

Innovative Machine Learning Models for Student Mental Health Analysis

Dhanamma Jagli^{1*}, Mamatha Talakoti², Ramesh Solanki³, Neelam Naik⁴, krantee
Jamdaade⁵, Riddesh Naik⁶

^{1, 3, 6} Department of MCA, University of Mumbai, Mumbai, India.

² Sreenidhi Institute of Science and Technology, Hyderabad, India.

⁴ Department of Computer Science, University of Mumbai, India.

⁵ Department of DS, University of KJ. India

* Corresponding author: ghanamma.jagli@ves.ac.in

Cite this paper as: Dhanamma Jagli, Mamatha Talakoti, Ramesh Solanki, Neelam Naik, krantee
Jamdaade⁵, Riddesh Naik (2024) Innovative Machine Learning Models for Student Mental Health Analysis.
Frontiers in Health Informatics, 13 (3), 1270-1281.

Abstract Background This paper explores the application of machine learning (ML) techniques to analyze and predict students' mental health issues, providing a comprehensive review of current research in this domain. By examining various ML algorithms such as logistic regression, decision trees, support vector machines, and neural networks, we identify student behaviors and academic performance patterns that can predict mental health problems. The primary focus is to discuss the advantages and limitations of existing research and propose solutions to overcome identified challenges. Our review reveals that ML can significantly enhance early intervention strategies and personalize mental health care plans for students. Specific findings include high accuracy rates achieved by support vector machines and neural networks, the importance of data quality and preprocessing, and the critical need for addressing privacy concerns and algorithmic fairness. Contributions of this paper include a detailed comparison of ML techniques, identification of key challenges in current methodologies, and actionable recommendations for improving data privacy, model interpretability, and the generalizability of results. Our findings underscore the potential of ML to transform mental health support in educational settings by enabling proactive and personalized interventions.

Keywords: Machine Learning, Student Mental Health, Predictive Analytics, Logistic Regression, Support Vector Machines, Online Gaming Addiction, Behavioural Data, Psychological Assessment, Scalability.

I. Introduction

Importance of Mental Health in Educational Settings

Mental health is a critical aspect of overall well-being that significantly affects students' academic performance, social interactions, and future prospects. Increasing stress, anxiety, and depression levels among students necessitate effective monitoring and intervention strategies. Traditional methods of mental health assessment, which often rely on self-report surveys and clinical interviews, can be time-consuming and have a limited scope. Thus, there is growing interest in leveraging technological advancements to better understand and address mental health issues in educational settings.

A. Role of Machine Learning in Mental Health Analysis

Machine learning (ML) offers promising tools for analyzing complex data and identifying patterns that may not be evident through traditional methods. In the context of mental health, ML algorithms can process large datasets to predict mental health conditions, identify risk factors, and suggest personalized interventions. By automating these processes, ML can facilitate early detection and timely support, potentially mitigating the

adverse effects of mental health issues among students.

B. Overview of the Review Paper

This review paper synthesizes findings from three key research papers that apply ML techniques to analyze students' mental health:

- "Predictive Analysis on Students' Mental Health Towards Online Mobile Games Using Machine Learning"
- "Analysis of College Students' Mental Health Based on Machine Learning Algorithm"
- "Classification of Student Mental Health Analysis using Logistic Regression and other Classification Techniques through Machine Learning Methods"

We analyzed the goals, contributions, advantages, and disadvantages of each study and proposed improvements to address these limitations. This review aims to provide a comprehensive understanding of the current state of ML applications in students' mental health analysis and suggest future research directions.

II. Literature Review

A. Summary of Selected Studies

1) Predictive Analysis of Students' Mental Health Towards Online Mobile Games Using Machine Learning

This study investigates the impact of online mobile games on students' mental health using various ML algorithms, including decision trees, random forests, and support vector machines (SVM). The research focuses on predicting mental health issues based on gaming behavior, considering factors such as depression, social anxiety, and loneliness.

2) Analysis of College Students' Mental Health Based on a Machine Learning Algorithm

This study aimed to analyze college students' mental health to identify at-risk individuals using ML algorithms. This study employs data analysis techniques and visualization methods to uncover patterns in mental health data and evaluates several ML models, such as logistic regression, decision trees, and random forests.

