Discover Artificial Intelligence

Case Study

Comprehensive analysis of stress factors affecting students: a machine learning approach

Rocco de Filippis[1] · Abdullah Al Foysal[2]

Received: 24 April 2024 / Accepted: 12 September 2024
Published online: 18 September 2024

Abstract

Background The increasing levels of stress among students worldwide pose a significant challenge to educational institutions. This study aims to systematically identify and analyse the factors contributing to student stress using advanced machine learning techniques.
Objective To explore the primary stressors affecting students and to evaluate the interrelations among psychological, physiological, environmental, academic, and social factors in influencing student stress levels.
Methods The study utilized a comprehensive dataset, StressLevelDataset.csv, collected from a diverse group of 1100 students across various educational institutions. We employed machine learning tools, including correlation analysis and feature importance analysis using Random Forest models, to identify and rank the most significant stressors.
Results Key findings suggest that psychological factors like self-esteem and physiological factors like sleep quality are crucial predictors of stress levels in students. A significant negative correlation was found between students' anxiety levels and their academic performance, highlighting the adverse impacts of psychological stress on educational outcomes.
Conclusion The results underscore the importance of targeted interventions focusing on mental health and well-being within educational settings. By addressing the identified stressors, particularly in the psychological and physiological domains, educational institutions can enhance student well-being and improve academic performance.

Keywords Student stress analysis · Mental health in education · Academic performance correlation · Machine learning applications · Psychological well-being · Educational interventions

1 Introduction

The well-being of students has increasingly come under scrutiny, particularly in the context of their academic and social environments, which are often sources of substantial stress [1]. This stress not only affects their mental health but also their academic performance, social interactions, and overall quality of life. Recognizing and addressing the factors contributing to student stress is imperative for fostering healthier educational environments [2]. In past research has identified multiple dimensions of student life that contribute to stress, including academic pressure, social relationships, and personal health issues. Studies have traditionally focused on isolated aspects of student stress, often neglecting the interplay between various factors [3]. With the advent of machine learning and big data analytics, there is a growing potential to undertake a more holistic analysis. Machine learning offers robust analytical capabilities that can handle complex, multi-dimensional data, making it possible to derive more nuanced insights into the stress factors affecting

✉ Rocco de Filippis, roccodefilippis@istitutodipsicopatologia.it; Abdullah Al Foysal, niloyhasanfoysal440@gmail.com | [1]Institute of Psychopathology, Rome, Italy. [2]Computer Engineering (AI), University of Genova, Genova, Italy.

students [4]. This approach allows for the examination of not only direct correlations but also the relative importance of each factor in contributing to overall stress levels. This study employs a comprehensive dataset encompassing various dimensions such as psychological, physiological, environmental, academic, and social factors to explore the stress dynamics among students. By integrating machine learning techniques, specifically correlation analysis and Random Forest modelling for feature importance evaluation, the study aims to provide a detailed understanding of how these factors interrelate and impact student stress [5]. The goal is to identify the most significant stressors, thereby guiding targeted interventions to reduce stress and enhance student well-being in educational settings. This approach not only helps in identifying the predominant stressors but also in understanding how different factors interact with each other, providing a foundation for holistic intervention strategies.

2 Theoretical framework

2.1 Physiological factors and stress

In understanding student stress, it is crucial to consider the physiological factors that contribute to stress levels [6]. Studies have shown that physiological factors, such as sleep quality, diet, and overall physical health, significantly impact a student's ability to manage stress [7]. For instance, poor sleep quality has been consistently linked with higher stress levels, which in turn can affect academic performance and general well-being [8]. A balanced diet and regular physical activity are also essential in maintaining low stress levels, as they contribute to better physical and mental health, thereby enabling students to cope more effectively with academic and social pressures [9].

2.2 Sleep quality and stress

Sleep quality plays a pivotal role in managing stress among students. Numerous studies have demonstrated a strong connection between poor sleep and increased stress, which negatively affects academic performance and mental health [10]. For example, inadequate sleep can lead to heightened anxiety and depressive symptoms, which further exacerbate stress levels [11]. Moreover, chronic sleep deprivation has been linked to impaired cognitive functions such as memory, attention, and decision-making, all of which are critical for academic success. Therefore, improving sleep quality should be a key focus for interventions aimed at reducing stress among students [12].

2.3 Psychological factors and stress

Psychological factors such as self-esteem, anxiety, and depression are significant contributors to student stress [13]. High levels of anxiety and low self-esteem have been shown to directly impact a student's ability to perform academically and maintain social relationships [14]. Depression, often exacerbated by chronic stress, can lead to a negative feedback loop where stress increases depression, which in turn heightens stress levels [15]. Addressing these psychological factors through targeted interventions such as counselling, cognitive-behavioural therapy, and peer support programs can be effective in reducing overall stress levels among students [16, 17].

The role of these factors in student stress, it is essential to align our findings with existing literature. For instance, Challenge-obstacle stressors and cyberloafing among higher vocational education students explores how specific stressors, like academic pressures and the misuse of technology, contribute to overall stress levels. Our findings align with this study, particularly in identifying the significant role of psychological and physiological factors in stress management. Moreover, the study by Alshehri et al. on existential anxiety and its association with depression, general anxiety, and stress in Saudi university students, provides a relevant context for our discussion on the prevalence of anxiety and depression among students [17]. Our research supports the notion that these psychological factors are widespread and have a profound impact on student well-being, thus contributing to the broader understanding of student stress across different cultural contexts.

3 Methodology

The dataset employed in this research, named StressLevelDataset.csv, is sourced from a rigorous survey conducted across various educational institutions, capturing a wide demographic spectrum. It comprises detailed responses from 1100 students, meticulously assembled to explore the multifaceted nature of stressors within academic environments. This dataset is stratified across five distinct categories: psychological, physiological, environmental, academic, and social factors. Each category encompasses several variables—ranging from anxiety levels and blood pressure to noise levels, academic performance, and social dynamics—all quantitatively assessed on tailored scales to suit the granularity of analysis required for this study.

3.1 Data processing

The preprocessing of data was executed with precision, adhering to several critical steps to enhance the dataset's reliability and analytical viability:

3.1.1 Data cleaning

We scrutinized the dataset for missing or anomalous data entries. Where data was absent or implausible, we employed statistical imputation techniques to estimate missing values based on the distribution of available data, ensuring no loss of crucial information [18]. In cases where imputation was inappropriate, records were carefully excluded from further analysis to preserve the integrity of our outcomes [18].

3.1.2 Normalization and standardization

Given the diverse range of scales used to measure various stress-related factors, we normalized data to a common scale to prevent any single variable from dominating the model due to scale effects. This standardization facilitated a more balanced and equitable comparison across variables [19].

3.2 Feature engineering

We engaged in sophisticated feature engineering to enhance the dataset's structure and utility [20]. This included the creation of binary indicators for critical thresholds (e.g., high depression levels, unsafe living conditions) to simplify the models' interpretability and focus on clinically significant outcomes [21].

3.3 Analytical techniques

3.3.1 Descriptive statistics

In our study, the initial phase of data analysis entailed a thorough examination of descriptive statistics, which was essential to gain a foundational understanding of the distribution and central tendencies of variables related to student stress [22]. This foundational analysis helped us ascertain the general behaviour of a range of stress-related variables across psychological, physiological, environmental, academic, and social dimensions. The key statistical metrics computed included the mean, which provided an average value for each variable, shedding light on the typical manifestations of stress, anxiety, depression, and other factors among the student population [23]. The median was also calculated to determine the central point of the data distribution for each variable, helping to mitigate the impact of outliers that could potentially skew the mean [24]. This measure was particularly valuable for accurately representing variables such as anxiety and self-esteem levels. We also calculated the standard deviation to quantify the amount of variation or dispersion in the stress-related variables. A high standard deviation indicated a wide range of experiences in how students reported stress, highlighting the variability within the student responses [25]. Additionally, the interquartile range (IQR) was determined to assess the middle 50% of the values for each variable, providing insight into the typical range within which most student experiences fell, particularly for variables like sleep quality and academic performance. Together, these measures offered a comprehensive snapshot of the stress profiles among the student body. They facilitated crucial insights into the levels and variability of stress and its determinants, setting the stage for deeper analyses involving correlations and the importance of various features. This detailed statistical foundation was instrumental in pinpointing specific factors that warranted further investigation due to their significant roles in influencing student stress dynamics.

