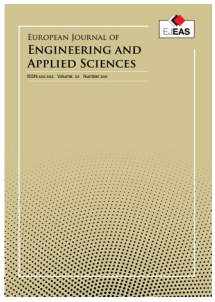
European Journal of Engineering and Applied Sciences

ISSN: 2651-3412 (Print) & 2667-8454 (Online)
Journal homepage: http://dergipark.gov.tr/EJEAS
Published by Çorlu Faculty of Engineering, Tekirdağ Namık Kemal University

European J. Eng. App. Sci. 7(2), 135-142, 2024



Research Article

Analyzing Factors Influencing Vocational High School IT Program Students' University Choices Using Association Rule Mining

Esma Türk[1], Erkan Özhan[2,*]

[1] Tekirdağ Namık Kemal University, Çorlu Faculty Engineering, Department of Computer Engineering, Tekirdağ, Türkiye, 59860
[1]esmaturk5834@gmail.com, [2]erkanozhan@gmail.com

Received: 24.12.2024 Accepted: 27.12.2024 DOI: 10.55581/ejeas.1606948

Abstract: The complex masses of data that have emerged with increasing data generation and storage have increased the need for computers and software with more advanced computing capabilities to process this data. However, extracting meaningful information from complex data remains a challenge. Data mining, particularly in collaboration with artificial intelligence algorithms, works to uncover intricate relationships within data. One of the complex problems to be solved is guiding high school students toward university departments that will optimize their performance. This study investigates the factors influencing the university department preferences of vocational high school information technology students and graduates in the field of computer science. Unlike previous research, has typically focused on academic performance and current educational contexts, this study explores the connections among students' past educational experiences, preferences, habits, and hobbies, tracing these back to primary and secondary education. As a case study, the research centers on the computer engineering department, revealing that students who wish to study or are studying computer engineering show a greater interest in activities related to design and game development, have a preference for the C# programming language, and exhibit a particular interest in chemistry, while demonstrating less affinity for street games. These findings underscore the relationship between students' higher education preferences in computer science and their prior learning experiences and social preferences, offering deeper insights into the decision-making process.

Keywords: *Artificial intelligence, Association rules, Career, Data mining, University department choice.*

Meslek Lisesi Bilişim Programında Üniversite Bölüm Tercihini Etkileyen Faktörlerin Birliktelik Kuralı ile Analizi

Öz. Artan veri üretimi ve depolamasıyla birlikte ortaya çıkan karmaşık veri yığınları, bu verilerin işlenmesi için daha gelişmiş hesaplama yeteneklerine sahip bilgisayarlar ve yazılımlara olan ihtiyacı artırdı. Ancak, karmaşık verilerden anlamlı bilgiler çıkarmak hala bir zorluktur. Veri madenciliği bilimi, özellikle yapay zeka algoritmaları ile iş birliği içinde, verilerdeki karmaşık ilişkileri ortaya çıkarmak için çalışmaktadır. Çözülmesi gereken karmaşık sorunlardan biri de lise öğrencilerinin üniversitede en yüksek verimi sağlayacak bölüme yönlendirilmesidir. Bu çalışma, meslek lisesi bilişim teknolojileri öğrencilerinin ve mezunlarının bilgisayar bilimleri alanındaki üniversite bölüm tercihlerini etkileyen faktörleri araştırmaktadır. Genellikle akademik performans ve mevcut eğitim bağlamlarına odaklanan önceki araştırmalardan farklı olarak bu çalışma, öğrencilerin geçmiş eğitim deneyimleri, tercihleri, alışkanlıkları ve hobileri arasındaki bağlantıları ilk ve orta öğretime kadar izleyerek araştırmaktadır. Bir vaka çalışması olarak bilgisayar mühendisliği bölümüne odaklanan araştırma, bilgisayar mühendisliği okumak isteyen veya okumakta olan öğrencilerin tasarım ve oyun geliştirmeyle ilgili faaliyetlere daha fazla ilgi gösterdiğini, C# programlama dilini tercih ettiğini ve kimyaya özel bir ilgi gösterirken sokak oyunlarına daha az yakınlık gösterdiğini ortaya koymaktadır. Bu bulgular, öğrencilerin bilgisayar bilimleri alanındaki yüksek öğrenim tercihleri ile önceki öğrenme deneyimleri ve sosyal tercihleri arasındaki ilişkinin altını çizmekte ve karar verme sürecine dair daha derin bilgiler sunmaktadır.

