International Journal of Mathematics and
Computer Science, **17**(2022), no. 3, 955–967

Hybrid Deep Neural network and Long Short term Memory Network for Predicting of Sunspot Time Series

Safwan O. Hasoon[1], **Muzahem M. AL-Hashimi**[2]

[1]Department of Software
Faculty of Computer Science and Mathematics
University of Mosul
Mosul, Iraq

[2]Department of Statistics and Informatics
Faculty of Computer Science and Mathematics
University of Mosul
Mosul, Iraq

email: Dr.safwan1971@uomosul.edu.iq,
muzahim_alhashime@uomosul.edu.iq

(Received December 30, 2021, Accepted February 4, 2022)

Abstract

Time series analysis and prediction are important statistical and technological topics in studying the behavior of different phenomena. The sun's climate, which follows an eleven-year cycle of activity, is a very important strange fact that must be predicted, as its effects will affect life on Earth and around. In this paper, we propose the use of deep neural networks with LSTM for the time-series prediction of sunspots and we compare the performance of the hybrid model with deep neural networks and RNN while keeping the experimental conditions as constant as possible. The results show that both the LSTM periodic neural network and the fully connected deep neural network can predict sunspots more accurately, which shows the advantages of the LSTM deep neural network for periodic sunspot prediction.

Key words and phrases: Sunspot time series, deep neural network, long and short-term memory network.
AMS (MOS) Subject Classifications:
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

1 Introduction

Nonlinear science research involves various fields of social and natural sciences such as finance, electric power, hydrology, and weather [1],[2]. In these dynamic systems, exhibiting sunspot phenomena, the sunspot time series produced by any variable evolving contains rich dynamics information of the system, which often contains important content. Sunspots are an interesting answer to studying the sun [3]. The study of sunspots is necessary to predict many related events such as coronal mass emissions, solar wind, and energetic particles driven by the solar magnetic field. Throughout science, living organisms, electronic technologies, and aeronautics, and astronautics are affected by solar activity [4]. One of the main indicators for estimating sunspot values is the sunspot number (SSN) used in many scientific fields [5]. The time series of sunspots are recorded at specific times, and building a model to predict future values to build future series is necessary for many technical and research fields [6]. Researchers have developed many predictive statistical and intelligence models to build models and simulate the expected chains. Intelligent neural networks proved to be more efficient than statistical methods in prediction tasks [5],[6]. RNN is one of the Neural Network algorithms used to build predictive models for time series. The long short-term memory network was first proposed by [7] And later developed into a Recurrent Neural Network (RNN). In this study, a model using LSTM in DNN architecture was developed, and the performance of the proposed model was compared with RNN and deep neural networks.

2 Deep learning

Deep learning is a technique that builds a neural network with multiple hidden layers that need a large amount of data to train the models to model complex data, find the internal structural characteristics of the data, and then discover the true relationship form between the variables [8]. Deep learning is a branch of machine learning, and it is also the fastest-growing direction in machine learning [9]. In addition to finding complex models, it can automatically extract features to solve complex nonlinear problems. In traditional shallow machine learning, feature extraction needs to be done by domain experts. How to transform the input signal to extract features usually requires a lot of domain knowledge, and the dependence on domain knowledge is very high. In deep learning, what is learned is that the hidden nodes of the network automatically extract and transform features from the

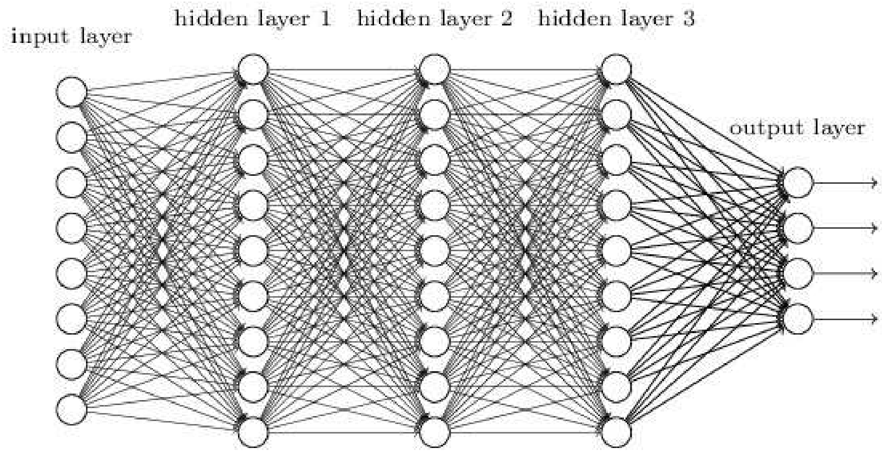

Figure 1: Fully connected with three hidden layer.