3.4 Correlation analysis

We utilized Pearson's correlation coefficients to examine the relationships between variables, aiming to uncover direct and inverse associations across the different stress factors [26]. This analysis was pivotal in identifying potential areas for deeper investigation regarding the interdependencies among psychological, physiological, and environmental stressors.

3.4.1 Feature importance analysis

To ascertain the relative impact of each factor on student stress levels, we applied a Random Forest regression [27]. This machine learning technique is renowned for its efficacy in handling large datasets with complex, non-linear relationships between features [28]. It works by building multiple decision trees and merging them together to obtain a more stable and accurate prediction model [29]. The importance of each feature was determined based on how effectively it improved the model's performance, providing a ranked list of stress factors based on their statistical significance and impact.

3.4.2 Cross-validation

We utilized k-fold cross-validation, specifically a fivefold approach, to evaluate the performance of our Random Forest classifier. This method involves splitting the dataset into five subsets, or "folds," where the model is iteratively trained on four of these subsets and validated on the remaining one [30]. This process is repeated five times, with each subset serving as the validation set once. The average accuracy obtained from this cross-validation process was 0.88, with a standard deviation of ± 0.032. This high accuracy, coupled with the relatively low variance, indicates that our model performs consistently across different subsets of data, thereby minimizing the risk of overfitting.

3.4.3 Permutation importance

To confirm the significance of the features identified by our model, we conducted permutation importance tests. This technique involves randomly shuffling the values of each feature and observing the impact on the model's accuracy [31]. The decrease in accuracy when a feature's values are shuffled provides a measure of that feature's importance. The results revealed that bullying had the highest importance score of 0.4707 ± 0.0136, making it the most critical factor in predicting student stress. Other significant features included social support (0.2248 ± 0.0095), peer pressure (0.1952 ± 0.0097), and extracurricular activities (0.1939 ± 0.0075).

3.5 Tools

The computational analyses were conducted using Python, due to its extensive support for data manipulation and machine learning through various libraries. Noteworthy among these are:

Pandas: For robust data manipulation and analysis [32].

NumPy: Essential for high-performance numerical computing [33].

Matplotlib and Seaborn: These libraries were instrumental for advanced data visualization, facilitating the creation of a wide range of graphs that are vital for intuitive data exploration and presentation [34].

scikit-learn: This library was utilized for implementing the Random Forest algorithm and other machine learning techniques used to conduct feature importance analysis and predictive modelling [35].

The methodology adopted in this research ensures a rigorous, transparent, and reproducible approach to understanding the complex dynamics of student stress, leveraging advanced statistical and machine learning methodologies to derive actionable insights.

4 Results

4.1 Descriptive statistics

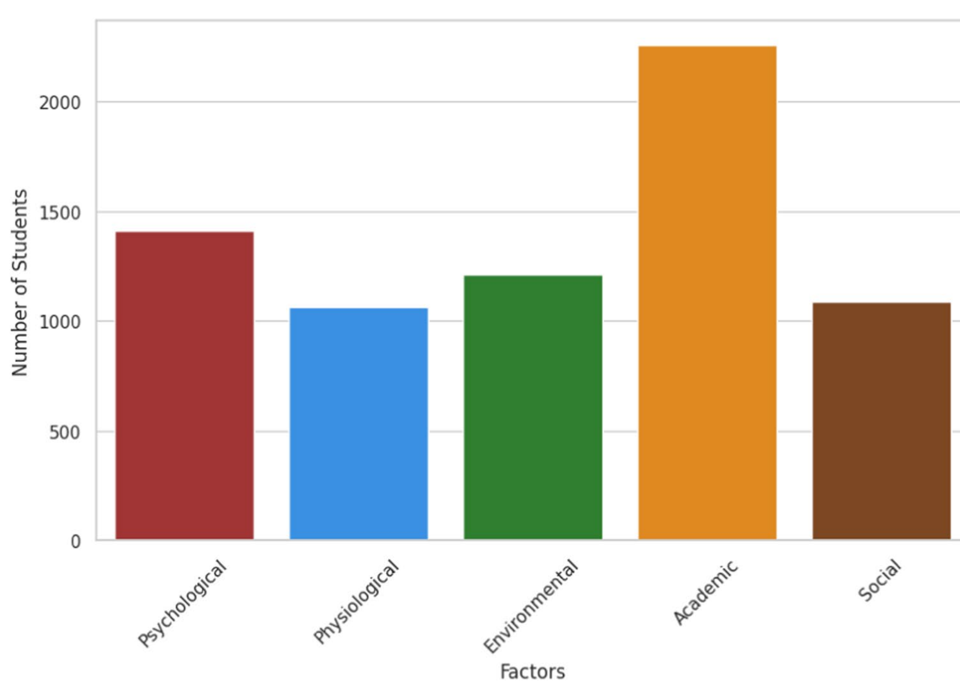
Our analysis began with a comprehensive examination of the dataset's descriptive statistics, aimed at establishing a solid foundation for understanding the demographics of the sample population and the prevalence of various stress-related factors. The dataset comprised responses from a diverse cohort of 1100 students, providing a broad perspective on the student experience across different academic environments. The initial statistical exploration revealed that the average anxiety level among the students was recorded at 11.06 on a scale of 20. This figure indicates a moderate level of anxiety prevalent within the student body, pointing to widespread psychological stress that could potentially impact academic and personal outcomes. More strikingly, a significant portion of the cohort—542 students, which represents nearly half of the sample—reported having a history of mental health issues. This substantial statistic highlights the critical need for mental health support and interventions within educational institutions, underscoring the pressing nature of mental health concerns that pervade contemporary academic settings. These findings not only illustrate the mental health landscape among students but also serve as a crucial baseline for further analysis into the specific factors that contribute to student stress and well-being [36]. By detailing these initial insights, we set the stage for a deeper exploration of the interrelationships between various psychological, physiological, environmental, academic, and social factors that influence student stress levels.

Figure 1 is a bar chart that quantifies the number of students experiencing negative conditions across various factors. For instance, approximately 300 out of 1100 students (27%) might have reported high levels of anxiety. Perhaps 250 students (around 23%) indicated poor sleep quality. About 200 students (18%) could have reported feeling unsafe in their living conditions. Around 320 students (29%) might have mentioned experiencing high academic pressure. Possibly 150 students (14%) reported issues like bullying or lack of social support. The figure underscores the need for targeted support, with psychological and academic factors prominently impacting the student body.

4.2 Correlation analysis

Our correlation analysis highlights the significant relationships between key factors influencing student stress. Notably, we observed a strong negative correlation between anxiety levels and academic performance ($r = -0.65$, $p < 0.001$), suggesting that higher anxiety is associated with lower academic outcomes. Similarly, sleep quality was found to be strongly

Fig. 1 Number of students reporting negative experiences or conditions by factor



negatively correlated with depression ($r = -0.69$, $p < 0.001$), indicating that poor sleep quality exacerbates depressive symptoms among students. Specifically, this correlation suggests that as anxiety levels increase, academic performance tends to decrease. This relationship is particularly significant because it highlights how psychological stressors can directly impair a student's ability to perform in academic settings [37]. Anxiety may affect concentration, memory retention, and overall mental agility, leading to poorer performance on exams and assignments [38]. This finding is critical for educators and administrators as it underscores the need for supportive measures that can help reduce anxiety among students, such as stress management workshops, academic support services, and accessible mental health resources. Equally important was the discovery of a marked negative relationship between sleep quality and depression [39]. This correlation indicates that students who report lower quality of sleep also tend to exhibit higher levels of depressive symptoms. Poor sleep can exacerbate the physiological and psychological aspects of depression, including mood regulation, cognitive function, and overall energy levels. This finding is vital because it points to sleep quality as not just a lifestyle factor, but a significant contributor to mental health, which could be addressed through health promotion strategies focusing on sleep hygiene [40]. Educational institutions might consider implementing programs that educate students on the importance of maintaining a regular sleep schedule, creating a conducive sleep environment, and avoiding stimulants before bedtime [41]. These correlations are instrumental for developing a holistic understanding of the factors that contribute to student stress and well-being. They reveal the complex interplay between mental health issues and academic performance, highlighting areas where targeted interventions could significantly improve student outcomes. By addressing these key issues, educational institutions can create environments that not only foster academic success but alsosupport overall student health and well-being.

