IUST Iran University of Science and Technology

Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Deep learning for Robust EEG Signal Forecasting using Long Short Time Memory Neural Network

Duaa A. Kareem*, Zaineb M. Alhakeem**(C.A.) , Nawar Hayder Tawfeeq Alqrnawi***, Batool Dahham Al-Ali*** and Heba Hakim****

Abstract: Signal forecasting in the medical field has many applications such as signal correction and anomaly detection. According to this application, robust forecasting is required to obtain a signal identical to the original signal. This study proposes a forecasting technique that obtains a robust signal that can be used in different applications. A long short-term memory neural network (LSTM-NN) was used to predict future samples from present and past samples. An Electroencephalography (EEG) dataset was used to test this technique. Four channels were used as input examples, one of which was the predicted output. All four channel samples were fed into the four networks to predict the future samples. To decrease complexity, only one hidden layer is used for this purpose. The statistical results are promising for applications that require an almost perfectly predicted signal. The number of hidden cells is first very low (five cells only), which gives a Root Mean Square Error of less than 20, whereas when the number of hidden cells is increased to 100, the Root Mean Square Error (RMSE) is approximately 7.5 for all four channels.

Keywords: EEG signal, Robust Forecasting, LSTM, GRU, Deep Learning.

1 Introduction

Information from the visual sensory system travels in the form of electrical signals to various brain areas. Different features (color, shape, and size) were initially processed separately. However, the brain must integrate these features to form an overall perception of objects [1].Understanding the importance of binding in individuals with brain injuries, strokes, or other conditions that affect brain function. Often, individuals with brain injuries have difficulty associating features, and thus, poor perception of things correctly. The complex dynamic behavior of brain activity is represented by recording electrical activity using an electroencephalogram [2].Smart health sensors and communication technology have advanced dramatically, enabling healthcare to provide faster and more accurate services such as remote monitoring. Electroencephalography (EEG) was used to measure and analyze the electrical activity of the brain. It is a non-invasive method that records electrical signals in the brain when examining senses such as hearing and vision, allowing healthcare providers to monitor brain function in real time [3]. The use of devices that integrate sensors

Iranian Journal of Electrical & Electronic Engineering, YYYY.
Paper first received DD MONTH YYYY and accepted DD MONTH YYYY.
* The author is with the Polymers and Petrochemicals Engineering Department, College of Oil and Gas Engineering ,Basrah University for Oil and Gas, Iraq, Basrah.
E-mail: doaa.abas@buog.edu.iq.
** The authors are with the Department of Chemical and Petroleum Refining Engineering, College of Oil and Gas Engineering ,Basrah University for Oil and Gas, Iraq, Basrah.
E-mails: zainebalhakeem@buog.edu.iq.
*** The authors are with the Department of Oil and Gas Engineering, College of Oil and Gas Engineering ,Basrah University for Oil and Gas, Iraq, Basrah.
E-mails: nawar.hayder@buog.edu.iq, batool.dahham@buog.edu.iq.
**** The authors are with the Computer Engineering Department, College of Engineering,Basrah University of Basrah, Iraq, Basrah.
E-mails: hiba.abdulzahrah@uobasrah.edu.iq.

Corresponding Author: Zaineb M. Alhakeem.

such as pulse and heart rate monitors, in addition to some vital signs, has increased owing to the rapid development and widespread use of wearable computing devices. These devices are characterized by their ability to record and store large amounts of data and reuse them in other fields [4]. Electroencephalography (EEG) is widely used to monitor neural activity in the brain in a non-invasive manner. An important subfield of EEG signal analysis is the prediction task, in which one or more future values $\{x_t,..., x_{t+m}\}$ of an EEG time series are estimated based on past values $\{x_{t-n},...,x_{t-1}\}$ of the same series. Recently, it has gained increasing attention owing to the increasing demand for real-time EEG. Real-time EEG is used in applications such as brain-computer interfaces (BCI) and closed-loop neurostimulation, where a prediction model is sometimes required to allow time for algorithmic decision-making. Prediction methodologies rely on modern deep learning-based methods, such as recurrent neural networks (RNN) and long short-term memory (LSTM), which have an advanced ability to handle nonlinear temporal relationships and accurately predict future patterns in brain signals [5], [6], [7], [8]. Studies conducted by many researchers using machine learning-based algorithms and methods to predict EEG signals have demonstrated their ability to analyze temporal biometric data [9]. Researchers have used different algorithms and machine learning techniques to predict EEG signals. Deep Learning: Long-term memory is the best technique for analyzing time-series datasets. Because EEG recordings are collected continuously over time, long-term memory is the best method for analysis [10]. We first focus on RNNs because they are a type of RNN, and because RNNs are simpler systems, the intuition gained from analyzing RNNs also applies to LSTMs. Importantly, the standard RNN equations, which we derived from differential equations, serve as a starting model that provides a clear logical path to eventually arrive at the LSTM system architecture [11]. EEG signals are increasingly coupled with deep learning and machine learning techniques to perform various tasks such as human response prediction, human visual feature representation, representation learning, emotion detection, and motor imagery signal classification. EEG signals are also used in the development of brain-computer interface systems, where external devices are controlled by intellectual commands through real-time automated analysis of EEG signals [12], [13], [14]. An advanced type of RNNs is the LSTM network, which was designed to address the gradual fading problem faced by traditional recurrent networks when working with long-term temporal data. LSTM networks can store data for long periods and are, therefore, suitable for analyzing and predicting nonlinear electrical signals in the brain. In a study published in 2018, LSTM networks were used to predict epileptic seizures by using EEG signals. The results showed that the model could improve the quality of life of patients and predict seizures long before they occur [15].

This study aims to use LSTM networks to improve the accuracy of EEG signal prediction, which is a deep-learning technique. This research focuses on utilizing the capabilities of LSTM to develop an efficient and robust model (robust) to predict EEG signals that handle complex and continuous temporal data, improve the accuracy of predicting neurological conditions, such as epileptic seizures, sleep, or other mental activities, and compare LSTM with traditional models.

To understand the complex patterns in EEG signals and improve predictions for medical and scientific applications, the performance of LSTM was compared with other techniques such as linear models (ARIMA, NARX) or traditional RNNs.

This research also aims to support doctors in early diagnosis and prediction before an appropriate time period for neurological diseases such as epilepsy and Parkinson's disease, improve brain-computer interface (BCI) systems through a better understanding of neural signals, address challenges in EEG analysis, and reduce noise in neural signals.

The expected results are higher prediction accuracy of EEG signals compared to traditional models, the possibility of applying the model in health systems and neurological research related to epilepsy, and improving brain-computer interfaces through a deeper understanding of EEG signals.

2 Literature Review

Brinkman et al. (2015) suggested in their study a method using a logistic regression machine learning algorithm with spectral power in conventional Berger bands as features and achieved greater prediction rates of seizures in the foreground (>4 h a part) than the time-matched chance predictor in 2/3 dogs. This study illustrates the development and validation of a Support Vector Machine (SVM) approach to predict seizures using precordial bowers (PIB) and interelectrode synchronization features computed from prolonged and moving IEEG recordings from naturally occurring epileptic dogs. The optimal pre-seizure time window for seizure prediction was investigated, and an analysis was performed on the effect of multiple PIB features and individual electrode pairs on inter-electrode synchronization features [16].

Munoz-Almaraz (2017) proposed a new method for EEG signal pre-processing based on supervised filters. These filters are used in machine learning algorithms, such as K-Nearest Neighbor (KNN), to improve seizure prediction. These preprocessing methods were analyzed in detail in terms of the Receiver Operating

Characteristic (ROC) curve and the area under the ROC curve. These results show significant statistical improvements compared with typical baseline spectral Power Band Filtering (PBF). Supervised filters provide better information than traditional PBF filters as the dataset grows in terms of monitored variable electrodes and length of time [17].

Nguyen et al. (2020) in their work, they show the discrepancy between the real and model-estimated EEG signals can be significantly reduced if the EEG model is not generated from a time series collected from randomly selected neurons of the neural network, but rather using a compressed set of independent vectors generated by Principal Component Analysis (PCA) from the entire data series generated by the network. We also show that networks of coupled Kuramoto phase oscillators can be used to model EEG data in addition to relying on neural model networks. In addition, the results confirm that the best prediction EEG signal models were obtained when the oscillators were weakly connected, satisfying the conditions for the existence of a CAS system. This study demonstrates the ability to reproduce and predict healthy and epileptic EEG signals, as well as the characteristics of EEG signals, Hurst exponent, and power spectrum of experimental EEG signals [18].

