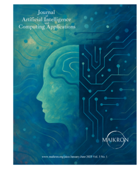
Journal of Artificial Intelligence and Computing Applications (2025) 3(1): 1–12

MAIKRON

Journal of Artificial Intelligence and Computing Applications



Research article

Analysis of the sleep quality of college students from different knowledge areas using a data mining approach

Andrea Morales-Robles[1,*], **Víctor Menéndez-Domínguez**[1], **and Héctor Rubio-Zapata**[1]

[1]Universidad Autónoma de Yucatan

ABSTRACT

Sleep is an essential physiological process involved in memory consolidation, metabolic and endocrine homeostasis, and immune system regulation. Therefore, good sleep quality is vital for maintaining physiological homeostasis. Poor sleep quality is prevalent among college students and affects both physical and mental health. Conventional statistical methods, such as logistic regression, are commonly used to generate predictive models of sleep quality and have been extensively applied to Health Sciences students, but their use has been less studied among students from other disciplines, such as Engineering and Exact Sciences. Data mining can help overcome certain limitations of these conventional methods, such as multicollinearity, by uncovering associations that might otherwise have gone unnoticed. In this study, we separately analyzed two samples of students from Health Sciences and Engineering and Exact Sciences. We found significant correlations between sleep quality and attributes such as perceived sleep quality, sleep latency, sleep duration, drug use, and the use of medication for depression and anxiety. Decision trees identified different predictive attributes between the two samples. These findings offer a novel insight into sleep quality among college students and may support informed decision-making and targeted interventions.

Keywords: sleep quality, college students, data mining

1. Introduction

Sleep is a physiological process that occupies nearly one-third of the human lifespan [1]; it is characterized by its reversibility and circadian periodicity, during which immobility and muscle relaxation are accompanied by a decrease in consciousness and responsiveness to external stimuli [2]. Given its crucial role in essential functions, such as memory consolidation, metabolic, endocrine and immune regulation [2], maintaining good sleep quality is vital.

College students are among the most extensively researched populations in relation to sleep, with Health Sciences students being the most studied subgroup. Poor sleep quality is highly prevalent in this population, affecting over 50% of students [3, 4, 5]. In contrast, students from other disciplines, such as Engineering and Exact Sciences, have been scarcely researched. However, literature reports a prevalence of poor sleep quality in this population between 48.7% and 67.8% [6, 7, 8].

Sleep quality is a multifactorial construct influenced by elements such as age, sex, the sleep environment, and comorbidities, among others [9]. Some of these factors may increase the risk of poor sleep quality, while others may serve as protective factors. Alghwiri et al., through logistic regression analysis, identified electronics usage hours, neck pain, headache, and other systemic diseases as risk factors for sleep quality [10]. Similarly, Li et

E-mail address: andreamoralesrobles@outlook.com
https://doi.org/10.5281/zenodo.16819159

al. found that college students in lower academic years (e.g., freshmen and sophomores), those who consumed alcohol or engaged in gambling, exercised less than once per week, or felt pressured during the academic year were at higher risk of experiencing poor sleep quality [11]. Although these approaches provide valuable information regarding students' sleep quality and may be the basis of therapeutic management, they may be limited by issues such as multicollinearity or imprecise estimates when some variable categories have a small number of cases. Data mining can help address these limitations, especially in the development of predictive models and in supporting clinical decision-making [12].

The aim of this study was to analyze the sleep quality of students from two campuses —Health Sciences and Engineering and Exact Sciences— using a data mining approach.

2. Method

2.1 Study design and participants

A cross-sectional study was conducted using probabilistic sampling. A total of 805 students from the Health Sciences Campus and 92 students from the Engineering and Exact Sciences Campus of the Universidad Autónoma de Yucatán (UADY) were recruited. The distribution of students by academic program on each campus was as follows:

- Health Sciences Campus: 32% medical students, 9.2% Nutrition students, 8.6% Rehabilitation students, 20.4% Odontology students, 18% Pharmaceutical Chemist Biologist (PCB) students, 8.1% Nursing students, and 3.7% Social Work Students.
- Engineering and Exact Sciences Campus: 19.6% Computer Sciences students, 2.8% Computer Engineering students, and 78.3% Software Engineering students.

All regular students of both sexes, aged 18 or older, enrolled in the first to sixth school semester, who agreed to participate voluntarily, were considered for inclusion. Students who were pregnant or working night shifts were excluded to minimize potential bias in sleep quality assessment.

2.2 Instruments

2.2.1 Sociodemographic data

Data regarding sociodemographic information was collected through a self-administered 28-item survey. As with the other instruments, this questionnaire was applied using Microsoft Forms, which students could access by scanning a QR code.

2.2.2 Sleep quality

The Pittsburgh Sleep Quality Index (PSQI) was administered to assess students' sleep quality; the Spanish version by Jimenez-Genchi et al. was used, with a Cronbach's alpha reliability coefficient of 0.78 [13]. PSQI is a self-administered 19-item questionnaire developed by Buysse in 1989, that evaluates seven components of sleep quality: subjective (perceived) sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction [14]. A global score, given by the sum of each component's score, of more than five points indicate poor sleep quality [13].

2.2.3 Daytime sleepiness

Since sleep quality is a multifactorial phenomenon, we evaluated some confounding factors, daytime sleepiness being one of them, which was evaluated by the Epworth Sleepiness Scale (ESS). This self-administered 8-item questionnaire was applied to evaluate students' tendency to fall sleep in different situations; the Spanish version by Sandoval-Rincón et al. was used, with a Cronbach's alpha reliability coefficient of 0.89. A total score of 10 to 12 points is suggestive of marginal sleepiness, while more than 12 points indicate excessive sleepiness [15].

2.2.4 Internet addiction

Other confounding factor was the presence of Internet addiction, assessed by the Internet Addiction Test (IAT); the Spanish version by Puerta-Cortés et al. was administered, with a Cronbach's alpha reliability coefficient of 0.91 [16]. This 20-item questionnaire distinguishes between four different levels of Internet addiction: no addiction (total score of 0 to 30 points), mild addiction (total score of 31 to 49 points), moderate addiction (total score of 50 to 79 points) and severe addiction (total score of 80 to 100 points) [17].