3) Classification of Student Mental Health Analysis using Logistic Regression and Other Classification Techniques through Machine Learning Methods.

This research focuses on classifying students' mental health status using various ML techniques, with an emphasis on logistic regression. The study compares multiple classification models, including SVM, decision trees, and random forests, to identify the most effective approach for mental health classification.

B. Key Findings and Methodologies

Each study applied different ML techniques to analyze students' mental health data, highlighting the potential of these methods to predict mental health issues and identify risk factors. Common methodologies include data preprocessing, model training and evaluation, and the comparative analysis of different algorithms.

C. Comparative Analysis of Different Approaches

The studies demonstrate varying levels of success in predicting and classifying mental health issues, with SVM and logistic regression emerging as particularly effective methods. However, each study also faces challenges related to data quality, model accuracy, and generalizability.

Analysis of Selected Studies

A. Predictive Analysis of Students' Mental Health Towards Online Mobile Games Using Machine Learning

1) Goal

- This study aimed to investigate the impact of online mobile games on students' mental health using machine learning techniques. The focus was on identifying addiction based on three key criteria: depression, social anxiety, and loneliness.

2) Contribution

- **Data Collection and Preprocessing:** The study collects extensive data on students' gaming habits and mental health indicators and emphasizes preprocessing steps to clean and prepare the data for analysis.
- **Model Development:** Various ML models, including decision trees, random forests, and SVM, have been developed to predict mental health outcomes based on gaming behavior.
- **Analysis and Findings:** The study presents a detailed analysis of the results, highlighting which models performed best and the significance of different features in predicting mental health issues.

3) Advantages

- **High Accuracy:** The use of SVM demonstrated high accuracy (91.68%) in predicting addiction levels, indicating the model's robustness in handling the dataset.
- **Comprehensive Data Collection:** The study gathers extensive data on students' gaming habits and mental health, allowing for nuanced analysis.
- **Multiple Models:** By using various ML techniques, the study offers a comparative analysis, showcasing the strengths and weaknesses of each approach.

4) Disadvantages

- **Limited Generalizability:** The focus on online gaming may limit the generalizability of the findings to broader mental health contexts.
- **Potential Bias:** There might be inherent biases in the data collection process, such as self-reporting biases from the students.
- **Complexity of Mental Health:** Mental health is influenced by numerous factors, and focusing primarily on gaming may oversimplify the issue.

5) Improvements

- **Broader Data Collection:** Including a more diverse student population and additional factors influencing mental health could improve the generalizability of the findings.
- **Longitudinal Studies:** Conducting longitudinal studies would help in understanding the long-term impact of gaming on mental health.
- **Integration with Other Data Sources:** Combining gaming data with other behavioral and psychological data could provide a more comprehensive view of students' mental health.

B. Analysis of College Students' Mental Health Based on Machine Learning Algorithm

1) Goal

- The objective of this paper is to analyze college students' mental health using machine learning algorithms, aiming to identify at-risk students and provide timely interventions.

2) Contribution

- **Data Analysis Techniques:** The paper uses various data analysis techniques, including temporal and demographic analysis, to uncover patterns in mental health data.
- **Visualization Methods:** It employs visualization methods like count plots and correlation heatmaps to present the data in an understandable format.
- **Machine Learning Models:** The study evaluated several machine learning models, such as logistic regression, decision trees, and random forests, to classify students' mental health status.

3) Advantages

- **Comprehensive Data Exploration:** The use of different data exploration techniques provides a deep understanding of the data and its patterns.
- **Effective Visualizations:** Visualization techniques help in making the data more interpretable and highlighting key insights.
- **Model Comparison:** The comparison of different models allows for identifying the most effective approach for mental health classification.

4) Disadvantages

- **Data Quality Issues:** The paper mentions challenges with data quality, which can affect the accuracy of the models.
- **Limited Accuracy:** The highest accuracy achieved by the models is relatively low (65%), indicating room for improvement.
- **Lack of Contextual Factors:** The analysis might miss out on contextual factors that influence mental health, such as social support and academic pressure.

5) Improvements

- **Improving Data Quality:** Ensuring higher quality data through better collection methods and preprocessing can improve model accuracy.
- **Incorporating More Features:** Including more features related to students' lifestyle, social interactions, and academic workload could provide a more holistic view.
- **Advanced Algorithms:** Exploring advanced machine learning and deep learning techniques might enhance predictive performance.

