medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

Machine learning-based prediction of depression severity among medical students

Dianhao Liu[a]; Zequn Chen[a]; Wesley J. Marrero[a*]; Thomas Thesen[b]

[a] Thayer School of Engineering, Dartmouth College, Hanover, NH, USA

[b] Department of Medical Education, Geisel School of Medicine, Dartmouth College

***Corresponding author**: Wesley J. Marrero
wesley.marrero@dartmouth.edu
Phone: (603) 646-3457
Address: 15 Thayer Dr, Hanover, NH 03755

Dianhao Liu; dianhao.liu.th@dartmouth.edu; (603) 646-3457; 15 Thayer Dr, Hanover, NH 03755

Zequn Chen; zequn.chen.th@dartmouth.edu; (603) 646-3457; 15 Thayer Dr, Hanover, NH 03755

Wesley J. Marrero; wesley.marrero@dartmouth.edu; (603) 646-3457; 15 Thayer Dr, Hanover, NH 03755

Thomas Thesen; thomas.thesen@dartmouth.edu; (603) 646-5888; 1 Rope Ferry Rd, Hanover, NH 03755

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

Abstract

Depression is a significant global cause of disability, yet it often remains undetected and untreated. Medical students have a higher likelihood of experiencing depression compared to many other age-matched groups. Previous studies have shed light on factors contributing to depression among college students. However, medical students face understudied unique experiences and demands. This gap hinders the ability to apply previously developed models to predict depression in medical students, limiting the effectiveness of mental health interventions in this population. Our study addresses this gap by presenting a machine learning model that captures the complex interplay of factors influencing the severity of depression in this population. Using data from the Healthy Minds Study, we found that a sequence of binary XGBoost models performed best at predicting the severity of depression as recorded through the Patient Health Questionnaire-9. Our modeling approach identified that prior diagnosis of depression and eating disorder, current financial stress, and younger age had the highest influence on predicted depression severity. The findings of this study hold great promise for enhancing medical education by more effectively addressing students' depression. Our predictive model may enable educators and administrators to identify students at higher risk, facilitating timely interventions and support.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

1. Introduction

Depression is a significant global cause of disability, yet it often remains undetected and untreated.[1] Medical students have a higher likelihood of experiencing depression compared to many other age-matched groups.[2,3] Starting in medical school, students confront an increased risk of depression, a pattern that persists and increases through each stage of medical training and into medical practice.[2] This is often accompanied with an increase in relationship and substance abuse issues.[2,4]

Previous studies have shed light on factors contributing to depression among college students.[5] Some common factors highly correlated with depression are: smoking[6], diminished social contact[7], adverse childhood experiences[8], poor academic performance[9], lack of financial resources and support[10], and daytime drowsiness.[11] However, medical students face understudied unique experiences and demands.[12] This gap hinders the ability to apply previously developed models to predict depression in medical students, limiting the effectiveness of mental health interventions.

Aiming to address this knowledge gap, this article adopts a novel approach to build predictive models of depression tailored to medical students. By integrating diverse variables and employing machine learning techniques, our study presents the development of a comprehensive model that captures the complex interplay of factors influencing the severity of depression in this population. Our approach promises to provide a deeper understanding of the critical predictors of depression, paving the way for tailored mental health support and interventions to enhance the well-being and academic performance of medical students.

The organization of this paper is as follows: firstly, the methodology section details the novel approach and predictive model construction, encompassing data collection, variable selection, and statistical analysis techniques. The results obtained from the predictive models are presented and interpreted, emphasizing the key factors influencing depression scores among

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

medical students. A thorough discussion of the study's findings, their implications, limitations, and recommendations for future research and mental health interventions for medical students follows. Finally, the conclusion provides a concise summary of the paper's contributions and their potential impact on promoting mental well-being among medical students.

2. Methods

2.1 Study population

The study population for this research is built upon the Healthy Minds Study (HMS) performed by the Healthy Minds Network, a nonprofit organization dedicated to mental health issues among college students in the United States.[13] Our dataset encompasses the mental health and demographic information of undergraduate and graduate students from various universities in the United States, spanning the academic years 2009-2022. Within the dataset, depression scores ranging from 0 to 27 are assigned to each student, based on their responses to Patient Health Questionnaire-9 (PHQ-9). Any samples lacking depression scores are excluded from this study.

2.2 Factors and response variables

The HMS dataset comprises numerous variables, many of which are poorly associated with depression. Including all these variables in the model may significantly increase computational burden and impede the interpretation of variables' effects. To address these concerns, we include variables in our models based on their relative importance in predicting depression.

For categorical input data, two commonly employed model agnostic variable importance methods are the chi-square statistic and the mutual information statistic.[14] In this dataset, several

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

variables are highly correlated, rendering the chi-square method unsuitable as it assumes independence between variables. Consequently, we assess the association between each variable and depression using the mutual information criterion. The scores derived from the mutual information criterion signify the degree of information shared between each variable and a response variable. Higher scores indicate a stronger relationship between the variable and depression. We empirically choose a threshold value representing a significant leap in mutual information scores to select the variables included in our model. Variables derived from depression (e.g., "Have you ever lost interest in things you liked for two weeks?") are excluded from this analysis.

The degree of depression in the HMS dataset is recorded on a scale from 0 to 27. [13] We use the depression score directly as an ordinal response variable (i.e., higher scores represent more acute depression levels) to model the severity of depression.

2.3 Data preprocessing

The HMS database consists of a small fraction of medical students, a small number of medical students with higher depression scores can affect the accuracy of predictive models, because the model cannot capture sufficient information from small dataset. To address this issue, we consider the enhancement of the medical student training set with other graduate students and the general (undergraduate and graduate) student population. That is, three different training sets are used: the medical student training set, graduate student training set, and general student training set. To maintain the accuracy of the prediction results, the test set only includes medical students without any controlled cluster sampling.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

2.4 Algorithms and model evaluation

To obtain an ordinal classification model with high predictive accuracy, we employ and compare the performance of various machine learning algorithms. Specifically, evaluate the ordinal extensions of binary classification models commonly used algorithms in the mental health literature.[15–19] We first build an ordinal logistic model and a support vector ordinal regression model to predict the severity of depression.[20] Since most machine learning models have been developed for binary tasks, we use the approach by Frank and Hall to convert our ordinal problem into a sequence of binary classifications.[21,22] This conversion allows us to assess the performance of XGBoost and neural networks, two widely used algorithms in the mental health literature.[23,24] Lastly, a multinomial logistic regression model is developed to predict depression scores as if they were unordered categories for comparison purposes.[21]

As shown in Figure 1, five-fold cross-validation is used to prevent overfitting by tuning and evaluating the models multiple times on different training and validation datasets, respectively.[25] This approach ensures that our models do not excessively fit the training data during their tuning process.

