Deep Neural Networks based Short Term Wind Power Forecasting

Project Guide:

Mr. Anu G Kumar
Assistant Professor

Team Members:

Akhil Ganesh M CB.EN.U4EEE18102
Anish Chandra Rayaprolu CB.EN.U4EEE18104
Akilan A CB.EN.U4EEE18103
V S Iniyavan CB.EN.U4EEE18123

Department of Electrical & Electronics Engineering

Contents

1. Introduction
2. Objectives
3. Scaling the Dataset
4. Multilayer Perceptron(MLP)
5. Convolutional Neural Networks (CNN)
6. Long Short Term Memory (LSTM)
7. Bi-Directional LSTM
8. Stacked LSTM
9. Performance Analysis
10. Conclusion
11. Time Schedule
12. References

1.Introduction

- This project aims to implement Deep Neural Networks (DNN) based models for forecasting of Power Output of a Wind Turbine.
- Compared to conventional ANNs, DNN models can hold and store more information within the neurons [2]
- It focuses on exploring various DNN models including variations of LSTM.
- It also explores if Hybrid DNN models can perform more accurately over DNN models.

Overview of 1st Review

- Dataset used: 1 year SCADA load data from Turkey, sampled at 10 minute intervals
- Necessary EDA was performed for the model prediction.
- LSTM model was developed and was concluded to have a RMSE of 57.8.

2.Objectives

To implement different Deep Learning architectures for wind power forecasting

- Implementation of different variations of LSTM
- Implementation of CNN
- Observing the results that have been obtained from the models
- Comparing the predictions of the different LSTM and CNN Architecture

3. Scaling the Dataset

- MinMax Scaler was used for the data preprocessing
- MinMax Scaler is very common when dealing with continuous data
- Normalization of data occurs where the values are between the given bounds
- Usually defined as [0,1], but in this case [-1,1], as tanh is the activation function of the LSTM network

4. Multi-layer Perceptron Model (MLP)

- MLP is the basic building block used in many Deep Learning architectures
- It is a class of feedforward neural networks and consists of several layers of nodes: one input layer, one or more hidden layers, and one output layer [13]
- Every node is connected to all the nodes of its layer, of the previous layer and of the next layer, hence called fully connected model
- Each node utilizes an activation function

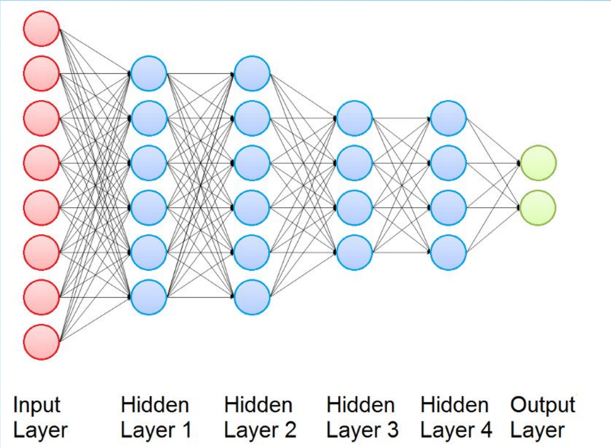

Fig 1.[13]

Model Parameters

Parameter	Value
Number of Layers	7
Layer type	Dense
Optimizer used	Adam
Activation Function	ReLU
Train/Test Split	70%

