IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Anxiety Depression and Stress Prediction among College Students using Machine Learning Algorithms

Aditi Borade[1], Mansi Ugale[2], Om Dighe[3], Prathamesh Warkhede[4], Vidya Kale[5]
Department of Information Technology[1,2,3,4,5]
Matoshri Aasarabai Polytechnic, Eklahare, Nashik, Maharashtra, India

Abstract: *The prevalence of mental health issues, specifically anxiety, depression, and stress, among college students has become a growing concern. Identifying early signs and risk factors is crucial for timely intervention and support. This research aims to develop a predictive model using machine learning algorithms to forecast anxiety, depression, and stress levels among college students based on various input features. Data will be collected through surveys, incorporating a range of demographic, academic, and lifestyle factors. The dataset will be preprocessed to handle missing values, normalize variables, and ensure compatibility with machine learning algorithms. Feature engineering techniques will be applied to extract relevant patterns and relationships from the data. Several machine learning algorithms, including but not limited to Decision Trees, Random Forest, Support Vector Machines, and Neural Networks, will be employed to build predictive models. The models will be trained on a portion of the dataset and validated using another portion to assess their accuracy, sensitivity, and specificity.*

The study aims to explore the effectiveness of different algorithms in predicting mental health outcomes and to identify the most influential features contributing to these predictions. Additionally, an interpretable model will be pursued to enhance the understanding of the relationships between input features and mental health indicators. The ultimate goal is to develop a reliable and practical tool that can assist universities and mental health professionals in identifying at-risk students early on, enabling targeted interventions and support services. This research contributes to the growing field of digital mental health by leveraging machine learning techniques to address the pressing issue of mental health challenges among college students.

Keywords: IMental Health, Anxiety, Depression Stress, College Students, Early Intervention, Predictive Modeling Machine Learning Algorithms, Data Collection, Surveys, Demographic Factors

I. INTRODUCTION

In recent years, the mental health of college students has emerged as a significant concern, with rising instances of anxiety, depression, and stress impacting academic performance and overall well-being. The transition to higher education, coupled with academic pressures, social challenges, and lifestyle adjustments, creates a complex environment that can contribute to mental health issues. Recognizing the importance of early intervention, this research endeavors to employ cutting-edge solutions in the form of machine learning algorithms to predict and understand the dynamics of anxiety, depression, and stress among college students. The prevalence of these mental health challenges necessitates a proactive approach to identify at-risk individuals and provide timely support. Traditional methods of assessment often rely on subjective self-reporting or retrospective analysis, potentially missing critical early warning signs. Machine learning, with its capacity to analyze large datasets and discern intricate patterns, offers a promising avenue to enhance our understanding and prediction of mental health outcomes among college students. This study seeks to leverage a comprehensive dataset, encompassing a spectrum of demographic, academic, and lifestyle factors relevant to college life. Through meticulous data preprocessing and feature engineering, the research aims to extract meaningful insights from the collected information. By employing a range of machine learning algorithms, including Decision Trees, Random Forest, Support Vector Machines, and Neural Networks, the study endeavors to build robust

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

predictive models capable of forecasting anxiety, depression, and stress levels. The exploration of different algorithms aims to uncover the most effective methodologies for predicting mental health outcomes in a college setting. Additionally, the pursuit of an interpretable model is driven by the desire to not only predict outcomes but also to understand the influential features and relationships that contribute to these predictions. Ultimately, the overarching goal of this research is to develop a practical and reliable tool that can assist universities and mental health professionals in identifying at-risk students early in their college journey. Such a tool could pave the way for targeted interventions and support services, contributing to the broader field of digital mental health and addressing the pressing issue of mental health challenges faced by college students.

Background:

College life, characterized by academic pressures, social transitions, and newfound independence, can be both exhilarating and challenging. However, the intensification of stressors has contributed to a surge in mental health issues, particularly anxiety and depression. These challenges not only impact individual students but also pose a broader risk to the educational community and society at large. Traditional approaches to identifying and addressing mental health concerns often rely on self-reporting or periodic assessments. However, these methods may not capture subtle changes or provide timely insights into emerging issues. The integration of machine learning offers a promising avenue to enhance the precision and efficiency of mental health prediction by leveraging diverse datasets and advanced analytical techniques.

Objective of the Study:

The primary objective of this research is to develop a predictive model that utilizes machine learning algorithms to forecast anxiety, depression, and stress levels among college students. By incorporating a comprehensive range of input features encompassing demographic, academic, and lifestyle factors, the study aims to create a nuanced understanding of the multifaceted nature of mental health challenges in the college population.

