

PAPER • OPEN ACCESS

## Forecast Energy Consumption Time-Series Dataset using Multistep LSTM Models

To cite this article: S. Nazir *et al* 2021 *J. Phys.: Conf. Ser.* **1933** 012054

View the [article online](#) for updates and enhancements.

You may also like

- [Neuromorphic deep spiking neural networks for seizure detection](#)  
Yikai Yang, Jason K Eshraghian, Nhan Duy Truong et al.
- [Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional LSTM networks for South Korea](#)  
Jong-Min Yeom, Ravinesh C Deo, Jan F Adamowski et al.
- [Dual temporal attention mechanism-based convolutional LSTM model for industrial dynamic soft sensor](#)  
Jiarui Cui, Yuyu Shi, Jian Huang et al.



The Electrochemical Society  
Advancing solid state & electrochemical science & technology

**247th ECS Meeting**  
Montréal, Canada  
May 18-22, 2025  
*Palais des Congrès de Montréal*

**Showcase your science!**

**Abstracts due December 6th**

# Forecast Energy Consumption Time-Series Dataset using Multistep LSTM Models

S. Nazir<sup>1\*</sup>, Azlan Ab Aziz<sup>1</sup>, J. Hosen<sup>1</sup>, Nor Azlina Aziz<sup>1</sup>, G. Ramana Murthy<sup>2</sup>

<sup>1</sup> Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia

<sup>2</sup> Vignan's Foundation for Science, Technology and Research, Vadlamudi, Guntur, 522213 Andhra Pradesh, India

\*syednazir13@gmail.com

**Abstract.** Smart grid and smart metering technologies allow residential consumers to monitor and control electricity consumption easily. The real-time energy monitoring system is an application of the smart grid technology used to provide users with updates on their home electricity consumption information. This paper aims to forecast a month ahead of daily electricity consumption time-series data for one of the real-time energy monitoring system named beLyfe. Four separate multistep Long Short Term Memory LSTM neural network sequence prediction models such as vanilla LSTM, Bidirectional LSTM, Stacked LSTM, and Convolutional LSTM ConvLSTM has been evaluated to determine the optimal model to achieve this objective. A comparison experiment is performed to evaluate each multistep LSTM model performance in terms of accuracy and robustness. Experiment results show that the ConvLSTM model achieves overall high predictive accuracy and is less computationally expensive during model training than remaining models.

## 1. Introduction

Smart grids have recently drawn growing interest due to their stability, durability, sustainability, and performance. A traditional smart grid comprises different components such as smart meters, electricity management systems, energy storage systems, and renewable energy resources [1], [2]. This research focuses on smart grid applications such as a real-time energy monitoring system named beLyfe, which helps users monitor real-time electricity consumption to manage electricity cost. Incorporating predictability features in such systems would allow users and power suppliers to manage demand and electricity costs [3].

Forecasting electricity time-series data can be classified into three main categories: short-term, medium-term, and long-term [3]. The short and medium terms of electricity forecasting are mainly required for cost management and demand and response management [4], whereas the long-term forecasting is usually done to identify a specific power generation plant's production capacity for the future [5]. The selection of an appropriate forecasting model has a significant influence on forecasting ranges. In long-term forecasting, statistical models such as ARIMA and SARIMA are frequently studied due to predicting data containing trends and seasonal components. These models could have reasonable prediction outcomes where data patterns involve linearity and are less computationally expensive than recent machine learning models. In short-term electricity forecasting, data trends involve non-linear abrupt changes that are complex to forecast [6]. The deep neural network models can forecast time-series data, including non-linear patterns, through automatic learning of temporal dependency and mapping from input to output [7].



In recent studies, the deep neural network model, particularly the recurrent neural network RNN, primarily developed to predict sequence data, has received considerable attention to predicting short-term electricity load time-series data [8], [9]. RNN is often suffered due to vanishing and exploding gradient problem; however, this issue has been overcome by the improved version of the RNN model named long short-term memory LSTM model. The LSTM model works well in sequence data prediction and could be used for different purposes such as text translation, sentiment analysis, time-series forecasting. Several architecture-related changes have been made in the LSTM model to achieve prediction accuracy on different data characteristics.