Anahtar kelimeler: *Birliktelik kuralları, Kariyer, Üniversite bölüm tercihi, Veri madenciliği, Yapay zeka.*

∗Corresponding author
E-mail address: erkanozhan@gmail.com (E. Özhan)

1. Introduction

Vocational high schools are a type of secondary education institution. These schools provide practical training rather than theoretical knowledge and aim to produce skilled workers in various professional fields. Vocational high schools enable students to acquire basic knowledge and skills in a specific occupational field and offer options for both direct participation in the labor force and university transition. Students attending Vocational and Technical Anatolian High Schools (Mesleki ve Teknik Anadolu Liseleri-MTAL) take all the cultural subjects like those in other high schools, while also receiving vocational education related to their chosen field. Vocational high schools enhance students' chances of finding employment by offering practical training and providing opportunities for professional experience through internships. Vocational high schools increase students' chances of finding a job by providing opportunities for professional experience through practical training and internships. In addition, students are also provided with advantages such as being insured according to the social security requirements that are compulsory during the internship and increasing the additional points and multiplier coefficient when they apply to associate degree programs in their fields after graduating from vocational high schools. In addition, focusing intensely on vocational courses can lower students' university exam performance, and the limited number of academic courses can leave students inadequately prepared. On the other hand, vocational high schools aim to provide students with practical knowledge and skills in a specific occupational field. Vocational high schools are designed to provide students with the necessary skills for direct entry into business life and to contribute to the qualified labor force in specific sectors. Accordingly, academic courses have a supporting role in vocational high schools and the focus is on vocational education. Therefore, the aim of vocational high schools is to focus intensively on vocational courses. These schools aim to provide vocational specialization rather than university preparation; therefore, exam success is not the primary goal of these schools. Previously, companies were reluctant to employ uninsured interns and were unwilling to cover internship insurance costs, leading to difficulties in accepting interns and limiting the number of available placements. This situation made it challenging for students to secure internship positions. To address this issue, the government began covering students' insurance costs during internships, allowing students to gain work experience under social security protection and insurance coverage against workplace accidents, while companies were able to employ interns without incurring additional expenses. The rationale behind awarding extra points is that, in Turkey, vocational high schools have fewer academic courses geared toward university entrance exams, which hinders students' preparation for these exams. As a result, many students prefer high schools with a more comprehensive academic curriculum, such as Anatolian high schools, which significantly reduces the preference for vocational schools and makes it difficult for these schools to attract students. In response to the demand for mid-level professionals supplied by vocational high school graduates, the institution responsible for the student selection and placement exams has introduced incentives, such as additional points and increased multipliers, to encourage enrollment in vocational schools.

Data mining techniques allow educational institutions to forecast student performance and discern correlations between attributes [1]. In recent years, the development of data mining techniques has enabled the creation of more detailed student models, capable of representing and detecting a broader range of student behaviours than was previously feasible [2].

Young people often make the most important and critical decisions of their lives during their high school years. They shape their careers based on the courses they take and the choices they make during this period. The decisions they make can have lifelong consequences. Career choice is one of the significant events that affects a person's life. The process of career planning encompasses a number of key elements, including the determination of career direction and goal, the selection of an appropriate career path, and the formulation of an effective educational plan [3]. When an individual chooses a profession, they also determine a lifestyle and work environment that suits them [4]. A study by Kurt & Fidan [5] revealed that students at a state university often make significant decisions based on short-term considerations, leading to dissatisfaction with their education and career prospects. The researchers recommend that universities prioritize practical training, expand internship opportunities, and support social and career development activities to enhance graduates' employability. They also emphasize the importance of providing career counseling and educational support, as well as collecting feedback from graduates and industry professionals to improve university programs.

