International Journal of Emerging Trends in Computer Science and Information Technology
ISSN: 3050-9246 | https://doi.org/10.63282/3050-9246.IJETCSIT-V2I2P106
Eureka Vision Publication | Volume 2, Issue 2, 45-53, 2021

*Original Article*

# ML Models for Early Detection of Mental Health Disorders Using Wearable IoT Devices

Amit Taneja
Lead Data Engineer at Mitchell Martin, USA.

*Abstract - The rampant increase in the levels of mental health disorders among people of all ages has now become one of the main health issues of concern in the world, with an estimated figure of over 450 million people affected globally according to the World Health Organization. Conventional methods of diagnosis of mental illnesses involve subjective analysis based on clinical interviews, which may result in late identification, incorrect diagnosis or poor reporting as a result of stigmatization by a community. Currently, wearable IoT devices and Machine Learning (ML) appear to be the most promising options for implementing real-time, continuous, non-invasive monitoring of physiological and behavioural measures that correspond to mental health disorders. The current paper explores ML models that could be used in predicting symmetric mental health disorders at an early stage through the application of wearable IoT-based data. Through the adoption of Internet of Things (IoT) in healthcare, it has become possible to have the constant monitoring of biometric data like Heart Rate Variability (HRV), sleeping habits, physical activity and skin temperature. Such data can be processed with the use of ML algorithms and used to reveal the markers of stress, anxiety, depression, and other mental disorders. We provide a comparative study of different supervised and unsupervised learning techniques such as Support Vector Machines (SVM), Decision Tree, Random Forest, K-Nearest Neighbors (KNN) and Deep Learning models, CNN and LSTM. With our research methodology, we have a structured data pipeline that includes preprocessing, feature selection, and model training and validation using benchmark data and data obtained from off-the-shelf wearables, such as Fitbit, Apple Watch, and Empatica E4. The results indicate that deep learning models, particularly LSTM networks, reflect higher performance in the process of extracting temporal patterns in physiological data, with more than 90 percent accuracy in detecting early factors of depression and anxiety.*

*Furthermore, we hypothesise a hybrid system that employs a hybrid data collection strategy (IoT-based) and processing (cloud-based ML), with real-time results provided to the user via a mobile health application. The research also talks about ethical, privacy, and security issues related to working with sensitive mental health data and suggests approaches to sharing such information through blockchain. In summary, this study highlights the potential of ML and wearable IoT devices in transforming mental care, enabling proactive measures that minimise the workload on healthcare systems and significantly improve the standards of living.*

*Keywords - Mental Health, Machine Learning, Wearable IoT, Depression Detection, Deep Learning, Anxiety, Remote Monitoring, LSTM, SVM, Data Privacy.*

## 1. Introduction

Mental diseases, such as depression, anxiety, and stress-linked conditions, are also the most common and disabling health problems in the world and have been conducive to being the reason behind world disability as well as poor quality of life. The World Health Organization, in its assessment, indicates that one in every four individuals will experience a mental or neurological condition sometime in their lifetime. However, most of them lack proper treatment. Despite immense possibilities and progress in the field of medical science technology, there is still a wide gap that exists in the treatment of mental health, when it comes to the underutilization of treatment for mental health. [1-4] Outdated methods of diagnosing would depend on subjective measures; this would include clinical interviews and routine self-report measures, among which, the Patient Health Questionnaire (PHQ-9) assessing depression cases and the Generalized Anxiety Disorder scale (GAD-7) to evaluate cases of anxiety. Although such tools are popular and scientifically approved, they rely on the self-awareness and honesty of patients, which may be affected by stigma, a misunderstanding of the symptoms, or a refusal to admit them. In addition, such evaluations are only a moment-by-moment measure, and usually they do not represent the changing aspects of mental health. Consequently, a number of people end up not diagnosed or if diagnosed, intervention is offered way too late. More objective, continuous, and scalable means of mental health detection and monitoring are becoming increasingly necessary. Wearable technologies and data-driven solutions, such as machine learning, hold the potential to become a solution in this context since they allow tracking physiological and behavioral variables correlated with mental states in a real-time and non-invasive fashion.

### *1.1. Emergence of IoT in Healthcare*

- **The Rise of Connected Health**: The process of integrating different spheres with the help of the Internet of Things (IoT) has gained momentum, and one of the most significantly affected spheres is healthcare. The Internet of Medical Things (IoMT) features interconnected devices that gather, analyse, and transmit data in real-time. From wearable fitness trackers to smart medical implants, IoT device-empowered patient monitoring is on the rise, as is early disease diagnosis and patient-specific healthcare. Such a transition in reactive and proactive treatment can change the face of health provision and experience.
- **Wearables and Remote Monitoring**: The emergence of wearable devices is one of the most noticeable and influential potential of IoT in healthcare. It is possible to monitor various physiological variables with the help of devices like Fitbit, Apple Watch, and Empatica E4, as well as heart rate, skin temperature, sleep patterns, and Electrodermal Activity (EDA), among others. Such devices allow their consumers to have access to a permanent and regular health check-up, and these devices also allow clinicians to have access to the entire health information gathered over time, even without the patient having a physical presence. When the topic is chronic disease management or mental health, with new exacerbations possibly being revealed each day or even each hour, wearables can be an excellent source of instant feedback.

