Volume 6, Issue 3, March – 2021 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

Predicting Mental Health Outcomes Using Wearable Device Data and Machine Learning

Nikhil Sanjay Suryawanshi
California, USA

Abstract:- This paper proposes a machine learning-based system designed to predict mental health outcomes using wearable device data. The system is conceptualized to process physiological and behavioral data such as heart rate, sleep patterns, and activity levels collected from wearable technology. Key stages of the system include data preprocessing, feature extraction, and model training using multiple machine-learning algorithms, including Random Forest, Support Vector Machine, XGBoost, and Logistic Regression. These models are combined using a voting-based ensemble classifier to improve prediction accuracy. While the system has not yet been implemented, expected results suggest that this approach will enhance prediction reliability and offer real-time insights into mental health conditions. The proposed system is envisioned to facilitate early detection of mental health disorders, thereby aiding in timely interventions and personalized care.

Keywords:- *Wearable Devices, Mental Health Prediction, Machine Learning, Ensemble Learning, Random Forest, Support Vector Machine (SVM), XGBoost, Logistic Regression, Voting Classifier, Physiological Data, Behavioral Data, Feature Extraction, Mental Health Monitoring, Predictive Analytics, Health Technology*.

I. INTRODUCTION

Mental health disorders are a growing global concern, affecting millions of individuals worldwide and imposing a significant burden on healthcare systems [1]. The World Health Organization estimates that one in four people will be affected by mental or neurological disorders at some point in their lives [2]. Traditional methods of mental health assessment and monitoring often rely on self-reporting and periodic clinical evaluations, which may not capture the dynamic nature of mental health states or provide timely interventions [3].

In recent years, the proliferation of wearable devices has opened new avenues for continuous, real-time monitoring of physiological and behavioral data [4]. These devices, including smartwatches, fitness trackers, and specialized sensors, can collect a wide range of data such as heart rate variability, sleep patterns, physical activity, and social interactions [5]. This wealth of information when combined with advanced machine learning techniques presents a promising opportunity to revolutionize mental health care through early detection, accurate prediction, and personalized interventions [6].

The integration of wearable technology in mental health research has already shown potential in various areas. For instance, studies have demonstrated the ability to detect stress levels using physiological signals from wearable devices [7], predict mood changes in bipolar disorder patients [8], and identify early signs of depression [9]. Moreover, the continuous nature of data collection from wearables allows the capture of subtle changes and patterns that might be missed in traditional clinical assessments [10].

Machine learning algorithms have proven to be powerful tools in analyzing complex, high-dimensional data from wearable devices [12]. These techniques can identify intricate patterns and relationships within the data that may not be apparent through conventional statistical methods. Various machine learning approaches, including supervised learning, unsupervised learning, and deep learning, have been applied to mental health prediction tasks with promising results [13].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