This paper aims to forecast month ahead daily electricity consumption time-series data of beLyfe real-time energy monitoring system. Predicting short-term electricity time-series data is difficult as it involves a non-linear abrupt shift in the data pattern. To accomplish this task, we evaluate four separate multistep LSTM sequence prediction models, such as Vanilla LSTM, Bidirectional LSTM, Stacked LSTM, and Convolutional ConvLSTM. A comparison experiment is performed between these models using four separate household quarterly datasets of electricity consumption obtain from application programming interface APIs of beLyfe energy monitoring system. Each model significant hyper-parameter selection scanning is performed through manual tuning since neural network performance depends on data characteristics [10]. This experiment performs hyper-parameter tuning on each model through different combinations of neuron sizes, filter sizes, activation functions, and optimization algorithms. Experimental results indicate that the ConvLSTM multistep sequence prediction model outperformed the remaining models in overall predictive accuracy and model fitness on input data.

## 2. Literature Review

The LSTM model has been extensively discussed in the short-term load forecasting STLTF in the last few years. Yunpeng et al. [11] evaluate three different time-series data patterns in forecasting multistep ahead time-series data. The study indicates that the LSTM model accuracy is acceptable; however, the computational complexity is high compared to statistical models. Choi et al. [9] proposed a novel framework by combining Residual network ResNet with LSTM model to improve forecasting accuracy by incorporating latent features of input historical load data through ResNet.

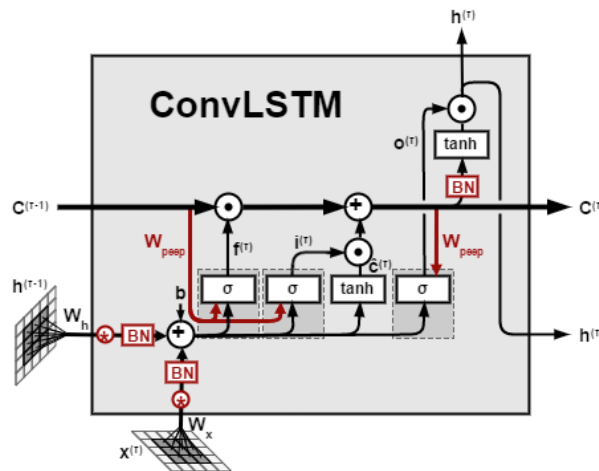
Kong et al. [12] proposed an LSTM-based deep learning forecasting framework by determining residential activities volatile behaviour features regarding the electricity consumption for the individual appliances time-series data. Kong et al. [13] proposed a similar LSTM based forecasting framework; in this study, a density-based clustering approach is used to determine the aggregated load and individual load profiles for consistency analysis. Essien et al. [14] propose a novel deep learning modelling-based framework using convolutional ConvLSTM stacked autoencoders (SAE). The author incorporates wavelet transforms (WT), 2-dimensional CNNs and LSTM stacked autoencoders (SAE) towards single-step time series prediction in this framework.

Yan et al. [15] proposed a hybrid framework by incorporating the stationary wavelet transformation SWT with LSTM models to forecast household five-minute incremental electricity load time-series data. Kim et al. [16] proposed a novel multi-channel LSTM forecasting model incorporating three different channels: power consumption, time location, and customer behaviour to extract each channel features individually by a single LSTM parallel structure to forecast multistep power consumption time-series data with a multiple output strategy.

The hybrid CNN-LSTM has also been frequently utilized in several research articles on STLTF [6], [17]. The CNN-LSTM model can extract spatial and temporal features from electricity load time-series data for prediction. Kim et al. [18] recently proposed a novel hybrid CNN-LSTM model for forecasting residential house energy consumption. A significant drawback of this work is discovering optimal hyper-parameters of the CNN-LSTM model. The paper is organized as follows; at first, we briefly introduce the architecture of Convolutional LSTM ConvLSTM, then describe experimental setup, then discussed obtain results finally conclusion is drawn.