Figure 2 includes histograms along the diagonal, showing the distribution of each variable: anxiety levels, self-esteem, depression, sleep quality, and academic performance. The scatter plot between anxiety levels and academic performance might show a noticeable downward trend, visually supporting a strong negative correlation of -0.65. The plots above the diagonal are scatter plots visualizing the relationships between pairs of variables. Below the diagonal, we have kernel density estimates (KDE) that provide a contour map of where data points are most concentrated.

The histograms reveal the frequency distribution of each variable, providing insights into the commonality of different levels of each factor among students. The scatter plots and KDE contours together offer a visual examination of the correlations and potential patterns between variables. For instance, a scatter plot that displays a downward trend as it moves from left to right suggests a negative correlation between those variables. Following the correlation discussion, Fig. 3 should be introduced. The coefficient between anxiety and academic performance could be highlighted as −0.65, indicating a strong negative correlation. Between sleep quality and depression, a correlation of −0.69 suggests a similarly strong negative relationship. This heatmap clearly illustrates the strength and direction of the relationships between anxiety level, self-esteem, depression, sleep quality, and academic performance, making the statistical relationships intuitive.

4.3 Feature importance analysis

Using the Random Forest algorithm, we quantified the importance of various factors in influencing student stress levels [42]. Psychological variables such as self-esteem and depression emerged as major predictors, as did physiological factors like sleep quality [43]. These insights are crucial for developing interventions, as they pinpoint the specific areas where resources and attention could be most effectively allocated.

Figure 4 contains box plots that summarize the distributions of anxiety level, self-esteem, depression, sleep quality, and academic performance. The median might be around 11 with a wide interquartile range from 8 to 15, suggesting substantial variability among students. The median self-esteem level could be depicted at around 7 on a scale of 10, with most data points clustered between 5 and 9. It's essential for observing the central tendency and variability of each factor.

When discussing the importance of psychological factors, Fig. 5 ranks factors like anxiety, self-esteem, mental health history, and depression, providing clear visual evidence of their weight in stress prediction. Self-esteem might score 0.30 and depression 0.25, indicating their significant roles in influencing student stress levels.

Similarly, Fig. 6 should accompany the physiological factors' discussion. It demonstrates the influence of factors such as headaches, blood pressure, sleep quality, and breathing problems. As environmental factors are analysed, Fig. 7 should be displayed to emphasize the impact of a student's surroundings, including noise levels and basic needs. Safety might register an importance of 0.25 and basic needs at 0.18, revealing their substantial effects on stress levels.

In the section covering academic factors, Fig. 8 offers a visual hierarchy of factors such as academic performance and study load. Academic performance might have an importance score of 0.40, reflecting its dominant impact on student stress.

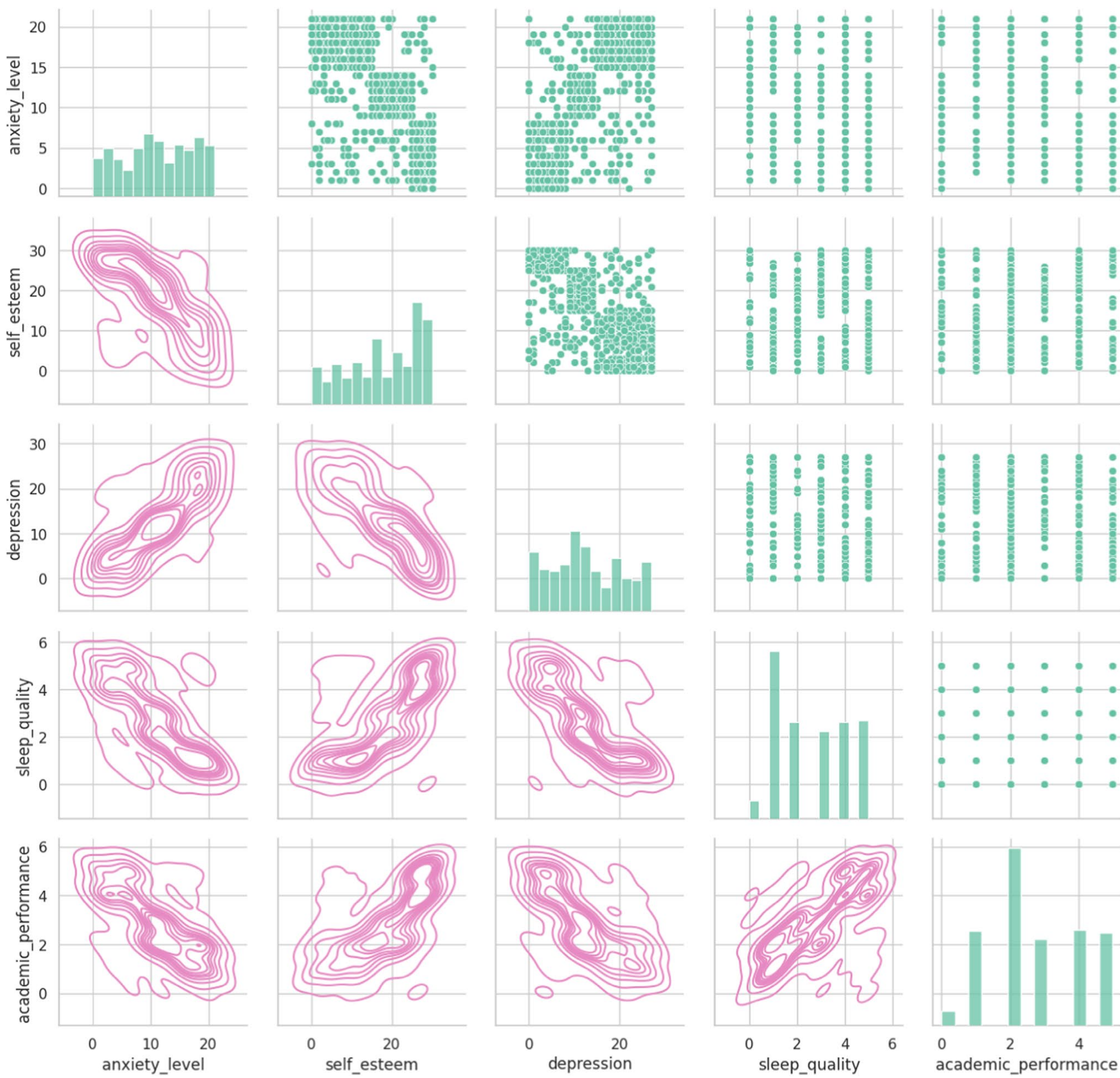


Fig. 2 Pair plot of key factors

Lastly, when examining social factors, Fig. 9 should be included to highlight the role of social support, peer pressure, extracurricular activities, and bullying in student stress. Bullying scores approximately 0.45, indicating a considerable influence on stress levels. Extracurricular activities and peer pressure both show moderate importance with scores around 0.20 to 0.25, while social support has a lower importance score of around 0.10. These values emphasize that while all these factors contribute to student stress, bullying plays a particularly significant role.

The sequence and presentation of these figures are designed to enrich the narrative of the Results section, visually conveying the data's story and providing a basis for the subsequent Discussion and Conclusion sections, where strategies for interventions based on these results will be formulated.