Shakeel et al. (2020) proposed an approach for the forward prediction of a time series by comparing an autoregressive (AR) model and an adaptive method (Least Mean Square (LMS)-based AR model) to reduce the error because its parameters can be dynamically adjusted and the AR model performs better in real time for long prediction periods. The results indicate that for shorter data segments of approximately 128ms, the AR model outperformed the LMS-based AR model, whereas for longer prediction lengths of 256ms, the LMS-based AR model outperformed the AR model. Low-computational-cost learning management systems can help predict the EEG phase (alpha oscillations) [19].

Mathur et al. (2021) conducted a comparative experiment between two stages of visual processing to identify the distinct EEG signals associated with each stage. Changes associated with color and shape coordination were detected after (100ms and 1500ms). The two stages were identified using the corresponding study and test intervals, where features were extracted from the time and frequency domains. These features were used to train a set of machine learning classifiers separately. The frequency-domain temporal representation of the signal was applied during training to a convolutional neural network (CNN), and the obtained results achieved promising performance. This study contributes to the literature in two ways: First, it presents a data analysis using deep learning techniques to classify whether a trial belongs to the (100 ms or 1500ms) category, while the second links these findings to identify different stages of visual processing in the human brain and linking visual features.

Thara et al. (2022) developed a simple prediction method consisting of long short-term memory (LSTM) with a windowing technique using EEG. The window size was increased from (5 to 20 steps), which increased the captured temporal information and, thus, the prediction accuracy. The neural network learns the past 20time steps to predict the future EEG signals in two stages. In the univariate method, only one feature was used as the basis for predicting the future values. In the multivariate method, 42 features were used to predict future EEG. Consequently, the multivariate method is more robust and provides a prediction almost equal to the actual target value. In the univariate method, the accuracy was approximately 70%, whereas in the multivariate method, it was 90%)[10]. Hanna et al. (2024) enhanced the prediction of resting-state EEG signals by applying deep learning to obtain robust, long-term predictions. Methods: They applied the deep neural network model WaveNet to predict resting-state EEG signals in the theta (4-7.5 Hz) and alpha (8-13 Hz) frequency bands. They also compared WaveNet with an AR model, which has been widely used in real-time EEG applications. The results showed that WaveNet reliably predicted EEG signals at both the theta and alpha frequencies. It outperformed the AR model in terms of signal amplitude and phase estimation. They concluded for the first time that deep learning can be used to predict resting-state EEG time series over 100ms [5].

The technique suggested in this study is illustrated in Figure 1. The contributions of this study are as follows:

1. High performance forecasting model with different parameters

2. Robust predicted Signal to be used in different applications

3. Robust model with low size training model

This paper consists of the following sections: the first is the discretion of the dataset used, the second is the method used for forecasting, after which the performance evaluation indices finally provide the results and discussion.

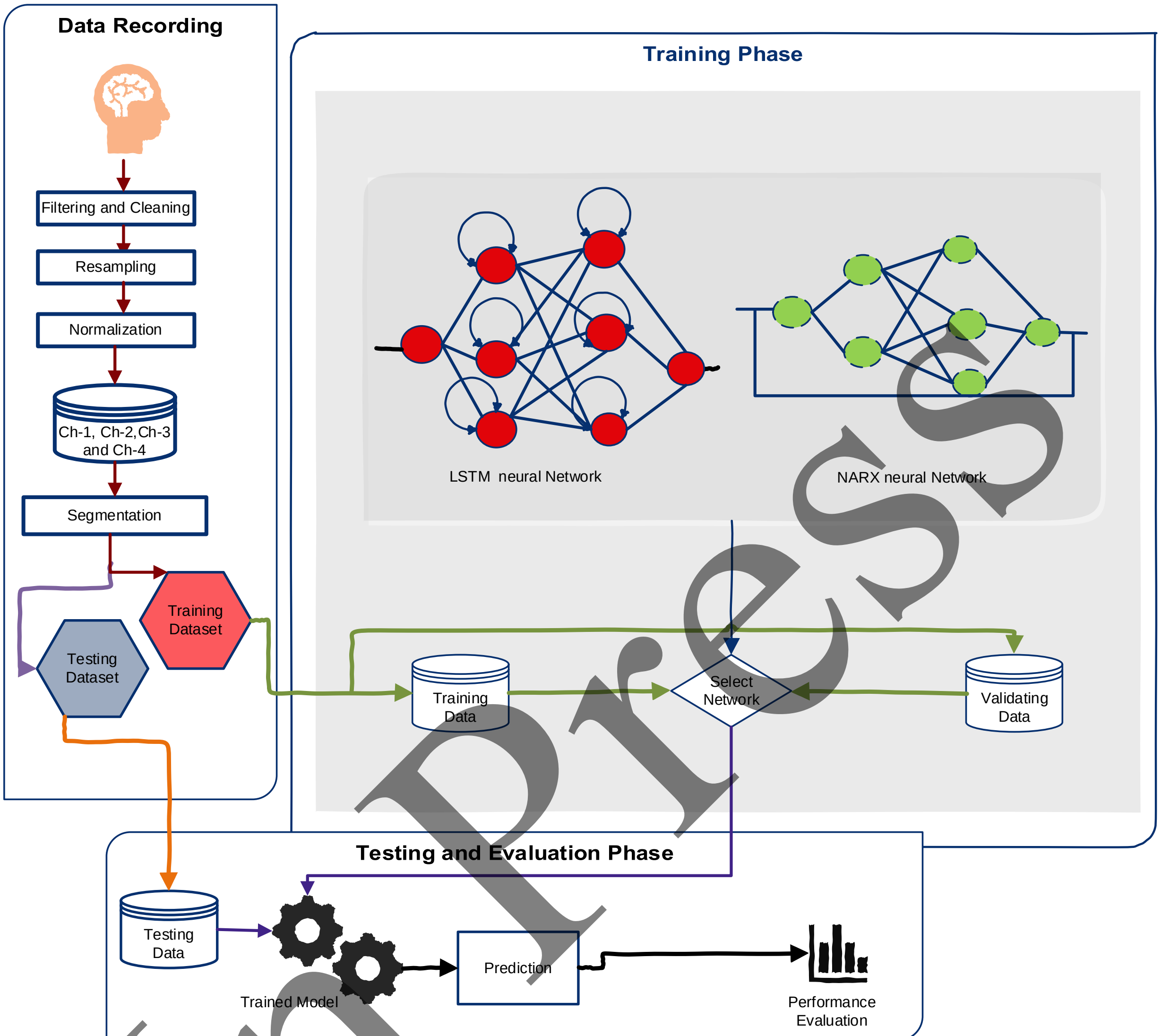


Fig 1. The overall system discretion

3 EEG Dataset Characteristics

The data used in this study were obtained from single subject recordings. This dataset was recorded by Professor Wang Guangjun at the Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences in Beijing, China [20]. This dataset was recorded for a thirty-nine-year-old right-handed male who had corrected eyesight and no history of mental or physical disease that affected neurological signals. This person avoided any material that could lead to anomalies in the brain signals, such as caffeine and alcohol, for at least 24 h before the recording session.

During the recording, the subject sat comfortably on a seat in front of a display that showed a red cross to start the recording, during which the subject did not blink or close their eyes. Approximately 60 records were recorded for this patient, using the same procedure. The recording sensors were 32 wet electrodes, all of which were in the standard locations (10-20) standard as shown in Figure 2. The sampling rate of the signal is 20 kHz. This dataset records two other types of data for the same person and task: magnetic resonance imaging (MRI) and electrocardiogram (ECG) signals.

4 Methodology

Deep learning is used in many fields such as classification and prediction. In this study, a long short-term memory network (LSTM) was used to predict the medical signal of one sample ahead.