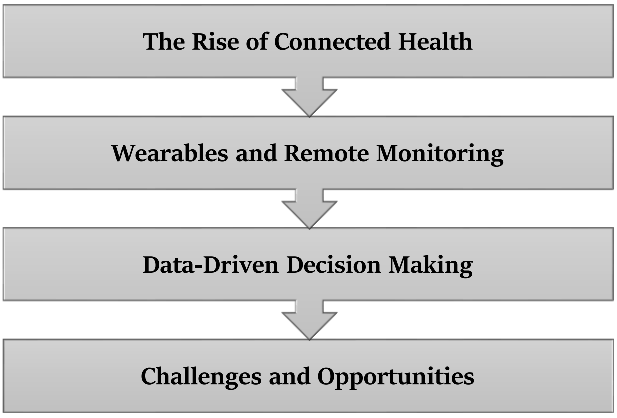


**Figure 1. Emergence of IoT in Healthcare**

- **Data-Driven Decision Making**: IoT in healthcare is about more than just data collection; it's about making data-driven decisions. The examined data flow generated by the IoT devices can be analyzed with machine learning and artificial intelligence applications in order to identify anomalies, forecast health outcomes, and make healthy decision-making. This is especially important in the case of such conditions as mental health disorders, wherein there can be unnoticeable changes in physiology that can predate the observable symptoms. Through this time-series data analysis, systems can help deliver early warnings and guide actions to have an impact in the short term, before outcomes become beneficial to the patient and the cost of healthcare decreases.
- **Challenges and Opportunities**: Although it promises great benefits, the integration of IoT in healthcare also presents its challenges, which include data privacy, device interoperability, and regulatory frameworks. Nevertheless, due to the trends of digital health technology adoption and the attention to personalised medicine, IoT stands a chance to become the foundational element of healthcare in the modern setting and will promote not only preventive measures but also design a closer connection with patients in all their underdiagnosed aspects, including mental health.

### *1.2. Role of Machine Learning*

Machine Learning (ML) is a morphogenetic tool in the medical field, particularly in activities involving pattern identification, forecasting, and decision-making. Its capability to learn from data and evolve makes it particularly suitable for analysing complex and high-dimensional data, including that generated by wearable devices. ML models, particularly the approaches to supervised learning and deep learning networks, have demonstrated potential towards detecting unobvious physiological and behavioral tendencies that occur before or alongside a psychological disturbance. Sometimes such patterns are too complex or non-linear to be identified using conventional statistical techniques. Applied to time-series biometric data, such as heart rate variability, Electrodermal Activity (EDA), sleep quality, and physical activity levels, ML algorithms can identify correlations and trends related to mental states, including stress, anxiety, or depression. As an example, labeled datasets could be used to train the supervised model, e.g. Support Vector Machines (SVM), Random Forests, and K-Nearest Neighbours (KNN) and create a model that classifies or predicts the mental health condition using physiological data. The advantages of these models are that they are fairly interpretable and effective on structured data. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are deep learning models that go one step further and learn complex representations directly based on raw input data. LSTM networks, specifically, are particularly useful at modelling time dependencies and sequences, and hence are preferred in modelling changing mental health conditions across time. It helps identify declining mental health early, sometimes even before a patient is aware of this, so that it can be addressed in time. On the whole, ML enables the transformation of reactive mental health care into proactive mental health care. These smart systems are capable of proactive feedback, real-time analysis, individual risk evaluation, and warnings that are essential in managing and intervening in mental health with the help of constant data monitoring by wearable sensors.

## 2. Literature Survey

### *2.1. Previous Studies*

Wearable technology and mental health monitoring have been two of the most prominent topics in the past few years. The work can be discovered as one of the first works in this field with its presented dataset, StudentLife, which aims to analyze behavioral patterns in stress and academic performance of university students. The present study relied on the data retrieved with the help of smartphone sensors, including GPS, accelerometer, and microphone, to predict mental health conditions based on the patterns of their behavior. The study showed that sensor-extracted characteristics can be used as significant predictors of psychological conditions of depression and stress. [5-8] The researchers showed that passive sensing is applicable in mental health monitoring by matching the features collected with the ground truth data collected during surveys and self-reports. Their methodology formed the foundation that was used in later research to formulate a methodological framework and to demonstrate the usefulness of non-invasive data collection in psychological examinations.