input signal [10]. These features will become more and more abstract layer by layer, thereby establishing a high-dimensional mapping relationship to solve a complex non-complex Linear problem [11].

2.1 Fully connected deep neural network

In a conventional, fully connected deep neural network (DNN), from the input layer to the hidden layer and then to the output layer, the neurons between the layers are connected, but the neurons in the layer are not connected [12]. As shown in Figure 1, this paper constructs a fully connected deep neural network model with three hidden layers. The network structure often determines the complexity of the algorithm and the stimulability of the model. The number of hidden layers and the number of nodes is an important content in the selection of deep learning hyperparameters [13][14],[15]. The greater the number of nodes and the number of layers, the more adjustable the model and the finer the granularity of the prediction results, but it also means the higher the computational complexity [13],[16]. Generally, the primary approach for calculating the number of hidden layers and nodes is to choose a structure that is as compact and as feasible while still achieving the accuracy standards; that is, to choose as few hidden layers and nodes as possible [17]. In this paper, the above parameters are finally selected in many experiments.

2.2 Long-term and short-term memory network (LSTM)

Long-term and short-term memory networks are widely employed in the field of deep learning and have become a frequently used neural network model in today's rapid growth of neural networks [18]. Compared with basic neural networks, the characteristics of long-term and short-term memory networks for processing time series are not only in the layer. A right connection is established with the layer, and a right link is established between the same layer [18],[19]. A type of cyclic neural network model is the long-term and short-term memory network model [20]. The cyclic neural network is a recurrent neural network based on a sequence that recurs in the sequence's evolution direction and connects all nodes in a chain. This kind of chain link the characteristics reveal the close relationship between the sequences [21]. The long-term and short-term memory network solves the problem that the gradient disappears when the recurrent neural network processes long-term data. As a result, the long-term and short-term memory networks are well suited to processing time-series events with relatively lengthy intervals and delays, such as speech recognition, machine translation, and time series prediction [22].

As illustrated in Figure 2, the working unit in the long and short-term memory network receives the current input information, the hidden state, and the unit state from the previous instant. The persistence and suppression of information are accomplished by three gates (input gate, forget gate, and output gate) [23],[24],[25].

The forget gate determines how much information is "forgotten" about the hidden state at the previous moment through the activation function. The input gate consists of two parts:
(1) Choose what value to be retained through the function, (2) Generate candidate vector values through the activation function.
The product of the two is taken as a part of the state quantity and the sum of the product of the state quantity generated in the forget gate and the previous time as the current unit state. Finally, the output gate generates the candidate vector, selects the retained information through the function, and transmits the result as the current hidden state to the next unit and the unit at the same time in the upper layer [24],[26].

As shown in Figure 2, we construct an LSTM network prediction model (taking the mud content prediction as an example) that includes one LSTM cyclic layer and two fully connected layers (dense layers). The number of nodes in the two fully connected layers is 20 and 10, respectively. On the

