湖南大学学报（自然科学版）
Journal of Hunan University (Natural Sciences)

第 50 卷 第 12 期
2023 年 12 月

Available online at http://jonuns.com/index.php/journal/index

Vol. 50 No. 12
December 2023

Open Access Article

https://doi.org/10.55463/issn.1674-2974.50.12.9

Prediction of Time to Failure (TTF) of Power Systems Using a Deep Learning Technique

Muhammad Zohaib Sohail[1]*, Taimoor Zafar[1], Talha Ahmed Khan[2], Muhammad Asim[3], Sadique Ahmad[3], Tariq Mairaj[4], Mohammed A. ELAffendi[3]

[1] Department of Electrical Engineering, School of Engineering and Applied Sciences, Bahria University, Karachi, Pakistan

[2] Department of Computer Science, School of Engineering and Applied Sciences, Bahria University, Karachi, Pakistan

[3] Data Science and Blockchain Laboratory, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia

[4] Department of Electronics and Power Engineering, Pakistan Navy Engineering College, National University of Sciences and Technology, Karachi, Pakistan

* Corresponding author: mzohaib.bukc@bahria.edu.pk

Received: September 16, 2023 / Revised: October 11, 2023 / Accepted: November 9, 2023 / Published: December 29, 2023

Abstract: Power distribution systems consist of various components, including transformers, generators, bus bars, switch gear, utility electric poles, and power distribution cables. A high reliability of a power distribution system is extremely required for the economic growth of any country. The reliable system ensures continuous power supply to the metropolis and industrial sector. Sudden power failures in residential areas and industries will affect the economy of any country. Similarly, unwanted inspection and maintenance/replacement activities will also increase the cost. The research aimed to forecast the time to failure (TTF) of a power distribution system to ensure condition-based predictive maintenance. This will help the power sector to plan maintenance activities when they are highly required. In this study, a prognostic framework to estimate TTF based on long short-term memory (LSTM) is proposed. Prior information on the TTF will ensure continuous power supply to the metropolis and industrial sector, resulting in improved economic growth of the country. To address the issues of sudden power failure, the authors proposed a prognostic framework based on deep learning Long Short-Term Memory (LSTM) approach to forecast the TTF of power grid stations for the first time. The data sets of TTF of the power grid due to equipment failure and environmental failures are used for training the LSTM algorithm. In this approach, different layers and optimizers of the LSTM scheme are used, including input, output, and hidden layers, and adaptive moment estimation (ADAM) and stochastic gradient descent moment (SGDM) optimizers that control the accuracy of the prediction step. In addition, LSTM can handle long-term dependencies, which is unlikely for other recurrent techniques. The results obtained through the LSTM approach using ADAM are found to be quite satisfactory in terms of the root mean square error (RMSE) between the actual values and predicted (computed) values compared to the SGDM. Accurate results show that the proposed prognostic framework will help power companies to plan maintenance and replacement activities well in time before the time of failure to ensure continuous power supply to the country.

Keywords: power systems, reliability, root mean square error.

使用深度学习技术预测电力系统的故障时间(TTF)

摘要：配电系统由各种组件组成，包括变压器、发电机、母线、开关设备、公用电线杆

和配电电缆。任何国家的经济增长都极其需要配电系统的高可靠性。可靠的系统确保了大都市和工业部门的持续供电。居民区和工业的突然停电都会影响任何国家的经济。同样，不必要的检查和维护/更换活动也会增加成本。该研究旨在预测配电系统的故障时间(TTF)，以确保基于状态的预测性维护。这将有助于电力部门在非常需要时规划维护活动。在这项研究中，提出了一种基于长短期记忆（长短期记忆网络）来估计TTF的预后框架。TTF的事先信息将确保大都市和工业部门的持续供电，从而改善该国的经济增长。针对突发停电问题，作者首次提出了基于深度学习长短期记忆（长短期记忆网络）方法的预测框架来预测电网站点的TTF。采用因设备故障和环境故障导致的电网TTF数据集来训练长短期记忆网络算法。在这种方法中，使用了长短期记忆网络方案的不同层和优化器，包括输入层、输出层和隐藏层，以及控制预测步骤精度的自适应矩估计(亚当)和随机梯度下降矩(SGDM)优化器。此外，长短期记忆网络可以处理长期依赖性，这对于其他循环技术来说是不可能的。与SGDM相比，通过使用亚当的长短期记忆网络方法获得的结果在实际值和预测（计算）值之间的均方根误差(均方根误差)方面非常令人满意。准确的结果表明，所提出的预测框架将帮助电力公司在故障发生之前及时计划维护和更换活动，以确保国家的持续供电。

关键词： 电力系统、可靠性、均方根误差.

1. Introduction

The power system consists of different components such as transformers, generators, bus bars, and switch gears. We often come across failures in power systems resulting in breakdowns that cause huge loss of production for industries and other related sectors. Breakdowns normally occur because of malfunctioning or failures in components installed in a power system [1].

Power is essential for all sectors and individuals alike; without it, life is wretched. Power system failures can be caused by various circumstances, including wear-out failures, malfunctions in various pieces of equipment or system components, and environmental impacts [2]. Strong and reliable power system components ensure continuous power supply for all sectors in a country.

Causes of power failures can be minimized by different preventive measures for the prediction of failures in power systems, and timely repair action can be performed to overcome the system's failure and restore it to working mode. To minimize failures and their frequent restoration, there is a concept of RAM parameters called reliability, availability, and maintainability [1].

Power failures are increasing day by day. The main reason behind failures of a power system is that the fault arises in the installed components of a power system. To overcome the failures of power systems, the behavior of each component, i.e. its TTF (Time to Failure) is a significant feature for reliability improvement. Predicting time to failure, i.e., future states of a power system, is a challenging task to improve the reliability of the system.

Two approaches for load curtailment and reduction in failure cost were developed in [1]. The first one is the global steady technique comprising the development of a model to cater to the faulty transmission power system. This model helps to reduce the time and cost resulting in unexpected failures. The other technique is the dynamic iterative technique, which tests all the components of the power system and highlights the components having frequent failures. The whole procedure continues iteration until the load curtailment may be minimized or even equal to zero. In these approaches, the combination of two methods, i.e, fuzzy and Monte Carlo simulations were implemented to track the randomness and behavior of different components of a power system.

