Deep Learning Approach for Forecasting Water Quality in IoT Systems

Nguyen Thai-Nghe[1]
College of ICT
Can Tho University
Can Tho City, Vietnam

Nguyen Thanh-Hai[2]
College of ICT
Can Tho University
Can Tho City, Vietnam

Nguyen Chi Ngon[3]
College of Engineering Technology
Can Tho University
Can Tho City, Vietnam

Abstract—**Global climate change and water pollution effects have caused many problems to the farmers in fish/shrimp raising, for example, the shrimps/fishes had early died before harvest. How to monitor and manage quality of the water to help the farmers tackling this problem is very necessary. Water quality monitoring is important when developing IoT systems, especially for aquaculture and fisheries. By monitoring the real-time sensor data indicators (such as indicators of salinity, temperature, pH, and dissolved oxygen - DO) and forecasting them to get early warning, we can manage the quality of the water, thus collecting both quality and quantity in shrimp/fish raising. In this work, we introduce an architecture with a forecasting model for the IoT systems to monitor water quality in aquaculture and fisheries. Since these indicators are collected every day, they becomes sequential/time series data, we propose to use deep learning with Long-Short Term Memory (LSTM) algorithm for forecasting these indicators. Experimental results on several data sets show that the proposed approach works well and can be applied for the real systems.**

Keywords—*Forecasting model; deep learning; Long-Short Term Memory (LSTM); water quality indicators*

I. INTRODUCTION

Global climate change and Global Warming effects have caused many problems to the farmers and the managers. Climate change not only causes great impacts on all regions, resources, environment and socio-economic, but also affects water resources, agriculture and rural development, health and the coastal areas. Economic losses due to the impact of climate change combined with the costs of recovering from the damage reduce the economic growth of many countries and globally. Especially, Agricultural/Aquaculture countries have been suffering great losses because of Global climate change. For example, fishery is one of the key economic sectors of Vietnam's agriculture, with billions of dollars in exports. It also creates jobs for millions of workers, contributing to stabilizing social security and developing the country. Shrimp is a seafood product in the top export of billion US dollars. However, shrimp nursery is too difficult because they are particularly sensitive to water. Polluted water environment is a great concern for numerous fishermen and Agricultural managers. However, these farmers do not have enough technique and media to forecast environment factors, so they cannot manage risks which they are facing. Thus, how to manage and monitor the water quality, especially in aquaculture and fisheries such as fish raising management is really necessary.

Recently, techniques in the Internet of Things (IoT) can help users to build the monitoring systems. IoT refers to the millions or even billions of physical devices around the world currently connected to the internet, collecting and sharing data together. With advancements in internal processors and wireless networks, we now can make everything more proactive and intelligent [1]. There are many applications of IoT such as automatic door systems (smart-home), aircraft and self-driving cars (smart-city), etc. which have become a common part of IoT. This adds a level of digital intelligence to passive devices, allowing us to communicate real-time data without human involvement, effectively integrating the digital and physical world. When something is connected to the internet, it is able to send or receive information, or both. Leveraging IoT techniques, the ability to send or receive this information makes everything smarter, and that is also our target.

This work[1] proposes an architecture for the IoT systems to monitor and forecast the water quality (such as indicators of salinity, temperature, pH, and dissolved oxygen - DO) in aquaculture and fisheries. These data are collected every day and ordered by time, they can be considered as time series or sequential data. Previous works show that deep learning with LSTM (Long-Short Term Memory) could be an appropriate method for time series data [2], thus, we propose deep learning with LSTM models for forecasting these water quality indicators.

The remaining of this study is presented as follows. We exhibit some state-of-the-art approaches in sensor data processing in Section II. Section III introduces the proposed architectures for the forecast problems. In Section IV, we present the data sets used in the experiments and do some investigations, comparisons, explain the experimental results. Finally, we discuss and summarize our work in Section V.

II. RELATED WORK

A complete IoT system relates to several sides such as sensor networks, hardware circuits, energy, etc., and software for analysis. One of important tasks in the environmental monitoring IoT system is the data analyzing and forecasting methods. In this work, we focus on reviewing the literature which relates to the forecasting methods. There are several researches in this field.

The authors in [3] introduced a method to develop a service oriented architecture for the weather information systems which forecast weather using these data mining techniques

---

[1]Corresponding author: Nguyen Chi Ngon

(e.g., Artificial Neural Network and Decision tree). The study in [4] proposed a deep neural network for weather forecasting with a set of weather-related variables. The study in [5] used recurrent neural networks (RNN) with its Long Short-Term Memory (LSTM) architecture to predict the ambient temperature (TA). The prediction is based on meteorological data (e.g., temperature, humidity) retrieved from IoT stations and send them to the basic station with LoRa protocol. They formulate the TA prediction problem as a time series regression problem. The proposed approach uses two types of hidden layers: LSTM layer and full connected dense layer. Experimental results show that the model work well.

In [6], the water, temperature, pH and DO levels are measured and integrated with aerating and water supply pumps using Arduino. The user could receive information at predetermined intervals on preferred communication. The authors tested on a sample of two days measurements of temperature, pH and DO levels. Results show that with this integration system, farmer need not hire worker at their site, consequently drive down operating costs and improve efficiency.

The authors in [7] introduced an online water quality monitoring system for intensive fish culture in China. Based on historical data, the system forecasts the water quality using artificial neural networks (ANNs) and control the water quality in time to reduce catastrophic losses. Their results demonstrate that multi-parametric, long-distance and online monitoring for water quality information can be accurately acquired and predicted.