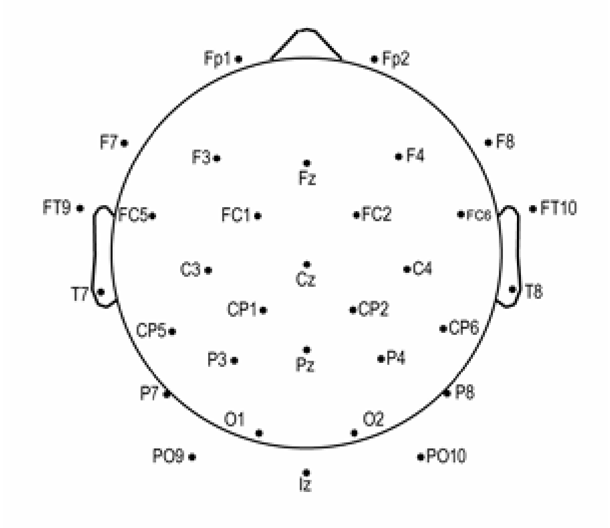


Fig 2. The EEG sensors distribution [20]

4.1 Long Short-Term Memory (LSTM) Networks

LSTM refers to long short term memory which are the enhanced variation of RNNs. These networks are designed to handle sequential data. thus, LSTM can integrate customized methods to retain information unlike the standard RNNs. this unique feature provides the ability use old data to predict upcoming outcomes.

a vital challenge for RNNs is the vanishing of gradient issues which happens during back propagation. here, the model keeps updating the weight of the neural network by gradients. but, when facing with long sequences, the gradients become very small which leads to very little updates in the network, stopping deeper layers from learning. resulting the traditional RNNs to fail the maintenance of information over time, and unsuitable to preform complex sequences [21].LSTM holds three gates to dictate the information processed, Firstly the Input Gate which determines whether the input data should be retained or left, only relevant data shall be carried forward. Input Gate: processes new incoming information, to be decided later if the data should be added to the memory state. Secondly the Output Gate: refines and forwards updated data to Next state, which later helps in predictions via the structure of gates, LSTM manages the memory, letting it handles long-term dependencies in better demeanor than traditional RNNs [22].

One of the recurrent networks is the long short-term memory, which is used to solve the backflow problem. Four vectors were introduced into the network: Input Activation, Output Activation, Forget Activation. The final vector was the candidate vector. Figure 3 shows the structure of a single cell of the LSTM, where the input sequences are H_k,X_k and C_k are the cell output vector, input sequence, and cell state vector to the network cell, respectively. The vectors are applied to the activation function (sigma), and the function (hyperbolic tangent (tanh)) computes the states of cells C [23].

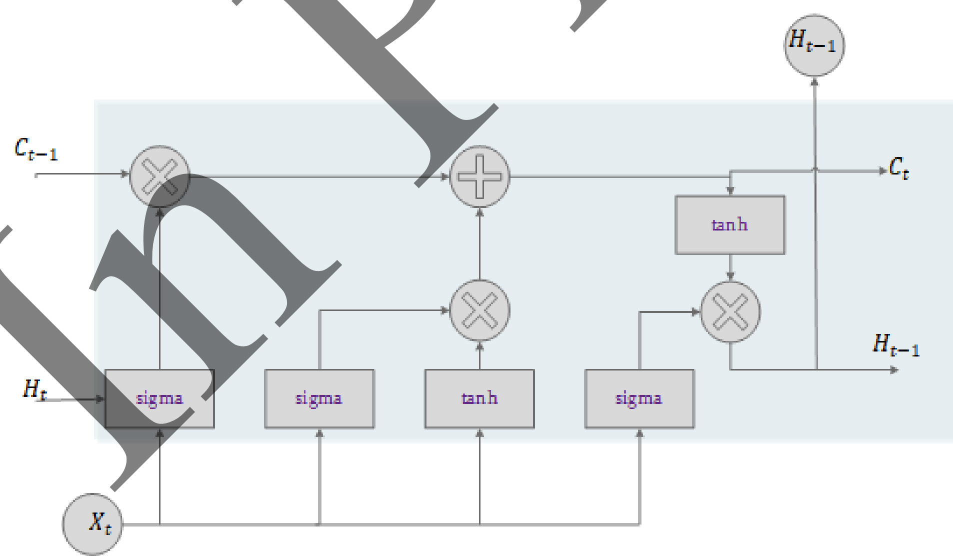


Fig 3. The hidden cell of the network

Some the LSTM applications is in the Internet of Things (IoT) industrial in manufacturing, sensors, devices, and numerous machines generates a large among of operational information, if these data were overlooked in most cases it will lead to failure, shutdown, and expensive downtime. By accessing the historical data, LSTM models are able to aid in the scheduling and predicting maintenance plans and anomalies before failures. The crucial parameters including temperature, energy consumption and vibration levels are analyzed over time to predict potential challenges in the process. a proactive approach is always recommended to shrink the downtime, and increase productivity and lifespan of the intended equipment [24].

Moreover, in the energy sector, LSTM will play a crucial part in the forecasting procedures. examining previous data and patterns anticipating future energy requirements for each location or zone. by this method, a much better power distribution, power handling and overall enhancements are achieved.

Unlike the traditional RNNs, that fails in the long term dependencies, LSTM can be an advanced memory mechanism to retain vital data over period frames. LSTM are able to use selectivity to store, edit, and discard information which makes them highly reliable for tasks including sequential processing, like speed recognition, financial prediction and maintenance. LSTM provides more reliability for complex machines, ensuring advanced accuracy and performance in a wide range of industries unlike the limitations of standard RNNs

In this study, a single hidden-layer LSTM network was used for signal regression. LSTM is employed to predict a single sample using a sequence of fixed-length segments from the dataset. The LSTM used in this study has one hidden layer and four other layers (input, output, fully connected, and regression layers). Figure 4 illustrates the structure of the NN, which consists of five layers. This network is used to forecast the Next sample of EEG signal using the history of the signal.

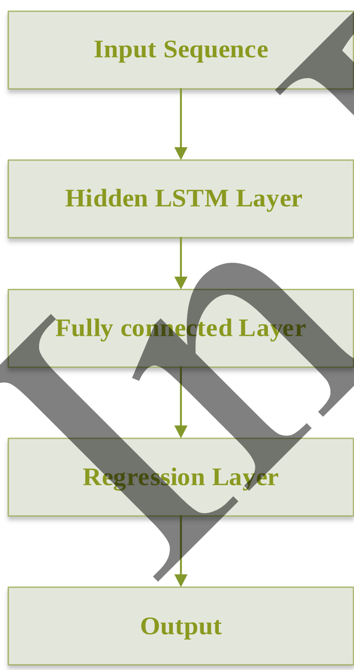


Fig 4. The proposed LSTM structure

5 Performance Evaluation Tools

In this work there are two signals the original signal and the predicted signal to compare between these signals the following sstatistical indices are used: Maximum Absolute Error (*MxAE*), Sum of Squar Error (*SSE*), *MAE*, Root Mean Squared Error (*RMSE*, Total Sum of Squares (*SST*) and Relative Absolute Error (*RAE*). These indices are defined as follows:

$$\mathrm{SSE}=\sum_{i=1}^{n}(y_i - yp_i)^2 \quad (1)$$

where y_i is the original signal, y_{pi} is the predicted signal, and n is the number of samples in both the original and the predicted signals.

$$\mathrm{SST}=\sum_{i=1}^{n}(y_i - \bar{y})^2 \quad (2)$$

$$\mathrm{Max.\ AE} = \max|y_i - yp_i| \quad (3)$$

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(y_i - yp_i)^2}{n-2}} \quad (4)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - yp_i| \quad (5)$$

$$\mathrm{RAE}=\frac{\sum_{i=1}^{n}|Y_i - Yp_i|}{\sum_{i=1}^{n}|Y_i - \bar{Y}|} \quad (6)$$

6 Simulation Results and Discussion

The EEG signal can be corrupted during transformation from the sensors (source) to the processing device (destination). Therefore, robust prediction of the signals is important to keep the signals corrected. The following sections explain the results of this study:

6.1 Data pre-processing

The data were preprocessed to make them suitable for this method. preprocessing started from the filtering process to remove artifacts that contaminate the EEG signal during recording. Some of these artifacts are signals in the electric lines near recording sensors. First, data were resampled at a low sampling rate of 500 Hz. The first filter was a high-pass filter (0.5 Hz), and then, a low-pass filter at 70 Hz was applied to the same signal.

Subsequently, a cleaning step was applied to the filtered data to denoise them. The reference was averaged, and then a band-pass (2-30) Hz filter was applied to the signal. After the cleaning step, the data were segmented into a 1 second window and bad windows were nominated from the recorded data. Any possible artifacts such as eye blinking and jaw movements were removed using independent component analysis (ICA). Figure 5 shows the steps involved in the preprocessing phase.

6.2 Data evaluated

The dataset included one subject who recorded three different signals (EEG, ECG, and MRI signals). In this study, only four EEG signals recorded by 32 sensors are used to test the proposed model. First, the signals were resampled and averaged to expand and enhance them according to the noise.

