Artificial Intelligence Evolution
http://ojs.wiserpub.com/index.php/AIE/

UNIVERSAL WISER PUBLISHER

Research Article

Deep Learning Approaches for Electroencephalography (EEG)-Based User Response Prediction

Greeshma Sharma[1*], Vishal Pandey[2], Ayush Chauhan[3], Sushil Chandra[4]

[1]Department of Design, Indian Institute of Technology Delhi, Delhi, India
[2]Department of Computer Science and Engineering, Indian Institute of Technology Roorkee, Roorkee, India
[3]Vivekananda Institute of Professional Studies, Delhi, India
[4]Institute of Nuclear Medicine and Allied Sciences, Defence Research and Development Organization, Delhi, India
E-mail: greeshmacct@gmail.com

Received: 7 June 2023; **Revised:** 7 November 2023; **Accepted:** 17 November 2023

Abstract: In deep learning, finding the best algorithms for time series data can be challenging due to its stochastic and nonlinear nature. This study endeavours to address the challenges posed by a 10-class classification and binary classification problem employing deep learning algorithms. We collected Electroencephalography (EEG) data from participants engaged in the Corsi Block Tapping Task, utilizing various combinations of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM) models across multiple layers to achieve the highest accuracy across different frequency bands - namely, beta band (14-30 Hz), alpha band (8-13 Hz), theta band (4-8 Hz), and delta band (0.5-4.35 Hz). Our findings in the context of the 10-class classification problem highlight the superior performance of the 1 CNN + 4 linear layers model, boasting an accuracy of 64.47%. In the realm of binary classification, the 1 LSTM + 4 linear layers model emerged as the top performer, achieving an impressive accuracy of 93.30%. Notably, the beta wave exhibited enhanced predictive capabilities. These results hold promising implications for the design of brain-computer interface experiments, where specific brain regions can predict responses with heightened accuracy. Furthermore, future applications may encompass the development of cognitive systems where both time and accuracy play pivotal roles.

***Keywords*: deep learning, CNN, LSTM, BiLSTM, EEG**

1. Introduction

Automatic and continuous prediction of human behavior with minimum latency is a critical factor for various autonomous as well as Brain-Computer Interface (BCI) systems. Deep learning has recently gained widespread attention in the field of signal processing. However, it has not been fully explored for EEG signal classification. The direct goal of this paper is to identify a deep learning algorithm that predicts behavior with higher accuracy with minimum errors. Due to the sheer volume of EEG data points per second (depending on sampling frequency as well), the number of trials, and the number of participants, identifying and implementing machine learning tools become a daunting task. Since EEG is a time-series signal where adding input variables increases the complexity for prediction, deep learning models could provide solutions for such complexity. One review paper explored the effect of deep learning on the amount of EEG data

DOI: https://doi.org/10.37256/aie.4220233179

used [1]. About 40% of the studies used Convolutional Neural Networks (CNNs), while 14% used Recurrent Neural Networks (RNNs), most often with a total of 3 to 10 layers. Finally, the median gain in accuracy of Deep Learning (DL) approaches over traditional baselines was 5.4% across all relevant studies. Along with CNN and RNN, a new model of Long Short-Term Memory (LSTM) has emerged to apply to EEG signals for a range of activities and tasks such as motor imagery [2], human decision prediction [3], prediction of epileptic seizures [4], confusion detection [5], and sleep stage classification [6]. LSTM models demonstrate their versatility in processing EEG data, which is a complex and time-varying signal. The ability to capture temporal dependencies within EEG signals makes LSTMs a valuable tool for understanding various cognitive processes and developing applications ranging from medical diagnostics to user experience optimization.