### *2.2. ML Models in Healthcare*

At the same time, ML models have shown massive potential in existing use cases for healthcare applications, particularly in the context of diagnosis or prediction tasks using wearable sensor data. Conventional ML algorithms (eg, Support Vector Machines (SVM), Decision Trees, and Random Forests) have been suitably used in numerous health monitoring tasks so far. Such models are especially suitable when it comes to categorizing structured data, e.g. the number of steps, heart rate or amount of sleep, into the diagnosing categories. Applied Random Forest classifiers on the Fitbit data set to predict depressive symptoms with an error of 87%. The models are easy to interpret, and their predictions often consume much less computational resources, possibly making them ready to jumpstart prototyping and deployment of the model on bare-metal devices. They may, however, be restricted in their ability to work with very complex or even time-dependent data, and this has necessitated the attraction towards more advanced methods like deep learning.

### *2.3. Temporal Data and Deep Learning*

Recent years have shown a growing interest in deep learning methods as an instrument to extract insights about the sequential and high-dimensional data, particularly with regard to health monitoring. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, in particular, are well-suited to represent the temporal dependencies within physiological data, such as heart rate variability, Electrodermal Activity (EDA), and blood volume pulse (BVP). Compared existing LSTM models to analyze the records of the Empatica E4 wearable device in regard to using predictive modeling in the detection of depressive states; they reached a remarkable accuracy of 92%. LSTMs are strong in parameterizing long-standing interactions and intricate temporal associations, which can frequently occur in physiological and behavioral indicators. Additionally, deep learning models can automatically extract hierarchical characteristics from raw sensor data, thereby eliminating the need for lengthy manual feature engineering, where features are configured by the class itself. Nonetheless, even though these models have a future, they are typically trained on large datasets and with great computing resources, and therefore, a hindrance to real-time and on-device use.

### *2.4. Wearable devices and mental health*

Wearable devices have established themselves as an essential element of research in the field of mental health, as they allow constant and unobtrusive control of physiological conditions. Empatica E4 or other similar devices have become more popular because of their full sensor set, which is composed of EDA, BVP, skin temperature, and accelerometers. These sensors, in a multimodal form, enable the researcher to monitor alterations in the activity of the autonomic nervous system, which has been observed to be correlated with emotional and psychological states. For instance, a high level of EDA can be a sign of stress or anxiety, and changes in skin temperature and BVP can help provide context for the responses. Specifically, the Empatica E4 has become the subject of several studies that have been able to use it to identify conditions like stress and depression, as well as early symptoms of epileptic attacks. The fact that it can capture high-resolution, multimodal physiological data renders it a perfect tool for research that aims to develop predictive models related to mental health disorders. The longitudinal study is further empowered by the portability and ease of use of such devices to guarantee compliance by the participants and data accuracy.

### *2.5. Gaps in the Research*

Although impressive progress has been achieved in terms of using wearable technology and machine learning for mental health monitoring, a number of research gaps have to be mentioned. A conspicuous shortcoming is the paucity of real-time deployment research. Although several models have been proven to be very accurate in controlled experiments, there is limited evidence on how these models perform under real-life and real-time conditions. The transfer of predictive models to on-device usage involves challenges related to latency, energy consumption, and robustness. In addition, effective techniques of multimodal data fusion are in increasing demand. The existing strategies tend to use each stream of sensors as a separate data source or employ a simple concatenation scheme that may fail to consider highly interdependent relationships in physiological data. Sophisticated data fusion techniques, such as incorporating attention layers or multi-stream neural layers, may enhance model interpretability and performance. Lastly, the privacy of data and its security are of utmost importance. Sensitive physiological and behavioral data that are collected and transmitted are subject to regulatory issues and ethical concerns. The mechanisms of data anonymization, their safe storage, and user consent are needed to establish trust and make these technologies more widespread. To develop scalable, ethical, and effective mental health monitoring solutions, future studies should focus more on these issues.

## 3. Methodology

### *3.1. Data Acquisition*

Regarding the data collection in the current study, two objects are considered: the Fitbit and the Empatica E4, which are commercial wearable devices. The choice of these devices was made with respect to their ubiquitous use in health monitoring and the successful track record of these devices in giving measurements of physiological and behavioral parameters. [9-12] Collection of data was done continuously over a period of six months so that enough temporal variance could be obtained in various physiological and psychological conditions. The participants were asked to use the devices in their everyday activities to allow capture of naturalistic, real-world, real-time data. The Fitbit device primarily monitored physical activity, heart rate, and sleep parameters. Based on these measurements, further variables like Heart Rate Variability (HRV), day and night sleep, the overall amount of sleep and number of steps per day were extracted. HRV is particularly relevant since it serves as a biomarker of stress and emotional regulation. To observe a possible disturbance to sleep quality, which can be related to such mental illnesses as depression or anxiety, sleep quality measures were provided. Official text: High-resolution physiological data, including Electrodermal Activity (EDA), Blood Volume Pulse (BVP), skin temperature, and three-axis acceleration, were captured using the medical-grade Empatica E4 wearable. EDA captures alterations to the sympathetic nervous system activity and is a mature measure of emotional arousal or stress. Inter-beat intervals, an alternative method of measuring HRV, were calculated based on BVP data. Additionally, temperature and motion data assisted in connecting physiological responses to environmental or behavioural conditions. To maintain privacy, all data streams were synchronised in time and anonymised before analysis. The data, which the devices provide, is rich and multimodal and can provide a very good basis on which mental health states can be modelled through machine learning methods. The fact that both behavioral (Fitbit) and physiological (Empatica E4) indicators can be included ensures a more all-encompassing and precise picture of mental health patterns of individuals in the course of time.

