scoring variable for each community is different. Indeed, this variable in Community A is "Q7. The gender gap is a passing fad" (1.52), in Community B is "Q9. The gender gap is a problem that only affects women" (2.48) and for Community C is "Q4. Women have more problems than men in finding jobs in Engineering/Technology" (1.50). In addition, the standard deviations within each community offer valuable insights into the variability of responses for each item. The standard deviations provide an indication of the dispersion or spread of responses around the mean. Upon examination of the standard deviations, it is evident that there is relatively low dispersion in the distribution of the data within each community for most variables. However, for items Q4 and Q5, there are notable differences observed across the communities. Particularly in Community A, the standard deviations for these items are notably higher compared to the other communities, indicating greater variability in the responses within Community A for these specific items. This suggests a more diverse range of perspectives or opinions within Community A regarding gender-related issues in Engineering/Technology. Conversely, in

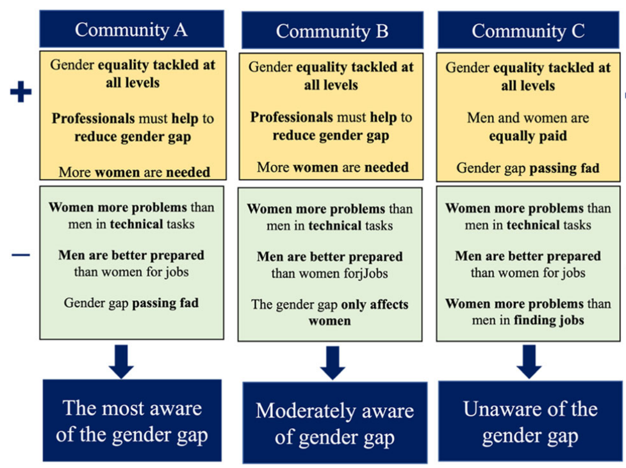

FIGURE 3 Summary of the main features obtained in Communities A, B, and C.

Community C, the standard deviation for items Q6, Q7, Q8, and Q9 is substantially higher compared to the other communities, indicating greater variability in responses within Community C for these items. This suggests a diverse range of perspectives within Community C regarding these questions.

Figure 3 summarizes the most important findings from Figure 2 where the yellow rectangles represent the highest-rated questions and the green rectangles represented the lowest-rated questions. According to the results, it can be concluded that Community A is the most aware of the gender gap in STEM, Community B is aware but in a more moderate way, and Community C is the most skeptical about the existence of the gender gap.

Furthermore, Table 2 shows the differences between the three communities using the Bonferroni test. Significant differences are found between all pairs of communities for all variables ($p < .05$). However, an exception is

TABLE 2 Bonferroni test between all communities for each variable.

Variable	Community comparison	Mean difference	Sig.
Q1	A vs. B	0.527	0.000
Q1	A vs. C	1.040	0.000
Q1	B vs. C	0.513	0.000
Q2	A vs. B	−0.575	0.000
Q2	A vs. C	−0.033	1.000
Q2	B vs. C	0.542	0.000
Q3	A vs. B	−0.736	0.000
Q3	A vs. C	−0.192	0.089
Q3	B vs. C	0.545	0.000
Q4	A vs. B	0.399	0.000
Q4	A vs. C	1.680	0.000
Q4	B vs. C	1.281	0.000
Q5	A vs. B	−0.674	0.000
Q5	A vs. C	−1.916	0.000
Q5	B vs. C	−1.242	0.000
Q6	A vs. B	1.050	0.000
Q6	A vs. C	1.644	0.000
Q6	B vs. C	−0.594	0.000
Q7	A vs. B	−1.000	0.000
Q7	A vs. C	−2.048	0.000
Q7	B vs. C	1.048	0.000
Q8	A vs. B	0.900	0.000
Q8	A vs. C	1.730	0.000
Q8	B vs. C	−0.829	0.000
Q9	A vs. B	0.051	1.000
Q9	A vs. C	0.446	0.000
Q9	B vs. C	−0.396	0.001

observed between Community A and Community C in variables "Q2. Women have more problems than men when it comes to technical tasks in Engineering/Technology" ($p = .849$) and "Q3. Men are better prepared than women for jobs in Engineering/Technology" ($p = .129$). This means that there is no significant difference in these variables between these two communities. A similar pattern is seen between Community A and Community B in the variable "Q9. The gender gap is a problem that only affects women" ($p = .879$). However, for the remaining variables, all pairs of communities show significantly different results, indicating divergent opinions among the communities. From these statistical results, it can be concluded that the algorithm shows a good performance among the detected communities, where the patterns of behavior and beliefs are clearly defined among the different communities of STEM university students.

4.2 | Results of the STEM undergraduate' beliefs and opinions on the gender gap in their professional fields

This subsection will describe the opinion and perception of university students in STEM degrees on the gender gap in their professional environment, followed by an explanation of the correlations between all variables. All results will be explained taking into account the gender of the students. Figure 4 shows the descriptive frequency (percentage) study of all Likert-type variables. It can be observed that the questions "Q2. Women have more problems than men in technical tasks in Engineering/Technology" (68.8%) and "Q3. Men are better prepared than women for jobs in Engineering/Technology" (63.4%) are the two questions that receive the highest percentage as "strongly disagree", while "Q1. Gender equality is an important issue that needs to be tackled at all levels" (66.4%) receives the highest percentage as "Totally agree". It is noteworthy that a high percentage of students do not perceive a great need for women in technological fields (Q6), with 41% responding "Neither agree nor disagree." In other questions, there is much more dispersion in the answers.