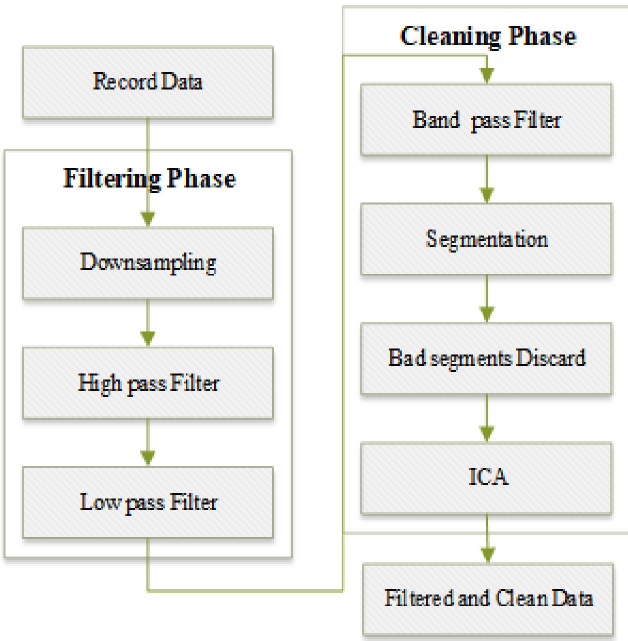


Fig 5. Detailed Preprocessing phase

After all enhancements, the signal was segmented into multiple sliding windows, each of which had 400 samples shifted back to one sample, which was used to predict the current sample of only one channel, as shown in Figure 6. All the channels were segmented using the same technique, and each segment consisted of 400 samples. The four segments are fed into the network to predict a new sample for a specific channel.

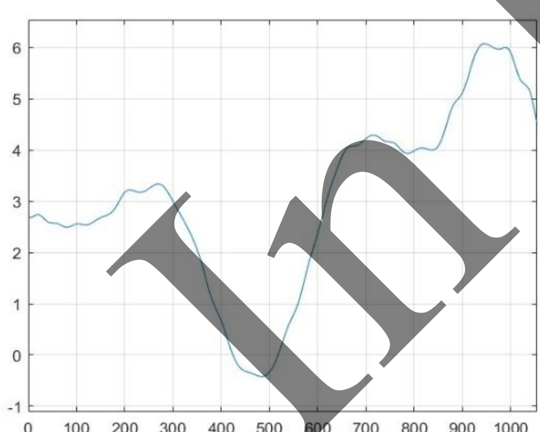


Fig 6. The segmentation of the EEG ch-1

6.3 Model training

To train the model, only four EEG channels were used; after the filtering, cleaning, and segmenting steps, the training step was started. The model was trained using the LSTM network. The input of the network consisted of four segments with 400 samples and the output was one sample of a single channel. Figure 7 illustrates the steps of the overall model. The dataset was divided into two sets (50% training and 50% testing). The training set consisted of four segments of equal length that were the inputs of the neural network. The output of the signal is one sample of a specific channel, which is the current sample, whereas the inputs end with the previous samples. Therefore, the network was not trained on the same samples at the same time.

As we used four channels for training and testing the network, four networks were used for training to output four new samples for each channel. Figure 8 shows the inputs and outputs of the network.

The LSTM parameters are listed in Table 1. The initial learning rate of the network was 0.01 with a drop period every two epochs, with a factor of dropping equals to 0.1. The mini-batch size was set to 20. Training was conducted for only ten epochs to determine the luts with the minimum training time. Only one hidden layer of 100 cells was used to reduce the complexity of the model. Stochastic Gradient Descent with Momentum (SGDM) optimizer was used. Figure 9 shows the training trends of RMES for the fourth channel. The error started high at approximately 2 in the first epoch and then decreased to less than 0.5 in the second epoch. This means that the model could be used with the minimum number of epochs and there was no need to increase the training time

The loss of the training was close to zero from the beginning of the training for the ch-4 of EEg signal, as shown in Figure 10.

The prediction results are presented in Figures (11–14). In general, the results of the regression using LSTM showed very high performance. In Figure 11, the predicted signal of Channel 1 is very close to the real signal, as shown in the magnified section. The difference between the two signals has a very small value, close to zero, as shown in the distribution of the error in the histogram. The number of samples with an error of approximately zero was approximately 500 for the entire signal and the maximum error was approximately 2.5, with a very low number of samples in this region.

Figure 12 illistrate the prediction of the second channel that also shows the same behavior as the first channel, with very low error value with the same distribution of error like the first channel. Number of samples with error around zero is less than that of the first channel.

The results shown in Figure 13 exhibit the same regression quality as those of the previous two channels, with an error distribution in the ±1 range. Figure 14 shows the same behavior as that of the third-channel prediction. These results are promising but not very clear, and statistics are needed for the final decision.

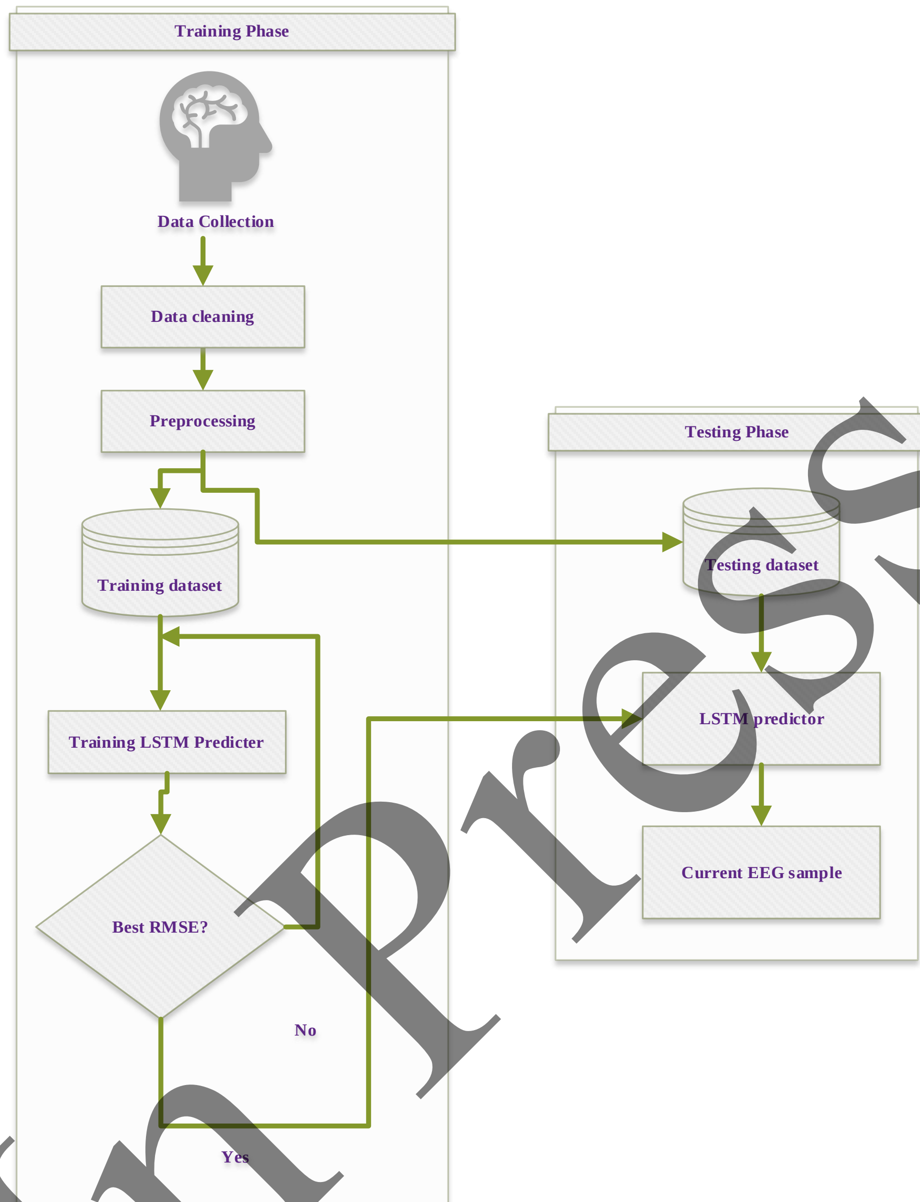


Fig 7. The steps of proposed prediction model

Table 1. LSTM hyperparameters.