2.3 Data analysis

Data was analyzed through a data mining approach using the Waikato Environment for Knowledge Analysis (WEKA), version 3.9.6 (Figure 1).

A separate dataset was constructed for each campus, enabling independent analysis of the two populations. Initially, both datasets underwent preprocessing, which included removing inconsistent or incongruent information. A total of 55 attributes were selected for analysis in each dataset. The Health Sciences dataset included 805 instances, while the Engineering and Exact Sciences dataset comprised 92 instances. Then, a discretization process was employed, to transform continuous attributes into nominal ones, making data suitable for data mining algorithms [18].

Three data mining techniques were used to analyze data: tree-based classification, clustering, and association rule mining. For the classification analysis, the J48 decision tree algorithm was applied, with 10-fold cross-validation and the Sleep Quality variable as the reference attribute. For clustering, the SimpleKMeans algorithm was applied, using the Sleep Quality attribute

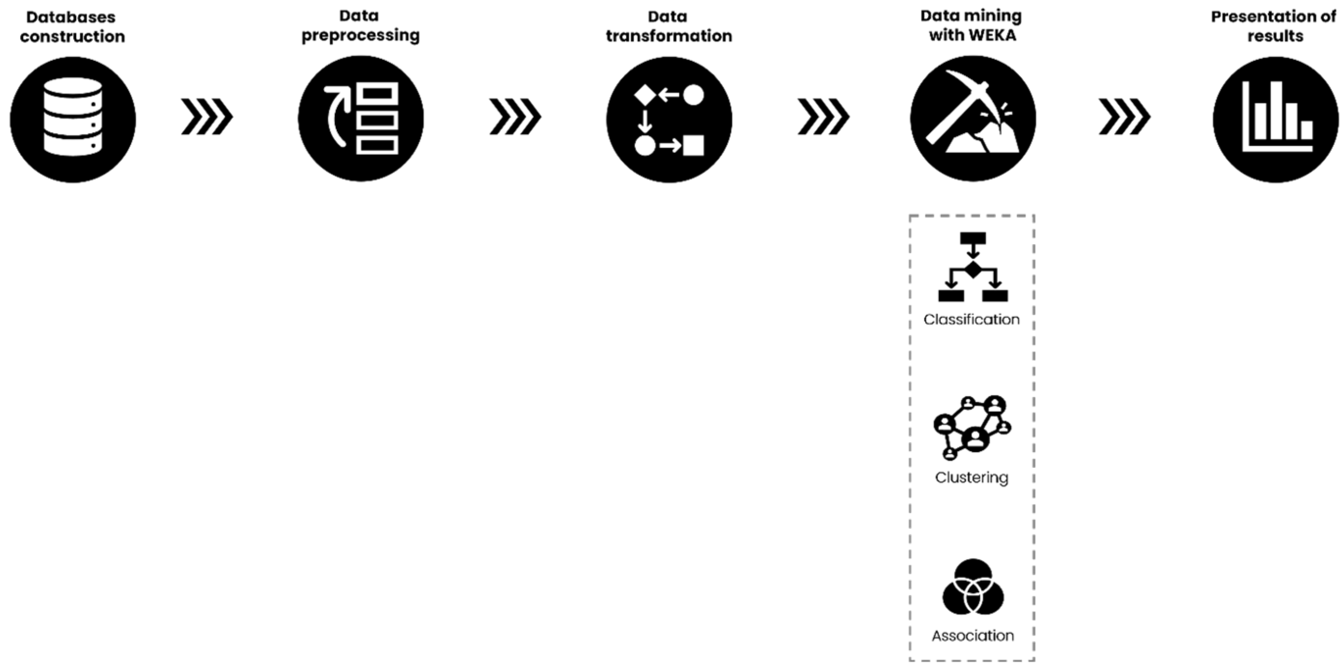


Figure 1. Study methodology. Data mining techniques used were tree-based classification (J48 algorithm), clustering (SimpleKMeans algorithm) and association rule mining (Apriori algorithm).

to evaluate cluster alignment. For association analysis, the Apriori algorithm was applied. All algorithms were unsupervised, except for the J48 algorithm.

3. Results

From the total sample of 897 students, 89.7% were from the Health Sciences Campus ($n = 805$), and 10.3% were from the Engineering and Exact Sciences Campus ($n = 92$). The prevalence of poor sleep quality among Health Sciences students —61.1% of whom were female— was 74.4%. Among Engineering and Exact Sciences students —18.5% of whom were female— the prevalence was 67.4%.

3.1 Data mining

All algorithms were applied separately for each campus dataset in the following order: tree-based classification to generate a decision tree for sleep quality prediction, clustering analysis to characterize students with good and poor sleep quality, and association rule mining to extract the most relevant rules for each campus.

3.1.1 Tree-based classification algorithm

As previously mentioned, the J48 algorithm was used to generate decision trees.

The algorithm was first applied to the Health Sciences dataset. Sleep Quality was used as the reference attribute, which has two possible values, according to the PSQI: good (equal to or less than five points) or poor (more than five points). A 10-fold cross-validation was performed. To optimize model performance, the MinNumObj parameter was set to 5, and to reduce overfitting, the confidence factor was set to 0.25. The accuracy of the resulting decision tree was 84.9689%, correctly classifying 684 out of 805 instances. Model performance metrics were as follows: Kappa statistic = 0.5903, mean absolute error = 0.2208, root mean squared error = 0.3463, relative absolute error = 57.9151%, and root relative squared error = 79.351% (Figure 2).

Then, the algorithm was applied to the Engineering and Exact Sciences dataset, again using Sleep Quality as the reference attribute and 10-fold cross-validation. For this analysis, the tree was left unpruned to maximize accuracy and minimize error. The MinNumObj parameter was set to 2. The resulting decision tree achieved an accuracy of 81.5217%, correctly classifying 75 out of 92 instances. Model performance metrics were as follows: Kappa statistic = 0.5969, mean absolute error = 0.1964, root mean squared error = 0.4033, relative absolute error = 44.5295%, and root relative squared error = 86.0252% (Figure 3).