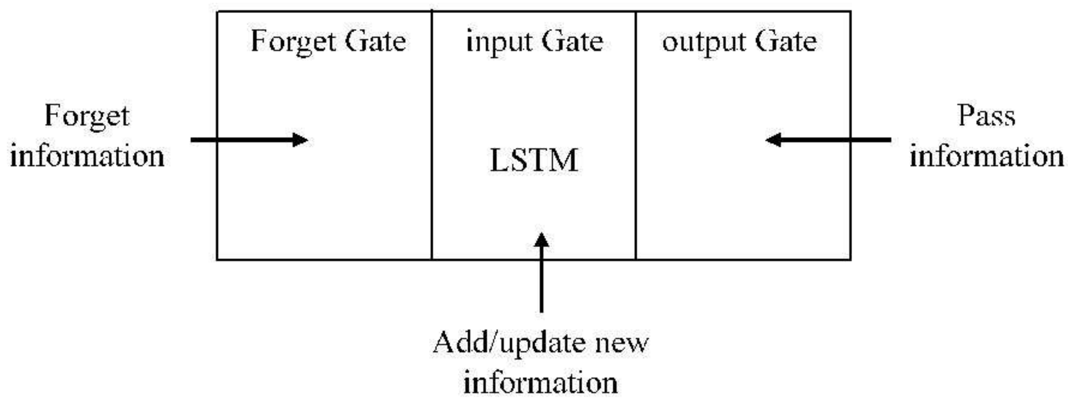

Figure 2: Parts of an LSTM cell.

one hand, it is consistent with the last two hidden layers of the previously designed fully connected deep neural network model for easy comparison. On the other hand, it is because of the complexity of the actual data. The degree (respectively five-dimensional features and four-dimensional features) is not particularly complicated and it is more reasonable to choose this number of nodes to improve the calculation efficiency. An LSTM loop layer considers that the complexity of the actual data is not high. The structure of the loop body itself Deeper, an LSTM layer corresponding to the first three layers of the fully connected deep neural network will also have a certain contrast. To prevent over-fitting, the optimization process uses the Adam algorithm and the Dropout regularization approach. The LSTM cyclic neural network's prediction framework is similar to that of a fully connected deep neural network, except that the LSTM cyclic structure replaces part of the hidden layer. The Dropout regularization approach, which may be utilized in the loop body of LSTM or the final fully linked layer, can be applied in various ways in the LSTM network training.

3 Experiments

The time series of sunspots is one of the longest time series that has been collected and preserved so far. It is a series from 1749 for daily, monthly, and yearly in CSV format. Available on the Sunspot Index and Long-term

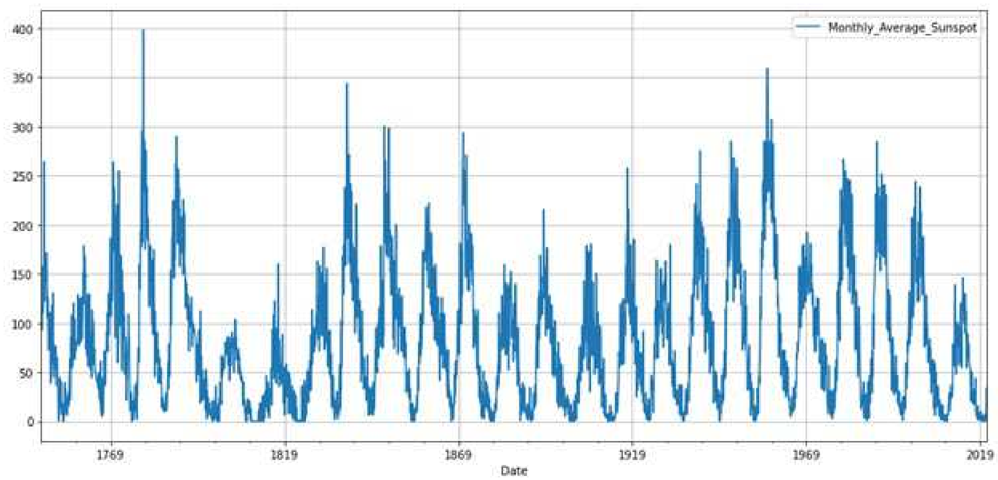

Figure 3: Sunspot cycles.

Solar Observations [26]. This period is sufficient to study the features and characteristics of this cosmic phenomenon and predict the number of sunspots for 3,266 months. Figure 1 shows the average monthly number of sunspots as it is noticed from the series of sunspots that they are stable in the middle and have a non-stable variance that constantly changes over time. In other words, there is a pattern of small oscillations and a pattern of large oscillations. It is also noted that there is a periodic nature of the series and it appears that it is a regular periodic behavior and the absence of a clear linear trend in the series. This analysis shows that the sunspot series has a strong cycle for 129 months represented by the reference line that shows the position of the main peak in the spectral density.

