PROCEEDINGS OF ECOS 2023 - THE 36TH INTERNATIONAL CONFERENCE ON
EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS
25-30 JUNE, 2023, LAS PALMAS DE GRAN CANARIA, SPAIN

Modeling of Submerged Membrane Bioreactor Filtration using Deep Learning Neural Networks

Nur Sakinah Ahmad Yasmin[a], Norhaliza Abdul Wahab[b], Kumerasan A. Danapalasingam[c] and R. Vilanova[d]

[a] *Faculty of Electrical Engineering Universiti Teknologi Malaysia, 81310 Skudai Johor Bahru, Malaysia, nursakinahyasmin@gmail.com*
[b] *Faculty of Electrical Engineering Universiti Teknologi Malaysia, 81310 Skudai Johor Bahru, Malaysia, norhaliza@utm.my, CA*
[c] *Faculty of Electrical Engineering Universiti Teknologi Malaysia, 81310 Skudai Johor Bahru, Malaysia, kumeresan@fke.utm.my*
[d] *Department Telecommunications and Systems Engineering Universitat Autònoma de Barcelona, Spain, Ramon.Vilanova@uab.cat*

Abstract:

Wastewater treatment (WWTP) is one of the major challenges due to the growing global population. Despite the complexity of the non-linearity and dynamic of environmental data, deep learning technologies should be created for WWTP. In this study, deep learning of long short-term memory (LSTM) is adopted to forecast permeate flux in membrane bioreactor of WWTP with five parameters involved such as permeate flux, transmembrane pressure (TMP), air flow, pump (voltage) and backwash. Three deep learning models derived from LSTM namely vanilla LSTM, bidirectional LSTM, and stacked LSTM including recurrent neural network (RNN) were constructed. The proposed LSTM's models provide promising results with 90 % accuracy of the predicted permeate flux model and can aid in establishing fouling backwashing process strategies, leading to reduced capital, energy consumption, and operational cost. Therefore, the proposed LSTM model is an adequate interpolation tool to predict the permeate flux of membrane bioreactor process in WWTP systems.

Keywords:

LSTM; Deep Learning; Neural Network; WWTP.

1. Introduction

One of the promising 21st-century technologies is the membrane bioreactor, which combines a biological process with permeable membranes. Since then, this technology has experienced constant advancement, and its uses have covered a wide range of industries, including water treatment, wastewater reclamation, juice concentration, dairy production, medical use, and cell harvesting [1-5]. But a serious issue that continues to restrict this technique's potential is membrane fouling. Fouling lead to an increase in operational costs due to more labour expenses for maintenance, increase energy demand, increase cleaning chemical expenses that will lead to shorter membrane life [6, 7]. It needs to be controlled and minimized through effective and efficient approaches.

The formation of fouling includes several mechanisms including adsorption, pore blockage and the formation of cake on the membrane [6]. Adsorption takes place when there is certain interaction between solutes/particles with the membrane. It results from the surface energy process and thermodynamic equilibrium process. Internal fouling caused by in-pore adsorption can contribute to overall flux drop as well as rise in transmembrane pressure (TMP) since the pore of membrane are equivalent to those of many macromolecules [8] [9]. Meanwhile the pore blockage is caused by entire or partial blockage by colloids and particles that cause pore obstruction [10]. Cake formation is the process by which particles accumulate on a membrane outside surface layer by layer, increasing the flow of permeates resistance. The additional resistance is known as a cake resistance, and the process is frequently referred to as the development of a fouling cake.

In order to maintain the membranes' ability to function sustainably, routine operations must include cleaning the membranes—both physically and chemically. Chemical cleaning is recommended to be avoided or limited in frequency in the full-scale application of submerged membranes due to the detrimental effects it has on the membrane [11]. However, by using efficient physical cleaning techniques, which would lengthen the membrane lifetime, the use of cleaning agents can be decreased. Procedures for backwashing are essential to the longevity and effective operation of low-pressure membranes. As a result, this operating technique has become essential to ensuring reliable, consistent, and high-quality water [12]. Backwashing is a vital

https://doi.org/10.52202/069564-0135

component of physical cleaning and is done by pushing a reversed flow through a membrane from the permeate side to the feed side while utilizing a specific kind of medium. Backwashing causes foulants that were deposited or adsorbed on the membrane surface or pores to become loosened or detached.

The widespread application of artificial intelligence (AI) in areas including healthcare, smart cities, intelligent search, big data, and pattern recognition, as well as its rapid development, present a huge chance to accomplish this goal. One of the most popular AI method used for water quality prediction is artificial neural network (ANN) [13]. ANNs are feedforward neural networks consisting of multiple layers of interconnected nodes, or neurons. They are primarily used for supervised learning tasks, such as image classification and regression, and are not well-suited for handling sequential data. As a typical representative of ANN, the conventional feed-forward neural network (FFNN) and its upgraded algorithm are a classic example of neural networks and have been successfully used to predict water quality [14-16]. Another neural network that has been extensively used in model prediction in various water environment is radial basis function (RBF), as it has a straightforward structure, quick training time and capacity to estimate any functions globally with arbitrary precision [17-19]. However, the above ANN's model are not designed to consider the sequential nature of data, and thus are not well-suited for tasks such as time-series prediction problems. Hence, the so-called time series is a collection of observations that have been made in a time. Compared to automatic water quality monitoring plant, automatically collects the quality water parameters at a defined time interval such as once per day and uploads it into a server to reflect the variations of water quality. As a result, time series data is presented for water quality metrics.