In this way, Figure 5 shows the mean and standard deviation of the results given by the whole population, female and male students. All questions have been averaged to analyze the result, but questions Q2, Q3, Q5, Q7, and Q9 have been rotated, that is, what used to be "1" is now "5", what used to be "5" is now "1" and so on

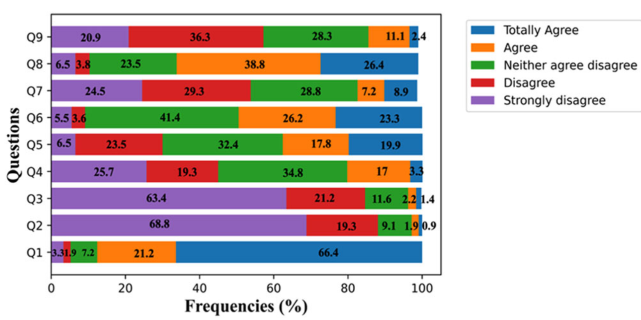

FIGURE 4 Descriptive frequency study of all Likert-type questions.

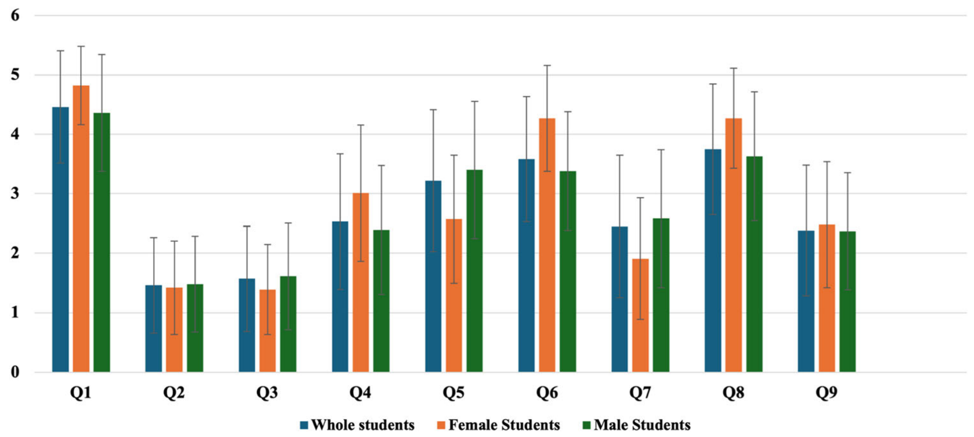

FIGURE 5 Descriptive study by mean and standard deviation of the Likert-type questions.

TABLE 3 Moderate Pearson correlations between different variables.

Variables		Pearson correlation	Sig. (bilateral)
Q1	Q6	0.441	0.000
Q1	Q7	−0.434	0.000
Q2	Q12	0.447	0.000
Q2	Q3	0.434	0.000
Q12	Q3	0.405	0.000
Q4	Q10	0.401	0.000
Q4	Q7	−0.433	0.000
Q4	Q11	−0.456	0.000
Q5	Q6	0.464	0.000
Q5	Q11	0.488	0.000
Q5	Q8	−0.485	0.000

TABLE 4 Strong Pearson correlations between different variables.

Variables		Pearson correlation	Sig. (bilateral)
Q10	Q6	0.558	0.000
Q10	Q8	0.591	0.000
Q5	Q4	−0.551	0.000
Q5	Q7	0.534	0.000
Q6	Q10	0.558	0.000
Q6	Q11	0.560	0.000
Q6	Q8	0.588	0.000

for all answers, as a lower value for these questions shows a worse perception of the gender gap.

Regarding the existence of correlations between the variables, Table 3 shows moderate correlations close to 0.5 (quite high) and Table 4 shows strong correlations. It can be observed that some variables correlate with other variables not specific to the professional domain, but to the educational domain (included in the global questionnaire but not in this study). However, it has been decided to add them as the results were decisive in this study. Specifically, these variables correspond to questions Q10, Q11, and Q12 which have been explained previously in the methodology.

Finally, a statistical analysis was conducted to examine the differences between male and female STEM students. Table 5 shows the results of Mann–Whitney *U* tests, which show that only two variables, "Q2. Women have more problems when it comes to technical tasks" and "Q9. The gender gap is a problem that only affects women," do not exhibit significant differences between both groups. However, there are significant differences in the rest of the variables, indicating a contrast of opinions between male and female students.

5 | DISCUSSION

This section is divided into two subsections. The first discusses how the community detection algorithm identifies patterns among STEM students. The second focuses on analyzing STEM students' opinions and beliefs about the gender gap.

5.1 | Discussion of the Louvain Community Detection algorithm

The community detection algorithm has been modeled to automatically obtain different communities of students with specific perceptions and beliefs of the gender gap in

TABLE 5 Mann–Whitney *U* test for both groups (male and female students).

	Mann–Whitney *U*		
	Value	***Z***	**Sig.**
Q1	21,071.000	−6.662	0.000
Q2	28,875.500	−1.082	0.279
Q3	26,003.00	−2.981	0.003
Q4	21,459.5000	−5.480	0.000
Q5	18,160.000	−7.464	0.000
Q6	15,255.000	−9.436	0.000
Q7	19,551.500	−6.319	0.000
Q8	18,959.000	−6.684	0.000
Q9	28,211.000	−0.960	0.337

STEM professional fields. The algorithm, combined with the statistical analysis presented in Figure 2, has successfully identified the key distinguishing characteristics of each community, enabling us to differentiate them based on their perspectives. As a result, the algorithm has categorized the communities into three distinct groups: Community A, characterized by a high awareness of the gender gap in STEM; Community B, exhibiting a moderate level of awareness; and Community C, expressing opposition to the existence of the gender gap. Notably, Community C is noteworthy in size, comparable to the other two communities, consisting of 170 students. This indicates a significant number of students who do not agree with the notion of a gender gap in STEM. This is also observed in other studies [69] which found gender biases in the selection decisions of scientific evaluation committees in a nationwide competition for elite research positions. Communities A and B strongly agree that engineering/technology professionals should contribute to narrowing the gender gap. This idea is in line with some studies that affirm the importance of outreach activities led by STEM professionals (science exhibitions, company/university visits, company workshops) to increase students' motivation toward STEM [63, 72, 85]. Both groups (A, B) are also aware of the real statistics on the low participation of women in STEM and the need to integrate more women, contrary to the non-perception of Community C [92]. Likewise, the C community holds the belief that the gender gap is merely a transient trend, but they are unaware of the substantial disparities that persist in specific fields. For instance, only 1.7% of women are enrolled in Technology (compared to 8.2% of men) and a mere 6.6% in Engineering (in contrast to 24.6% of men). Finally, Community C has a distorted perception of the gender pay gap in STEM, despite evidence that the gender pay gap ranges from 21% to 36% in the engineering sectors [7, 76].