3.1 Network training

The data set is divided into (3000) for training and (265) for testing and building three prediction models (RNN, DNN, DNN-LSTM), As shown in Figure 5, the standardized training data is input into the designed fully connected deep neural network (DNN) and long-short-term memory network (LSTM) algorithms to train the network model. In the process of neural network training, the loss function is usually used to measure the ability of the network model to fit the training data. The neural network algorithm adjusts the model parameters (that is, the weight value of the network) through multiple training iterations. Finally, through the decline of the loss function

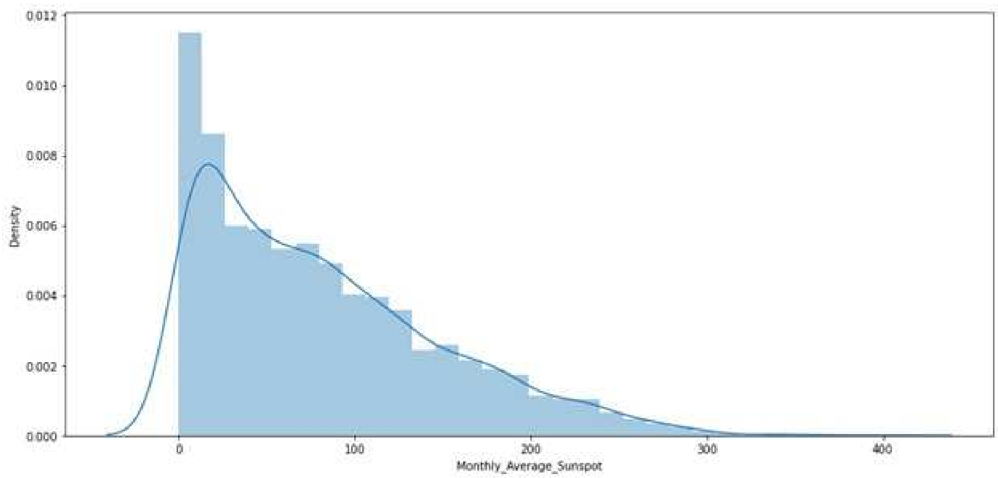

Figure 4: Strong of sunspot series cycle.

Trend to judge whether the algorithm has achieved the ideal deep neural network model. Figure 5 shows the training process of the fully connected deep neural network model for predicting sunspots. It can be seen from the figure that the network model has reached convergence when the iteration reaches 100 times, and the total time is 2 min 44 s. Then the trained network model is applied to the test data (after standardized processing) to verify its performance.

3.2 Metrics for Evaluation

Mean Squared Error (MSE) and Mean Absolute Error (MAE) are generally used to measure the deviation between the observed value and the true value. Therefore, we use the root mean square error to evaluate and compare the prediction results of the LSTM network and the DNN network. The calculation formula is:

$$MSE = \frac{1}{n}\sum(y - \widehat{y})^2 \text{ ....(1)}$$

$$MAE = \frac{1}{n}\sum|y - \widehat{y}| \text{ ....(2)}$$

where:
n = number of items.
y = original or observed y-value.
$y\widehat{=}y$-value from regression.

