INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND ANALYSIS

ISSN(print): 2643-9840, ISSN(online): 2643-9875

Volume 08 Issue 09 September 2025

DOI: 10.47191/ijmra/v8-i09-29, Impact Factor: 8.266

Page No. 5168-5181

Predictive Modeling of Mental Health Disorders: AI Applications in Biomarkers and Behavioral Data

AbdulBasir Momand[1], Sohaib Ahmad Khalil[2], Mohsin Mahmood[3]

[1]Rana University, Afghanistan

[2]Abasyn University, Peshawar

[3]City University, Peshawar

ABSTRACT: Mental health disorders represent a growing global health challenge, often characterized by complex, multifactorial symptoms that complicate timely diagnosis and treatment. Advances in artificial intelligence (AI), particularly in machine learning and data-driven analytics, offer promising tools for early detection and personalized interventions. This study explores predictive modeling approaches that leverage both biological markers (biomarkers) and behavioral data to enhance diagnostic accuracy for mental health conditions such as depression, anxiety, and bipolar disorder. We investigate the integration of physiological signals (e.g., EEG, heart rate variability) with behavioral indicators (e.g., speech patterns, social activity, digital footprints) using supervised learning models. The results indicate that combining multimodal datasets significantly improves the performance of AI models in classifying mental health states, achieving improved precision and sensitivity across diverse patient profiles. Our findings underscore the potential of AI-enhanced frameworks to support clinicians in objective diagnosis, individualized care planning, and early intervention strategies. Future work will focus on longitudinal validation and ethical deployment in real-world healthcare settings.

KEYWORDS: AI-based Models, Predicting, Mental Health Disorders, Biomarkers, Social Media Data, Behavioral Patterns.

1. INTRODUCTION

The current crisis in mental health costs billions of dollars in lost earnings and productivity, in addition to the immeasurable burden on the well-being of individuals and their families. To treat mental health effectively and promptly, it is necessary to identify conditions as early as possible and use the best treatment options. (Knapp & Wong, 2020). There are significant barriers that prevent those who suffer from mental illness from receiving proper care. The combination of the social stigma associated with mental illness, the high cost of counseling, the shortage of therapists and professionals, the poor quality of services and treatments for the mentally ill, and the lack of availability of professional mental health services increases the time and risk necessary to treat and cure a patient diagnosed with a mental illness. (Aguirre et al.2020). The principles of physical health can be extended to mental health, enabling the integration of appropriate structured diagnostic and clinical decision-making tools to include machine learning-based algorithms as a means to phenotypically and biologically classify mental illness. The decision-supporting system case shows that the newly developed prediction architecture is able to effectively differentiate subjects. (Rashid & Calhoun, 2020).

1.1 Background of the study

Understanding the early signs of mental health is crucial to improving the quality of life for people all over the world. With increasing digital media usage habits, several studies have been conducted to predict early signs of different types of disorders using the digital habits of a person as a parameter. (Santamaría-García et al.2020). Digital data, social media activities, videos, audios, and other behavioral elements provide a variety of data with which a person's mental health can be predicted quite accurately. This work focuses on available studies that help in providing a comprehensive review of AI-based models and theoretical frameworks that use digital markers to predict different types of mental health disorders, including but not limited to anxiety, depression, schizophrenia, pain, fear, attention deficit hyperactivity disorder, acute stress, and post-traumatic stress disorder. (Wies et al., 2021). Two prediction techniques, such as fusion models and transfer learning models, are also reviewed in our work. We have also explored different types of emotions that play a crucial role in determining mental health disorders, such as sadness, anger, anxiety, valence, arousal, and dejection, with related work and modeling techniques. (Breaux et al.2021).

Predictive Modeling of Mental Health Disorders: AI Applications in Biomarkers and Behavioral Data

We are currently experiencing a remarkable digital era during which enormous amounts of data are generated every single second. This vast and ever-growing data can take on various forms, such as biodata including personal identification details, complex medical data containing patient histories, and an array of social media data reflecting our interactions online. (Deepa et al.2022). Additionally, there exists data that is generated effortlessly by users through their smartphones, activity trackers, and various other digital devices and applications. Recent studies show that the digital data of a person will have a significant impact on their mental health. (Odgers and Jensen2020). Using activity trackers, people can track their behavior, including physical activities, sleep duration, time in front of the screen, and so on. Using smartphones, people can track the use of different apps, such as social media apps, calls, shopping, and so on. (Vuorre et al.2021). Increments in the usage of such apps show that the person may have some mental health issues, including stress and depression. Researchers also believe that there is a relationship between people's emotional states and their calls and smartphone use. (Gianfredi et al., 2021).

1.2 Problem statement

Mental health disorders have been growing as a major concern worldwide through the years. One of the problems of mental health disorders is the difficulty in tracking them and engaging patients, especially due to their episodic nature. (2019 Mental Disorders Collaborators, 2022). Nowadays, it is a very expensive and time-consuming process to diagnose, which is done by psychiatrists analyzing patient and relative reports. The use of artificial intelligence models to predict mental health disorders carries significant potential to overcome current limitations in forecasting these diseases through the analysis of social media data and behavioral patterns. (Lee et al.2021). It could also make it possible to assess the patient's condition continuously, not only over time but also in different situations and settings, and to evaluate patients at different stages of the same disease or those who suffer from different illnesses at similar moments. (Karunarathna et al., 2024). The early detection of mental health issues could lead to illness management based on real-time interventions that can be more effective than the current treatment models; most psychiatric patients have seasonal appointments with their psychiatrist instead of a continuous follow-up throughout the year due to human resource limitations. (Galea et al., 2020).

1.3 Research questions

Advancements in AI-based predictive modelling methodologies revealed opportunities for implementing large-scale studies for predicting mental health disorders using a broad array of sources including genetic, brain imaging biomarkers, physiological responses, speech signatures, unstructured text data, social media information, electronic health records, digital footprints, and wearable device records. However, various challenges include developing prospective models utilizing varied datasets collected over a long period to address predictive and clinical validity issues. (Koutsouleris et al.2022). There are also data privacy and ethical concerns associated with using certain data sources. As researchers start to tackle related challenges. In this text, we aim to help researchers from different domains develop predictive AI models for mental disorders, make informed decisions based on loopholes in the related literature, and enhance future studies. In summary, the predominant and challenging goal of this field is to develop predictive models for various disorders using large and diverse datasets collected over a long period, with good clinical validity, and meaningful predictive power. (Sui et al., 2020). These models could complement or replace the clinical diagnostic process or result in novel insights and diagnostic strategies, including treatment decisions. A related effort is to determine whether real-world and wearable data could advance these models. The prospective models are much more difficult to develop, as the same person's various data sources must be combined and maintained over a period of time. (Nahavandi et al.2022). In light of recent advancements in this field, this work aims to answer the following primary questions:

A. What are the different types of mental health disorders that can be predicted by digital data?
B. What is the range of mental health disorders that can be predicted using digital markers?
C. What are the available studies to predict mental disorders using digital markers for each disorder?
D. What type of emotion can predict the different types of mental health disorders?
E. What are the available modeling techniques used by previous studies to predict mental health disorders?