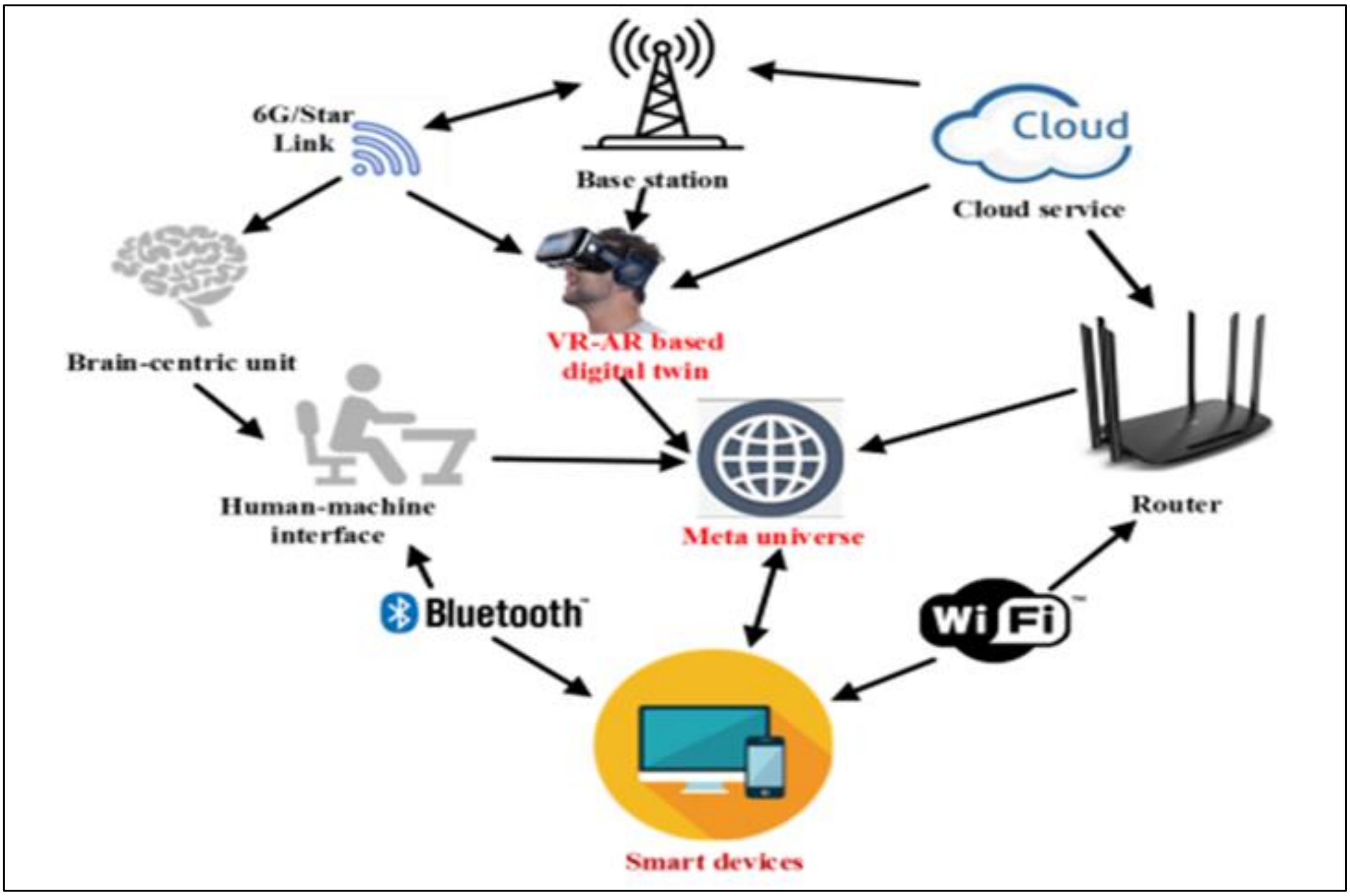

Fig 1: Intelligent Ecology Flowchart of AI-Based Wearable Devices with Data Storage, Transaction, Interaction, and Communication Networks [11].

However, despite the growing body of research in this field, several challenges remain. These include ensuring the privacy and security of sensitive health data [14], addressing the interpretability of complex machine-learning models in clinical settings [15], and validating the generalizability of predictive models across diverse populations [16]. Additionally, there is a need for standardization in data collection protocols and feature extraction methods to facilitate comparability across studies and enable the development of robust, widely applicable predictive models [17].

This research paper aims to contribute to the evolving field of mental health prediction using wearable device data and machine learning. We will explore novel approaches to feature engineering, investigate the efficacy of various machine learning algorithms, and propose a framework for integrating these predictive models into clinical practice. By integrating continuous, multi-modal data from wearables with advanced analytics, we aim to enhance our understanding of mental health dynamics and improve patient outcomes through proactive intervention and personalized care strategies.