Prediction Plot

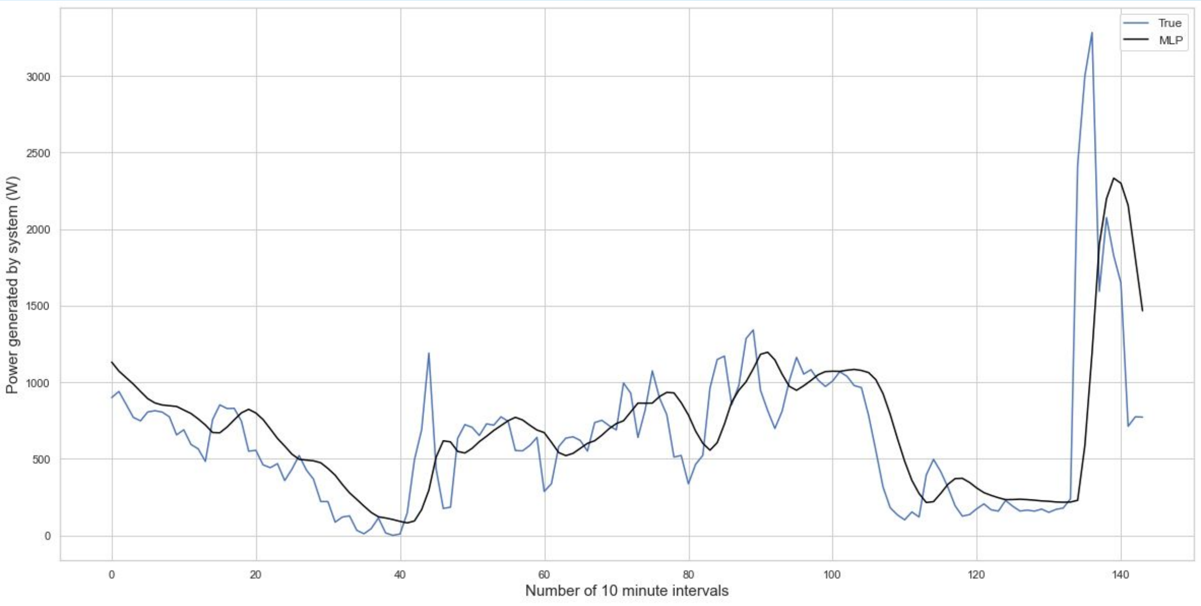

Fig 2. Prediction plot of the MLP model

MAE	MSE	RMSE
0.058	0.008	0.090

5. Convolutional Neural Network (CNN)

- CNNs use image recognition and classification in order to detect objects, recognize faces, etc and differentiate one from the other. [2]
- CNN, although popular in image datasets, can also be used on time series data [2]
- CNNs use convolution operations that can handle spatial information available in image [1][2][6]
- Through CNN we can use 1D convolutions to extract information along the time dimension
- CNN's learn patterns within the time window. When there is missing data, CNN's are very useful.[6]

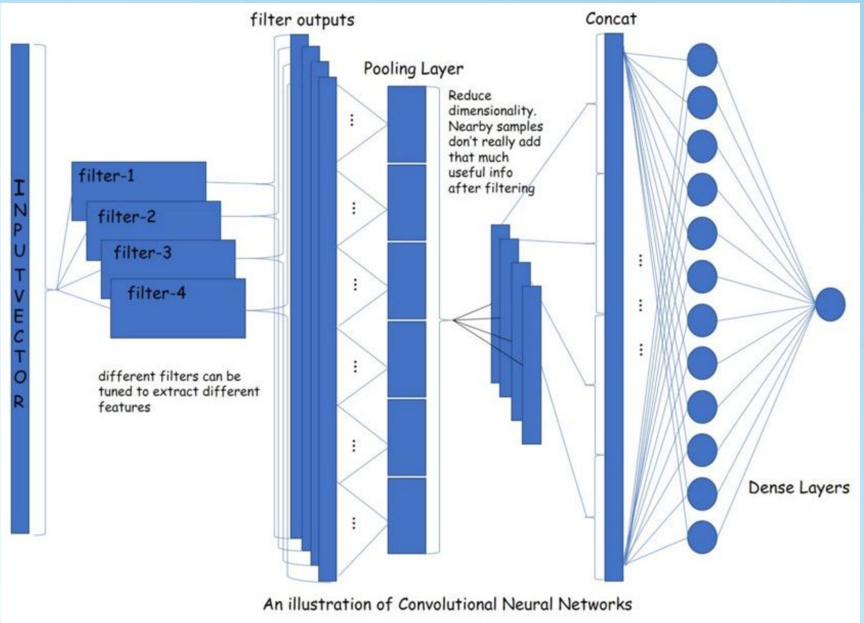

Fig 3. [11]

Model Parameters

Parameter	Value
Number of Layers	9
Layer types	Dense(4), Conv1D(2), MaxPool(2), Flatten(1)
Optimizer used	Adam
Activation Function	ReLU
Train/Test Split	70%
Time Step	144