The authors in [9] proposed the detection and monitoring of emotion applications. The authors introduce a deep learning approach for emotion classification through an iterative process by adding and removing large number of sensor signals from different modalities. They uses Convolutional Neural Network and Long Short-term Memory Recurrent Neural Network on the raw sensor data. The results show that their proposed approach is effective in human emotion classification when large number of sensors input is utilised and the hybrid models outperform traditional fully connected deep neural network.

In [10], the authors proposed a deep learning and image-based model for air quality estimation. The model extracts feature information from scene images captured by camera equipment and then classifies them to estimate air quality levels. The authors tested on several methods such as Support Vector Machines and Deep Residual Network on a high-quality outdoor air quality data set. Experimental results show that the proposed method produces more accurate results than other methods.

The authors in [11] used deep-learning technologies in the environmental field to predict the status of pro-environmental consumption. The authors predicted the pro-environmental consumption index based on Google search query data, using a recurrent neural network (RNN) model. They compared the prediction accuracy of the RNN model with that of the ordinary least square and artificial neural network models. The RNN model predicts the pro-environmental consumption index better than any other model.

In [12], the authors presented the Public agencies aiming to enforce environmental regulation have limited resources to achieve their objectives. They demonstrate how machine-learning methods can inform the efficient use of these limited resources while accounting for real-world concerns, such as gaming the system and institutional constraints. They predict the likelihood of a facility failing a water-pollution inspection and propose alternative inspection allocations that would target high-risk facilities.

The authors of [13] developed a web-based system which monitor the pH and salinity values. The workings of the monitoring and control system send data on pH and salinity sent by the sensor and then stored and in database hosting. They said that for the value of a dangerous salt content worth less than 160 and more than 210, if for a pH of less than 6.5 and more than 7.5.

Darmalim et al. in [14] proposed an IoT solution to automatically monitor the environmental factors. The system had five sensors to measure each parameter. It is developed using a Pythonframework with a web application to present information from the IoT device. The work in [16] presented a feasible model for the daily average temperature on the area of Zhengzhou and apply it to weather derivatives pricing. Other researches can be found in [8], [17], [18], [19], [20].

Although many researches have been done in the topic of environmental forecasting, using LSTM for forecasting water quality indicators is not seen recently and this problem still has room for improvements. In this work, we propose an architecture for the IoT systems to monitor water quality (such as indicators of salinity, temperature, pH, and DO). Since these data are collected every day (ordered by time), we propose deep learning with Long-Short Term Memory (LSTM) models for forecasting these time series data.

III. PROPOSED METHOD

A. System Architecture

The system architecture is presented in Fig. 1. In this system, we have implemented the sensors for monitoring water quality indicators such as salinity, temperature, pH, and dissolved oxygen (DO) in the fish/shrimp ponds. Data from sensors is daily transferred to a cloud database (in this work, we have used the cloud database provided by the DigitalOcean[2]). This data is automatically transferred to a server for analyzing and forecasting.

Using this system, the users can visualize all of the indicators via mobile devices or pc/laptop. The system automatically sends the messages to the users (farmers) when the value of an indicator exceeds the limit (threshold) value. Moreover, based on historical data, the system can forecast the values of each indicator for the next dates (or other selected time).

The system infrastructure is implemented by another group in our team. In this paper, our task is to develop a forecasting model which can provide the forecast values to send early warning to the users. Thus, in the rest of this work, we focus on developing the forecasting models, especially Deep Learning with LSTM for forecasting time series data.

---

[2]https://www.digitalocean.com

Fig. 1. System Architecture for Shrimp/Fish Pond Monitoring

B. The Proposed LSTM Architectures

LSTM (Long Short-Term Memory) is well-known deep learning method which widely-used in forecast models and can be used to avoid the long-term dependency problem [2], [15].

We propose the LSTM architecture as presented in Fig. 2. For testing purposes, the proposed model includes a single LSTM layer with four units followed by a Fully Connected layer using ReLU as the activation function (the leftside of Fig. 2) and another network containing two recurrent hidden states with 16 units for the first LSTM layer and 4 units for the second LSTM layer which we call stacked LSTM (the rightside of Fig. 2). The input to these models are a sequence values (based on timestamp) of each indicator (e.g., salinity, pH, DO, and Temperature) and the output are their forecast values. We implemented the LSTM with Adam optimizer [21] as the optimized function. The learning rate used is initiated at 0.001. In the experiments (will be presented in Section IV-C) we will validate to choose which LSTM model is appropriate for the kind of sensor data.

When we look inside the LSTM (shown in Fig. 3 as an example), its major and important part of a LSTM network is cells that provide a bit of memory to the LSTM so it can remember the past. LSTM usually contains three types of gates including Input Gate, Output Gate and Forget Gate. Gates in LSTM consist of sigmoid activation functions which produce output ranging from 0 to 1. In most of the cases, the value is either 0 or 1. The sigmoid function (Equation 1) is used commonly for gates because we want a gate to give only positive values and should be able to give us a clear cut answer whether, we need to keep a particular feature or we need to discard that feature.