1.4 Significance of the study

Studying various psychological and mental health disorders in their early stages can be challenging due to the unavailability of an easy, scalable diagnostic system. Early diagnosis of attention-deficit/hyperactivity disorder, suicidality, major depressive disorder, bipolar disorder, and schizophrenia is essential to control their progression and social burden. Human blood contains valuable information that can be used to identify or rule out health problems. (Colizzi et al.2020). There are different medical tests available that doctors use to monitor mental health disorders. However, the frequent possibility of negative outcomes for patients and expensive tests limits the scalability and applicability of these tests in primary care. Development of cheaper, less intrusive, and scalable early diagnostic systems for these disorders is essential. It provides an opportunity for increased awareness about the disorders on the part of the involved stakeholders, who might access existing medical and psychological care systems. (Torous et al.2021). Social media data provide easy, safe, and scalable access to individuals' day-to-day behavior and experiences. Mental health conditions are known to be risk factors for attempted or completed suicides. There are only two previously published studies that used social media data to predict attempted or completed suicides. (Chancellor & De Choudhury, 2020). They both took surveillance-based approaches to assess social media self-disclosures and preexisting medical records or claims data as a positive or negative indicator of the self-enhanced suicides. Their prediction systems achieved fairly good postdictive performance on web content of psychiatric hospitalizations and incident suicide attempts. Unlike prior researchers, we aim at predicting the suicidality of social media users long before they create an alarming self-disclosure message. In order to do this, we studied the behavior of individuals on social media networks weeks before their first incident diagnosed as non-fatal self-harm, suicide attempt, or suicide completion. (Drouin et al.2020).

1.5 Scope and limitations

This work mainly focuses on the utilization of AI-based models for predicting mental health disorders using various types of biomarkers, including but not limited to social media data and behavioral patterns, presenting the effectiveness of these models according to various factors such as the sample size. Furthermore, applications of these models in research and practice related to mental health and illness management are also introduced and discussed. However, due to the overlap issue introduced in the previous section, for those special topics that will be covered, we may combine with the introduction part only to avoid repetitive topics. (Tutun et al.2023). Then, the biomarkers issue will also be introduced. Based on the related work, social media data is surely part of these biomarker features. Therefore, related work will be included in the next part without the need to define feedback and measure again. (Abi-Dargham et al.2023). Additionally, unlike the related work, the main emphasis of our work is put on how we can use AI-based models to help researchers and medical professionals who are not so familiar with AI know which AI-based methods they can employ to achieve their similar objectives according to the desired results. This work also serves as an initial guide for educational purposes on the AI-based application in the review-annotated research. (Koutsouleris et al.2022). This guide aims to outline the potential applications of AI in mental health research while also addressing the ethical considerations and limitations inherent in utilizing these technologies. Furthermore, it highlights the need for cautious interpretation of AI-driven predictions, particularly in cases where data quality and sources may vary significantly. (Albahri et al.2023). Other additional benefits will also be discussed in the conclusions. Moreover, a limitation of the proposed idea is that only the following six popular biomarker types of features in medical applications are discussed in this exploratory study for concreteness:

1.5.1 Demographic

Identifying depression and other mental health disorders at early stages can positively impact mental well-being. Given the prevalence of mental health conditions, relying on traditional methods for categorizing people into healthy and unhealthy groups is not practical. Predicting mental health disorders remains challenging due to the complex nature of social media behaviors and limitations of existing models. In this paper, we discuss the demographic characteristics of participants in the MHDW Challenge (Meehan et al.2022). The presented demographic statistics not only can help identify biases in the utilized datasets but also provide insights into understanding the population dominance of a given social media platform. In this section, we describe the dataset and the sample at a high level. We then provide detailed statistics of the participants' attributes. We also examine the research questions and research goals, and describe the empirical analysis used to investigate our research questions. Based on the analysis, we discuss the key insights. Finally, we discuss our study's limitations, note future directions, and provide a high-level overview of the paper.

1.5.2 Physiological health

Physiological health is a measure of the overall condition of an organism at a given time. It is an important feature of an individual and is closely related to their psychological state. This affects how well body organs and cells operate, especially in the central nervous system. Physiological measures include heart rate, heart rate variability, and skin conductance response, which have been proven effective in predicting mental state (Mansi et al.2021). Many studies have made successful use of these measures in predicting psychological distress, excitement, sleep quality, psychological stress, and physical activity. Furthermore, some studies that focus on children and adolescents even base their measures on such physiological signals to fully understand their mental states (Chung and Teo2022). Obviously, we are no longer in the era when we are only able to use closed-ended rating scales for examining mental problems. These self-reports, while valuable, are sometimes too subjective and unreliable, while the physiological measures are more reliable. As a result, people have started to treat physiological data as second only to individuals' subjective feelings.

1.5.3 Social media

Social media has emerged as a cardinal component of an individual's daily routine. A vast number of people spend hours each day on these platforms, seeking opportunities to interact with others and share their personal experiences. The premise behind this technology is that the more information we know about one another, the more we can understand and help one another. Given that one-third of our lives is spent at work and much of our social interactions are quite visible on social media, the information available on social media can capture our life and personality quite comprehensively (Moisescu et al., 2022). Through this comprehensive view, work activity and social interaction on social media can be used to predict and assess individuals for physical and mental health. Not only can our activity on social media be used as behavior markers to reflect our health state, but through our use of personal content and communication with others on social media, we may also share information about our physical and mental health with others. This can help overcome the barriers to seeking help and treatment that many face. A supportive online community can grant patients access to individuals willing to listen and assist during hard times and may hold the potential for a unique ability to provide personalized recommendations or interventions because a large number of users can be easily influenced at the same time (Zhong et al., 2021).

On the other hand, it is easy to overinterpret online behavior that is more indicative of superficial behavior patterns over time, making it a harder task to truly understand the actual psychological state of a specific individual. Moreover, many of these profile-based attributes are self-reported, which narrows the scope and potential insights that limited data can afford. Ideally, profiles are based on historical data that long predate the clinical onset of mental illness; however, it is challenging to identify these behaviors in profiles, particularly considering that users can interact with social media in wildly different ways depending on their individual predictors (Brauer et al., 2024). Our general population is composed of thousands of patient cohorts, each with unique responses to a variety of diseases and dysfunctions. Consequently, it is challenging to reconcile the complexity of our patient ecosystem when detecting and accurately treating mental health conditions across this diverse social media landscape. (Zhang et al.2021).