Prediction Plot

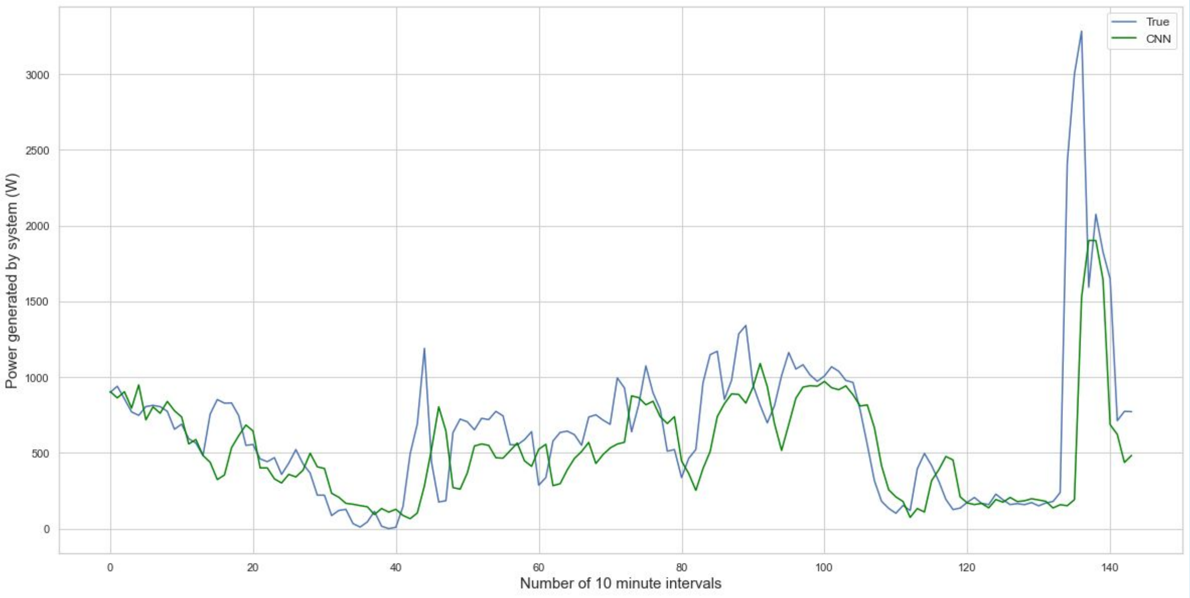

Fig 4. Prediction plot of CNN model

MAE	MSE	RMSE
0.073	0.011	0.107

6. Long Short Term Memory Model (LSTM)

- LSTM – special kind of RNN, capable of learning long-term dependencies.[3]
- The cell state act as a transport highway that transfers relative information all the way down the sequence chain.[3][7]
- LSTM has the ability to remove or add information to the cell state, carefully regulated by structures called gates.The gates can learn what information is relevant to keep or forget during training.[3][7]

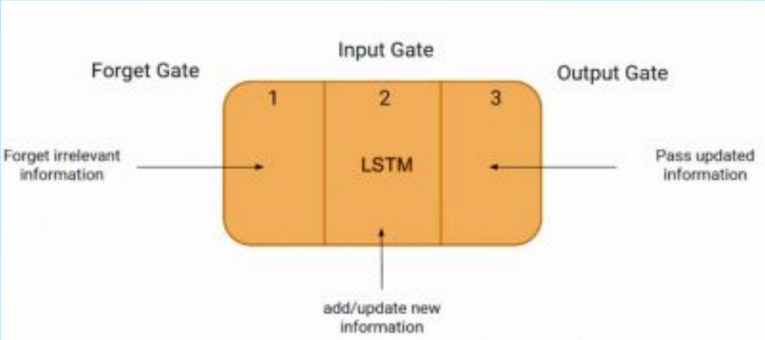

Fig 5. [8] LSTM Architecture

- LSTM unit consists two main parts:
  - The combination of the past with the present information.
  - Interaction of the combination after being processed with the present.[7]
- The information extraction of the past-present combination c_t regulated by the output switch.[10]
- The output to along with its c_t and the next input x_{t+1} will be passed to the next LSTM cell in the next time step to operate in the same way as the previous ones.[10]
- LSTM model made least error scores for majority of the months between LSTM, MLP and MA.[3][10]