Reliability and maintainability analyses for strudel production lines for different machine workstations and entire line levels were proposed in [1] and [2]. Statistics about the failure and repair data were analyzed and the significant index parameters were examined. Additionally, different modes of failure and hazard rates for the whole system were calculated in [2]. In [3], descriptive statistical analysis of failure and repair data for food production lines were provided. A comparative study of the performance and evaluation between the production lines is conducted by using different statistical techniques. Failure and repair data of various production lines were collected, and suitable distribution fitting is applied using the goodness of fit test through Minitab software [3]. Comparison of probability plots and hazard rate plots of different production lines are examined and analyzed.

In [4], a literature review was made on reliability, availability, maintainability engineering, and related parameters. All of them were divided into different groups, namely, the assessment of RAM parameters, the methodologies required to obtain various RAM parameters, design, support, life cycle cost, etc., and simulation and modeling for reliability, availability, and maintainability. In [4], an additional parameter called Supportability has also been introduced, which is also beneficial during the study of RAM parameters. In supportability, there is a concept of information technology (IT), which is effective and a key factor for implementation in different systems and other related sectors.

According to [5], long short-term memory (LSTM) is an updated and modified type of traditional recurrent neural network that has been implemented on the available data of reliability, i.e., time to failures, to compute and predict the reliability of the system for future time instants where data is not available. The proposed LSTM technique was applied to an online service-oriented system composed of several individual components. The authors conducted several experiments and tests to analyze the overall performance and efficiency of the proposed approach and compared it with related techniques for validation.

The deep learning-based LSTM scheme was widely used in various industries and applications [6]-[8] because of LSTM's high accuracy and learning capabilities. If the database has all of the system properties, it can produce highly effective outcomes. The effectiveness of LSTM is also dependent on the most recent information on the health of the system.

In [9], the authors implemented the LSTM technique on different medical time series data. Various optimizers are used by the authors to compare the efficiency of algorithms during training. The problem of exploding and vanishing gradients during prediction has also been discussed, and various techniques based on classification accuracy measurement have been proposed to overcome it [9].

In this study, for the first time, a deep learning-based prognostic framework is offered to anticipate the time to failure (TTF) of a power system using the historical record of power system components. The prognostic framework is based on the LSTM algorithm under the implementation of two optimizers, adaptive moment estimation (ADAM) and stochastic gradient descent moment (SGDM) optimizer. There are various optimizers among them, but only ADAM and SDGM optimizers are selected because they identify the decay rate for the gradient and for squared gradient moving averages [6]-[8].

The datasets used in this study are given in reference [10]. It is also reported in ref. [10] that power grid outages are mostly caused by equipment failures or by environmental circumstances. The main objectives of this study are as follows:

- To propose a high-performance prognostic framework based on a deep learning scheme to forecast future failure rates of power systems.
- To evaluate the performance of the proposed framework using two different optimizers based on the LSTM algorithm.

This paper is divided into five sections. The first portion includes an introduction with context, motivation, significance of power systems and their components, a literature review (i.e, past work on reliability improvement), a problem description, and the study objective. The second section comprises a case study and a description of the dataset. The methodology portion of the third section includes the research methodology, the proposed algorithm, and its implementation. The fourth section includes the results, discussion, computation, and analysis of the obtained results, and the fifth section includes the conclusion, future work, and recommendations.

2. Case Study

The occurrence of failures in power grids has become a common factor in power outages that must be overcome. There are multiple causes of power failures such as equipment faults, environmental effects, lightning, operating faults, and damage in transmission and distribution lines. Power failures result in huge loss of production (for industrial consumers) and other related sectors if they persist for a longer duration. On analyzing the dataset [10], the frequent causes of power grid failure are equipment and environmental faults. Therefore, to overcome this issue or minimize power failures, prediction-based approaches have been proposed in this study to predict the time to failure of power grids.

2.1. Dataset Description

The dataset comprises failures in power grids installed in different regions of Australia [10]. Outages (failure) data consists of multiple parameters (features) such as failure region, failure's start date and time, average outage (shutdown/repair) time in minutes (TTR), event ID, number of customers affected, and reason or root cause of power failure in the power grid.

Since dataset failures in the power grids are mostly caused by equipment failures or environmental conditions, whereas failures due to other elements such as operating faults and lightning are rare. Quarterly failure data of six years from 2013 to 2018 are available in the dataset.

As per the given source, there were some constraints during the compilation of the dataset. The data contained information about a failure duration of more than 5 minutes and at least 50 or more customers were affected by failures. Out of many local government regions (areas) of Australia, the

Canterbury region is the focus of the analysis throughout this study. Table 1 represents the sample dataset starting from 2013 and having different features.

Table 1 Australian Power Grid Dataset sample (Developed by the authors based on [10])

S #	LGA	Start Date	TTF (hrs.)	Start Time	TBF (hrs.)	Customers Interrupted	TTR (min)	Reason
1	Canterbury	5/1/2013	5	3:19 PM	0	1000	32	Equipment fault
2	Canterbury	8/1/2013	8	10:44 PM	79	99	105	Environmental
3	Canterbury	13/01/2013	13	5:05 PM	114	100	160	Equipment fault
4	Canterbury	16/01/2013	16	3:11 PM	70	188	219	Third party
5	Canterbury	18/01/2013	18	8:49 PM	53	66	356	Environmental
6	Canterbury	20/01/2013	20	11:34 AM	38	172	79	Equipment fault
7	Canterbury	25/01/2013	25	2:30 PM	122	64	75	Environmental
8	Canterbury	26/01/2013	26	6:03 PM	27	63	56	Equipment fault
9	Canterbury	19/02/2013	50	12:13 PM	570	80	87	Environmental
10	Canterbury	1/3/2013	60	10:25 AM	238	70	65	Environmental
11	Canterbury	17/03/2013	77	8:14 AM	381	85	55	Environmental
12	Canterbury	28/03/2013	88	5:16 PM	273	65	43	Third party
13	Canterbury	5/4/2013	95	5:14 AM	179	71	106	Equipment fault
14	Canterbury	5/4/2013	95.5	9:12 PM	15	79	93	Equipment fault
15	Canterbury	15/04/2013	105	8:56 AM	227	1128	85	Equipment fault
16	Canterbury	27/04/2013	117	9:55 AM	228	2102	44	Equipment fault
17	Canterbury	1/5/2013	121	1:24 PM	99	6326	25	Environmental
18	Canterbury	2/5/2013	122	9:38 PM	20	2088	126	Equipment fault
19	Canterbury	31/05/2013	151	12:10 AM	686	61	72	Environmental
20	Canterbury	31/05/2013	151.5	10:38 AM	10	62	297	Equipment fault

3. Methodology

Out of the various strategies covered in the literature review, this study is the first to use a long short-term memory (LSTM) algorithm to estimate the time to failure of a power system based on component failure information.