We assess our multi-class (i.e., ordinal, and multinomial) models using the mean absolute error (MAE) of its predictions, which represents the average absolute difference between predicted values and actual observations. A model demonstrating consistent low MAE across different validation sets in our cross-validation framework indicates suitability for generalizing to previously unseen data. The MAE among models across the cross-validation framework are statistically compared using Wilcoxon rank-sum tests with the Bonferroni correction. The best-performing multi-class model is used to estimate the severity of depression in the test set. The overall significance level of our statistical analyses is set to 0.05.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

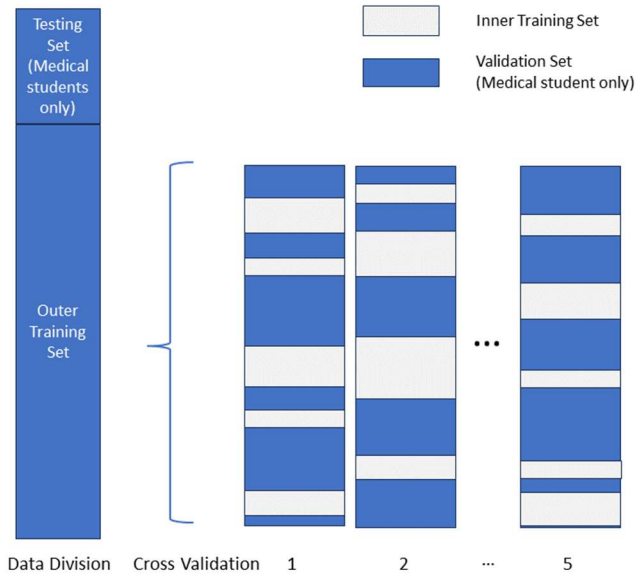


Figure 1. Initial data division and Monte Carlo cross-validation process.

2.5 Variable importance and partial dependence

We use model-dependent variable importance metrics to measure the relative contribution of each feature to the best-performing model's predictions. Permutation importance is employed, involving the shuffling of single variable values while keeping the remaining variables values as before.[26] A noticeable gap in MAE between the original and shuffled input shows a factor plays an important role in the predictive model.

To get a clearer picture of how changes in variables affect the output of the best-performing model, we use partial dependence analysis.[27] This analysis offers visualizations of the influence of specific variables on our model's predictions, while holding other variables constant. After selecting variables with high importance, we systematically vary their values. The resulting predictions from our trained models are plotted against the corresponding variable values, enabling us to observe trends and relationships.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

3. Results

3.1 Descriptive results

The records from 2009 to 2021 contain information of 484,653 students, including 171,800 (35.45%) graduate students and 4,923 (1.02%) medical students. Following the definition from the Healthy Minds Network for depression (i.e., a PHQ-9 $\geq$ 14),[13] the number of individuals with depression is 72,059 (14.87%) among all students, 17,453 (3.60%) among graduate students, and 319 (0.07%) among medical students. Figure 1 and Supplemental Figure 1 show the distributions of the number of overall students and students with depression by year, respectively. Supplemental Table 1 in the Appendix displays the demographic as well as physical and mental health characteristics of the students recorded. We select 19 variables with a mutual information score above a 10% importance threshold. Supplemental Table 2 presents a full listing of students' characteristics.

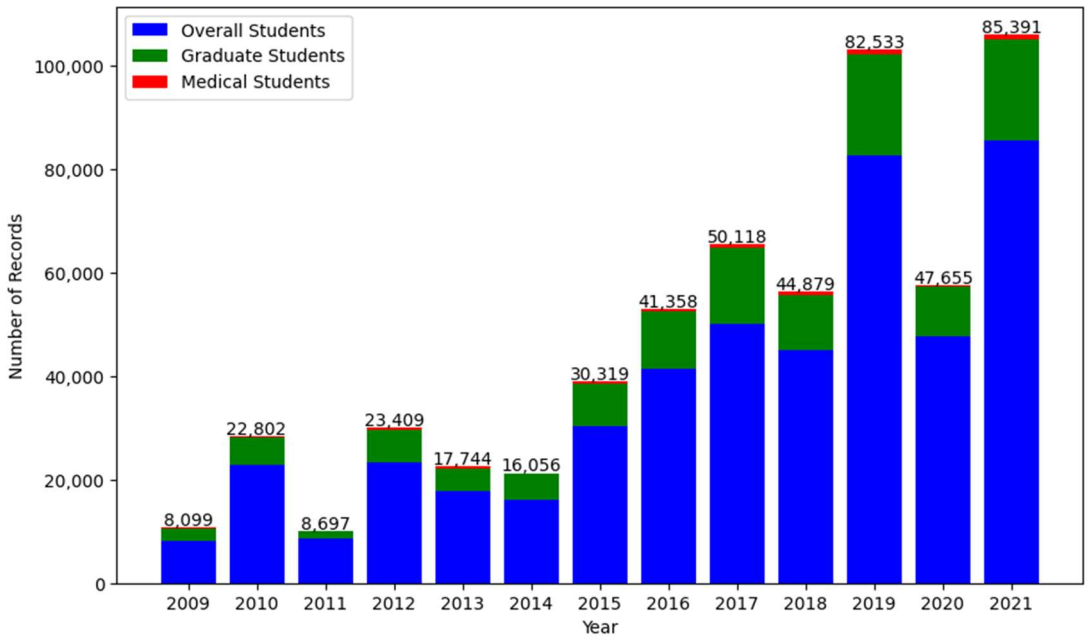


Figure 2. Number of records included in different years.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

3.3 Multi-class model evaluation

Based on the various training sets and algorithms considered, 27 different models are trained to predict depression severity. The performance of our ordinal models on the cross-validation framework is presented in Figure 3. We find that ordinal logistic regression and support vector ordinal regression perform significantly better than the multinomial logistic regression ($p<0.0001$). Moreover, the sequence of binary XGBoost models outperformed the remaining models (95% CI, 3.08 to 3.39, $p< 0.0024$). The Wilcoxon rank-sum test with the Bonferroni correction across all pairs of models are shown in Supplemental Table 3 of the Appendix. Our best-performing method, the sequence of binary XGBoost models, presents an MAE of 3.32 when trained in the outer training set and applied to the test set.