1.5.4 Activity engagement

Correct mental health or prevention of mental health is the process of maintaining and enhancing good mental health and well-being. It encompasses different key components including emotional resilience and problem solving. In the current study, we utilize the term activity engagement to indicate the applied activities in which individuals engage and/or initiate, governed by personal interest, talents, and hobbies with the intention to improve mental wellness and well-being. There is much research suggesting that positive emotions and activities could be useful to assist people in managing their mental health, strong well-being, and assist society in combating some of the impacts of COVID-19 (Waters et al.2022). However, describing helpful and unhelpful activities for corporate mental health well-being has mainly elided those types of complexity in favor of broad generalizable research. Moreover, the discrepancy of clinical and professional factors has resulted in staff's ability to understand and assist employees with issues regarding mental illness, stress, or overall mental well-being being suboptimal. Notably, the mental health literacy of help-seeking staff predicted their self-efficacy and likelihood to assist employees struggling with mental health challenges, pointing to the importance of helping staff improve their strategies for work-related prevention and intervention. The application of activity engagement to encourage employees to find activities that enrich their mind and body and foster a culture of mental wellness is sorely needed (Moss et al.2022).

1.5.5 Contextual

The global mental health care crisis is caused not only by a lack of resources but by a lack of understanding of how to treat patients effectively. Although their etiology is still widely unclear, many psychiatric disorders are characterized by hypo- or hyper-connective brain patterns. Artificial intelligence techniques have gained significant ground in different areas of healthcare, including the mental health domain, and have indeed demonstrated the capacity to predict symptom severity of these disorders through models focused on patterns of brain connectivity and processing multimodal data. This text extends the previous work and presents multimodal predictive models that incorporate not only neurobiological data but also behavioral outputs and social media data, reaching significant accuracy from the combination of multimodal data sources (Nemesure et al., 2021). Autonomic nervous system data and the patterns of human interactions can inform about both the structural and functional state of the brain. Besides established attribute-symptom links, the quantity of reported communication within the social platforms, regarded as a pattern of social activity, with clinically relevant topics can provide markers about mental health conditions. Our aim is to create a platform to interactively promote insights among people at large and predict mental health symptom scores. The creation of a chatbot-based solution allows for a close and guided interaction of users with specific interventions and individual feedback sessions. The solution is made as a SaaS solution commercially available (Ríssola et al.2021).

2. LITERATURE REVIEW

Mental health disorders are prevalent worldwide and cause substantial personal suffering and economic burden. Risk prediction using prognostic modeling can be incorporated at the early screening stage, paving the way to new research solutions for preventing and treating mental disorders as well as contributing to savings in medical costs. (Herrman et al.2022). An AI-based early risk-prediction model accelerates screening initiatives and health policy-making. This paper aims to provide a comprehensive overview of AI-based risk-prediction models for mental health. With the perspective of clinical research, we summarize current progress and discuss ongoing and future directions, motivated by the fact that no systematic summary exists in the current literature. (Health Organization, 2022). The search engines for all relevant articles were used, with the following means of search: ("mental health" OR "mental disorder" OR "psychiatric disorder" OR "psychiatric diagnosis") AND ("risk prediction" OR "early detection" OR "early diagnosis" OR "prognostic model" OR "prediction model" OR "machine learning" OR "artificial intelligence") AND ("biomarkers" OR "social media" OR "behavioral patterns" OR "deep phenotyping"). The related articles were also collected manually. The focus of the present survey is on the last two decades according to the years of publication and the clinical interest. All articles were filtered through the titles, abstracts, and main texts separately to ensure that the topics and contributions were included. (Pant et al.2020).

2.1 Theoretical framework

In the majority of mental health diagnosis and treatment sites, trained psychiatrists and health professionals are used in the two-stage process: first, to weigh the overall context and experiences of the clients and patients; and second, to diagnose mental health disorders by assessing the severity of mental stress-induced symptoms. (Carleton et al.2020). Using various technological developments in the manufacture of sophisticated and advanced medical devices, another promising way to assess mental health disorders at an early stage appears to be feasible. For example, by using smartphone applications and wearable sensor devices that measure various biomarkers, early alarm and accurate prediction of mental health disorders can be assured. (Bickman2020). When artificial intelligence-based machine learning and deep learning algorithms are used, diagnosis accuracy is evaluated to be around 90% for a number of training data, implying higher accuracy. With the rapid increase of big sequenced data from high-throughput medical equipment, innovation in the construction of models based on pre-processed data from novel AI methods can afford an early-stage diagnosis and open up opportunities for personalized treatment learning in order to fit the diverse and unique chief complaints and needs of individual clients. (Kaur et al.2020). Consequently, the higher diagnosis accuracy and earlier prediction of mental health disorders can lead to a major improvement in the client's quality of life and reduction of medical costs in the long term. Assuming that further machine learning research is conducted to enable the cell as the smallest form of a computer to permit mental health diagnosis and treatment, then the usability, affordability, and accessibility research of mental health diagnosis and treatment could be affected by the existing factors that affect the adoption of services specifically. (Olatunji et al.2024).

2.2 Previous research

AI models for predicting mental health issues have commonly relied on socio-demographic and socio-economic data, such as age group, gender, marital status, employment status, educational level, and housing conditions. With advances in biotechnology, researchers have collected blood, serum, plasma, and urine samples from patients and identified blood biomarkers such as the enzyme butyrylcholinesterase, which has been used in studies for predicting suicidal ideation. (Xu et al.2022). Apart from blood markers, researchers have managed to predict psycho-behavioral disorders by analyzing urine and stool samples. The identification of urine markers for people with mood disorders such as bipolar disorder could yield accuracy ranging from 75% to 80%. AI can also play a part in biomarker research. (Cai et al.2024). Previous researchers have employed AI to optimize predictive models using molecular data, such as blood gene expression, to effectively predict the antidepressant response. Both blood and urine samples have been utilized as biomarkers to improve the effectiveness of suicide prevention. (Lin et al., 2020). The widespread use of social media platforms has raised the potential to use AI algorithms to predict depression. The feasibility of monitoring the emotions, linguistic, or social network structures of individual users on social media platforms also means that a large number of studies or model algorithms could be developed. Models using data have found that the prediction accuracy of predicting depression by analyzing user activities is roughly 58%. (Opoku et al.2021). Moreover, instead of harvesting textual communication from social media platforms, a detection modeling study implemented a comprehensive surveillance framework that associated a wide range of records such as prescription history and diagnosis with depression and anxiety syndrome, without using social media data. (Hochman et al.2021).