Wang et al. [2] have used EEG data of eight subjects to train the proposed aggregate Approximation (AX)-LSTM network, which is then implemented to predict motor imagery labels of EEG data of the remaining subjects. Results achieved prediction accuracies above 60% in experiments of all the subjects. Another study used a novel hierarchical LSTM model to achieve significant improvement in classifying human decisions in a guard duty experiment [3]. It suggests that implementing the LSTM model with an attention mechanism successfully captures discriminate EEG features contributing to robust decision classification. In addition, deep learning approaches deliver competitive performance without the need for handcrafted features, enabling end-to-end classification [7]. Sun et al. [8] proposed an EEG-based Computational Prediction Approach. The paper proposed a Convolutional Neural Network for EEG (ConvEEGNN) to predict subsequently remembered and forgotten events from EEG recorded during the memory process. An average accuracy of 72.07% was observed during the recording process. Deep learning models have found applications in forecasting venue categories within social networks [9], predicting user responses to advertisements in marketing [10], and estimating conversion ratios for mobile users [11]. Nevertheless, the realm of user behavior prediction within non-deterministic systems, like medical and emerging technologies, remains largely unexplored, and we aim to address this gap. The novelty of this paper lies in the development of a novel framework that harnesses the Corsi Block Test as a predictive tool for user response. This innovation holds considerable promise for the fields of Brain-Computer Interface (BCI) technology and rehabilitation studies.

This study introduces a novel framework that leverages one-second EEG data collected following the onset of a target stimulus to construct a robust classification model. The primary objective of this research is to employ deep learning algorithms to predict user behavioral responses, categorizing them as either 'correct' or 'incorrect'.

Within this research framework, we pursue two key objectives. Firstly, we address the binary classification challenge, aiming to predict the accuracy of user responses, classifying them as either 'correct' or 'incorrect'. Secondly, we delve into the realm of a more intricate task, the 10-class classification problem, where our objective is to predict the specific number sequence to which the user is responding.

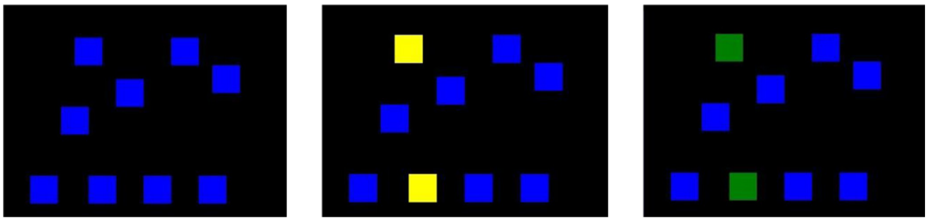
2. Methodology

2.1 *Participants*

Thirty participants were recruited for the study out of which 17 were male and 13 were female (mean age = 22 years; SD = 2.2). All participants had near-perfect (n = 25) or corrected to near-perfect vision (n = 05). Written, informed consent was collected from each participant after a detailed explanation of the task study conducted. The participants were requested not to take any caffeine, alcohol, or any medications for a minimum of 6 hours prior to the experiment [12]. They had no history of neurological or psychiatric illnesses.

2.2 *Experiment design*

The participants performed the Corsi Block Tapping test which was adopted from its standardized version (Figure 1). The task was designed using OpenSesame [13-15], an open-source program to create an experiment for social sciences. The test consisted of nine blue squares (blocks) of size 100 × 100 px on a black background. The stimuli sequences were randomized for each trial. An arbitrary sequence of blocks was flashed on the screen and the participants were required to repeat the sequence in the order in which they had flashed. The stimuli sequence (Corsi span) was increased step-by-step in order from 1 to 9. By incrementing the Corsi-span the visuospatial memory of an individual could be quantified. Initially, an instruction screen containing a detailed explanation of the task at hand was shown. The participant started the task with a Corsi span of 2. The sequence length was gradually increased from 2 to 9 step-wise. One successful attempt leads to an increase in a level while two unsuccessful attempts lead to a decrease in a level. A successful trial was represented as subjects pointed at the blocks in the order they were presented. Two trials of the same sequence length were presented to the participants. Every participant was required to perform the complete Corsi sequence irrespective of the accuracy of the prediction. Self-corrections were not permitted on the computer screen.