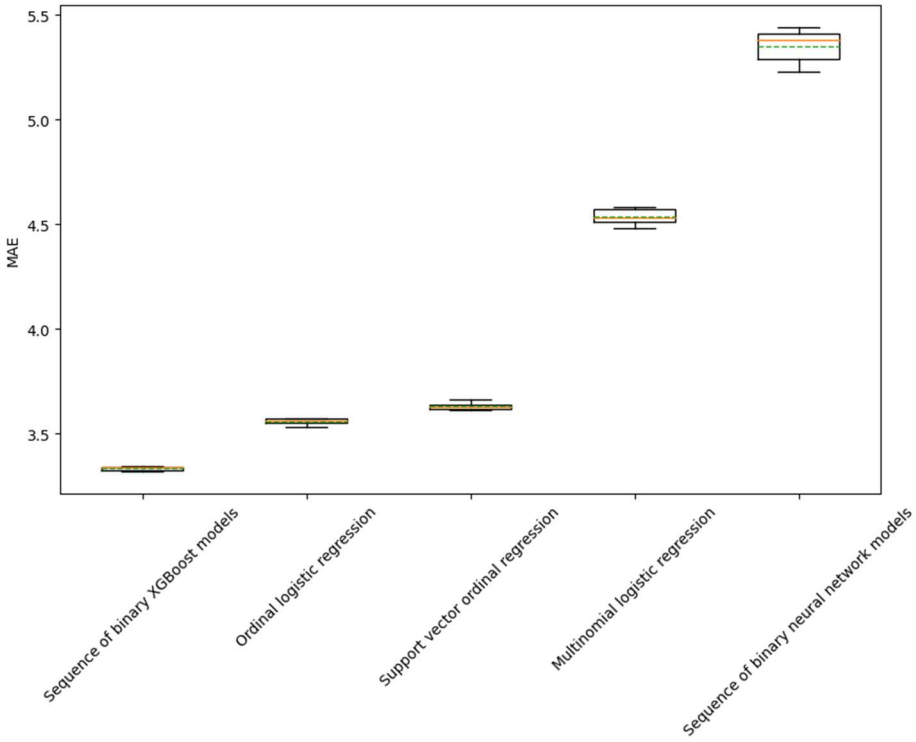


Figure 3. MAE of potential multi-class models to predict overall depression.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

3.4 Variable importance

We analyze the importance of variables in our sequence of binary XGBoost models since it has the highest accuracy. Our variable importance results in Figure 4 reveal that history of depression has the highest influence on predicted depression severity. Other notable factors include eating disorder (SCOFF[28] score), financial status, and age group. Figure 5 shows the marginal impact of different levels of each of the four most important variables (i.e., variables with an importance score of at least 10%) on the predicted depression score with partial dependence plots. The rest of the partial dependence plots are shown in Supplemental Figure 2. History of depression, higher SCOFF scores, lower financial status, and younger age are strongly associated with higher depression scores.

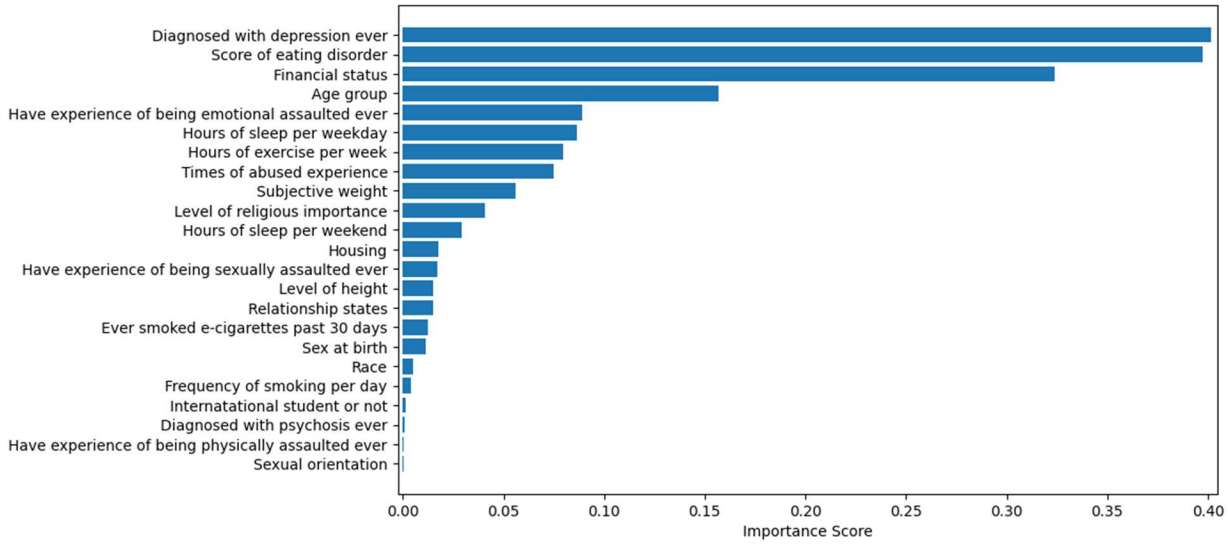


Figure 4. Permutation importance score in MAE of each predictor in sequence of binary XGBoost models.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