3.1.2 Clustering algorithm

The SimpleKMeans algorithm was employed for this analysis. Using the classes to cluster evaluation mode, Sleep Quality served as the reference for evaluating cluster alignment. For each dataset, two clusters were generated: one predominantly associated with good sleep quality and the other with poor sleep quality. The most relevant attributes for each cluster are presented in Figures 4 and 5, respectively.

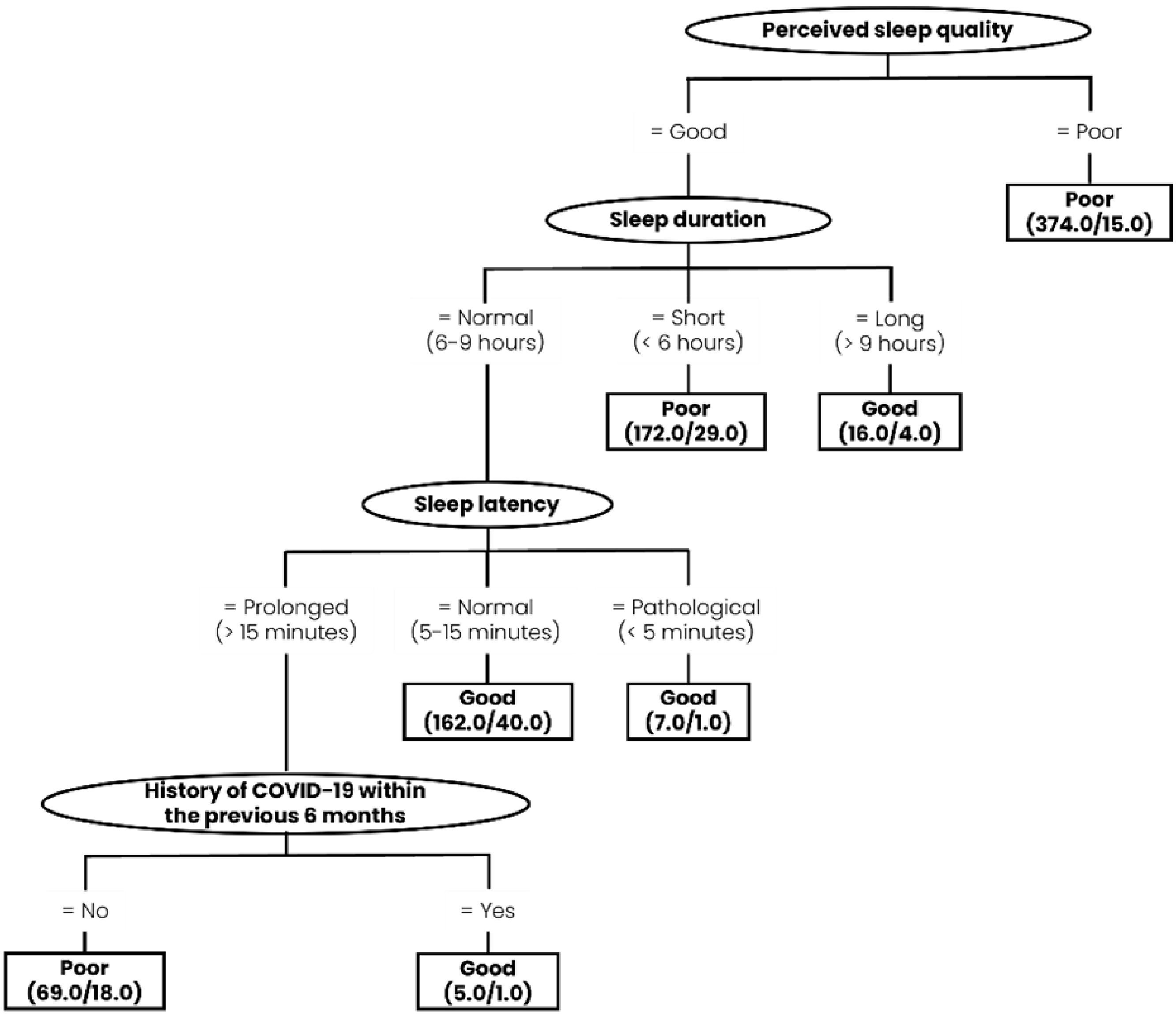


Figure 2. Decision tree (J48 algorithm) for sleep quality among Health Sciences students. To optimize model performance, the MinNumObj parameter was set to 5, with a confidence factor of 0.25.

