Automatic Depression Screening Through CNN And EEG Data

[1*]MunjiSobha Latha, [2]Ch. Kodandaramu , [3]A Srinivasa Babu
Student, Dept. Of Master of Computer Applications, at Miracle Educational Society Group of Institutions
Associate Professor , Miracle Educational Society Group of Institutions
Associate Professor , Miracle Educational Society Group of Institutions
*sobha2001munji@gmail.com

ABSTRACT
Depression remains a challenge worldwide owing to subjective assessment techniques. In this particular project, we experiment with an automated depression detection system based on EEG signals using Convolutional Neural Networks (CNN). The required EEG data is pulled from a public Kaggle data set, where it is preprocessed and entered into a CNN model that classifies individuals as either depressed or non-depressed. The CNN model demonstrated a 93% accuracy rate, significantly outperforming traditional Support Vector Machines (SVM) which offered a 63% accuracy rate. The focus of the proposed solution is on early identification, which minimizes the subjectivity associated with diagnosis and enables the efficient scale of mental health screening and timely intervention within clinical environments.
Keywords: SVM, Depression Detection, EEG Signals, Convolutional Neural Networks, Mental Health Screening

INTRODUCTION:

As one of the leading causes of disability worldwide in terms of population, depression impacts the lives of tens of millions while crippling their productivity, welfare, and overall quality of life. Described the nature and characteristics of depression, it will be noted that standard methods of its evaluation heavily depend on a patient's self-description of the experienced symptoms alongside the psychological evaluation, both of which are highly subjective and error-prone. Oftentimes, individuals fail to explain their emotions, which leads to missed or late diagnoses. In recent years, brain activity analysis through Electroencephalogram (EEG) signals has emerged as a promising objective method for detecting mental disorders, including depression. EEG provides insight into the brain's electrical activity, making it a valuable tool for identifying neural irregularities associated with depressive states. Machine learning, especially Convolutional Neural Networks (CNNs), allows automated and efficient analysis of such complex, high-dimensional EEG data. This project proposes a CNN-based system to detect depression using EEG signals, aiming to enhance prediction accuracy, reduce misclassification, and enable timely diagnosis. By leveraging modern AI techniques, the system offers a scalable and non-invasive alternative to traditional depression screening.

GAP IDENTIFIED BASED ON LITERATURE SURVEY:

While numerous studies have demonstrated the potential of EEG signals for depression detection, significant gaps still exist in practical implementation and prediction accuracy. Traditional machine learning models such as SVM, KNN, and Decision Trees depend heavily on manual feature extraction, which is time-consuming and prone to human bias. These models also show limited performance when handling noisy, high-dimensional EEG datasets.

Several researchers, including Ay et al. (2019) and Farruque et al. (2020), noted that traditional classifiers struggle to learn deep, hierarchical representations of EEG signals, often achieving accuracy below clinical utility standards.

Moreover, EEG preprocessing inconsistencies have been reported, making it difficult to compare results across studies. Preprocessing pipelines like PREP (Bigdely-Shamlo et al., 2015) are not yet universally adopted, which impacts reproducibility and reliability. Additionally, there is a lack of real-time, user-friendly diagnostic tools based on EEG and AI integration for depression screening.

The key gaps identified include:

- Inconsistent preprocessing protocols
- Reliance on manual feature engineering
- Low accuracy and generalization in traditional models
- Lack of practical, deployable solutions for real-time detection

PROBLEM STATEMENT:

Current depression detection methods lack objectivity, scalability, and precision, often leading to delayed or missed diagnoses. There is a critical need for an accurate, automated system that can analyze EEG signals for early depression detection using deep learning models.

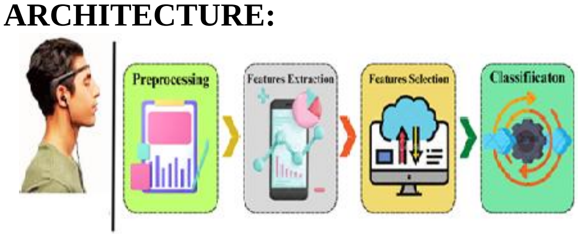
Key Challenges:

- Complexity and noise in EEG signals affecting model accuracy
- Manual preprocessing and feature extraction introducing subjectivity
- Traditional classifiers (e.g., SVM) perform poorly with high-dimensional data
- Ensuring real-time, user-friendly deployment in clinical or remote settings
- Balancing model accuracy with computational efficiency