In LSTM, which is a specific type of recurrent neural network, it can handle long-term dependencies problems as well. There are multiple layers and gates in the single-cell memory of LSTM, and each of them is interconnected. Initially the input $x\tau$ and hidden state $h\tau{-}1$ are fed to the sigmoid function for the forget gate layer. Afterwards, the input and modulation gate layer which computes $i\tau$ and g_τ respectively through sigmoid and tangent activation functions. The output layer $o\tau$ computes the final output $h\tau$ after multiplication of cell state c_τ with output layer.

3.1. Algorithm

RNNs are the type of neural networks having a feed-forward network usually used for predicting time series data. As recurrent neural networks have a deficiency that could not model data or series having long-term dependencies. Therefore, to overcome this problem another deep learning technique called long short-term memory (LSTM) could be applied to model data having long-term dependencies in series [5].

In LSTM, there are multiple small units (memory cells) in hidden layers. Normally, in any type of neural network, a minimum of three types of layers are observed, namely an input layer, a hidden layer, and an output layer, which will form a network. In contrast, LSTM has different layers and gates such as forget gate layer, input gate layer, output layer, cell state, input modulation gate, and output modulation gate [11]-[12].

These gates are activated by different trigonometric functions such as the sigmoid function, which results in values ranging from 0 to 1, while the hyperbolic tanh function returns values ranging from -1 to +1. The LSTM network comprises multiple cells interconnected with each other while taking the previous output of the cell, which becomes the recurrent input for the next cell, and the current input value will be fed. The weights for different layers are learned during the training process of the algorithm, which also depends upon multiple parameters such as the choice of optimizer, number of hidden layers, learning rate, and epochs. There must be a reasonable selection for these hyperparameters, because if parameters are not properly selected, then it results in computation complexity, training time of algorithm may also increase and error in the resultant values.

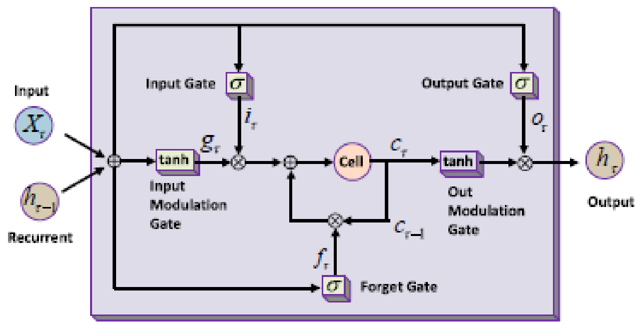

Fig. 1 Unit cell memory of the LSTM [11]

In Figure 1, the basic memory cell block or unit of the LSTM network is depicted.

3.2. Implementation of the LSTM Algorithm

Initially, the historical Time to Failure (TTF) data will be taken and then standardized to perform pre-processing steps for better performance. Then, the selection of different layers of the network will be defined, i.e., the input layer, hidden layers (which would be responsible for computation of weights learned during the training process), and the output layer for the computation of results. Now, there are various parameters such as epochs, optimizer, learning rate period, scheduler, and many more.

Afterwards, training of the LSTM model will begin with the help of different gates and layers of the LSTM model. The weights are learned during the training process; therefore, the computation complexity of the LSTM training phase is greater than that of the incoming testing phase. The reason is that in the testing (prediction) phase of the LSTM model, the already learned weights will be used to predict the future states of TTF.

The root mean square (RMSE) will also be computed for comparison of different dataset combinations. Figure 2 describes the block diagram of the entire algorithm.

A forget gate is in the LSTM initially, as shown in Figure 3. The purpose of the forget gate layer is to segregate information, i.e., identify whether to store or keep the information or neglect (forget) it. Because there is a sigmoid (σ) activation function that results in the values in the range from 0 to 1. Values near 0 indicate that the value of the cell state will be ignored, while 1 indicates that information in the cell state will be kept (hold). The mathematical equation for the forget gate layer is as follows:

$$f_{\tau=}\ \sigma\,(W_{xf}\,x_\tau\ +\ W_{hf}\,h_{\tau-1}\ + b_f\,) \qquad (1)$$

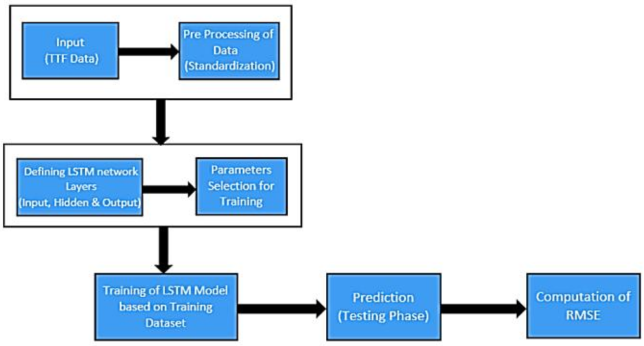

Fig. 2 Block diagram for the implementation of the algorithm (Developed by the authors)

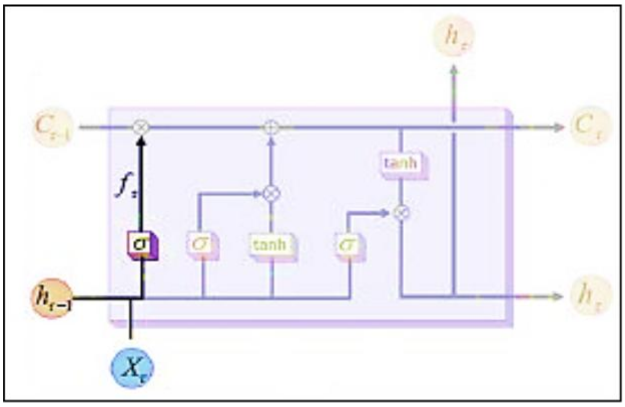

Fig. 3 Forget layer for LSTM [13]