2.3 Gaps and literature

In terms of AI-based machine learning and physiological approaches to detect mental health disorders, literature is still at a relatively narrow level. To summarize the literature, we have the following meta-analysis type of papers. First, it was interesting to report that an extensive review showed a meager pattern of evidence related to the third wave of cognitive behavioral therapy, acceptance and commitment therapy, and mindfulness intervention for common mental conditions in universities, from 2000 to 2015. (Khan & Javed, 2022). A recent study seems to portray the same kind of situation. Also, different kinds of narrative reviews carried out were found to be mainly based on simple models with mainly medications or psychological therapies as intervention and the psychological self-reports as the dependent variable, with relatively small samples and lack of blinding. Such narrowness of the literature is seen as a major barrier for the progression of the field, conditioned to centralization on the same type of model. (Navarro et al.2021). Second, a total of 18 recent review studies on early accurate diagnosis of Autism Spectrum Disorder (ASD) itself showed that the type of ad-hoc subject is diagnosis, carrying a total of 46 recognition methods based on the following features as inputs, as we have ad hoc studied. Minibio and literature information boost is undoubtedly important to help progress by bringing deeper intelligent discussion, both about advances in the field itself and how it can be applied to go beyond the current application scenarios. (McCarty & Frye, 2020).

2.4 Conceptual framework

The overall objective of the AI-based model for predicting mental health disorders using biomarkers, social media data, and behavioral patterns is an effective, proactive digital mental health support system that can increase access to effective mental health services. By analyzing relevant literature and the advantages that frontier technologies bring to mental health research and application, we propose a clear conceptual framework using these technologies. (Lee et al.2021). This research provides an insightful understanding of cutting-edge technologies in mental health support and, particularly, the combination of these technologies into an integrated model. We hope to attract academic researchers and commercial investors to further explore technical directions, and we also appreciate nonprofit organizations that can effectively apply our research in real-world scenarios. (Lattie et al., 2022). Our overall objective is an effective, proactive digital mental health support system that can increase access to effective mental health services. This system continuously collects user multimodal data; analyzes, aggregates, and derives users' mental well-being indicators; predicts user mental health disorders; and performs real-time proactive intervention for negative events. (Xu et al.2020). These steps could identify both the negative emotions users have expressed in a wearable device, such as the emotions in a user's conversation, or the physiological health data, such as the data trends and abnormal values in a user's wearable devices. (Mullick et al.2022).

3. METHODOLOGY

The high prevalence of mental health disorders and the demand for effective screening of large populations call for scalable and accessible diagnostic processes. Recent advances in sensing technologies provide an opportunity to develop cost-effective and low-burden methods for monitoring and diagnosing mental health disorders. In this work, we present a machine learning-based system for predicting mental health disorders using physiological signals and speech components extracted from voice recordings collected from smartphone sensors. To validate the robustness of the ML models and their overall performance, we measured the prediction accuracy on a dataset containing voice recordings of subjects obtained using smartphones. We measured the prediction accuracy for five common mental health disorders, which include depression, mental distress, stress, anxiety, and bipolar disorder. Our results indicate that non-verbal information such as voice recordings contains predictive patterns that can be of significant value in the screening, monitoring, and diagnosis of patients with mental health disorders. This work focuses on utilizing physiological and speech signals obtained using smartphone sensors in predicting mental health disorders. Our study employs AI tools and methodologies to design predictive models that leverage powerful non-verbal cues in voice data. We use the phone's sensors to obtain voice recordings of subjects whose mental health status is known. We then analyze the quality of these voice recordings to obtain relevant information and features from the data. Our feature extraction process utilizes tools in speech processing and computational biology to extract annotated components from the segment of the voice signals of the subjects. We then develop predictive ML models that make use of the extracted features from the voice recordings. The predicted mental health outcomes using our ML models are compared against the known mental health status of the subjects. These experiments make use of a number of AI-based models such as unsupervised deep learning, supervised learning for dimension reduction, decision trees, and other supervised models. Following traditional ML model building approaches, to optimize prediction performance, we use nested cross-validation and grid search.

3.1 Research design

We conducted a longitudinal prospective cohort study to understand, classify, and predict mental health outcomes based on patients' behaviors, social media communications, and use of a monitor. Our study analyzed a de-identified copy of two datasets containing longitudinal data on over 1,000 hospital visits across 171 individuals, consisting of data from pre-admission, during hospitalization, and post-discharge phases. The datasets collected were response variables for 10 psychiatric disorders, including depression, suicidal ideation, PTSD, bipolar disorder, and eating disorders, using structured interviews following established criteria through well-validated instruments. Behavioral data such as logs of patients' electronic health record use, verbal intensity, and device use of wearable and mobile computing devices, including speech device sensor monitoring data, were also collected.

The study focused on analyzing the complete and de-identified dataset to establish a structured approach towards modeling mental health disorders, a step crucial for developing electronic mental health monitoring and response systems. We utilized this data to address two central questions: which biomarker-sourced, individual-level patient data systems' features were likely to improve predictions of major psychiatric disorders and represented building blocks for future early warning systems, and what models yielded the most accurate forecasts for a short-term future risk of these mental health disorders. Device use, sociodemographic data, and socioeconomic status observed at admission were chosen as input features based on their feasibility and interpretability of implementation in a hospital setting. This EHR-derived information was fed into individual-level logistic regression models predicting the mental health disorders observed at subsequent hospital admission. The logistic regression models' predictions were assembled into a single forecast of each disorder, guided by prior estimates of comorbidity. Overall, the clinical value of individual features, such as the use of ambulatory monitors and EHR utilization data, is deemed to promote detectable signals distinguishing those at risk for developing psychiatric disorders and serve as integral stay-ahead sensors.

3.2 Sampling method

In this study, we proposed and evaluated several different sampling methods for addressing the multiple challenges of modeling high-skew distributional data on a binary classification task. The advantages, disadvantages, and benefits of using these methods and their possible implications for similarly challenged tasks are also discussed. Three over-sampling and four under-sampling methods were evaluated according to sensitivity, specificity, precision, recall, F1-score, receiver operating characteristic curve, and the area under the ROC curve for both the control and imbalanced internal validation data. When the imbalance ratio was set at 1–3, the ADASYN method had the best performance in terms of precision, recall, F1, and AUC on the internal validation data, both in the case of 5-fold cross-validation and the random 70%–30% data split for 10 independent runs. If we consider the time cost and data scale, down-sampling in a balanced way can achieve good precision, recall, F1, and AUC. In the final evaluation, down-sampling can achieve robust and good performance in terms of sensitivity, specificity, precision, recall, F1, and AUC for both internal and external validation. We argue that ADASYN may be a useful tool for data from a similar high-skew distributional binary classification task, but it may require more time to execute. These findings can provide a theoretical and methodological guide for achieving a robust model and translating the model’s topology into a practical solution in the real world or for use in clinical applications regarding social media data, mental health disorder prediction, and distributional data.

It is difficult to obtain reliable clinical data for people with mental health disorders. Social media data may be valuable for predicting mental health disorders. However, there are several challenges associated with using social media for predicting mental disorders because the data might be high-skew distributional for the specific task. The dataset we used herein was highly skewed. Most participants presented with low social media activity, and only a minority of the participants were diagnosed with mental health disorders by physicians. Given the high quality of the data provided for the use of social media detailed behavioral patterns, the results could contribute to the development of better models and ideas. Furthermore, it is beneficial to translate the best performance model based on the internal validation into practical solutions in the real world, and with different external validation data, ensuring that they are reusable in real-world clinical applications.

