Identification of Real-Time Maglev Plant using Long-Short Term Memory Network based Deep Learning Technique

Amit Kumar Sahoo[1]*, Rudra Narayan Pandey[2], Sudhansu Kumar Mishra[2] and Prajna Parimita Dash[3]

[1]Department of EEE, Centurion University of Technology and Management, Odisha, India

[2]Department of EEE, [3]Department of ECE, Birla Institute of Technology, Mesra, Ranchi, India

Deep neural network has emerged as one of the most effective networks for modeling of highly non-linear complex real-time systems. The long-short term memory network (LSTM) which is a one of the variants of recurrent neural network (RNN) has been proposed for the identification of a highly nonlinear Maglev plant. The comparative analysis of its performance is carried out with the functional link artificial neural network- least mean square (FLANN-LMS), FLANN-particle swarm optimization (FLANN-PSO), FLANN-teaching learning based optimization (FLANN-TLBO) and FLANN-black widow optimization (FLANN-BWO) algorithm. The proposed LSTM model is a feed forward neural network trained by a simple iterative method called the ADAM algorithm. The obtained results indicate that the proposed network has better performance than the other competitive networks in terms of the MSE, CPU time and convergence rate. To validate the dominance of the proposed network, a statistical tests, i.e. the Friedman test, is also applied.

Keywords: FLANN, Maglev system, Mean Square Error, Recurrent Neural Network, System identification

Introduction

The identification of any system is found to be an essential requirement in different applications in the area of control systems, power systems, communication engineering etc. Precise and fast identification of a highly non-linear complex system is a challenging task. To achieve this, variants of artificial neural network (ANN) have emerged as one of the effective adaptive approaches. The ANN has been efficaciously implemented in the field of classification, prediction, estimation and modeling due to its universal approximation and automatic learning capability.[1] The training of ANNs is carried out by using either derivative or derivative free based algorithms. The derivative based decent gradient algorithms, such as LMS, RLS, BP etc. may be trapped in local minima or face over-fitting problems.[2] To overcome such bottlenecks, derivative free evolutionary and swarm intelligence algorithms, such as GA, PSO, TLBO etc. are utilized to update the weights of neural networks.[3–6] This hybridization of neural networks with the artificial intelligence technique got an edge over others for the identification of different complex non-linear systems.

The ANNs perform well in the modeling of non-linear system, but, still have some shortcomings due to the absence of operational information at the hidden layers and nodes. In recent years, a deep learning technique has been implemented by many researchers to avoid the above shortcomings, and to obtain improved results.[7] The long-short term memory network (LSTM) network is considered as the most important recurrent neural network (RNN) based deep learning technique.[8] In this technique, the output of the neural network is feedback to the input, which follows the data sequence or chain of information. The LSTM uses different gated structures to estimate each state to avoid the above listed issues, and it is widely used in different fields like, identification, prediction and speech recognition.

Subudhi *et al.*[9] used a memetic algorithm i.e. a differential evolution (DE) based approach to avoid the shortcomings of the gradient based training approach of the multilayer perceptron (MLP) network for non-linear system identification. Adaptive robust identification of complex non-linear dynamic plants has been proposed by Majhi *et al.*[4] In this model, the robust identification is carried out by applying a low complexity single layer functional link artificial neural network (FLANN) trained by a swarm

*Author for Correspondence
E-mail: amitkumar@cutm.ac.in

intelligence based particle swarm optimization (PSO) technique. Satapathy *et al.*[10] have proposed an improved teaching-learning based optimization (ITLBO) to solve global numerical optimization problem. From simulation results, it is observed that the ITLBO shows better performance than the traditional TLBO on various benchmark problems in terms of accuracy, convergence, speed, stability and robustness. Kumar *et al.*[11] have introduced a metaheuristic based socio evolution & learning optimization algorithm (SELO) inspired by the social learning behavior of humans. The SELO is implemented to solve 50 benchmark problems, and its performance is compared with the other competitive algorithms. Hayyolalam *et al.*[12] have proposed a novel black widow optimization algorithm (BWO) in 2020, which is inspired by mating behavior of black widow spiders. The efficacy of the BWO algorithm is determined by taking 51 different benchmark functions. Ogunmolu *et al.*[13] have introduced deep dynamic neural networks for the system identification of non-linear systems. Here, the authors investigate the effectiveness of a deep neural network for the identification of a non-linear system with complex behavior. The overall contribution of this paper is as follows:

 (i) An LSTM network has been proposed for the identification of the highly non-linear Maglev plant. The proposed network is a combination of LSTM and the feed-forward neural network trained by a simple iterative method called the ADAM algorithm to identify the Maglev plant.[14]

 (ii) A FLANN model is also implemented for the identification of the Maglev plant, and the weights of the FLANN model have been updated by using different nature inspired optimization techniques, i.e. PSO, TLBO and recently proposed BWO.

 (iii) A comparative analysis has been carried out by considering the MSE, CPU time and the convergence rate, and a non-parametric Friedman test is performed to obtain the performance of the above network.

 The organization of the paper is as follows: in section 2, the principle and structure of the Maglev plant is explained. The proposed LSTM model for the identification of the Maglev Plant is illustrated in section 3 and in section 4, the simulation study and performance analysis is discussed. Finally, the concluding remarks are given in section 5, and the scope of the future work is outlined.

The Maglev Plant

The Maglev system has extensive real world application in various fields, like frictionless bearings and high-speed trains due to its low noise and low friction characteristics.[15] The laboratory setup of the Maglev plant are shown Fig. 1. The Maglev plant available in our laboratory has an interface of two blocks, i.e. a physical Maglev plant and a computer interfaced with the Maglev plant. The Maglev device contains an electromagnet, power amplifier, photoelectric sensors, transmitter and receiver set, and control elements. The experimental set-up of the Maglev system is capable of controlling the steel ball to move up and down with respect to an equilibrium position. The purpose of controlling the steel ball is to trace the ball in a desired trajectory, irrespective of the disturbances.

When the current flows through the winding, then a magnetic force generated, which helps to attract the steel ball upward, and it balances the ball in an equilibrium position with respect to the gravitational force. The position of the Maglev ball is controlled by the electromagnetic force represented by 'F'. The laboratory setup used for research is manufactured by Feedback Instrument Ltd., and it is operated in a MATLAB environment.

The maglev plant parameters are shown in Table 1, and its transfer function is as follows.

Fig. 1 — Maglev laboratory setup

Table 1 — Maglev plant parameters

Name of the Parameter	Symbol	Value
Mass of steel ball	m	0.02 kg
Control voltage to current gain (Constant)	k_1	1.05 A/V
Sensor gain (Constant), offset	k_2, η	143.48 V/m, −2.8 V
Input voltage to the Controller	V_i	±5 V
Output voltage of sensor	V_o	+1.25 V to−3.75 V
Equilibrium position of steel ball	x_0	0.009 m
Current at equilibrium position	i_0	0.8 A
Gravitational constant	g	9.81 m/s^2

$$G_p(s) = \frac{\Delta V_O}{\Delta V_i} = \frac{-3518.85}{s^2 - 2180} \qquad \ldots (1)$$

From the above Eq. (1), it is clearly seen that the Maglev plant is a non-linear and complex system. Therefore, it is difficult to identify the Maglev plant parameters, which can be successfully controlled.

Proposed RNN based LSTM Model for Identification of the Maglev Plant

In this paper, a deep learning approach for MAGLEV plant identification has been proposed. In the last few years, many researchers have employed deep learning for system identification. Deep learning eliminates the drawbacks of many other machine learning based approaches because of its inherent advanced features. It uses a large number of hidden layers, and thereby feature extraction and feature matching is carried out in one go, that leads to minimal human intervention.