On the other hand, one of deep learning neural networks that are designed to handle sequential data, such as time series-data is known as long short-term memory (LSTM) is highly recommended for time-series prediction [20-22]. LSTMs are specifically designed to address the problem of vanishing gradients in recurrent neural networks (RNN), which can make it difficult to train models that rely on long-term dependencies in the input data. LSTMs use a gating mechanism to selectively remember or forget information over time, which makes them particularly useful for tasks such as speech recognition, machine translation, and sentiment analysis. In summary, while ANNs are used for a wide range of tasks and have a general-purpose architecture, LSTMs are specifically designed to handle sequential data and have a unique architecture that enables them to model long-term dependencies in the input data. Due to those characteristics of LSTM, the research has grown significantly as a result of their successful time-series prediction ability [23]. In our knowing, just a few research have used LSTM models to predict permeate flux in membrane for drinking water quality process [20, 24].

The main objective of the current study is to develop and validate the capabilities of deep learning neural network models based on LSTMs, including vanilla-LSTM, bi-directional LSTM, and stacked LSTM, in predicting the permeate flux in a submerged membrane bioreactor. In addition to LSTM models, a comparison will be made with another neural network model called RNN. The study utilizes a dataset consisting of approximately 4021 samples, encompassing five key parameters: permeate flux, transmembrane pressure (TMP), air flow, pump voltage, and backwash. These parameters are utilized to construct the predictive models. The results of the study demonstrate the potential of deep learning LSTM models in accurately predicting time series data, particularly within the context of a submerged membrane filtration system. The ability of LSTMs to capture long-term dependencies and model complex dynamics makes them suitable for accurately forecasting the permeate flux over time in this specific application.

2. Modeling

2.1. Long short-term memory (LSTM)

Long short-term memory (LSTM) is a type of recurrent neural network (RNN) that specifically designed for time-series problem and their long-range dependencies more accurately than conventional RNNs [25]. Figure 1(a) shows the architecture of RNN network. An RNN can be viewed as the design of a normal feedforward MLP network with loops added. Recurrent neural networks have cycles that feed previous time step activations into the network as inputs to influence predictions at the current time step. The network's internal states, which are theoretically capable of storing long-term temporal contextual information, store these activations. As a result of this process, RNNs can take use of a contextual window that changes dynamically over the course of the input sequence history [25].

A LSTM memory cell contains one tanh layer in addition to three sigmoid layers. The input, forget, and output gates of the LSTM are used to maintain and refresh the memory cells while filtering out extraneous data. The forget gate first chooses whether data should be kept or removed from the model. The input gate also manages the reserve of data on the current cycle input. The input gate has two responsibilities. Finding the state of the cell that needs to be updated is the first task; the sigmoid function chooses the value that needs to be updated. The second duty is updating the data to reflect the current condition of the cell. The output gate is the final gate. The following hidden state is determined by the output gate.

2.2. Vanilla LSTM

https://doi.org/10.52202/069564-0135

A vanilla LSTM is a most basic LSTM model as it has a straightforward LSTM configuration as shown in Figure 1(b). LSTM have a input layer and output layer that directly connected to single hidden layer of LSTM memory cells to create predictions [26]. The LSTM architecture described in the original LSTM study from 1997 is the one that will perform well on the majority of minor sequence prediction challenges. Contrary to its extensions, vanilla LSTM still performs brilliantly on a range of datasets more than 20 years after its launch, according to Greff and K. et al. [27]. Nelson et al. [28] used Vanilla LSTM to predict stock prices for the first time, and they were encouraged by the findings.

2.3 Stacked LSTM

Stacked LSTM were introduced by Graves et al. [29], in their application of LSTMs to speech recognition, beating a benchmark on a challenging standard problem. In the same work, it was found that the depth of the network was more crucial to accurately modelling rather than the number of memory cell. A stacked LSTM architecture is made up from multiple LSTM layers. Instead of sending a single value to the LSTM layer below, an LSTM layer above sends a sequence of values. Instead of having one output time step for all input time steps, specifically, one output per input time step [30]. Figure 1© shows the architecture of stacked LSTM.

2.4 Bidirectional LSTM

In order to maximise the use of the input sequence, bidirectional LSTMs walk through input time steps in both the forward and backward directions [31]. In order to implement this design, the first recurrent layer of the network is duplicated so that there are now two layers side by side. The input sequence is then provided as-is to the first layer as input, and a reversed copy is provided to the second layer. This method was created in the past as a broad method for improving the effectiveness of recurrent neural networks (RNNs). Figure 1(d) shows the architecture of bidirectional of LSTM network.