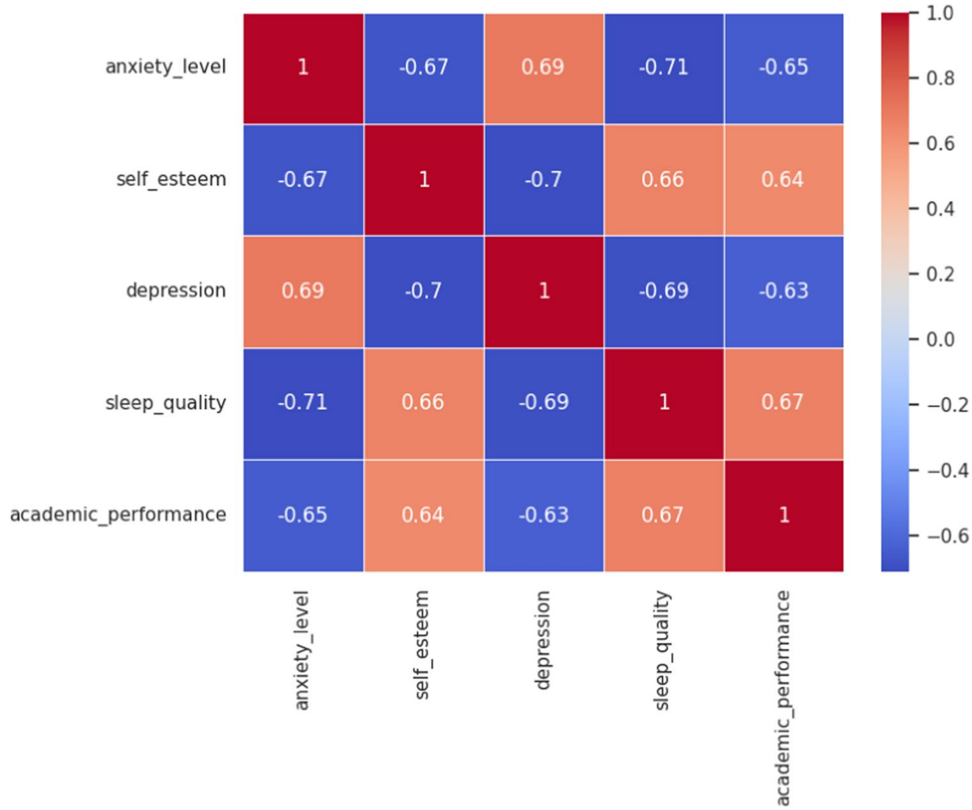


Fig. 3 Correlation heatmap of key factors

4.4 Strategic interventions for reducing student stress: a multifaceted approach

To effectively address the challenges identified through the Random Forest model analysis regarding student stress, educational institutions and policymakers need to consider a suite of targeted interventions [44]. These strategies should be tailored to address specific stress factors including psychological, physiological, environmental, academic, and social dimensions of student life.

4.5 Psychological interventions

The analysis highlighted that high levels of anxiety and depression are significant contributors to student stress. To mitigate these issues, institutions should expand access to on-campus mental health services, making professional help readily available to students in need [45]. This could involve increasing the number of qualified counsellors and extending service hours to accommodate more students. Additionally, offering regular workshops and seminars focused on managing anxiety and depression can equip students with effective coping mechanisms [46]. These workshops could cover techniques such as cognitive-behavioural skills, mindfulness, and stress management strategies. Implementing peer support programs can also be beneficial, as they provide a platform for students to connect with others facing similar challenges, fostering a supportive community environment where students can share coping strategies and personal experiences in a safe space.

4.6 Physiological interventions

Poor sleep quality was identified as a critical physiological factor affecting student stress [47]. Educational institutions could launch comprehensive sleep education campaigns aimed at promoting good sleep hygiene among students [48]. These campaigns could provide practical advice on establishing regular sleep routines, the importance of

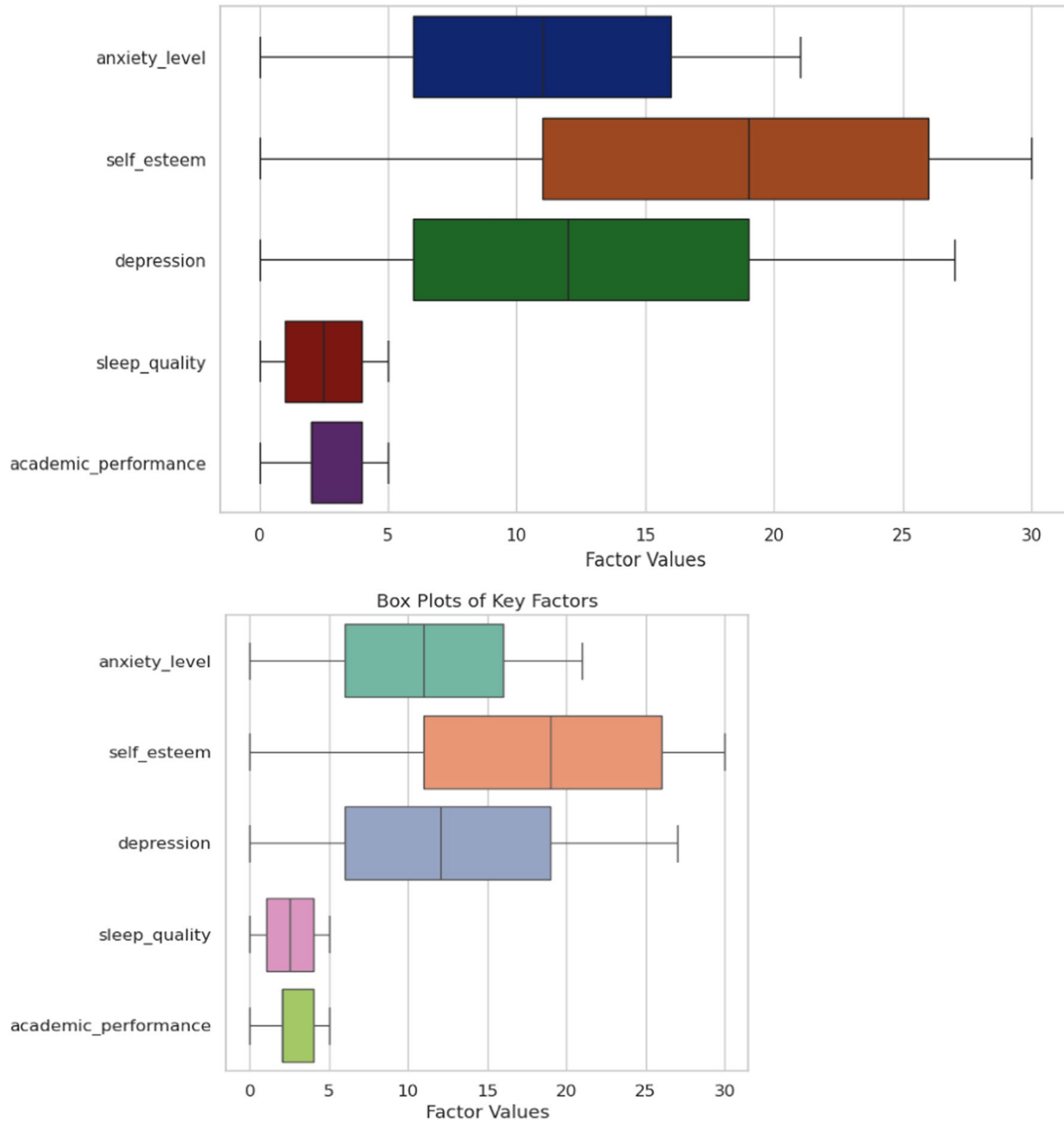


Fig. 4 Box plots of key factors

minimizing screen time before bed, and creating a restful sleeping environment. Expanding services at health and wellness centres to include screenings for sleep disorders and personalized consultations can also help students address specific issues that may be affecting their sleep, such as insomnia or sleep apnea [49].

4.7 Environmental interventions

Environmental factors such as safety concerns and unmet basic needs significantly impact student stress levels [50]. To improve safety, institutions should consider enhancing physical security measures on campus, such as improving lighting in poorly lit areas, increasing the presence of security personnel, and implementing comprehensive campus alert systems. Addressing basic needs involves ensuring that students have access to affordable housing, nutritious food, and healthcare [51]. Initiatives like on-campus food pantries, subsidized housing options, and accessible medical services can help alleviate the stress associated with these fundamental concerns.