In the next stage, the purpose is to compute a new contribution or addition to the cell state. Now the input $x\tau$ and hidden state $h\tau{-}1$ are fed to the sigmoid function in the in previous step. Subsequently, using the tangential (tanh) function creates a vector of new values (information) to the cell state, as shown in Figure 4. Then, the multiplication of two incoming gates, i.e., $i\tau$ and g_τ will be performed, and these will be added to the cell state through the addition operation. The mathematical equation is given.

$$i_{\tau=}\ \sigma\,(W_{xi}\,x_\tau\ +\ W_{hi}\,h_{\tau-1}\ + b_i\,) \qquad (2)$$

$$g_{\tau=}\ tanh\,(W_{xg}\,x_\tau\ +\ W_{hg}\,h_{\tau-1}\ + b_g\,) \qquad (3)$$

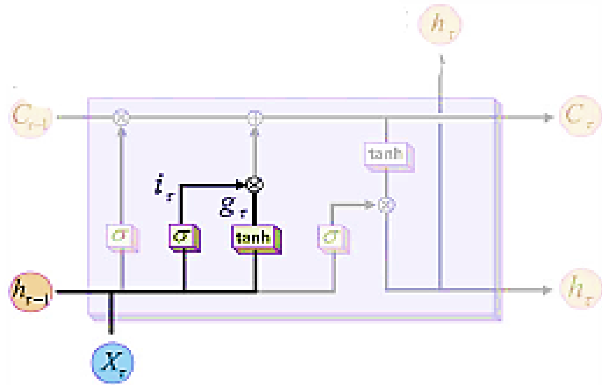

Fig. 4 Input and modulation gate layer for the LSTM [13]

In the next step, as shown in Figure 5, the top line indicating the LSTM unit is the cell state. Therefore, update the previous value of cell state $c_{\tau-1}$ to the new state, i.e, c_τ or in other terms, we interpolate new cell values. The output of the forget gate is multiplied by the previous cell state and then added with two incoming quantities i_τ and g_τ since these are the new candidate values. The cell state is important as far as the output of LSTM is concerned. It can be expressed as follows:

$$c_{\tau=}\ f_\tau\ \otimes\ c_{\tau-1}\ + i_\tau\ \otimes\ g_\tau \qquad (4)$$

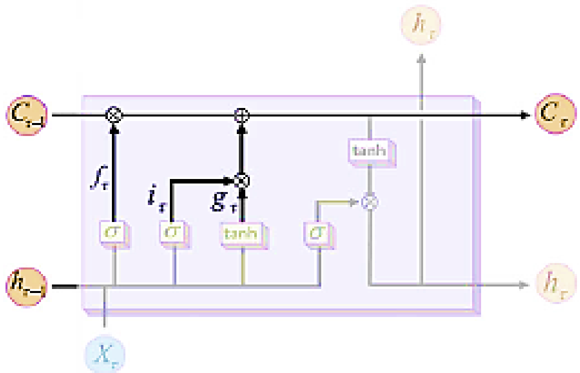

Fig. 5 Cell state for the LSTM [13]

In the last step of the output, the input and the previous hidden state will gain and feed to the sigmoid function to obtain o_τ as shown in Figure 6. The updated cell state will now pass through the tanh activation function (which gives the values in the range from -1 to 1). This result will be multiplied by the previous obtained output from o_τ, and then the final result, i.e, h_τ will be achieved. The mathematical expression of output gate h_τ and output layer o_τ is,

$$o_{\tau} = \sigma\,(W_{xo}\,x_\tau + W_{ho}\,h_{\tau-1} + b_o\,) \quad (5)$$

$$h_{\tau} = o_\tau \otimes tanh\,(c_\tau\,), \quad (6)$$

where o_τ is the output layer, $h_{\tau-1}$ is the previous output modulation gate, and c_τ is the cell state.

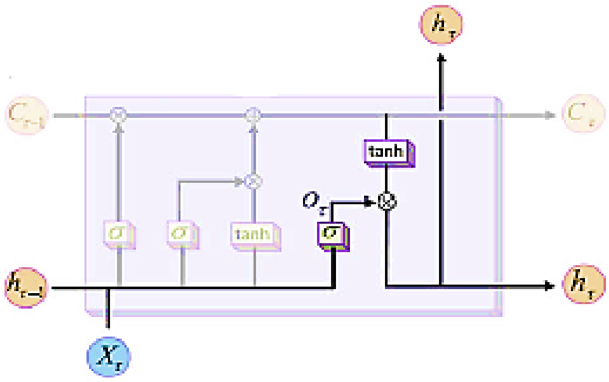

Fig. 6 Output layer of the LSTM [13]

4. Results and Discussion

4.1. Prediction of TTF using the LSTM Technique

Time-to-failure data as the time sequence values for LSTM input to model for the training of algorithm. The dataset in the case study provides information about the time to failure for power grids having 280 failure samples. There are two phases, one is training and the other is testing (prediction). Therefore, initially, 80% of the samples have been used to train the network, while the remaining 20% (each of the same number of samples will be used for prediction/testing purposes. To make the LSTM model, the input size (size of sequence for input layer) is considered to be 1 since there is one temporal TTF feature. In addition, the output layer (number of fully connected layers) is also considered to be 1. Finally, the main and paramount hidden layer (which corresponds to number of neurons in Neural Networks) is taken to 200. For standardization of data, the mean and standard deviation are computed for the train data for normalization.

There are some additional training parameters selected for better results during the training process and prediction. The first parameter is the selection of the optimizer for training. There are different options for the optimizer, i.e., ADAM, SGDM, and RMSPROP, but ADAM (adaptive moment estimation) and SGDM (stochastic gradient descent momentum) options are adopted from them because they identify the decay rate for the gradient and for squared gradient moving averages. The second parameter is the number of epochs (time instants for training). Then, a few more parameters are the gradient threshold. Initial learning rate, learn rate schedule, learn rate drop period, learn rate drop factor, and verbosity. The values selected for training parameters are listed in the following table.