Nowadays, deep learning is used in almost every field like computer vision, speech recognition, forecasting, system identification etc. The training of a deep learning based network may be supervised learning or unsupervised one. In this research work, the job of system identification of the Maglev plant is done by using a supervised learning based Deep learning approach because of the availability of input and output data. The basic recurrent neural networks (RNNs) were mostly implemented for system identification a few years back.[13] As it is a memory based network, it memorizes/remembers the provided sequential input data, which have been utilized in successive iterations. The LSTM which is a type of RNN has a nature of remembering information for a long period of time as its default behavior. The information is entered in sequential order in the LSTM network through three step gate processes.[16] The cell state of LSTM network is controlled and protected by three gates, i.e. the forget gate, input gate, and output gate, as shown in Fig. 2.[17]
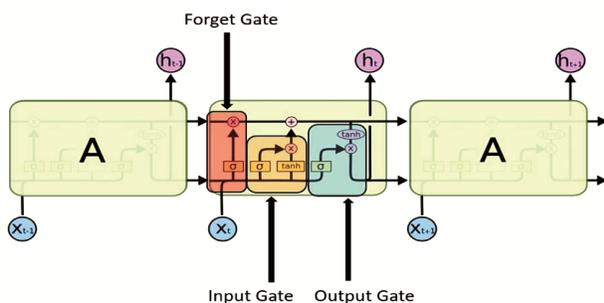

Fig. 2 — LSTM with three-step gate process

Forget Gate

This gate mainly decides, which information is important to store or omit from the past. Mathematically, the *sigmoid function* is held responsible for the above process, i.e., 0 to omit and 1 to keep the data.

$$y_o = \sigma(W_f.[f_{t-1}, x_i]) \qquad \ldots (2)$$

where, x_i is the input, f_{t-1} indicates the output, σ is the Sigmoid function, y_o is the input to the cell state C_{t-1} and W_f is the weight.

Update Gate/Input Gate

Here, the gate decides which parts of the unit should be added to the current state. It is mathematically expressed as:

$$i_t = \sigma(W_i.[f_{t-1}, x_i]) \qquad \ldots (3)$$

$$\tilde{C}_t = tanh(W_c.[f_{t-1}, x_i]) \qquad \ldots (4)$$

where, C_t is the current cell state, i_t is the inner state and *tanh* function level the all values ranging from-1 to 1.

Output Gate

This gate decides on the current cell to be available for the output, then

$$S_t = \sigma(W_o.[f_{t-1}, x_i]) \qquad \ldots (5)$$

$$f_t = S_t * tanh(C_t) \qquad \ldots (6)$$

where, S_t is the output of Sigmoid.

Simulation Study and Performance Analysis

The algorithms were executed in the A*cerAspire*V system, Window 10 OS, Intel® Core™ i5-3337U CPU @ 1.80GHz processor, RAM of 8 GB and in the MATLAB platform. The proposed and other competitive approaches are implemented for the identification of the Maglev plant in real time. The whole available input and output data are divided into two parts, i.e. the training data set and testing data set in a ratio of 8:2. Then the training set is utilized for training the LSTM model, and the performance of the model is validated using the testing set. The LSTM model parameters are given in Table 2.

Here, the ADAM optimization algorithm is implemented to train the proposed network for 200 epochs. To restrict the range of gradients from going out of bounds, the gradient threshold and learning rate is considered to be 1 and 0.005 initially,