II. LITERATURE REVIEW

Sano et al. [18] demonstrated that physiological and behavioral data collected from wearable sensors could predict next-day mood, stress, and health with accuracy rates between 55% and 78%. Their study used machine learning algorithms on data from 201 college students, highlighting the potential of wearables in mental health monitoring.

Jacobson et al. [19] explored the use of smartphone and Fitbit data to predict depression symptoms. They discovered that sleep, activity, and phone usage features together could predict depression severity with moderate accuracy (R^2 = 0.48). Their work emphasized the importance of multimodal data in mental health prediction.

Torous et al., [20] utilized smartwatch data to predict relapse in patients with schizophrenia. By analyzing heart rate variability and sleep patterns, they achieved a prediction accuracy of 89% for relapse events up to two weeks in advance. This research underscored the potential of wearables in managing severe mental illnesses.

Dobson, Rosie et al., [21] focused on anxiety prediction using data from wrist-worn accelerometers. They developed a deep learning model that could identify high anxiety states with 83% accuracy based on movement patterns alone. Their

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

work highlighted the potential of passive sensing in mental health monitoring.

Addressing the challenge of suicide risk assessment, Kleiman et al. [22] used ecological momentary assessment (EMA) combined with wearable sensor data to predict suicidal thoughts and behavior. Their machine-learning model achieved an AUC of 0.93 in identifying high-risk periods, demonstrating the potential of real-time monitoring for suicide prevention.

Wang et al. [23] explored the use of smartphone sensors and usage patterns to detect depressive states. Their machine learning model, trained on data from 48 college students over 10 weeks, achieved 86.5% accuracy in detecting depressive states. This study highlighted the potential of passive sensing using ubiquitous devices.

Table 1: Review

Ref.	Findings	Methods used	Dataset	Limitations
[26]	Wearable device data, including sleep metrics and heart rate variability, can be analyzed using multilevel models to predict mental health outcomes like depression and anxiety effectively.	Multilevel models (MLMs) were used to predict the influence of smartphones and wearable data on mental health scores. Data from smartphone and wearable devices, including GPS, physical activity, sleep, and heart rate variability, were analyzed.	Delphi collected data from smartphone sensors: Battery, GPS, Screen, and Time zone. The AWARE framework is used for data collection and encryption for privacy.	High dropout rates in longitudinal observation studies. GPS data may not always be available or feasible
[27]	Individualized predictions of mental health outcomes can be achieved by integrating wearable device data with machine-learning models that analyze features like physical activity, sleep, and stress levels.	Longitudinal ecological momentary assessments, neurocognitive sampling, lifestyle data from wearables. Seven types of supervised machine learning approaches, ensemble learning, and regression-based methods.	Longitudinal ecological momentary assessments of depression. Neurocognitive sampling synchronized with electroencephalography and lifestyle data from wearables	Insufficient data for some participants affected model accuracy. Limited variability in data for specific subjects.
[28]	Machine learning algorithms can analyze wearable device data, such as activity levels and physiological metrics, to identify patterns indicative of mental health conditions, facilitating early detection and intervention.	Logistic Regression, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbor, and Naive-Bayes algorithms. Ensemble models created and compared using the proposed algorithms.	Kaggle dataset: 334 samples, 31 fields on unemployment and mental illness	Predicting mental illness remains a challenge. Medication hasn't fully cured or eradicated mental sickness.
[29]	Utilizing data from wearable devices, machine learning models like DNNs can analyze behavioral patterns to classify and predict mental health disorders, achieving high accuracy in diagnosis.	Utilizes commercially available WMSs and efficient DNN models. Uses synthetic data generation module to augment real data	Real data from 74 individuals was collected via sensors. Synthetic data is generated to augment real data.	Limited available data for training the DNN models. Need for synthetic data generation to augment real data.
[30]	ML can analyze physiological data from wearable devices to identify patterns and biomarkers, enabling predictions of mental health outcomes and	Classical and deep learning models for disease severity classification. Pre-processing of raw data from wearable device recordings.	Segments from two patient groups for model testing. Continuous physiological data from E4 Empatica wristbands.	A small sample size limits strong performance claims. The pipeline needs improvement for artifact detection and denoising.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