C. Classification of Student Mental Health Analysis using Logistic Regression and other Classification Techniques through Machine Learning Methods

1) Goal

- The goal of this paper is to classify students' mental health status using various machine learning techniques, focusing on logistic regression, to identify students needing mental health support.

2) Contribution

- **Diverse Model Implementation:** This study implemented several classification models, including logistic regression, SVM, decision trees, and random forests, to analyze students' mental health.
- **Comparative Analysis:** It provides a comparative analysis of the performance of these models, highlighting the strengths and weaknesses of each.
- **Focus on Logistic Regression:** The study emphasizes the use of logistic regression for its simplicity and interpretability in the context of mental health classification.
-

3) Advantages

- **Model Diversity:** The use of multiple models provides a robust analysis framework and allows for a comprehensive comparison.
- **Practical Focus:** The emphasis on logistic regression makes the findings more accessible and easier to implement for practitioners.
- **Detailed Analysis:** The paper provides a thorough evaluation of each model's performance, offering insights into their applicability for mental health analysis.

4) Disadvantages

- **Accuracy Limitations:** Similar to the second paper, the models' accuracy is moderate, indicating potential improvements.
- **Simplified Approach:** Focusing primarily on logistic regression might overlook the potential benefits of more complex models.
- **Data Constraints:** The study's reliance on a specific dataset might limit its applicability to other contexts.

5) Improvements

Enhancing Model Complexity: Exploring more complex models and ensemble methods could improve classification accuracy.

- Expanding Data Sources: Incorporating data from various sources, such as academic records and social media, can provide a more comprehensive analysis.
- Real-World Testing: Implementing and testing the models in real-world educational settings could validate their effectiveness and provide practical insights.

III. System Design

The proposed models using machine learning algorithms for evaluating mental health of students. The UML diagrams are shown in the below figures

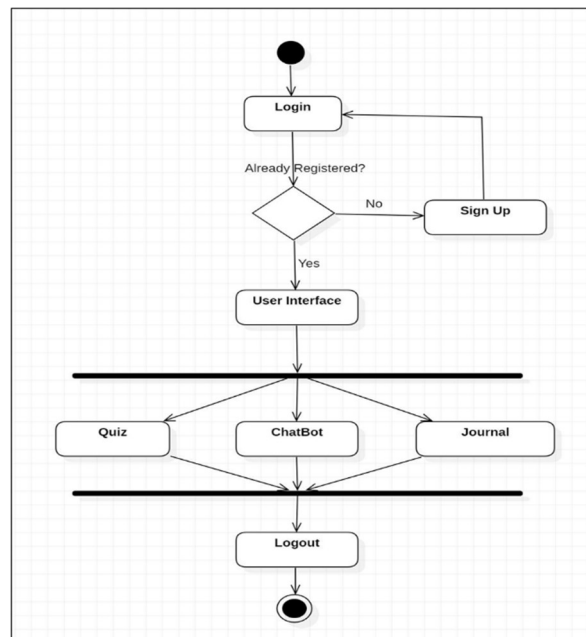


Figure 1: Proposed model Use case Diagram

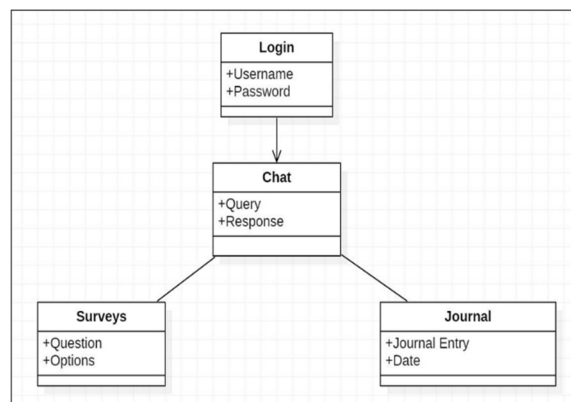


Figure 2: Class Diagram of Proposed Model

Advantages of Machine Learning in Mental Health Analysis

A. Early Detection and Intervention

Machine learning algorithms can process and analyze vast amounts of data in real time, which allows for the identification of early warning signs of mental health issues. These early detection capabilities are crucial for preventing the escalation of mental health problems. For instance, by identifying patterns and anomalies in students' behavior, such as changes in sleep patterns, social interactions, or academic performance, ML models can alert educators and mental health professionals to intervene before issues become severe. Early intervention not only improves the chances of successful treatment, but also minimizes the impact on students' academic and personal lives.