When analyzing the communities according to gender, Community A, which shows the highest level of awareness of the gender gap, has the highest percentage of female students (48.33%). In contrast, Community C, the least aware, has the lowest percentage of female students (10.59%). Therefore, it can be deduced that female students are more concerned about the gender gap than their male counterparts. Regarding the degree (Telecommunications, Industrial, Computer Science), there are no differences in the percentage of students in the communities, so the results are independent of the type of engineering degree. Finally, the year to which the students belong does not seem to influence their perception of the gender gap, even the youngest students (1st and 2nd) have the same opinion as students of the last years (3rd, 4th, and Master).

On the other hand, this study has also investigated whether there are correlations between communities (Table 2) and if there is a relationship between variables considering the entire population in general (Tables 3 and 4). Some correlations coincide with the conclusions of our community detection algorithm. Indeed, correlations were found between students who think that gender equality should be addressed at all levels also think that more women need to be integrated into STEM jobs and students who think that STEM professionals should help to reduce the gender gap also agree that more women are needed in STEM (Community A and B). In this way, students who fully agree that gender equality must be addressed at all levels do not see the gender gap as a fad (Community A). As for Community C, it can be found a strong correlation between men and women receiving equal pay and the gender gap being a trend. Thus, this group of students also thinks that women do not have more problems than men in finding jobs in STEM sectors and also agrees that the gender gap is a trend (Community C). Hence, it can be concluded that the community detection algorithm performs optimally, as it is able to automatically and efficiently group a large number of students into communities from surveys. Subsequently, relationships within and outside each community can be obtained from a more detailed statistical analysis by applying descriptive and group comparison tests. Consequently, this methodology shows great potential as the algorithm could be used in other surveys and it is also possible to modulate the quantification of the weights that mark the relationships between users, as well as the number of questions.

5.2 | Discussion of the STEM undergraduates' beliefs and opinions on the gender gap in professional fields

The students' opinions have been described in Figures 4 and 5. This research shows that a very high number of students (around 50%) do not perceive the need for more women in STEM, but it is a fact that today there is still an under-representation of women in these professional fields, so students should be more aware of these worrying recent statistics [27, 29, 54, 69, 89]. It is also important to note that a relatively high percentage of students (around 43%) agree or neither agree nor disagree that the gender gap in STEM is a problem that only affects women, a worrying fact, as the gender gap has been shown to greatly affect social and economic dimensions of a country. In fact, the EU estimates that closing the gender gap in STEM would increase GDP per capita by about 3.0% by 2050 [29]. This worrying perception is also seen in the fact that men and women are equally paid in STEM, with a high number of students (around 40%) agreeing with this statement. The latest figures say otherwise, with men in the European technology sector earning 19% more than women [27] and in engineering sectors, the gender pay gap ranges from 21% to 36% [46, 67]. Indeed, in Sterling et al. [77] authors argue that the pay gap is a reality and is due to cultural beliefs about the worth of women in STEM professions or their motivation and self-esteem regarding their self-efficacy.

At the other end of the spectrum, it is pleasing to note that a large percentage of students perceive that gender equality should be addressed at all levels (close to 90%). These findings exhibit certain parallels with those observed in the study conducted by Bert et al. [10], which revealed that undergraduate medical students showed high gender sensitivity and low gender stereotypes. Indeed, this trend of gender sensitivity appears to hold true as students strongly disagree with the notions that women face more difficulties than men in technical tasks or that women are less prepared for STEM jobs, with approximately 90% expressing disagreement. However, recent research studies suggest that female students experience lower levels of confidence in their studies despite outperforming their male counterparts [17, 57]. Moreover, our results show that many students (around 66%) think that Engineering and Technology professionals should help to reduce the gender gap, so we could think that they will be future professionals aware of this problem. In fact, some studies point to the importance of outreach activities led by STEM professionals to boost young people's motivation [63, 72, 85, 56]. It, therefore, confirms that gender biases related to women's skills and qualifications in STEM jobs are no longer as widespread as they were years ago [17, 61, 74].

On the other hand, this study has analyzed existing correlations between all variables (Tables 3 and 4). Regarding the gender gap and university curricula, students who agree that STEM professionals should contribute to reducing the gender gap, that more women professionals are needed, and that women have more problems than men in finding jobs in STEM, also agree that this problem should be addressed in the curricula. In this regard, some recent studies also suggest designing gender-sensitive university curricula [83, 90] or increasing policy interventions on gender equality within academia [90]. In this way, Olatundun et al. [64] assert the need for changes in engineering curricula to include humanities and nontechnical content to attract girls to STEM. Besides, the integration of soft robotics into engineering degrees has been shown to increase interest and mitigate gender differences [59]. On the contrary, students who believe that the gender gap should not be addressed in curricula think that there is no pay gap and that women do not have more problems finding STEM jobs (they deny the existence of a gender gap). However, many data show lower participation of women in STEM professions [6, 27, 29, 54] and the main reasons are related to low and/or inequitable salaries, poor working conditions, or lack of recognition and promotion opportunities [31]. Lastly, students who perceive gender as a factor influencing the completion of engineering studies also hold the belief that women encounter more difficulties in technical tasks and that men possess superior preparation for STEM jobs. These notions are closely related to distorted gender stereotypes, as highlighted by some research which states that female engineering students often feel a disconnect with their studies and professional experiences while perceiving significant male dominance in their sectors [32, 70]. Furthermore, some studies reveal that female students perceive mathematics and physics subjects as heavily masculine in nature [44].