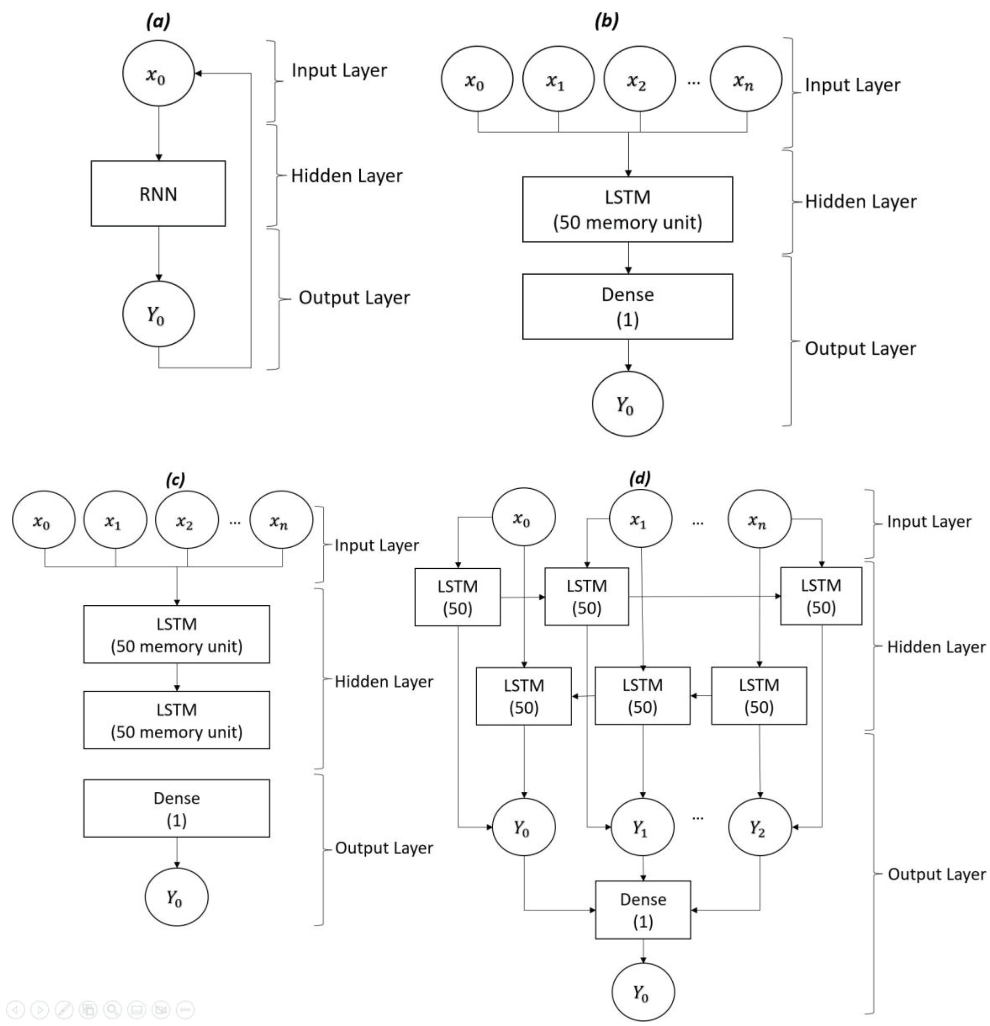

Figure. 1. Architecture of: (a) RNN (b) Vanilla LSTM (c) Stacked LSTM (d) Bidirectional LSTM.

https://doi.org/10.52202/069564-0135

3. Methods

3.1. Pilot source water

A schematic diagram of submerged membrane bioreactor (SMBR) with four configurations of inlet stream, permeate stream, aeration and backwashing stream is illustrated in Figure 2. The permeate stream used diluted palm oil mill effluent (POME) collected from Mahamurni Plantation palm oil Sdn. Bhd., Wastewater treatment plant, located in Sedenak Johor Bahru as its source. The plant comprises of a single bioreactor tank that has been fitted inside with submerged hollow fibre and Polyethersulfone (PES) material is used to construct the hollow fibre membrane.

The POME was supplied into the 20 L bioreactor tank through the inlet stream. When the level reaches the specified level, the influent supply to the bioreactor will be stopped off by a mechanical level controller. The filtered influent is then pumped to the effluent tank down the pipe using a peristaltic pump (P-101). During permeate stream, the P-101 is activated to open the valve (SV-101). Electronic flow sensors are used to detect the permeate flux flow rate (FM-101). This part employs an electronic pressure sensor to monitor the transmembrane pressure (TMP) (P-101).

In order to study the long-term effects on transmembrane pressure (TMP) and membrane flux, the membranes were cleaned using just physical cleaning which are backwashing and aeration stream. The membrane was cleansed from the inside out in the backwash stream. When pump (P-102) is activated and valve (SV-102) is opened, the backwash procedure is complete. The membrane was cleaned by the backwash using filtered tap water that was kept in the backwash tank. The air that was pumped into the bioreactor using an air compressor (C-101). Electronic flow sensors are used to measure the air velocity under proportional valve (PV-101) control (FA-101). To produce bubbles, the air from the compressor is sent through the air diffuser.