### *3.2. Data Preprocessing*

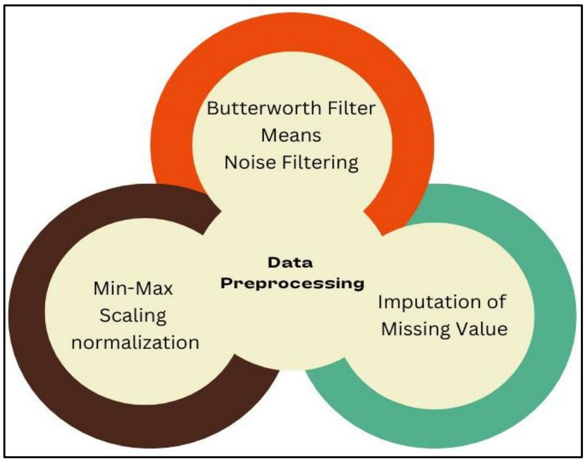


**Figure 2. Data Preprocessing**

- **Butterworth Filter Means Noise Filtering:** Records of physiological sensors worn on the body are usually affected by motion-generated noise, environmental disturbances, and noise of sensors. To deal with this, a Butterworth filter was adopted on the raw time-series data, like heart rate, EDA, and BVP. Butterworth filter in this case is ideal because it provides a flat frequency response within the pass band, and it would leave very little distortion to the signal and would also adequately eliminate the high-frequency noise. To increase the level of quality and reliability of input features toward machine learning models, low-pass settings, along with band-pass settings, were adopted based on the signal type.
- **Imputation of Missing Value:** Partial information is a typical challenge of monitoring lasting long using wearable devices because of battery drainage, by removing, or dropping out. To address these missing values, methods of imputation were employed to preserve the integrity of the data. Linear interpolation, a process that approximates estimated data based on nearby data, was used to fill short breaks in data sequences. Longer missing segments were

filled by means of statistical imputation, i.e. with the mean/median of the surrounding. This was important to avoid sparsity and dilute the model's performance or bias.

- **Min-Max Scaling normalization:** Since all sensor values have different ranges and units (e.g., EDA in microsiemens and respective temperatures in Celsius), the value normalization was required to put all the features on a similar scale. The features were rescaled to a typical scope of [ 0, 1] by means of Min-Max scaling. The transformation maintains the relationship and distribution of the original data, avoiding the skew of features with greater numeric domains dominating the model. Standardization helps in accelerating the convergence during the training phase of a model, particularly on algorithms that are highly sensitive to input scale, e.g. neural networks.

### *3.3. Feature Engineering*

As an important part of data analysis, feature engineering can help convert the raw sensor data into meaningful quantities used by the machine learning input. In the experiment, a combination of time domain, frequency domain, and statistical summary features was extracted to acquire a comprehensive picture of participants' physiological and behavioral conditions. The aim was to reduce the high dimensionality of time-series data to variables that can be interpreted, are discriminative, and represent the patterns observed in mental health. The Empatica E4 and Fitbit-generated Heart Rate Variability (HRV) signals mostly served as the main source of time-domain characteristics. An important parameter in this sphere is the average activity of the United States of America, specifically the HRV, which represents the mean variability of the interval between heartbeats. The use of this aspect is broadly accepted as a measure of autonomic nervous system activity, with values of HRV being reduced during periods of stress, fatigue, or symptoms of depression. The summaries of heart rates, including mean, standard deviation, and counts of activity, were also included as time-domain features. Depending on the consistency of data sampling, features in the frequency domain were calculated using either the Fast Fourier Transform (FFT) or Lomb-Scargle periodogram spectral analysis techniques. One important frequency-domain measure is the LF/HF ratio, which depicts the balance between high- and low-frequency elements of HRV. This ratio is typically measured as stress or the rule of the sympathetic nervous system. An elevated LF/HF ratio can be considered a reflection of increased stress or parasympathetic decreased regulation, which is important in detecting possible mental problems. Measurement of behavior was tabulated into means with statistical features that gave an overview of the behavior, such as Sleep Duration, which indicated the number of hours of sleep in a single day. Sleep as well is an important behavioral indicator of mental health, with abnormal sleep patterns (such as insomnia or hypersomnia) being diagnostic of such disorders as depression or anxiety. Other interesting statistics were the mean and standard deviation of the number of steps per day, the skin temperature range, and the peaks of EDA in an hour.