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \qquad (1)$$

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \qquad (2)$$
$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \qquad (3)$$
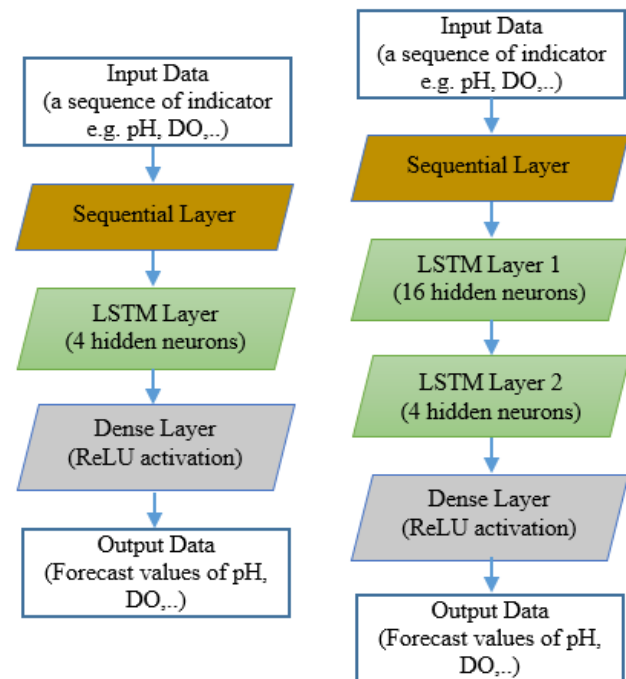$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \qquad (4)$$

Where:

Fig. 2. The Proposed Single LSTM (Left Side) and Stacked LSTM (Right Side) Models
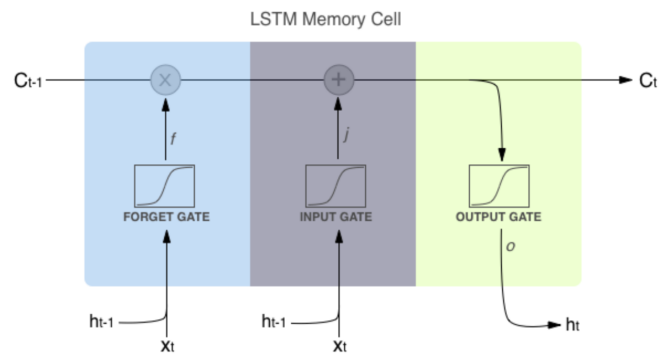

Fig. 3. An Illustration of LSTM Cell [15].

- i_t denotes input gate.

- f_t reveals output gate.

- o_t represents forget gate.

- w_m shows weights for the respective $gate(m)$.

- h_{t-1} denotes the output of the preceded LSTM block at the timestamp t-1.

- x_t is the input at the current timestamp.

- b_m exhibits the biases of the respective $gate(m)$.

Equation 2 reveals what new information will be stored in the cell state while Equation 4 aims to provide the activation to the final output of the LSTM block at timestamp t. The forget gate at Equation 3 give us the information on which information to throw away from the cell state.

IV. EXPERIMENTAL RESULTS

A. Data Sets for the Experiments

In the proposed architecture in Fig. 1, the sensors' implementation is deployed by another group in our team. However, this system is just started and we do not have enough data for training the model. Luckily, there are several published data sets (presented in Table I) which have the structures similar to our data structure, so we have used them to test the proposed forecasting approach.

The first data set is published by Australian Government, which is the sensor data of Temperature, pH and salinity at New South Wales Estuary.

The second data set is published by the BCO-DMO, located at the Woods Hole Oceanographic Institution, US. This data set relates to water quality parameters (e.g., pH, salinity, temperature, and dissolved oxygen - DO) in Tomales Bay, CA (38.148N, 122.898W) which is a monthly time series through sampling every 2 km.

The third data set is published by the University of Melbourne, Australia. This data set relates to an IoT system which provides the Sensor data including of temperature, light level, humidity every 5 minutes at 8 locations from 2014 to 2015.

The fourth and fifth data sets are published by the Government of Canada. These data relate to the Surface temperature and salinity measured along the track of commercial ships, mostly between Montreal (Quebec) and St. John's (Newfoundland) from 1999 to 2018. We have just used the data of 2017 and 2018 for testing.

The sixth data set is published by the European Multidisciplinary Seafloor and water column Observatory. This dataset contains temperature and dissolved oxygen concentrations and the associated sensor raw data acquired between July 2017 and August 2018 on EMSO-Azores observatory by the EGIM.

After pre-processing by removing missing data, the remaining number of observations are presented in the column #Records in Table II.

TABLE I. PUBLISHED DATA SET RESOURCES

No.	Data set name	Link (accessed July, 2020)
1	NSW Estuary Temperature, pH and salinity data	https://data.gov.au/dataset/ds-nsw-c2041218-0496-453a-ac22-9e974fdebf89/distribution/dist-nsw-5154e497-647f-4e2a-a2ad-151fa293b0da/details?q=
2	Water quality - Tomales Bay, CA	https://www.bco-dmo.org/dataset-deployment/455476
3	Melbourne - Sensor readings	https://data.melbourne.vic.gov.au/Environment/Sensor-readings-with-temperature-light-humidity-ev/ez6b-syvw
4	Fisheries & Oceans Canada 2018	https://open.canada.ca/data/en/dataset/8a3dc9e5-f3af-4270-8c09-43fa2c25848b
5	Fisheries & Oceans Canada 2017	https://open.canada.ca/data/en/dataset/8a3dc9e5-f3af-4270-8c09-43fa2c25848b
6	Temperature and DO from the EGIM, EMSO-Azores observatory, 2017-2018	https://www.seanoe.org/data/00453/56501

B. Evaluation Metrics and Baseline

Root mean squared error (RMSE) is used to evaluate the models. The RMSE for n observations is calculated by equation (5).

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - \hat{p}_i)^2} \qquad (5)$$

where, p_i is the true value, and $\hat{p}_i$ is the forecasted value.