Model Parameters

Parameter	Value
Number of Layers	3
Layer types	Dense, LSTM, Dropout
Optimizer used	Adam
Activation Function	Tanh
Train/Test Split	70%
Time Step	144

Prediction Plot

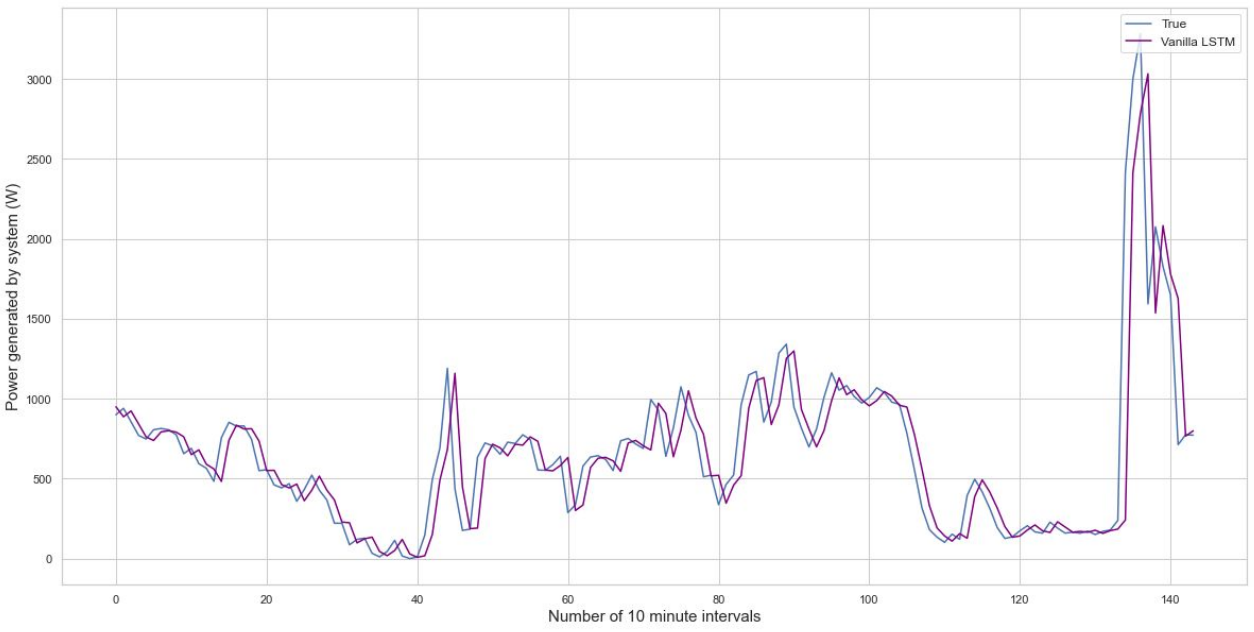

Fig 6. Prediction plot of LSTM model

MAE	MSE	RMSE
0.046	0.005	0.069

7. Bi-Directional Long Short Term Memory (Bi-LSTM)

- Bi-LSTM is the process of making any neural network have the sequence information in both directions backwards (future to past) or forward(past to future).
- In LSTM, we can make input flow only in one direction, either backwards or forward.
- In Bi-LSTM, input flows in two directions, to preserve the future and the past information.
- BI-LSTM is usually employed where the sequence to sequence tasks are needed.
- It is used in text classification, speech recognition and forecasting models.