As technology advances, students are required to equip themselves with new knowledge in various fields. With the growing prevalence of electronics and computers in industry, there is a significant need for well-trained individuals with high cognitive skills, not only in the usage but also in the production and maintenance stages. MTAL serve as foundational institutions where students first encounter computer science, just as they do for many other fields. In these schools, students explore the deeper workings of computers, going beyond their everyday uses. While some students continue their education in higher institutions or pursue careers in the information technology field after graduation, others may choose different professions that they find more suitable.

This study analyzes which informatics or other departments students in the informatics programs of vocational and technical high schools prefer for their university education after their training, and the factors influencing these choices, using association rules from data mining methods. Additionally, the post-secondary process is examined through the experiences of graduates.

1.1. Related Works

Previous studies have highlighted various factors influencing university students' career choices. For instance, [4] conducted a study at Ege University and found that 41.6% of students chose their profession due to positive opinions about the field, while 34.4% felt compelled by a lack of alternatives, and

39.6% believed the profession offered certain advantages. Additionally, 33.5% were influenced by recommendations from others. These findings underscore the complexity of career decision-making processes among students, influenced by personal preferences, external pressures, and perceived benefits.

Wang et al. [6] examined the factors influencing student academic performance by applying association rule mining to analyze student behavioral data. The study found that factors such as consumption level, breakfast frequency, daily online hours, canteen meal frequency, and book borrowing amounts impacted student performance. For instance, low consumption level and high breakfast frequency were found to be associated with good academic performance. While Wang et al. identified behavioral factors affecting academic performance, this study aims to address the issue in a more comprehensive manner by examining the role of both academic habits and extracurricular interests in students' decisions regarding their chosen field of study. Our study attempts to address the issue more thoroughly by investigating the role of both academic habits and extracurricular interests in students' decisions regarding their chosen field of study, even though Wang et al. identified behavioral factors affecting academic performance.

Wang & Bai [7] used the Apriori algorithm to explore factors influencing university students' career choices. Their analysis revealed significant relationships between students' educational attributes and employment decisions. The study concluded that this method improved the accuracy of career guidance by reducing errors and enhancing data mining efficiency, helping to better prepare students for the job market.

Saa [8] explored the relationship between students' personal and social factors and their academic performance. Using data from Ajman University, the study applied data mining techniques like Naïve Bayes and decision trees to predict students' GPA. The results showed that both personal and social factors, along with academic efforts, significantly impact performance. The research highlights the value of regularly analyzing student data to benefit both universities and students.

Kim et al. [9] analyzed students' online class preferences based on gender and school levels using the Apriori algorithm. The study revealed that female students generally preferred hands-on, maker-based classes, where they could engage in creative projects using digital tools, while male students showed a stronger preference for virtual experience-based classes, where they could interact with virtual environments for learning. These findings highlight the importance of offering personalized learning environments that align with students' interests and learning styles. The research suggests that identifying such preferences can help guide students towards the right academic paths, improving both engagement and educational outcomes.

Sodanil et al. [10] aimed to investigate the factors affecting career path selection of students studying in the field of information technology. They utilized an association rule mining approach to analyze students' grades in 25 core courses. The study identified 14 association rules that potentially influence students' career path choices. These rules can be used to identify factors affecting career path selection and guide students.

The authors [11] analyzed the relationships between users' interests, professional skills and career backgrounds using a large sample of LinkedIn profiles with the aim of investigating their influence on career alignment. The study demonstrated that association rules between interests and skills mirror users' career alignment, and it proposed an association rule-based classifier for career alignment. While Si et al. examined individuals' historical interests, including past hobbies or preferred activities, our study aims to ascertain influential factors in selecting a specific university department by analyzing both past habits and the comprehensive scope of academic life.

In a study published in 2023 [12], developed a model for recommending appropriate elective courses to students using the Apriori algorithm and association rule mining. The study revealed several challenges, including students' limited ability to comprehend the relationship between elective courses and potential career pathways, their difficulty in identifying the skills that can be acquired through these courses, and their lack of sufficient knowledge about the range of career options available to them.