(a) (b) (c)

Figure 1. The computerized Corsi Block Tapping task. (a), (b) and (c) show the preparation, stimuli, and user input stages of a trial, respectively

2.3 *EEG Acquisition*

EEG data were recorded through Ag/AgCl electrodes from 64 electrode locations in line with the international 10-20 system using the eego™ mylab EEG system (ANT Neuro, Enschede, Netherlands). Impedances were kept below 5 KΩ. EEG data was collected reference-free on a tablet running ANT eego software. The signals were continually digitized at a rate of 2,048 samples per channel during the normal baseline and experimental phase. EEG data were recorded while the participant performed the Corsi Block Tapping Task and at rest (termed as Baseline).

2.4 *EEG signal processing methods*

We imported the raw EEG data recorded from the eego™ mylab EEG system and saved it in a EDF+ format. The raw EEG data was recorded at 1,024 Hz, but we downsampled it to 256 Hz using ASA software (ANT Software BV, Enschede, Netherlands). The acquired EEG data were preprocessed using the EEGLAB toolboxes (9.0.4) [28] with custom scripts. During EEG recording, the signals are interrupted by artifacts caused by eyeblinks, heart rate, and muscular movements such as jaw clenches. We eliminated these artifacts from the resampled signal using one Finite Impulse Response (FIR) filter of order 250 using the Kaiser window method. We used one Butterworth Infinite Impulse Response (IIR) filter to eliminate artifacts from the resampled signal caused by eyeblinks, heart rate, and muscular movements. All EEG signals were bandpass filtered using a FIR filter with 0.1 Hz to 40 Hz bandwidth. The power line interference was removed using a 50 Hz notch filter. The EEG data was re-referenced to the computed average reference. The data mean subtraction was performed to remove the baseline and Direct Current (DC) offset. We conducted checks to remove faulty channels by examining channels with no EEG activity for more than 5 seconds, channels with high noise (high standard deviation relative to other channels), and channels with low correlation with other channels (low correlation to other channels using a rejection threshold of 0.70). The Electromyogram (EMG) and ocular artefact segments were reduced using Independent Component Analysis (ICA) employing the runica function provided by EEGLAB [10-11]. Data was then visually inspected for any residual eye blinks and muscular artefacts. Different frequency bands were obtained by applying Discrete Wavelet Transform (DWT). The waves were beta (13-30 Hz), alpha (8-13 Hz), theta (4-8 Hz), and delta (0.5-4 Hz). The data that was fed into the system was obtained by concatenating the EEG data of all 30 subjects. The data obtained was then divided into batches of size 64. Normalization was performed on the data by scaling the EEG values according to the minimum and maximum values using preprocessing. To further perform data augmentation, the concept of window sliding (15 s) was implemented.

2.5 *Analysis*

For 2-class and 10-class problems, the following different models were implemented. Single Long Short-Term Memory (LSTM), Multi-LSTM Layer, Convolutional Neural Network (CNN) followed by LSTM layer, a Bidirectional LSTM (BiLSTM) sandwiched between two LSTM layers were the models. A sigmoid activation function was employed to feed information into the last dense layer. The hit-and-trial method was utilized throughout the entire process of configuring the parameters. The model's objective function was set as "binary_crossentropy" and the optimizer was "Adam". The other fitting parameters epoch (200) and batch_size (50) were chosen based on the hit-and-trial approach. All the models were implemented with SmoothL1loss and Mean Square Error (MSE) with learning rate = 0.1, 0.01, 0.001.

In these models, various combinations of different deep learning algorithms were tested, and different hyperparameters like learning rate, batch size, sequence length, number of hidden layers, activation functions, number of epochs, loss function, hidden dimensions, etc. were tuned. For the evaluation of our deep learning models, we employed a systematic approach that included hyperparameter tuning and model architecture variations. In this section, we provide a detailed account of our methodology for each model type: Single LSTM, Multi-LSTM Layer, CNN followed by LSTM, and BiLSTM sandwiched between two LSTM layers.