PROPOSED METHOD:

The proposed system utilizes a Convolutional Neural Network (CNN) to detect depression based on EEG signals. The methodology involves acquiring an EEG dataset from Kaggle, which includes labeled data of depressed and non-depressed individuals. The raw data undergoes preprocessing—normalization, label correction, and missing value handling. Then, the data is reshaped into a 3D tensor suitable for CNN input. The CNN model, designed with multiple convolutional and pooling layers, extracts deep features and classifies the data. The system is trained using 80% of the dataset and evaluated on the remaining 20%. A comparison with an SVM model highlights the CNN's superior performance, achieving 93% accuracy. The final model includes real-time prediction capability through a user-friendly GUI.

ARCHITECTURE:



DATASET:

The dataset used for this study was sourced from Kaggle and comprises EEG signal recordings from both depressed and non-depressed individuals. Each record contains 989 columns representing features extracted from multi-channel EEG readings. The last column holds the class label: 0 for Normal and 1 for Depressed. The data is provided in CSV format and includes readings from various brain regions, allowing for comprehensive analysis of neural activity. Before model training, the data is cleaned, normalized, and reshaped into a format compatible with CNN input, ensuring effective learning of patterns associated with depressive symptoms.

METHODOLOGY:

Data Preprocessing:

EEG signals are inherently noisy and may contain missing values or irregularities. Preprocessing was a crucial step involving multiple tasks:

- **Handling Missing Values:** Null values in the dataset were replaced using zero-padding to maintain uniform structure.
- **Normalization:** The EEG signal amplitudes were scaled to a uniform range to reduce variance and ensure even model learning.
- **Shuffling:** Dataset entries were randomized to avoid bias during model training.
- **Reshaping:** The EEG features were reshaped into 3D tensors of shape (18, 18, 3), emulating image input suitable for CNN architecture.

3. Feature Extraction Using CNN:

CNN was selected for its ability to automatically extract hierarchical and spatial features from EEG signals. Unlike traditional models that require manual feature engineering, CNN captures complex signal patterns by learning directly from the raw reshaped data.

4. Splitting the Dataset:

The dataset was divided into an 80:20 ratio for training and testing, respectively. This ensures that the model is evaluated on unseen data and helps in measuring generalization capability.

5. Model Design and Training:

The CNN architecture used in the project included the following components:

- Two convolutional layers with 32 filters (3×3) followed by ReLU activations.
- Two max-pooling layers (2×2) for dimensionality reduction.
- A flatten layer to convert feature maps into a single vector.
- A dense (fully connected) layer with 256 neurons and ReLU activation.
- An output layer with softmax activation for binary classification.

The model was compiled using the Adam optimizer and trained using categorical cross-entropy as the loss function. The batch size was set to 16, and training was run for 10 epochs.

6. Performance Evaluation:

After training, the model was tested using the 20% reserved dataset. Performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix were computed. These metrics provided a comprehensive evaluation of the CNN's classification capabilities.

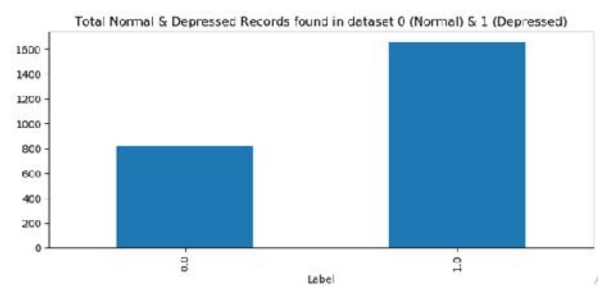
7. SVM Benchmarking:

A traditional Support Vector Machine (SVM) model was trained on the same data for comparative analysis. Results showed the CNN significantly outperformed the SVM, validating the effectiveness of deep learning for EEG-based depression detection.

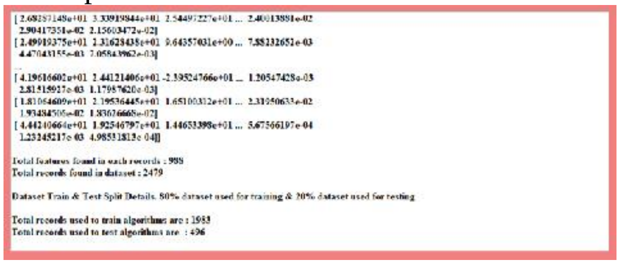
8. Real-Time Prediction Interface:

The system also includes a GUI that allows users to upload new EEG data and receive real-time predictions. This user-friendly interface makes the system practical for deployment in clinical or remote health monitoring environments.