Another aim of this study was to find out the differences between the opinions of male and female students. The results suggest that both groups think that the gender gap is not only a problem for women and that women are equally capable of performing technical tasks, which is very positive. According to Eagly [23], fewer people now believe that women are less intelligent or skilled than men. In contrast, there are significant gender differences for the rest of the variables. It is very remarkable that female students perceive more than male students the need for women professionals in these sectors, with differences of almost one point. The same tendency is observed for men are better prepared than women or gender equality is an important issue that needs to be addressed at all levels, although it has been demonstrated that educational institutions [14, 17, 18, 29] as well as families [4, 61] and government policies [76] are being effective in achieving gender equality in STEM. Regarding STEM jobs, female students perceive that they have more problems than men in finding STEM jobs or that they do not receive the same pay, which is a reality as recent statistics reveal that fewer women than men work in STEM sectors [29, 30, 54, 89] and that there is a large pay gap [7, 27, 76]. Finally, significant differences are found in that the gender gap is a fad (perhaps the most worrying), where men are closer to neither agree nor disagree. In contrast, the gender gap has been analyzed since the 1970s [71] and is a recent and worrying reality that is being addressed by many institutions [27, 29, 92]. Consequently, female students are more aware of the gender gap than male students, which implies the need for more training and raising awareness among engineering students, as they are the future professionals in these sectors.

6 | CONCLUSIONS

The community detection algorithm has showcased excellent performance, adaptability, and potential by successfully uncovering and identifying distinct student groups with shared beliefs and perceptions concerning the gender gap. The algorithm has revealed two communities of students who are aware of the gender gap, with one community exhibiting a higher level of awareness than the other. Additionally, there is a third group, comparable in size to the other two, which expresses skepticism regarding the existence of the gender gap. This finding is concerning and underscores the urgency to implement strategies within STEM programs that effectively raise awareness about the gender gap. This study revealed a lack of awareness of the gender gap among university students on issues of major concern that should be addressed: a high number of students do not perceive the need for more professional women, a relatively high percentage of students believe that the gender gap is a problem that only affects women and, finally, a high percentage think that men and women are equally paid. The study also shows that women are more aware than men of the gender gap in STEM professional sectors. While it is positive that students (boys and girls) reject the idea that women are less prepared than men for STEM jobs and that they have more problems performing technical tasks, the same is not true for other variables, and the idea that men perceive the gender gap as a temporary fad is particularly worrying.

This study holds significant implications for engineering education as it contributes to understanding university students' perceptions, motivations, and awareness of the gender gap in their future work environments. This understanding enables universities to develop strategies that foster more egalitarian perceptions among students within university settings, with particular attention to the areas identified as requiring improvement in this study. These areas include the low perception that women are needed in these sectors or that this gender gap is temporary. On the contrary, this disparity has persisted over time, and there is now an urgent need for more women in these fields in our digitized societies (professionals in AI, Big Data, Cybersecurity, Cloud Computing, or IoT). In addition, there is a pressing need to address the skepticism expressed by a high number of students regarding the existence of the gender gap. Strategies should focus on dispelling misconceptions and promoting awareness-raising initiatives about STEM disciplines. On a broader scale, technology and engineering companies can benefit from implementing policies that create more equitable working environments, forging partnerships with universities to leverage synergies. Furthermore, governments can utilize the insights from this study to guide their efforts in generating public awareness campaigns to improve perceptions of the gender gap in STEM. The findings from this study provide valuable guidance on where to concentrate these collective efforts. Finally, it should be noted that this study is novel, as there is not much literature analyzing the beliefs of university students on the gender gap, especially in the Engineering and Technology sectors; so it is of great interest to the scientific community, also incorporating novel analysis techniques related to AI. Indeed, the application of community detection algorithms, that is, the integration of AI in these fields, opens a new research horizon that has not yet been widely explored and is highly pioneering. Indeed, community detection algorithms can unveil underlying patterns (hidden patterns) in complex data sets that are not easily discernible at first glance. In our context, they can identify clusters of students who share similar characteristics such as gender, academic interests, skill level, and more. By examining these communities, researchers can uncover gender biases in terms of participation, academic performance, and access to resources, among other factors. Furthermore, community detection algorithms can segment the student population into homogeneous groups based on various characteristics, including gender. By analyzing these communities, researchers can identify significant differences between gender groups in terms of behaviors, attitudes toward STEM education, social interactions, and academic experiences. This provides valuable information for understanding the factors contributing to gender biases in the educational sphere and designing effective interventions to address them. Therefore, our proposal establishes the methodological foundations and describes the procedure for using AI algorithms in areas of education and qualitative research of great significance, such as the gender gap in the STEM field. This aspect is of utmost importance for digitized societies and their evolution.

In considering the potential of our study, we acknowledge limitations and suggest avenues for future research. While our sample size is sufficient, broadening the diversity of participants, particularly in terms of race, ethnicity, and socioeconomic status, would improve the breadth and generalizability of our results, highlighting the intersectionality of gender beliefs. Another limitation is related to the inclusion of the university faculty's perspective on gender and STEM issues to achieve more comprehensive results that allow us to propose more universal and effective strategies, taking into account the key educational stakeholders.

ACKNOWLEDGMENTS

This work has been supported by the University of Valladolid as a Teaching Innovation Project "InGenias: Women as a Precursor to Technological and Scientific Vocations (2019-2021)". We are grateful to the primary and secondary school teachers who participated in this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data sets used during the current study are available from the corresponding author upon reasonable request.