Methodology:

Data collection will involve administering surveys designed to capture a holistic view of students' lives, including their academic performance, social interactions, and daily routines. Following data collection, rigorous preprocessing techniques will be applied to handle missing values, normalize variables, and optimize the dataset for compatibility with machine learning algorithms. Feature engineering will then extract valuable patterns and relationships from the data.An array of machine learning algorithms, such as Decision Trees, Random Forest, Support Vector Machines, and Neural Networks, will be employed to construct predictive models. These models will undergo training and validation processes to assess their accuracy, sensitivity, and specificity in predicting mental health outcomes.

Significance of the Study:

The findings of this research hold significant implications for the development of a reliable and practical tool that can assist universities and mental health professionals in identifying at-risk students early on. This, in turn, enables targeted interventions and support services, contributing to the broader field of digital mental health.By merging the realms of mental health research and machine learning, this study represents a pioneering effort to address the pressing issue of mental health challenges among college students. The insights gained from this research have the potential to reshape current practices and create a more proactive and informed approach to supporting the mental well-being of the college community.

II. METHOD

Data Collection:

Design and administer surveys to collect data on anxiety, depression, and stress levels among college students.

Include a diverse set of demographic variables (age, gender, ethnicity), academic factors (GPA, course load), and lifestyle variables (sleep patterns, exercise habits).

Ensure ethical considerations and obtain informed consent from participants.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Dataset Preparation:
Clean and preprocess the collected data to handle missing values and outliers.
Normalize and standardize variables to bring them to a consistent scale.
Explore the dataset to understand its distribution and characteristics.

Feature Engineering:
Identify relevant features that may contribute to mental health outcomes.
Create new features or transformations that enhance the predictive power of the model.
Utilize domain knowledge to inform feature selection.

Machine Learning Algorithms:
Select a variety of machine learning algorithms suitable for classification tasks, such as Decision Trees, Random Forest, Support Vector Machines, and Neural Networks.
Divide the dataset into training and validation sets for model development and evaluation.

Model Training:
Train each selected algorithm on the training dataset.
Adjust hyperparameters to optimize model performance.
Evaluate models using appropriate metrics, considering accuracy, sensitivity, specificity,

Model Validation:
Validate the models on an independent dataset not used during training to assess their generalization capability.
Employ techniques like cross-validation to ensure robust performance evaluation.

Interpretability Analysis:
For selected models, conduct feature importance analysis to understand which features contribute most to mental health predictions.
Choose interpretable models or use techniques (such as SHAP values) to explain predictions.

Comparison of Algorithms:
Compare the performance of different machine learning algorithms to identify the most effective approach for predicting anxiety, depression, and stress levels among college students.

Ethical Considerations:
Ensure the ethical handling of data, protecting participant privacy and confidentiality.
Obtain necessary approvals from institutional review boards (IRBs) or ethics committees.

Results Analysis and Interpretation:
Analyze the results comprehensively, considering the strengths and limitations of the chosen algorithms.
Draw conclusions about the predictive power of the models and the identified risk factors.

III. ADVANTAGES OF ANXIETY, DEPRESSION, AND STRESS AMONG COLLEGE STUDENTS USING MACHINE LEARNING ALGORITHMS

Early Detection and Intervention:
Machine learning models can analyze a diverse set of factors to identify early signs of anxiety, depression, and stress.
Early detection allows for timely intervention and support, potentially preventing the escalation of mental health issues.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Personalized Risk Assessment:
Machine learning algorithms enable personalized risk assessments by considering individual characteristics and lifestyle factors.
Tailored predictions can guide targeted interventions that address specific needs and risk factors for each student.

Comprehensive Analysis of Multifactorial Influences:
Machine learning can handle large and complex datasets, considering a wide range of demographic, academic, and lifestyle variables.
This enables a more holistic understanding of the multifactorial influences contributing to mental health outcomes in college students.

Improved Prediction Accuracy:
Machine learning models, such as ensemble methods (e.g., Random Forest), often outperform traditional statistical methods in terms of prediction accuracy.
The ability to analyze intricate patterns in data enhances the accuracy of identifying students at risk.

Continuous Monitoring:
Machine learning algorithms can facilitate continuous monitoring of mental health indicators over time.
This provides a dynamic perspective on students' well-being, allowing for adaptive interventions based on changing circumstances.

Objective and Standardized Assessment:
Machine learning ensures objective and standardized assessments, reducing the potential for biases inherent in subjective evaluations.
Consistent evaluation criteria enhance the reliability of mental health predictions.