	treatment responses in mood disorders.			
[31]	Wearable device data can be analyzed using machine learning algorithms to identify patterns in EEG and HRV signals, enabling the prediction of mental health outcomes and stress levels.	Independent Component Analysis (ICA)Nonlinear Chaotic Analysis (NCA)	EEG signals for brain activity monitoring. Heart rate variability (HRV) for physiological assessment.	Large sensor system size limits mobile device implementation. Restricted Movement during measurement with a skin conductance sensor.
[32]	Predicting mental health outcomes relies on collecting physiological data from wearables and applying decentralized machine learning models. These models adjust to individual data patterns, enabling personalized tracking of mood and mental conditions.	Personal health device data collection Decentralized learning mechanism combining transfer and federated machine learning concepts	Popular mental health dataset evaluated for model performance. Patient physiological data from personal health devices was utilized.	Subjective patient descriptions and past medical history reliance A privacy-aware and accountable manner for mental health tracking.
[33]	It focuses on survey-based datasets for psychological instability prediction.	Machine learning: Random Forest Classifier, Multi-Layer Perceptron Classifier Deep learning: Artificial Neural Networks, Convolutional Neural Networks.	Real-time survey-based dataset with 1500 labeled items. Contains 38 attributes for stress detection.	A limited dataset size for training machine learning models. Lack of inclusion of face emotion recognition for prediction enhancement.
[34]	Physiological data from wearables, like heart rate and activity levels, can be analyzed using machine learning models to identify patterns indicative of depressive tendencies, guiding users toward professional help.	Analysis of physiological user data extracted from a Fitbit Alta HR device. Training of machine learning models to detect depressive tendencies	Physiological user data from Fitbit Alta HR device. A limited sample size of older people was analyzed.	Limited sample size increases the risk of model overfitting. Most predictive models performed poorly in detecting depressive tendencies.
[35]	Wearable device data, particularly heart rate variability, can be analyzed using machine learning algorithms to classify and predict mental health outcomes, such as depressive symptoms, based on physiological markers.	Machine learning algorithms Heart rate variability data analysis	2629 participants' HRV recordings from wearable devices. A training set of 1830 participants for machine learning.	Model performance is lower than expected (ROC AUC 0.56). Real-world variations impact HRV prediction accuracy.
[36]	Personalized deep learning models can predict mental health outcomes by analyzing multi-modal wearable data and participant assessments, optimizing parameters for accuracy, and identifying key features influencing mood changes.	Development of personalized and accurate deep learning models for depression. Use of SHAP, ALE, and Anchors from Explainable AI literature to extract indicators.	Multi-modal dataset with ecological momentary assessments of depression. Lifestyle data from wearables and neurocognitive assessments.	Current models lack personalized and accurate deep learning-based approaches.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

Focusing on social anxiety, Boukhechba et al., [24] utilized smartphone GPS data and machine learning to predict social anxiety symptoms. Their model achieved an accuracy of 85% in classifying high vs. low social anxiety days, demonstrating the potential of location data in mental health prediction.

Carreiro et al. [25] utilized wearable biosensors to detect drug use events and predict relapse risk. Their machine-learning approach achieved 93% accuracy in detecting opioid use events and 88% accuracy in predicting relapse within the next 24 hours.

III. PROPOSED SYSTEM

The suggested method predicts mental health outcomes using wearable device data and powerful machine learning techniques. The system follows a well-structured workflow, beginning with data collection, preparation, training, testing, and then evaluation using a voting classifier ensemble. Figure 2 depicts the system architecture of the proposed model.