ORCID

Noemi Merayo http://orcid.org/0000-0002-6920-0778
Alba Ayuso http://orcid.org/0000-0002-0740-7212

REFERENCES

1. M. Adraoui, A. Retbi, M. K. Idrissi, and S. Bennani, *A new approach to detect at-risk learning communities in social networks, EMENA-ISTL 2019: innovation in information systems and technologies to support learning research*. Proc. EMENA-ISTL, 75–84, 2020.
2. P. Aeby, R. Fong, M. Vukmirovic, S. Isaac, and R. Tormey, *The impact of gender on engineering students' group work experiences*, Int. J. Eng. Educ. **35** (2019), 756–765.
3. V. Alejos, M. G. Sanchez, M. P. Milagros, F. Falcone, P. Sanchís, and A. López-Martín, The influence of gender in the adoption of engineering studies, 2014 XI Tecnologias Aplicadas a la Ensenanza de la Electronica (Technologies Applied to Electronics Teaching) (TAEE), Bilbao, Spain, 2014, pp. 1–4. https://doi.org/10.1109/TAEE.2014.6900178.[nm1]
4. J. Almeida and A. D. Daniel, *Women in engineering: developing entrepreneurial intention through learning by doing approach*, 2021 IEEE Global Eng. Educ. Conf. (EDUCON), 2021, pp. 116–121.
5. M. E. Andrews, A. D. Patrick, and M. Borrego, *Engineering students' attitudinal beliefs by gender and student division: a methodological comparison of changes over time*, Int. J. STEM Educ. **8** (2021), 13.
6. Australian Government, *STEM Equity Monitor. Explore Australia's data on women in STEM*, Department of Industry, Science and Resources, Canberra, Australia, 2022. https://www.industry.gov.au/publications/stem-equity-monitor
7. Australian Government, *Gender pay gaps by industry*, Department of Industry, Science and Resources, Canberra, Australia, 2021. https://www.industry.gov.au/publications/stem-equity-monitor/workforce-data/gender-pay-gaps-stem-and-other-industries
8. A. Ayuso, N. Merayo, I. Ruiz, and P. Fernandez, *Challenges of STEM vocations in secondary education*, IEEE Trans. Educ. **65** (2022), 713–724.
9. C. F. Baizan, L. P. Jimenez, and V. R. Biarge, *A general overview of the professional activities of women in the computer and electrical engineering fields in Spain*, IEEE Trans. Educ. **28** (1985), 233–235.
10. F. Bert, E. Boietti, S. Rousset, E. Pompili, E. Franzini Tibaldeo, M. Gea, G. Scaioli, and R. Siliquini, *Gender sensitivity and stereotypes in medical university students: an Italian cross-sectional study*, PLoS One **17** (2022), e0262324.
11. J. M. Bland and D. G. Altman, *Statistics notes: Cronbach's alpha*, BMJ **314** (1997), 572. https://doi.org/10.1136/bmj.314.7080.572
12. V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, *Fast unfolding of communities in large networks*, J. Stat. Mech. Theory Exp. **2008** (2008), P10008.
13. B. Bloodhart, M. M. Balgopal, and A. M. A. Casper, *L.B. Sample McMeeking and E. V. Fischer, Outperforming yet undervalued: undergraduate women in STEM*, PLoS One **15** (2021), 1–10.
14. J. Bloomfield and M. Fisher, *Quantitative research design*, J. Australas. Rehabil. Nurses' Assoc. **22** (2019), 27–30.
15. C. C. Bridges, *Hierarchical cluster analysis*, Psychol. Rep. **18** (1966), 851–854.
16. P. Cantrell and J. Ewing-Taylor, *Exploring STEM career options through collaborative high school seminars*, J. Eng. Educ. **98** (2013), 295–303.
17. P. Carnemolla and N. Galea, *Why Australian female high school students do not choose construction as a career: a qualitative investigation into value beliefs about the construction industry*, J. Eng. Educ. **110** (2021), 819–839.