2.5.1 *Hyperparameter tuning*

Hyperparameter tuning is a crucial aspect of model optimization, influencing the performance and generalization capability of deep learning models. To fine-tune our models, we adopted a grid search strategy, systematically exploring a range of hyperparameter combinations to identify the optimal settings. Specifically, we varied hyperparameters such as the learning rate, batch size, sequence length, number of hidden layers, activation functions, number of epochs, loss function, and hidden dimensions. We experimented with learning rates in the range of 0.1, 0.01, and 0.001 to find the most suitable rate for each model. Batch sizes 32, 50, and 64 were tested to assess their impact on training efficiency and model convergence. Sequence lengths, representing the duration of the EEG input window, were varied to determine the optimal length for capturing relevant patterns. We explored different configurations of hidden layers to strike a balance between model complexity and performance. Various activation functions, including ReLU, sigmoid, and tanh, were evaluated to identify the most suitable functions for each layer. The number of training epochs was tuned to avoid overfitting while ensuring model convergence. We experimented with values ranging from 100 to 200 epochs. We employed two loss functions during experimentation: SmoothL1Loss and Mean Square Error (MSE), assessing their impact on model performance. Our iterative approach to hyperparameter tuning allowed us to uncover the optimal combination for each model architecture.

2.5.2 *Model architectures*

To investigate the performance of different deep learning architectures for EEG-based user response prediction, we explored four main model configurations:

1. Single LSTM: This model consists of a single LSTM layer followed by a dense output layer. We employed this model to evaluate the effectiveness of a basic recurrent architecture in capturing temporal dependencies within the EEG data.

2. Multi-LSTM Layer: This model comprises multiple LSTM layers stacked sequentially, followed by a dense output layer. Stacking LSTM layers allows for hierarchical feature extraction, enabling the model to capture intricate patterns in the EEG data.

3. CNN followed by LSTM: In this model, EEG data passes through a 1D Convolutional Neural Network (CNN) before being fed into an LSTM layer and a subsequent dense output layer. Combining CNN and LSTM layers leverages the spatial and temporal characteristics of EEG data, potentially enhancing prediction accuracy.

4. Bidirectional LSTM (BiLSTM) sandwiched between two LSTM layers: This complex model configuration involves two LSTM layers, with a BiLSTM layer in the middle, followed by a dense output layer. The Bidirectional LSTM layer allows the model to consider context from both past and future time steps, potentially improving the prediction of user responses. These extensive experiments with varying hyperparameters and model architectures

ensured a comprehensive assessment of our deep learning models' capabilities and allowed us to select the most effective configurations for predicting user responses from EEG data.

3. Results

Various combinations were tried, out of which the three best-performing models were selected. These various combinations were tried with a dataset of different waves that is alpha, beta, theta, and on the combined dataset as well. The criterion for selecting the best model was accuracy and F1 score. For the 10-class Problem, results show that, out of all these models, best-performing models are: 2 CNN + 4 linear layers (64.47%), 1 LSTM + 4 linear layers (63.35%), and 1 LSTM + 1 BiLSTM + 1 LSTM + 4 linear (59.47%). And the best prediction occurred with the beta wave (Figure 2).

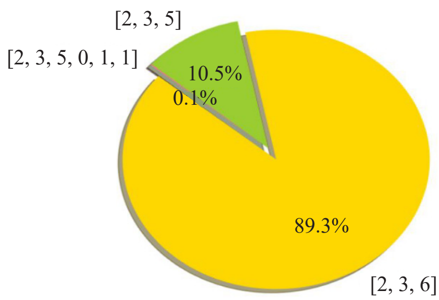


Figure 2. The results of the 10-class problem. Here, the chosen model predicted the Corsi block number sequence [2, 3, 6] with an accuracy of 89.3%, and the [2, 3, 5] with an accuracy of 10.5%

For binary classification, results show that, out of all these models, best-performing models are: 1 LSTM + 4 linear layers (93.30%), 2 LSTM + 4 linear (92.64 %), and 1 CNN + 4 linear layers (63.98%). Also, the beta wave performed better with the prediction (Figure 3).

