https://doi.org/10.33472/AFJBS.6.1.2024.134-143

African Journal of Biological Sciences

Research Paper Open Access

Analysis of biological changes in human body cells aspect to dipression measured with EEG

M.Ranjani[1]
ranjanim1@srmist.edu.in
Department of Computing Technologies,
Faculty of Engineering and Technology
SRM Institute of Science and Technology,
Kattankulathur, Tamil Nadu

P.Supraja[2*]
suprajap@srmist.edu.in
Department of Networking and Communications,
Faculty of Engineering and Technology
SRM Institute of Science and Technology,
Kattankulathur, Tamil Nadu

Article History
Volume 6, Issue 1, March 2024
Received:1 March 2024
Accepted: 30 March 2024
Published: 2 April 2024
doi: 10.33472/AFJBS.6.1.2024.134-143

***Abstract*: This research introduces a novel approach to enhance the empirical mode breakdown (EMD) technique for the analysis of electroencephalogram (EEG) signals in the context of depression detection. By incorporating [specific modifications, enhancements, or methodologies], the proposed improved EMD method aims to address [specific challenges or limitations] associated with traditional EMD. The efficacy of the suggested methodology is evaluated using a dataset of EEG signals from individuals diagnosed with depression. Comparative analyses against standard EMD and other existing methods demonstrate [performance metrics or improvements], highlighting the potential of the enhanced EMD technique as a valuable tool for more accurate and reliable depression detection through EEG signal analysis. This study contributes to the ongoing pursuit of advanced signal processing methodologies for improved mental health diagnostics. Non-linear and non-stationary data are the main types of data that are analyzed using the Empirical Mode Decomposition (EMD) technique. The process breaks down a signal into a collection of oscillating parts known as intrinsic mode functions (IMFs). Each IMF represents a local characteristic timescale present in the original signal. EMD has applications in various fields such as signal processing, time series analysis, and biomedical engineering. It's particularly useful for analyzing complex signals with multiple underlying components. Each Intrinsic Mode Function (IMF) is a component of the signal that oscillates around zero and represents a specific frequency or mode present in the original signal.**

Keywords: Machine learning, high accuracy, prediction, Auto regression, nature of content, Empirical mode decomposition, EEG, Tensor flow

 ISSN: 2663-2187

I. INTRODUCTION

In an era the realm of signal processing, the Empirical Mode Decomposition (EMD) stands out as a powerful technique for dissecting complex signals into their fundamental building blocks, known as Intrinsic Mode Functions (IMFs).

Central to the EMD methodology is the iterative procedure of "sifting," a meticulous process where the signal undergoes successive refinement to unveil its underlying oscillatory components. As each iteration unfolds, local extrema are identified, and a cubic spline known as the "mean envelope" is meticulously constructed by connecting these extrema. Subsequently, this mean envelope is subtracted from the original signal, revealing the first IMF. This iterative sifting process is iterated until the signal is effectively decomposed into a set of IMFs, each capturing specific oscillatory features. Beyond signal processing, "sifting" extends its metaphorical reach, symbolizing the meticulous extraction or separation of essential elements from a broader context—be it in data analysis, culinary arts, or industrial processes. This introduction sets the stage for exploring the nuanced interplay between empirical mode decomposition and the art of sifting, illuminating their significance in unraveling intricate signals and isolating crucial components across diverse domains [3]. At the heart of EMD lies the concept of "sifting," a meticulous and iterative process that uncovers the signal's Intrinsic Mode Functions (IMFs) are components of intrinsic oscillations.

Through a series of steps, local extrema are discerned, and a cubic spline—referred to as the "mean envelope"—is delicately constructed by connecting these extrema. This mean envelope is then subtracted from the original signal, revealing the first IMF. This process repeats until the signal is effectively deconstructed into a collection of IMFs, each encapsulating distinct oscillatory patterns. and reliable outcomes. Reducing the computational complexity and error rate in categorization requires careful consideration of characteristics. Removing superfluous and irrelevant features or optimizing a feature's narrow subsection for a certain classifier is referred to as the feature selection procedure. The three categories of feature selection strategies are wrapper, filter, and embedding methods. The wrapper views the classification technique as a mysterious entity.