We have used regression version of the Support Vector Machines (SVM) as a baseline for comparison.
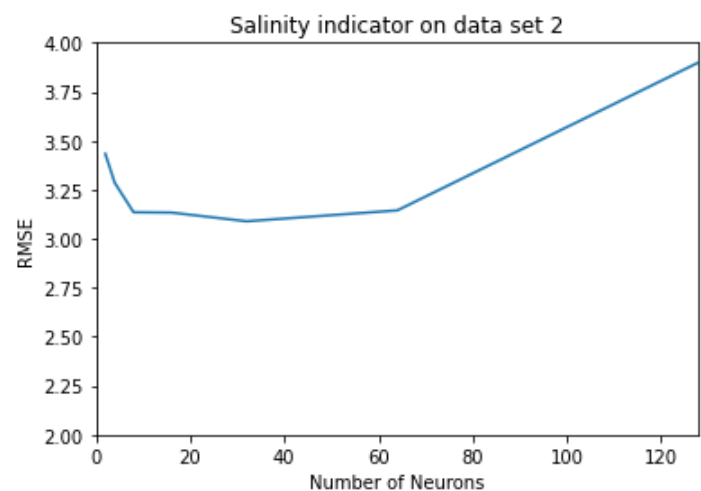
C. Experimental Results

Fig. 4. Relationship between the RMSE and the Number of Neurons (Salinity Indicator on Data Set 2) using LSTM. Other Data Sets are nearly Similar

In the first group of experiments, a single layer of the LSTM (model in the left side of Fig. 2) is performed to select the number of hidden neurons. Results show that, on these data sets, the RMSE is not much different when using the number of hidden neurons from 4 to 60. An example is presented in

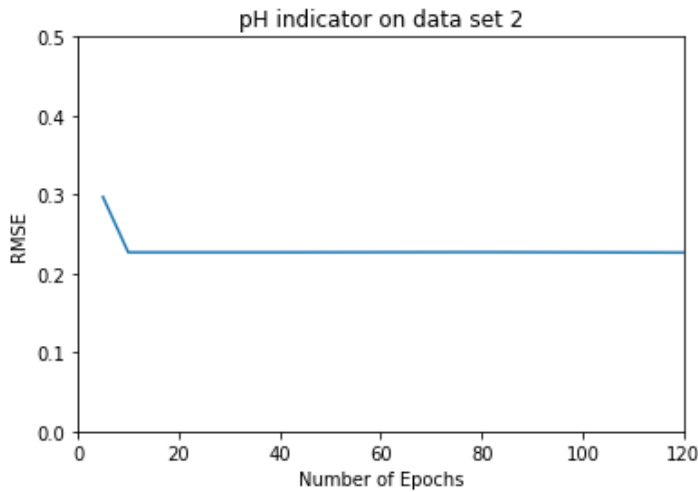Fig. 4. Thus, we have selected four hidden neurons for faster training.

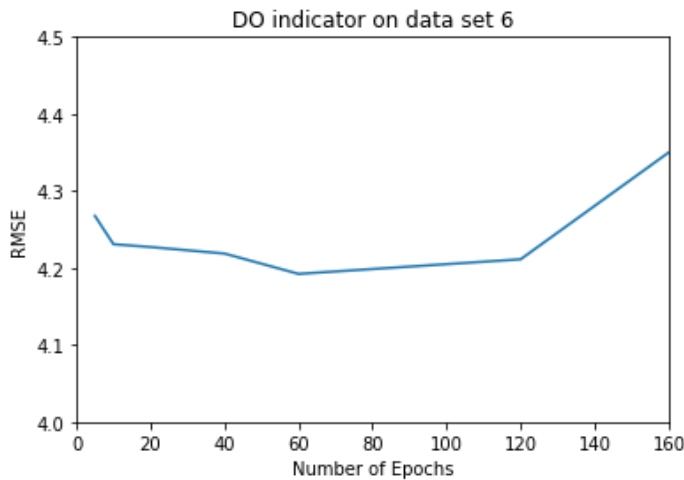Fig. 5. Relationship between the RMSE and the Number of Epochs (on Small Data Sets) using LSTM.

Fig. 6. Relationship between the RMSE and the Number of Epochs (on Larger Data Sets) using LSTM.

The second group of experiments is performed to choose the number of optimizing epochs. Results show that the RMSE is the best with 10 epochs for small data sets (data set 1 and data set 2) as presented in Fig. 5. However, the number of epochs should be increased for larger data sets (data sets 3,4,5, and 6), i.e., from 60 epochs to 80 epochs as an example in Fig. 6.

In the third group of experiments, we have revised the single LSTM model by adding more layers, which is called stacked LSTM (e.g., model in the right side of Fig. 2). Results show that the RMSE can be improved when using stacked LSTM as presented in Fig. 7. However, the RMSE is not significantly improved by adding more than 2 layers to the LSTM model, thus we have used the model with 2 layers as proposed in the right side of Fig. 2 for the rest experiments.
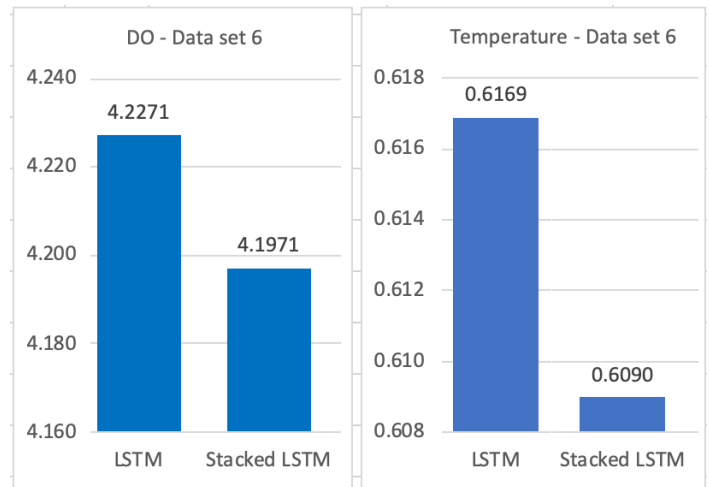