3.3 Data collection methods

The sites approved by the hospital were accessed, such as portals, groups, and communities. This step was essential for data retrieval and collection. The main objective in order to access the data from social media sites and private areas was to look for posts from patients that contained complaints related to depressive symptoms. Queries about sleep disturbances or pains, among others, in the forums would help collect this data and make it possible to find patients, what complaints they had, what time of the day they usually asked for help, accessing places, and sharing qualitative information that could be used in the development of an app with the posts and comments of the patients themselves.

The patient forums were easily accessed, and the posts, along with the comments, could be accessed. Since the data was used through the consultation, it was possible to store the retrieved posts and build execution scripts. However, there were complexities with the acquired database to be used post-queries. Most of these issues were related to the commands being executed that appeared to be correctly inserted but were failing to perform the search. Understand that the methodology employed for this project was carried out in three phases: search, data collection, and data storage.

3.4 Data analysis methods

The data used will be collected at different experimental evaluation environments and from available datasets. In all studies, the data collected will be normalized before classification, using existing methods. Moreover, the analysis itself aims at new methods and procedures that may contribute to the scientific community, as the results presented will be submitted to experiments with organizers and groups of volunteers with common characteristics. The hypothesis tests will be conducted and evaluated scientifically against benchmarks in the areas considered in the tasks. New and additional evaluations may require the pre-processing of metadata. The subsequent aggregation and integration processes will consider the classification tasks. Since the forecasts involve few articles and an insignificant amount of data in comparison with the new proposals, they need to be studied in sufficient quantities to justify their uses in benchmarks related to each of the main tasks. The sample population data will, therefore, need to be constructed.

3.5 Ethical considerations

Ethical considerations are at the core of any work involving the handling of sensitive information and demographic groups. They are particularly relevant when using AI/ML models to analyze mental health traits, behaviors, and labels, as often the methods can lie behind a 'black box' that delivers final predictions that cannot be easily explained. The performance of these models will vary greatly with the method used. We are encouraged by the recent advancements that utilize molecular biomarker data for mental health predictions and highlight the ethical priorities and considerations for developing AI/ML models for behavioral mental health predictions. The list is formatted as a set of recommendations for industry practitioners, with an emphasis on issues particular to the latest generation of AI/ML models using data of clinical relevance.

We note that more general sensitivity and privacy concerns need to be addressed too, especially for models utilizing behavioral data and social media analysis. These include de-anonymization and re-identification, invasive monitoring and untoward influence, and exacerbation of hidden biases and stereotypes. In instances where sensitive data is accessed in an uncontrolled manner, practitioners should also consider the increased desirability of these models and the increased potential risk of privacy violations. Additionally, ethical safeguards are not only about avoiding negative consequences but also about ensuring that predictions obtained from AI/ML models are robust and can be trusted. This ensures the evaluation process is conducted rigorously, that the models are generalized and validated for use on data from various populations, especially before being utilized for clinical decisions. In any data-driven subfield, the new results should support unthinking reliance on numbers to the exclusion of knowledge about human nature, urge for interrogation of the theoretical underpinnings when necessary, and advocate for transparent dissemination and replicability for the AI/ML models.

4. DATA ANALYSIS AND FINDINGS

Based on the prevalence of diseases in data from a general insurance company, it was found that the insurance participants were women accounting for nearly 70% and predominantly from 30 to 50 years old. Because the age range of customers was predominantly between 30 and 50 years, it is a typical distribution according to general insurance purchasing in Vietnam. The socio-professional information of these participants gives an overview that tends to avoid white-collar workers more than other jobs. Life, health, and stability of mental health for workers are extremely important, especially in today’s stressful society. This study has raised the probability of a participant suffering from mental illness compared to the general population, the risk index based on lifestyle, health information, and other socio-professional information from general insurance participants. According to the different contributions of factors and different levels of mental disorder risk exposure, the insurance company provides alerts to each customer group intelligently and specifically. The appropriate consultation can provide information about employers and trends if there are objective problems. Information that aims at the discovery of customers suffering from mental disorders is an attractive service or can be a channel for health organizations to promote the treatment of mental disorders to people with mental illnesses through the appeal of insurance service packages.

4.1. Presentation of results

AI-based models for predicting mental health disorders using biomarkers, social media data, and behavioral patterns: 4.1. Presentation of results

The results of the different classifications and regression models for the selected prediction tasks, in terms of their respective performance metrics: accuracy, F1 score, precision, recall, area under the PR curve, and area under the ROC curve, are presented. The results for the prediction of major depressive disorder are shown where we employed three types of input data, the BFI-I test, the EMA data, and the Twitter data, using different machine learning models for binary classification. We observe that the SVM model, in our case an RBF kernel, achieved the highest F1 score of 0.8052 and the highest AUC value of 0.89 using the EMA data as input. We also observe that LR regression and decision tree models achieved comparable results, while the other classification models suffered from lower predictive power. Our model clearly separated the population based on the sum of PANAS scores, which strongly demonstrates the effectiveness of T-pattern metrics for MDD prediction.

These results suggest that the specific behavioral patterns and other emotional aspects captured by the EMA data in our study have significant value for MDD detection when compared to general personality traits demonstrated over a long period of time. The proposed model may thus contribute to the screening of high-risk individuals and reduce the costs of diagnosis, which are associated with multiple appointments with a professional therapist. In the case of data collection involving the BFI-I behavioral test, the XGBoost model demonstrated the best F1 score of 0.76. Surprisingly, the non-linear model XGBoost had the best performance, which indicates that the relationships between BFI-I test answers and the depression symptoms, measured using the patient’s mood score, in the EMA may be complex. This is consistent with the fact that the relations between BFI-I answers and actual mental health are not directly determined by any biological measures but by the patient’s subjective emotional state.

4.2 Analysis of data

We collect data from five different sources, and the pre-processing steps used for constructing the dataset are explained in this subsection. Our dataset comprises neuroimaging, electroencephalography, electrocardiography, eye-blink, and self-reported psychological measures for subjects belonging to different age groups. Pre-processing of these diverse data leads to a highly structured dataset that can be used for performing a wide variety of experiments, and in this work, we focus on the problem of predicting major depressive disorder. The dataset thus constructed has been validated and used to evaluate the efficacy of various machine learning algorithms and detect the optimal algorithms for predicting MDD. In conclusion, a succinct representation of our dataset pre-processing techniques on each of the data types and the construction of the labeled dataset that can be used for training various machine learning models is presented. Extensive preprocessing on these diverse data sources has been done to come up with a dataset amenable to predictions. Real data procured from many sources allows our machine learning models to be deployable, assisting researchers in overlaying data for many other medical conditions and helping medical practitioners with more accurate diagnoses. The subsequent computational bottleneck would be architectural and parameter translation for each of the data segments, and unlike the case of other classifiers, our deep neural network that we utilize would not have architecture derived from intuitive human specifications, making it less interpretable.