Hyperparameters	Value
Learning rate	0.01
Epochs	10
Solver	SGDM Optimizer
Number of layers	1
Number of hidden cells by layer	5
Mini-Batch Size	20
Training, Testing set ratio	50:50

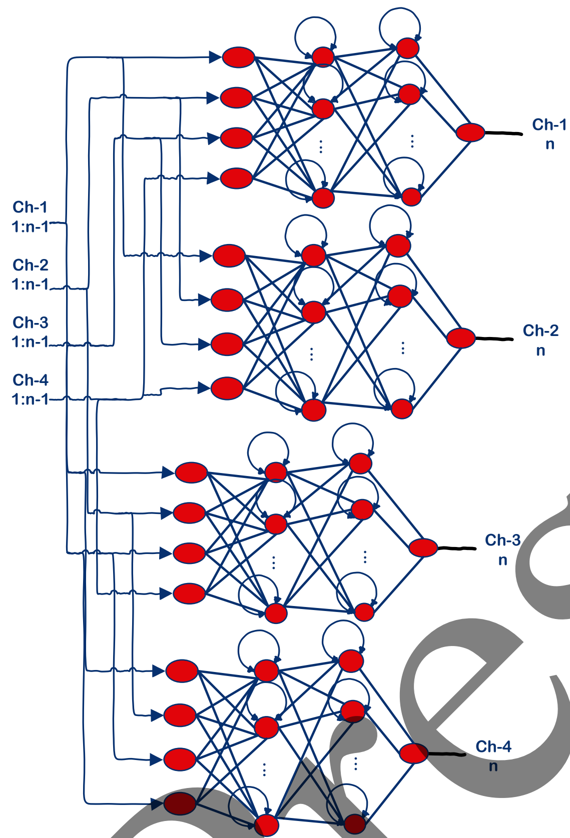


Fig 8. The training LSTM Neural Network

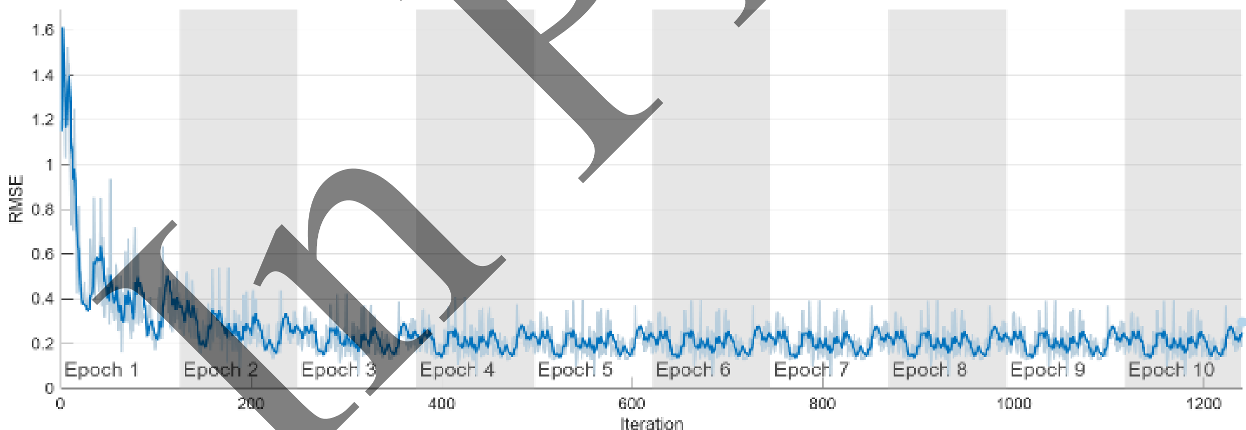


Fig 9. RMSE of training process during the iterations for prediction Ch-4 by LSTM

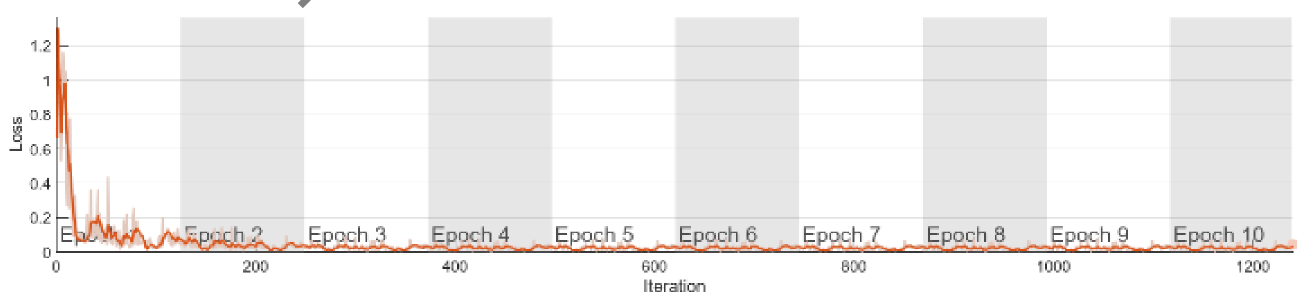
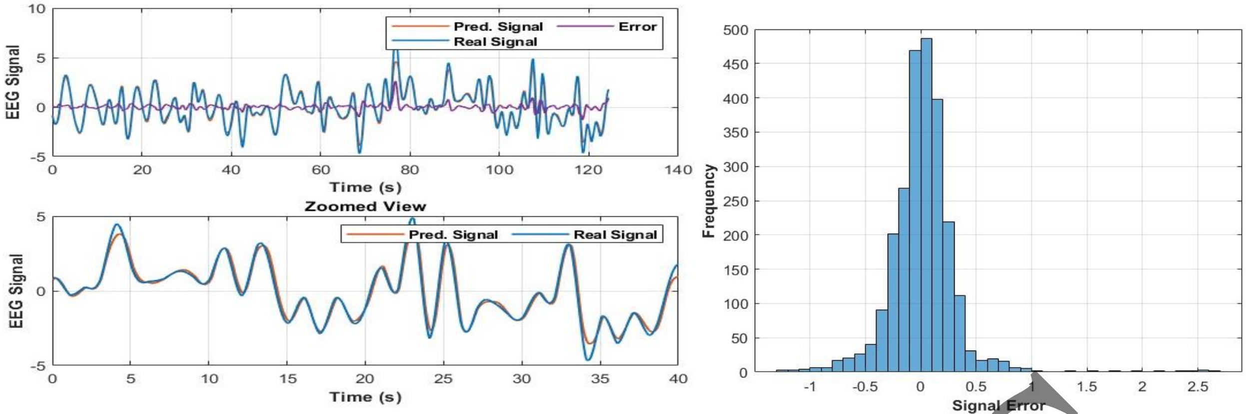


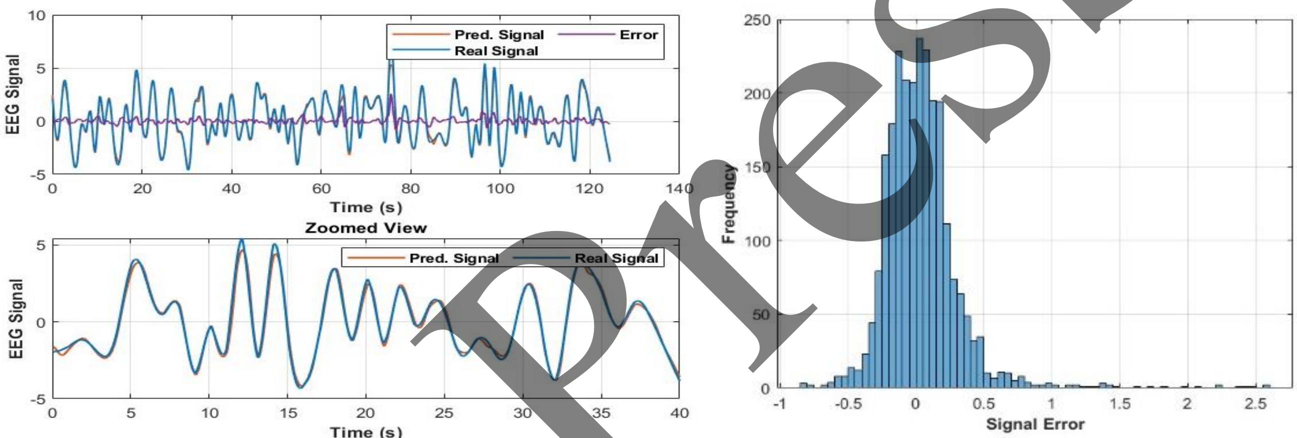
Fig 10. Loss of training process during iteration for prediction Ch-4 by LSTM