Fig. 7. RMSE Comparison between LSTM and Stacked-LSTM on Data Set 6. Other Data Sets are also Similar

The final results are presented in Table II. There are four considered indicators which include salinity, pH, temperature, and dissolved oxygen (DO). As shown from Table II, most of the cases, the LSTM has lower error (RMSE) than the baseline SVM, thus it can produce better forecasting results.

Fig. 8 and Fig. 9 visualize the true values and the forecast values of the salinity and the pH indicators on small data set while Fig. 10 and 11 visualize the forecast values of the temperature and the DO indicators on lager data set, respectively. The result shapes of other datasets are also similar when visualizing.

These results show that the forecast values of the LSTM model are nearly approached to the true values, thus the proposed model could be appropriate for sensor data with time series, and it could be applied to the real world. These results also consist with other works in this domain [5], [7], [9], [11].

V. CONCLUSION

In this work, we proposed an IoT system for monitoring water quality in aquaculture and fisheries, especially the model for forecasting the quality indicators such as salinity, pH, temperature, and DO. Experimental results on the considered data sets with four indicators show that the proposed approach can be applied for the real system. By monitoring these real-time indicators and getting early warning, the system can help the farmers to manage the quality of the water, thus collecting both quality and quantity in shrimp/fish raising.

In the future, we will collect the data provided in our system to retrain the model for implementing to the real world. In this work, we deployed a simple LSTM to perform the forecasting tasks. Some sophisticated models should be taken into account to improve the forecast performance in further researches. The LSTM model in this work uses one variable (the indicator) for forecasting, however, by utilizing other features/variables in the dataset to adapt the LSTM to Multi-Variate LSTM forecasting models could be an approach for improving the forecasting performance.

TABLE II. RMSE COMPARISON BETWEEN LSTM AND SVM. THE RESULTS FORMATTED IN BOLD AND RED ARE THE BEST RESULTS CORRESPONDING THE EACH INDICATOR.

No.	Data set	Indicator	#Records	RMSE	
				SVM	LSTM
1	NSW Estuary Temperature, pH and salinity data	Salinity	6,244	9.0773	**8.2611**
		pH	3,358	0.4709	**0.4396**
		Temperature	6,261	1.9212	**1.8062**
2	Water quality - Tomales Bay, CA	pH	246	0.1953	**0.1922**
		Salinity	229	3.0744	**2.8755**
		Temperature	229	1.6553	**1.6150**
		Percent DO (O2Sat)	246	11.7498	**11.7002**
		DO (O2Mg)	246	1.0571	**1.0482**
3	Melbourne - Sensor readings	Temperature	56,570	1.9700	**1.1922**
		Humidity	56,570	**7.0900**	7.5474
4	Fisheries & Oceans Canada 2018	Salinity	33,835	0.3786	**0.3640**
		Temperature	34,047	**0.3419**	0.3423
5	Fisheries & Oceans Canada 2017	Salinity	30,714	0.4350	**0.4274**
		Temperature	30,714	0.3112	**0.3104**
6	Temperature and DO from the EGIM, EMSO-Azores observatory, 2017-2018	DO	38,657	4.2445	**4.1971**
		Temperature	38,657	0.6095	**0.6079**

ACKNOWLEDGMENT

REFERENCES

[1] Shafique, Kinza & Khawaja, Bilal & Sabir, Farah & Qazi, Sameer & Mustaqim, Muhammed. (2020). Internet of Things (IoT) For Next-Generation Smart Systems: A Review of Current Challenges, Future Trends and Prospects for Emerging 5G-IoT Scenarios. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2970118.

[2] Lim, B., & Zohren, S. (2020). Time Series Forecasting With Deep Learning: A Survey. arXiv preprint arXiv:2004.13408.

[3] ZhanJie Wang and A. B. M. Mazharul Mujib. (2017). The Weather Forecast Using Data Mining Research Based on Cloud Computing. Phys.: Conf. Ser. 910 012020.

[4] Aditya Grover, Ashish Kapoor, and Eric Horvitz. (2015). A Deep Hybrid Model for Weather Forecasting. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15). Association for Computing Machinery, New York, NY, USA, 379–386. DOI:https://doi.org/10.1145/2783258.2783275

[5] Ben Abdel Ouahab Ikram, Boudhir Anouar Abdelhakim, Astito Abde-lali, Bassam Zafar, and Bouhorma Mohammed. (2019). Deep Learning architecture for temperature forecasting in an IoT LoRa based system. In Proceedings of the 2nd International Conference on Networking, Information Systems & Security (NISS19). Association for Computing Machinery, New York, NY, USA, Article 43, 1–6. DOI:https://doi.org/10.1145/3320326.3320375

[6] Harun, Zambri & Reda, Eslam & Hashim, Harris. (2018). Real time fish pond monitoring and automation using Arduino. IOP Conference Series: Materials Science and Engineering. 340. 012014. 10.1088/1757-899X/340/1/012014.