4.3 Comparison with previous studies

There are several significant distinctions between our study and previous studies applying machine learning techniques. To begin with, the dataset was obtained from the primary healthcare recruitment clinic. Unlike previous studies that have relied on self-guided recruitment with no method to verify the user, we found a large pool of verifiable and diverse patient data from our clinic-based recruitment. As a result, we were able to obtain a study population that is representative of various mental health symptomatology and ethnic backgrounds, as opposed to other studies that have a limited geographic and ethnic representation. Second, we were able to control potential confounding factors by controlling for demographic information, clinical measures, and behavioral records. Lastly, we used multiple sources of data, such as biological measures, in addition to typical social media and texting patterns of users.

Some studies have suggested similar levels of prediction performance using much smaller data and minor modeling differences. Additionally, some studies differ from our approach by focusing solely on using social media data. They observed not fewer strengths of precursors. While these studies indicated a high discriminative power without using outcome measures, their lack of clinical gold standard profiling assessments or model performance comparison to alternative prediction models, including a traditional questionnaire, limits the reviewers' ability to determine the practical usefulness of the proposed models in a real-world clinical setting. Our study attempted to address some of these shortfalls in previous research and validate results in a secondary cohort, albeit with lower prediction accuracy than the original study.

4.4 Discussion of key findings

This review reveals that there is a knowledge gap in developing a modeling pipeline that is highly predictive and provides insights into which biomarkers, social media data, and behavioral patterns are important. This lack of predictive and interpretative capability of deep learning models is also apparent in the field of computer vision. An important point is made in the understanding of predictive model interpretability, "One could argue that whether a model is interpretable often largely depends on our problem context and the particular user, reasoner, or context in which it is being employed or evaluated." Further research also needs to be conducted to utilize MRI, fMRI, and EEG data to develop a real-time mental health monitoring system based on single raw input

EEG signals. A real-time deep learning pipeline is developed that is sensitive to real-time single raw input EEG signal features with a predictive performance of an AUC of 0.768 using a communication scenario, a language voice reasoning task to study the natural character of human communication through speech, and to study how this can be exploited for continuous emotion recognition in humans.

Despite a large volume of publications on computational methods for mood disorders or mental health, only a few methods discussing their evaluation of real-life deployment exist. The vast majority of the scientific narrative that underpins efforts to harness social media data as mental health proxies is based on algorithms developed by a relatively small network of researchers who established networks and data from social media companies before their access to real-world data was significantly curtailed by ethical, legal, and commercial considerations.

5. CONCLUSION AND RECOMMENDATION

The importance of early detection of mental health problems has led to several studies aimed at predicting mental health disorders such as depression, bipolar disorder, and schizophrenia using risk biomarkers, social media data, and behavioral patterns. Each of these diagnostic models has its limitations. Therefore, in this paper, we conducted a review of AI-based models for predicting mental health problems using social media data and biomarkers and highlighted the advantages, drawbacks, and recommendations of each of these models. Our review showed that existing AI models have been shown to effectively predict some mental health disorders by analyzing risk biomarkers, social media usage, and behavioral patterns. The findings, imperfections, and recommendations in this study will help in paving a means of filling identified gaps in diagnostic models, facilitating the development of multi-component diagnostic models, and promoting the growth of the area.

We recommend:

- Conducting more studies on AI-based models for predicting different mental health disorders using social media data.
- Integrating more detailed aspects of personality into AI-based models.
- Deploying and processing multi-modal data to improve model performance.
- Employing longitudinal and group experimental designs in model evaluation.
- Self-corroboration and validation of self-reported depression or other mood disorders.
- Considering factors that may interfere with the evaluation of predictive models in the real world, such as the experience of confronting the social science problems of investigators.
- Building privacy assurance systems, particularly in contexts wherein the information to generate models is sourced from social media platforms.

5.1 Summary of findings

We have employed different datasets and prediction models for predicting mental health conditions, including generalized anxiety, depression, and post-traumatic stress disorder using patients' biomarker data or social media data. Our biomarker data included data from 120 subjects identified as healthy controls from a comprehensive investigation of the human brain in PTSD military war fighters, and subjects from a breadth of PTSD transcriptomics where participants were either without PTSD diagnoses or had current PTSD. We have identified driver genes or driver biological pathways as these driver biological signatures provide better accuracy and potentially help predict genetic and epigenetic features of certain groups in a population, which, in turn, may provide better understanding and effective treatment for the related mental health conditions. Our social media data came from a social media platform that allows a specific group of people to share their thoughts on topics related to mental health conditions.

5.2 Conclusion

In this chapter, we extensively described various AI-based models that aim to predict anxiety, depression, and PTSD by utilizing data harvested from social media platforms, analyzing behavior patterns, and incorporating certain biomarkers. The features that contributed to the formulation of these models were meticulously derived from both textual content and associated metadata, effectively reflecting distinct patterns that can be observed between healthy controls and individuals who are suffering from mental health disorders during the execution of various tasks. In these specific tasks, it becomes evident that those individuals who exhibit symptoms of anxiety, depression, and PTSD—conditions that are recognized by AI-based models—display certain features that signal significant impairments in both verbal and nonverbal communication. Furthermore, these individuals reveal differences in their activity regularities compared to their healthy counterparts. Achieving an adequate sample size and determining the types of tasks performed stand out as key challenges in the ongoing development of these predictive models. If these AI-based models can be validated as effective, they hold considerable promise in functioning as tools for the discovery of impairments, as well as serving pivotal roles in preventive measures and assessing treatment responses. The work discussed in this chapter contributes to this burgeoning field.

5.3 Implications of the study

Understanding the implications of the study that aims to predict onsets of mental health issues with a deep understanding of neuroinformatics, data science, and behavioral informatics using novel artificial intelligence methods for biomarkers, posture and walking patterns, and linguistic patterns from multiple modalities of data derived from smartphones is significant. By performing this study, the research community benefits from the following pioneering approaches. It may raise awareness of mobile sensing apps and encourage the use of bioacoustics as well as thermal dynamics and haptic patterns for future sensing devices, enabling them to be combined with mobile sensing to improve pre- and post-screening of mental health issues. Another approach would be to extract large-scale data from participants with various demographic distributions across diverse geographic locations. More demographic features mean more prominent hierarchies, diversity, and hierarchical distributions. Creating such a novel dataset needs to be better understood by asking well-defined research questions and developing robust tools. For instance, when data is split subjectively into cross-sectional and longitudinal groups by human assessors for mental health diagnostics, the sensitivity and specificity of the networks are crucial.

5.4 Recommendations for the future research

Fusing different types of data seems to be the most promising way to generate precise and robust models. Apart from mentioning and investigating more types of data and their coherencies, it is also interesting to investigate modality distributions, data coherencies, and model design. Deep learning algorithms have shown a great privileged status in the field due to their good performance in data aggregation and knowledge generalization. Exploring new types of deep learning would also be beneficial for present and future research. In addition to developing models and progressing the process to meet the performance gaps, exploring new tasks that may better present the potential power of models also receives attention. Finally, the sole usage of benchmark data is considered a limitation of the current research, and multi-source data fusing and comparison is our future research work.