Ahmed et al. [13]using the academic data of 582 students from the Department of Computer Science and Engineering of Bangladesh University of Engineering and Technology, aimed to reveal the hidden information that affects factors such as student performance and retention/retention by association rule mining. In the study, a dataset containing data such as students' personal information, course grades and attendance was analyzed using the Apriori algorithm. The results showed that there are relationships between gender, dormitory residence status and courses taken that affect student achievement. Similarly, students' previous academic experiences and social characteristics could be investigated by extracting attributes associated with their behavior in dropping out of their university programs.

Mashiloane's [14] research involved the analysis of data from first-year computer science students using the Apriori algorithm to identify the optimal combinations of courses that would best support students in both computer science 1 and mathematics 1. His findings revealed a notable correlation between computer science 1 and various other academic fields, including biology, chemistry, psychology, and economics.

Wang et al. [15] conducted a study in which they used a machine learning technique called XGBoost to predict the post-university career choice of university students. Analysing a real-world dataset of 18,000 graduates ‘education and career choice information, it was found that academic data such as entrance exam score, scholarship amount and especially first semester GPA play an important role in students’ career choice. The XGBoost algorithm was able to predict students' career choice with 89.1% accuracy, 85.4% recall rate and 0.872 F1 value. Although Wang et al. ‘s study shows that data mining methods such as our study can be used on students’ career choices, our study focuses on the university department choice of vocational high school IT department students by

using association rules instead of classification and different attributes are used.

In this research, a unique approach is adopted by incorporating factors like students' favorite street games, subjects they enjoyed during primary and secondary education, personal hobbies, and habits, to explore their relationship with career choices. These often-overlooked characteristics may play a significant role in influencing career decisions.

The investigation of the factors that shape students' career decisions represents a pivotal area of inquiry within the field of education. Studies have sought to identify effective factors and determine career choice by examining students' habits and academic performance in their existing educational contexts, using data mining techniques. This study differs from the existing literature in that it examines the impact of students' social and emotional choices, interests, and habits in their previous educational experiences, as well as their academic interests and childhood habits in a holistic manner, on their career choice using association rules.

- This study contributes to the existing literature in the following ways:
- This study addresses a gap in the existing literature by evaluating how students' past educational experiences and social and emotional preferences influence their choice of university department.
- In this study, a dataset containing 20 attributes collected through a survey was used and more and different attributes were included in the analysis compared to previous studies.
- This study contributes to career guidance research by examining how habits and previous academic experiences influence students' academic interests and career choices.

2. Materials and Methods

Our study, which examines the factors influencing student career choices, utilizes association rules to analyze the factors that affect student behavior in making career decisions. Association rules are a used method in the literature for understanding student behavior [6].

In this study, a survey was conducted with the participation of current students and graduates from the Information Technology department of a MTAL. Approximately 200 individuals took part in the survey. The data collected from the survey were analyzed using the "Apriori Algorithm" as part of the "Association Rules Analysis," one of the techniques in data mining.

2.1 Data Set

The dataset for this study was collected from responses to a 20-question online survey distributed to 217 current and former students of a MTAL in Istanbul. To collect the data, in-class announcements were made, and the survey link was shared in online student groups. The variables in the dataset were determined after discussions and consultations regarding potentially influential characteristics. Association rule mining was used to analyze and compare survey responses, creating models related to university program preferences of MTAL students and graduates. Since standardized scales were not used in the survey, reliability and validity analyses were not conducted.

The dataset includes responses from a total of 118 high school students and 99 high school graduates to the administered survey. The attributes of the dataset are detailed in Appendix 1.

The dataset comprises a total of 18 attributes. The statistical ratios of the different values for these 18 attributes can be calculated across the entire dataset. Prior to performing these calculations, the records are standardized to ensure consistency, categorized appropriately, and formatted for analysis. Subsequently, the analysis of these standardized records will be conducted using the Apriori algorithm implemented in R Studio.

The Apriori algorithm is characterized by its ability to identify recurring patterns and is used to uncover frequent itemsets in databases containing various types of information [16].