18. M. Chiquito, R. Castedo, A. P. Santos, L. M. López, and C. Alarcón, *Flipped classroom in engineering: the influence of gender*, Comput. Appl. Eng. Educ. **28** (2020), 80–89.
19. A. Clauset, M. E. J. Newman, and C. Moore, *Finding community structure in very large networks*, Phys. Rev. E **70** (2004), 066111.
20. Community API, *Community detection for NetworkX*, Thomas Aynaud Revision, 2022. https://python-louvain.readthedocs.io/en/latest/api.html
21. V. Da Fonseca, C. Ribeiro, N. F. Favilla, and A. Gonçalves, *Performance evaluation of modularity based community detection algorithms in large scale networks*, Math. Probl. Eng. **2014** (2014), 1–15.
22. R. DeVellis, Scale development: theory and applications: theory and application, Sage, Thousand Oaks, CA, 2003.
23. A. H. Eagly, *Have gender stereotypes changed? Yes and no*, Presentation at INSEAD Women at Work Conference. https://www.youtube.com/watch?v=ewOsOtHB-I8
24. Elinext, *The IT talent gap: regions under the risk*, Elinext Group, 2002. https://www.elinext.com/blog/worldwide-it-talent-gap-in-2022/
25. Emerson, *Emerson Survey: 2 in 5 Americans Believe the STEM Worker Shortage is at Crisis Levels*, Emerson Electric Co., United States, 2018. https://www.emerson.com/en-us/news/corporate/2018-stem-survey
26. European Commision, *The Digital Economy and Society Index (DESI)*, Directorate-General for Communications Networks, Content and Technology, Bruxelles/Brussel, Belgium, 2024. https://digital-strategy.ec.europa.eu/en/policies/desi
27. European Commission, *Women in Digital Scoreboard*, Directorate-General for Communications Networks, Content and Technology, Bruxelles/Brussel, Belgium, 2021. https://digital-strategy.ec.europa.eu/en/news/women-digital-scoreboard-2021
28. European Commission, *Developments and forecasts of changing nature of work*, Directorate-General for Communications Networks, Content and Technology, Bruxelles/Brussel, Belgium, https://knowledge4policy.ec.europa.eu/foresight/topic/changing-nature-work/developments-forecasts-changing-nature-work_en
29. European Institute for Gender Equality, *Economic Benefits of Gender Equality In the European Union*, Directorate-General for Communications Networks, Content and Technology, Bruxelles/Brussel, Belgium, 2022. https://eige.europa.eu/gender-mainstreaming/policy-areas/economic-and-financial-affairs/economic-benefits-gender-equality
30. P. Fatourou, Y. Papageorgiou, and V. Petousi, *Women are needed in STEM: European policies and incentives*, Commun. ACM **62** (2019), 52.
31. N. A. Fouad, W. H. Chang, M. Wan, and R. Singh, *Women's reasons for leaving the engineering field*, Front. Psychol. **8** (2017), 1–12.
32. A. Garcia-Holgado, C. S. Gonzalez-Gonzalez, and A. Peixoto, *A comparative study on the support in engineering courses: A case study in Brazil and Spain*, IEEE Access **8** (2020), 125179–125190.
33. A. García-Holgado, A. Vázquez Ingelmo, J. J. Mena Marcos, F. J. García-Peñalvo, C. S. González González, M. C. Sánchez Gómez, and S. Verdugo Castr, *Estudio piloto sobre la percepción de la brecha de género en estudios de ingeniería informática*, Actas del V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2019 (M. L. Sein-Echaluce Lacleta, Á. Fidalgo-Blanco, and F. J. García-Peñalvo, eds.), Servicio de Publicaciones Universidad de Zaragoza, Zaragoza, Spain, 2019, pp. 698–703.
34. M. Grañeras, M. E. Moreno, and N. Isidoro, Ministerio de Educación y Formación Profesional. *Radiography of the gender gap in STEAM*, 2011. https://alianzasteam.educacionyfp.gob.es/datos-steam/observatorio-steam/universidad/alumnado-nivel.html
35. E. K. Henriksen, F. Jensen, and J. Sjaastad, *The role of out-of-school experiences and targeted recruitment efforts in Norwegian science and technology students' educational choice*, Int. J. Sci. Educ. **5** (2015), 203–222.
36. Huawei, *Addressing the ICT talent shortage and providing support for the digital transformation*, Huawei Technologies Co., 2019. https://e.huawei.com/es/publications/global/product-insights/enterprise-services/global-vision/201908291143
37. N. Iam-On and T. Boongoen, *Generating descriptive model for student dropout: a review of clustering approach*, Hum. Centric Comput. Inform. Sci. **7** (2017), 1–10.
38. A. Jackson, N. Mentzer, and R. Kramer-Bottiglio, *Increasing gender diversity in engineering using soft robotics*, J. Eng. Educ. **110** (2021), 143–160.
39. A. Khaled, S. Ouchani, and C. Chohra, *Recommendations-based on semantic analysis of social networks in learning environments*, Comput. Human. Behav. **101** (2019), 435–449.
40. V. S. Kumaran, A. Sankar, and K. Kiruthikaa, *Group recommendation in online learning environment using community detection*, Int. J. Comput. Syst. Eng. **2** (2015), 51–59.
41. J. Leskovec, A. Rajaraman, and J. D. Ullman, Mining of massive data sets, Cambridge University Press, 2020.
42. L. Mahnane, *Recommending learning activities in social network using data mining algorithms*, J. Educ. Technol. Soc. **20** (2017), 11–23.
43. T. Mainhard, S. Oudman, L. Hornstra, R. J. Bosker, and T. Goetz, *Student emotions in class: the relative importance of teachers and their interpersonal relations with students*, Learn. Instr. **53** (2018), 109–119.
44. E. Makarova, B. Aeschlimann, and W. Herzog, *The gender gap in STEM fields: the impact of the gender stereotype of math and science on secondary students' career aspirations*, Front. Educ. **4** (2019), 1–11.
45. J. Maldonado-Mahauad, M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, and J. Munoz-Gama, *Mining theory-based patterns from Big data: identifying self-regulated learning strategies in massive open online courses*, Comput. Human. Behav. **80** (2018), 179–196.
46. F. D. Malliaros and M. Vazirgiannis, *Clustering and community detection in directed networks: a survey*, Phys. Rep. **533** (2013), 95–142. https://doi.org/10.1016/j.physrep.2013.08.002
47. S. Mateos-Sillero and C. Gómez-Hernández, *Libro Blanco de las mujeres en el ámbito tecnológico*, Ministerio de Economía y Empresa, 2019.
48. A. Menacho, P. Plaza, E. Sancristobal, C. Perez-Molina, M. Blazquez, and M. Castro, *Halloween educational robotics*, IEEE Trans. Educ. **64** (2021), 406–412.