In this chapter, we present a review of recent works in developing AI-based models for the prediction of mental health problems. In principle, we begin by presenting data types and related works that show their significance in the prediction model. In these case reports and empirical findings, we make the link between these candidate data and the related mental health problems. Following that, recent works using the candidate data types are discussed. Selected models, data types and tasks, and present and future limitations are also presented in comparison. Not only do we try to gain a better understanding and compare the successes and limitations, but we also propose some potential trends for future AI-based prediction models for mental health problems.

5.5 Practical recommendation

Step I: Data Preprocessing Data preparation and preprocessing constitute the first steps for applying machine learning models to healthcare problems. Preprocessed data are the foundation for training robust and efficient models. Practical tips are discussed below to preprocess biomarker, health app log data, or social media data.

Biomarkers: A common preprocessing approach for biomarkers, including family history, blood, hormone, stress, brain pathology, and neurobehavioral tests, is feature scaling. We used the z-score for this. Patient identity is used to identify participants in different analyses, but it is not a target. For ML, the presence of a large range of male and female participants and different desired outcomes is important.

SMD: Our study used daily diary entries collected from platforms. Each post format is utilized to optimize collection. Daily diaries direct when the entry must be completed; the entry also includes birthday, age, and the day's emotional state using a seven-item scale. The focus in our study was negative emotion, which used the sadness and anxiety measurements, but additional questions are included. Responses—examples are birthday, age, z-score tables, daily image files, and the table of daily emotional states collected/used in our study—are typical entries for SMD preprocessing using standard practice.

The features were scaled and centered by calculating the z-score for measuring noise, normality, and distribution in the sample. Data are visualized using line plots, count plots, scatter plots, area plots, violin plots, and box plots to check signal periodicity, associations, distribution, centering, and outliers. Importantly, the entries are balanced, and feature selection and correlation exploration are performed for dimension reduction and computational efficiency.

REFERENCES

1) Knapp, M. & Wong, G. (2020). Economics and mental health: the current scenario. World Psychiatry. wiley.com
2) Aguirre Velasco, A., Cruz, I. S. S., Billings, J., Jimenez, M., & Rowe, S. (2020). What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. BMC psychiatry, 20, 1-22. springer.com
3) Rashid, B. & Calhoun, V. (2020). Towards a brain-based predictome of mental illness. Human brain mapping. wiley.com

4) Meehan, A. J., Lewis, S. J., Fazel, S., Fusar-Poli, P., Steyerberg, E. W., Stahl, D., & Danese, A. (2022). Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. Molecular psychiatry, 27(6), 2700-2708. nature.com
5) Mansi, S. A., Barone, G., Forzano, C., Pigliautile, I., Ferrara, M., Pisello, A. L., & Arnesano, M. (2021). Measuring human physiological indices for thermal comfort assessment through wearable devices: A review. Measurement, 183, 109872. sciencedirect.com
6) Chung, J., & Teo, J. (2022). Mental health prediction using machine learning: taxonomy, applications, and challenges. Applied Computational Intelligence and Soft Computing, 2022(1), 9970363. wiley.com
7) Moisescu, O. I., Dan, I., & Gică, O. A. (2022). An examination of personality traits as predictors of electronic word-of-mouth diffusion in social networking sites. Journal of Consumer Behaviour. [HTML]
8) Zhong, B., Huang, Y., & Liu, Q. (2021). Mental health toll from the coronavirus: Social media usage reveals Wuhan residents' depression and secondary trauma in the COVID-19 outbreak. Computers in human behavior. nih.gov
9) Brauer, K., Sendatzki, R., & Proyer, R. T. (2024). Exploring the acquaintanceship effect for the accuracy of judgments of traits and profiles of adult playfulness. Journal of Personality. wiley.com
10) Zhang, T., McCoy, T. H., Perlis, R. H., Doshi-Velez, F., & Glassman, E. (2021, October). Interactive cohort analysis and hypothesis discovery by exploring temporal patterns in population-level health records. In 2021 IEEE Workshop on Visual Analytics in Healthcare (VAHC) (pp. 14-18). IEEE. github.io
11) Waters, L., Algoe, S. B., Dutton, J., Emmons, R., Fredrickson, B. L., Heaphy, E., ... & Steger, M. (2022). Positive psychology in a pandemic: Buffering, bolstering, and building mental health. The Journal of Positive Psychology, 17(3), 303-323. tandfonline.com
12) Moss, R. A., Gorczynski, P., Sims-Schouten, W., Heard-Laureote, K., & Creaton, J. (2022). Mental health and wellbeing of postgraduate researchers: exploring the relationship between mental health literacy, help-seeking behaviour, psychological distress, and wellbeing. Higher Education Research & Development, 41(4), 1168-1183. tandfonline.com
13) Nemesure, M. D., Heinz, M. V., Huang, R., & Jacobson, N. C. (2021). Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. Scientific reports. nature.com
14) Ríssola, E. A., Losada, D. E., & Crestani, F. (2021). A survey of computational methods for online mental state assessment on social media. ACM Transactions on Computing for Healthcare, 2(2), 1-31. google.com
15) Santamaría-García, H., Baez, S., Gómez, C., Rodríguez-Villagra, O., Huepe, D., Portela, M., ... & Ibanez, A. (2020). The role of social cognition skills and social determinants of health in predicting symptoms of mental illness. Translational psychiatry, 10(1), 165. nature.com
16) Wies, B., Landers, C., & Ienca, M. (2021). Digital mental health for young people: a scoping review of ethical promises and challenges. Frontiers in digital health. frontiersin.org
17) Breaux, R., Dvorsky, M. R., Marsh, N. P., Green, C. D., Cash, A. R., Shroff, D. M., ... & Becker, S. P. (2021). Prospective impact of COVID-19 on mental health functioning in adolescents with and without ADHD: Protective role of emotion regulation abilities. Journal of Child Psychology and Psychiatry, 62(9), 1132-1139. wiley.com
18) Deepa, N., Pham, Q. V., Nguyen, D. C., Bhattacharya, S., Prabadevi, B., Gadekallu, T. R., ... & Pathirana, P. N. (2022). A survey on blockchain for big data: Approaches, opportunities, and future directions. Future Generation Computer Systems, 131, 209-226. [PDF]
19) Odgers, C. L., & Jensen, M. R. (2020). Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions. Journal of Child Psychology and Psychiatry, 61(3), 336-348. nih.gov
20) Vuorre, M., Orben, A., & Przybylski, A. K. (2021). There is no evidence that associations between adolescents' digital technology engagement and mental health problems have increased. Clinical Psychological Science, 9(5), 823-835. sagepub.com
21) Gianfredi, V., Provenzano, S., & Santangelo, O. E. (2021). What can internet users' behaviours reveal about the mental health impacts of the COVID-19 pandemic? A systematic review. Public health. nih.gov
22) Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H. C., ... & Jeste, D. V. (2021). Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 6(9), 856-864. nih.gov
23) Karunarathna, I., De Alvis, K., Gunasena, P., & Jayawardana, A. (2024). Perioperative and ICU care for COPD patients: Enhancing outcomes through multidisciplinary approaches. researchgate.net
24) Galea, S., Merchant, R. M., & Lurie, N. (2020). The mental health consequences of COVID-19 and physical distancing: the need for prevention and early intervention. JAMA internal medicine. jamanetwork.com