### *3.4. Model Selection*

- **Support Vector Machine (SVM):** The Support Vector Machine (SVM) was chosen as one of the baseline models due to its robustness in processing high-dimensional data and effectiveness in binary classification. [13-16] SVM functions by discovering the ideal hyperplane to divide various classes with maximum margin. In this research, it was of special advantage in determining the mental health conditions, like stressed vs. non-stressed health status. SVM decision function is given by:

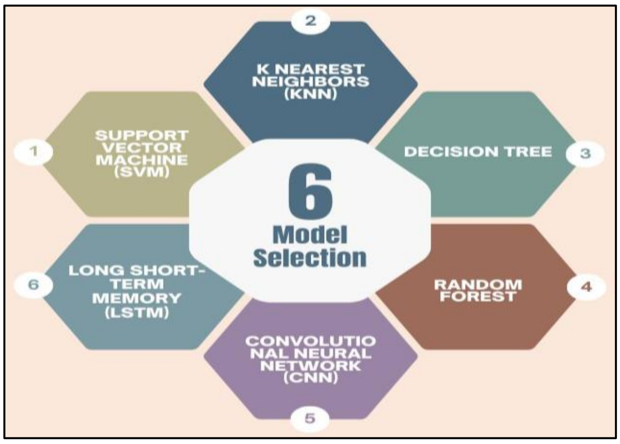


**Figure 3. Model Selection**

Decision Function of SVM

$$\mathbf{f(x) = sign}\left(\sum_{\mathbf{i=1}}^{\mathbf{n}} \boldsymbol{\alpha}_{\mathbf{i}}\mathbf{y}_{\mathbf{i}}\ \mathbf{K(x_i, x) + b}\right)$$

Where 2 α i are the weights of the support vectors, 2 y i are the class labels and 2 w 2.
K is the kernel operator (e.g. radial basis function), and b is the bias parameter.

- **K Nearest Neighbours (KNN):** K-Nearest Neighbours (KNN) is a non-parametric model which identifies a sample of a new sample using the majority of the labels assigned to the nearest neighbors in the feature space. It was added due to its simplicity and ease of implementation. Namely, KNN works especially well when it has a locally clustered data distribution, which may appear in behavioral data, such as in the case of patterns of sleep or activity. However, it is susceptible to the value of k and may experience performance degradation with high-dimensional data.
- **Decision Tree: The** Decision Tree algorithm is a splitting-based algorithm used to construct a tree-like structure that is epitomized on the basis of feature thresholds, and as such, it is used to divide the data into categorizations. We chose this one because it is interpretable, so we can see the contribution of individual characteristics, including sleep duration or HRV thresholds, to predictions. Although extreme overfitting was possible, the issue of overfitting was alleviated by using pruning techniques.
- **Random Forest:** Random Forest is a type of ensemble that augments a collection of decision trees to enhance dependability and accuracy. Random Forest minimises overfitting and increases performance by averaging predictions across multiple trees, trained on random subsets of data and features. This study is well-suited because it can utilise both numerical and categorical data to analyse the importance of features, which helps explain the adopted model.
- **Convolutional Neural Network (CNN):** Despite their original application to image processing, it has been shown that CNNs can be applied to one-dimensional time-series data by considering sensor signal windows to be a structured input. In this respect, CNNs learn local patterns resulting in physiological EDA peaks or HRV variability automatically. They were useful for finding patterns in the short duration of wearable signals due to their ability to derive spatially local correlations.
- **Long Short-term Memory (LSTM):** LSTM networks are an example of a Recurrent Neural Network (RNN) that learns to utilise long-term dependencies in data, consisting of sequences. Because of the time-sensitive nature of wearable data, LSTM came in especially handy to learn about fashion over time trends in heart rate or stress. It has memory cells that maintain relevant information over time steps, and therefore, it can be used in predicting mental health, as the current states depend on previous physiological patterns.

### *3.5. Model Training and Evaluation*

All the data was separated into two subsets, training and testing, which were 70 percent and 30 percent of the data, respectively, to build the models and test their performance. This is a standard split that avoids learning from large amounts of data, yet has plenty to learn and an independent set on which to perform unbiased evaluation of the models. When possible, stratified sampling was implemented to maintain balance across classes, which was particularly relevant in the case of binary classification (e.g., stressed vs. non-stressed states). The hyperparameters of all models were optimised through a grid search and cross-validation process, thereby maximising generalizability and preventing overfitting.