These algorithms perform multiple scans over the entire dataset to detect large itemsets. During the first scan, the support value of each item is calculated and compared against the minimum support threshold defined by the user. It is then determined whether each item’s support value meets or exceeds this threshold. In subsequent scans, new itemsets are generated from the data that exceeded the support count in the previous scan. These itemsets are referred to as candidate itemsets. In each subsequent scan, the support values of the itemsets are recalculated, starting from the candidate itemsets determined in the previous scan, until the final combination of items in the database is reached. This process continues until no new candidate itemsets are found [17].

Agrawal & Srikant [18] have described the working principle and pseudo-code of the Apriori algorithm as follows:

- During the first scan of the data, all items are counted to find large itemsets.
- The next scan, let it be the k-th scan, consists of two steps:
- Using the Apriori-gen function, candidate itemsets (C_k) are generated with the itemsets $L_{\{k-1\}}$ obtained in the $(k-1)$-th scan.
- Then, the database is scanned to count the support of the candidates in (C_k).
- For a fast count, it is necessary to accurately determine the candidates forming (C_k) in a given transaction.

One of the traditional applications of the Apriori algorithm is to identify and define relationships between products sold in markets. Similarly, it can be utilized by manufacturers to optimize products or goods that are consumed or whose stock levels are declining in warehouse systems. Arranging products that are transported together on nearby shelves can reduce warehouse movements and transportation costs. This method can provide benefits such as preparing high-probability ordered items in cafes and restaurants in advance or grouping related products together in the menu.
In this study, the relationship between the vocational fields of MTAL students in the information technology departments,

their preferred university departments, and the courses, hobbies, and habits they are interested in has been revealed using association rule analysis, a data mining method, and the Apriori algorithm.

The primary focus of traditional statistical methods is not on uncovering associations or co-occurrences within data sets. Association rule mining algorithms, on the other hand, are firmly situated within the field of data mining, despite their mathematical and statistical underpinnings. This is due to their emphasis on identifying co-occurring patterns and their confidence levels. These methods have been changed to data mining extract meaningful associations and their support values from data, a task that is generally beyond the scope of traditional statistical approaches.

3. Research and Findings

Initially, the dataset was tested using R Studio with the inclusion of libraries specific to the Apriori algorithm. Based on the rules obtained, the top 100 rules with high reliability were listed. No filtering was applied and no attributes were selected during the rule calculation process. The initial 100 rules generally consisted of binary rules. As the number of attributes included in a rule increases, the probability of co-occurrence decreases. Correspondingly, the confidence values of associations with low co-occurrence probabilities also decrease. The fundamental parameters of the Apriori algorithm are support and confidence values. The total number of rules discovered before determining the support and confidence values was calculated to be 8663. The first 10 values of the 100 rules mentioned in Table 1 are listed.

Table 1 The top 10 rules with the highest confidence

Rule No	Connected Column 1	Connected Column 2	Sup. Ratio	Conf.
1	Type of software liked=Compiler software	Is there a computer at home=Yes	12%	96%
2	Hobby=Reading books	Likes doing puzzles=Yes	13%	93%
3	Hobby=Reading books	Is patient=Yes	12%	90%
4	Hobby=Reading books	Likes regular life=Yes	13%	93%
5	Hobby=Reading books	Is there a computer at home=Yes	13%	93%
6	Favorite science subject=Physics	Is a university student or graduate=No	11%	83%
7	Favorite science subject=Physics	Gender=Male	12%	90%
8	Favorite science subject=Physics	Is there a family member working in IT=No	12%	86%
9	Favorite science subject=Physics	Is there a computer at home=Yes	11%	83%
10	Desired university department=Computer Engineering	Is a high school graduate=No	14%	100%

Continuing with the study, five popular university departments in the field of information technology were selected, and filtering was applied to the data for re-analysis. The selected programs are as follows: Computer Engineering, Computer Programming, Computer Education and Instructional Technology, Web Design and Programming, and Software Engineering. Considering that these programs fall within the scope of higher education, no distinction was made between undergraduate and graduate levels in the analysis. The proportions of these selected departments in the dataset are presented in Table 2.