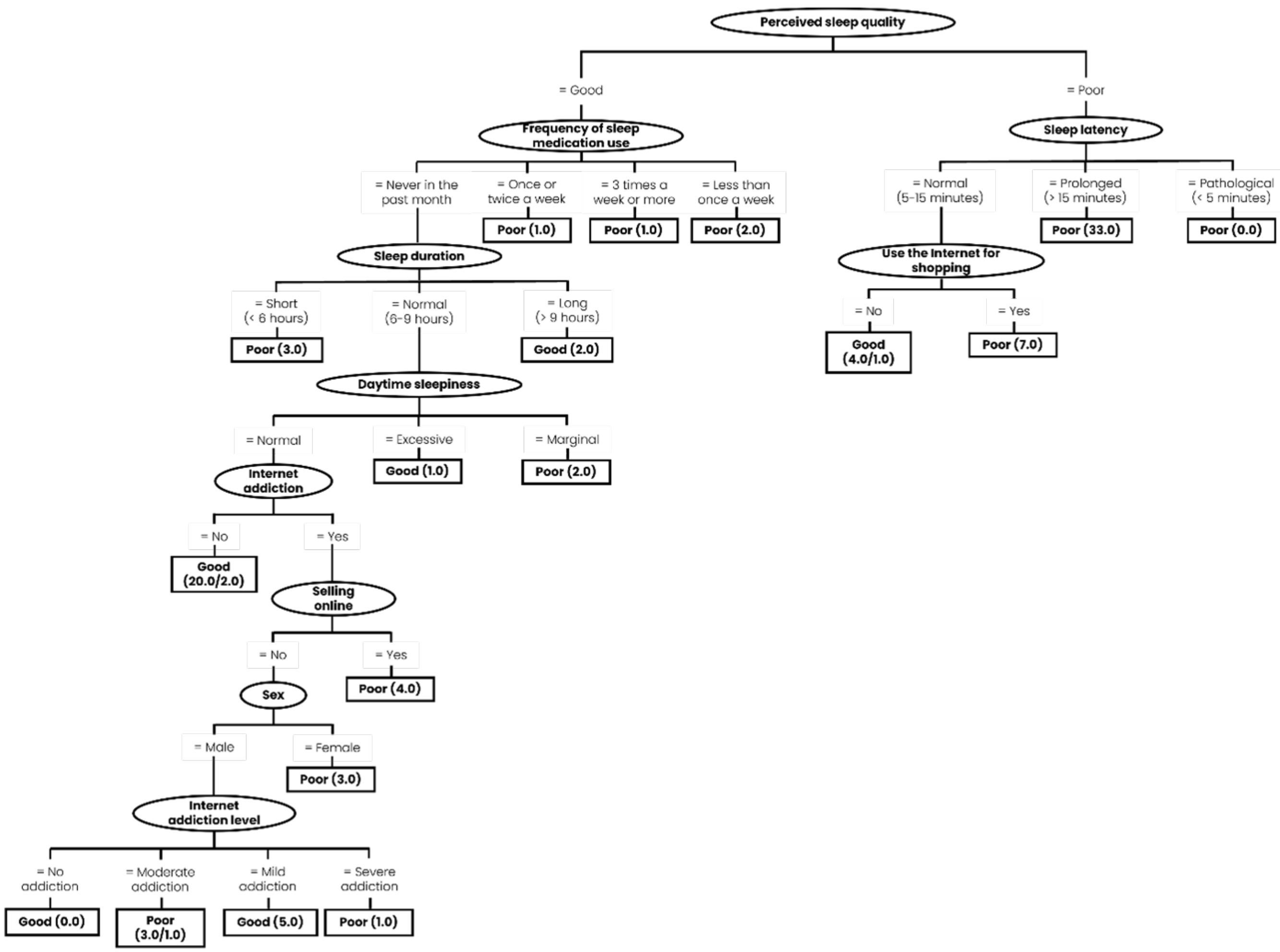


Figure 3. Decision tree (J48 algorithm) for sleep quality among Engineering and Exact Sciences students. The tree was left unpruned to maximize accuracy and minimize error; the MinNumObj parameter was set to 2.

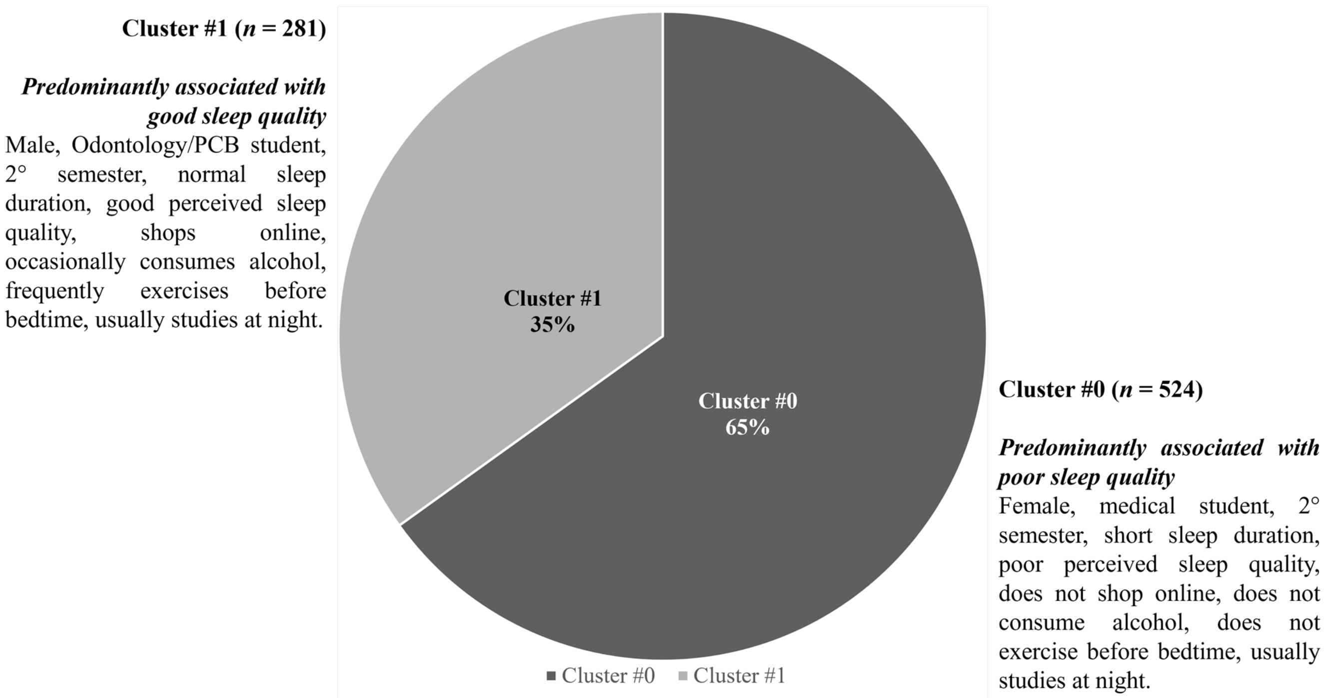


Figure 4. Generated clusters (SimpleKMeans algorithm) for the Health Sciences dataset, based on students' sleep quality. The most relevant attributes are presented.