49. P. Mengoni, A. Milani, and Y. Li, *Community graph elicitation from students' interactions in virtual learning environments*, Lect. Notes. Comput. Sci. **10962** (2018), 414–425.
50. Microsoft, *Why don't European girls like science or technology*? Microsoft, 2017, https://news.microsoft.com/europe/features/dont-european-girls-like-science-technology/
51. Ministerio de ciencia e innovación del Gobierno de España, *Científicas en cifras*, Ministerio de ciencia e innovación del Gobierno de España, Madrid, Spain, 2019. https://www.ciencia.gob.es/site-web/Secc-Servicios/Igualdad/cientificas-en-cifras.htm
52. S. Moll, J. A. Moraño, M. D. Roselló, L. M. S. Ruiz, Is there a gender difference in Maths competencies achievement between Aerospace Engineering students in Spain? *IEEE Front. Educ. Conf. (FIE), El Paso*, TX, USA, 2015, 1–4. https://doi.org/10.1109/FIE.2015.7344112
53. M. M. Msambwa, K. Daniel, C. Lianyu, and A. A. Fute, *A systematic review of the factors affecting girls' participation in science, technology, engineering, and mathematics subjects*, Comput. Appl. Eng. Educ.**32** (2023), e22707. https://doi.org/10.1002/cae.22707
54. National Science Foundation, *The STEM labor force of today: scientists, engineers, and skilled technical workers*, Science and Engineering Indicators National Science Foundation, Alexandria, Virginia, 2021. https://ncses.nsf.gov/pubs/nsb20212#utm_medium=email&utm_source=govdelivery
55. C. Navarro-Guzmán and A. Casero-Martínez, *Analysis of gender differences in the choice of university studies*, Estudios Sobre Educación **22** (2012), 115–132. https://doi.org/10.15581/004.22.2075
56. Network Analysis in Python Algorithms. *Communities. Louvain_communities.*, NetworkX Developers. https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.louvain.louvain_communities.html
57. B. A. Nosek and F. L. Smyth, *Implicit social cognitions predict sex differences in math engagement and achievement*, Am. Educ. Res. J. **48** (2011), 1125–1156.
58. J. Nunnally and L. Bernstein, Psychometric theory, McGraw-Hill Higher, INC, New York, 1994.
59. E. A. Olatundun, F. E. Archibong, and C. A. Ikutegbe, *Strategies to stimulate young female in engineering*, Int. J. Eng. Res. Adv. Technol. **5** (2019), 15–21.
60. J. Oliver and B. García, *Students' attitudes to engineering education*, Proc. 2012 IEEE Global Eng. Educ. Conf.(EDUCON), Marrakech, Morocco, 2012, p. 1. https://doi.org/10.1109/EDUCON.2012.6201059.
61. N. Olmedo-Torre, F. Sanchez Carracedo, M. N. Salan Ballesteros, D. Lopez, A. Perez-Poch, and M. Lopez-Beltran, *Do female motives for enrolling vary according to STEM profile?* IEEE Trans. Educ. **61** (2018), 289–297.
62. Organisation for Economic Cooperation and Development, *Programme for the International Student Assessment (PISA)*, Organisation for Economic Cooperation and Development, Boulogne-Billancourt, 2021. https://www.oecd.org/pisa/publications/pisa-2018-results.htm
63. T.-A. Palmer, P. F. Burke, and P. Aubusson, *Why school students choose and reject science: a study of the factors that students consider when selecting subject*, Int. J. Sci. Educ. **39** (2017), 645–662.
64. A. Patrick, C. Riegle-Crumb, and M. Borrego, *Examining the gender gap in engineering professional identification*, J. Women Minor. Sci. Eng. **27** (2021), 31–55.
65. J. V. Peña-Calvo, M. Inda-Caro, C. Rodríguez-Menéndez, and C. M. Fernández-García, *Perceived supports and barriers for career development for second-year STEM students*, J. Eng. Educ. **105** (2016), 341–365. https://doi.org/10.1002/jee.20115
66. G. Pérez-Artieda, E. Gubia, E. Barrenechea, P. Sanchis, A. López-Martín, D. Astrain, D. Morato, J. López-Taberna, and I. Matías, Analysis of women enrollment in Engineering programs at the Public University of Navarre, IEEE Front. Educ. Conf. (FIE) Proc., Madrid, Spain, 2014, pp. 1–8. https://doi.org/10.1109/FIE.2014.7044066
67. C. Pradana, S. S. Kusumawardani, and A. E. Permanasari, *Comparison clustering performance based on moodle log mining*, IOP conference series: materials science and engineering, vol. **722**, IOP Publishing, 2020, p. 012012.
68. F. O. Ramirez, N. Kwak, K. Seely-Gant, M. Ogawa, L. M. Frehill, S. Y. Atwaters, and R. Akimana, *Women's enrollments in STEM in higher education: Cross-national trends, 1970–2010*, Advancing women in science: an international perspective (W. Pearson, L. M. Frehill, and C. L. McNeely, eds.), Springer International Publishing, Cham, 2015, pp. 9–49.
69. I. Régner, C. Thinus-Blanc, A. Netter, T. Schmader, and P. Huguet, *Committees with implicit biases promote fewer women when they do not believe gender bias exists*, Nat. Hum. Behav. **3** (2019), 1171–1179.
70. M. S. Ross, J. L. Huff, and A. Godwin, *Resilient engineering identity development critical to prolonged engagement of Black women in engineering*, J. Eng. Educ. **110** (2021), 92–113.
71. A. S. Rossi, *Women in science: why so few?: social and psychological influences restrict women's choice and pursuit of careers in science*, Science **148** (1965), 1196–1202.
72. H. Salmi, H. Thuneberg, and M.-P. Vainikainen, *How do engineering attitudes vary by gender and motivation? attractiveness of outreach science exhibitions in four countries*, Eur. J. Eng. Educ. **41** (2016), 638–659.
73. D. Singh and R. Garg, *Issues and challenges in community detection algorithms*, Int. J. Manag. Technol. Eng. **8** (2018), 1375–1382.
74. C. R. Starr, *I'm not a science nerd!": STEM stereotypes, identity, and motivation among undergraduate women*, Psychol. Women Q. **42** (2018), 489–503.
75. STEM education guide, 2024. https://stemeducationguide.com/stem-education-statistics/
76. STEM women: The Gender Pay Gap in STEM, 2021. https://www.stemwomen.com/the-gender-pay-gap-in-stem. Accessed March 01, 2023.
77. A. D. Sterling, M. E. Thompson, S. Wang, A. Kusimo, S. Gilmartin, and S. Sheppard, *The confidence gap predicts the gender pay gap among STEM graduates*, Proc. Natl. Acad. Sci., **117**(2020), 30303–30308.
78. G. Stoet and D. C. Geary, *The gender-equality paradox in science, technology, engineering, and mathematics education*, Psychol. Sci. **29** (2018), 581–593.