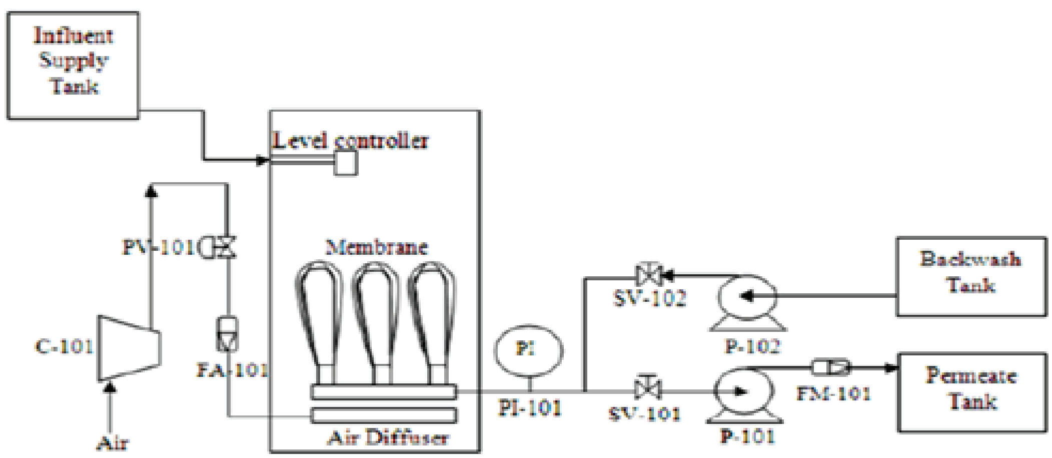

Figure. 2. Schematic diagram of submerged membrane pilot plant.

The pilot plant system was equipped with full monitoring and control using a data acquisition of National Instrument. Supervisory Control and Data Acquisition (SCADA) software is a type of graphical language-based programme that can be used with LabVIEW. For the aim of modelling, the software was created to record the collected data in the pilot plant. In this research, POME is used to generate fouling in the MBR filtration process. Figure 3 shows the actual sample used in the filtration process.

https://doi.org/10.52202/069564-0135

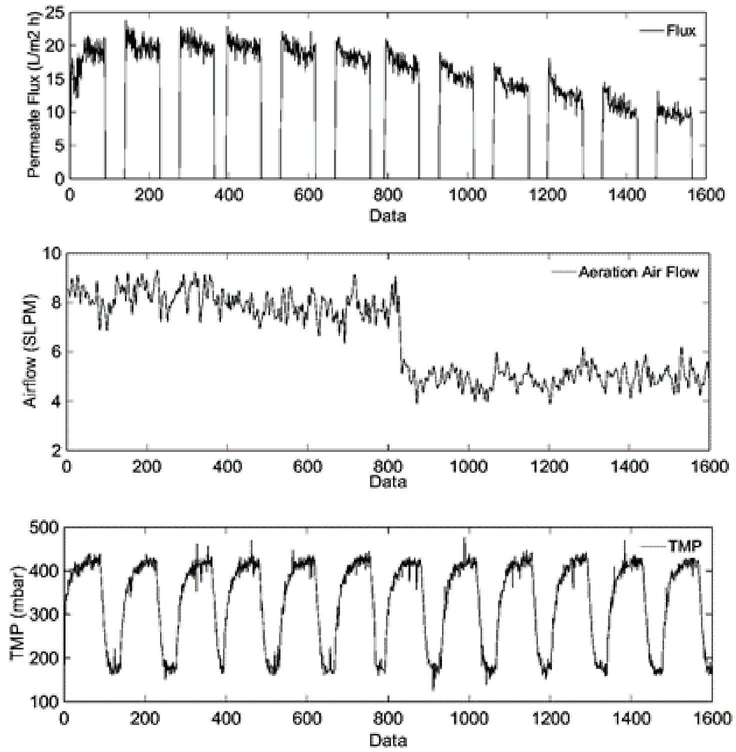

Figure. 3. Samples of the collected data using submerged membrane bioreactor plant.

3.2. Data pre-processing

Data pre-processing involved normalization or standardization of the data before they were used to train models. The normalization data is a method of scaled down the collected data that in the original range, so all values fall between 0 and 1. Meanwhile, standardize means rescaling the distribution values to have a mean of 0 and a standard deviation of 1. This can be thought of as subtracting the mean value or centering the data. Like normalization, standardization can be useful and required in machine learning algorithm when the data contains input values with different scale. Standardization makes the underlying assumption that your data has a Gaussian distribution (bell curve) with a well-behaved mean and standard deviation. If this assumption is not satisfied, an accurate result might not be able to achieve. The collected data was normalise using scikit-learn object MinMaxScaler and standardization can both be achieved using scikit-learn machine learning library in Python programming [32].

3.3. Transform time series data

A key function in Pandas library that is crucial tool for converting time series data into a supervised problems is Pandas shift () function. The shift() function can be used to make copies of columns that are pushed forward (rows of NaN values are added to the front) or pulled back (rows of NaN values added to the end). This is to generate columns of lag observations as well as columns of prediction observations. In terms of time series forecasting, forecasts are made using prior observations (t-1, t-n) and the current time (t), as well as future times (t+1, t+n). We can see how a time series containing sequences of input and output patterns can be utilised to generate a new DataFrame for a supervised learning task [32].