Efficient Resource Allocation:
Targeted interventions based on machine learning predictions allow for more efficient allocation of mental health resources.
Universities can prioritize support services for students identified as high-risk, optimizing the use of available resources.

Identification of Underlying Patterns:
Machine learning algorithms can uncover complex patterns and relationships in data that may not be apparent through traditional analyses.
Understanding these patterns can contribute to a deeper comprehension of the factors influencing mental health outcomes.

Digital Mental Health Advancements:
Integrating machine learning into mental health research contributes to the evolving field of digital mental health.
Technological advancements provide scalable and accessible solutions that align with the preferences of the digital-native college student population.

Interpretability and Insights:
Some machine learning models offer interpretability features, allowing researchers and practitioners to gain insights into the factors driving predictions.
Interpretable models enhance the trust and understanding of the decision-making process.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

IV. DISADVANTAGE OF ANXIETY, DEPRESSION, AND STRESS AMONG COLLEGE STUDENTS USING MACHINE LEARNING ALGORITHMS

Ethical Concerns:
The use of sensitive mental health data raises ethical considerations, including issues related to privacy, consent, and the responsible handling of personal information.
Ensuring data security and obtaining informed consent become critical to address potential ethical dilemmas.

Bias in Data and Models:
If the training data used to develop machine learning models contains biases, the algorithms may perpetuate or even amplify those biases.
Biases in the data may lead to unfair predictions, especially when considering diverse populations of students.

Interpretable Black-Box Models:
Some advanced machine learning models, such as deep neural networks, operate as complex black-box systems, making it challenging to interpret how specific predictions are made.
Lack of interpretability can limit the understanding of the underlying factors contributing to mental health predictions.

Limited Generalization:
Machine learning models may be trained on specific datasets that do not fully represent the diversity of college student populations.
The resulting models might have limited generalization capabilities when applied to different demographics or institutions.

Overemphasis on Quantitative Data:
Machine learning models often rely on quantitative data, potentially overlooking qualitative aspects of mental health experiences that are essential for a comprehensive understanding.
Emotional nuances and subjective experiences may not be adequately captured through quantitative measures alone.

Informed Consent Challenges:
Obtaining informed consent for mental health data collection can be challenging, as individuals may be hesitant to disclose sensitive information, leading to potential selection bias in the dataset.

Resource Intensiveness:
Developing and implementing machine learning models requires significant computational resources and expertise.
Smaller institutions or research projects with limited resources may face challenges in adopting and maintaining such technology.

Dynamic Nature of Mental Health:
Mental health is inherently dynamic, and it may change over time due to various factors.
Machine learning models, if not regularly updated and adapted, may struggle to keep pace with the evolving mental health needs of college students.

Dependency on Self-Reported Data:
Many mental health studies rely on self-reported data, which may be subject to biases and inaccuracies.
Students may underreport or overreport their mental health symptoms, impacting the reliability of the collected information.

Stigmatization and Labeling:
The use of predictive models may inadvertently lead to stigmatization if individuals are labeled based on predictions.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Students identified as high-risk may face unintended consequences, such as increased scrutiny or limited opportunities.

V. CONCLUSION

The project on Anxiety, Depression, and Stress among college students using machine learning algorithms represents a significant step towards addressing the pressing mental health challenges faced by this demographic. The culmination of extensive data collection, preprocessing, and model development has provided valuable insights into the predictive capabilities of machine learning in identifying early signs of psychological distress. As we conclude this project, several key points emerge:

1. Understanding and Early Detection:

The utilization of machine learning algorithms has deepened our understanding of the intricate interplay between various demographic, academic, and lifestyle factors contributing to anxiety, depression, and stress among college students. The project highlights the potential for early detection, providing a proactive approach to supporting students at risk.

2. Predictive Power of Machine Learning:

The project demonstrates the predictive power of machine learning algorithms in assessing mental health outcomes. By leveraging diverse datasets and employing advanced algorithms, the models developed showcase the capability to identify patterns that might go unnoticed through traditional assessment methods.

3. Targeted Interventions and Support:

The predictive models offer the opportunity for universities and mental health professionals to implement targeted interventions and support services. This shift towards personalized assistance ensures that resources are efficiently allocated to those who need them the most, optimizing the impact of mental health initiatives on college campuses.

4. Ethical Considerations and Responsible Use:

Throughout the project, a keen awareness of ethical considerations has guided the collection and analysis of sensitive mental health data. The responsible use of machine learning models emphasizes the importance of privacy, consent, and the ethical handling of information, ensuring the well-being and trust of the student population.