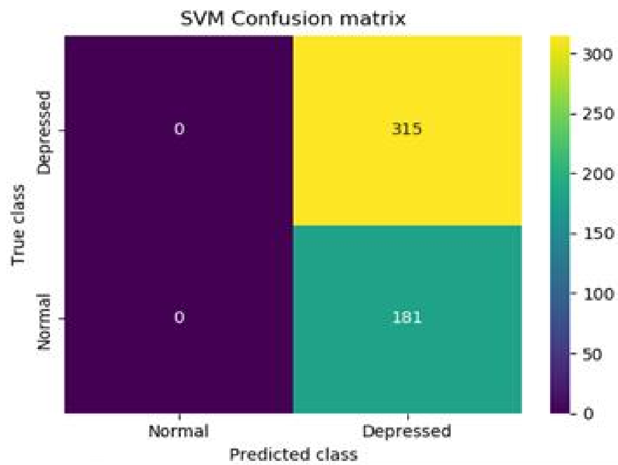
RESULTS:



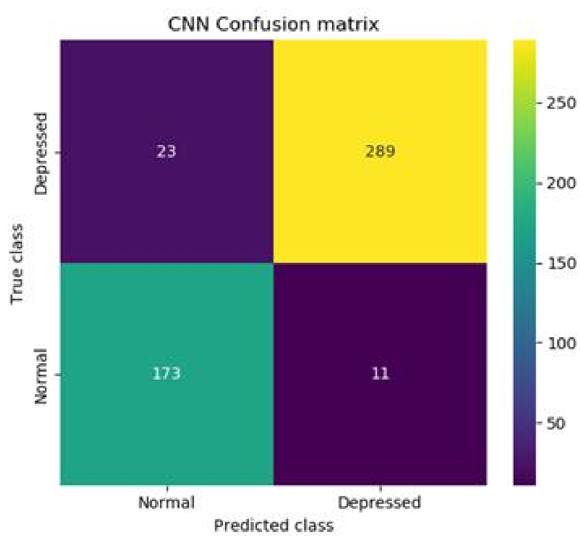
X-axis represents labels as 0 or 1 where 0 means Normal and 1 means Depressed and y-axis represents counts of records



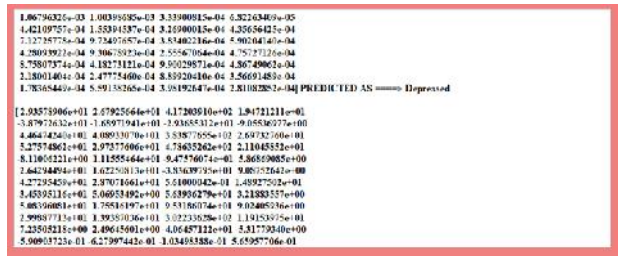
FeaturesExtraction



Training SVM Algorithm got 63% accuracy



Training CNN Algorithm got 93% accuracy



Predicting Depression from Test Signals

CONCLUSION

This project introduces an efficient, accurate, and scalable approach to depression detection using EEG signals and CNN. By automating feature extraction and leveraging deep learning, the system overcomes limitations of traditional diagnostic methods. The CNN model significantly outperformed conventional SVMs, achieving 93% accuracy and demonstrating strong potential for real-world clinical applications. With a user-friendly interface and real-time prediction feature, this solution can enhance mental health diagnostics by providing timely and objective insights. Future improvements may focus on expanding the dataset and integrating multimodal data to further boost performance and applicability across diverse populations.

REFERENCES:

1. Ay, B., Yildirim, O., Talo, M., Baloglu, U.B., Aydin, G., Puthankattil, S.D., Acharya, U.R., 2019. Automated depression detection using deep representation and sequence learning with EEG signals. J. Med. Syst. 43 (7), 205. https://doi.org/10.1007/ s10916-019-1345-y.
2. Bakhshali, M.A., Khademi, M., Ebrahimi-Moghadam, A., Moghimi, S., 2020. EEG signal classification of imagined speech based on riemannian distance of correntropy spectral density. Biomed. Signal Process Control 59 (2), 101899–101911. https:// doi.org/10.1016/j.bspc.2020.101899.