Table 2 Parameters for the training process of all power failures (Developed by the authors)

S. No.	Parameter	Value
1	Optimizer	ADAM
2	Max epochs	200
3	Gradient Threshold	1
4	Initial Learn Rate	0.003
5	Learn Rate Schedule	Piecewise
6	Learn Rate Drop Period	125
7	Learn Rate Drop Factor	0.2
8	Verbose	0

Table 3 Parameters for the training process of equipment failure (Developed by the authors)

S. No.	Parameter	Value
1	Optimizer	SGDM
2	Max epochs	200
3	Gradient Threshold	1
4	Initial Learn Rate	0.004
5	Learn Rate Schedule	Piecewise
6	Learn Rate Drop Period	125
7	Learn Rate Drop Factor	0.2
8	Verbose	0

Afterwards, for the training of model train network command is applied having parameters such as train data, test data, number of layers, and training options as discussed earlier. To predict future time steps, the first model will be trained by using the training dataset. After training, the prediction of future time steps for test data will be computed by using the previously trained network. To predict the values of the number of time steps in fore-coming future states, a function called Predict and Update State will be applied to forecast the time step one at a time and then update the network state for each prediction.

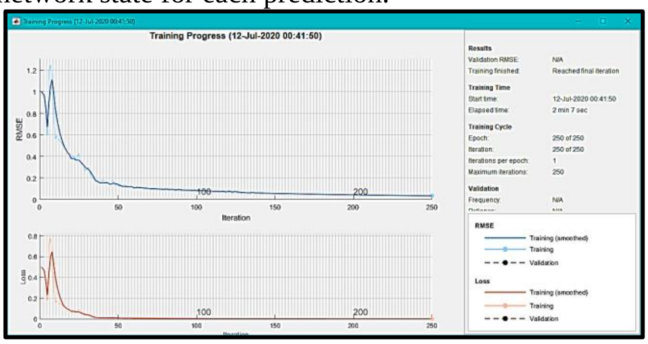

Fig. 7 ADAM optimizer training progress (Developed by the authors)

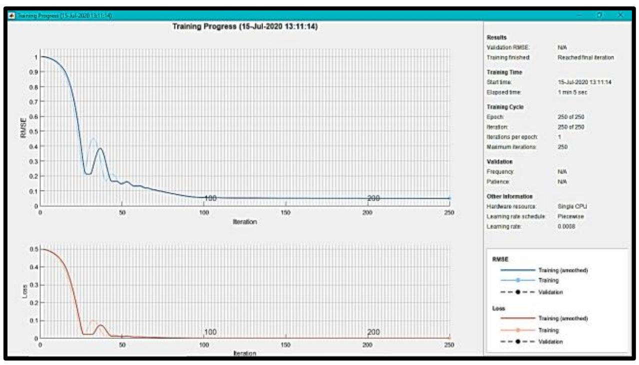

Fig. 8 SGDM optimizer training progress (Developed by the authors)

Therefore, for every prediction of TTF values, usage of the previous prediction as an input value to the function is applied. Figure 9 shows the plot for the training time series with the predicted (forecast) values.

To calculate the accuracy of prediction, the root mean square error (RMSE) between the test and the predicted data must be computed. The purpose of the RMSE is to find the difference (error) between the actual test value and the predicted value. First, there is a need to unstandardized the data before finding the RMSE. The mathematical expression for the root means square error is as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} [(Predicted)_i - (Actual)_i]^2}{N}} \tag{7}$$

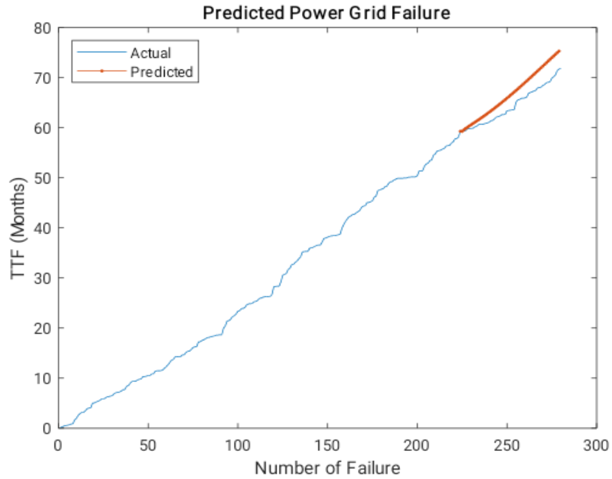

Fig. 9 Actual and predicted TTF values without update (80% training) (Developed by the authors)

Figure 9 describes the plots of the predicted and actual test values of time to failure of the power grid, while in the subplots of Figure 10, the calculated root mean square error (RMSE) and its plot between the predicted and the observed value of time to failure are shown because this is only a prediction, i.e., having no update phenomena, which results in a high value of RMSE between the actual and predicted TTF.

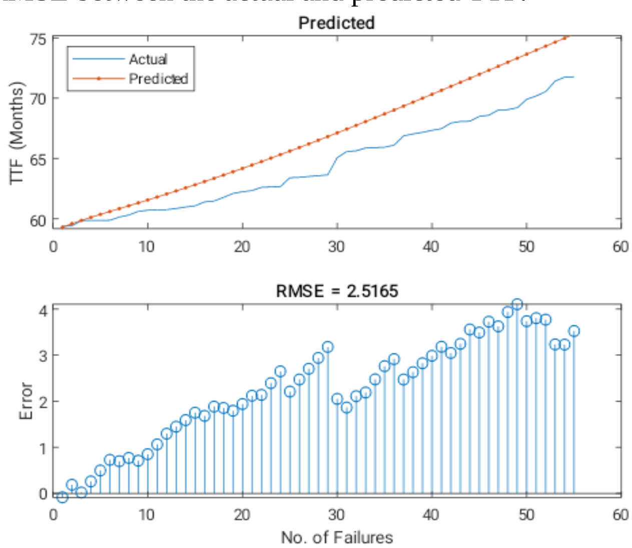

Fig. 10 Predicted TTF values without updating with RMSE (Developed by the authors)

For the update of the prediction model, we need to consider observed values instead of predicted values obtained before. First, the network is reset to prevent the previously obtained predictions from altering the predictions for the new data. It can be observed in Figure 11 that after reset and updating the network with the observed values, the prediction becomes better, which results in the desired output. We can now obtain the root mean square error between the observed and newly predicted values using equation (7). Since it can be observed that the observed values are close to the new predicted values, the RMSE decreases, which shows less error between them and comparatively better results, as shown in Figure 11.