[7] Li D., Liu S. (2013). Remote Monitoring of Water Quality for Intensive Fish Culture. In: Mukhopadhyay S., Mason A. (eds) Smart Sensors for Real-Time Water Quality Monitoring. Smart Sensors, Measurement and Instrumentation, vol 4. Springer, Berlin, Heidelberg

[8] Dieu Tien Bui, Khabat Khosravi, John Tiefenbacher, Hoang Nguyen, Nerantzis Kazakis. (2020). Improving prediction of water quality indices using novel hybrid machine-learning algorithms, Science of The Total Environment, Volume 721, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.137612.

[9] Eiman Kanjo, Eman M.G. Younis, and Chee Siang Ang. (2019). Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. Information Fusion, Volume 49, Pages 46-56, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2018.09.001.

[10] Qiang Zhang, Fengchen Fu, and Ran Tian. (2020). A deep learning and image-based model for air quality estimation. Science of The Total Environment, Volume 724, 138178, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.138178.

[11] Lee D, Kang S, Shin J. (2017). Using Deep Learning Techniques to Forecast Environmental Consumption Level. Sustainability. 9(10):1894.

[12] Hino, M., Benami, E. & Brooks, N. (2018). Machine learning for environmental monitoring. Nat Sustain 1, 583–588. https://doi.org/10.1038/s41893-018-0142-9

[13] Preetham, K., Mallikarjun, B. C., Umesha, K., Mahesh, F. M., & Neethan, S. (2019). Aquaculture monitoring and control system: An IoT based approach. International Journal of Advance Research, Ideas and Innovations in Technology, 5(2).

[14] Darmalim, U., Darmalim, F., Darmalim, S., Hidayat, A. A., Budiarto, A., Maheswora, B., & Pardamean, B. (2020). IoT Solution for Intelligent Pond Monitoring. E&ES, 426(1), 012145.

[15] Harini Suresh. The Vanishing Gradient Problem. http://harinisuresh.com/2016/10/09/lstms/, accessed on August 20, 2020.

[16] Wang, Z., Li, P., Li, L., Huang, C., & Liu, M. (2015). Modeling and forecasting average temperature for weather derivative pricing. Advances in Meteorology, 2015.

[17] Licata, R.J., Tobiska, W.K., & Mehta, P.M. (2020). Benchmarking Forecasting Models for Space Weather Drivers. arXiv: Space Physics.

[18] Dong, Jianhua & Wang, Guoyin & Yan, Huyong & Xu, Ji & Zhang, Xuerui. (2015). A survey of smart water quality monitoring system. Environmental Science and Pollution Research. 22. 10.1007/s11356-014-4026-x.

[19] Kachroud, Moez & Trolard, Fabienne & Kefi, Mohamed & Jebari, Sihem & Bourrié, Guilhem. (2019). Water Quality Indices: Challenges and Application Limits in the Literature. Water. 11. 1-26. 10.3390/w11020361.

[20] Siddiqui, Nihal & Tauseef, S.M. & Dobhal, Rajendra. (2020). Advances in Water Pollution Monitoring and Control Select Proceedings from HSFEA 2018: Select Proceedings from HSFEA 2018. 10.1007/978-981-32-9956-6.