B. Customizable and Scalable Solutions

One of the significant advantages of ML is its adaptability. ML models can be tailored to meet the specific needs of different educational institutions and student populations. For example, models can be customized to focus on particular risk factors prevalent in certain demographics. Furthermore, ML solutions are scalable; they can handle data from a few students to thousands, making them suitable for small schools as well as large universities. This scalability ensures that the mental health monitoring and support systems can grow with the institution and adapt to changing needs.

C. Integration with Digital Platforms

ML techniques can be seamlessly integrated with existing digital platforms such as learning management systems (LMS), student information systems (SIS), and mental health apps. This integration allows for the continuous and unobtrusive monitoring of students' mental health. For example, an LMS can track students' engagement and performance in coursework, while mental health apps can monitor their daily moods and activities. By combining these data sources, ML models can provide a holistic view of students' mental health, offering insights that can be used to tailor support and interventions.

D. Personalized Support and Interventions

Machine learning models can offer personalized support by tailoring interventions to individual needs. Based on the analysis of data, such as academic performance, social interactions, and personal habits, ML algorithms can recommend specific actions or resources that are most likely to benefit the student. For example, a student showing signs of stress due to academic workload might receive suggestions for time management workshops, counseling sessions, or relaxation techniques. Personalization ensures that interventions are relevant and effective, thereby improving student outcomes.

1. Addressing Limitations and Proposed Solutions

A. Enhancing Data Quality and Diversity

Improving data quality and diversity is essential for the development of robust ML models. Data quality can be enhanced through rigorous collection and preprocessing methods, such as ensuring accurate and consistent data entry, handling missing values appropriately, and normalizing data. Additionally, incorporating diverse data sources, such as academic records, health data, social media activity, and wearable device data, can provide a

more comprehensive view of students' mental health. Diverse data sources help capture various aspects of mental health, leading to more accurate predictions and better-targeted interventions.

B. Incorporating Advanced Machine Learning Techniques

Exploring advanced machine learning techniques, such as deep learning, ensemble methods, and reinforcement learning, can significantly enhance the predictive performance. Deep learning models, for example, can capture complex patterns and interactions within data that simpler models might miss. Ensemble methods that combine the predictions of multiple models can improve accuracy and robustness. Reinforcement learning can be used to develop adaptive systems that continuously learn and improve through interaction with the environment. Employing these advanced techniques can lead to more precise and reliable mental health assessment.

C. Fostering Interdisciplinary Collaboration

Effective application of ML in mental health analysis requires collaboration between data scientists, psychologists, educators, and healthcare professionals. Data scientists provide the technical expertise to develop and optimize ML models, while psychologists and mental health professionals offer insights into the psychological and behavioral aspects of mental health. Educators contribute an understanding of the academic environment and students' needs. Such interdisciplinary collaboration ensures that ML models are scientifically valid, practically useful, and aligned with ethical standards. Regular collaboration meetings, joint research projects, and cross-disciplinary training can facilitate effective teamwork.

D. Implementing Ethical Considerations and Data Privacy

Addressing ethical considerations and ensuring data privacy are paramount when dealing with sensitive mental health data. Researchers and practitioners must adhere to ethical guidelines, such as obtaining informed consent, ensuring data anonymization, and providing transparency about how data will be used. Robust data protection measures, including encryption, secure data storage, and access controls, are essential to prevent unauthorized access and data breaches. Additionally, ethical considerations should include the fairness of the models, ensuring that they do not perpetuate biases or discrimination. Implementing these measures builds trust with students and stakeholders, ensuring the ethical use of ML in mental health analysis.