4.1.1. *Prediction of the TTF of Power Grid Failures*

Now, as discussed in the dataset, 280 time to failure values are present. Therefore, 80% of values among them were considered for training the model, i.e., 225 values. The remaining 55 steps were then predicted through the trained model (by the weights which were learned during the training process). Figure 11 shows actual TTF values and predicted and updated TTF values. Reasonable results in terms of low RMSE were obtained.

Figure 11 comprises 55 predicted values (testing period) compared with the actual TTF values. Also, RMSE has been calculated between the actual and predicted values. The RMSE value is found to be 0.337, which means that the predicted value is fine.

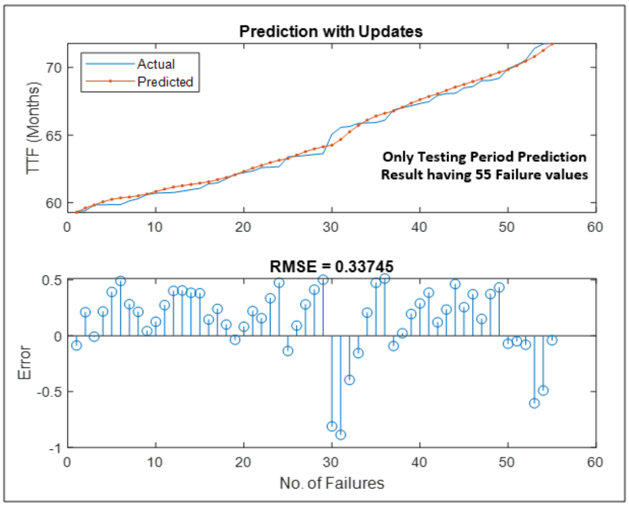

Fig. 11 Predicted TTF values with update having RMSE (80% training) (Developed by the authors)

We are more interested in analyzing the proposed framework while gap between the training data and estimated data to evaluate the strengths of the proposed framework. Therefore, to achieve this, the percentage of the training database gradually decreases from 80% to 60%. In the next stage, 70% (195 values) were used for training the algorithm, as shown in Figure 12.

4.1.2. Prediction of the TTF of Equipment Failures

As discussed earlier, the prediction of time to failure is computed using all component failure databases, including environmental, equipment faults, and lightning. Now, we are more interested in the TTF due to equipment failure. Therefore, we processed the datasets manually to create a new database using only equipment failure data reported in [10].

Currently, 160 equipment failure values have been found in the datasets in ref. [10]. Equipment failures are the most frequent failure causes in power systems; therefore, the prediction of failures has been computed in upcoming results. Again, first, 80% equipment failure data (i.e., around 128 values) were considered for training the LSTM model and then the remaining values were predicted. Thus, 32 values were computed (predicted) for power grid equipment failures.

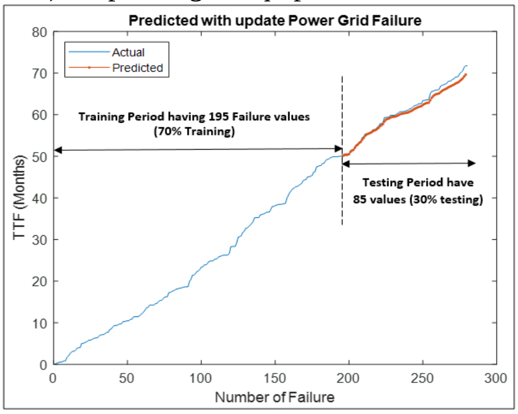

Fig. 12 Actual and predicted TTF values with update (70% training) (Developed by the authors)

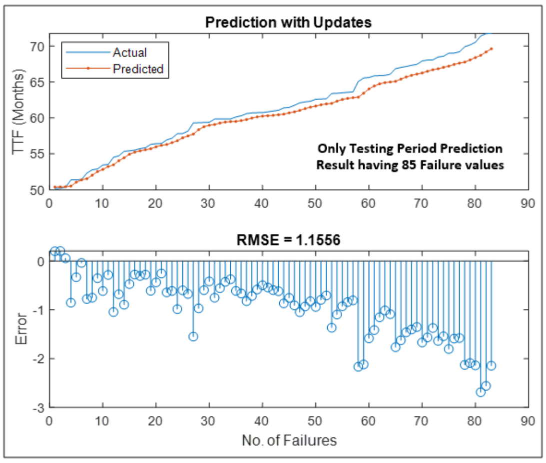

Fig. 13 Predicted TTF values with update having RMSE (70% training) (Developed by the authors)

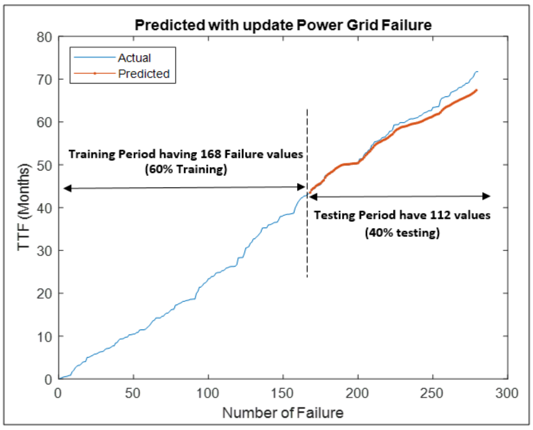

Fig. 14 Actual and predicted TTF values with update (60% training) (Developed by the authors)

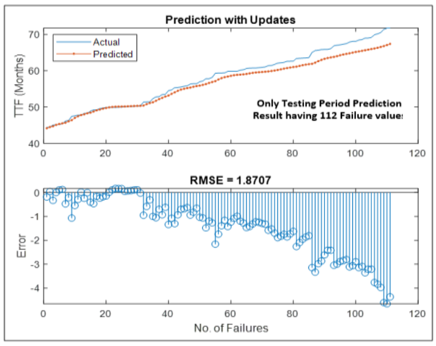

Fig. 15 Predicted TTF values with update having RMSE (60% training) (Developed by the authors)

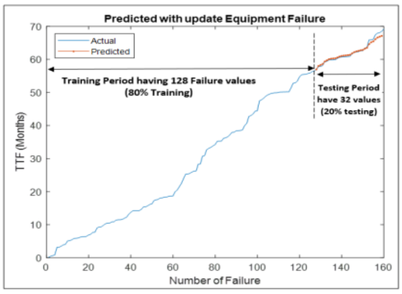

Fig. 16 Actual and predicted equipment failure values (80% training) (Developed by the authors)