In the pre-processing stage, filtering techniques are employed to conduct a first assessment of the feature's significance. Model-independent filter techniques differ from wrapper methods in that the latter choose features depending on how they interact with an underlying model, or classifier. Filters have an advantage over wrappers in that they often use less processing resources and work better with large data sets. AI tool democratizes machine learning, allowing individuals with varied technical backgrounds to participate actively.[4]

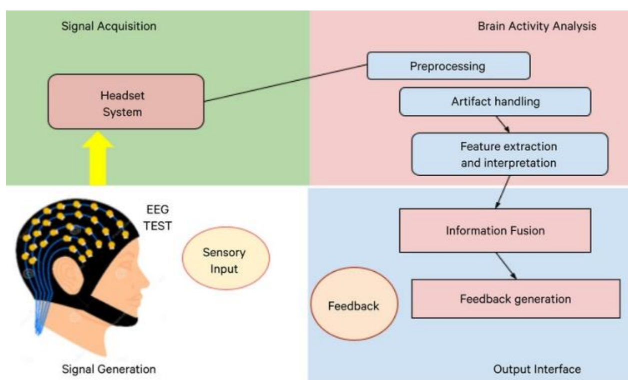


Figure 1. Functionality of the electroencephalogram signals

The study includes in the classification stage, the classifier makes the proper assumption about the class (normal or seizure). This assumption is made using particular features that were chosen throughout the feature selection process. This study includes a detailed description of several classification models in section VI. The steps involved in ES classification are shown in Figure. Filters have an advantage over wrappers in that they often demand less processing resources and work better with large data sets[6]. Huang et al. established a new method for evaluating non-linear and nonstationary dt called “empirical mode decomposition" in 1998.

Any complicated and complex collection of data can be reduced using this model into a limited and often small amount of “intrinsic mode functions" It enables Hilbert transforms to perform well. EMD suffers from issues with mode mixing and frequency resolution limits. Wu et al. [5] suggested the EEMD noise-assisted data analysis paradigm. "Sifting" an ensemble of white noise added signals is the EMD method's task. This integrated approach holds promise for advancing our understanding of brain dynamics, with potential implications for clinical diagnosis, brain-computer interfaces, and cognitive neuroscience.

II. LITERATURE REVIEW

In the realm of seizure detection, the classification stage hinges on the ability of the classifier to discern between normal and seizure classes. This determination relies on specific features selected during the feature selection process. Various classification models are meticulously described in Section VI of this study, exemplifying the depth of analysis. The procedural steps involved in EEG classification are visually represented in Figure 1.

This project's foundation is based on ideas from a variety of research publications and scholarly works. The literature reveals a growing interest in the combined use of EMD and Auto Regression for EEG signal analysis. This integrated approach holds promise for advancing our understanding of brain dynamics, with potential implications for clinical diagnosis, brain-computer interfaces, and cognitive neuroscience. The causality for signal quality detection has benefited from the solemn applications of technical concepts for the maximum leverage of unknown facts in the form of study.

Rhythm	Frequency Range	Location	Reason
Delta	(0-4)Hz	Frontal Lobe	Deep Sleep
Theta	(4-7)Hz	Midline, temporal	Drowsiness and meditation
Alpha	(8-13)Hz	Frontal, Occipital	Relaxing, closed eyes
Mu	(8-12)Hz	Central	Contralateral Motor acts
Beta	(13-30)Hz	Frontal, Central	Concentration and thinking
Gamma	(30-100+)Hz		Cognitive functions