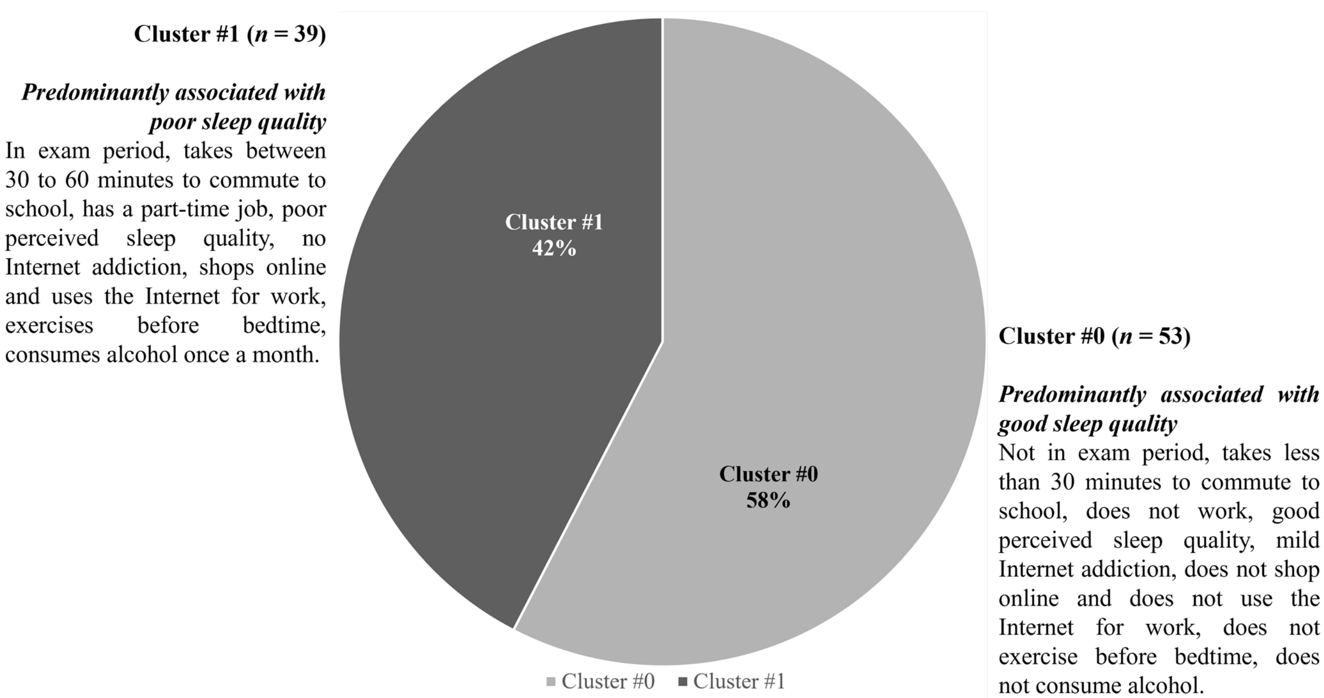


Figure 5. Generated clusters (SimpleKMeans algorithm) for the Engineering and Exact Sciences dataset, based on students' sleep quality. The most relevant attributes are presented.

In terms of clustering accuracy, 34.4099% of instances were incorrectly clustered in the Health Sciences dataset. In contrast, only 53.26% of instances were correctly clustered in the Engineering and Exact Sciences dataset. These results suggest that additional, unmeasured attributes may be influencing students' sleep quality and limiting the clustering model's discriminative power.

To optimize clustering model performance, several parameter configurations were tested. The most accurate results were obtained using the following settings: Euclidean Distance as the distance function, k-means++ as the initialization method, a maximum of 500 iterations, two clusters, preservation of instances order enabled, and a random seed of 100. This configuration was applied to both datasets.

3.1.3 Association algorithm

The Apriori algorithm was applied for this analysis. For the Health Sciences dataset, the algorithm performed 79 cycles, using a minimum support threshold of 0.92 (741 instances) and a minimum confidence threshold of 0.4. For the Engineering and Exact Sciences Campus dataset, 60 cycles were performed with a minimum support threshold of 0.94 (86 instances) and the same confidence threshold of 0.4.

In both datasets, the algorithms were run with the RemoveUseless filter enabled, and parameters were set as follows: number of rules = 50, delta = 0.001, upper bound of minimum support = 1.0, lower bound of minimum support = 0.3, and significance level = -1.0. The most meaningful rules based on relevance, confidence, lift, leverage and conviction metrics are presented in Tables 1 and 2. Conviction values for all rules are greater than 1, indicating reliable associations between attributes. Additionally, lift values above 1 and leverage values greater than 0 suggest positive correlations among the associated attributes.

4. Discussion

According to the tree-based classification algorithm, perceived sleep quality emerged as a key variable in evaluating a student's overall sleep quality, as it served as the root node in both decision trees. Since perceived sleep quality is a component of the PSQI, including this variable into a traditional logistic regression model may introduce multicollinearity among independent variables, limiting its value as a predictive factor. However, in a real-world clinical context, this attribute can provide useful information and help guide the assessment toward a more informed evaluation.

When analyzing the decision tree for the Health Sciences dataset, only three additional attributes —besides the root node— were identified as predictors of sleep quality: sleep duration, sleep latency and history of COVID-19 within the previous six months. As in the previous cases, the predictive value of sleep duration and sleep latency is limited. However, longer sleep duration and shorter sleep latency contribute to a lower PSQI overall score, which is interpreted as better sleep quality [13].

Regarding the history of COVID-19 within the previous six months, current evidence presents an ambiguous relationship with sleep quality, potentially linked to the role of sleep in modulating the immune response [19]. In their evaluation of the long-term effects of COVID-19, Mekhael et al. found out that patients who had recovered from the disease exhibited altered sleep architecture, shorter sleep duration and longer sleep latency [20], features commonly associated with poor sleep quality. Interestingly, the decision tree for Health Sciences students indicated that those who had COVID-19 within the previous six months were more likely to report good sleep quality. However, given the limited evidence regarding the association between COVID-19 and sleep quality, it cannot be concluded that a positive history of COVID-19 predicts good sleep quality. Yet, despite the evidence suggesting a negative impact of COVID-19 on sleep, our results indicate that students who had a positive history of COVID-19 within the previous six months were more likely to report good sleep quality, contradictory to the existing evidence.