4. Result and discussion

4.1 Model performance

In this study, three different architecture of LSTM models such as Vanilla LSTM, Stacked LSTM and Bidirectional LSTM including RNN were developed to predict the permeate flux in submerged membrane bioreactor. During training these models, two loss function was used which are mean absolute error (MAE) and mean square error (MSE) to evaluate the performance of the models as it is suitable and recommended for time series problems. In this work, 100 epoch and 50 batch size were used. An epoch is an iteration over the entire training dataset. The model was iterating over the training data, multiple times until it achieves 100 epochs and updating the model’s parameters at each time. The accuracy metric will be reported at each training epoch to reflect the performance of the model in addition to the loss function. The dataset consists of 4021 samples, and it were divided into 60/40 for each training and testing dataset. Figure 4 shows the loss function graph of RNN, Vanilla LSTM, Stacked LSTM and Bidirectional LSTM models that were iterates for 100 epochs. From the graph, we can clearly see that the models learned the problem by achieving towards zero error. A train (blue) and test (orange) line is created showing the MSE over the training epochs converged up to three decimal places which is a good sign for training models. The performance of the model suggests that the MAE and MSE error is a good match for a deep learning neural network LSTM's model as compared to RNN model which achieved MAE slightly lower than the others LSTM's model.

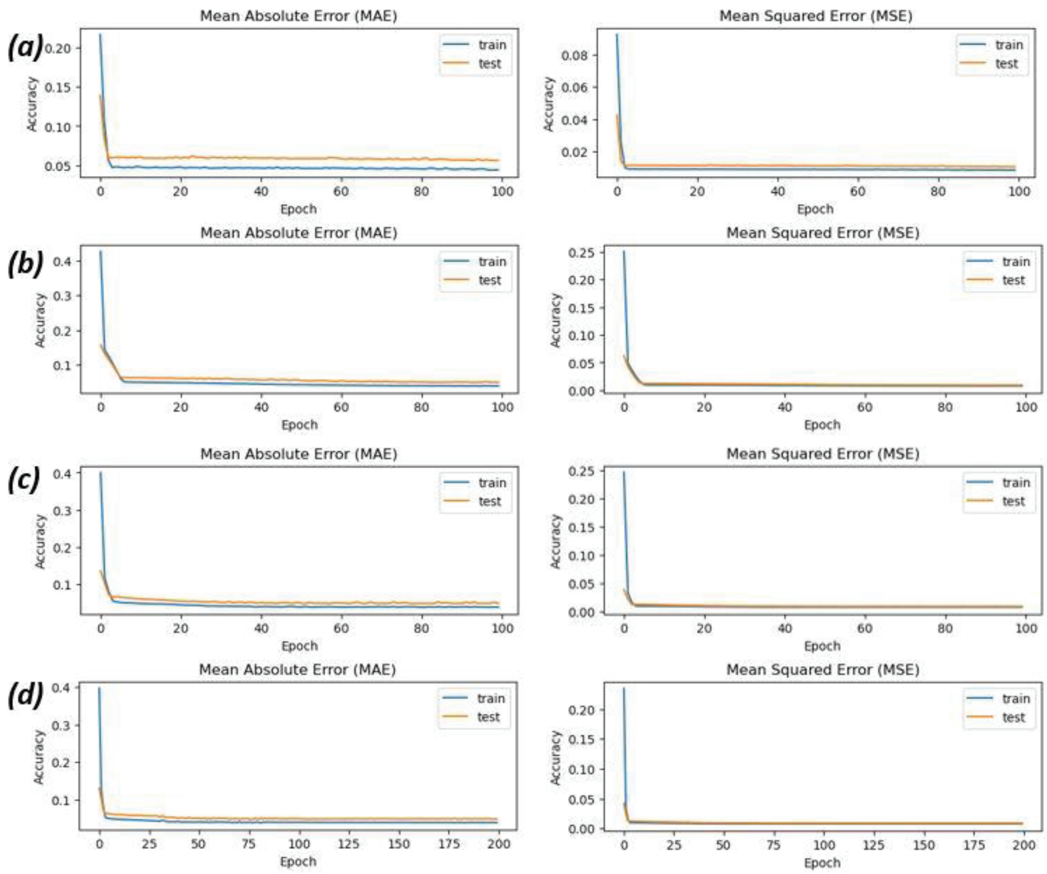

Figure. 4. (a) RNN, (b) Vanilla LSTM (c) Stacked LSTM (d) Bidirectional LSTM

After training using 60 % of dataset, the rest 40 % of the samples were used to predict the models. Figure 5 shows the plotted graph of training and testing dataset for RNN model. RNN model has the simplest architecture with 50 memory units, and it takes inputs from previous time steps direct to the output. From the training graph, RNN are able to forecast the training data set with 92.4 %. Meanwhile, the testing dataset only able to forecast the testing dataset with 89.0 %. Figure 6 shows the graph of vanilla LSTM. The architecture of Vanilla LSTM used 50 memory unit with 1 dense and batch size of 32. For training and testing dataset, it achieves 93.2% and 90.6 % respectively. As compared to RNN, Vanilla LSTM are able to improve the accuracy performance better when predicting the unknown model. Figure 7 shows the training and testing graph of stacked LSTM. For stacked LSTM, 50 memory unit were stacked with another 100 memory unit and 1 dense for its architecture. The results shows that 93.2% were achieve during training plot while 90.7% during the testing plot. The same result was obtained for bidirectional LSTM for both training and testing graph as shown in Figure 8. The overall accuracy performance results were also evaluated using root mean square error (RMSE) and it was tabulated in Table 1.