25) Koutsouleris, N., Hauser, T. U., Skvortsova, V., & De Choudhury, M. (2022). From promise to practice: towards the realisation of AI-informed mental health care. The Lancet Digital Health, 4(11), e829-e840. thelancet.com
26) Sui, J., Jiang, R., Bustillo, J., & Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. Biological psychiatry. sciencedirect.com
27) Nahavandi, D., Alizadehsani, R., Khosravi, A., & Acharya, U. R. (2022). Application of artificial intelligence in wearable devices: Opportunities and challenges. Computer Methods and Programs in Biomedicine, 213, 106541. [HTML]
28) Colizzi, M., Lasalvia, A., & Ruggeri, M. (2020). Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care?. International journal of mental health systems, 14, 1-14. springer.com
29) Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., ... & Firth, J. (2021). The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry, 20(3), 318-335. wiley.com
30) Chancellor, S. & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: a critical review. NPJ digital medicine. nature.com
31) Drouin, M., McDaniel, B. T., Pater, J., & Toscos, T. (2020). How parents and their children used social media and technology at the beginning of the COVID-19 pandemic and associations with anxiety. Cyberpsychology, Behavior, and Social Networking, 23(11), 727-736. parkviewhealth.org
32) Tutun, S., Johnson, M. E., Ahmed, A., Albizri, A., Irgil, S., Yesilkaya, I., ... & Harfouche, A. (2023). An AI-based decision support system for predicting mental health disorders. Information Systems Frontiers, 25(3), 1261-1276. springer.com
33) Abi-Dargham, A., Moeller, S. J., Ali, F., DeLorenzo, C., Domschke, K., Horga, G., ... & Krystal, J. H. (2023). Candidate biomarkers in psychiatric disorders: state of the field. World Psychiatry, 22(2), 236-262. wiley.com
34) Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. Information Fusion, 96, 156-191. google.com
35) Herrman, H., Patel, V., Kieling, C., Berk, M., Buchweitz, C., Cuijpers, P., ... & Wolpert, M. (2022). Time for united action on depression: a Lancet–World Psychiatric Association Commission. The Lancet, 399(10328), 957-1022. lse.ac.uk
36) Health Organization, W. (2022). World mental health report: Transforming mental health for all. google.com
37) Pant, M., Zaheer, H., Garcia-Hernandez, L., & Abraham, A. (2020). Differential Evolution: A review of more than two decades of research. Engineering Applications of Artificial Intelligence, 90, 103479. [HTML]
38) Carleton, R. N., Afifi, T. O., Turner, S., Taillieu, T., Vaughan, A. D., Anderson, G. S., ... & Camp, R. D. (2020). Mental health training, attitudes toward support, and screening positive for mental disorders. Cognitive Behaviour Therapy, 49(1), 55-73. tandfonline.com
39) Bickman, L. (2020). Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. Administration and Policy in Mental Health and Mental Health Services Research, 47(5), 795-843. springer.com
40) Kaur, S., Singla, J., Nkenyereye, L., Jha, S., Prashar, D., Joshi, G. P., ... & Islam, S. R. (2020). Medical diagnostic systems using artificial intelligence (ai) algorithms: Principles and perspectives. IEEE Access, 8, 228049-228069. ieee.org
41) Olatunji, A. O., Olaboye, J. A., Maha, C. C., Kolawole, T. O., & Abdul, S. (2024). Revolutionizing infectious disease management in low-resource settings: The impact of rapid diagnostic technologies and portable devices. International Journal of Applied Research in Social Sciences, 6(7), 1417-1432. researchgate.net
42) Xu, Y., Liang, J., Gao, W., Sun, Y., Zhang, Y., Shan, F., ... & Xia, Q. (2022). Peripheral blood cytokines as potential diagnostic biomarkers of suicidal ideation in patients with first-episode drug-naïve major depressive disorder. Frontiers in public health, 10, 1021309. frontiersin.org
43) Cai, H., Du, R., Yang, K., Li, W., & Wang, Z. (2024). Bibliometric Analysis of the Research Status and Global Trends in Behavioral and Psychological Symptoms of Dementia in Alzheimer's Disease from 2002 to 2022. Current Neuropharmacology, 22(10), 1720-1732. nih.gov
44) Lin, E., Lin, C. H., & Lane, H. Y. (2020). Precision psychiatry applications with pharmacogenomics: artificial intelligence and machine learning approaches. International journal of molecular sciences. mdpi.com
45) Opoku Asare, K., Terhorst, Y., Vega, J., Peltonen, E., Lagerspetz, E., & Ferreira, D. (2021). Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. JMIR mHealth and uHealth, 9(7), e26540. jmir.org

46) Hochman, E., Feldman, B., Weizman, A., Krivoy, A., Gur, S., Barzilay, E., ... & Lawrence, G. (2021). Development and validation of a machine learning-based postpartum depression prediction model: A nationwide cohort study. Depression and anxiety, 38(4), 400-411. wiley.com
47) Khan, N. Z. & Javed, M. A. (2022). Use of artificial intelligence-based strategies for assessing suicidal behavior and mental illness: A literature review. Cureus. nih.gov
48) Navarro, C. L. A., Damen, J. A., Takada, T., Nijman, S. W., Dhiman, P., Ma, J., ... & Hooft, L. (2021). Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. bmj, 375. bmj.com
49) McCarty, P. & Frye, R. E. (2020). Early detection and diagnosis of autism spectrum disorder: Why is it so difficult?. Seminars in Pediatric Neurology. sciencedirect.com
50) Lattie, E. G., Stiles-Shields, C., & Graham, A. K. (2022). An overview of and recommendations for more accessible digital mental health services. Nature Reviews Psychology. nature.com
51) Xu, Z., Pérez-Rosas, V., & Mihalcea, R. (2020, May). Inferring social media users' mental health status from multimodal information. In Proceedings of the twelfth language resources and evaluation conference (pp. 6292-6299). aclanthology.org
52) Mullick, T., Radovic, A., Shaaban, S., & Doryab, A. (2022). Predicting depression in adolescents using mobile and wearable sensors: multimodal machine learning–based exploratory study. JMIR Formative Research, 6(6), e35807. jmir.org
53) 2019 Mental Disorders Collaborators, G. B. D. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of …. The Lancet Psychiatry. sciencedirect.com