The Engineering and Exact Sciences Campus decision tree also identified perceived sleep quality as the root node. In addition to sleep duration and sleep latency, other relevant predictors for sleep quality included the frequency of sleep medication use, daytime sleepiness, selling online and Internet addiction. Our results suggest that any frequency of sleep medication use is associated with poor sleep quality. This aligns with Alamir et al. findings, who reported that individuals consuming anxiolytics and antidepressants experienced nocturnal awakenings, daytime sleepiness and insomnia, factors associated with poor sleep quality [21].

Regarding daytime sleepiness, de Sousa et al., using an explanatory model, predicted a higher probability of daytime sleepiness among college students with poor sleep quality [22]. In contrast, our findings showed that those with excessive daytime sleepiness were more likely to report good sleep quality. This apparent contradiction may be explained by the possibility of a distorted perception of sleep as being restorative when the individual experiences greater fatigue.

In this study, female students showed a higher probability of experiencing poor sleep quality, consistent with the findings of Sánchez-Sánchez, who reported that college-aged women had a 13% higher risk of developing poor sleep quality [23]. According to Regal et al., this may be related to the influence of female hormones, a higher prevalence of stress and the onset of disorders such as depression or anxiety [24].

In relation to Internet addiction, our results indicate that students with no or mild addiction are more likely to experience good sleep quality, whereas those with moderate or severe addiction tend to experience poor sleep quality. These findings are consistent with a meta-analysis by Alimoradi et al., which concluded that individuals with Internet addiction have twice the risk of experiencing sleep disorders, daytime sleepiness,

Table 1. Generated association rules (Apriori algorithm) for Health Sciences students.

Rule	Description	Confidence	Lift	Leverage	Conviction
1	If a student never consumes drugs, then they do not consume drugs.	1.00	1.06	0.05	42.48
2	If a student does not consume drugs, then they never consume drugs.	1.00	1.06	0.05	42.48
3	If a student is single and never consumes drugs, then they do not consume drugs.	1.00	1.06	0.05	41.42
4	If a student is single and does not consume drugs, then they never consume drugs.	1.00	1.06	0.05	41.42
5	If a student never consumes drugs and does not take medication for depression, then they do not consume drugs.	1.00	1.06	0.05	41.42
6	If a student does not consume drugs and does not take medication for depression, then they never consume drugs.	1.00	1.06	0.05	41.42
7	If a student never consumes drugs, then they are single and do not consume drugs.	0.97	1.06	0.05	3.02
8	If a student does not consume drugs, then they are single and never consume drugs.	0.97	1.06	0.05	3.02
9	If a student never consumes drugs, then they do not consume drugs and do not take medication for depression.	0.97	1.06	0.05	3.02
10	If a student does not consume drugs, then they never consume drugs and do not take medication for depression.	0.97	1.06	0.05	3.02
11	If a student uses the Internet for communication but does not take medication for anxiety, then they do not take medication for depression.	0.99	1.02	0.02	2.78
12	If a student has normal sleep efficiency and does not take medication for depression, then they do not take medication for anxiety.	0.99	1.02	0.02	2.57
13	If a student does not take medication for anxiety, then they do not take medication for depression.	0.99	1.02	0.02	2.55
14	If a student is single and does not take medication for anxiety, then they do not take medication for depression.	0.99	1.02	0.02	2.48
15	If a student has normal sleep efficiency and does not take medication for anxiety, then they do not take medication for depression.	0.99	1.02	0.02	2.45
16	If a student does not take medication for depression, then they do not take medication for anxiety.	0.99	1.02	0.02	2.38
17	If a student is single and does not take medication for depression, then they do not take medication for anxiety.	0.99	1.02	0.02	2.31
18	If a student uses the Internet for communication but does not take medication for depression, then they do not take medication for anxiety.	0.99	1.02	0.02	2.27

Table 2. Generated association rules (Apriori algorithm) for Engineering and Exact Sciences students.

Rule	Description	Confidence	Lift	Leverage	Conviction
1	If a student never consumes drugs, then they do not consume drugs.	1.00	1.06	0.05	4.73
2	If a student does not consume drugs, then they never consume drugs.	1.00	1.06	0.05	4.73
3	If a student is single and never consumes drugs, then they do not consume drugs.	1.00	1.06	0.05	4.67
4	If a student is single and does not consume drugs, then they never consume drugs.	1.00	1.06	0.05	4.67
5	If a student never consumes drugs and has no history of COVID-19 within the previous six months, then they do not consume drugs.	1.00	1.06	0.05	4.67
6	If a student does not consume drugs and has no history of COVID-19 within the previous six months, then they never consume drugs.	1.00	1.06	0.05	4.67
7	If a student never consumes drugs, then they are single and do not consume drugs.	0.99	1.06	0.05	2.84
8	If a student does not consume drugs, then they are single and never consume drugs.	0.99	1.06	0.05	2.84
9	If a student never consumes drugs, then they do not consume drugs and have no history of COVID-19 within the previous six months.	0.99	1.06	0.05	2.84
10	If a student does not consume drugs, then they never consume drugs and have no history of COVID-19 within the previous six months.	0.99	1.06	0.05	2.84
11	If a student does not take medication for anxiety, then they do not take medication for depression.	1.00	1.02	0.02	1.93
12	If a student does not take medication for anxiety and has no history of COVID-19 within the previous six months, then they do not take medication for depression.	1.00	1.02	0.02	1.91
13	If a student is single and does not take medication for anxiety, then they do not take medication for depression.	1.00	1.02	0.02	1.89
14	If a student uses the Internet for entertainment but does not take medication for anxiety, then they do not take medication for depression.	1.00	1.02	0.02	1.87
15	If a student is single, does not take medication for anxiety and has no history of COVID-19 within the previous six months, then they do not take medication for depression.	1.00	1.02	0.02	1.87

and consequently, poor sleep quality [25]. It is likely that individuals with mild addiction share more characteristics with those who have no addiction than with those who have moderate or severe addiction, making them more likely to experience good rather than poor sleep quality, offering valuable insights not only for the clinical approach to these individuals but also for their management. Additionally, this attribute may have been particularly relevant among Engineering and Exact Sciences students, where the frequent use of electronic devices and constant Internet access is practically unavoidable due to the nature of their academic programs.