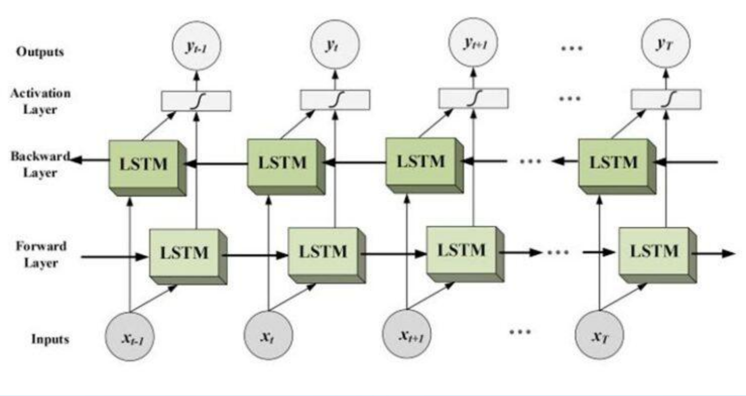

Fig 7. [14]

Model Parameters

Parameter	Value
Number of Layers	3
Layer types	Dense, Bidirectional LSTM, Dropout
Optimizer used	Adam
Activation Function	Tanh
Train/Test Split	70%
Time Step	144

Prediction Plot

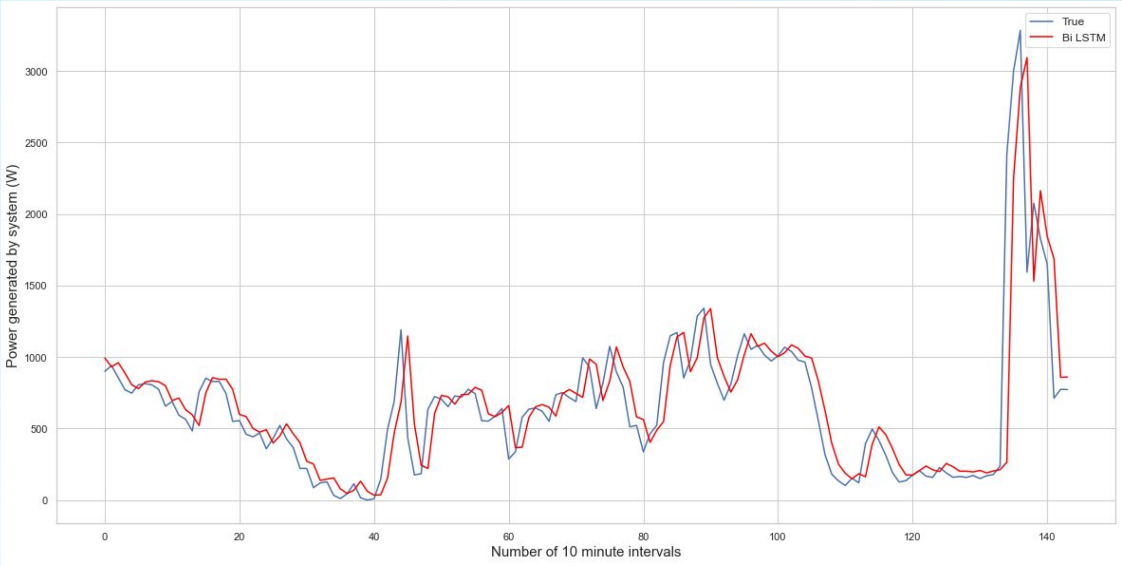

Fig 8. Prediction plot of Bi-LSTM model

MAE	MSE	RMSE
0.043	0.0045	0.067

8. Stacked LSTM

- The original LSTM model is comprised of a single hidden LSTM layer followed by a standard feedforward output layer.
- The Stacked LSTM is an expansion of this model that includes multiple hidden LSTM layers with numerous memory cells in each layer.
- LSTMs work with sequence data, which implies that adding layers over time adds degrees of abstraction to the input observations. Essentially, chunking data throughout time or expressing the problem on several time periods.

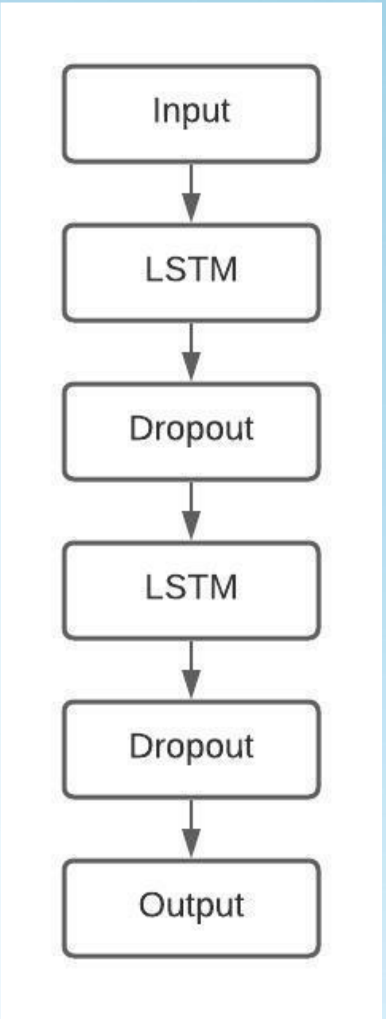

Fig 9 [11]