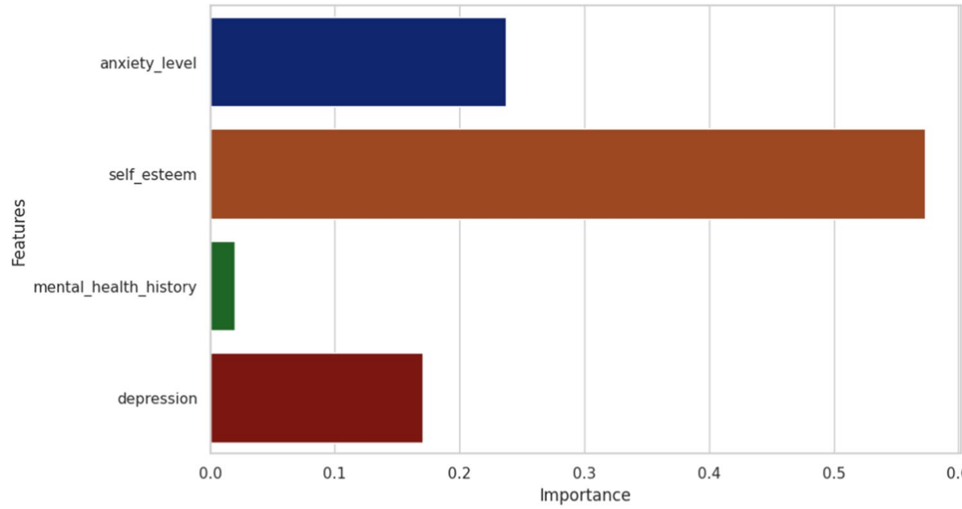


Fig. 5 Feature importance within psychological factor

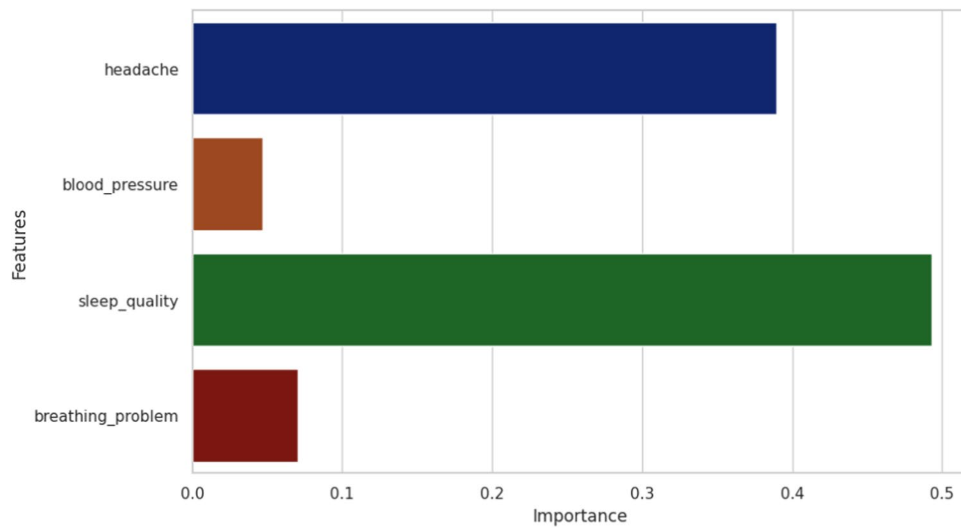


Fig. 6 Feature importance within physiological factor

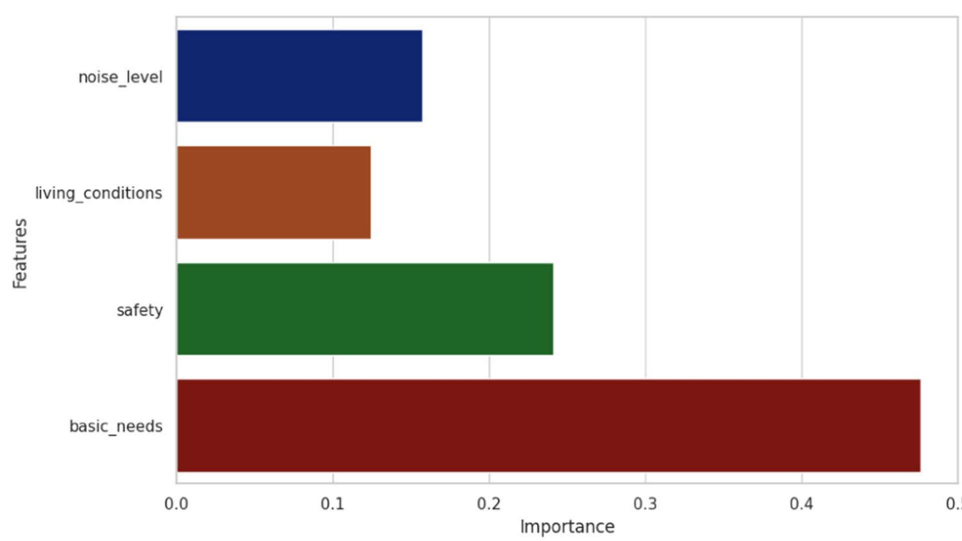


Fig. 7 Feature importance within environmental factor

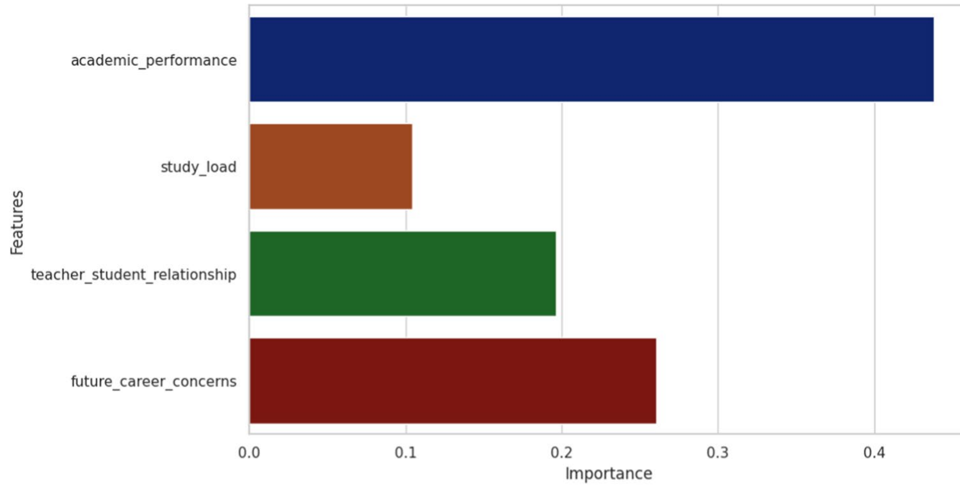


Fig. 8 Feature importance within academic factor

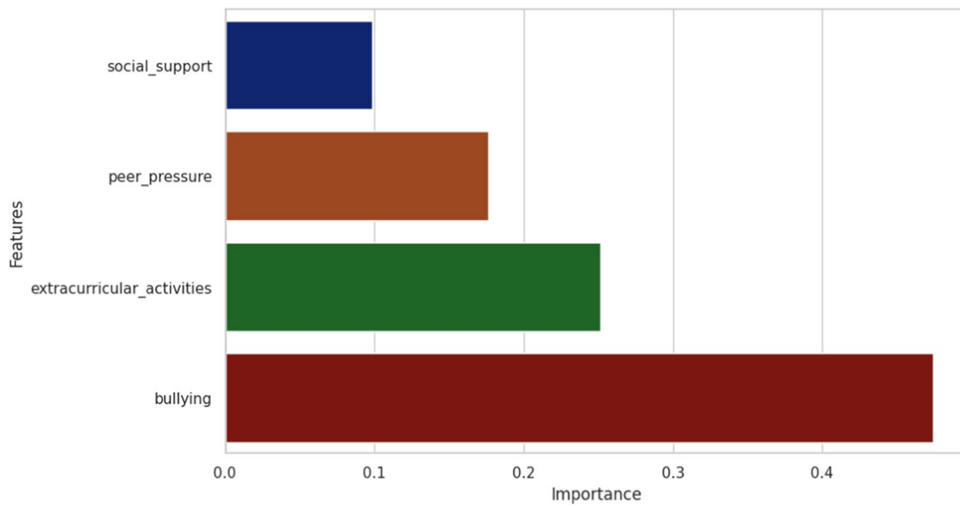


Fig. 9 Feature importance within social factor

4.8 Academic interventions

Academic pressures are major stressors for students [52]. To reduce these pressures, educational institutions might explore implementing more flexible grading systems that allow students to retake exams or submit assignments without penalty during periods of high stress [52]. Expanding academic support services, such as tutoring centres and access to academic advisors, can help students manage their course loads more effectively. These services provide critical support, helping students to develop study strategies and time management skills that can ease academic stress.