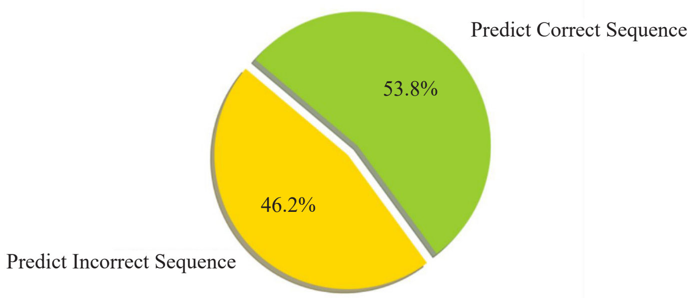


Figure 3. The results of the binary classification problem. Here, the chosen model predicted that the subject clicked the correct Corsi block number sequence with an accuracy of 53.8% and the incorrect Corsi block number sequence with an accuracy of 46.2%

Table 1 shows the accuracy for the afore-mentioned models, which was obtained from the test data as a validation of the models for both the binary as well as 10 class classification problems. This data was created as a split from the complete dataset of 30 subjects and was kept aside for validation.

In all the cases, data of the beta waveform performed better than the other waves. The reason might be that the beta wave is associated with the normal waking state of consciousness when attention is directed toward cognitive tasks and the outside world [16]. Beta has "fast" activity, present when a person is alert, attentive, engaged in problem-solving, judgment, decision-making, or focused mental activity [17, 18]. Furthermore, pattern recognition by the model in the case of combined data was quite difficult than in the case of when the data of each wave was taken individually.

Table 1. Model comparison results for test data

Model	Accuracy (%)	Precision	Recall	F1 Score
10 Class classifications				
2 CNN + 4 Linear Layers 2 CNN + 4 Linear Layers	45.54	0.70	0.85	0.57
1 LSTM + 4 Linear Layers	43.57	0.62	0.83	0.58
1 LSTM + 1 BiLSTM + 1 LSTM + 4 Linear Layers	42.79	0.61	0.80	0.57
Binary classifications				
2 CNN + 4 Linear Layers	62.65	0.75	0.85	0.66
1 LSTM + 4 Linear Layers	63.79	0.82	0.87	0.72
2 LSTM + 4 Linear Layers	60.93	0.80	0.80	0.71

4. Conclusion

The present study introduces an innovative framework aimed at developing deep learning algorithms designed to predict user responses based on fundamental spatial memory tasks. We explored various combinations of deep learning models, ultimately selecting the top-performing models based on their accuracy and F1 score. Our work holds significant practical implications for various domains. The accurate prediction of user responses from EEG data has applications in fields such as human-computer interaction, neurorehabilitation, and assistive technology. The ability to foresee user behaviour with high precision can lead to the development of more responsive and intuitive interfaces, improving the user experience. While our research primarily focuses on technological advancements, it is crucial to emphasize ethical considerations in the use of EEG data for predicting user responses. As EEG data is sensitive and can reveal insights into an individual's cognitive processes, privacy and informed consent must be respected. Future applications should adhere to ethical guidelines and ensure transparency in data collection and usage. Building upon our findings, future research can explore several directions. Firstly, investigating the robustness and adaptability of our deep learning models across different EEG acquisition setups and diverse participant demographics would enhance the generalizability of our approach. Additionally, incorporating real-time processing and feedback mechanisms could extend the utility of our models in dynamic applications, such as BCIs. Furthermore, the integration of multimodal data sources, such as EEG and eye-tracking, may provide a more comprehensive understanding of user behavior and cognition. Despite the promising results, our study has inherent limitations. The use of EEG data is subject to noise, and our models may be sensitive to variations in electrode placement and experimental conditions. Additionally, our dataset's size and participant pool could be expanded to increase the models' robustness. Furthermore, our models' performance may vary for individuals with neurological conditions or other health factors, which warrant further investigation. In conclusion, our work contributes to the growing body of research on predicting user responses from EEG data. We have addressed reviewer feedback and enhanced the clarity and completeness of our paper, ensuring that our deep learning-based framework can be better understood, replicated, and applied in practical applications. As technology continues to advance, the accurate prediction of user behavior from neural signals remains a promising avenue for improving human-computer interaction and brain-computer interfaces.

Acknowledgement

The research was supported by the director INMAS by providing all facilities for the execution of this study.

Conflict of interest

Authors have no conflict of interest to declare.