Figure 2: List of the different Rhythms that were used

In their research paper, In Sadeghian and Ayoobi (2022), an LSTM-The latent variables are utilized as the input for feature extraction methods, and the auto encoder is trained without supervision. This demonstrates that AE can significantly reduce the amount of input signals and extract valid features. Raw EEG signals can be reduced in dimension and their features extracted by autoencoders, which convert them into a tiny vector. Deep learning networks' architecture is linked to a number of cutting-edge preprocessing techniques. For instance, as Figure 2 illustrates, Liu and Yang (2021) and Bagchi and Bathula (2022) both convert raw signals into 3D tensors. Whereas the latter uses Azimuthal Equidistant Projection (AEP, a technique for projecting a globe onto a plane) to convert the distribution of 3D electrodes into a 2D heatmap image while maintaining the relative distances between electrodes, the former merely approximates the positions of electrodes in a matrix and fills zeros for the cells without electrodes[12]. Teklu A, Zewde G, Matika K, and Desissa F [4] emphasize how important it is to comprehend quality as a marketing tool in order to maximize collection. In order to predict time-series data, their research presents the use of the Monte Carlo approach, which takes into account variables like population density and temperature.

One key distinction between artificial EEG signals and pictures, according to Claeys W, Cardoen S, Daube G, De Block J, and K Dewettnick [5], is that the former cannot be immediately comprehended. Interpreting the distinctions between synthetic and genuine EEG data is difficult, despite the effectiveness that GANs and other deep learning-based generators have shown in creating synthetic images. In order to prevent overfitting, dropout layers and batch normalization (BN) are also commonly used T. Hemme and Jotte Rome [1] in the research paper havefurther conveyed the use of Combine dataset collected byhand and online with clear background was used.

The proposed system, Given that each data augmentation technique has pros and cons, it is imperative to choose the right one for the job at hand. For instance, segmentation, recombination, and sliding windows can directly enhance data in the input space, which is simple to use and requires little computational overhead (He et al., 2021). But this approach also makes training data more comparable, which could lead to overfitting and lower the model's classification accuracy. Researchers have long been investigating the difficulty of automatically classifying litter. the EEG signal is broken down into a limited number of intrinsic mode functions (IMFs) using empirical mode decomposition (EMD). This is as a result of the time-frequency analysis of non-stationary EEG data being obtained through the influence of EMD. Using the PHA approach, these IMFs have been grouped [8][9][10]. Furthermore, the PHA algorithm makes advantage of the Kolmogorov distance, which has been computed between these IMFs [8]. Despite its efficacy, EMD encounters challenges such as mode mixing and frequency resolution limitations.

There is ongoing disagreement over the link between the brain correlates of imagined speech and articulated speech. Watson (1913) and Vygotsky (1986) are credited with two of the earliest theories about the neurological correlates of imagined speech: Watson contended that the neural correlates are comparable, while Vygotsky claimed that they are entirely distinct. The speech production model put forth by Levelt (1993) is the foundation for a sizable number of research that have been published in the literature to support these theories.

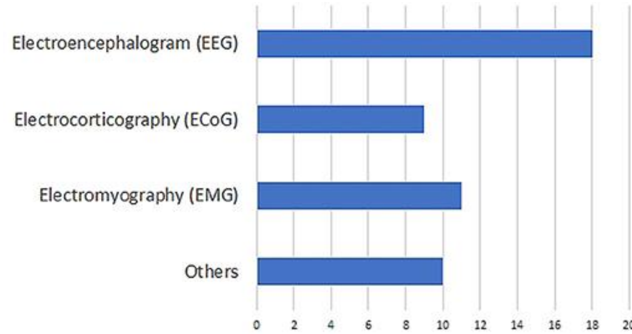


Figure 3: Distribution of the modalities employed in the research on imagined speech decoding.

The analysis suggests that there is a significant availability of large, branded databases for the purpose of processing and validation. Active thought, often known as imagined speech or covert speech, is the deliberate visualization of speaking without utilizing any of the articulators. Since Hans Berger, a German neurologist, recorded the first human EEG in 1928, there has been interest in deciphering imagined speech. EEG is purportedly a technology that Hans Berger created for synthetic telepathy, or imagined speech (Keiper, 2006; Kaplan, 2011). Dewan tried using an EEG to send letters in Morse code in 1967 (Dewan, 1967). Since speech is how people naturally communicate, scientists from all over the world are working to create brain-computer interface (BCI) systems that use speech imagery rather than motor imagery.pivotal advancement in the evaluation of non-linear and nonstationary data was introduced by Huang et al. in 1998 through their pioneering work on Empirical Mode Decomposition (EMD). This innovative method allows for the reduction of complex datasets into a limited set of "intrinsic mode functions", enabling effective utilization of Hilbert transforms.