a) Original Signal Vs. Predicted signal

b) Error Distribution

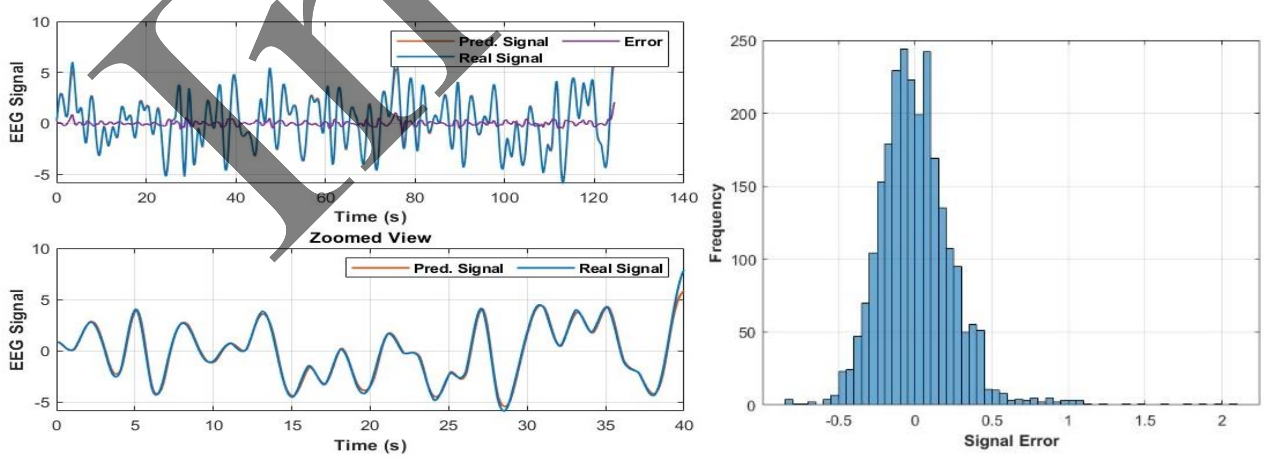
Fig 11. Forecasting EEG-Ch1 using LSTM



Original Signal Vs. Predicted signal

b) Error Distribution

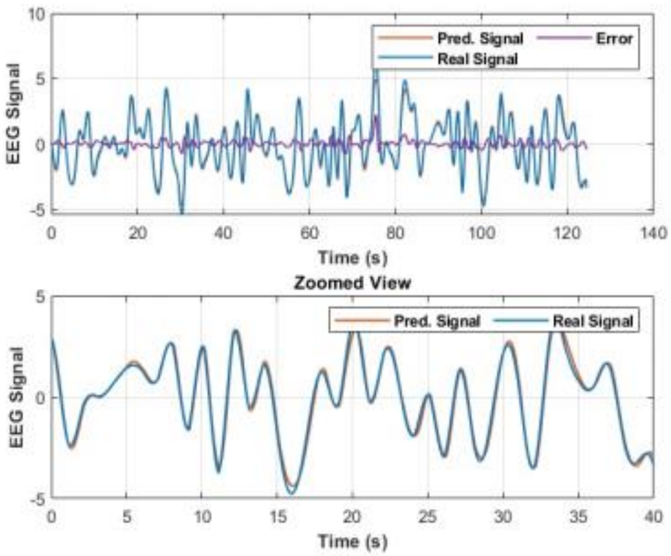
Fig 12. Forecasting EEG-Ch2 using LSTM



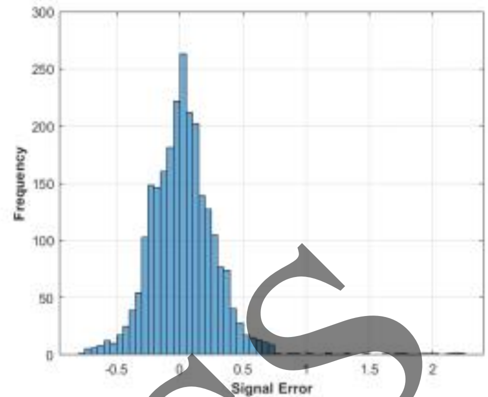
a) Original Signal Vs. Predicted signal

b) Error Distribution

Fig 13. Forecasting EEG-Ch3 using LSTM



a)Original Signal Vs. Predicted



b) Error Distribution

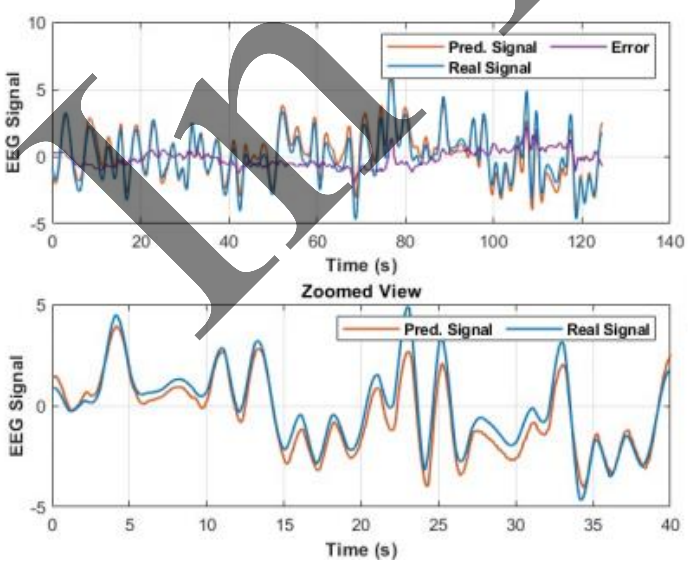
Fig 14. Forecasting EEG-Ch4 using LSTM

6.4 Comparison results of the LSTM model with Different Models

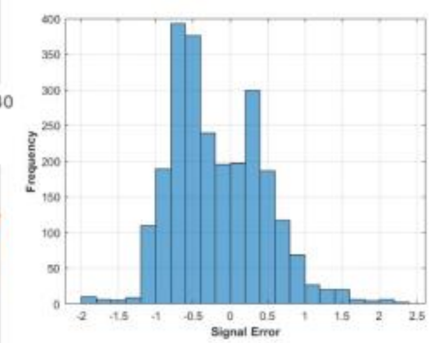
For an in-depth comparison, a CNN was used to predict the same signals to evaluate the proposed model. The CNN was trained using the same dataset with the same ratio as the training and testing groups. The training optimizer of the CNN is Adaptive Moment Estimation (ADAM) with 11 layers. The network was trained using 50 epochs to obtain results. The same parameters of LSTM could not be used to train the CNN because the results were not acceptable for comparison.

Figures (15–18) show the regression results using CNN for the four channels. It is obvious that the error is very high, and the prediction quality is lower than that of the LSTM.

In Figure 15, the prediction of the signal is close, but not robust, and can be considered as a replacement for the original signal when necessary to correct corrupted samples. Sudden changes could not be followed correctly by the CNN over time. The distribution of error is within the same range as LSTM, but the number of samples is in the range of more than 0.5, because the absolute value is higher than that of LSTM.



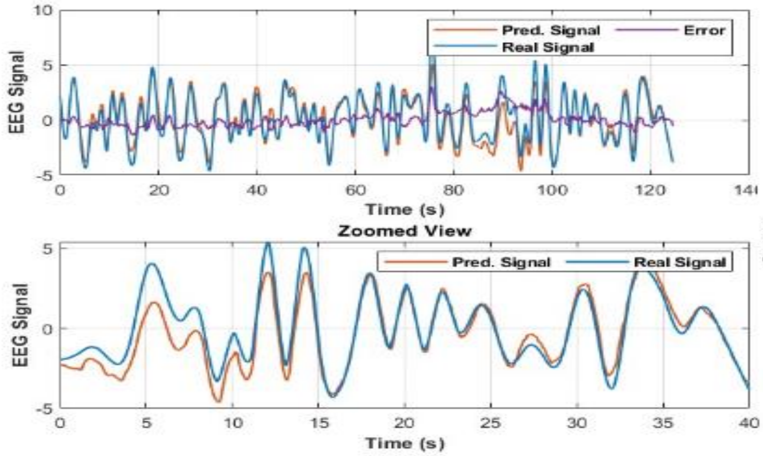
a) Original Signal Vs. Predicted signal

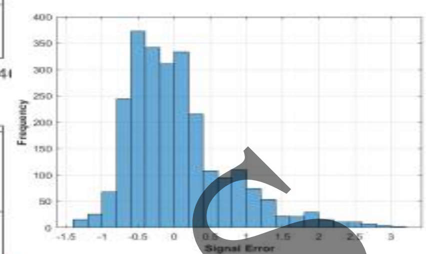


b) Error Distribution

Fig 15. Forecasting EEG-Ch1 using CNN

Figure 16 shows a worse prediction behavior than that of ch-1 for more samples in the range of ±1. In the time period (80-100) time sample is the worst case of prediction. Ch-3 in Figure 17 exhibits a better prediction, but is not similar to that of LSTM. The hills and valleys of the signals could not be predicted well because of the sudden changes that the CNN could not detect during regression, as shown in Figure 18.