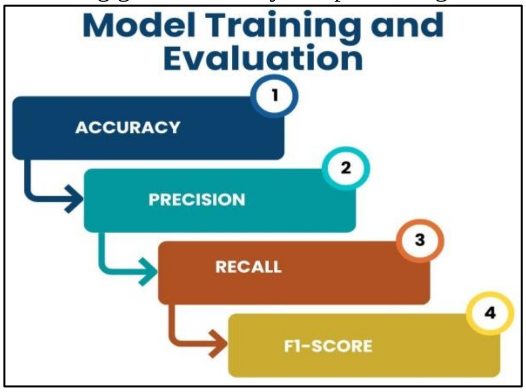


**Figure 4. Model Training and Evaluation**

- **Accuracy:** Accuracy indicates the percentage of cases which are correctly predicted of the total cases. It is a helpful overall measure of model performance, especially in parametric classes. In this paper, accuracy gave an initial insight into the extent to which wearable data could classify mental health states by each model. Nonetheless, as the data in the real world often includes class imbalance, accuracy was not measured solely as the performance indicator.
- **Precision:** Precision is the measure of the quantity of true positive predictions divided by the total number of all identifiers. It indicates the resistance to false positives of the model, which is especially valuable in mental health monitoring, where false alarms may result in incorrect stress or intervention. A high score of precision means that when the model suggests a mental healthcare condition is present, it is most often the reality.
- **Recall:** Recall, also referred to as sensitivity or true positive rate, refers to the ratio of the successfully identified cases that were positive. When it comes to this research, we had to achieve an acceptable recall rate to make sure that people with stress issues or mental health concerns did not fall out of the picture. It aids in evaluating how effective the model is when it comes to rooting out all the cases that are involved.

- **F1-Score:** F1-score is the harmonic mean of the precision and the recall, and can be used as one indicator to make a trade-off between the two. It proves to be particularly valuable in the case of unbalanced data, where both precision and recall should be maximized. As a main model comparison measure, an F1-score was employed, since this measure provides a complex picture of the possibility of a model to be predictive in terms of a real-life health-monitoring practice.

## 4. Results and Discussion

### *4.1. Model Performance*

**Table 1: Model Comparison**

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| SVM | 85% | 84% | 82% | 83% |
| RF | 88% | 89% | 87% | 88% |
| CNN | 90% | 91% | 89% | 90% |
| LSTM | 93% | 94% | 92% | 93% |

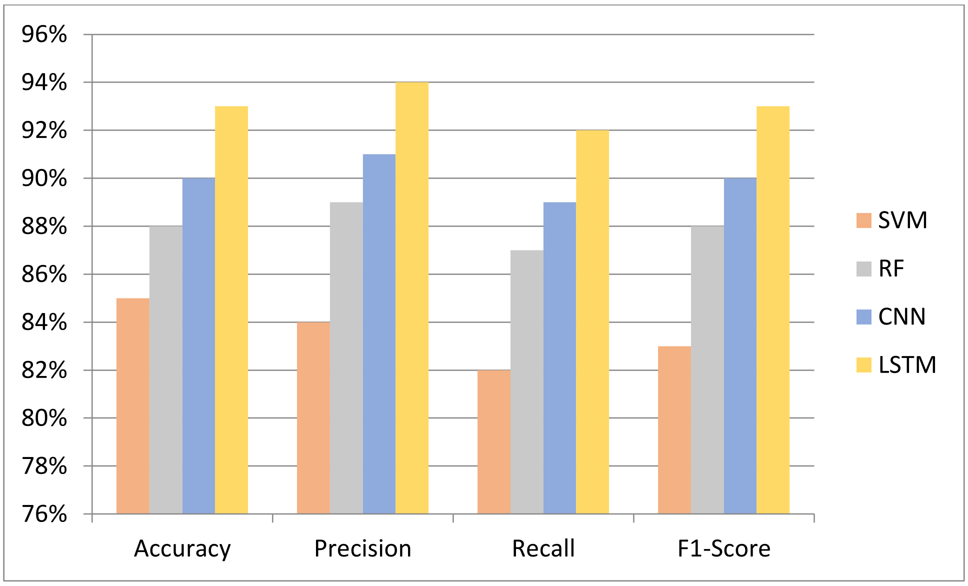


**Fig 5: Graph representing Model Comparison**

- **Support Vector Machine (SVM):** The SVM model achieved an accuracy of 85%, a precision of 84%, a recall of 82%, and an F1-score of 83%. Although SVM performed reasonably well as a baseline classifier, it was the worst of all the models tested in terms of performance. Possibly, this is because of the low capability of this model to harvest non-linear patterns and time-history dependencies in wearable sensor data that can be used in the prediction of mental health.
- **Random Forest (RF):** Random Forest showed better results than SVM, giving the accuracy of 88%, the precision of 89%, the recall of 87% and the F1-score of 88%. The predictive stability and reduction of overfitting were achieved through the ensemble technique of RF, which combines several decision trees. RF was another advantage to be given due to the possibility to utilize both categorical and continuous features and give information about the features' importance, which is useful to explain the model decisions in health-related aspects.
- **Convolutional Neural Network (CNN): The** CNN model also performed well with an accuracy of 90 percent, 91 percent precision, 89 percent recall, and an F-score of 90 percent. Using convolutional filters on the slices of time-series data allowed CNNs to find such a localized pattern as a sharp change in the heart rate or increased shifts in electrodermal activity. This is one of the reasons why CNNs proved particularly good at extracting high-level, complex features of wearable data that may otherwise be disregarded by other models.
- **Long Short-Term Memory (LSTM):** The highest performance differed on all evaluation metrics as LSTM has an accuracy of 93%, precision of 94%, recall of 92%, and F1-score of 93. The architecture and its peculiarities to work exclusively with sequential data enabled it to represent the temporal relationship of sensor data on a long-term scale. This allowed the LSTM to pick up subtle patterns indicative of physiological and behavioural changes, which is why it can be used to predict mental health effectively using wearable data.