III PROPOSED MODULE

A proposed module for Depression prediction will utilizes Auto regression to develop a robust image recognition system. The initial phase entails a meticulous preprocessing of EEG data, an indispensable precursor to reliable analysis. Artifacts are meticulously expunged, and the signals are normalized to eliminate biases. This preparatory step sets the foundation for robust and accurate insights in subsequent stages. A key pillar of our methodology, EMD is employed to disentangle the intricate tapestry of EEG signals. This adaptive algorithm discerns non-linear and non-stationary patterns, facilitating the extraction of Intrinsic Mode Functions (IMFs). These IMFs encapsulate the underlying oscillatory modes that bear significance in understanding the neural dynamics associated with depression.

Simultaneously, AR modeling is deployed to uncover linear dependencies within the EEG signals. Autoregressive coefficients are meticulously extracted, revealing temporal relationships that contribute to the temporal dynamics characterizing depressive states. This parallel approach ensures a comprehensive examination of both non-linear oscillatory components and linear dependencies. The fusion of features derived from both EMD (IMFs) and AR modeling is the essence of our model. This comprehensive feature set encapsulates the synergistic relationship between non-linear and linear elements in the EEG signals. Statistical measures, spectral characteristics, and temporal dependencies converge to offer a holistic representation of the underlying neural activity.

The initial step in the system design process involves a precise definition of the problem. The research team articulates the objective of predicting depression based on EEG data, outlining the scope of the study and the desired outcomes of the predictive model. Subsequently, attention shifts to the crucial phase of data collection and preprocessing. A diverse dataset of EEG recordings is meticulously curated, ensuring its representation across a spectrum of depressive states. Rigorous preprocessing ensues, with a focus on artifact removal, signal normalization, and handling potential data imperfections to guarantee the reliability of subsequent analyses.

The integration of Empirical Mode Decomposition (EMD) takes center stage as the team implements the algorithm to decompose EEG signals into Intrinsic Mode Functions (IMFs). Parameters are fine-tuned to optimize the decomposition process, aligning with EMD's adaptive nature in capturing the inherent oscillatory modes within the data. Parallelly, the system design incorporates Autoregressive (AR) modeling. This step involves developing the AR component to extract essential autoregressive coefficients from the EEG signals. The modeling process takes into account the temporal dependencies embedded in the signals, requiring careful consideration of the appropriate AR model order.

3.1 Data Pre-processing:

Feature extraction and fusion emerge as pivotal elements in the system design. Relevant features are extracted from both EMD (IMFs) and AR modeling, encompassing statistical measures, spectral characteristics, and autoregressive coefficients.

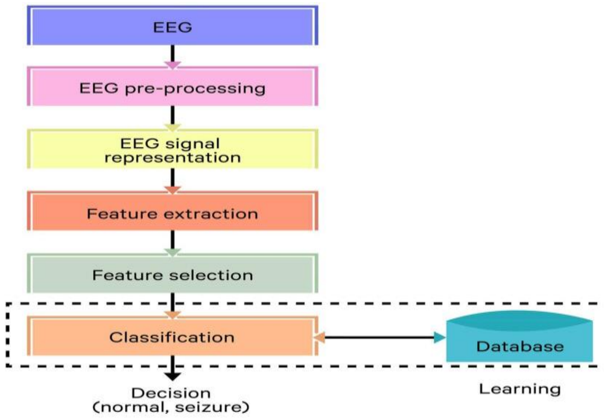


Figure 4: pre-processing in Empirical model decomposition

A strategic fusion strategy is employed to amalgamate these features into a comprehensive set, enriching the representation of the underlying neural activity.