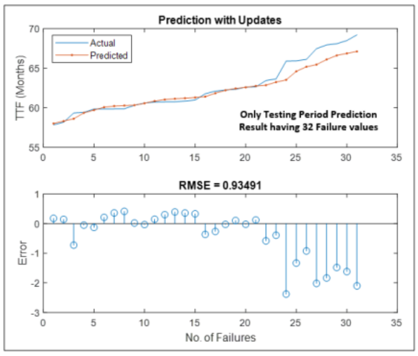

Fig. 17 Predicted equipment failure values with RMSE (80% training) (Developed by the authors)

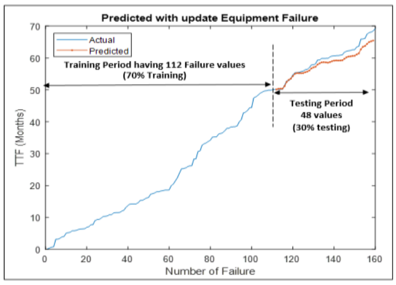

Fig. 18 Actual and predicted equipment failure values (70% training) (Developed by the authors)

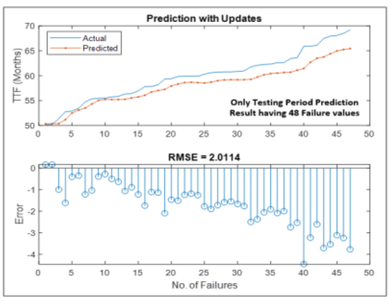

Fig. 19 Actual equipment failure values with RMSE (70% training) (Developed by the authors)

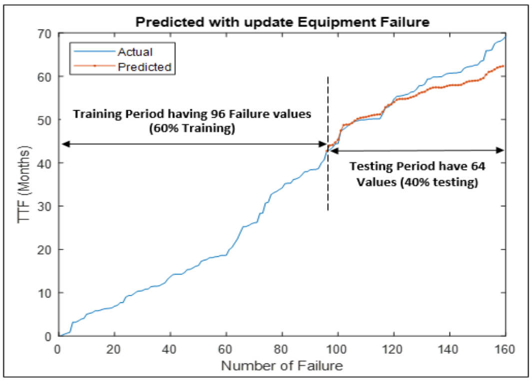

Fig. 20 Actual and predicted equipment failure values (60% training) (Developed by the authors)

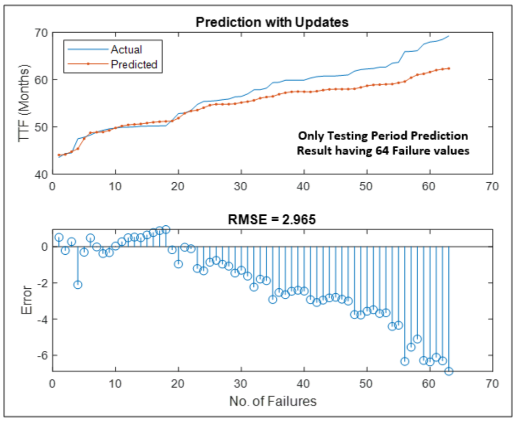

Fig. 21 Actual equipment failure values with RMSE (60% training) (Developed by the authors)

4.2. Prediction of TTF using the LSTM Technique

4.2.1. Using the ADAM Optimizer

Considering ADAM optimized in hyperparameters for training the LSTM model, the initial updating of the exponential moving average for the gradient was done and the square of the gradient corresponded to the estimation (computation) of the 1st and 2nd moments. Table 4 contains the RMSE values for different training periods for overall power grid failures and equipment failures.

Table 4 Computed RMSE values using ADAM optimizer in training process (Developed by the authors)

RMSE	80% Training	70% Training	60% Training
Power Grid	0.337	1.156	1.871
Equipment	0.935	2.011	2.965

Table 5 Computed RMSE values using SGDM optimizer in training process (Developed by the authors)

RMSE	80% Training	70% Training	60% Training
Power Grid	1.207	2.693	3.155
Equipment	1.563	2.119	6.171

4.2.2. Using the SGDM Optimizer

In addition, a stochastic gradient descent moment (SGDM) optimizer was also considered for learning the algorithm (model) during the training process. Table 5 gives root mean square error values between actual and predicted values with different training periods for overall power grid failures and only equipment failures.

5. Conclusion

In this comprehensive study, a deep learning algorithm called Long Short-Term Memory (LSTM) was implemented to compute the time to failure of a power grid. Results achieved from LSTM techniques having adaptive moment estimation (ADAM) optimized during training were found to be satisfactory, and better prediction and computation of TTF were achieved. The root mean square error (RMSE) values obtained through ADAM are minimum compared to SGDM optimized values. These include (a) the obtained RMSE using LSTM-ADAM optimizers for power grid using all component data is 0.337 to 1.871 for 80% to 60% training datasets, respectively; (b) the RMSE using Equipment database for 80% to 60% training is 0.935 to 2.965; (c) the resulted RMSE using LSTM-SGDM optimizers for power grid components database for 80% to 60% training data is 1.207 to 3.155, respectively; (d) finally, the obtained RMSE using SGDM optimizer for power system TTF using component failure information for 80% to 60% training is 1.563 to 6.171, respectively.

Therefore, the proposed LSTM-based prognostic framework of the ADAM optimizer can be used anywhere in any condition for power system failure analysis. It will help caretaker companies to perform predictive maintenance activities to ensure continuous power supply. It will also help reduce maintenance costs by avoiding unwanted inspections and maintenance activities. Prediction of time to failure of power grids will lead to reduction in power outages because the reliability of the power system will be improved. Afterwards, timely measures will be taken to overcome or even reduce any future failure in power systems.

5.1. Recommendations and Future Research

In this study, the deep learning-based LSTM algorithm is applied to historical datasets of power systems for failure trend forecasting. The results could be further improved using the real-time datasets of the power system components because LSTM is a long-term short-term algorithm.

The algorithm could be implemented on local available power failure data to minimize the power failures and challenges faced by power sectors within the country. Similarly, modeling of maintainability, i.e., time to repair action, can be implemented with the prediction of failure to make the system more robust.