79. G. Sun and S. Bin, *Topic interaction model based on local community detection in MOOC discussion forums and its teaching*, Educ. Sci. Theory Pract.**18** (2018), 2922–2931.
80. K. S. Taber, *The use of Cronbach's alpha when developing and reporting research instruments in science education*, Res. Sci. Educ. **48** (2018), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2
81. V. Tandrayen-Ragoobur and D. Gokulsing, *Gender gap in STEM education and career choices: what matters*, J. Appl. Res. High. Educ. **14** (2022), 1021–1040. https://doi.org/10.1108/JARHE-09-2019-0235
82. V. A. Traag Faster, *unfolding of communities: speeding up the Louvain algorithm*, Phys. Rev. E **92** (2015), 1–12.
83. C. S. Tu, M. J. Dupaya, and T. B. Buenaobra, *An investigation of gender-based sensitivity of college students*, SDCA Asia-Pac. Multidiscip. Res. J.**2** (2020), 14–17.
84. U.S. Bureau of Labor Statistics, Employment in STEM occupations, 2022. https://www.bls.gov/emp/tables/stem-employment.htm
85. J. Vennix, P. den Brok, and R. Taconis, *Do outreach activities in secondary STEM education motivate students and improve their attitudes towards STEM?* Int. J. Sci. Educ. **40** (2018), 1263–1283.
86. T. Verge, *Gender equality policy and universities: feminist strategic alliances to re-gender the curriculum*, J. Women Politics Policy **42** (2021), 191–206.
87. J. Wang and Y. Zhang, *Clustering study of student groups based on analysis of online learning behavior*, Proc. 2019 Int. Conf. Mod. Educ. Technol., 2019, pp. 115–119.
88. Q. Wang. H. Wang, *Study on MOOC withdrawal rate based on graph community detection model*, 2019 10th Int. Conf. Inform. Technol. Med. Educ. (ITME), 2019, pp. 526–529.
89. WES, *Useful Statistics. As reported in March 2022 by Engineering UK*, Women's Engineering Society, Stevenage, UK. https://www.wes.org.uk/content/wesstatistics
90. K. M. Whitcomb, Z. Y. Kalender, T. J. Nokes-Malach, C. D. Schunn, and C. Singh, *Comparison of self-efficacy and performance of engineering undergraduate women and men*, Int. J. Eng. Educ. **36** (2020), 1996–2014.
91. A. F. Wise and Y. Cui, *Learning communities in the crowd: characteristics of content related interactions and social relationships in MOOC discussion forums*, Comput. Educ. **122** (2018), 221–242.
92. World Economic Forum, *Global gender gap report*, World Economic Forum, Cologny, Switzerland, 2023. https://www.weforum.org/publications/global-gender-gap-report-2023/in-full/
93. Y. J. Xu, *Attrition of women in STEM: examining job/major congruence in the career choices of college graduates*, J. Career Dev. **44** (2017), 3–19. https://doi.org/10.1177/0894845316633787
94. S. Yassine, S. Kadry, and M. A. Sicilia, *Application of community detection algorithms on learning networks. The case of Khan Academy repository*, Comput. Appl. Eng. Educ. **29** (2021), 411–424.
95. T. T. Yuen and T. A. Pickering, Investigation of social networks and discussions in stem education communities on twitter, 2015 Int. Conf. Learn. Teach. Comput. Eng., 2015, pp. 128–133.
96. J. Zhang, J. Fei, X. Song, and J. Feng, *An improved Louvain algorithm for community detection*, Math. Probl. Eng. **2021** (2021), 1–14.
97. X. Zhang, Z. Ma, Z. Zhang, Q. Sun, and J. Yan, *A review of community detection algorithms based on modularity optimization*, J. Phys. Conf. Series **1069** (2018), 012123. https://doi.org/10.1088/1742-6596/1069/1/012123

AUTHOR BIOGRAPHIES

Noemi Merayo received a Telecommunication Engineer degree from the Valladolid University, Spain, in February, 2004, and a PhD degree in the Optical Communication Group at the Universidad de Valladolid, in July 2009. She works as Lecturer at the Universidad de Valladolid. She is a coauthor of more than 36 articles in peer-reviewed journals and 90 conference publications. She has also been a Visiting Research Fellow at the University of Hertfordshire in the Optical Networks Group, Science and Technology Research Institute, the TOyBA research group (University of Zaragoza), and the Technology University of Munich (TUM). Her research focuses on the design and performance evaluation of optical networks and the application of artificial intelligence techniques. Dr. Merayo is currently coordinating the Master in Physics and Technology of lasers at the University of Valladolid and the University of Salamanca.

Alba Ayuso was awarded her PhD in October 2021 from the University of Valladolid (Spain). She got a degree in speech therapy from the University of Valladolid in 2014; and a master's degree in neuropsychology and education from the University of La Rioja in 2016. Since 2014, she has been an Associate Professor at the Department of Pedagogy, at the University of Valladolid; and a Collaborating Professor at the University of La Rioja since 2017. Her research interests include language development studies, speech-language therapy interventions, and education. Ms. Ayuso is a member of the Association of Speech Therapists of Spain.

How to cite this article: N. Merayo, and A. Ayuso, *Identifying beliefs about the gender gap in engineering professions among university students using community detection algorithms and statistical analysis*, Comput. Appl. Eng. Educ. **32**, (2024), e22751. https://doi.org/10.1002/cae.22751