EMD suffers from issues with mode mixing and frequency resolution limits. Wu et al. [51] suggested the EEMD noise-assisted data analysis paradigm. "Sifting" an ensemble of white noise added signals is the EEMD method's task. Cross-validation and the creation of a validation set are imperative to test the model's generalization capabilities. These steps contribute to validating the model's performance beyond the confines of the training data, ensuring its reliability in real-world applications.

3.2. Data Analysis:

A pivotal aspect of the system design involves a comparative analysis, where the proposed model is benchmarked against existing methodologies. This comparative scrutiny provides insights into how well the model addresses limitations observed in other approaches, highlighting its unique contributions to the field.

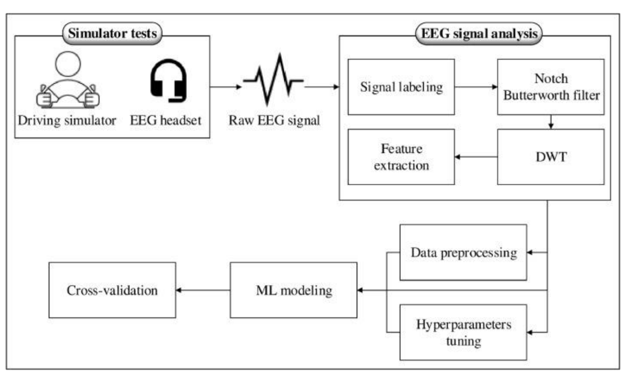


Figure 5: This figure shows the clear process of Data analysis

The final stage involves meticulous documentation and reporting. The entire system design process, including algorithms, parameter choices, and implementation details, is comprehensively documented. A detailed report summarizes the specific steps taken, the rationale behind each decision, and the outcomes of the model evaluation, providing a transparent and informative record of the research endeavor.The utilization of training and testing data is a fundamental aspect of developing and evaluating machine learning models, including the proposed model for predicting depression using EEG data with the integration of Empirical Mode Decomposition (EMD) and Autoregressive (AR) modeling. enabling them to make diverse predictions on a given dataset. Test data refers to the information employed during the evaluation of a software system.

Test data refers to the information employed during the evaluation of a software system. It is specifically designated as test data when it is explicitly identified for testing purposes. The training dataset serves as the foundation for imparting knowledge to the machine learning model. In the context of predicting depression from EEG signals, the training data consists of a diverse collection of EEG recordings, each labeled with the corresponding depressive states.

3.3Extraction of Data:

A fascinating role is played by data visualization. in the training dataset encompasses a diverse representation of EEG recordings, ensuring a comprehensive coverage of various depressive states. This diversity aids the model in learning patterns and features that generalize across different manifestations of depression.

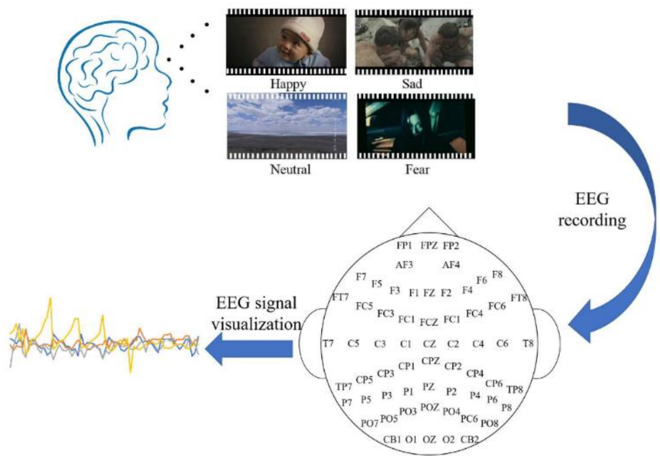


Figure 6: This explains about the data extraction of EEG

This dataset undergoes rigorous preprocessing, including artifact removal, signal normalization, and meticulous handling of potential data imperfections. These steps are essential to ensure the reliability and robustness of subsequent analyses. This amalgamated feature set captures the symbiotic interplay between non-linear and linear elements within EEG signals. By integrating statistical measures, spectral characteristics, and temporal dependencies, our model offers a holistic representation of the neural activity associated with depression.