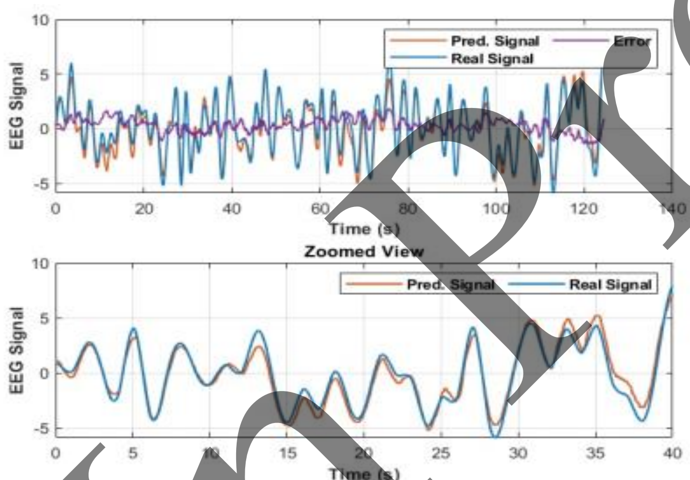


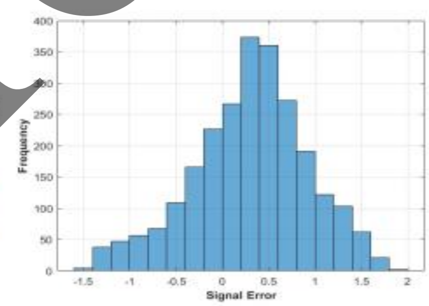


b) Error Distribution

a) Original Signal Vs. Predicted signal

Fig 16. Forecasting EEG-Ch2 using CNN





a) Original Signal Vs. Predicted signal

b) Error Distribution

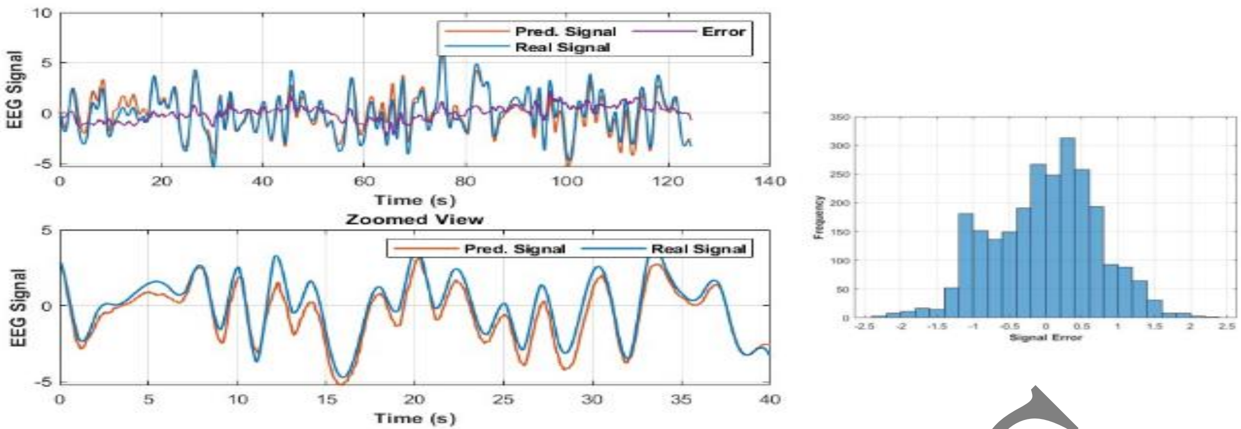
Fig 17. Forecasting EEG-Ch3 using CNN

Although CNN has more hidden layers and training epochs, the results do not compete with those of LSTM. Therefore, CNN was nominated from the comparison because of the complexity of the model.

Table 2 shows the statistics of the prediction using the multiple types of networks (LSTM, NARX, GRU, Bi-LSTM, and CNN) with the same training parameters. This table shows that the LSTM generally has a lower RMSE. NARX shows a lower Mx.AE for all channels compared to all other NN. This means that NARX gives the same peaks for the signals, but other types of errors are not satisfactory, especially RMSE, which has a higher value. A high RMSE indicates that NARX cannot predict the signal correctly.

Bi-LSTM shows the worst case of Mx.AE values for all channels, which makes this NN out of the competition among the other types because of the high value of error. Other types of performance metrics were used for compression. The RMSE shows lower values than LSTM, but still better than the worst case of NARX.

Finally, the strong comparison between GRU and LSTM shows that the results are very close to each other. For GRU, ch-2 exhibited a lower RMSE value than the other three channels, whereas in LSTM, all other channels had the best results. The GRU shows a higher maximum value of the absolute error than the LSTM, despite the better RMSE values in some channels.



a) Original Signal Vs. Predicted signal

b) Error Distribution

Fig 18. Forecasting EEG-Ch4 using CNN

Table 2. Outline of statistical indices results of EEG forecasting using different types of NN with five hidden cells

Model	Chan. No.	MxAE	RMSE	MAE	RAE
LSTM	1	2.6768062	15.812100	0.200235	0.030634
	2	2.5685658	14.427481	0.188807	0.0187714
	3	2.1621144	13.067456	0.193952	0.011361
	4	2.2210586	13.982861	0.195552	0.019178
NARX	1	1.6310832	28.116042	0.292031	0.04187575
	2	1.78791479	42.500152	0.518325	0.05570975
	3	1.19158608	32.591729	0.388232	0.03510196
	4	1.96668688	49.101973	0.514902	0.11363676
GRU	1	2.8111789	15.473086	0.18679503	0.029335005
	2	2.6547027	12.598344	0.15735331	0.014313413
	3	2.5734432	17.286928	0.26051018	0.019883901
	4	2.4877760	14.731150	0.19636081	0.021286083
BI-LSTM	1	3.0792029	16.949270	0.18097998	0.035199314
	2	3.5989432	15.303142	0.15776423	0.021119202
	3	2.7565539	11.846808	0.15699200	0.0093383184
	4	3.0758986	16.283195	0.20641261	0.026007688
CNN	1	2.2152280	32.366886	0.54837	0.128361
	2	5.3532526	65.3640127	1.014795	0.3852953
	3	3.0361113	47.305483	0.768114	0.1488979
	4	2.3099112	36.4934157	0.589113	0.1306325

The MAE values are approximately the same for both GRU and LSTM, which means that the average errors for both are in the same range, that is (0.15-0.26). This shows that the overall error value was approximately 0.2. The relative error is almost the same as the superiority of LSTM over the GRU. These results lead to the selection of LSTM for enhancement to achieve better results.

To enhance the results of the LSTM model, the number of hidden cells was increased to 100, which decreased the maximum error by half. All performance indices were enhanced with an increase in the number of cells in the hidden layer, as shown in Table 3.

Table 3. Outline of statistical indices results of EEG forecasting using LSTM-100

Model	Chan. No.	MxAE	RMSE	MAE	RAE
LSTM	1	1.28709	7.348488	0.085942	0.006616
	2	1.282771	7.102759	0.092088	0.00455
	3	1.118791	7.729703	0.106111	0.003975
	4	0.816335	7.379883	0.104236	0.005342

7 Conclusion and Future Work

This paper proposes a medical signal forecasting technique based on deep learning. LSTM was used for signal regression and consisted of five layers with one hidden layer to reduce the complexity of the prediction model. The number of hidden cells ranged from low to medium (range 5–100). This method is based on feeding a network of current and past samples to predict future samples. This technique provides a robust predicted signal with both five and one hundred cells, which means that LSTM can estimate high-frequency signals with peaks and hills, such as the EEG signal, which has a very low amplitude and high frequency. LSTM was compared with other types of neural networks such as CNN, GRU, Bi-LSTM, and NARX. LSTM was superior in most the performance indices calculated for all four channels used for testing. As mentioned previously, the training-to-testing ratio was 50:50, indicating that LSTM could predict the correct signal with a low-size training dataset. LSTM obtained a prediction RMSE of less than 16, whereas the GRU had the best RMSE of approximately 28. LSTM has an Mx.AE that is close to that of the GRU. To improve the results of LSTM, the number of hidden cells was increased to 100, and all the performance indices were enhanced accordingly. This enhancement ensures the high quality of the predicted signal. In Future work, this technique will be applied to a larger dataset with multiple subjects for more generalization and will be improved to become a single-network model, rather than a multiple-network model.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Duaa A. Kareem(First draft writing), Zaineb M. Alhakeem(Methodology) , Nawar Hayder Tawfeeq Alqrnawi(First Draft Writing), Batool Dahham Al-Ali(Editting) and Heba Hakim(Reviewing)

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability

The used data is published and cited in the manuscript.