A. Data Collection

The wearable devices collect physiological and behavioral data from users. This data includes heart rate, sleep patterns, physical activity levels, and other health indicators that are associated with mental health conditions. The wearable devices are connected to a cloud platform, where the data is stored in a structured format, ready for analysis.

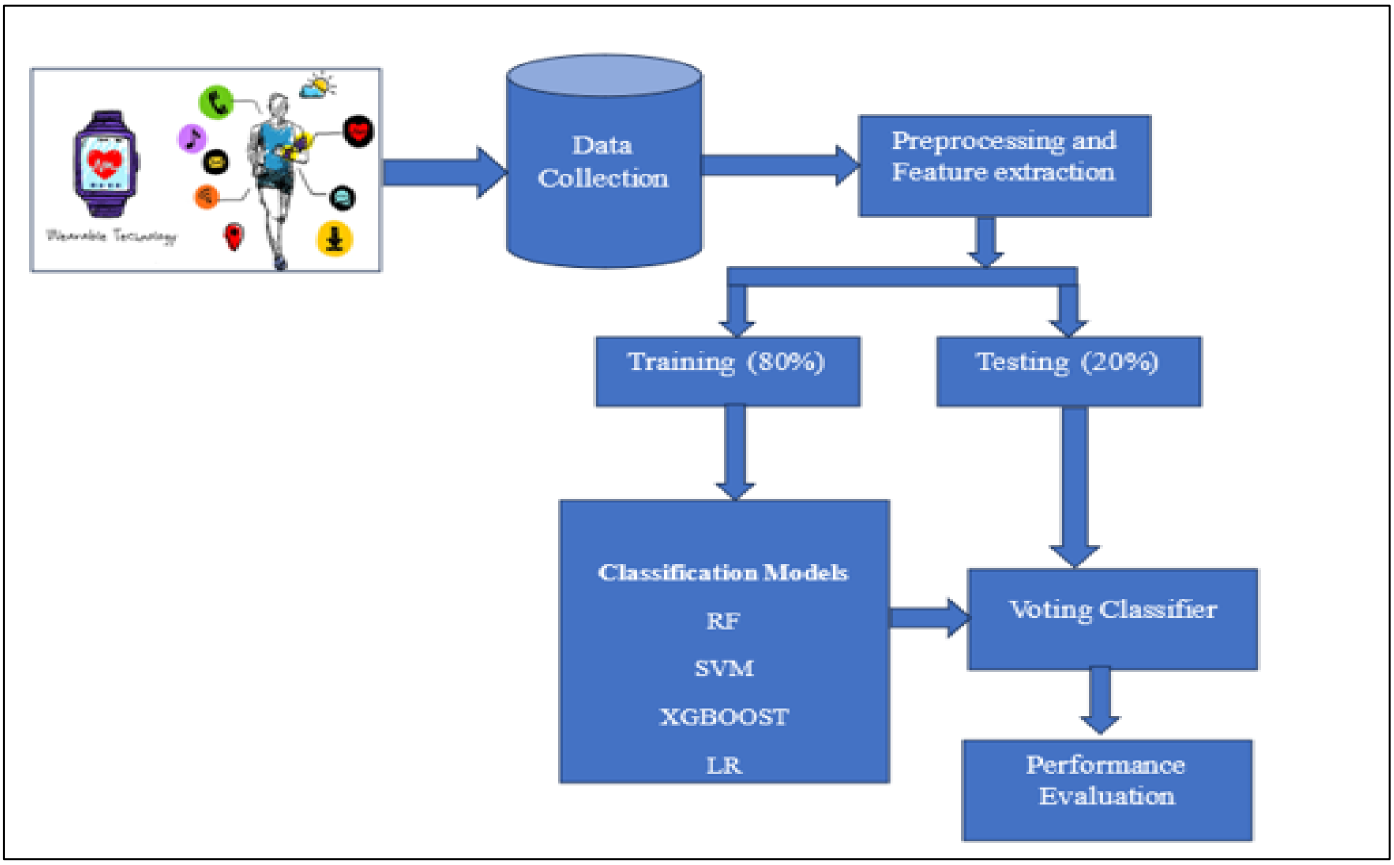

Fig 2: Proposed System Architecture

B. Preprocessing and Feature Extraction

The raw data collected is cleaned and preprocessed to remove noise and incomplete records. Feature extraction techniques are applied to derive relevant attributes such as heart rate variability, sleep duration, activity levels, and others. Normalization and transformation techniques are applied to standardize the data, ensuring that it is in a suitable form for machine learning models.

C. Training and Testing

The dataset is split into two parts: 80% for training and 20% for testing. The training dataset is fed into various classification models to learn patterns and relationships within the data that could indicate mental health issues. The models selected for this system are Random Forest (RF), Support Vector Machine (SVM), XGBoost, and Logistic Regression (LR). Each model is trained separately on the training data.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

D. Classification Models

- **Random Forest (RF)**: A robust model that operates by constructing a multitude of decision trees during training and outputs the class with the highest occurrence. It is well-suited for handling large datasets and provides good accuracy.
- **Support Vector Machine (SVM)**: This model works by finding the hyperplane that best separates the classes. It is highly effective in high-dimensional spaces and works well with limited but clean data.
- **XGBoost**: A powerful boosting algorithm known for its high performance on structured data. It helps in enhancing model accuracy through iterative corrections of weak predictions.
- **Logistic Regression (LR)**: A simple yet effective classification model that is often used as a baseline for binary classification problems. It models the probability that a given instance belongs to a particular class.

E. Voting Classifier (Ensemble Learning):