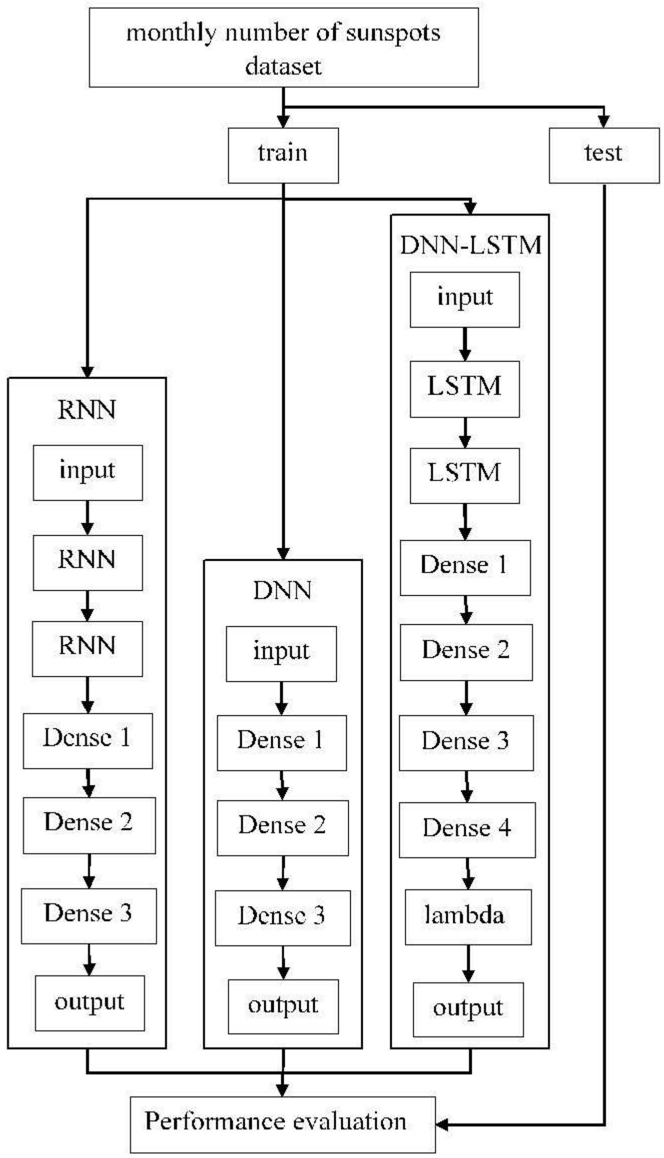

Figure 5: Stacked DNN-LSTM Architecture.

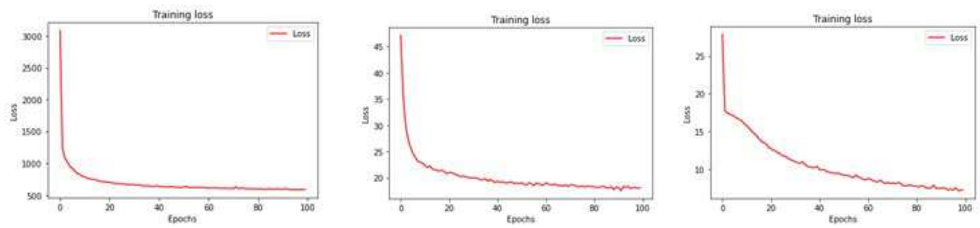

Figure 6: Loss function decline curve: (a) DNN-LSTM (b) DNN (c) RNN

4 Experimental results and analysis

Three models have been created using the PyTorch framework to develop models (DNN, RNN, DNN-LTSM). The design of the DNN uses three network layers, and RNN uses five network layers and uses Adam as the optimizer and DNN-LTSM, as shown in figure 6. MAE as the loss function for all models and Visdom tool to detect the convergence of the loss function in real-time and the models are trained for a total of 100 cycles.

Table 1: Comparison of the performance of prediction models using (MAE, MSE)

Model	MAE	MSE
DNN	16.832048	642.3686
RNN	15.332574	477.70837
DNN-LTSM	13.747599	413.61658

Figures 7,8,9 represent the comparison between the real-time series and the predicted value based on the model used.