Code Availability

The code is available after requesting from the corresponding author.

Acknowledgments

The authors thank their collage and Departments for their support to finish this work.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work the authors used Grammarly and Paperpal in order to edit and grammars check the language. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

[1] N. Mathur, A. Gupta, S. Jaswal, and R. Verma, "Deep learning helps EEG signals predict different stages of visual processing in the human brain," Biomed Signal Process Control, vol. 70, p. 102996, Sep. 2021, doi: 10.1016/j.bspc.2021.102996.

[2] K. J. Blinowska and M. Malinowski, "Non-linear and linear forecasting of the EEG time series," Biol Cybern, vol. 66, no. 2, pp. 159–165, Dec. 1991, doi: 10.1007/BF00243291.

[3] S. Kamil Gatfan, "A Review on Deep Learning For Electroencephalogram Signal Classification," Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 16, no. 1, pp. 137–151, Mar. 2024, doi: 10.29304/jqcsm.2024.16.11453.

[4] D. Sivan, "Big Data Analytics in Healthcare: A Survey," May 2015.

[5] H. Pankka, J. Lehtinen, R. J. Ilmoniemi, and T. Roine, "Enhanced EEG Forecasting: A Probabilistic Deep Learning Approach," Jan. 16, 2024. doi: 10.1101/2024.01.16.575836.

[6] A. Gharabaghi et al., "Coupling brain-machine interfaces with cortical stimulation for brain-state dependent stimulation: enhancing motor cortex excitability for neurorehabilitation," Front Hum Neurosci, vol. 8, Mar. 2014, doi: 10.3389/fnhum.2014.00122.

[7] M.-I. Stefanou et al., “Brain State-dependent Brain Stimulation with Real-time Electroencephalography-Triggered Transcranial Magnetic Stimulation,” Journal of Visualized Experiments, no. 150, Aug. 2019, doi: 10.3791/59711.

[8] C. Zrenner, D. Desideri, P. Belardinelli, and U. Ziemann, “Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex,” Brain Stimul, vol. 11, no. 2, pp. 374–389, Mar. 2018, doi: 10.1016/j.brs.2017.11.016.

[9] X. Du, K. Liang, Y. Lv, and S. Qiu, “Fast reconstruction of EEG signal compression sensing based on deep learning,” Sci Rep, vol. 14, no. 1, p. 5087, Mar. 2024, doi: 10.1038/s41598-024-55334-9.

[10] Thara D. K., Premasudha B. G., Murthy T. V., and S. A. C. Bukhari, “EEG Forecasting With Univariate and Multivariate Time Series Using Windowing and Baseline Method,” International Journal of E-Health and Medical Communications, vol. 13, no. 5, pp. 1–13, Dec. 2022, doi: 10.4018/IJEHMC.315731.

[11] A. Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network,” Physica D, vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[12] N. Mathur, A. Gupta, S. Jaswal, and R. Verma, “Deep learning helps EEG signals predict different stages of visual processing in the human brain,” Biomed Signal Process Control, vol. 70, p. 102996, Sep. 2021, doi: 10.1016/j.bspc.2021.102996.

[13] E. H. Houssein, A. Hammad, and A. A. Ali, “Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review,” Neural Comput Appl, vol. 34, no. 15, pp. 12527–12557, Aug. 2022, doi: 10.1007/s00521-022-07292-4.

[14] M. K. Titkanlou, D. T. Pham, and R. Mouček, “Classification of EEG Signal Using Deep Learning Architectures Based Motor-Imagery for an Upper-Limb Rehabilitation Exoskeleton,” SN Comput Sci, vol. 6, no. 3, p. 193, Feb. 2025, doi: 10.1007/s42979-025-03743-6.

[15] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, “A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals,” Comput Biol Med, vol. 99, pp. 24–37, Aug. 2018, doi: 10.1016/j.compbiomed.2018.05.019.

[16] B. H. Brinkmann et al., “Forecasting Seizures Using Intracranial EEG Measures and SVM in Naturally Occurring Canine Epilepsy,” PLoS One, vol. 10, no. 8, p. e0133900, Aug. 2015, doi: 10.1371/journal.pone.0133900.

[17] F. J. Muñoz-Almaraz, F. Zamora-Martínez, P. Botella-Rocamora, and J. Pardo, “Supervised filters for EEG signal in naturally occurring epilepsy forecasting,” PLoS One, vol. 12, no. 6, p. e0178808, Jun. 2017, doi: 10.1371/journal.pone.0178808.

[18] P. T. M. Nguyen, Y. Hayashi, M. D. S. Baptista, and T. Kondo, “Collective almost synchronization-based model to extract and predict features of EEG signals,” Sci Rep, vol. 10, no. 1, p. 16342, Oct. 2020, doi: 10.1038/s41598-020-73346-z.

[19] A. Shakeel, T. Tanaka, and K. Kitajo, “Time-Series Prediction of the Oscillatory Phase of EEG Signals Using the Least Mean Square Algorithm-Based AR Model,” Applied Sciences, vol. 10, no. 10, p. 3616, May 2020, doi: 10.3390/app10103616.

[20] Q. Liu et al., “Open access EEG dataset of repeated measurements from a single subject for microstate analysis,” Sci Data, vol. 11, no. 1, p. 379, Apr. 2024, doi: 10.1038/s41597-024-03241-z.

[21] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua, “Hierarchical Multimodal LSTM for Dense Visual-Semantic Embedding,” in 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, Oct. 2017, pp. 1899–1907. doi: 10.1109/ICCV.2017.208.

[22] A. Oellrich et al., “The digital revolution in phenotyping,” Brief Bioinform, vol. 17, no. 5, pp. 819–830, Sep. 2016, doi: 10.1093/bib/bbv083.

[23] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[24] G. Zhang, V. Davoodnia, A. Sepas-Moghaddam, Y. Zhang, and A. Etemad, “Classification of Hand Movements From EEG Using a Deep Attention-Based LSTM Network,” IEEE Sens J, vol. 20, no. 6, pp. 3113–3122, Mar. 2020, doi:10.1109/JSEN.2019.2956998.

Biographies (All Authors)

Duaa Abbas Kareem was born in Basrah, Iraq in 1990. She received the B.S. and M.S. degrees in Computer engineering from the University of Basrah, in 2011 and 2015 respectively. Her teaching interests covering wide areas of modules across the Polymers and Petrochemicals Engineering Department, College of Oil and Gas Engineering, Basrah University for Oil and Gas, includes Dynamics and Process Control. Her research interests include Biomedical Engineering, Digital signal processing, Control Systems.

Zaineb M. Alhakeem was born in Basrah, Iraq in 1983. She received the B.S. and M.S. degrees in Computer engineering from the University of Basrah, in 2005 and 2008 respectively. She is also had her Ph.D degree within the Electrical Engineering Department, College of Engineering in 2020, University of Basrah, Basrah, Iraq. Her teaching interests covering wide areas of modules across the Department of Chemical and Petroleum Refining Engineering, College of Oil and Gas Engineering, Basrah University for Oil and Gas, include Optimization, Dynamics and Process Control. Her research interests include Biomedical Engineering, Digital signal processing, Control Systems.

Nawar Hayder Tawfeeq Alqrnawi was born in Iraq Basrah, He received his B. S degree in telecommunication engineering from IRAQ UNIVERSITY COLLEGE in IRAQ -BASRA and his MSc in electrical electronics engineering from istanbul gelisim university in Turkey . He worked as a planner manager in YILDIZLAR Company . and as a Project manager at ALWAQIA company which specializes in Electricity protection systems. lastly, he works as a lecturer in oil and gas at Basra University, his current research interest is in renewable energy, systems, and network monitoring Internet of Things (IoT).

Batool Dahham Al-Ali was born in Basrah, Iraq. She received the B. A degree in Biology Science, University of Basrah 2010 and B. A degree in English Language, University of Basrah 2014. M.A. in Arts- Department of English in IUL in Lebanon 2019. Her teaching interests covering wide areas of modules across the Department of Oil and Gas Engineering, College of Oil and Gas Engineering, Basrah University for Oil and Gas, include Languages.

Heba Hakim was born in Basrah, Iraq in 1980. She received the B.S. and M.S. degrees in Computer Engineering Department from the University of Basrah, in 2002 and 2004 respectively. She is also had the Ph.D. degree in Electrical Engineering Department in 2020 from University of Basrah, Basrah, Iraq. Her research interests include Artificial Intelligence, Navigation System, Image Encryption and Steganography