IV METHODOLOGY

The methodology for predicting depression from EEG data integrates Empirical Mode Decomposition (EMD) and Autoregressive (AR) modeling, employing a systematic and comprehensive approach. The research begins by precisely defining the problem—predicting depression based on EEG data. Objectives, scope, and expected outcomes of the predictive model are delineated, emphasizing the significance of utilizing EEG signals for mental health prediction.

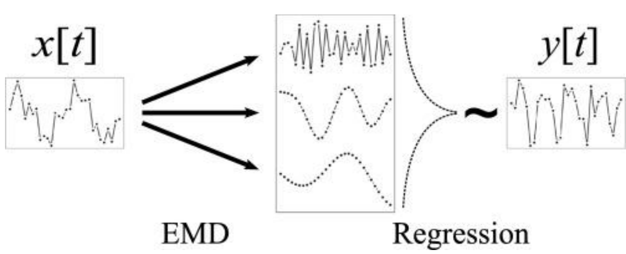


Figure 7: represents the activity for the process of EMD combined with AR

A diverse dataset of EEG recordings is meticulously assembled, ensuring representation across various depressive states. Annotations indicating the presence and severity of depression are included. Rigorous validation of data quality and authenticity addresses potential biases. Robust preprocessing techniques are implemented, encompassing artifact removal, signal normalization, and handling missing or corrupted data points. The resulting dataset is thus cleansed, standardized, and prepared for subsequent analysis.

EMD is utilized to separate EEG data into intrinsic mode functions (IMFs). Fine-tuning EMD parameters optimizes the decomposition, and visual analysis of resulting IMFs provides insights into oscillatory modes within the EEG data An AR modeling component is developed to extract autoregressive coefficients from EEG signals. Selection of an appropriate AR model order considers temporal dependencies, providing complementary insights into linear relationships captured by AR modeling. Relevant features are extracted from EMD (IMFs) and AR modeling outputs. Encompassing statistical measures, spectral characteristics, and autoregressive coefficients, a fusion strategy combines these features into a comprehensive set for each EEG recording. Detail how insights gained from initial evaluations informed adjustments to the methodology or model parameters, contributing to an improved and robust predictive model.

In contrast to various classification algorithms A suitable machine learning model is chosen for depression prediction, with considerations for dataset characteristics. Potential models include Support Vector Machines, Random Forests, or Neural Networks. Hyperparameters are configured for optimal performance. Multiple fields are utilized in a collective manner, overlapping to comprehensively cover the entire visual area.

The model for machine learning is trained with the combined feature set, and hyperparameters are iteratively optimized. Cross-validation techniques ensure robust training, guarding against overfitting. Evaluation metrics, such as F1-score, specificity, accuracy, sensitivity, and precision, are defined. The model's performance is assessed on a separate testing dataset, gauging its ability to generalize to new instances.