5. Challenges and Considerations:

Acknowledging the advantages of machine learning, the project also underscores potential challenges, including biases, interpretability issues, and the dynamic nature of mental health. These challenges highlight the need for ongoing refinement and adaptation of models, along with a commitment to addressing ethical and societal considerations.

6. Contributions to Digital Mental Health:

The project contributes to the evolving field of digital mental health by showcasing the potential of technology-driven solutions in addressing mental health challenges among college students. This work serves as a foundation for future advancements and collaborations in the intersection of mental health and technology.

7. Call for Further Research:

As the project concludes, it becomes evident that the exploration of mental health using machine learning is an evolving and dynamic field. Further research is encouraged to refine models, improve interpretability, and address emerging challenges. Continuous collaboration between academia, healthcare professionals, and technology experts is essential for advancing the understanding and treatment of mental health issues.

In conclusion, the project signifies a crucial step towards leveraging technology for the betterment of mental health outcomes among college students. By embracing the potential of machine learning, this work contributes to a growing body of knowledge aimed at creating supportive, data-driven solutions that enhance the overall well-being of the student community.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

REFERENCES

[1] Sau, A., Bhakta, I. (2017)"Predicting anxiety and depression in elderly patients using machine learning technology."Healthcare Technology Letters 4 (6): 238-43.
[2] Tyshchenko, Y. (2018)"Depression and anxiety detection from blog posts data."Nature Precis. Sci., Inst. Comput. Sci., Univ. Tartu, Tartu, Estonia.
[3] https://adaa.org/understanding-anxiety/depression/symptoms
[4] https://www.webmd.com/balance/stress-management/stress-symptoms-effects_of-stress-on-the-body#1
[5] Oei, T. P., Sawang, S., Goh, Y. W., Mukhtar, F. (2013) "Using the depression anxiety stress scale 21 (DASS-21) across cultures." International Journal of Psychology 48 (6): 1018-1029.
[6] Kroenke, K., Spitzer, R. L.,Williams, J. B. (2001) "The PHQ-9: validity of a brief depression severity measure."Journal of general internalmedicine 16 (9): 606-613.
[7] Sau, A., Bhakta, I. (2018) "Screening of anxiety and depression among the seafarers using machine learning technology."Informatics in Medicine Unlocked :100149.
[8] Saha, B., Nguyen, T., Phung, D., Venkatesh, S. (2016) "A framework for classifying online mental health-related communities with an interest in depression."IEEE journal of biomedical and health informatics 20 (4): 1008-1015.
[9] Reece, A. G., Reagan, A. J., Lix, K. L. M., Dodds, P. S., Danforth, C. M., Langer, E. J.(2016) "Forecasting the Onset and Course of Mental Illness with Twitter Data."Scientific reports 7 (1): 13006.
[10] Braithwaite, S. R., Giraud-Carrier, C., West, J., Barnes, M. D., Hanson, C.L. (2016) "Validating machine learning algorithms for Twitter data against established measures of suicidality."JMIR mental health 3 (2): e21.
[11] Du, J., Zhang, Y., Luo, J., Jia, Y., Wei, Q., Tao, C., Xu, H. (2018) "Extracting psychiatric stressors for suicide from social media using deep learning."BMC medical informatics and decision making 18 (2): 43.
[12] Al Hanai, T., Ghassemi, M. M., Glass, J.R. (2018) "Detecting Depression with Audio/Text Sequence Modeling of Interviews." InInterspeech : 1716-1720.
[13] B. Singh, "Retina Recognition Using Support Vector Machine", International Journal of Engineering and Computer Science, Vol 5, No. 8, pp. 17605-17608, 2016.
[14] R. Mehta and K. Egiazarian, "Dominant Rotated Local Binary Patterns (DRLBP) for texture classification", Pattern Recognition Letters, Vol. 71, No. 3, pp. 16-22, 2016.
[15] K. Addakiri and M. Bahaj, "On-line Handwritten Arabic Character Recognition using Artificial Neural Network", International Journal of Computer Applications, Vol. 55, No. 13, pp. 42-46, 2012.
[16] V. Karbhari, D. Prapti, V. Sriniwas, and M. Majharoddin, "Zernike Moment Feature Extraction for Handwritten Devanagari (Marathi) Compound Character Recognition", International Journal of Advanced Research in Artificial Intelligence, Vol. 3, No. 1, pp. 68-76, 2014.