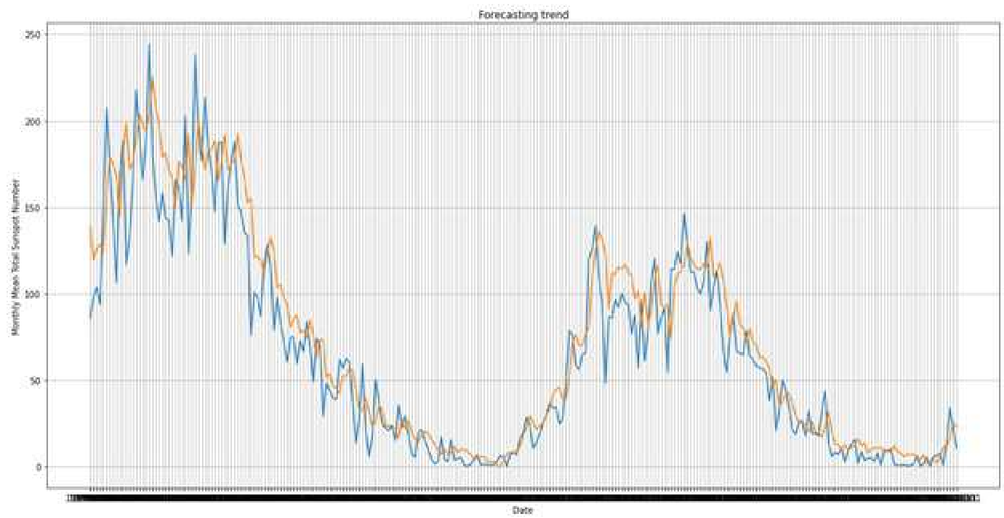

Figure 7: Test dataset of the DNN-LSTM model applied.

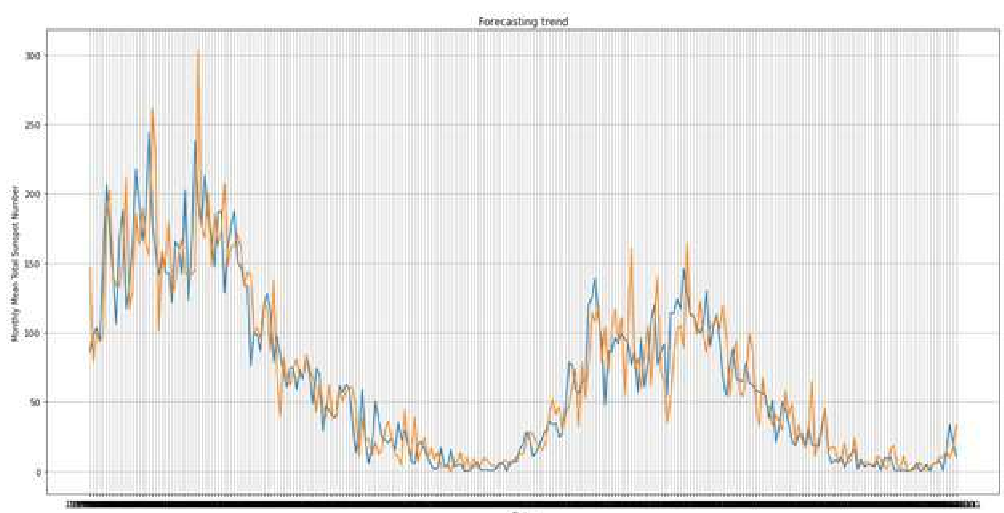

Figure 8: Test dataset of the DNN model applied.

5 Conclusion

In this paper we proposed a prediction of sunspots using the DNN-LSTM hybrid neural network. Unlike the traditional (DNN and RNN) algorithms that only deals with local information, ours use the information of the previous frame and works on the resolution of the current frame. At the same time, based on the DNN-LSTM architecture, we used a cost function based on the monthly sunspot rate to train the network. The results showed that the algorithm based on the hybrid structure has higher detection performance than the traditional (DNN and RNN) algorithms. Moreover, we studied the effect of the cost function based on context information and the traditional cross-entropy cost function on the performance of the proposed method. Our results showed that the cost function and the error ratio are more suitable for the DNN-LSTM structure.

References

[1] Elena Malyutina, Vladimir Shiryaev, Time Series Forecasting Using Nonlinear Dynamic Methods and Identification of Deterministic Chaos, Procedia Computer Science, **31,** (2014), 1022–1031. 10.1016/j.procs.2014.05.355.

[2] Panagiotis Papaioannou, Ronen Talmon, Serafino Di, Ioannis Kevrekidis, Constantinos Siettos, Time Series Forecasting Using Manifold Learning, (2021).

[3] Samir Hamouda, Sunspots Production and Relation to Other Phenomena: A Review, (2020).

[4] Jie Zhang et al., Earth-affecting solar transients: a review of progress in solar cycle, Progress in Earth and Planetary Science, **24,** (2021), 8. 10.1186/s40645-021-00426-7.