https://doi.org/10.52202/069564-0135

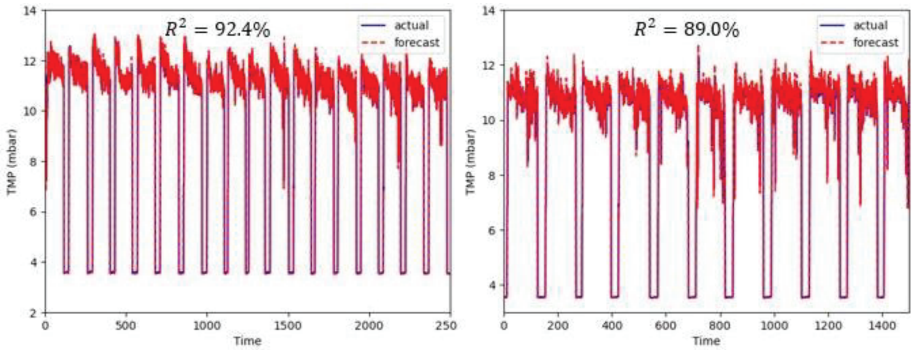

Figure. 5. The graph of training and testing dataset for RNN model.

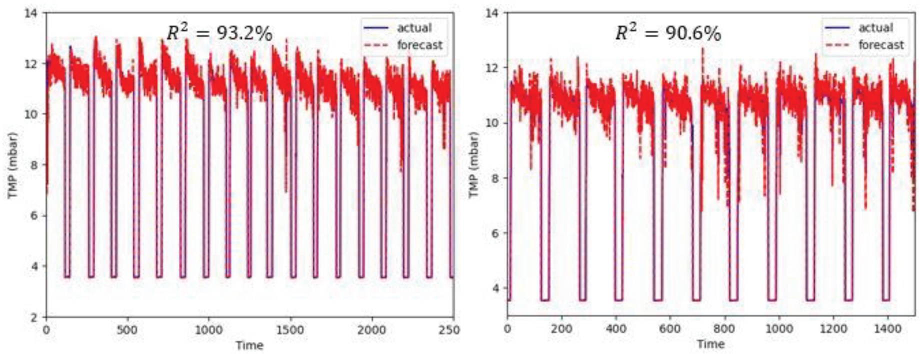

Figure. 6. The graph of training and testing dataset for Vanilla LSTM model.

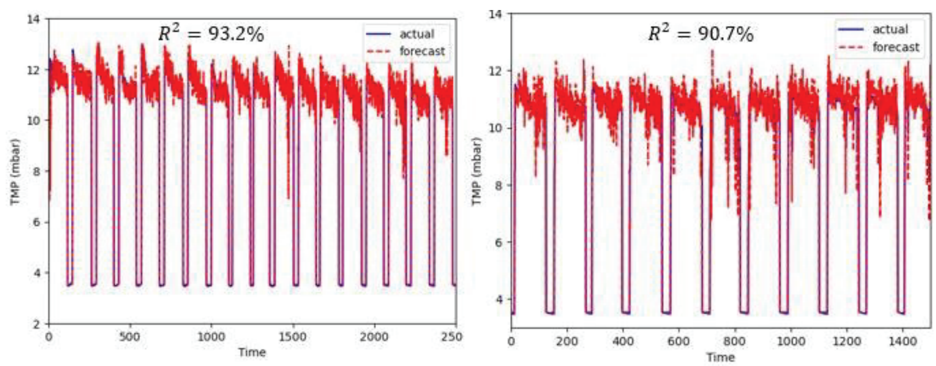

Figure. 7. The graph of training and testing dataset for Stacked LSTM model.

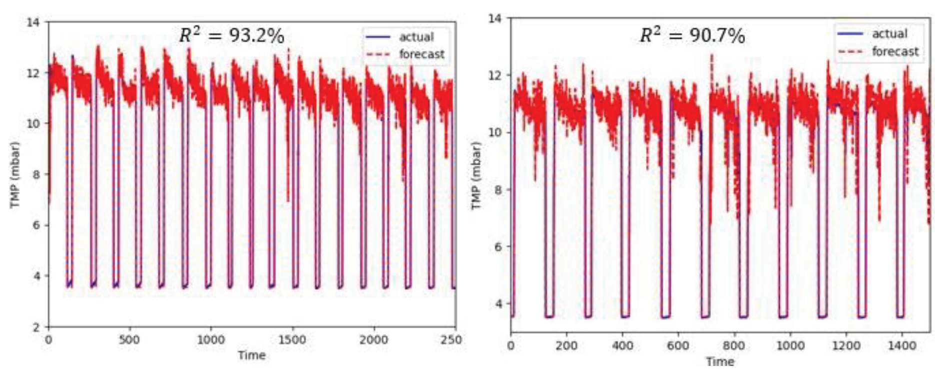

Figure. 8. The graph of training and testing dataset for Bidirectional LSTM model.