References

[1] Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*. 2019; 16(5): 051001. Available from: doi: 10.1088/1741-2552/ab260c.

[2] Wang P, Jiang A, Liu X, Shang J, Zhang L. LSTM-Based EEG Classification in motor imagery tasks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2018; 26(11): 2086-2095. Available from: doi: 10.1109/TNSRE.2018.2876129.

[3] Hasib MM, Nayak T, Huang Y. A hierarchical LSTM model with attention for modeling EEG non-stationarity for human decision prediction. In: *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. Las Vegas, NV, USA: IEEE; 2018. p. 104-107. Available from: doi: 10.1109/BHI.2018.8333380.

[4] Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI. A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. *Computers in Biology and Medicine*. 2018; 99: 24-37. Available from: doi: 10.1016/j.compbiomed.2018.05.019.

[5] Ni Z, Yuksel AC, Ni X, Mandel MI, Xie L. Confused or not confused? In: *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. New York, NY, USA: ACM; 2017. p. 241-246. Available from: doi: 10.1145/3107411.3107513.

[6] Michielli N, Acharya UR, Molinari F. Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals. *Computers in Biology and Medicine*. 2019; 106: 71-81. Available from: doi: 10.1016/j.compbiomed.2019.01.013.

[7] Fedjaev J. *Decoding EEG Brain Signals using Recurrent Neural Networks*. Germany: GRIN Verlag; 2019.

[8] Sun X, Qian C, Chen Z, Wu Z, Luo B, Pan G. Remembered or forgotten? - An EEG-based computational prediction approach. *PLoS One*. 2016; 11(12): e0167497. Available from: doi: 10.1371/journal.pone.0167497.

[9] Khan NU, Wan W, Riaz R, Jiang S, Wang X. Prediction and classification of user activities using machine learning models from location-based social network data. *Applied Sciences*. 2023; 13(6): 3517. Available from: doi: 10.3390/app13063517.

[10] Gharibshah Z, Zhu X, Hainline A, Conway M. Deep learning for user interest and response prediction in online display advertising. *Data Science and Engineering*. 2020; 5(1): 12-26. Available from: doi: 10.1007/s41019-019-00115-y.

[11] Matos LM, Cortez P, Mendes R, Moreau A. Using deep learning for mobile marketing user conversion prediction. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. Budapest, Hungary: IEEE; 2019. p. 1-8. Available from: doi: 10.1109/IJCNN.2019.8851888.

[12] Liguori A. Absorption and subjective effects of caffeine from coffee, cola and capsules. *Pharmacology Biochemstry and Behavior*. 1997; 58(3): 721-726. Available from: doi: 10.1016/S0091-3057(97)00003-8.

[13] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. 2004; 134(1): 9-21. Available from: doi: 10.1016/j.jneumeth.2003.10.009.

[14] Sharma G, Daniel R, Chandra S, Singh R. Effect of complexity on frontal event related desynchronisation in mental rotation task. *Applied Psychophysiology and Biofeedback*. 2019; 44: 235-245. Available from: doi: 10.1007/s10484-019-09436-0.

[15] Sharma G, Anto A, Singh V. Sex differences in mental rotation: Cortical functional connectivity using direct transfer function. *Biomedical Signal Processing and Control*. 2018; 40: 425-432. Available from: doi: 10.1016/j.bspc.2017.10.007.

[16] Dustman RE, Boswell RS, Porter PB. Beta brain waves as an index of alertness. *Science*. 1962; 137(3529): 533-

534. Available from: doi: 10.1126/science.137.3529.533.

[17] Hosťovecký M, Babušiak B. Brain activity: beta wave analysis of 2D and 3D serious games using EEG. *Journal of Applied Mathematics, Statistics and Informatics*. 2017; 13(2): 39-53. Available from: doi: 10.1515/jamsi-2017-0008.

[18] Soltani A, Vakili AM, Sho'ouri N. Beta wave activity analysis of EEG during mental painting reflects influence of artistic expertise. In: *2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)*. Tehran, Iran: IEEE; 2019. p. 225-229. Available from: doi: 10.1109/ICBME49163.2019.9030399.