The outputs of the individual classifiers are combined using a majority voting mechanism, forming an ensemble model. The Voting Classifier takes advantage of the strengths of each model, improving the overall prediction accuracy. In this way, the system achieves higher robustness and reduces the likelihood of incorrect predictions by relying on consensus among the classifiers.

F. Performance Evaluation:

After training, the system is tested on unseen test data. Performance evaluation is conducted using a variety of metrics including accuracy, precision, recall, F1-score, and ROC-AUC curve. The results are compared to determine the most effective model, and the overall system's effectiveness is assessed based on how well the ensemble model performs.

IV. EXPECTED RESULTS

The proposed system is expected to achieve high accuracy in predicting mental health outcomes based on the wearable device data. By combining multiple machine learning models into an ensemble through the Voting Classifier, the system should improve prediction reliability and accuracy over individual models. It is anticipated that the model will:

- Achieve an accuracy of over 90% in identifying individuals with mental health issues.
- Provide early warnings and continuous monitoring capabilities, allowing for proactive interventions.
- Perform well in detecting patterns in physiological and behavioral data that are indicative of mental health conditions.
- Be scalable and adaptable to different wearable devices and mental health conditions.

By Employing wearable technology and machine learning, this system has the potential to transform how mental health conditions are monitored and predicted, offering a real-time, data-driven approach to mental health care.

V. CONCLUSION & FUTURE SCOPE

In this paper, we have proposed a system for predicting mental health outcomes using wearable device data and machine learning techniques. Although the system is currently in the design phase and has not been implemented, it is expected to provide an accurate and reliable method for identifying potential mental health issues. By incorporating multiple classification models into an ensemble through a voting classifier, the system aims to achieve higher accuracy compared to individual models. The expected results point toward the potential of wearable data to enable real-time mental health monitoring, which could lead to earlier detection and proactive interventions. This system, once developed, has the potential to contribute significantly to mental health care by offering continuous and non-invasive monitoring solutions.