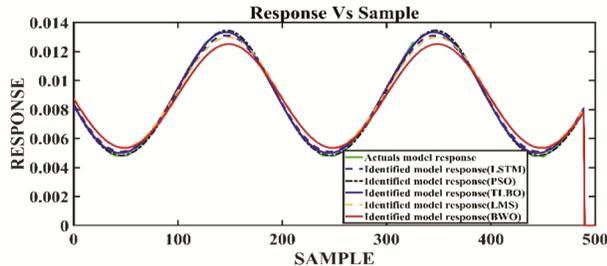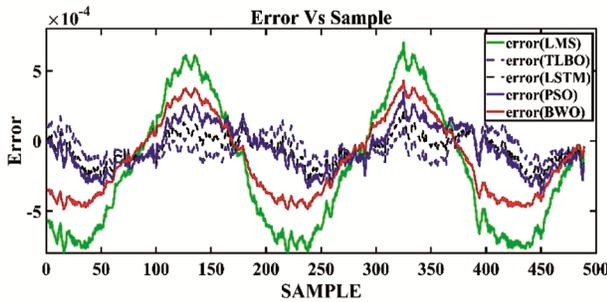| Table 2 — The parameters of LSTM model					
No. of Features	No. of Responses	No. of Hidden Layer	No. of Epoch	Starting Learning Rate	Optimization Algorithm
1	1	200	200	0.005	ADAM

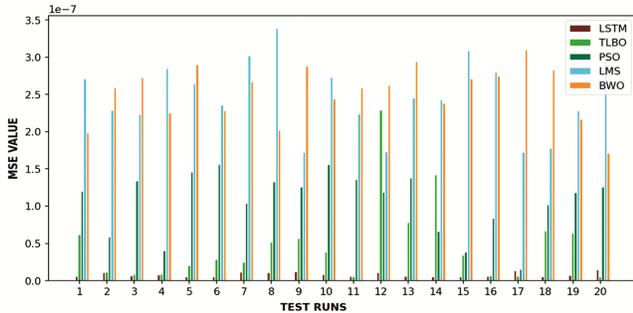Fig. 3 — Comparative identified model response

Fig. 4 — Comparative Error plot

Fig. 5 — Comparative plot of MSE in various test runs

Table 3 — Performance analysis table			
Model	No. of iteration	MSE	CPU Time(in sec)
LSTM	200	9.5293E-09	1269.72
TLBO	30	2.7498E-08	462.02
PSO	20	1.3945E-08	782.43
BWO	100	2.28E-07	382.422
LMS	10	2.47E-07	4.15

Table 4 — Friedman test rank table				
Methods	LSTM	TLBO	PSO	BWO	LMS
Mean Ranks	4.6	8	11.6	17.8	18

Table 5 — Friedman test parameters				
Source	Sum of Square (SS)	Degree of Freedom (DOF)	Mean Square (MS)	Chi-Square	Critical value (p)
Columns	175.7	4	43.925	70.28	1.98098E-14
Error	24.3	76	0.3197		
Total	200	99			

The MSE error and computational time taken by each model for training is given in Table 3, and from the table, it can be concluded that the LSTM model outperforms others, as the MSE error of LSTM is significantly lower than the other state of the art techniques under consideration.

To study the supremacy and repeatability of the obtained response of the network a non-parametric Friedman test is performed by using the MATLAB. The average rank of the different networks used for identification is given in Table 4, which signifies that the lower rank network has higher accuracy and performance. The Friedman test parameter is given in Table 5 and the critical value is obtained as 1.98098E-14 from the Friedman test. A null hypothesis concept comes, if the critical value is less than the significance level, i.e. $\alpha = 0.05$ and it can be rejected.

Conclusions

In this article, the LSTM network is analyzed, and suitably implemented for the identification of a non-linear Maglev plant. From the simulation results, it is observed that the LSTM network exhibits superior performance in terms of MSE and convergence rate as compared to the other competitive models, such as FLANN-LMS, FLANN-PSO, FLANN-TLBO and FLANN-BWO model. The proposed approach overcomes the local minima and

and then the learning rate is reduced after 125 epochs by multiplying it with a factor of 0.2.