[5] Mohammad Nazari-Sharabian, Moses Karakouzian, Relationship between Sunspot Numbers and Mean Annual Precipitation: Application of Cross-Wavelet Transform-A Case Study, (2020), J. 3. 10.3390/j3010007.

[6] H. Abdel-Rahman, Beshir Marzouk, Statistical method to predict the sunspots number, NRIAG Journal of Astronomy and Geophysics, (2018), 7. 10.1016/j.nrjag.2018.08.001.

[7] Sepp Hochreiter, Jürgen Schmidhuber, Long Short-term Memory, Neural computation, (1997), 9. 1735-80. 10.1162/neco.1997.9.8.1735.

[8] Rajesh Singh, Anita Gehlot, Mahesh Prajapat, Bhupendra Singh, Deep Learning, (2021), 10.1201/9781003245759-5.

[9] Albert Liu, Oscar Law, Deep Learning, (2021), 10.1002/9781119810483.ch2.

[10] Miroslav Kubat, Deep Learning, (2021), 10.1007/978-3-030-81935-4_16.

[11] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Deep Learning, (2021), 10.1007/978-1-0716-1418-1_10.

[12] Michael Kohler, Sophie Langer, Approximation properties of fully connected deep neural networks, Journal of Multivariate Analysis, (2019), 10.1016/j.jmva.2019.06.005.

[13] Michael Kohler, Sophie Langer, On the rate of convergence of fully connected deep neural network regression estimates, The Annals of Statistics, (2021), 49. 10.1214/20-AOS2034.

[14] Jinkyu Kim, Juyeob Kim, Byungjo Kim, Miyoung Lee, Joohyun Lee, Hardware design exploration of fully-connected deep neural network with binary parameters, (2016), 305–306, 10.1109/ISOCC.2016.7799799.

[15] Jiarui Liu, Ruofei Hu, Hongyu Hu, A Hybrid Classification of Driver's Style and Skill Using Fully-Connected Deep Neural Networks, (2020), 10.4271/2020-01-5107.

[16] Yize Tang, Shenshen Gu, An Algorithm of Dynamically-Connected Deep Neural Network, (2019), 141–144. 10.1109/IHMSC.2019.00040.

[17] Helmut Bölcskei, Philipp Grohs, Gitta Kutyniok, Philipp Petersen, Optimal Approximation with Sparsely Connected Deep Neural Networks, SIAM Journal on Mathematics of Data Science, (2017), 1. 10.1137/18M118709X.

[18] Fattun Fattah et al., A Comparative Analysis Between LSTM and LSTM Ensembled Strategy for Short-term Load Forecasting, (2021), 10.36227/techrxiv.16807213.

[19] Joos Korstanje, LSTM RNNs, (2021), 10.1007/978-1-4842-7150-6_18.

[20] Ximin Wang et al., (2021). LSTM-Exploit: Intelligent Penetration Based on LSTM Tool, 10.1007/978-3-030-78615-1_8.

[21] Jiwoo Choi, Hyungjin Yoo, Sangil Choi, Taewon Kang, Identification of Gait Cycle using LSTM, The Journal of Korean Institute of Information Technology, **1,** (2021), 27–39, 10.14801/jkiit.2021.19.8.27.

[22] Grósz, Tamás, Mikko Kurimo, LSTM-XL: Attention Enhanced Long-Term Memory for LSTM Cells, (2021), 10.1007/978-3-030-83527-9_32.

[23] Emil Hvitfeldt, Julia Silge, Long short-term memory networks, 10.1201/9781003093459-14.

[24] Federico Succetti, Andrea Ceschini, Francesco Luzio, Antonello Rosato, Massimo Panella, Time Series Prediction with Autoencoding LSTM Networks, 10.1007/978-3-030-85099-9_25.

[25] Li Zhang, Rebecca Strachan, Prediction Using LSTM Networks, (2019), 1–8, 10.1109/IJCNN.2019.8852206.

[26] "SILSO World Data Center for the production, preservation and dissemination of the international sunspot number", Wwwbis.sidc.be, (2021) [Online]. Available: https://wwwbis.sidc.be/silso/. [Accessed: October 31, 2021].