### 3. ConvLSTM

The ConvLSTM neural network is an LSTM [19] variant containing a convolution operation inside the LSTM cell, as shown in Fig 1. The ConvLSTM is a fully-connected (FC-LSTM) extension to capture spatial similarity by integrating convolutional structures in both the transitions from input to state and state to state [20]. The distinguishing characteristics of the ConvLSTM model are; the inputs, cell outputs, hidden states and gates are 3D tensors in which the last two dimensions are spatial such as a grid of rows and columns. The ConvLSTM defines a particular cell's future state in the grid from its nearby neighbour's inputs and past states to get a clearer view of the inputs and states. Using a convolution operator in the state-to-state and input-to-state transitions will effectively achieve this. The layers of the ConvLSTM are stacked together to form an encoding-forecasting structure.



**Figure 1.** A ConvLSTM cell structure.

The equations of ConvLSTM representing the internal cell functionality are shown in (1) below, where (\*) denotes the convolution operator and ( $\circ$ ) denotes the Hadamard product as before:

$$\begin{aligned}
 i_t &= \sigma(W_{xi} * X_t + W_{hi} * h_{t-1} + W_{ci} \circ C_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf} \circ C_{t-1} + b_f) \\
 C_t &= f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo} * X_t + W_{ho} * h_{t-1} + W_{co} \circ C_t + b_o) \\
 h_t &= o_t \circ \tanh(C_t)
 \end{aligned} \tag{1}$$

In this equation (1), the  $i_t$  represents the equation of newly inputs in the ConvLSTM cell, and  $f_t$  represents the forget gate equation used to forget past cell status,  $C_{t-1}$ , and  $C_t$  represents latest cell output which would propagate to the final state  $h_t$ , and  $o_t$  represents the output gate controlled by the state  $h_t$ .

### 4. Experiment

The experiment is conducted on python 3.3.7 with Keras deep learning library. Deep learning models training must involve discovering significant hyper-parameters to demonstrate high-performance accuracy and computing time [21]. The significant hyper-parameters use in the training of each multistep LSTM sequence prediction models are; the number of neurons or units and convolutional filter sizes, activation functions and optimization learning algorithms. The chosen hyper-parameter combination for an experiment is mentioned below in Table 1.

**Table 1.** Scalar Decomposition

<b>Hyper-parameter</b>	<b>Vanilla</b>	<b>Bidirectional</b>	<b>Stacked</b>	<b>ConvLSTM</b>
Neurons	100	100	100	100
Filters	-	-	-	64
Activation Function	ReLU	ReLU	ReLU	ReLU
Optimization Algorithm	Adam	Adam	Adam	Adam

The batch size and the number of epochs consider significant parameters carefully selected to achieve the best fit train model on the input data set within the least training computation time. For this experiment, the chosen batch size and epochs are 32 and 500, respectively. An early stopping method [22] is used in this experiment to avoid overfitting and stop training once the model performance stops improving. A split validation parameter with a value between 0 and 1 required for early stopping is defined as 0.67, indicating that the models use 67.7% input data for training and 33.3% for validation.

## 5. Results and Discussion

The four different multistep LSTM sequence models prediction errors are measure through the root mean squared error (RMSE) standard deviation, as shown in equation 2.