Table 1. Overall performance results of the models.

Model	Accuracy	Training	Testing
Vanilla LSTM	MSE	0.679	0.818
Epoch: 100	$RMSE$	0.824	0.905
Batch Size: 32	MAE	0.360	0.465
	R^2	0.932	0.906
Stacked LSTM	MSE	0.683	0.808
Epoch: 100	$RMSE$	0.826	0.899
Batch Size:32	MAE	0.363	0.466
	R^2	0.932	0.907
Bi-Directional LSTM	MSE	0.676	0.806
Epoch: 200	$RMSE$	0.822	0.898
Batch Size: 64	MAE	0.367	0.467
	R^2	0.932	0.907
RNN	MSE	0.762	0.955
Epoch: 100	$RMSE$	0.873	0.977
Batch Size: 32	MAE	0.414	0.534
	R^2	0.924	0.890

4.2 Conclusion

The good prediction model of submerged membrane filtration process can be a good useful in order to analysing backwashing strategies and helps to gaining some understanding on the fouling phenomena. The main objective of the study is to compare the performance of two popular neural network architectures: RNN (Recurrent Neural Network) and LSTM (Long Short-Term Memory). Specifically, the researchers explored different types of LSTM models, including Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM, to assess their effectiveness in comparison to RNN. Upon conducting the experiments, the researchers found that all three types of LSTM models Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM yielded similar results. Despite their architectural differences, these variations of LSTM exhibited comparable performance, indicating that the core LSTM mechanisms were effective in capturing the underlying patterns in the data. However, when comparing these LSTM models with the RNN architecture, it was evident that the different types of LSTM consistently outperformed RNN. This suggests that LSTM's specialized design, which incorporates memory cells and gates to selectively retain and forget information, offers distinct advantages over the basic RNN architecture.

The improved performance of LSTM models can be attributed to their ability to capture and remember long-term dependencies in sequential data, thereby mitigating the vanishing gradient problem that hampers RNN performance. Vanilla LSTM, Stacked LSTM (which includes multiple LSTM layers), and Bidirectional LSTM (which processes input sequences in both forward and backward directions) all leverage these capabilities to achieve better results compared to RNN. These findings underscore the significance of employing LSTM architectures, in their various forms, for tasks involving sequential data or temporal dependencies. By utilizing the memory cells and gating mechanisms, LSTM models are better equipped to handle complex patterns and long-range dependencies in the data, leading to improved performance and more accurate predictions.

The performance of LSTM models can be further enhanced through various optimization strategies, including tuning the number of epochs, batch size, memory units, and dense architecture network. These adjustments aim to improve the accuracy and reliability of predictions. Additionally, the effectiveness of LSTM can be further amplified by integrating optimization techniques such as particle swarm optimization (PSO) or genetic algorithms (GA). These optimization methods assist in identifying the optimal architecture for the LSTM network, ensuring its efficiency and effectiveness in capturing complex patterns and delivering robust predictions. By leveraging these optimization approaches, LSTM models can be fine-tuned to achieve superior performance and more accurate results.

Acknowledgments

This work was financially supported by the Universiti Teknologi Malaysia High Impact Research Grant (UTMHR) vote Q.J130000.2451.08G74. Additionally, this work has received support from the Catalan Government under Project 2021 SGR 00197 and the Spanish Government under MICINN project PID2019-105434RBC33, which is co-funded with the European Union ERDF funds.

Nomenclature

MSE mean square error

RMSE root mean square error

$\dot{MAE}$ mean absolute error

R^2 correlation coefficient

References

[1] S. Vinardell *et al.*, "Advances in anaerobic membrane bioreactor technology for municipal wastewater treatment: A 2020 updated review," *Renewable and Sustainable Energy Reviews,* vol. 130, p. 109936, 2020.

[2] J. Ma, R. Dai, M. Chen, S. J. Khan, and Z. Wang, "Applications of membrane bioreactors for water reclamation: micropollutant removal, mechanisms and perspectives," *Bioresource technology,* vol. 269, pp. 532-543, 2018.

[3] S. Güneş-Durak, A. Ciggin, and N. Tüfekci, "Fabrication, characterization and treatment of polymeric membranes with submerged membrane bioreactor system: fruit juice industry wastewater," *International Journal of Environmental Science and Technology,* pp. 1-14, 2022.

[4] F. A. Fraga, H. A. García, C. M. Hooijmans, D. Míguez, and D. Brdjanovic, "Evaluation of a membrane bioreactor on dairy wastewater treatment and reuse in Uruguay," *International Biodeterioration & Biodegradation,* vol. 119, pp. 552-564, 2017.

[5] N. Wung, S. M. Acott, D. Tosh, and M. J. Ellis, "Hollow fibre membrane bioreactors for tissue engineering applications," *Biotechnology letters,* vol. 36, pp. 2357-2366, 2014.

[6] X. Shi, G. Tal, N. P. Hankins, and V. Gitis, "Fouling and cleaning of ultrafiltration membranes: A review," *Journal of Water Process Engineering,* vol. 1, pp. 121-138, 2014.