Model Parameters

Parameter	Value
Number of Layers	5
Layer types	Dense, LSTM, Dropout
Optimizer used	Adam
Activation Function	Tanh
Train/Test Split	70%
Time Step	144

Prediction Plot

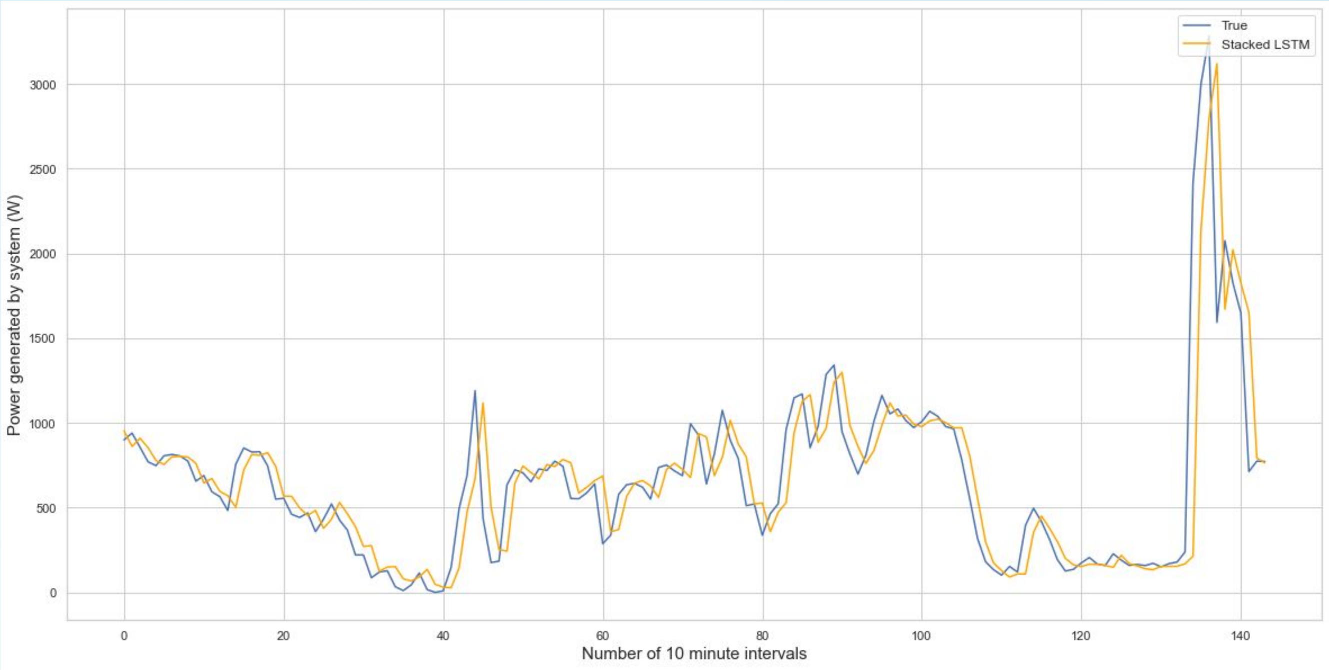

Fig 10. Prediction plot of Stacked LSTM model

MAE	MSE	RMSE
0.063	0.007	0.085

9. Performance Analysis

Error Parameters	Formulae
Mean Absolute Error (MAE)	$MAE = \sum_{i=1}^{n} \frac{\lvert y_i - x_i \rvert}{n}$
Mean Squared Error (MSE)	$M = \frac{1}{n} \sum_{t-1}^{n} \left\lvert \frac{y_i - x_i}{y_i} \right\rvert \times 10($
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum_{i-1}^{N} (y_i - x_i)^2}$