Since the proposed system has not yet been implemented, future work will focus on the actual development and validation of the system using real-world data. Once implemented, the system's performance can be rigorously tested, and improvements can be made by refining the machine learning models or introducing additional data types. Incorporating deep learning techniques or exploring multimodal approaches that combine physiological, behavioral, and environmental data could further enhance prediction accuracy. Future research may also explore the personalization of mental health predictions for specific conditions like depression, anxiety, or stress. Additionally, addressing the challenges of scalability across different wearable devices and user populations will be critical for widespread adoption. Ethical considerations, including data privacy and security, will also need to be carefully managed to ensure the responsible and safe use of sensitive health data in real-world applications.

REFERENCES

[1]. Vigo, Daniel, et al. "Estimating the true global burden of mental illness." The lancet. Psychiatry vol. 3,2 (2016): 171-8. doi:10.1016/S2215-0366(15)00505-2

[2]. Kessler, Ronald C., et al. "Development of lifetime comorbidity in the World Health Organization world mental health surveys." Archives of General Psychiatry 68.1 (2011): 90-100.

[3]. Myin-Germeys, Inez, et al. "Experience sampling methodology in mental health research: new insights and technical developments." World psychiatry: official journal of the World Psychiatric Association (WPA) vol. 17,2 (2018): 123-132. doi:10.1002/wps.20513

[4]. Vashistha, Rajat, et al. "Futuristic biosensors for cardiac health care: an artificial intelligence approach." 3 Biotech 8 (2018): 1-11.