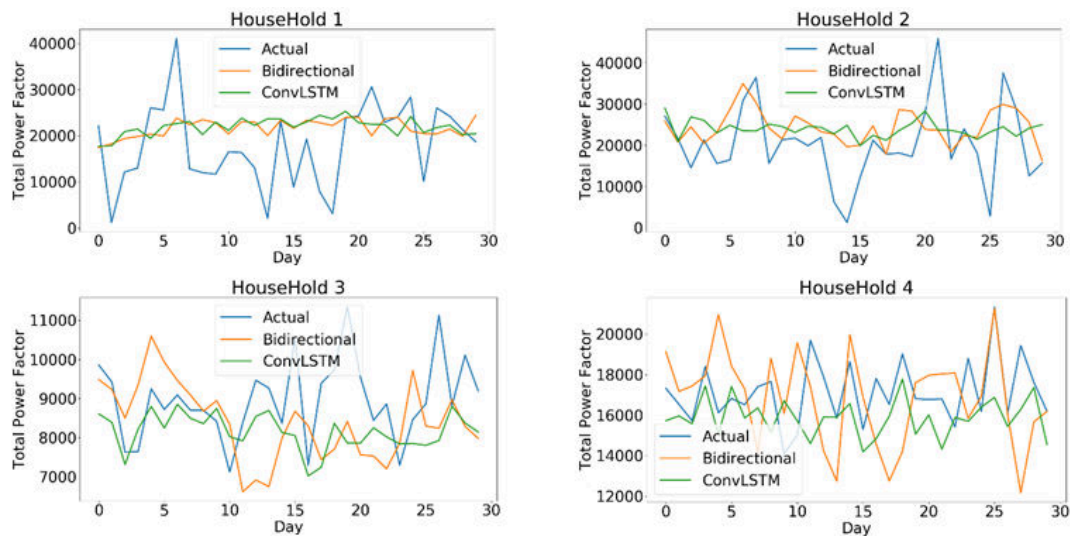
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

Table 2 represents each multistep LSTM model comparative experimental prediction results of future month electricity consumption time-series dataset of four separate household datasets.

**Table 2.** The RMSE error score of four different multistep LSTM sequence prediction models.

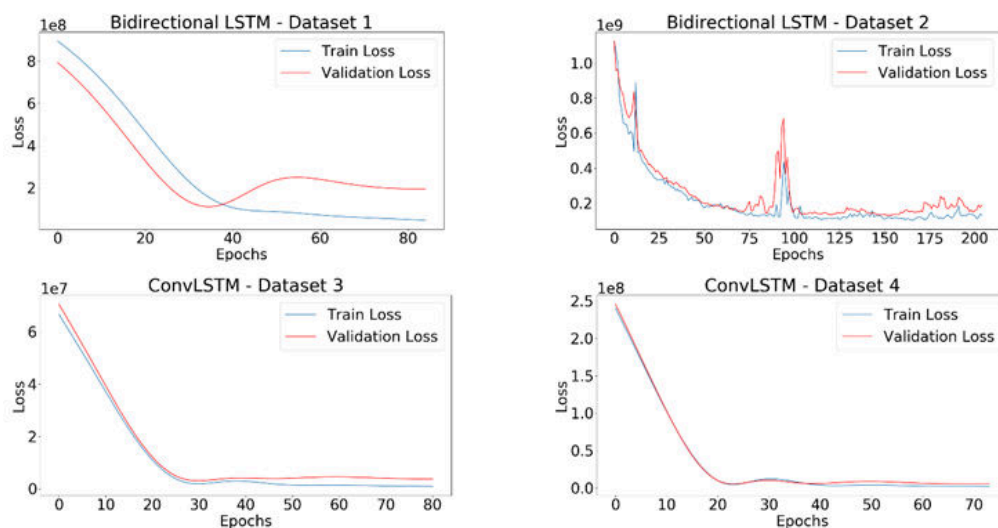
<b>Dataset</b>	<b>Vanilla</b>	<b>Bidirectional</b>	<b>Stacked</b>	<b>ConvLSTM</b>
Household 1	12282	9511	12101	9800
Household 2	10934	9731	10514	10109
Household 3	1555	1518	3181	1300
Household 4	2354	2750	2401	1993
<b>Average</b>	<b>27125</b>	<b>23510</b>	<b>28197</b>	<b>23202</b>

Table 2 shows that the convolutional ConvLSTM outperformed the remaining models on average. Bidirectional LSTM is performed slightly lower in terms of RMSE error measurement from ConvLSTM on average, although it generated fewer errors in household datasets 1, and 2. The ConvLSTM and Bidirectional multistep LSTM model monthly test dataset prediction results using four different quarterly household datasets are shown below in Figure 2.



**Figure 2.** Monthly test dataset predicted outcomes of the ConvLSTM and Bidirectional multistep LSTM models using four different household quarterly datasets.

We also examined each multistep LSTM model fitness in this experiment by evaluating the train and validation loss patterns over each epoch during training. Once the model fit function is called, the model training history is returned, including training and validation losses. These losses are calculated at each epoch to help diagnose the model fitness.