Table 2 Distribution of Students' Preferences for Computer Science-related Departments

Department	Num. of Preferences	Percentage
Computer Engineering	38	17 %
Computer Programming	20	10 %
Computer Education and Instructional Technology	36	16.5 %
Web Design and Coding	20	10 %
Software Engineering	20	10 %

3.1 Association Rules for the Computer Engineering Department

Association rules for students from Vocational and Technical Anatolian High Schools who wish to pursue a degree in Computer Engineering, as well as for participants who have graduated from the department, have been identified. The diagram of the top 4 rules with the highest confidence is presented in Fig. 1. A total of 1,003 rules were identified for the Computer Engineering Department, and the first four of these rules were shown in Fig. 1.

Based on the Apriori algorithm results, we can interpret the tendencies of students who want to study computer engineering as follows:

Interest in design applications and all important aspects: Students who favor design applications and consider all aspects (such as ease of use, functionality, aesthetics) important are highly likely to choose computer engineering. The rule:

$$\{Fav\_Software\_Type = Design\ applications, Imp\_Aspect = All\ of\ them\} => \{Univ\_Dept\_Wish = Computer\ Engineering\}$$

with a support of 0.0092 and a confidence of 1 suggests that students who prioritize design and its various aspects have a strong inclination toward this field.

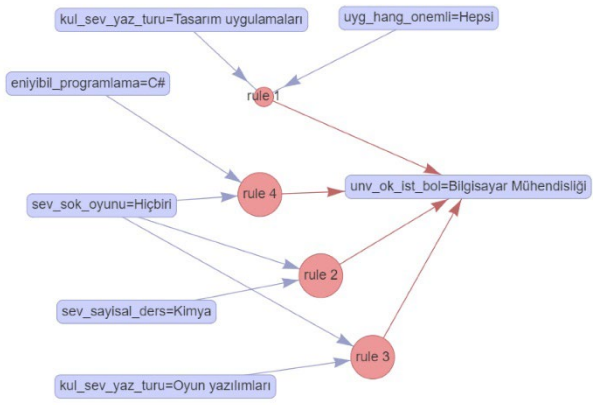


Figure 1. *Rule graph diagram for the Department of Computer Engineering*

Chemistry and lack of outdoor games: Students who select chemistry as their favorite science subject and do not engage in outdoor games tend to choose computer engineering. The rule:

$$\{Fav\_Sci\_Subject = Chemistry, Fav\_Outdoor\_Game = None\} => \{Univ\_Dept\_Wish = Computer\ Engineering\}$$

shows a support of 0.0092 and confidence of 1, indicating a clear connection between these interests and the choice of computer engineering.

Game software interest and lack of outdoor activities: Students who are interested in game software and do not play outdoor games show a strong inclination toward computer engineering. The rule:

$$\{Fav\_Software\_Type = Game\ softwares, Fav\_Outdoor\_Game = None\} => \{Univ\_Dept\_Wish = Computer\ Engineering\}$$

with a support of 0.0138 and confidence of 1 highlights that students with a passion for game software and minimal outdoor activity are likely to pursue computer engineering.

$$\{Best\_Prog\_Lang = C\#, Fav\_Outdoor\_Game = None\} => \{Univ\_Dept\_Wish = Computer\ Engineering\}$$

The results reveal a strong association between students' proficiency in the C# programming language and their preference for computer engineering, especially among those who do not participate in outdoor games. This highlights the role of specific programming skills in shaping students' academic choices. The mastery of C# appears to be a significant factor influencing the decision to pursue computer engineering, suggesting that targeted programming education can effectively guide students towards this field.

When examining the rules for students who successfully enrolled in and are currently studying in the computer engineering department, it is observed that those excelling in social sciences and chemistry and preferring design software tend to choose this major. Additionally, female students with strong social sciences backgrounds and family members working in IT are more likely to prefer computer engineering. Students who are proficient in mathematics and have an interest in compiler software, particularly those who do not play chess and have family members in the IT field, also show a tendency to select this major.

When comparing the rules for students who want to study computer engineering with those who are already studying it, some common traits emerge. Both groups highlight students excelling in social sciences and mathematics and having family members in the IT field. However, students who are currently enrolled show a more specific interest in software types (such as design or compiler software). Additionally, the fact that non-chess players stand out among the current students suggests that even hobbies can influence career choices.