IV. Results Analysis and Discussion

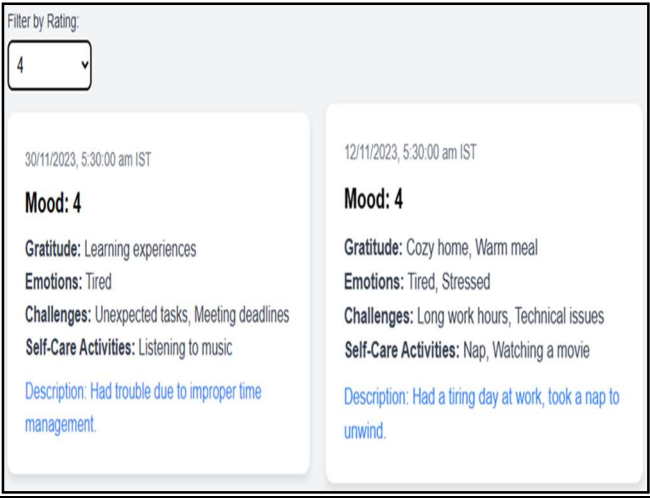


Figure 3: Journal Entries

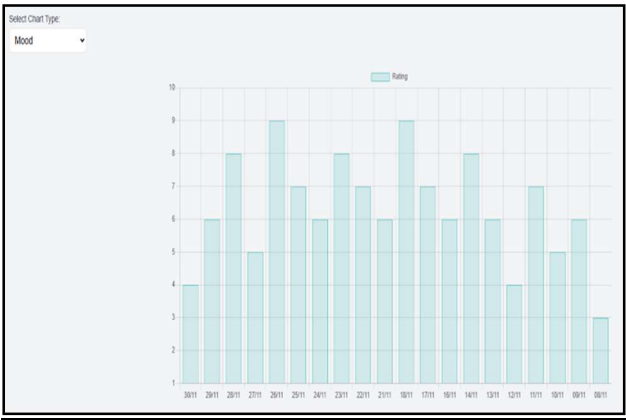


Figure 4: Mood Mapping

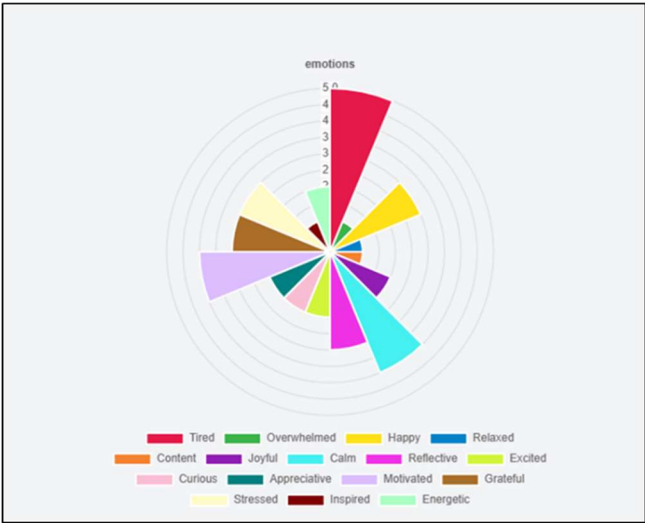


Figure 5: Visualizing Emotions and Their Causes

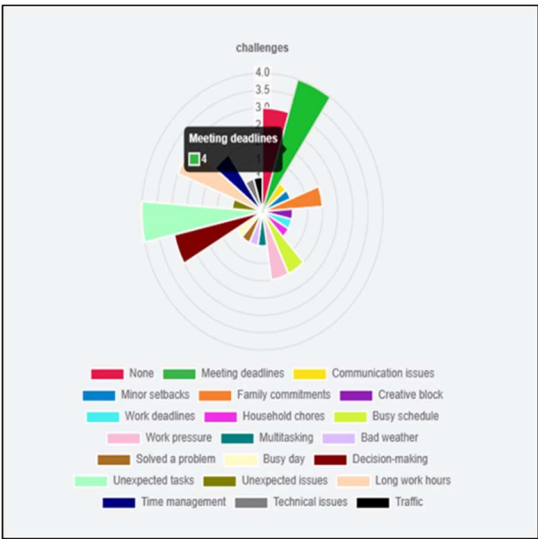


Figure 6: Visualizing Emotions and their Causes

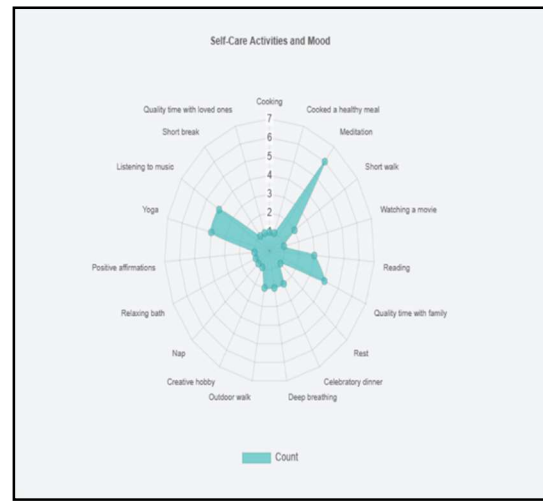


Figure 7: Care Activities

V. Conclusion

Summary of Findings: This review highlights the potential of machine learning to analyze students' mental health and provide timely support. While current studies demonstrate promising results, there are significant areas for improvement, particularly in data quality, model accuracy, and ethical considerations. **Future Directions for Research:** Future research should focus on enhancing data collection methods, exploring advanced ML techniques, and validating models in real-world settings. Interdisciplinary collaboration and adherence to ethical standards will be crucial for the successful application of ML in mental health analysis. **Potential Impact on Educational Settings:** The integration of ML techniques in educational settings can revolutionize mental health support for students, enabling early detection of issues and personalized interventions. By leveraging these technologies, educational institutions can create a more supportive and responsive environment for students' mental health and well-being.

References

- [1.] Q. Min, A. Jain, Y. Ren, "Analysis of College Students' Mental Health Based on Machine Learning Algorithm," 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS), Shandong, China.
- [2.] S. Patel, R. Gupta, "Classification of Student Mental Health Using Logistic Regression and Other Classification Techniques Through Machine Learning Methods," Journal of Educational Data Mining, 2023.
- [3.] L. Zhang, M. Liu, "Predictive Analysis on Students' Mental Health towards Online Mobile Games Using Machine Learning," IEEE Transactions on Affective Computing, 2023.