**Figure 3.** Training and validation loss graphs of ConvLSTM and Bidirectional LSTM.

In order to prevent under fitting during model training, we define a patience parameter with a value of 50 in the early stopping class function `Object() { [native code] }`, which specifies the number of epochs with no improvement during which the training will be stopped [23]. The patience parameter provides an extension to further optimize the model training and validation losses during the early stopping phase. An early stopping also reduces deep neural network model computational time complexity during training [24]. We have presented training and validation loss graphs of ConvLSTM and Bidirectional multistep LSTM models on four different household data samples on which these models mainly generate least prediction errors, as shown below in Figure 3. These models training and validation loss graphs demonstrate that training stops at less than 100 epochs, except for the Bidirectional LSTM at dataset 2, which takes 200 epochs during model training. It could be understood that, as the number of epochs increases, the computational training time of the model would also increase. The multistep ConvLSTM model also outperformed Bidirectional LSTM in terms



of computational time, according to Figure 3. However, the ConvLSTM model forecast errors require further improvement by incorporating meteorological factors into input electricity consumption time-series data to turn it into multivariate time-series data [25]. In this comparative study, we conclude that the multistep ConvLSTM model could be used in the prediction of monthly electricity consumption time-series data for beLyfe real-time energy monitoring system, which is quite challenging to forecast through a traditional statistical model such as exponential smoothing (ES) and ARIMA which require data to be stationary before forecast.

## 6. Conclusion

The electricity consumption forecasting gains considerable attention in the research field after smart grid technology development. Knowing the future electricity consumption would help electricity service providers effectively manage demand and cost. In this paper, we determine four different multistep LSTM forecasting models performance in terms of prediction accuracy and robustness. A comparison experiment is conducted between four different multistep LSTM forecasting models to determine the best one. In this experiment, four different household daily basis electricity consumption time-series datasets have been used to forecasting the future month electricity consumption for beLyfe real-time electricity monitoring system. We used quarterly data to train each multistep LSTM model which is further split into training and validation sets. The dataset split would help diagnose the model fitness on the training input dataset through providing training history in returns. The prediction residuals errors of each multistep LSTM model are measure through RMSE error measure. The experimental results indicate that the ConvLSTM model outperformed other remaining multistep LSTM in terms of prediction accuracy and efficiency on average.

## References

- [1] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 25, no. 2, pp. 303–315, 2014, doi: 10.1109/TNNLS.2013.2276053.
- [2] J. Moon, S. Park, S. Rho, and E. Hwang, "A comparative analysis of artificial neural network architectures for building energy consumption forecasting," *Int. J. Distrib. Sens. Networks*, vol. 15, no. 9, 2019, doi: 10.1177/1550147719877616.
- [3] L. Sehovac, C. Nesen, and K. Grolinger, "Forecasting building energy consumption with deep learning: A sequence to sequence approach," *Proc. - 2019 IEEE Int. Congr. Internet Things, ICIOT 2019 - Part 2019 IEEE World Congr. Serv.*, pp. 108–116, 2019, doi: 10.1109/ICIOT.2019.00029.
- [4] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Appl. Energy*, vol. 212, no. October 2017, pp. 372–385, 2018, doi: 10.1016/j.apenergy.2017.12.051.
- [5] R. Jamil, "Hydroelectricity consumption forecast for Pakistan using ARIMA modeling and supply-demand analysis for the year 2030," *Renew. Energy*, vol. 154, pp. 1–10, 2020, doi: 10.1016/j.renene.2020.02.117.
- [6] 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, 2018, doi: 10.3390/en11123493.
- [7] S. Hosein and P. Hosein, "Load forecasting using deep neural networks," *2017 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. ISGT 2017*, 2017, doi: 10.1109/ISGT.2017.8085971.
- [8] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," *Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017*, vol. 2017-Decem, pp. 511–516, 2017, doi: 10.1109/ICMLA.2017.0-110.
- [9] H. Choi, S. Ryu, and H. Kim, "Short-Term Load Forecasting based on ResNet and LSTM," *2018 IEEE Int. Conf. Commun. Control. Comput. Technol. Smart Grids, SmartGridComm 2018*, pp. 1–6, 2018, doi: 10.1109/SmartGridComm.2018.8587554.
- [10] C. Lang, F. Steinborn, O. Steffens, and E. W. Lang, "Electricity load forecasting - An