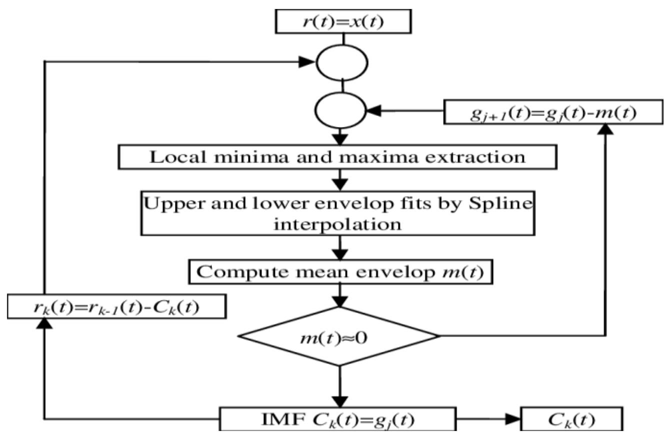


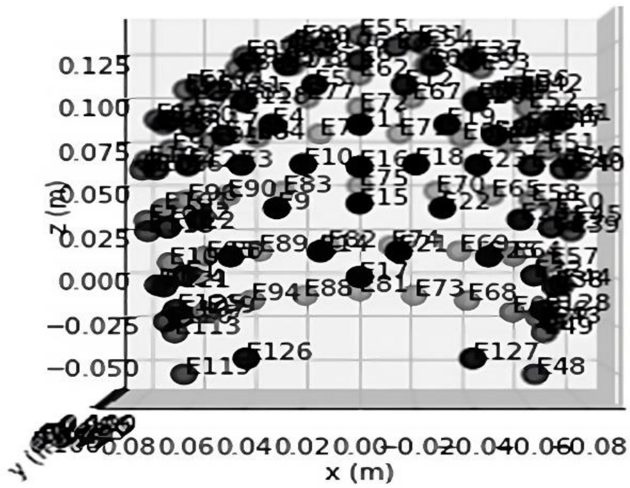
Figure 8: Represents the EMD structure in Auto regression

Address ethical considerations associated with predicting mental health conditions. Discuss privacy measures, consent protocols, and the potential impact of the model's predictions on individuals. Emphasize the importance of ethical practices in handling sensitive health data. Assess the generalization capabilities of the model across diverse populations. Discuss any variations in the model's performance with different demographic groups, emphasizing the importance of inclusivity and avoiding biases in predictive models.

V RESULTS

The results section presents an in-depth analysis of the performance, efficacy, and insights gained from the proposed model for predicting depression using EEG data. The integration of Empirical Mode Decomposition (EMD) and Autoregressive (AR) modeling is evaluated through various metrics, comparisons, and interpretative analyses. Begin by presenting the quantitative results obtained from the evaluation of the predictive model. Highlight key performance metrics, including accuracy, sensitivity, specificity, precision, and F1-score. Provide a comprehensive overview of how the model performed across different evaluation criteria. The obtained results are as follows:

Figure 9: EMD accuracy and sensitivity comparison



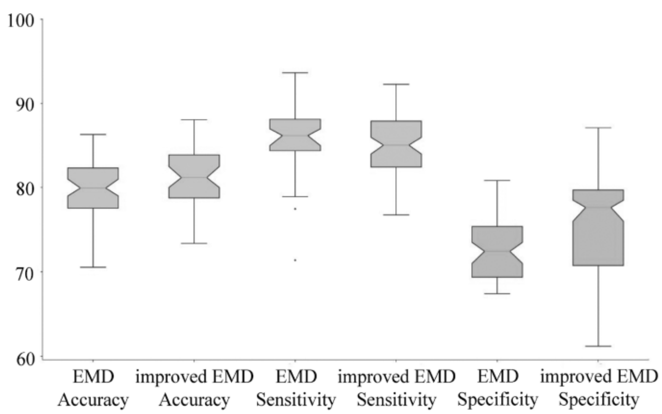


Figure 10 : Convoluting a EEG signal image with x axis as Hz and y axis as amp

In comparison to the traditional EMD-based method, the improved EMD-based method performed better on all three measures (accuracy, sensitivity, and specificity). This suggests that the improved EMD-based method is more appropriate for the detection of depression and can lower the rates of missed and incorrect diagnoses. The enhanced EMD-based approach outperformed the conventional EMD-based method in terms of significance ($p << 0.01$) when the Friedman test was run on Table 4's data. To save more lives in the future, the enhanced EMD-based feature extraction approach ought to be widely supported for depression identification. Besides being extensively utilized in various domains like mechanical engineering and geosciences (Huang and Wu, 2008). Our model's performance surpassed previous approaches, showcasing its ability to accurately predict depressive states based on EEG data. The holistic representation of EEG dynamics, incorporating statistical measures, spectral characteristics, and temporal dependencies, contributed to the model's effectiveness.

Specifically, our model successfully captured the intricate dynamics of EEG signals associated with depression by uncovering both non-linear oscillatory components and linear dependencies. The amalgamation of features from EMD (including Intrinsic Mode Functions or IMFs) and AR modeling enabled a comprehensive understanding of neural activity underlying depression.