9. Comparative Analysis

Architecture	MAE	MSE	RMSE
BI-LSTM	**0.043**	**0.0045**	**0.067**
LSTM	0.046	0.005	0.069
Stacked LSTM	0.063	0.007	0.085
MLP	0.058	0.008	0.090
CNN	0.073	0.011	0.107

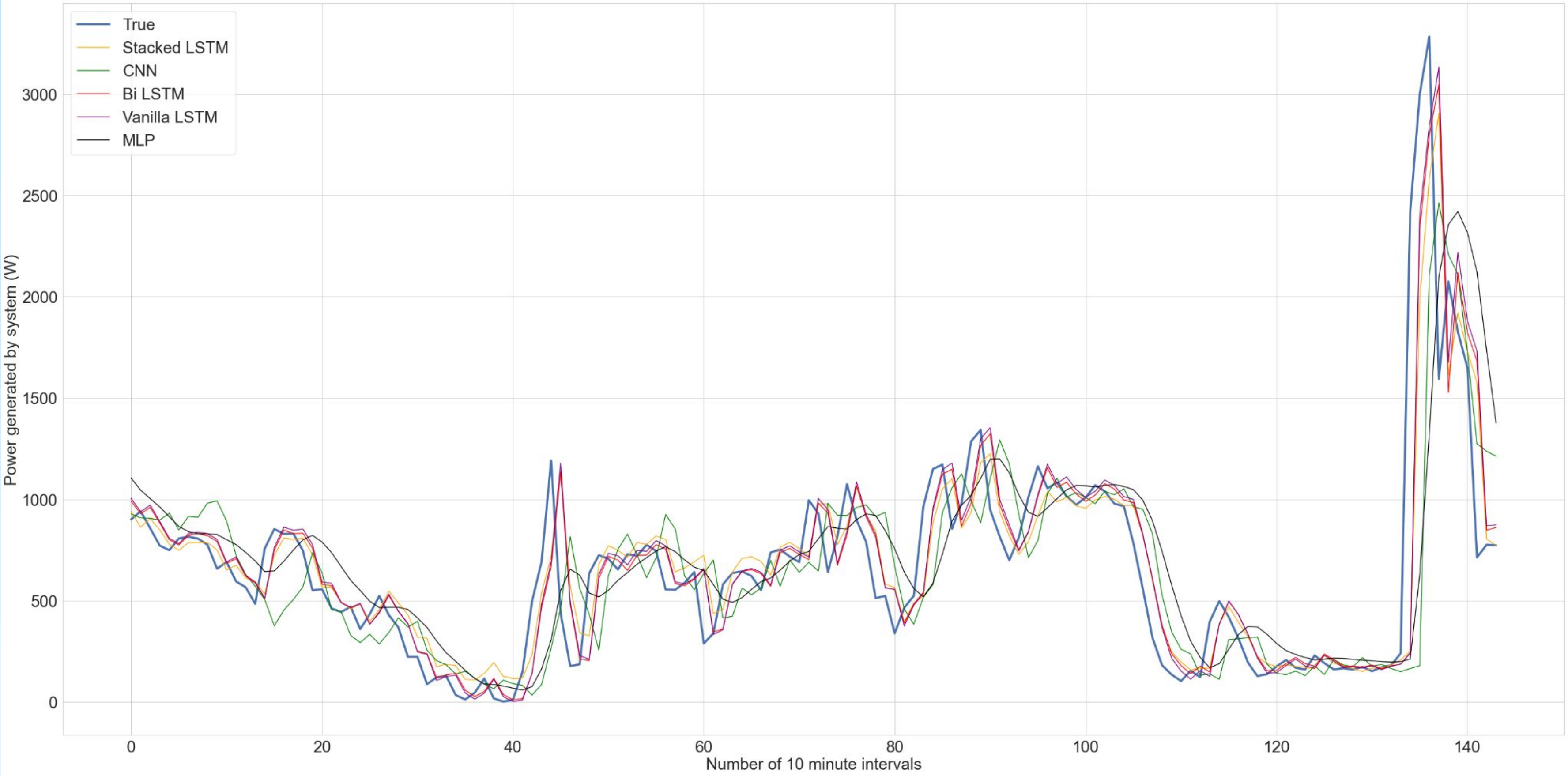

Fig 11. Comparative Prediction plot

11. Conclusion

- The MLP, CNN and different variations of the LSTM model were built and model training has been done.
- The Performance Analysis was completed for all the aforementioned models and the results were observed.
- A comparative study was done between the performance analysis of all the different models with Bi-LSTM performing the best