3. Beck, A.T., 1974. The development of depression: a cognitive model. In: Friedman, R.J., Katz, M.M. (Eds.), The Psychology of Depression: Contemporary Theory and Research. John Wiley, Oxford, pp. 318–330.
4. Beck, A.T., Beck, R.W., 1972. Screening depressed patients in family practice. A rapid technic. PGM (Postgrad. Med.) 52 (6), 81–85. https://doi.org/10.1080/ 00325481.1972.11713319.
5. Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.M., Robbins, K.A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front. Neuroinf. 9, 16. https://doi.org/10.3389/fninf.2015.00016.
6. Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebr. Cortex 19 (12), 2767–2796. https://doi.org/10.1093/cercor/bhp055.
7. Bradley, M.M., Lang, P.J., 1999. Affective norms for English words (ANEW): instruction manual and affective ratings. Technical Report C-1, the Center for Research in Psychophysiology. University of Florida.
8. Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis. 10, 433–436. https://doi. org/10.1163/156856897X00357.
9. Brinthaupt, T.M., 2019. Individual differences in self-talk frequency: social isolation and cognitive disruption. Front. Psychol. 10, 1088. https://doi.org/10.3389/ fpsyg.2019.01088.
10. Burns, D.D., 1980. Feeling Good: the New Mood therapy. New York, NY, USA, Signet. Chou, T., Deckersbach, T., Dougherty, D.D., Hooley, J.M., 2023. The default mode network and rumination in individuals at risk for depression. Soc. Cognit. Affect Neurosci. 18 (1), nsad032 https://doi.org/10.1093/scan/nsad032.
11. COVID-19 Mental Disorders Collaborators, 2021. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet (London, England) 398 (10312), 1700–1712. https:// doi.org/10.1016/S0140-6736(21)02143-7.
12. Daly, M., Robinson, E., 2022. Depression and anxiety during COVID-19. Lancet (London, England) 399 (10324), 518. https://doi.org/10.1016/S0140-6736(22)00187-8.
13. de Aguiar Neto, F.S., Rosa, J.L.G., 2019. Depression biomarkers using non-invasive EEG: a review. Neurosci. Biobehav. Rev. 105, 83–93. https://doi.org/10.1016/j. neubiorev.2019.07.021.
14. Eichstaedt, J.C., Smith, R.J., Merchant, R.M., Ungar, L.H., Crutchley, P., Preot¸iucPietro, D., Asch, D.A., Schwartz, H.A., 2018. Facebook language predicts depression in medical records. Proc. Natl. Acad. Sci. U.S.A. 115 (44), 11203–11208. https://doi. org/10.1073/pnas.1802331115.
15. Elliott, R., Rubinsztein, J.S., Sahakian, B.J., Dolan, R.J., 2002. The neural basis of moodcongruent processing biases in depression. Arch. Gen. Psychiatr. 59 (7), 597–604. https://doi.org/10.1001/archpsyc.59.7.597 .
16. Engels, A.S., Heller, W., Spielberg, J.M., Warren, S.L., Sutton, B.P., Banich, M.T., Miller, G.A., 2010. Co-occurring anxiety influences patterns of brain activity in depression. Cognit. Affect Behav.

Neurosci. 10 (1), 141–156. https://doi.org/ 10.3758/CABN.10.1.141.

17. Fahimi Hnazaee, M., Khachatryan, E., Van Hulle, M.M., 2018. Semantic features reveal different networks during word processing: an EEG source localization study. Front. Hum. Neurosci. 12, 503. https://doi.org/10.3389/fnhum.2018.00503.
18. Farruque, N., Zaiane, O., Goebel, R., 2020. Augmenting semantic representation of depressive language: from forums to microblogs. In: Brefeld, U., Fromont, E., Hotho, A., Knobbe, A., Maathuis, M., Robardet, C. (Eds.), Machine Learning And Knowledge Discovery In Databases. ECML PKDD 2019, Lecture Notes in Computer Science, vol. 11908. Springer, Cham. https://doi.org/10.1007/978-3-030-46133-1_22.
19. Ferdek, M.A., van Rijn, C.M., Wyczesany, M., 2016. Depressive rumination and the emotional control circuit: an EEG localization and effective connectivity study. Cognit. Affect Behav. Neurosci. 16 (6), 1099–1113. https://doi.org/10.3758/ s13415-016-0456-x.
20. Gabard-Durnam, L.J., Mendez Leal, A.S., Wilkinson, C.L., Levin, A.R., 2018. The harvard automated processing pipeline for electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. Front. Neurosci. 12, 97. https://doi.org/10.3389/fnins.2018.00097