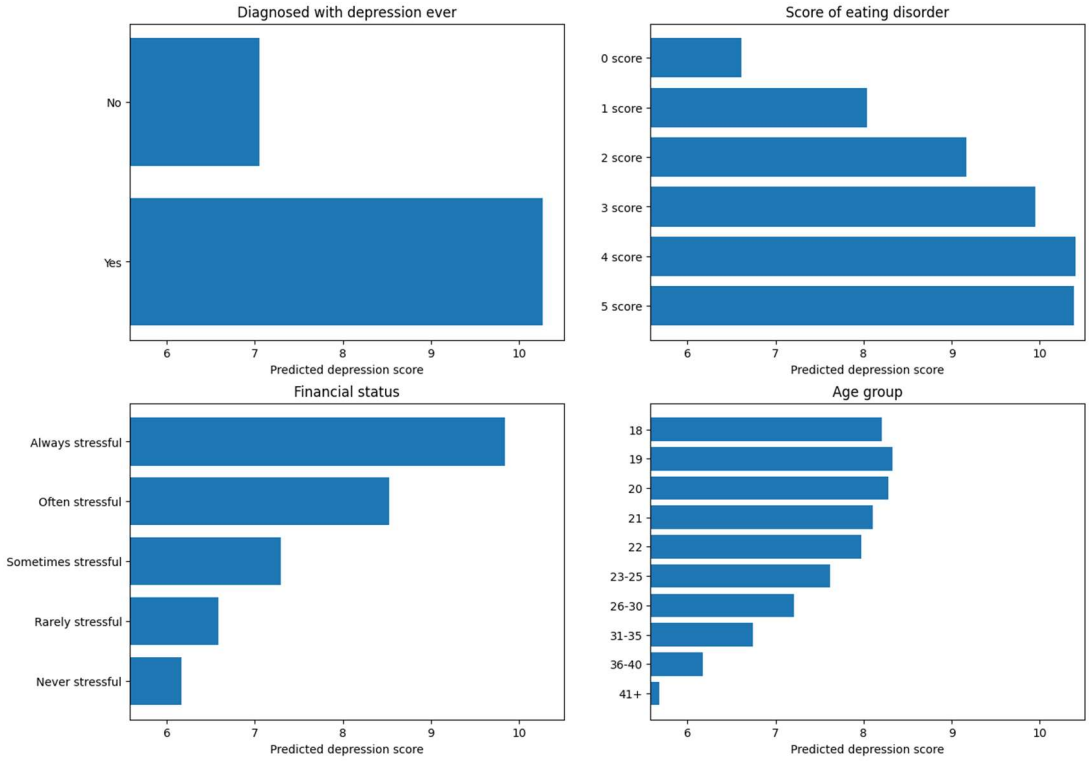


Figure 5. Partial dependence of depression history, eating disorder, financial status, and age group.

4. Discussion

Our study aimed to identify the key predictors of depression among US medical students, a demographic known to experience high levels of stress and mental health challenges. By exploring multiple machine learning models, we found that a sequence of binary XGBoost models performed best when predicting depression severity. The best-performing model identified several factors that most strongly predict the occurrence of depression in medical students. Using mean absolute error analysis, we were able to rank predictive factors by decreasing order of predictive power: 1. prior diagnosis of depression, 2. prior eating disorder, 3. current financial stress, 4. younger age, 5. prior experience of emotional assault, 6. current lack of sleep during the week, and 7. current lack of exercise.

Consistent with existing literature, a history of depression emerged as the strongest predictor of future depressive episodes. This finding aligns with previous research that shows

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

that about half of the individuals who overcome their initial episode of depression are likely to experience at least one more episode in their lifetime. This probability increases to 80% for those who have had two or more previous episodes.[29]

An association between disordered eating behaviors and mood disorders has been shown previously in the literature.[30,31] Individuals with eating disorders may experience depressive symptoms due to the psychological and physiological stressors associated with these disorders. Conversely, depression can lead to disordered eating as a coping mechanism. Our findings underscore the importance of considering the comorbidity between depression and eating disorders in medical students' holistic mental health.

Medical students often incur significant debt to finance their education and due to the high demands of medical school are unable to maintain gainful employment at the same time. This burden is disproportionately felt by students from communities Underrepresented in Medicine (URiM).[32] Our results highlight the role of financial stress as a predictor of depression among medical students and highlight the need for comprehensive support systems that address both the economic and psychological need of this population.

The finding between younger age and increased susceptibility to depression among medical students suggests that young students face unique challenges that predispose them to higher risks of depression in the demanding environment of higher education, a phenomenon previously noted in educational psychology research.[8] At the same time, it suggests that older students with more previous experience may display more resilience that helps them adapt better to the rigors of medical school.

Emotional assault, which can encompass various forms of emotional abuse and traumatic experiences, has been consistently linked in research to a heightened risk of developing mental health disorders, including depression. The significant association between the experience of emotional assault and depression in medical students found int his study

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

underscores the need for a trauma-informed approach in the support and care provided within medical education settings.

Sleep deprivation has a well-established link with mental health issues, particularly depression. Adequate sleep is crucial for cognitive functioning, emotional regulation, and overall psychological well-being.[33] In medical students, who often face irregular and extended hours due to academic and clinical responsibilities, sleep deprivation can be a significant issue. This lack of sleep can impair judgment, increase irritability, and exacerbate stress, all of which can contribute to depressive symptoms. The relationship between sleep and mental health is bidirectional; not only can poor sleep contribute to the development of mental health disorders, but existing mental health issues can further disrupt sleep patterns, creating a vicious cycle. Regular physical activity is known to have a positive impact on mental health, including depression. The absence of physical activity in the routine of medical students may leave these students more susceptible to stress and depression.[34]

The findings of this study hold great promise for enhancing medical education by more effectively addressing students' depression. Our predictive model may enable educators and administrators to identify students at higher risk, facilitating timely interventions and support. Mental health programs and resources tailored to meet the specific needs of students at risk can be implemented to create a healthier and more supportive learning environment.