When examining the generated clusters, poor sleep quality appears to be prevalent across both datasets, which is consistent with current evidence indicating a prevalence rate above 60% [3, 4, 5]. Although specific evidence regarding the Engineering and Exact Sciences students is limited, since most existing studies focus on college students in general or specifically in healthcare students, the literature review revealed a prevalence of poor sleep quality of 48.7% among students from the fields of Informatics or Mechanical Engineering [8], and between 65.1% and 67.8% among students from Sciences and Engineering disciplines [6, 7].

Although the four generated clusters share some attributes, it is important to highlight certain distinguishing features. For example, in the Health Sciences campus, poor sleep quality was associated with female students in their second school semester. This finding aligns with the results reported by Li et al., who stated that college students in lower academic years are at a higher risk of experiencing poor sleep quality [11]. In contrast, good sleep quality among these students was associated with occasional consumption of alcohol, which has a sedative effect at low doses. However, with repeated and chronic consumption, this substance may lead to sleep disorders and poor sleep quality [26]. Conversely, unlike the cluster predominantly associated with poor sleep quality in the Health Sciences dataset, students from the Engineering and Exact Sciences campus showed an association between poor sleep quality and alcohol consumption once a month. This contradictory finding may be due to unexamined factors related to alcohol use itself, such as quantity consumed or type of alcohol.

Finally, according to the association algorithm, for students of both campuses, particular attention should be paid to the consumption of drugs, and medication for depression and anxiety as this attributes where strongly correlated with a confidence ranging between 0.97 and 1. As mentioned before, poor sleep quality is related to the consumption of anxiolytics and antidepressants [21], but also psychoactive substances, like alcohol, caffeine, or drugs, especially in those who consume them chronically [26].

4.1 Proposal of a Decision Support System based on AI models

Given the relevance of the identified predictors of sleep quality and the structure of the generated models, we propose the development of a simple, interpretable Decision Support System (DSS). This system would align with AI-assisted mental health frameworks and serve as an early identification tool for students at risk of poor sleep quality, thereby promoting preventive strategies.

The proposed DSS would integrate the three modeling approaches used in this study —decision trees, clustering, and association rule mining— and operate in the following sequence:

1. Input module: The system receives data from a brief student self-report instrument, including key variables identified in this study, such as perceived sleep quality, sleep latency, use of medications and Internet usage habits.

2. Risk classification: Using a trained decision tree model, the system classifies each student as likely to have good or poor sleep quality, supported by interpretable rules derived from the tree structure.

3. Profiling and feedback: The student is matched to one of the identified clusters (e.g., low-risk profile, high-risk profile with medication use, etc.), and a brief narrative is provided to describe the student's risk profile, including associated lifestyle or academic habits.

4. Recommendation engine: Based on the most relevant association rules, the system generates tailored recommendations, such as seeking psychological counseling, limiting electronic device usage before bedtime, or referral to a health professional.

5. Follow-up planning: The system stores the classification result for follow-up or intervention planning.

This DSS would emphasize interpretability, transparency and utility in non-specialized settings such as universities, making it a valuable tool for early intervention in students' mental health.

5. Conclusions

Poor sleep quality is a highly prevalent problem among students in both Health and Exact Sciences disciplines. Current evidence indicates that several attributes are associated with and can predict poor sleep quality. However, analysis using a data mining approach revealed additional, potentially unnoticed attributes, including perceived sleep quality, sleep latency or sleep duration. The decision trees and the profiles of students with poor sleep quality differed between campuses, identifying key factors specific to each population.

These findings underscore the importance of alternative analytical approaches that can overcome the limitations of conventional statistical methods, such as multicollinearity, and generate insights that may be more applicable in real-world clinical scenarios for predicting diagnoses and supporting informed therapeutic decision-making.

The proposed DSS may help address this problem by integrating the data mining techniques used in this study, offering a useful tool for early intervention in students' mental health in alignment with AI-assisted frameworks.

Ethics Statement

This study was reviewed and approved by the Ethics Committee of the Faculty of Medicine at the Universidad Autónoma de Yucatán under protocol number 01-2023. All participants were informed about the study and provided voluntary consent prior to their involvement. This study used anonymized data collected via an online survey. No personally identifiable information was collected.

CRediT authorship contribution statement

Andrea Morales-Robles: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Software, Visualization, Writing – original draft, Writing – review & editing. **Víctor Menéndez-Domínguez:** Conceptualization, Methodology, Project administration, Supervision, Software, Writing – review & editing. **Héctor Rubio-Zapata:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Microsoft Copilot in order to improve readability. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

The authors would like to acknowledge the Health Sciences Campus and the Engineering and Exact Sciences Campus of the Universidad Autónoma de Yucatán for the facilities provided during the completion of this study. They also extend their gratitude to all the university students who voluntarily agreed to participate.

References

[1] L. Fabres and P. Moya, "Sueño: conceptos generales y su relación con la calidad de vida," *Revista Médica Clínica Las Condes*, vol. 32, no. 5, pp. 527–534, 2021.

[2] P. Carrillo-Mora, J. Ramírez-Peris, and K. Magaña-Vázquez, "Neurobiología del sueño y su importancia: antología para el estudiante universitario," *Revista de la Facultad de Medicina de la UNAM*, vol. 56, no. 4, pp. 5–15, 2013.

[3] M. J. Tahir, N. I. Malik, I. Ullah, H. R. Khan, S. Perveen, R. Ramalho, A. R. Siddiqi, S. Waheed, M. M. Mohamed Shalaby, D. De Berardis, S. Jain, G. L. Vetrivendan, H. Chatterjee, W. X. Gopar Franco, M. A. Shafiq, N. T. Fatima, M. Abeysekera, Q. Sayyeda, S. F. Shamat, W. Aiman, Q. Akhtar, A. Devi, A. Aftab, S. Shoib, C.-Y. Lin, and A. H. Pakpour, "Internet addiction and sleep quality among medical students during the covid-19 pandemic: A multinational cross-sectional survey," *PLoS ONE*, vol. 16, no. 11, 2021.

[4] A. Mohammadbeigi, R. Absari, F. Valizadeh, M. Saadati, S. Sharifimoghadam, A. Ahmadi, M. Mokhtari, and H. Ansari, "Sleep quality in medical students; the impact of over-use of mobile cell-phone and social networks," *Journal of Research in Health Sciences*, vol. 16, no. 1, pp. 46–50, 2016.