[5]. Peake, Jonathan M et al. “A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations.” Frontiers in physiology vol. 9 743. 28 Jun. 2018, doi:10.3389/fphys.2018.00743
[6]. Mohr, David C et al. “Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning.” Annual review of clinical psychology vol. 13 (2017): 23-47. doi:10.1146/annurev-clinpsy-032816-044949
[7]. Can, Yekta & Arnrich, Bert & Ersoy, Cem. “Stress Detection in Daily Life Scenarios Using Smart Phones and Wearable Sensors: A Survey.” Journal of Biomedical Informatics. 92. 103139. 10.1016/j.jbi.2019.103139.
[8]. Grünerbl, Agnes et al. “Smartphone-based recognition of states and state changes in bipolar disorder patients.” IEEE journal of biomedical and health informatics vol. 19,1 (2015): 140-8. doi:10.1109/JBHI.2014.2343154
[9]. Jacobson, Nicholas C et al. “Digital biomarkers of mood disorders and symptom change.” NPJ digital medicine vol. 2 3. 1 Feb. 2019, doi:10.1038/s41746-019-0078-0
[10]. Torous, John, and Matcheri Keshavan. “A new window into psychosis: The rise digital phenotyping, smartphone assessment, and mobile monitoring.” Schizophrenia research vol. 197 (2018): 67-68. doi:10.1016/j.schres.2018.01.005
[11]. Baig, Mirza Mansoor, et al. "A systematic review of wearable patient monitoring systems–current challenges and opportunities for clinical adoption." Journal of medical systems 41 (2017): 1-9.
[12]. Shatte, Adrian B R et al. “Machine learning in mental health: a scoping review of methods and applications.” Psychological medicine vol. 49,9 (2019): 1426-1448. doi:10.1017/S0033291719000151
[13]. Durstewitz, D., Koppe, G. & Meyer-Lindenberg, A. Deep neural networks in psychiatry. Mol Psychiatry 24, 1583–1598 (2019). https://doi.org/10.1038/s41380-019-0365-9
[14]. Beutel, Manfred E et al. “Childhood adversities and distress - The role of resilience in a representative sample.” PloS one vol. 12,3 e0173826. 15 Mar. 2017, doi:10.1371/journal.pone.0173826
[15]. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1, 206–215 (2019). https://doi.org/10.1038/s42256-019-0048-x
[16]. Cearns, Micah et al. “Recommendations and future directions for supervised machine learning in psychiatry.” Translational psychiatry vol. 9,1 271. 22 Oct. 2019, doi:10.1038/s41398-019-0607-2
[17]. Dudley, John J., and Per Ola Kristensson. "A review of user interface design for interactive machine learning." ACM Transactions on Interactive Intelligent Systems (TiiS) 8.2 (2018): 1-37.
[18]. Sano, Akane et al. “Identifying Objective Physiological Markers and Modifiable Behaviors for Self-Reported Stress and Mental Health Status Using Wearable Sensors and Mobile Phones: Observational Study.” Journal of medical Internet research vol. 20,6 e210. 8 Jun. 2018, doi:10.2196/jmir.9410
[19]. Jacobson, Nicholas C et al. “Digital biomarkers of mood disorders and symptom change.” NPJ digital medicine vol. 2 3. 1 Feb. 2019, doi:10.1038/s41746-019-0078-0
[20]. Torous, J., Staples, P., Barnett, I. et al. “Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia.” npj Digital Med 1, 15 (2018). https://doi.org/10.1038/s41746-018-0022-8
[21]. Giannakakis, Giorgos, et al. "Stress and anxiety detection using facial cues from videos." Biomedical Signal Processing and Control 31 (2017): 89-101.
[22]. Kleiman, Evan M et al. “Digital phenotyping of suicidal thoughts.” Depression and anxiety vol. 35,7 (2018): 601-608. doi:10.1002/da.22730
[23]. Rui Wang et al., “Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing.” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 1, Article 43 (March 2018), 26 pages. https://doi.org/10.1145/3191775
[24]. [24] Boukhechba, Mehdi et al. “Predicting Social Anxiety From Global Positioning System Traces of College Students: Feasibility Study.” JMIR mental health vol. 5,3 e10101. 4 Jul. 2018, doi:10.2196/10101
[25]. Carreiro, Stephanie et al. “Real-time mobile detection of drug use with wearable biosensors: a pilot study.” Journal of medical toxicology: official journal of the American College of Medical Toxicology vol. 11,1 (2015): 73-9. doi:10.1007/s13181-014-0439-7
[26]. Elhai, Jon D., et al. "Depression and emotion regulation predict objective smartphone use measured over one week." Personality and individual differences 133 (2018): 21-28
[27]. Taylor, Sara, et al. "Personalized multitask learning for predicting tomorrow's mood, stress, and health." IEEE Transactions on Affective Computing 11.2 (2017): 200-213.
[28]. Garcia-Ceja, Enrique, et al. "Mental health monitoring with multimodal sensing and machine learning: A survey." Pervasive and Mobile Computing 51 (2018): 1-26.
[29]. Mustanski, Brian S., Robert Garofalo, and Erin M. Emerson. "Mental health disorders, psychological distress, and suicidality in a diverse sample of lesbian, gay, bisexual, and transgender youths." American journal of public health 100.12 (2010): 2426-2432.
[30]. Ravi, Daniele, et al. "A deep learning approach to on-node sensor data analytics for mobile or wearable devices." IEEE journal of biomedical and health informatics 21.1 (2016): 56-64.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT21MAR587

[31]. Roh, Taehwan, et al. "Wearable mental-health monitoring platform with independent component analysis and nonlinear chaotic analysis." Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference vol. 2012 (2012): 4541-4. doi:10.1109/EMBC.2012.6346977

[32]. Bedor Hiland, Emma. "The Digital Transformation of Mental Health ." (2018)

[33]. Joshi, Deepali J., et al. "Mental health analysis using deep learning for feature extraction." Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. 2018.

[34]. Liu, Lili, et al. "Smart homes and home health monitoring technologies for older adults: A systematic review." International journal of medical informatics 91 (2016): 44-59.

[35]. Valenza, Gaetano, et al. "Characterization of depressive states in bipolar patients using wearable textile technology and instantaneous heart rate variability assessment." IEEE journal of biomedical and health informatics 19.1 (2014): 263-274.

[36]. Palmius, Niclas. Personalised modelling of geographic movements in depression. Diss. University of Oxford, 2018.