However, the study also revealed some limitations on the data size and modeling approach. The HMS dataset spans from 2009 to 2022 and exhibits variations in the number of records and variables over this time frame. To mitigate this issue, we excluded certain predictors related to depression that were inconsistently available. Additionally, the dataset contains a very limited number of depression cases. To address this limitation, we employed controlled random sampling to balance our training data. Furthermore, our dataset encompasses various student categories, with a majority not being medical students. We addressed this shortcoming by utilizing an expanded dataset including the general population of

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

students (undergraduate and graduate). Bias in our model can lead to inaccurate and unfair outcomes, which can have far-reaching consequences. This work may be extended to measure and address inequitable error rates across students' social positioning or intersectionalities. Future work may focus on developing fair predictive models to account for potentially sensitive attributes and capturing potential sources of bias.

Our model can help medical practitioners predict depression scores in medical students. These scores convey intuitive information regarding the severity of medical students' depression conditions. This work also helps identify new variables that impact the severity of depression among medical students, which are not found in previous studies. Understanding the significant impact of these variables can guide medical educators and mental health professionals in developing targeted interventions.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

5. References

1. Smith K. Mental health: A world of depression. *Nature*. 2014;515(7526):180-181. doi:10.1038/515180a

2. Puthran R, Zhang MWB, Tam WW, Ho RC. Prevalence of depression amongst medical students: a meta-analysis. *Med Educ*. 2016;50(4):456-468. doi:10.1111/medu.12962

3. Goebert D, Thompson D, Takeshita J, et al. Depressive Symptoms in Medical Students and Residents: A Multischool Study. *Acad Med*. 2009;84(2):236. doi:10.1097/ACM.0b013e31819391bb

4. Bahreini M, Hassanzade M, Akaberian S, Mirzaei K, Zaeri S. Prediction of Mental Health of students Based on their Parental Bonding. Published online 2016.

5. Liu XQ, Guo YX, Zhang WJ, Gao WJ. Influencing factors, prediction and prevention of depression in college students: A literature review. *World J Psychiatry*. 2022;12(7):860-873. doi:10.5498/wjp.v12.i7.860

6. Song SJ, Ziegler R, Arsenault L, Fried LE, Hacker K. Asian Student Depression in American High Schools: Differences in Risk Factors. *J Sch Nurs*. 2011;27(6):455-462. doi:10.1177/1059840511418670

7. Vanhalst J, Luyckx K, Teppers E, Goossens L. Disentangling the Longitudinal Relation Between Loneliness and Depressive Symptoms: Prospective Effects and the Intervening Role of Coping. *J Soc Clin Psychol*. 2012;31(8):810-834. doi:10.1521/jscp.2012.31.8.810

8. Kelifa MO, Yang Y, Carly H, Bo W, Wang P. How Adverse Childhood Experiences Relate to Subjective Wellbeing in College Students: The Role of Resilience and Depression. *J Happiness Stud*. 2021;22(5):2103-2123. doi:10.1007/s10902-020-00308-7

9. Ngin C, Pal K, Tuot S, Chhoun P, Yi R, Yi S. Social and behavioural factors associated with depressive symptoms among university students in Cambodia: a cross-sectional study. *BMJ Open*. 2018;8(9):e019918. doi:10.1136/bmjopen-2017-019918

10. Çelik N, Ceylan B, Ünsal A, Çağan Ö. Depression in health college students: relationship factors and sleep quality. *Psychol Health Med*. 2019;24(5):625-630. doi:10.1080/13548506.2018.1546881

11. Gonsalvez I, Li JJ, Stevens C, Chen JA, Liu CH. Preexisting Depression and Daytime Sleepiness in Women and Men. *Behav Sleep Med*. 2022;20(4):380-392. doi:10.1080/15402002.2021.1924720

12. Slavin SJ. Medical Student Mental Health: Culture, Environment, and the Need for Change. *JAMA*. 2016;316(21):2195-2196. doi:10.1001/jama.2016.16396

13. Explore our Data. Healthy Minds Network. Accessed September 18, 2023. https://healthymindsnetwork.org/research/data-for-researchers/

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

14. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E*. 2004;69(6):066138. doi:10.1103/PhysRevE.69.066138

15. Chung J, Teo J. Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges. *Appl Comput Intell Soft Comput*. 2022;2022:e9970363. doi:10.1155/2022/9970363

16. Cho G, Yim J, Choi Y, Ko J, Lee SH. Review of Machine Learning Algorithms for Diagnosing Mental Illness. *Psychiatry Investig*. 2019;16(4):262-269. doi:10.30773/pi.2018.12.21.2

17. Shin D, Lee KJ, Adeluwa T, Hur J. Machine Learning-Based Predictive Modeling of Postpartum Depression. *J Clin Med*. 2020;9(9):2899. doi:10.3390/jcm9092899

18. Lee C, Kim H. Machine learning-based predictive modeling of depression in hypertensive populations. *PLOS ONE*. 2022;17(7):e0272330. doi:10.1371/journal.pone.0272330

19. Lewis RJ. An Introduction to Classification and Regression Tree (CART) Analysis.

20. Chu W, Keerthi SS. Support Vector Ordinal Regression. *Neural Comput*. 2007;19(3):792-815. doi:10.1162/neco.2007.19.3.792

21. Perktold J, Skipper Seabold, Sheppard K, et al. statsmodels/statsmodels: Release 0.14.0. Published online May 5, 2023. doi:10.5281/ZENODO.593847

22. Frank E, Hall M. A Simple Approach to Ordinal Classification. In: De Raedt L, Flach P, eds. *Machine Learning: ECML 2001*. Lecture Notes in Computer Science. Springer; 2001:145-156. doi:10.1007/3-540-44795-4_13

23. Du C, Liu C, Balamurugan P, Selvaraj P. Deep Learning-based Mental Health Monitoring Scheme for College Students Using Convolutional Neural Network. *Int J Artif Intell Tools*. 2021;30(06n08):2140014. doi:10.1142/S0218213021400145

24. Jain T, Jain A, Hada PS, Kumar H, Verma VK, Patni A. Machine Learning Techniques for Prediction of Mental Health. In: *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*. ; 2021:1606-1613. doi:10.1109/ICIRCA51532.2021.9545061

25. Wong TT, Yeh PY. Reliable Accuracy Estimates from k-Fold Cross Validation. *IEEE Trans Knowl Data Eng*. 2020;32(8):1586-1594. doi:10.1109/TKDE.2019.2912815

26. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. *Bioinformatics*. 2010;26(10):1340-1347. doi:10.1093/bioinformatics/btq134

27. Cafri G, Bailey BA. Understanding Variable Effects from Black Box Prediction: Quantifying Effects in Tree Ensembles Using Partial Dependence. *J Data Sci*. 2016;14(1):67-95. doi:10.6339/JDS.201601_14(1).0005

28. Morgan JF, Reid F, Lacey JH. The SCOFF questionnaire. *West J Med*. 2000;172(3):164-165.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. (DSM-IV). *Am J Psychiatry*. 1995;152(8):1228-1228. doi:10.1176/ajp.152.8.1228

30. Godart N, Radon L, Curt F, et al. Mood disorders in eating disorder patients: Prevalence and chronology of ONSET. *J Affect Disord*. 2015;185:115-122. doi:10.1016/j.jad.2015.06.039

31. Turner H, Marshall E, Wood F, Stopa L, Waller G. CBT for eating disorders: The impact of early changes in eating pathology on later changes in personality pathology, anxiety and depression. *Behav Res Ther*. 2016;77:1-6. doi:10.1016/j.brat.2015.11.011

32. Roche R, Manzi J, Baker S, Ndubuizu T. Under-represented minority students and identification of obstacles to a career in medicine. *Clin Teach*. 2021;18(2):186-190. doi:10.1111/tct.13312

33. Alhola P, Polo-Kantola P. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatr Dis Treat*. 2007;3(5):553-567.

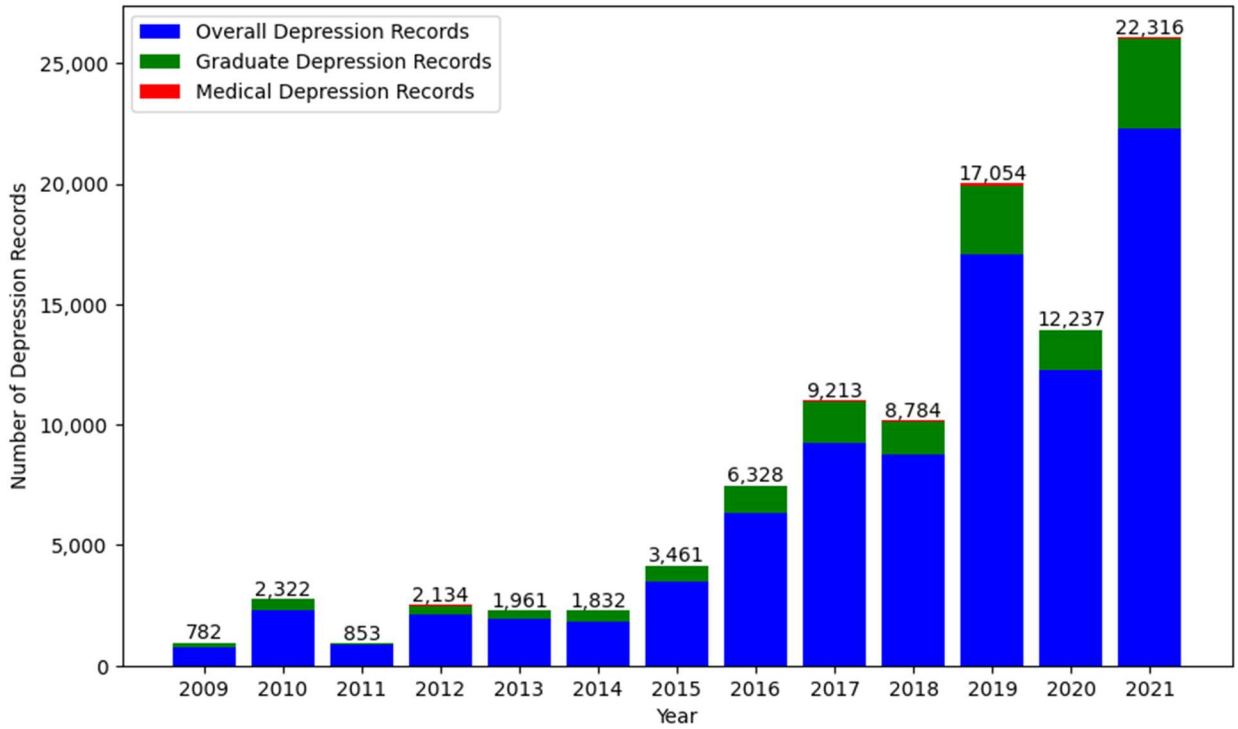
34. Mikkelsen K, Stojanovska L, Polenakovic M, Bosevski M, Apostolopoulos V. Exercise and mental health. *Maturitas*. 2017;106:48-56. doi:10.1016/j.maturitas.2017.09.003