[21] Diederik P. Kingma and Jimmy Lei Ba. Adam : A method for stochastic optimization. (2014). arXiv:1412.6980v9
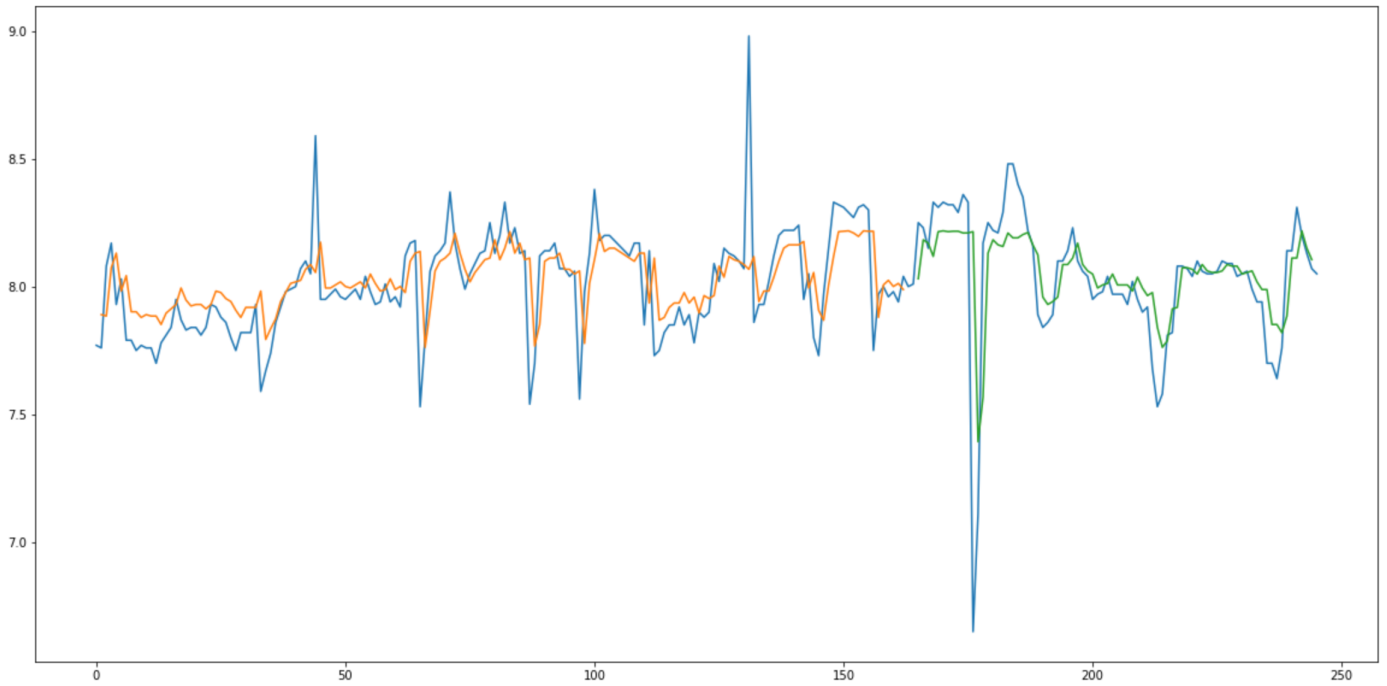
Fig. 8. Forecasting Visualization for the pH in Dataset 2. The Blue Line is the True Values, the Orange Line is the Predicted Values on Training Set (67%), and the Green Line is the Predicted Values on Test Set (33%)
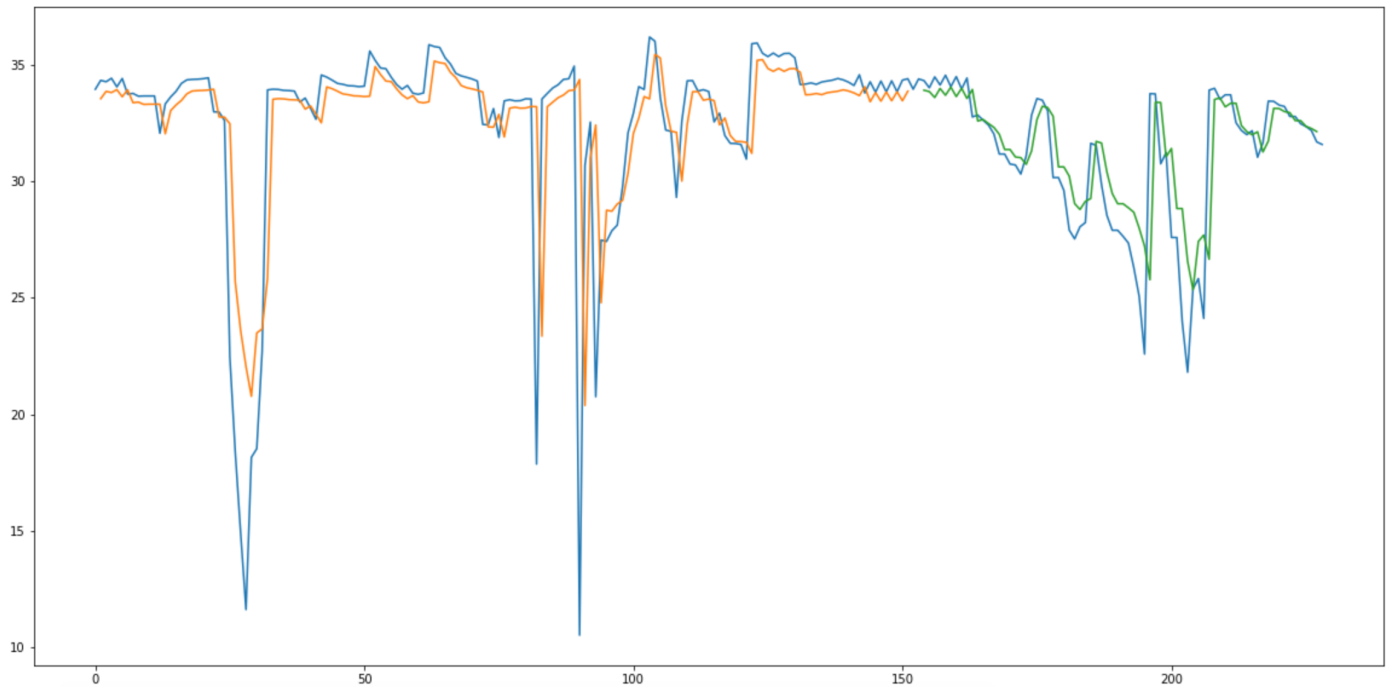