[5] A. Morales-Robles, "Asociación entre la calidad de sueño y la adicción al internet en estudiantes de diferentes licenciaturas del área de la salud," Master's thesis, Universidad Autónoma de Yucatán, Mérida, México, 2024.

[6] N. Hasim, J. Yusof, F. A. Ibrahim, M. S. Bungsu, N. A. Yakub, and N. I. Shari, "Stress and sleep quality among engineering students," *Sains Humanika*, vol. 15, no. 3, pp. 29–40, 2023.

[7] T. Kuhn, S. Karsan, J. J. Heisz, and L. E. Middleton, "The differing relationships between academic discipline, sleep hygiene, and dysfunctional sleep attitudes on sleep quality and duration in canadian university students," *Frontiers in Psychology*, vol. 15, 2024.

[8] M. J. Schmickler, S. Blaschke, R. Robbins, and F. Mess, "Determinants of sleep quality: A cross-sectional study in university students," *International Journal of Environmental Research and Public Health*, vol. 20, no. 3, 2023.

[9] J. C. Sierra, C. Jiménez-Navarro, and J. D. Martín-Ortiz, "Calidad del sueño en estudiantes universitarios: importancia de la higiene del sueño," *Salud Mental*, vol. 25, no. 6, pp. 35–43, 2002.

[10] A. A. Alghwiri, F. Almomani, A. A. Alghwiri, and S. L. Whitney, "Predictors of sleep quality among university students: the use of advanced machine learning techniques," *Sleep and Breathing*, vol. 25, no. 2, pp. 1119–1126, 2021.

[11] Y. Li, W. Bai, B. Zhu, R. Duan, X. Yu, W. Xu, M. Wang, H. Wanqing, W. Yu, W. Li, and C. Kou, "Prevalence and correlates of poor sleep quality among college students: a cross-sectional survey," *Health and Quality of Life Outcomes*, vol. 18, no. 210, 2020.

[12] W.-T. Wu, Y.-J. Li, A.-Z. Feng, L. Li, T. Huang, A.-D. Xu, and J. Lyu, "Data mining in clinical big data: the frequently used databases, steps, and methodological models," *Military Medical Research*, vol. 8, no. 44, 2021.

[13] A. Jiménez-Genchi, E. Monteverde-Maldonado, A. Nenclares-Portocarrero, G. Esquivel-Adame, and A. de la Vega-Pacheco, "Confiabilidad y análisis factorial de la versión en español del Índice de calidad de sueño de pittsburgh en pacientes psiquiátricos," *Gaceta Médica de México*, vol. 144, no. 6, pp. 491–496, 2008.

[14] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The pittsburgh sleep quality index: A new instrument for psychiatric practice and research," *Psychiatry Research*, vol. 28, pp. 193–213, 1989.

[15] M. Sandoval-Rincón, R. Alcalá-Lozano, I. Herrera-Jiménez, and A. Jiménez-Genchi, "Validation of the epworth sleepiness scale in mexican population," *Gaceta Médica de México*, vol. 149, no. 4, pp. 409–416, 2013.

[16] D. X. Puerta-Cortés, X. Carbonell, and A. Chamarro, "Análisis de las propiedades psicométricas de la versión en español del internet addiction test," *Trastornos Adictivos*, vol. 14, no. 4, pp. 99–104, 2012.

[17] K. S. Young and C. N. de Abreu, *Internet Addiction: A handbook and guide to evaluation and treatment.* Hoboken, New Jersey: John Wiley & Sons, Inc., 2011.

[18] A. Rajalakshmi, R. Vinodhini, and K. Fathima Bibi, "Data discretization technique using weka tool," *International Journal of Computer Science Engineering and Technology*, vol. 6, no. 8, pp. 293–298, 2016.

[19] L. Shaik, S. Boike, K. Ramar, S. Subramanian, and S. Surani, "Covid-19 and sleep disturbances: A literature review of clinical evidence," *Medicina*, vol. 59, no. 5, 2023.

[20] M. Mekhael, C. H. Lim, A. H. El Hajjar, C. Noujaim, C. Pottle, N. Makan, L. Dagher, Y. Zhang, N. Chouman, D. L. Li, T. Ayoub, and N. Marrouche, "Studying the effect of long covid-19 infection on sleep quality using wearable health devices: Observational study," *Journal of Medical Internet Research*, vol. 24, no. 7, 2022.

[21] Y. A. Alamir, K. J. Zullig, S. Wen, H. Montgomery-Downs, A. L. Kristjansson, R. Misra, and J. Zhang, "Association between nonmedical use of prescription drugs and sleep quality in a large college student sample," *Behavioral Sleep Medicine*, vol. 17, no. 4, pp. 470–480, 2017.

[22] E. A. de Sousa, L. A. Freires, G. D. Lopes Loureto, and J. C. Albuquerque da Costa, "Sleep quality and daytime sleepiness among college students: testing an explanatory model," *Ciencias Psicológicas*, vol. 17, no. 2, 2023.

[23] Z. Sánchez Sánchez, "Relación del estrés y la calidad de sueño en estudiantes universitarios: Revisión bibliográfica," *FACSA - Terapia Física*, vol. 1, no. 1, 2020.

[24] A. R. Regal, M. C. Amigo, and E. Cebrián, "Sueño y mujer," *Revista de Neurología*, vol. 49, no. 7, pp. 376–382, 2009.

[25] Z. Alimoradi, C.-Y. Lin, A. Broström, P. H. Bülow, Z. Bajalan, M. D. Griffiths, M. M. Ohayon, and A. H. Pakpour, "Internet addiction and sleep problems: A systematic review and a meta-analysis," *Sleep Medicine Reviews*, vol. 47, pp. 51–61, 2019.

[26] M. Sejbuk, I. Mironczuk-Chodakowska, and A. M. Witkowska, "Sleep quality: A narrative review on nutrition, stimulants, and physical activity as important factors," *Nutrients*, vol. 14, no. 9, 2022.