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

6. Appendix

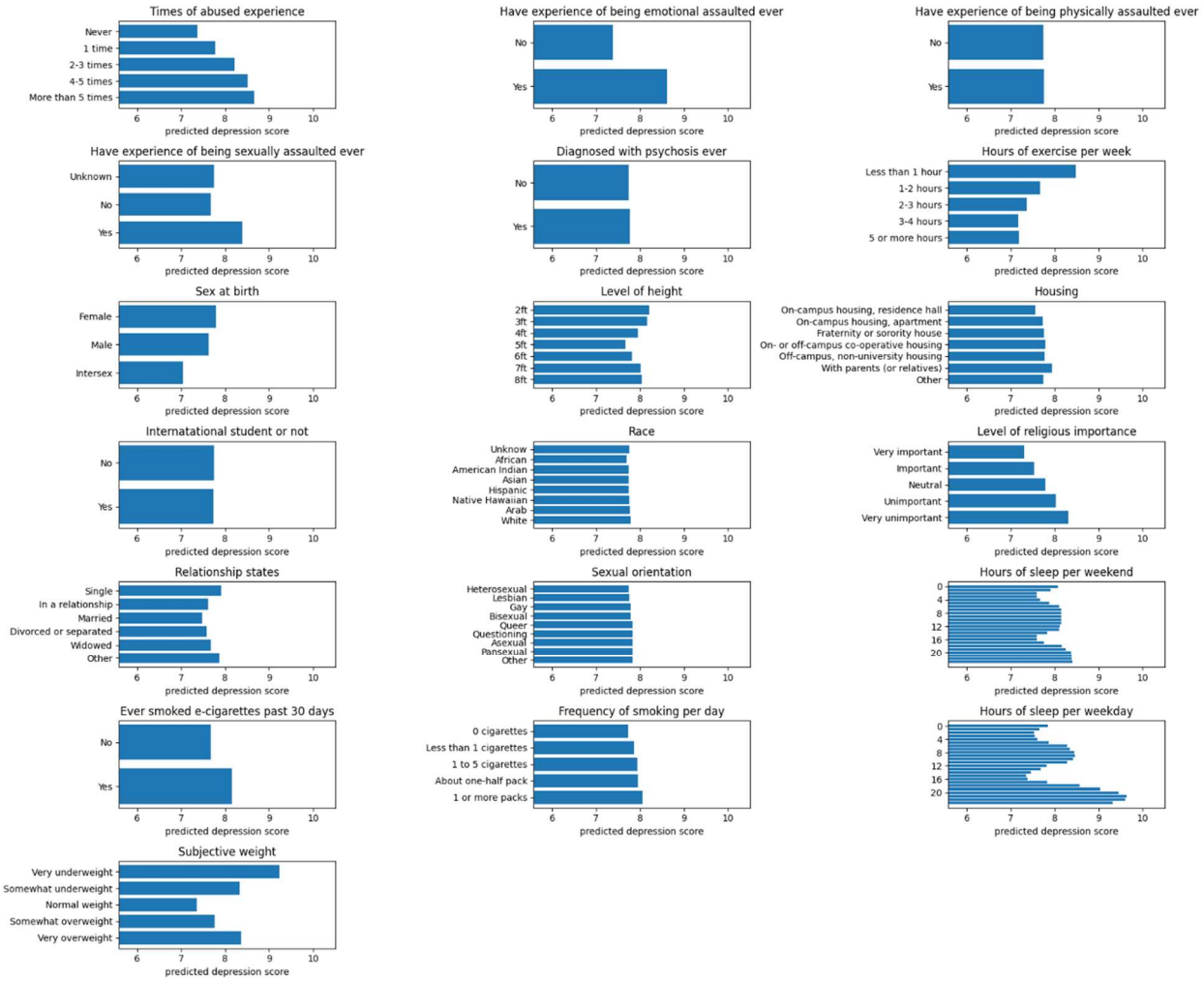
6.1 Supplemental Figures



Supplemental Figure 1. Number of depression records included in different years.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .



Supplemental Figure 2. Partial dependence plot for rest variables in sequence of binary XGBoost models.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

6.2 Supplemental Tables

Supplemental Table 1. Selected variables with their meanings and scores.

Demographic variables		
Variable name	Meaning	Mutual Information Score
age	Level of age	0.060806
sex_birth	Sex at birth	0.059762
race_black	Race: African American	0.068075
race_ainaan	Race: American Indian	0.079659
race_asian	Race: Asian	0.038503
race_his	Race: Hispanic	0.061001
race_pi	Race: Native Hawaiian	0.062572
race_mides	Race: Middle Eastern	0.029515
race_white	Race: White	0.073291
international	International student	0.007698
fincur	Financial status	0.112953
relship	Relationship status	0.006771
sexual_h	Sexual orientation: Heterosexual	0.005165
sexual_l	Sexual orientation: Lesbian	0.041546
sexual_g	Sexual orientation: Gay	0.031767
sexual_bi	Sexual orientation: Bisexual	0.022946
sexual _queer	Sexual orientation: Queer	0.018342
sexual_quest	Sexual orientation: Questioning	0.016848
sexual_other	Sexual orientation: Self-identify	0.009148
religios	Level of religious importance	0.071104
height_ft	Level of height	0.149735

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

Mutual information criterion variables		
Variable name	Meaning	Score
dx_psy	Diagnosed with psychosis ever	0.337642
lone_isolated	Frequency of feeling isolated currently	0.262182
smok_vape	Use of electronic cigarette or vape pen ever	0.198321
talk1_6	Prefer to seek emotional help from religion	0.194991
smok_freq	Smoking frequency (in 30 days) level	0.181191
sa_exp	Have experience of being sexually assaulted ever	0.170068
Q3_22	Times of heavy drinking in 2 weeks	0.163248
drug_coc	Use of cocaine in 30 days	0.161053
exerc	Hours of exercise per week	0.158574
body_sr	Ever worry about weight	0.155654
assault_phys	Have experience of being physically assaulted ever since coming to school	0.151466
ed_scoff	Eating disorder (0-5 SCOFF score)	0.142633
adjust_aca_1	Academic adjustments in college	0.139382
birthcontrol_3	Use of contraceptive pills in 30	0.129280

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

	days	
abuse_life	Times of abused experience	0.114874
assault_emo	Have experience of being sexually assaulted ever since coming to school	0.118799
sleep_wknight	Sleep time in weekdays	0.1098-0.1095

Supplemental Table 2. Detail of selected variables

Value	Count	Percentage
	Degree of depression (Score)	
0	28,644	5.91%
1	25,471	5.26%
2	36,952	7.62%
3	39,382	8.13%
4	37,782	7.80%
5	35,368	7.30%
6	33,258	6.86%
7	31,439	6.49%
8	28,594	5.90%
9	26,176	5.40%
10	21,388	4.41%
11	18,750	3.87%
12	16,722	3.45%
13	14,902	3.07%
14	13,397	2.76%
15	11,819	2.44%
16	10,767	2.22%
17	9,080	1.87%
18	8,542	1.76%
19	6,703	1.38%
20	5,734	1.18%
21	5,297	1.09%
22	4,374	0.90%
23	3,716	0.77%