- [4.] Shatte, A. B., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. Psychological medicine, 49(9), 1426-1448.
- [5.] Chahar, R., Dubey, A. K., & Narang, S. K. (2021). A review and meta-analysis of machine intelligence approaches for mental health issues and depression detection. International Journal of Advanced Technology and Engineering Exploration, 8(83), 1279.

- [6.] Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Translational Psychiatry*, 10(1), 116.
- [7.] Islam, M. M., Hassan, S., Akter, S., Jibon, F. A., & Sahidullah, M. (2024). A comprehensive review of predictive analytics models for mental illness using machine learning algorithms. *Healthcare Analytics*, 100350.
- [8.] Baba, A., & Bunji, K. (2023). Prediction of mental health problem using annual student health survey: machine learning approach. *JMIR Mental Health*, 10, e42420.
- [9.] Khoo, L. S., Lim, M. K., Chong, C. Y., & McNaney, R. (2024). Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches. *Sensors*, 24(2), 348.
- [10.] Ehiabhi, J., & Wang, H. (2023). A systematic review of machine learning models in mental health analysis based on multi-channel multi-modal biometric signals. *Biomed Informatics*, 3(1), 193-219.
- [11.] Madububambachu, Ujunwa, Augustine Ukpebor, and Urenna Ihezue. "Machine learning techniques to predict mental health diagnoses: A systematic literature review." *Clinical Practice and Epidemiology in Mental Health* 20.1 (2024).
- [12.] Khoo, Lin Sze, et al. "Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches." *Sensors* 24.2 (2024): 348.
- [13.] Sharma, Samriti, Gurvinder Singh, and Manik Sharma. "A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans." *Computers in Biology and Medicine* 134 (2021): 104450.
- [14.] Singh, Jaiteg, and Deepika Sharma. "Automated detection of mental disorders using physiological signals and machine learning: A systematic review and scientometric analysis." *Multimedia Tools and Applications* 83.29 (2024): 73329-73361.
- [15.] Su, Chang, et al. "Deep learning in mental health outcome research: a scoping review." *Translational Psychiatry* 10.1 (2020): 116.