Fig. 9. Forecasting Visualization for the Salinity in Dataset 2. The Blue Line is the True Values, the Orange Line is the Predicted Values on Training Set (67%), and the Green Line is the Predicted Values on Test Set (33%)
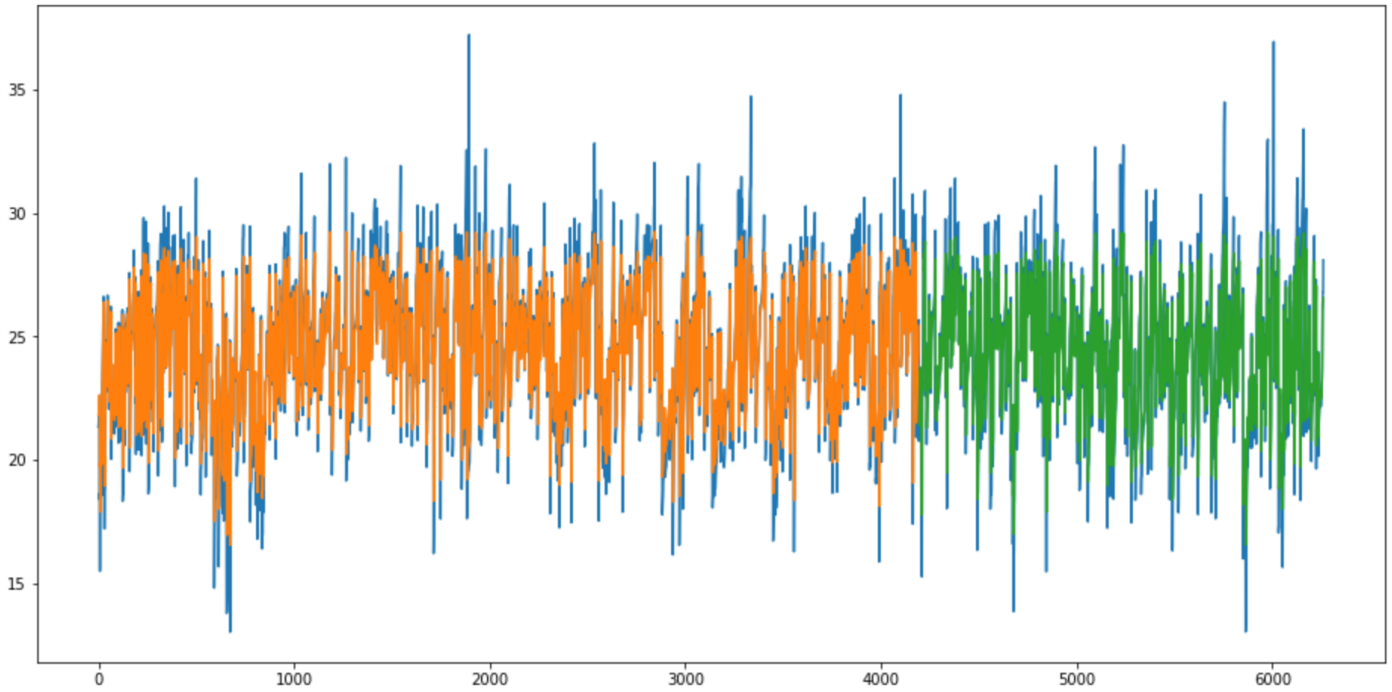
Fig. 10. Forecasting Visualization for the Temperature in Dataset 1. The Blue Line is the True Values, the Orange Line is the Predicted Values on Training Set (67%), and the Green Line is the Predicted Values on Test Set (33%)
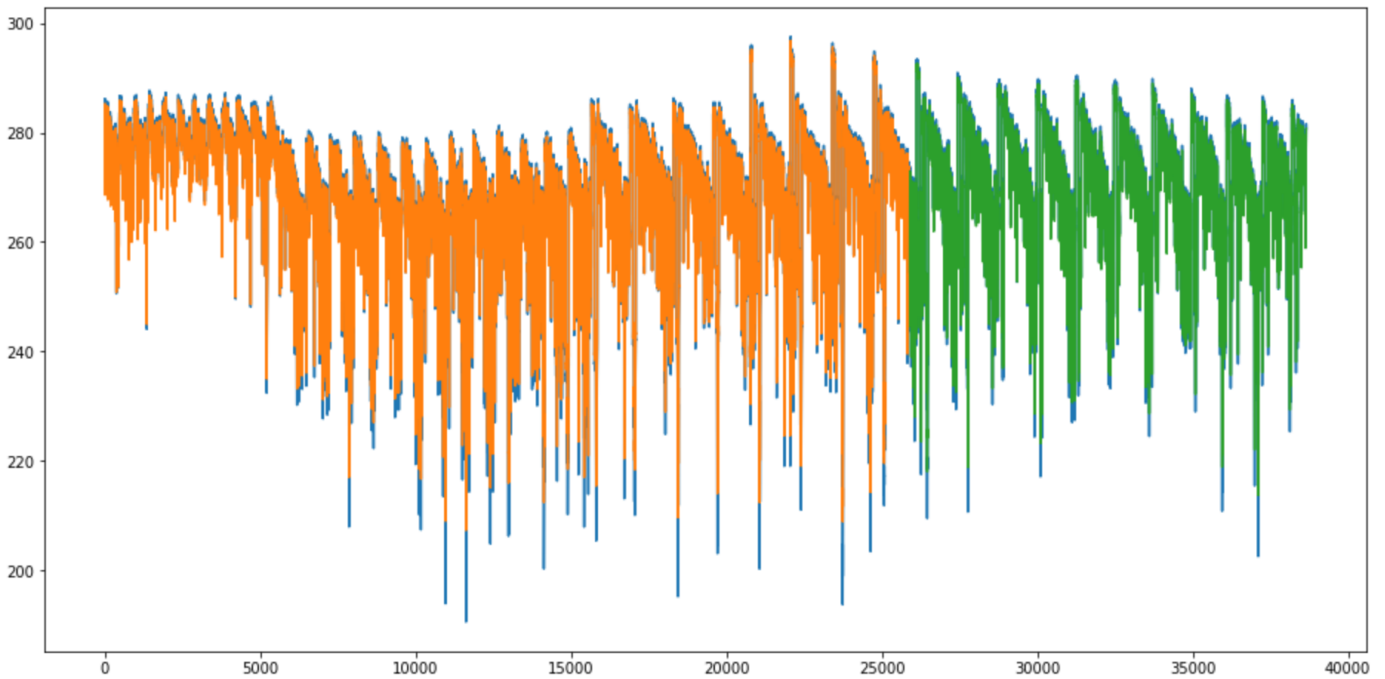

Fig. 11. Forecasting Visualization for the Dissolved Oxygen (DO) in Dataset 6. The Blue Line is the True Values, the Orange Line is the Predicted Values on Training Set (67%), and the Green Line is the Predicted Values on Test Set (33%)