### *4.2. Discussion*

The relative performance of the models emphasizes the significance of the selection of architectures that are congruent with the character of the data to be analysed. In this paper, deep learning models, particularly the Long Short-Term Memory

(LSTM) network, are shown to be considerably superior to traditional machine learning methods. The advantage of LSTM is that it has a distinct design with memory cells that can store relevant information over a long period of time. This enables the model to interpret temporal dependencies and changing trends in physiological data, such as heart rate variability, Electrodermal Activity (EDA), and sleep cycles. Such time trends may be very subtle, but they are an important predictor of the status of mental health, like the progressive build-up of stress, anxiety, or depression. Traditional models, such as Support Vector Machines (SVM) and even ensemble methods like Random Forest (RF), cannot be used effectively in modelling time dependencies. They gloss over the complex temporal dynamics of wearable data, where data points are independent and rely on a set of fixed features. Even though this was advantageous over SVM because of utilizing feature randomness and ensemble learning against SVM, Random Forest did not match deep learning methods when it came to modeling how physiological states changed across the time scans. Convolutional Neural Networks (CNNs) performed well, especially when it comes to detecting temporal patterns in short information. CNNs are also excellent at extracting local features from structured inputs, such as time-series windows, and this means they are well-suited to detecting sudden changes in physiological signals.

Nevertheless, they cannot take the same long-range memory instances as LSTMs, and this was probably the cause of the performance difference. Collectively, the results confirm the emerging literature that wearable IoT data have high potential to capture mental health data in real-time and without intrusion, when complemented with next-generation deep learning models. This integration not only increases the accuracy of diagnosis but also introduces opportunities for early intervention and customized mental health care on a large scale.

### *4.3. Limitations*

Although the findings of the present study have a positive promise of utilizing wearable data and machine learning (ML) to evaluate mental health, a number of limitations should be noted to provide a realistic approach and to inform future research. Among the most critical challenges associated with wearable devices, it is worth mentioning battery life, especially for more sophisticated ones such as the Empatica E4, which have several sensors (e.g., EDA, BVP, temperature). Such devices require regular charging of around 24 or 48 hours, which may interfere with continuous monitoring. This can cause loss of data, particularly at crucial periods when continuous data would be most ideal, because someone is straining their mind or emotions. Frequent recharging may also lead to participant fatigue, potentially resulting in decreased long-term compliance in real-life scenarios. The other peculiar restriction is the imbalanced training set, since there are significantly fewer cases of clinically significant mental health episodes compared to the other normal or mild stress cases. Such an imbalance may make the learning process in the model biased and choose the majority class, thus reducing the possibility of detecting the less frequent, and at the same time more serious, conditions such as panic attacks or depressive states. Data augmentation (or synthetic minority oversampling (SMOTE)) is one possible technique, but it is less than ideal and does not necessarily reflect the intricacy of actual mental health incidents. And finally, the study lasted six months, although it was sufficient to conduct a preliminary analysis; however, one cannot conclude over a longer period. The factor that increases the risk of mental illness include season, life transition, and even changing individual circumstances, all of which demand a model that is flexible and can last a long duration. Longitudinal research, with a duration of at least one year, preferably involving diverse audiences, should be conducted in the future to enhance the reliability and validity of this model in various life settings. These shortcomings will be instrumental in resolving the shift from research prototypes to clinically feasible mental health monitoring systems.

## 5. Conclusion

This paper brings out the great future of Machine Learning (ML) and Deep Learning (DL) in the prevention of mental health disorders by utilizing wearable Internet of Things (IoT) technologies. With the help of constant and non-intrusive data that are emitted throughout our active periods, including heart rate variability (HRV), Electrodermal Activity (EDA), temperature, and sleep patterns, models can distinguish between stress, anxiety, and depressive trends with a high degree of accuracy. When comparing all considered models, models with Long Short-Term Memory (LSTM) were the most efficient, with a 93% F1-score, because they were able to learn and represent long-time temporal patterns of the analysis of physiological data. The findings indicate that by making the right pairing of wearable technology and smart algorithms, it is highly feasible to construct mental health monitoring systems that enable reliable, yet scalable, real-time, and automated systems. Standard algorithms, such as Support Vector Machine (SVM) and Random Forest (RF), delivered good benchmarks, but Convolutional Neural Networks (CNN) demonstrated good results in detecting changes in the short term. Nevertheless, the outstanding performance of LSTM shows that to comprehend the behavioral and physiological signals throughout time, long-range memory modeling is essential.