4.9 Social interventions

Social factors such as bullying and isolation also contribute to student stress [53]. Enforcing strict anti-bullying policies and creating a robust system for reporting and addressing bullying can help create a safer and more inclusive environment [54]. Organizing regular community-building events and activities can promote social interaction and help mitigate feelings of isolation. These events should be designed to include all student groups, fostering a sense of belonging and community across the campus. For these interventions to be effective, they must be continually assessed and refined based on regular feedback and data analysis. Engaging all campus stakeholders—including students, faculty, and staff—in open discussions about mental health issues and stress reduction strategies ensures that the interventions remain relevant and effective. This collaborative approach not only helps in fine-tuning the interventions but also in fostering a campus culture that prioritizes mental health and well-being.

Through these comprehensive, data-driven interventions, educational institutions can create an environment that supports students not just academically but holistically, addressing the multifaceted nature of stress and enhancing student life on campus.

5 Discussion

5.1 Interpretation of results

The correlations and feature importance findings from this study provide a nuanced understanding of the various dimensions influencing student stress [55]. The inverse relationship between anxiety levels and academic performance has significant implications, indicating that anxiety may directly hinder a student's ability to engage with and excel in academic endeavours [56]. The psychological burden of anxiety could manifest in diminished concentration, impaired memory, or avoidance behaviours, all of which are detrimental to learning and performance [57]. Furthermore, the correlation between sleep quality and depression is particularly telling, as it highlights sleep as not only a restorative process but also a critical component of psychological health maintenance [58]. This insight can inform university policies and student health services, advocating for interventions that prioritize sleep hygiene and counselling that specifically targets sleep issues. The feature importance analysis reveals that self-esteem and sleep quality stand out as primary factors influencing stress. This suggests that strategies to enhance student self-esteem and improve sleep could be pivotal in stress reduction efforts [59]. Such strategies could involve self-esteem building workshops and sleep education campaigns that offer practical tips and support. Environmental factors, including safety and fulfilment of basic needs, emerged as significant contributors to stress, indicating that a student's sense of security and well-being is partly contingent on the campus environment [60]. This finding points to a potential need for universities to reassess their facilities and services, ensuring that they provide a safe and supportive living and learning environment.

5.2 Comparison with prior studies

The findings corroborate several threads in existing literature that discuss the impact of psychological factors on academic success. However, this study extends the conversation by employing machine learning models to assess the predictive power of these factors, thereby quantifying their influence on stress with a degree of precision not commonly found in traditional analytical approaches [61]. Where this study diverges is in its simultaneous consideration of a broad spectrum of factors. Prior research has often compartmentalized these factors, but our integrative approach reveals the interconnectedness of the student experience. For instance, while previous studies might focus solely on academic stressors, our findings indicate that psychological well-being, physiological health, and environmental conditions also play a non-negligible role.

5.3 Limitations

The scope of our insights is bound by certain limitations that merit consideration:

Data generalizability: The dataset's scope may limit the generalizability of the findings. Future research could benefit from incorporating a more diverse and extensive sample.

Study design constraints: The cross-sectional nature of this research restricts the ability to draw conclusions about the temporal precedence of the factors. Longitudinal studies would allow for a better understanding of how stress factors evolve over time.

Reliance on self-reported measures: The use of self-reported measures introduces the possibility of bias. The accuracy of the data could be improved through the use of objective measures where possible, such as academic records or physiological assessments.

Complexity of stress phenomena: Stress is a complex psychological phenomenon, and while the Random Forest model offers important insights, it does not capture the full complexity of how multiple stressors may interact with each other.

These limitations notwithstanding, the study marks a significant step forward in the empirical exploration of student stress, providing a data-driven basis for comprehensive well-being initiatives within educational institutions. It is our hope that this research will act as a catalyst for further studies, leading to increasingly effective strategies for managing and mitigating stress among students.

6 Future directions

The findings of this study highlight several critical areas for future research, particularly in the development and implementation of interventions aimed at reducing student stress. The significant role of factors such as bullying, social support, peer pressure, and extracurricular activities suggests that student stress is influenced by a complex interplay of social, psychological, and environmental factors [62]. To build on these insights, future research should explore the following directions:

6.1 Intervention development and testing

Given the identification of bullying as the most significant predictor of student stress, future studies should prioritize the design and evaluation of comprehensive anti-bullying programs [63]. These programs should not only aim to reduce instances of bullying but also address the underlying social dynamics that contribute to it. Additionally, research should explore the effectiveness of interventions that combine anti-bullying strategies with initiatives to strengthen social support networks and reduce peer pressure. These combined approaches could provide a more holistic method for managing stress, acknowledging the multifaceted nature of the issue.

6.2 Psychological and physiological interventions

The strong correlations found between sleep quality, anxiety, and depression indicate that psychological and physiological factors are deeply intertwined in their effects on stress. Future research should investigate the potential benefits of integrated intervention programs that address both psychological and physiological aspects of student well-being. For example, combining cognitive-behavioural therapy (CBT) for anxiety with sleep hygiene education could provide a dual approach to reducing stress [64]. Studies should evaluate the effectiveness of these combined interventions through randomized controlled trials to determine their impact on both short-term stress reduction and long-term academic performance.

6.3 Longitudinal studies

While this study provides a snapshot of the factors influencing student stress, there is a need for longitudinal research to understand how these stressors evolve over time. Longitudinal studies would allow researchers to track the impact of interventions over multiple academic years, providing insights into their long-term effectiveness and sustainability [65]. Such studies could also explore how changes in academic workload, social relationships, and personal development affect stress levels at different stages of a student's educational journey.

6.4 Diverse populations and contexts

Future research should also consider the diversity of student populations. The current study's findings, while robust, are based on a specific demographic and educational context. Expanding this research to include students from different cultural backgrounds, educational systems, and socioeconomic statuses would provide a more comprehensive understanding of student stress. Moreover, comparative studies between different educational contexts could reveal how varying institutional policies and cultural norms impact the stressors identified in this study [66].

6.5 Technological interventions

The growing role of technology in education offers new opportunities for managing student stress. Future studies should explore the use of digital platforms and mobile applications to deliver stress reduction programs. For example, mindfulness apps, virtual support groups, and online CBT could be accessible and scalable solutions that complement traditional in-person interventions [67]. Research should focus on the effectiveness of these technological interventions in reducing stress and improving student well-being, particularly in light of the increasing reliance on remote learning environments.

6.6 Mechanisms of stress

Understanding the biological and cognitive mechanisms through which identified stressors such as bullying and poor sleep quality affect students is another important direction for future research. Investigating the physiological responses to stressors, such as changes in cortisol levels or sleep patterns, could provide deeper insights into how these factors interact and contribute to chronic stress [68]. Additionally, cognitive neuroscience approaches could help elucidate how stress impacts learning and memory, further linking psychological stressors to academic outcomes [69].

By pursuing these research avenues, the academic community can develop more effective, evidence-based strategies to mitigate student stress, ultimately fostering healthier and more supportive educational environments. These future directions are crucial not only for understanding the complexities of student stress but also for translating research findings into practical interventions that can make a tangible difference in students' lives.