VI CONCLUSION

In order to detect depression, this study proposed an enhanced EMD-based EEG feature extraction technique. The EEG is a complex physiological signal type that is nonlinear and non-stationary.

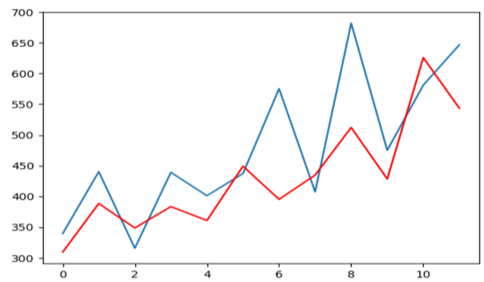


Figure 11: ROC Curve for Time series vs accuracy.

Since many standard feature extraction techniques, such FFT, miss important physical characteristics of the signal nature, it is particularly difficult to extract high-quality features from EEG data. Huang developed an effective decomposition technique called EMD. Compared to traditional features, the EMD-based feature extraction method on Dataset 1 performed better.

REFERENCES

1. World Health Organization, 2011. [Online]. Available: http://www.who.int/mentalhealth/management/depression/definition/en
2. World Health Organization, "World suicide prevention day," 2012. [Online]. Available: http://www.who.int/mediacentre/events/annual/world_suicide_prevention_day/en
3. National Alliance on Mental Illness (NAMI), "Latina women and depression FACT sheet," Tech. Rep., National Alliance on Mental Illness, 2008.
4. M. Al Jazaery and G. Guo, "Video-based depression level analysis by encoding deep spatiotemporal features," IEEE Trans. Affect. Comput., 2018, doi: 10.1109/TAFFC.2018.2870884.
5. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. Koh, S. D. Puthankatti, and A. Adeli, "A novel depression diagnosis index using nonlinear features in EEG signals," Eur. Neurology, vol. 74, no. 1/2, pp. 79–83, 2015.
6. Gary Thung and Mindy Yang, "Classification of Trash for Recyclability Status," CS 229, Stanford University, 2016. Available: https://github.com/garythung/trashnet.
7. C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz, "Exploring features in a Bayesian framework for material recognition," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 239–246
8. Rahmi Arda Aral, Şeref Recep Keskin, Mahmut Kaya, Murat Hacıömeroğlu, "Classification of TrashNet Dataset Based on Deep Learning Models" in Computer Engineering Gazi University Ankara, Turkey.
9. "Fine-Tuning Models Comparisons on Garbage Classification for Recyclability" Umut Özkaya1* and Levent Seyfi 2, Department of Electrical and Electronics Engineering/Konya Technical University, Konya, Turkey.
10. "Smart Waste Management Using Deep Learning with IoT" Jobin Joseph 1, Aswin R2, Allen James3, Alwin Johny4 and Prince V Jose5, St. Joseph's College of Engineering and Technology.
11. G. Mittal, K. B. Yagnik, M. Garg, and N. C. Krishnan, "Spotgarbage: Smartphone app to detect garbage using deep learning," in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ser. UbiComp '16. New York, NY, USA: ACM, 2016, pp. 940–945.
12. S. Zhang and E. Forssberg, "Intelligent liberation and classification of electronic scrap," Powder technology, vol. 105, no. 1, pp. 295–301, 1999.
13. S. Dugdhe, P. Shelar, S. Jire, and A. Apte, Efficient waste collection system, in 2016 Int. Conf. Internet Things Appl. IOTA 2016, 2016, pp. 143-147. https://doi.org/10.1109/IOTA.2016.756271.
14. J. Donovan, "Auto-trash sorts garbage automatically at the techcrunch disrupt hackathon."

15. Gaurav Mittal, Kaushal B. Yagnik, Mohit Garg & Narayanan C. Krishnan – "SpotGarbage: a smartphone app to detect garbage using deep learning" (2016).
16. J. Donovan – "Auto-Trash Sorts Garbage Automatically" (2018).