It is also possible to obtain similar results for other IT departments using the Apriori algorithm. This algorithm can reveal valuable insights into career choices by identifying relationships between various factors, such as students' academic success, software preferences, and family backgrounds.

4. Results

This study analyzed the factors influencing the career choices of both students who aspire to study computer engineering and those currently enrolled in the department. The Apriori algorithm-based analysis revealed that students' success in social science subjects, their preferred types of software, and familial factors play a significant role in their career choices. Particularly, students excelling in social sciences and showing an interest in design applications are more likely to choose computer engineering. Moreover, students with family members working in the IT sector and those who do not enjoy playing chess but are interested in compiler software also tend to select computer engineering.

For female students, success in social sciences and having family members in the IT field were observed as key determinants in their choice of computer engineering. For both prospective and current students of computer engineering, family influence, academic success, and personal interests were found to be significant factors in their career decisions. Similar analyses can be conducted for other IT departments using the Apriori algorithm, and these findings can be further expanded.

5. Discussion and Conclusion

When comparing the rules for the two groups of students, it is observed that success in quantitative sciences (especially chemistry and mathematics) and software preferences are prominent for both those who wish to study computer engineering and those already studying in the department. However, family-related factors, particularly the presence of family members in the IT sector, have a stronger influence on current computer engineering students. Moreover, success in social sciences and its effect on female students were more decisive for the current students. In addition, it may be useful for female students who want to study in a field related to computer science but need encouragement to meet with family members working in this field and benefit from their experiences, professional lives and advice.

One distinctive aspect of this study is its focus on elements that have been insufficiently addressed in the existing literature. Specifically, factors such as family structure, hobbies, favorite subjects in elementary and middle school, and social activities have been examined for their impact on career choices. This research is focused on individual preferences and social contexts, with the aim of highlighting how these factors influence career paths. The study emphasizes the significance of personal and social influences in career decisions, thereby contributing to a more comprehensive and enriched understanding of the field. While previous research has often concentrated on broader, more general factors influencing career selection, this study offers a detailed analysis of how personal interests and social contexts affect career decisions. This approach provides a more in-depth and comprehensive understanding compared to earlier studies. Future research could benefit from repeating these analyses with larger datasets and across different fields, which would further enhance our understanding of these dynamics. The influence

of peer pressure, family support, and expectations on career choices could be examined in further research using similar methodologies.

Acknowledgement

This study is part of the Master of Science thesis by Esma Türk, conducted in the Department of Computer Engineering within the Institute of Natural and Applied Sciences at Tekirdağ Namık Kemal University, under the supervision of thesis advisor Erkan Özhan. The authors would like to thank the Institute for its support and all survey participants for their valuable contributions.

Declaration of Ethical Standards

The authors declare that they comply with all ethical standards.

Credit Authorship Contribution Statement

Data curation – Esma Türk (ET), Erkan ÖZHAN (EO); Formal analysis – ET, EO; Investigation – ET, EO; Experimental performance – EO; Data collection – ET, EO; Processing – ET, EO; Literature review – ET, EO; Writing – ET, EO; Visualization – EO; Review and editing – EO; Supervision – EO.

Declaration of Competing Interest

The authors have no conflicts of interest to declare regarding the content of this article.

Data Availability Statement

The data used in this study are retained for use in future research and are not currently available for public access. For further information about the data, please contact the authors.

References

[1] Alangari, N., & Alturki, R. (2020). Association rule mining in higher education: A case study of computer science students. In R. Mehmood, S. See, I. Katib, & I. Chlamtac (Eds.), *Smart Infrastructure and Applications*, EAI/Springer Innovations in Communication and Computing (pp. 311–328). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-13705-2_13.

[2] Baker, R. S. J. D. (2010). Mining data for student models. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in Intelligent Tutoring Systems*, Studies in Computational Intelligence, vol. 308 (pp. 323–337). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14363-2_16.

[3] Nie, L. (2024). College students' career prediction model based on association rule mining algorithm. In Y. Zhang & N. Shah (Eds.), *Application of Big Data, Blockchain, and Internet of Things for Education Informatization*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 584 (pp. 378–384). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-63142-9_38.