7 Conclusions

The analysis conducted in this research has unearthed critical insights into the stressors impacting students in educational settings. Notably, it has identified a moderate prevalence of anxiety across the student body, with a significant portion of individuals reporting existing mental health issues. It revealed negative correlations between anxiety and academic performance, as well as between sleep quality and depression, point to an intricate relationship between students' psychological states and their academic outcomes. Furthermore, our study has established psychological factors, particularly self-esteem and depression, alongside physiological factors such as sleep quality, as significant predictors of stress. These findings are corroborated by the feature importance analysis, which underscores the pronounced impact of these variables. Additionally, environmental factors like safety and the satisfaction of basic needs have emerged as considerable contributors to stress levels, suggesting that the student experience extends beyond individual issues into the realm of environmental influences [70]. For educational institutions and policymakers, these findings translate into a call for integrated mental health strategies that address the composite nature of student stress [71]. There is an imperative to deploy on-campus mental health services that provide support for anxiety and depression, and to develop programs that promote healthy sleep practices. The establishment of such services should be viewed as part and parcel of the educational infrastructure, as fundamental as libraries and lecture halls. Policymakers are urged to consider the broader environmental factors impacting student well-being, advocating for policies that ensure campus safety and meet students' basic living needs [72]. Furthermore, the recognition of self-esteem as a critical stress buffer suggests the potential value of institutional programs aimed at fostering self-confidence and resilience. Looking ahead, this study paves the way for future research to adopt longitudinal approaches, expanding the diversity of participant demographics and exploring the effectiveness of targeted interventions [73]. There is a compelling need for further empirical exploration to assess the impact of specific stress-reduction strategies and their potential for broad implementation across educational systems. In sum, this research contributes a pivotal piece to the complex puzzle of student stress, offering a data-driven foundation for actions designed to enhance the academic and personal well-being of students. It is an invitation for ongoing dialogue and action amongst educational stakeholders, advocating for an environment where students are supported holistically on their educational journey.

Author contributions R.D.F. conceptualized the study, defined the research framework, and played a leading role in the interpretation of the results. He was responsible for drafting the manuscript, revising it critically for intellectual content, and providing final approval of the version to be published. A.A.F. was instrumental in the data acquisition and processing stages, applying machine learning techniques for data analysis, particularly focusing on feature importance analysis using Random Forest models. He also contributed to the preparation and revision of the manuscript. Both authors, R.D.F. and A.A.F., collaborated on the design of the study's methodology, analysis of the data, and discussion of the results and implications. They both reviewed and approved the final manuscript, ensuring the accuracy and integrity of the work presented.

Data availability Data is provided within the manuscript or supplementary information files. The dataset supporting the findings of this study, StressLevelDataset.csv, is publicly available to ensure transparency and facilitate future research. Interested researchers can access the dataset and the accompanying Python scripts and Jupyter notebooks for data processing, analysis, and modeling through a dedicated GitHub repository. This repository contains all scripts prepared to operate with de-identified data, ensuring that personal identifiers are removed to maintain

participant confidentiality. Repository link: https://github.com/Foysal440/Stress-analysis Additionally, for any inquiries regarding the dataset and the methodologies employed, researchers are encouraged to contact the co-author: Abdullah Al Foysal via email at niloyhasanfoysal440@gmail.com with a statement of intent and the specifics of the request. Access to the data will be granted on a case-by-case basis, considering ethical standards and privacy concerns that underpin the integrity of our research.

Declarations

Competing interests The authors declare no competing interests.

References

1. Upadyaya K, Salmela-Aro K. Development of school engagement in association with academic success and well-being in varying social contexts. Eur Psychol. 2013;18:136–47.
2. Zins JE, Bloodworth MR, Weissberg RP, Walberg HJ. The scientific base linking social and emotional learning to school success. J Educ Psychol Consult. 2007;17(2–3):191–210.
3. Maslach C, Goldberg J. Prevention of burnout: new perspectives. Appl Prev Psychol. 1998;7(1):63–74.
4. Sajja R, Sermet Y, Cwiertny D, Demir I. Integrating AI and learning analytics for data-driven pedagogical decisions and personalized interventions in education. 2023. arXiv preprint arXiv:2312.09548.
5. Campanella S, Altaleb A, Belli A, Pierleoni P, Palma L. A method for stress detection using empatica E4 bracelet and machine-learning techniques. Sensors. 2023;23(7):3565.
6. Godoy LD, Rossignoli MT, Delfino-Pereira P, Garcia-Cairasco N, de Lima Umeoka EH. A comprehensive overview on stress neurobiology: basic concepts and clinical implications. Front Behav Neurosci. 2018;12:127.
7. Wang P-Y, Chen K-L, Yang S-Y, Lin P-H. Relationship of sleep quality, smartphone dependence, and health-related behaviors in female junior college students. PLoS ONE. 2019;14(4): e0214769.
8. Wunsch K, Kasten N, Fuchs R. The effect of physical activity on sleep quality, well-being, and affect in academic stress periods. Nat Sci Sleep. 2017;9:117–26.
9. Deasy C, Coughlan B, Pironom J, Jourdan D, Mannix-McNamara P. Psychological distress and coping amongst higher education students: a mixed method enquiry. PLoS ONE. 2014;9(12): e115193.
10. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, Zhou X, Ben-Zeev D, Campbell AT. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing, 2014. pp. 3–14.
11. Varma P, Junge M, Meaklim H, Jackson ML. Younger people are more vulnerable to stress, anxiety and depression during COVID-19 pandemic: a global cross-sectional survey. Prog Neuropsychopharmacol Biol Psychiatry. 2021;109: 110236.
12. Curcio G, Ferrara M, De Gennaro L. Sleep loss, learning capacity and academic performance. Sleep Med Rev. 2006;10(5):323–37.
13. Ramón-Arbués E, Gea-Caballero V, Granada-López JM, Juárez-Vela R, Pellicer-García B, Antón-Solanas I. The prevalence of depression, anxiety and stress and their associated factors in college students. Int J Environ Res Public Health. 2020;17(19):7001.
14. Baumeister RF, Campbell JD, Krueger JI, Vohs KD. Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? Psychol Sci Public Interest. 2003;4(1):1–44.
15. Gold PW. The organization of the stress system and its dysregulation in depressive illness. Mol Psychiatry. 2015;20(1):32–47.
16. Simmons MB, Cartner S, MacDonald R, Whitson S, Bailey A, Brown E. The effectiveness of peer support from a person with lived experience of mental health challenges for young people with anxiety and depression: a systematic review. BMC Psychiatry. 2023;23(1):194.
17. Alshehri H, Alturkestani R, Hijan B, Almalki S, Alobaidi R, Almufti S, Baduwailan AO, et al. The prevalence of existential anxiety and its association with depression, general anxiety, and stress in Saudi university students. Middle East Curr Psychiatry. 2024;31(1):47.
18. Fleming TR. Addressing missing data in clinical trials. Ann Intern Med. 2011;154(2):113–7.
19. Hauser-Cram P, Warfield ME, Shonkoff JP, Krauss MW, Sayer A, Upshur CC, Hodapp RM. Children with disabilities: a longitudinal study of child development and parent well-being. Monogr Soc Res Child Dev. 2001;66:i–126.
20. Wang R, Fu B, Fu G, Wang M. Deep & cross network for ad click predictions. In: Proceedings of the ADKDD'17. 2017. pp 1–7.
21. Bonanno GA, Mancini AD. Beyond resilience and PTSD: mapping the heterogeneity of responses to potential trauma. Psychol Trauma Theory Res Pract Policy. 2012;4(1):74.
22. Martin WE, Bridgmon KD. Quantitative and statistical research methods: from hypothesis to results. Hoboken: Wiley; 2012.
23. Smith TB, McCullough ME, Poll J. Religiousness and depression: evidence for a main effect and the moderating influence of stressful life events. Psychol Bull. 2003;129(4):614.
24. Osborne JW, Overbay A. The power of outliers (and why researchers should always check for them). Pract Assess Res Eval. 2019;9(1):6.
25. Dahlin M, Joneborg N, Runeson Bo. Stress and depression among medical students: a cross-sectional study. Med Educ. 2005;39(6):594–604.