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

24	3,613	0.75%
25	2,489	0.51%
26	1,548	0.32%
27	2,746	0.57%
Eating disorder (SCOFF Score)		
0	203,763	50.45%
1	94,877	23.50%
2	59,537	14.75%
3	31,604	7.83%
4	11,398	2.82%
5	2,411	0.59%
Ever been abused		
Never	194,346	56.41%
1 time	29,720	8.62%
2-3 times	54,772	15.89%
4-5 times	16,366	4.75%
More than 5 times	49,301	14.31%
Age group		
18	57,025	11.77%
19	74,176	15.30%
20	73,069	15.08%
21	70,530	14.55%
22	44,253	9.13%
23-25	62,964	12.99%
26-30	52,101	10.75%
31-35	21,403	4.42%
36-40	11,215	2.31%
41+	17,875	3.69%
Ever emotionally assaulted past year		
No	227,706.00	66.44%
Yes	82,619.00	33.55%
Ever physically assaulted past year		
No	292,463.00	73.37%
Yes	17,664.00	26.62%
Ever sexually assaulted past year		
Unknow	257,727.00	53.18%
No	205,485.00	42.40%
Yes	21,441.00	4.42%
Medical student or not		
No	100,972.00	98.38%

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

Yes	5,324.00	1.61%
Ever psychosis		
No	254,734.00	99.34%
Yes	1,657.00	0.64%
Exercise hour per week		
Less than 1 hour	97,277	27.44%
1-2 hours	83,192	23.36%
2-3 hours	65,325	18.26%
3-4 hours	93,727	26.38%
5 or more hours	15,888	4.54%
Financial status		
Always stressful	49,626	10.24%
Often stressful	105,801	21.83%
Sometimes stressful	198,958	41.05%
Rarely stressful	104,152	21.49%
Never stressful	23,215	4.79%
Gender		
Female	324,283.00	67.04%
Male	159,406.00	32.86%
Intersex	440	0.09%
Graduate student or not		
No	299,500.00	61.80%
Yes	185,153.00	38.20%
Level of height		
1 ft	1,497.00	0.34%
2 ft	5,449.00	1.24%
3 ft	7,058.00	1.61%
4 ft	12,807.00	2.92%
5 ft	342,705.00	78.31%
6 ft	54,898.00	12.54%
7 ft	6,993.00	1.59%
8 ft	6,179.00	1.51%
Housing		
On-campus housing, residence hall	127,933	26.72%
On-campus housing, apartment	28,613	5.97%
Fraternity or sorority house	8,493	1.77%
On- or off-campus co-operative housing	52,320	10.93%

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

Off-campus, non-university housing	178,256	37.24%
With parents (or relatives)	72,383	15.12%
Other	10,634	2.221%
International student or not		
No	417,401.00	92.26%
Yes	35,043.00	7.74%
Race		
Unknow	296,656.00	61.21%
African	36,981.00	7.63%
American Indian	4,841.00	1.00%
Asian	52,841.00	10.90%
Hispanic	40,826.00	8.42%
Native Hawaiian	230	0.05%
Arab	1,385.00	0.29%
White	50,893.00	10.50%
Religious		
Very important	70,871.00	14.62%
Important	90,824.00	18.74%
Neutral	126,699.00	26.14%
Unimportant	102,067.00	21.06%
Very unimportant	93,656.00	19.32%
Relationship states		
Single	241,960.00	49.92%
In a relationship	175,551.00	36.22%
Married	58,953.00	12.16%
Divorced or separated	4,077.00	0.84%
Widowed	496	0.10%
Other	3,260.00	0.67%
Sexual orientation		
Heterosexual	393,391.00	92.91%
Lesbian	2,586.00	0.61%
Gay	1,936.00	0.46%
Bisexual	16,673.00	3.96%
Queer	2,381.00	0.57%
Questioning	3,082.00	0.73%
Asexual	1,169.00	0.27%
Pansexual	1,848.00	0.44%
Other	31	0.01%
Ever smoked e-cigarettes past 30 days		

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

No	204,121.00	84.80%
Yes	36,568.00	71519%
Frequency of smoking per day		
0 cigarettes	327,789.00	92.16%
Less than 1 cigarettes	12,525.00	3.52%
1 to 5 cigarettes	9,054.00	2.54%
About one-half pack	3,471.00	0.97%
1 or more packs	2,801.00	0.78%
Hours of weekday sleeping		
0	2,450.00	0.58%
1	1,265.00	0.30%
2	4,692.00	1.11%
3	13,977.00	3.31%
4	23,006.00	5.45%
5	23,477.00	5.57%
6	27,080.00	6.42%
7	26,657.00	6.32%
8	15,256.00	3.62%
9	2,840.00	0.67%
10	973	0.23%
11	407	0.09%
12	1,454.00	0.34%
13	4,484.00	1.06%
14	19,403.00	4.60%
15	55,881.00	13.25%
16	84,790.00	20.11%
17	62,788.00	14.89%
18	31,239.00	7.41.%
19	12,310.00	2.92%
20	4,799.00	1.13%
21	1,377.00	0.32%
22	458	0.10%
23	366	0.08%
Subjective weight		
Very underweight	4,053.00	1.07%
Somewhat underweight	32,615.00	8.66%
Normal weight	165,988.00	44.11%
Somewhat overweight	135,192.00	35.93%
Very overweight	38,410.00	10.20%
Year that data was recorded		

medRxiv preprint doi: https://doi.org/10.1101/2023.12.14.23299975; this version posted December 15, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-ND 4.0 International license .

2009	8,099.0	1.67%
2010	22,802.0	4.70%
2011	8,697.0	1.79%
2012	23,409.0	4.83%
2013	17,744.0	3.66%
2014	16,056.0	3.31%
2015	30,319.0	6.26%
2016	41,358.0	8.53%
2017	50,118.0	10.34%
2018	44,879.0	9.26%
2019	82,533.0	17.03%
2020	47,655.0	9.83%
2021	85,391.0	17.62%

Supplemental Table 3. P-values of Wilcoxon Rank-Sum Tests with Bonferroni Correction of the multi-class models. Values are shown using an upper triangular representation for simplicity purposes. P-values smaller than 0.001 are represented by dashes.

Model	Sequence of binary XGBoost models	Ordinal logistic regression	Support vector ordinal regression	multinomial logistic regression	Sequence of neural network models
Sequence of binary XGBoost models	1	0.008	-	-	-
Ordinal logistic regression		1	-	-	-
Support vector ordinal regression			1	-	-
multinomial logistic regression				1	-
Sequence of neural network models					1