[4] Sarıkaya, T., & Khorshid, L. (2009). Üniversite öğrencilerinin meslek seçimini etkileyen etmenlerin incelenmesi: Üniversite öğrencilerinin meslek seçimi. *TEBD, 7*(2), 393–423.

[5] Kurt, T., & Fidan, T. (2021). University for career construction: Expectations and realities. *Yükseköğretim Dergisi, 11*(2Pt2), 421–437. https://doi.org/10.2399/yod.20.591001.

[6] Wang, T., Xiao, B., & Ma, W. (2022). Student behavior data analysis based on association rule mining. *International Journal of Computational Intelligence Systems, 15*(1), 32. https://doi.org/10.1007/s44196-022-00087-4.

[7] Wang, L., & Bai, Y. (2022). Research on career guidance course system based on apriori algorithm and computer big data. In *2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE)* (pp. 136–140). Ottawa, ON, Canada: IEEE. https://doi.org/10.1109/CIPAE55637.2022.00036.

[8] Saa, A. A. (2016). Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications, 7*(5). https://doi.org/10.14569/IJACSA.2016.070531.

[9] Kim, J., Hwang, D., & Lee, S.-S. (2022). An analysis of students' online class preference depending on the gender and levels of school using Apriori Algorithm. *Journal of Digital Convergence, 20*(1), 33–39. https://doi.org/10.14400/JDC.2022.20.1.033.

[10] Sodanil, M., Chotirat, S., Poomhiran, L., & Viriyapant, K. (2019). Guideline for academic support of student career path using mining algorithm. In *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval* (pp. 133–137). Tokushima Japan: ACM. https://doi.org/10.1145/3342827.3342841.

[11] Si, H., Wu, H., Zhou, L., Wan, J., Xiong, N., & Zhang, J. (2020). An industrial analysis technology about occupational adaptability and association rules in social networks. *IEEE Transactions on Industrial Informatics, 16*(3), 1698–1707. https://doi.org/10.1109/TII.2019.2926574.

[12] Concha, C. R. V., & Fabregas, A. C. (2023). Apriori algorithm applied in job forecasting with elective tracking. In *TRANSPORT, ECOLOGY - SUSTAINABLE DEVELOPMENT: EKOVarna2022* (p. 020002). Varna, Bulgaria: AIP Publishing. https://doi.org/10.1063/5.0162456.

[13] Ahmed, S., Paul, R., & Hoque, A. S. M. L. (2014). Knowledge discovery from academic data using association rule mining. In *2014 17th International Conference on Computer and Information Technology (ICCIT)* (pp. 314–319). Dhaka, Bangladesh: IEEE. https://doi.org/10.1109/ICCITechn.2014.7073107.

[14] Mashiloane, L. (2014). Using association rule mining to find the effect of course selection on academic performance in computer science. In R. Prasath, P. O'Reilly, & T. Kathirvalavakumar (Eds.), *Mining Intelligence and Knowledge Exploration*, Lecture Notes in Computer Science, vol. 8891 (pp. 323–332). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-13817-6_31.

[15] Wang, Y., Yang, L., Wu, J., Song, Z., & Shi, L. (2022). Mining campus big data: Prediction of career choice using interpretable machine learning method. *Mathematics, 10*(8), 1289. https://doi.org/10.3390/math10081289.

[16] Sözen, E., Bardak, T., Peker, H., & Bardak, S. (2017). Apriori algoritması kullanılarak mobilya seçimde etkili olan faktörlerin analizi. *İleri Teknoloji Bilimleri Dergisi, 6*(3), 679–684.

[17] Öztemiz, F. (2017). *Apriori Algoritması ile Müşteri Bazlı Market Sepet Analizi ve Ürün Satış Tahmini* (MSc. Thesis). İnönü University, Malatya, Turkey. Retrieved from https://avesis.inonu.edu.tr/yonetilen-tez/b6a95449-a644-44f5-a254-fb3ddfad0953/apriori-algoritmasi-ile-musteri-bazli-market-sepet-analizi-ve-urun-satis-tahmini.

[18] Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94)* (pp. 487–499). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.