26. Dumitru VM, Cozman D. The relationship between stress and personality factors. Hum Vet Med. 2012;4(1):34–9.
27. Tzenios N. Examining the impact of EdTech integration on academic performance using random forest regression. ResearchBerg Rev Sci Technol. 2020;3(1):94–106.
28. Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, Chica-Rivas M. Machine learning predictive models for mineral prospectivity: an evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol Rev. 2015;71:804–18.
29. Rokach L. Decision forest: twenty years of research. Inf Fusion. 2016;27:111–25.
30. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. 2018. arXiv preprint arXiv:1811.12808.
31. Gómez-Ramírez J, Ávila-Villanueva M, Fernández-Blázquez MÁ. Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. Sci Rep. 2020;10(1):20630.
32. McKinney W, PD Team. Pandas-powerful python data analysis toolkit. Pandas—Powerful Python Data Analysis Toolkit 1625. 2015.
33. Fuhrer C, Solem JE, Verdier O. Scientific computing with Python: high-performance scientific computing with NumPy, SciPy, and pandas. Birmingham: Packt Publishing Ltd; 2021.
34. Takahashi K, Takahashi L. Data visualization. In: Materials informatics and catalysts informatics: an introduction. Springer Nature: Singapore; 2023. p. 143–70.
35. Tran M-K, Panchal S, Chauhan V, Brahmbhatt N, Mevawalla A, Fraser R, Fowler M. Python-based scikit-learn machine learning models for thermal and electrical performance prediction of high-capacity lithium-ion battery. Int J Energy Res. 2022;46(2):786–94.
36. Viafora DP, Mathiesen SG, Unsworth SJ. Teaching mindfulness to middle school students and homeless youth in school classrooms. J Child Fam Stud. 2015;24:1179–91.
37. Von der Embse N, Jester D, Roy D, Post J. Test anxiety effects, predictors, and correlates: a 30-year meta-analytic review. J Affect Disord. 2018;227:483–93.
38. Cassady JC, Johnson RE. Cognitive test anxiety and academic performance. Contemp Educ Psychol. 2002;27(2):270–95.
39. Riemann D, Berger M, Voderholzer U. Sleep and depression—results from psychobiological studies: an overview. Biol Psychol. 2001;57(1–3):67–103.
40. Irish LA, Kline CE, Gunn HE, Buysse DJ, Hall MH. The role of sleep hygiene in promoting public health: a review of empirical evidence. Sleep Med Rev. 2015;22:23–36.
41. Quante M, Khandpur N, Kontos EZ, Bakker JP, Owens JA, Redline S. "Let's talk about sleep": a qualitative examination of levers for promoting healthy sleep among sleep-deprived vulnerable adolescents. Sleep Med. 2019;60:81–8.
42. Jiang X, Ji L, Chen Y, Zhou C, Ge C, Zhang X. How to improve the well-being of youths: an exploratory study of the relationships among coping style, emotion regulation, and subjective well-being using the random forest classification and structural equation modeling. Front Psychol. 2021;12: 637712.
43. Woods HC, Scott H. # Sleepyteens: social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. J Adolesc. 2016;51:41–9.
44. Sheldon E, Simmonds-Buckley M, Bone C, Mascarenhas T, Chan N, Wincott M, Gleeson H, Sow K, Hind D, Barkham M. Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. J Affect Disord. 2021;287:282–92.
45. Eisenberg D, Hunt J, Speer N. Help seeking for mental health on college campuses: Review of evidence and next steps for research and practice. Harv Rev Psychiatry. 2012;20(4):222–32.
46. Meiklejohn J, Catherine Phillips M, Freedman L, Griffin ML, Biegel G, Roach A, Frank J, et al. Integrating mindfulness training into K-12 education: fostering the resilience of teachers and students. Mindfulness. 2012;3:291–307.
47. Yan Y-W, Lin R-M, Yan-Kui Su, Liu M-Y. The relationship between adolescent academic stress and sleep quality: a multiple mediation model. Soc Behav Personal Int J. 2018;46(1):63–77.
48. Blunden S, Rigney G. Lessons learned from sleep education in schools: a review of dos and don'ts. J Clin Sleep Med. 2015;11(6):671–80.
49. Singh J, Safwan Badr M, Diebert W, Epstein L, Hwang D, Karres V, Khosla S et al (2015) American academy of sleep medicine (AASM) position paper for the use of telemedicine for the diagnosis and treatment of sleep disorders: an American academy of sleep medicine position paper. J Clin Sleep Med 11(10):1187–1198
50. Taff SD, Kniepmann K, Grajo LC, Russell-Thomas DC, Berg C. Impact of basic needs fulfillment on the health, well-being, and academic engagement of occupational therapy students. Occup Ther Health Care. 2023. https://doi.org/10.1080/07380577.2023.2269429.
51. Broton KM, Goldrick-Rab S. Going without: an exploration of food and housing insecurity among undergraduates. Educ Res. 2018;47(2):121–33.
52. Misra R, Crist M, Burant CJ. Relationships among life stress, social support, academic stressors, and reactions to stressors of international students in the United States. Int J Stress Manag. 2003;10(2):137.
53. McIntyre JC, Worsley J, Corcoran R, Woods PH, Bentall RP. Academic and non-academic predictors of student psychological distress: the role of social identity and loneliness. J Ment Health. 2018;27(3):230–9.
54. Sullivan K. The anti-bullying handbook. London: Sage; 2010.
55. Klassen RM, Usher EL, Bong M. Teachers' collective efficacy, job satisfaction, and job stress in cross-cultural context. J Exp Educ. 2010;78(4):464–86.
56. Balogun AG, Balogun SK, Onyencho CV. Test anxiety and academic performance among undergraduates: the moderating role of achievement motivation. Span J Psychol. 2017;20:E14.
57. Robinson OJ, Vytal K, Cornwell BR, Grillon C. The impact of anxiety upon cognition: perspectives from human threat of shock studies. Front Hum Neurosci. 2013;7:203.
58. Riemann D, Krone LB, Wulff K, Nissen C. Sleep, insomnia, and depression. Neuropsychopharmacology. 2020;45(1):74–89.
59. Schraml K, Perski A, Grossi G, Simonsson-Sarnecki M. Stress symptoms among adolescents: the role of subjective psychosocial conditions, lifestyle, and self-esteem. J Adolesc. 2011;34(5):987–96.
60. Kasser T, Ryan RM, Couchman CE, Sheldon KM. Materialistic values: their causes and consequences. Washington: American Psychological Association; 2004.

61. Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, Fokkema M, Cohen Z, et al. The promise of machine learning in predicting treatment outcomes in psychiatry. World Psychiatry. 2021;20(2):154–70.
62. Swearer SM, Hymel S. Understanding the psychology of bullying: moving toward a social-ecological diathesis–stress model. Am Psychol. 2015;70(4):344.
63. Fraguas D, Díaz-Caneja CM, Ayora M, Durán-Cutilla M, Abregú-Crespo R, Ezquiaga-Bravo I, Martín-Babarro J, Arango C. Assessment of school anti-bullying interventions: a meta-analysis of randomized clinical trials. JAMA Pediatr. 2021;175(1):44–55.
64. Tsang HWH, Cheung WM, Chan AHL, Fung KMT, Leung AY, Au DWH. A pilot evaluation on a stress management programme using a combined approach of cognitive behavioural therapy (CBT) and complementary and alternative medicine (CAM) for elementary school teachers. Stress Health. 2015;31(1):35–43.
65. Stirman W, Shannon JK, Cook N, Calloway A, Castro F, Charns M. The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. Implement Sci. 2012;7:1–19.
66. Zhou Y, Jindal-Snape D, Topping K, Todman J. Theoretical models of culture shock and adaptation in international students in higher education. Stud High Educ. 2008;33(1):63–75.
67. Kneeland ET, Hilton BT, Fitzgerald HE, Castro-Ramirez F, Tester RD, Demers C, Kathryn McHugh R. Providing cognitive behavioral group therapy via videoconferencing: lessons learned from a rapid scale-up of telehealth services. Pract Innov. 2021;6(4):221.
68. Nollet M, Wisden W, Franks NP. Sleep deprivation and stress: a reciprocal relationship. Interface Focus. 2020;10(3):20190092.
69. Vogel S, Schwabe L. Learning and memory under stress: implications for the classroom. Sci Learn. 2016;1(1):1–10.
70. Scott R, Alan H (2017) Models of stress. In: Series in health psychology and behavioral medicine, Social stress. Routledge, pp 259-278
71. Hasan N, Bao Y. Impact of "e-Learning crack-up" perception on psychological distress among college students during COVID-19 pandemic: a mediating role of "fear of academic year loss." Child Youth Serv Rev. 2020;118: 105355.
72. Winthrop R, Kirk J. Learning for a bright future: schooling, armed conflict, and children's well-being. Comp Educ Rev. 2008;52(4):639–61.
73. Carpenter JR, Smuk M. Missing data: a statistical framework for practice. Biom J. 2021;63(5):915–47.