12. Time Schedule

Time Frame	Details of the task to be Completed
Review 0 - 18th August 2021	DNN literature survey
Review 1 - 29nd September 2021	Exploratory Data Analysis (EDA) DNN implementation - LSTM Model Performance Evaluation
Review 2 - 9th November 2021	DNN Implementation - Bi-LSTM and Stacked LSTM DNN implementation - CNN Model Performance Comparison
Phase 2 - January - February 2022	Hybrid DNN implementation (CNN-LSTM) (CNN-RBFNN-DGF)
Phase 2- March - April 2022	Ensemble DNN architecture

13. References

[1] Ying-Yi Hong, Christian Lian Paulo P. Rioflorido, “A hybrid deep learning-based neural network for 24-h ahead wind power forecasting” Applied Energy, vol. 250, pp. 530-539, September 2019.

[2] Runge, Jason, and Radu Zmeureanu, "A review of deep learning techniques for forecasting energy use in buildings" Energies, vol. 14, pp 608, 2021.

[3] R. Peña-Gallardo and A. Medina-Rios, "A comparison of deep learning methods for wind speed forecasting" 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), November 2020.

[4] D. A. Rosa de Jesús, P. Mandal, Y. -K. Wu and T. Senjyu, "Deep Learning Ensemble Based New Approach for Very Short-Term Wind Power Forecasting," 2020 IEEE Power & Energy Society General Meeting (PESGM), pp. 1-5, December 2020.

13. References

[5] Vaibhav Kumar and M L Garg, "Comparison of HDNN with other Machine Learning Models in Stock Market Prediction" International Journal of Computer Applications, vol. 182, pp. 1-9, July 2018.

[6] Bendong Zhao et al. “Convolutional neural networks for time series classification” Journal of Systems Engineering and Electronics vol. 28, no. 1, pp. 162 – 169, February 2017.

[7] C. Tian, J. Ma, C. Zhang, and P. Zhan, “A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network,” Energies, vol. 11, no. 12, pp. 3493, December 2018.

[8] Y. Qin et al., “Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal,” Appl. Energy, vol. 236, pp. 262–272, February 2019.

13. References

[9] T. Kaur, S. Kumar and R. Segal, "Application of artificial neural network for short term wind speed forecasting," 2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE), pp. 1-5, July 2016.

[10] I. Delgado and M. Fahim, “Wind Turbine Data Analysis and LSTM-Based Prediction in SCADA System,” Energies, vol. 14, no. 1, pp. 125, December 2020.

[11] Jason Brownlee, “Long Short-Term Memory Networks,” Medium, [Online]. https://machinelearningmastery.com/stacked-long-short-term-memory-networks/#:~:text=A%20Stacked%20LSTM%20architecture%20can,for%20all%20input%20time%20steps (Accessed: 10-Nov-2021)

[12] P. Sethuraman, “A comparison of DNN, CNN and LSTM using TF/Keras,” Medium, [Online].https://towardsdatascience.com/a-comparison-of-dnn-cnn-and-lstm-using-tf-keras-2191f8c77bbe. (Accessed: 8-Nov-2021)

13. References

[13] S. Verma, “Understanding input and output shapes in convolution neural network: Keras,” [Online]. https://towardsdatascience.com/understanding-input-and-output-shapes-in-convolution-network-keras-f143923d56ca. [Accessed: 9-Nov-2021].

[14] Yugesh Verma, “Complete Guide to Bidirectional LSTM,” [Online] https://analyticsindiamag.com/complete-guide-to-bidirectional-lstm-with-python-codes/ [Accessed: 2-Nov-2021].

Thank you