Acknowledgement

The authors would like to thank Prince Sultan University for their support. The authors would also like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

References

[1] CANIZES B, SOARES J, VALE Z, and KHODR H M. Hybrid fuzzy Monte Carlo technique for reliability assessment in transmission power systems. *Energy*, 2012, 45(1): 1007-1017. https://doi.org/10.1016/j.energy.2012.06.049

[2] PANAGIOTIS H, TSAROUHAS A, THEODOROS H, VARZAKAS B, et al. Reliability and maintainability analysis of strudel production line with experimental data. A case study. *Journal of Food Engineering*, 2009, 91(2): 250-259. https://doi.org/10.1016/j.jfoodeng.2008.09.002

[3] TSAROUHAS P H. A comparative study of performance evaluation based on field failure data for food production lines. *Journal of Quality in Maintenance Engineering*, 2011, 17(1): 26-39. https://doi:10.1108/13552511111116231

[4] SARASWAT S, and YADAVA G S. An overview on reliability, availability, maintainability and supportability (RAMS) engineering. *International Journal of Quality & Reliability Management*, 2008, 25(3): 330-344. https://doi:10.1108/02656710810854313

[5] WANG H, YANG Z, YU Q, and HONG T. Online Reliability Time Series Prediction via Convolutional Neural Network and Long Short Term Memory for Service-Oriented Systems. Knowledge-Based Systems, 2018, 159: 132-147. https://doi.org/10.1016/j.knosys.2018.07.006

[6] SALEH A R, DRISS M, and ALMOMANI I. CBiLSTM: A hybrid deep learning model for efficient reputation assessment of cloud services. *IEEE Access*, 2022, 10: 35321-35335. https://doi.org/10.1109/ACCESS.2022.3163246

[7] DURBHAKA GK, SELVARAJ B, MITTAL M, et al. Swarm-LSTM: Condition monitoring of gearbox fault diagnosis based on hybrid LSTM deep neural network optimized by swarm intelligence algorithms. *CMC-Computers, Materials and Continua*, 2021, 66(2): 2041-2059. https://doi.org/10.32604/cmc.2020.013131

[8] AMIN J, SHARIF M, RAZA M, et al. Brain tumor detection: a long short-term memory (LSTM)-based learning model. *Neural Computing and Applications*, 2020, 32: 15965-15973. https://doi.org/10.1007/s00521-019-04650-7

[9] RASHID T A, FATTAH P, and AWLA D K. Using Accuracy Measure for Improving the Training of LSTM with Metaheuristic Algorithms. Procedia Computer Science, 2018, 140: 324-233. https://doi.org/10.1016/j.procs.2018.10.307.

[10] https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Past-outage-data

[11] MEMON ML, MAHESHWARI MK, SHIN DR, et al. A framework for deep learning–based discontinuous reception in 5G wireless networks. *Transactions on Emerging Telecommunications Technologies*, 2019, 30(3): 1-13. https://doi.org/10.1002/ett.3579.

[12] LE X-H, HO HV, LEE G, and JUNG S. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water*, 2019, 11(7): 1387. https://doi.org/10.3390/w11071387.

[13] STAUDEMEYER RC, and MORRIS ER. Understanding LSTM--a tutorial into long short-term memory recurrent neural networks, 2019, https://doi.org/10.48550/arXiv.1909.09586

参考文:

[1] CANIZES B、SOARES J、VALE Z 和 KHODR H M。用于输电系统可靠性评估的混合模糊蒙特卡罗技术。能源，2012，45(1)：1007-1017。https://doi.org/10.1016/j.energy.2012.06.049

[2] PANAGIOTIS H、TSAROUHAS A、THEODOROS H、VARZAKAS B 等。利用实验数据对果馅卷生产线进行可靠性和可维护性分析。案例研究。食品工程, 2009, 91(2): 250-259。https://doi.org/10.1016/j.jfoodeng.2008.09.002

[3] TSAROUHAS P H. 基于食品生产线现场故障数据的性能评估比较研究。维修工程质量学报, 2011, 17(1): 26-39. https://doi:10.1108/13552511111116231

[4] SARASWAT S, 和 YADAVA G S。可靠性、可用性、可维护性和可支持性(随机存取存储器)工程概述。国际质量与可靠性管理杂志，2008，25 (3)：330-344。https://doi:10.1108/02656710810854313

[5] WANG H，YANG Z，YU Q，和 HONG T。面向服务系统的基于卷积神经网络和长短期记忆的在线可靠性时间序列预测。基于知识的系统，2018，159：132-147。https://doi.org/10.1016/j.knosys.2018.07.006

[6] SALEH AR、DRISS M 和 ALMOMANI I。CBiLSTM：一种用于云服务高效声誉评估的混合深度学习模型。IEEE访问，2022，10：35321-35335。https://doi.org/10.1109/ACCESS.2022.3163246

[7] DURBHAKA GK、SELVARAJ B、MITTAL M 等。群长短期记忆网络：基于群体智能算法优化的混合长短期记忆网络深度神经网络的变速箱故障诊断状态监测。羧甲基纤维素钠-计算机、材料和康体佳，2021, 66(2): 2041-2059。https://doi.org/10.32604/cmc.2020.013131

[8] AMIN J、SHARIF M、RAZA M, 等。脑肿瘤检测：基于长短期记忆（长短期记忆网络）的学习模型。神经计算与应用，2020，32：15965-15973。https://doi.org/10.1007/s00521-019-04650-7

[9] RASHID T A、FATTAH P 和 AWLA D K。使用准确度测量通过元启发式算法改进长短期记忆网络的训练。计算机科学，2018，140：324-233。https://doi.org/10.1016/j.procs.2018.10.307。

[10] https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Past-outage-data

[11] MEMON ML、MAHESHWARI MK、SHIN DR 等。5G无线网络中基于深度学习的不连续接收框架。新兴电信技术交易，2019，30(3)：1-13。https://doi.org/10.1002/ett.3579。

[12] LE X-H、HO HV、LEE G 和 JUNG S。长短期记忆(长短期记忆网络)神经网络在洪水预报中的应用。水，2019，11(7)：1387。https://doi.org/10.3390/w11071387。

[13] STAUDEMEYER RC, 和 MORRIS ER。了解长短期记忆网络——长短期记忆循环神经网络教程, 2019，https://doi.org/10.48550/arXiv.1909.09586