The fitting curves and the error plot of all the models are shown in Figs 3 and 4. It is observed that the LSTM, provides the improved performance as compared to other competitive approaches for the identification of the real time Maglev plant. It is confirmed that the LSTM network shows the best convergence rate among all the models. From Table 3, it is evident that the LSTM network has taken only 200 iterations of training time to obtain the best result compared to the other models. From the above comparison, it is evident that, the performance of the proposed model is better than that of other state-of-the-art neural network structures. The comparative MSE plot is given for various test runs in Fig. 5.

over-fitting problem of multilayer networks. The effectiveness of the proposed method is assessed for the identification of a real-time Maglev plant. A non-parametric statistical test, i.e., the Friedman test confirmed the dominance of the proposed LSTM network over the reported methods. In future, the weights of the proposed networks may be optimized by using more efficient and recent approaches, to enrich their robustness and efficacy.

Competing Interest

The authors declare that they have no competing interests.

Reference
1   Narendra K S & Parthasarathy K, Identification and control of dynamical systems using neural networks, *IEEE Trans Neural Networks*, **1** (1990) 4–26.
2   Majhi L, Borah M & Roy P, A new method for power system load modeling using a nonlinear system identification estimator, *IEEE Int Conf Adv Commun Control Comput Technol*, **61** (2016) 1–6.
3   Dhanalakshmi R, Vadivel A, Chandrasekaran R M & Parthiban P, Nature inspired heuristic approach to balance the resource in the multiple distribution centers, *J Sci Ind Res*, **74** (2015) 77–81.
4   Majhi B & Panda G, Robust identification of nonlinear complex systems using low complexity ANN and particle swarm optimization technique, *Expert Syst Appl*, **38** (2011) 321–333.
5   Rao R V, Savsani V J & Balic J, Teaching-learning-based optimization algorithm for unconstrained and constrained real-parameter optimization problems, *Eng Optim*, **44** (2012) 1447–1462.
6   Niranjan T & Parthiban P, Modelling and analysing an integrated multi channel food supply chain distribution of an indian dairy firm using modified TLBO algorithm, *J Sci Ind Res*, **78** (2019) 31–34.
7   Mallikarjuna B R, Viswanathan & Naib B B, Feedback-based gait identification using deep neural network classification, *J Crit Rev*, **7** (2020) 661–667.
8   Wan H, Guo S, Yin K, Liang X & Lin Y, CTS-LSTM: LSTM-based neural networks for correlatedtime series prediction, *Knowl-Based Syst*, (**in press**)(2020).
9   Subudhi B & Jena D, Nonlinear system identification using memetic differential evolution trained neural networks, *Neurocomputing*, **74** (2011) 1696–1709.
10   Satapathy S & Naik A, Improved teaching learning based optimization for global function optimization, *Decis Sci Lett*, **2** (2013) 23–34.
11   Kumar M, Kulkarni A J & Satapathy S C, Socio evolution & learning optimization algorithm: A socio-inspired optimization methodology, *Future Gener Comput Syst*, **81** (2018) 252–272.
12   Hayyolalam V & Kazem A A P, Black widow optimization algorithm: A novel meta-heuristic approach for solving engineering optimization problems, *Eng Appl Artif Intel*, **87** (2020) 103249.
13   Ogunmolu O, Gu X, Jiang S & Gans N, Nonlinear systems identification using deep dynamic neural networks, 2016.
14   Sakinah N, Tahir M, Badriyah T & Syarif I, LSTM with adam optimization-powered high accuracy preeclampsia classification, *Int Electron Symp*, (2019) 314–319.
15   Sain D, Swain S K & Mishra S K, Real time implementation of optimized i-pd controller for the magnetic levitation system using jaya algorithm, *IFAC-Papers On Line*, **51** (2018) 106–111.
16   Gonzalez J & Yu W, Non-linear system modeling using LSTM neural networks, *IFAC-Papers On Line*, **51** (2018) 485–489.
17   www.medium.com (11 June 2020)