[7] M. B. Asif and Z. Zhang, "Ceramic membrane technology for water and wastewater treatment: A critical review of performance, full-scale applications, membrane fouling and prospects," *Chemical Engineering Journal,* vol. 418, p. 129481, 2021.

[8] W. Guo, H.-H. Ngo, and J. Li, "A mini-review on membrane fouling," *Bioresource technology,* vol. 122, pp. 27-34, 2012.

[9] Q. Li and M. Elimelech, "Natural organic matter fouling and chemical cleaning of nanofiltration membranes," *Water Science and Technology: Water Supply,* vol. 4, no. 5-6, pp. 245-251, 2004.

[10] N. Delouche, B. Dersoir, A. Schofield, and H. Tabuteau, "Flow decline during pore clogging by colloidal particles," *Physical Review Fluids,* vol. 7, no. 3, p. 034304, 2022.

[11] G. Crozes, J. Jacangelo, C. Anselme, and J. Laine, "Impact of ultrafiltration operating conditions on membrane irreversible fouling," *Journal of Membrane Science,* vol. 124, no. 1, pp. 63-76, 1997.

[12] Z. Cui *et al.*, "Investigation of backwashing effectiveness in membrane bioreactor (MBR) based on different membrane fouling stages," *Bioresource technology,* vol. 269, pp. 355-362, 2018.

[13] K. B. Newhart, R. W. Holloway, A. S. Hering, and T. Y. Cath, "Data-driven performance analyses of wastewater treatment plants: A review," *Water research,* vol. 157, pp. 498-513, 2019.

[14] Y. Xie *et al.*, "Enhancing real-time prediction of effluent water quality of wastewater treatment plant based on improved feedforward neural network coupled with optimization algorithm," *Water,* vol. 14, no. 7, p. 1053, 2022.

[15] R. Huang, C. Ma, J. Ma, X. Huangfu, and Q. He, "Machine learning in natural and engineered water systems," *Water Research,* vol. 205, p. 117666, 2021.

[16] L. Zhao, T. Dai, Z. Qiao, P. Sun, J. Hao, and Y. Yang, "Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse," *Process Safety and Environmental Protection,* vol. 133, pp. 169-182, 2020.

[17] X. Meng, Y. Zhang, and J. Qiao, "An adaptive task-oriented RBF network for key water quality parameters prediction in wastewater treatment process," *Neural Computing and Applications,* pp. 1-14, 2021.

[18] M. Zeinolabedini and M. Najafzadeh, "Comparative study of different wavelet-based neural network models to predict sewage sludge quantity in wastewater treatment plant," *Environmental monitoring and assessment,* vol. 191, no. 3, p. 163, 2019.

[19] G. Wang, Q.-S. Jia, M. Zhou, J. Bi, J. Qiao, and A. Abusorrah, "Artificial neural networks for water quality soft-sensing in wastewater treatment: a review," *Artificial Intelligence Review,* vol. 55, no. 1, pp. 565-587, 2022.

[20] D. J. Kovacs *et al.*, "Membrane fouling prediction and uncertainty analysis using machine learning: A wastewater treatment plant case study," *Journal of Membrane Science,* vol. 660, p. 120817, 2022.

[21] F. Harrou, T. Cheng, Y. Sun, T. Leiknes, and N. Ghaffour, "A data-driven soft sensor to forecast energy consumption in wastewater treatment plants: A case study," *IEEE Sensors Journal,* vol. 21, no. 4, pp. 4908-4917, 2020.

[22] S. A. Khadem and A. D. Rey, "Nucleation and growth of cholesteric collagen tactoids: A time-series statistical analysis based on integration of direct numerical simulation (DNS) and long short-term memory recurrent neural network (LSTM-RNN)," *Journal of Colloid and Interface Science,* vol. 582, pp. 859-873, 2021.

[23] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter (PM2. 5) forecasting in smart cities," *Sensors,* vol. 18, no. 7, p. 2220, 2018.
[24] J. Shim, S. Park, and K. H. Cho, "Deep learning model for simulating influence of natural organic matter in nanofiltration," *Water Research,* vol. 197, p. 117070, 2021.
[25] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," 2014.
[26] J. Brownlee, *Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras*. Machine Learning Mastery, 2016.
[27] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey (2015)," *arXiv preprint arXiv:1503.04069,* 2016.
[28] D. M. Nelson, A. C. Pereira, and A. Renato, "de Oliveira. 2017. Stock market's price movement prediction with LSTM neural networks," in *International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, May*, pp. 14-19.
[29] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional LSTM," in *2013 IEEE workshop on automatic speech recognition and understanding*, 2013, pp. 273-278: IEEE.
[30] M. V. Sebt, S. Ghasemi, and S. Mehrkian, "Predicting the number of customer transactions using stacked LSTM recurrent neural networks," *Social Network Analysis and Mining,* vol. 11, pp. 1-13, 2021.
[31] J. Brownlee, *Long short-term memory networks with python: develop sequence prediction models with deep learning*. Machine Learning Mastery, 2017.
[32] "Ultrafiltration Membranes: Technologies and Global Markets," BCC PublishingJuly 2020 2020.

 https://doi.org/10.52202/069564-0135