- evaluation of simple 1D-CNN network structures,” arXiv, 2019.
- [11] L. Yunpeng, C. Software, and X. Jiaotonguniversity, “Multistep ahead time series forecasting for different data patterns based on LSTM recurrent neural network,” pp. 305–310, 2017, doi: 10.1109/WISA.2017.25.
  - [12] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, “Short-term residential load forecasting based on resident behaviour learning,” IEEE Trans. Power Syst., vol. 33, no. 1, pp. 2016–2017, 2018, doi: 10.1109/TPWRS.2017.2688178.
  - [13] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network,” IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 841–851, 2019, doi: 10.1109/TSG.2017.2753802.
  - [14] A. Essien and C. Giannetti, “A Deep Learning Framework for Univariate Time Series Prediction Using Convolutional LSTM Stacked Autoencoders,” IEEE Int. Symp. Innov. Intell. Syst. Appl. INISTA 2019 - Proc., pp. 1–6, 2019, doi: 10.1109/INISTA.2019.8778417.
  - [15] K. Yan, W. Li, Z. Ji, M. Qi, and Y. Du, “A Hybrid LSTM Neural Network for Energy Consumption Forecasting of Individual Households,” IEEE Access, vol. 7, pp. 157633–157642, 2019, doi: 10.1109/ACCESS.2019.2949065.
  - [16] C. S. O. O. Kim, “Multi-Step Short-Term Power Consumption Forecasting Using Multi-Channel LSTM With Time Location Considering Customer Behavior,” pp. 125263–125273, 2020, doi: 10.1109/ACCESS.2020.3007163.
  - [17] K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, and H. Zhou, “Multistep short-term power consumption forecasting with a hybrid deep learning strategy,” Energies, vol. 11, no. 11, pp. 1–15, 2018, doi: 10.3390/en11113089.
  - [18] T. Y. Kim and S. B. Cho, “Predicting residential energy consumption using CNN-LSTM neural networks,” Energy, vol. 182, pp. 72–81, 2019, doi: 10.1016/j.energy.2019.05.230.
  - [19] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
  - [20] Z. Chao, F. Pu, Y. Yin, B. Han, and X. Chen, “Research on real-time local rainfall prediction based on MEMS sensors,” J. Sensors, vol. 2018, pp. 1–9, 2018, doi: 10.1155/2018/6184713.
  - [21] M. Massaoudi, S. S. Refaat, I. Chihi, M. Trabelsi, H. Abu-Rub, and F. S. Oueslati, “Short-Term Electric Load Forecasting Based on Data-Driven Deep Learning Techniques,” IECON Proc. (Industrial Electron. Conf., vol. 2020-Octob, pp. 2565–2570, 2020, doi: 10.1109/IECON43393.2020.9255098.
  - [22] M. Li, M. Soltanolkotabi, and S. Oymak, “Gradient descent with early stopping is provably robust to label noise for overparameterized neural networks,” arXiv, vol. 108, 2019.
  - [23] Y. E. Cebeci, “A Recurrent Neural Network Model for Weather Forecasting,” UBMK 2019 - Proceedings, 4th Int. Conf. Comput. Sci. Eng., pp. 591–595, 2019, doi: 10.1109/UBMK.2019.8907196.
  - [24] S. Y. K. Wong, J. S. K. Chan, L. Azizi, and R. Y. D. Xu, “Time-varying neural network for stock return prediction,” arXiv, 2020.
  - [25] X. Cao, S. Dong, Z. Wu, and Y. Jing, “A data-driven hybrid optimization model for shortterm residential load forecasting,” Proc. - 15th IEEE Int. Conf. Comput. Inf. Technol. CIT 2015, 14th IEEE Int. Conf. Ubiquitous Comput. Commun. IUCC 2015, 13th IEEE Int. Conf. Dependable, Auton. Se, pp. 283–287, 2015, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.41.