There are a number of possibilities in terms of future development in perspective. Telehealth platforms may be one of the areas where such ML-based mental health systems may be integrated, which would increase their utility notably. Telehealth-integrated systems can enable the real-time transmission of mental health indicators, allowing a resource clinician, therapist, or caregiver to inform earlier intervention and concurrent care, even in low-coverage, remote regions. Second, real-life clinical trials would be necessary to ascertain the validity and efficacy of these models outside controlled conditions. The existing evidence, although promising, relies on a specific group of patients in a restricted time interval; its evidence will gain strength in the case of large-scale implementation within a variety of populations and integration into clinical practices. Finally, blockchain technology might also be used to deal with urgent data security and data privacy issues. Since the wearable devices capture sensitive information about personal health, the secure data transfer to and between users, the devices, and the health professionals is important. The blockchain provides an alternative, decentralized and tamper-proof approach to working with such data, supporting user and regulatory applications. Collectively, the following development lines will assist in filling the gap between scholarly research and real-life patient-focused mental health interventions to achieve more intelligent, humane healthcare networks.

## References


[1] Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., & Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ, 4, e2537.

[2] Ghandeharioun, A., Fedor, S., Sangermano, L., Ionescu, D., Alpert, J., Dale, C., ... & Picard, R. (2017, October). Objective assessment of depressive symptoms with machine learning and wearable sensor data. In the 2017 seventh international conference on affective computing and intelligent interaction (ACII) (pp. 325-332). IEEE.

[3] Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. Journal of Medical Internet Research, 13(3), e1838.

[4] Canzian, L., & Musolesi, M. (2015, September). Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis, in Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing (pp. 1293-1304).

[5] Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., ... & Campbell, A. T. (2014, September). StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In Proceedings of the 2014 ACM international joint conference on Pervasive and Ubiquitous Computing (pp. 3-14).

[6] Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2016, September). Continuous stress detection using a wrist device: in laboratory and real life. In Proceedings of the 2016 ACM international joint conference on Pervasive and Ubiquitous Computing: Adjunct (pp. 1185-1193).

[7] Smets, E., Rios Velazquez, E., Schiavone, G., Chakroun, I., D'Hondt, E., De Raedt, W., ... & Van Hoof, C. (2018). Large-scale wearable data reveal digital phenotypes for daily-life stress detection. NPJ digital medicine, 1(1), 67.

[8] Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. EPJ Data Science, 6(1), 15.

[9] Sano, A., & Picard, R. W. (2013, September). Stress recognition using wearable sensors and mobile phones. In 2013, the Humaine association conference on affective computing and intelligent interaction (pp. 671-676). IEEE.

[10] Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A survey of mobile phone sensing. IEEE Communications magazine, 48(9), 140-150.

[11] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[12] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. NPJ digital medicine, 3(1), 119.

[13] Singh, G., Garg, V., Tiwari, P., & Goel, R. (2020). Emergence of IoT and Big Data: A Study in the Healthcare Industry. In Applications of Deep Learning and Big IoT on Personalized Healthcare Services (pp. 42-54). IGI Global.

[14] Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). Artificial intelligence with a multi-functional machine learning platform for better healthcare and precision medicine. Database, 2020, baaa010.

[15] Noorbakhsh, J., Chandok, H., Karuturi, R. K. M., & George, J. (2019). Machine learning in biology and medicine. Advances in Molecular Pathology, 2(1), 143-152.

[16] Chen, P. H. C., Liu, Y., & Peng, L. (2019). How to develop machine learning models for healthcare. Nature materials, 18(5), 410-414.

[17] Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(5), e1379.

[18] Callahan, A., & Shah, N. H. (2017). Machine learning in healthcare. In Key advances in clinical informatics (pp. 279-291). Academic Press.

[19] Wang, S., Cao, J., & Philip, S. Y. (2020). Deep learning for spatio-temporal data mining: A survey. IEEE transactions on knowledge and data engineering, 34(8), 3681-3700.

[20] Chen, W., An, J., Li, R., Fu, L., Xie, G., Bhuiyan, M. Z. A., & Li, K. (2018). A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial–temporal data features. Future Generation Computer Systems, 89, 78-88.

[21] Jha, D. K., Srivastav, A., & Ray, A. (2016). Temporal learning in video data using deep learning and Gaussian processes. International Journal of Prognostics and Health Management, 7(4).

[22] Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020). A novel framework for spatio-temporal prediction of environmental data using deep learning. Scientific reports, 10(1), 22243.