LSTM Based Prediction of Total Dissolved Solids in Hydroponic System

Riky Dwi Puriyanto
Department of Electrical Engineering
Ahmad Dahlan University
Yogyakarta, Indonesia
rikydp@ee.uad.ac.id

Supriyanto
Department of Informatics Engineering
Ahmad Dahlan University
Yogyakarta, Indonesia
supriyanto@tif.uad.ac.id

Anton Yudhana
Department of Electrical Engineering
Ahmad Dahlan University
Yogyakarta, Indonesia
eyudhana@ee.uad.ac.id

Abstract—This paper discusses the implementation of long short term memory (LSTM) for forecasting the value of total dissolved solids (TDS). The TDS value in a hydroponic system represents the number of nutrients contained in water. The amount of water in the hydroponic system is important to note because optimal plant growth depends on the number of nutrients obtained by the plant. TDS data is sequential data, and one way to do forecasting is to use LSTM. This study uses a combination of epoch values of 100, 200, 300, 400 and 500. The RMSE values of on any combinations 57.41, 50.90, 57.81, 67.60 and 26.62. In general, the smallest RMSE value of each combination produces a graph that is close to except for a 70%: 30% combination. The greater use of training data compared to test data (90%: 10%) results in the smallest average RMSE value of 35.48.

Keywords—LSTM, forecasting, hydroponic, total dissolved solids

I. INTRODUCTION

Hydroponics is a technique of planting plants without using soil. Water is an essential component in hydroponic systems. The function of water is draining nutrients to the plant cells. We need to keep the plant nutrients need are met optimally. The benefits of the hydroponic system are the elimination of fungi, weeds, and diseases so that it can eliminate health risk [1].

The number of nutrients in water in a hydroponic system can be measured using total dissolved solids (TDS). TDS is the amount of substance dissolved in water. TDS refers to any mineral, salt, metal, cation or anion dissolved in water. TDS is the most important factor in water quality management. Many water quality plans have been implemented in recognition of TDS factor [2]. TDS value can represent water quality in terms of the amount or level of the nutrient in a hydroponic system. TDS shows the number of minerals, salts, metals, cations or anions dissolved in water. This includes anything in water other than pure water molecules (H2O) and solid waste.

TDS values are very important to be measured and predicted. It is because hydroponic plants will grow optimally with optimal nutrition. We must ensure sufficient nutrition for the plants. Measurement of the number of nutrients in a hydroponic system is generally done conventionally. Measurements are made directly to see directly the number of substances in the nutrient container. Estimates of the number of nutrients contained in hydroponic systems are also widely carried out without any measurement basis. The condition that is considered by conventional farmers is only the amount of water present in the nutrient container without paying attention to more details about the number of substances contained in the liquid nutrient.

Research on monitoring in agriculture has been done by [3][4][5]. They use the internet of thing (IoT) technology in monitoring, tracking and tracing, agriculture machinery, precision agriculture, and greenhouse production [4]. It also can be used to mineral content in the growing media. The resulting data is processed to optimize plant growth. Deep learning has been used by [5] to adjust some of the actuators based on measured input variables.

TDS values from water measurements in hydroponic installations are time-series data, which includes sequential data. Prediction of TDS needs to be done to determine water quality [6]. One method for sequential data forecasting is recurrent neural network (RNN). RNN is a family of neural networks that specifically handle sequential data [7]. In large amounts of data, the long short term memory (LSTM) which is a type of RNN can be used and produces better output. In this research, forecasting TDS values will be made with the LSTM method to get a small root mean square error (RMSE) value.

II. RESEARCH METHOD

A. System Hardware
Retrieval of TDS data on hardware is shown in the diagram block (Figure 1). TDS data is obtained from the TDS sensor. Analog data generated by the TDS sensor is processed in the microcontroller. Data is sent via the esp 8266 module and forwarded to the internet to be viewed via a computer or mobile phone. The microcontroller is also used to turn on the water pump so that water and nutrients can be channeled to hydroponic plants.

![Hardware block diagram](image)
**B. Long Short Term Memory (LSTM)**

LSTM is composed of cell memory, input gate, output gate, and forget gate. The cell memory serves to remember the value at each time interval, and the three gates function to regulate the outflow of information from the input, cell, and output. LSTM is very suitable for storing information based on time series data. The compiling diagram of the LSTM unit is shown in Figure 2.

The weights and biases to the input gates on the LSTM network control the extent to which new values flow into the cell. In addition, weights and biases to the forget gate and output gate control the extent to which the value remains in the cell and the extent to which the value in the cell is used to calculate the output activation of each LSTM block.

The next step is to decide which information will be stored in the cell. First, the sigmoid layer is called the input gate layer, where \( h_t - 1 \) and \( x_t \). \( h_t - 1 \) is the output value of the past or previous layer and \( x_t \) is the input value that will enter the layer. The forget gate diagram is shown in Figure 3(b). The forget gate function is written as a function \( f_t \) can be written with equation (1).

\[
f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)
\]

The next step is to decide which information will be stored in the cell. First, the sigmoid layer is called the input gate layer, deciding which value to update is written as \( i_t \), tanh layer creates a vector as the new candidate is written as \( C_t \), both will be combined to update the cell. The forget gate diagram is shown in Figure 3. The functions \( i_t \) and \( C_t \) can be written with equations (2).

\[
i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)
\]

\[
C_t = \tanh \left( W_c[h_{t-1}, x_t] + b_c \right)
\]

The next step is to update the old cell value \( C_{t-1} \) to the new cell value \( C_t \). The updating diagram \( C_t \) is shown in Figure 3. The function \( C_t \) can be written with the equation (3).

\[
C_t = f_t \cdot C_{t-1} + i_t \cdot C_t
\]

The final step is to decide which information will be released. The output will be based on cell but filtering needs to be done. First, run on the sigmoid layer which decides which part of the cell to output is written as \( o_t \). Next, enter the cell past the tanh and multiply it by \( o_t \) as \( h_t \). Output diagrams \( o_t \) and \( h_t \) are shown in Figure 3. Functions \( o_t \) and \( h_t \) can be written with equation (4).

\[
o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)
\]

\[
h_t = o_t \cdot \tanh(C_t)
\]

From the gate function, there is a notation \( Wf, Wc, Wi, Wo \) called weight and \( bf, bc, bi, bo \) called bias is the weight value obtained through the training process in training data or training data. \( \sigma \) is the sigmoid activation function and tanh.

For the regression purpose, the architecture of LSTM can be seen in Figure 4. Simulations will be carried out in MATLAB using the LSTM specification of one feature and 200 hidden layers. In the training option we use a number of settings including: the solver specified by adaptive moment estimation (ADAM); gradient threshold = 1; initial learning rate = 0.005; learn rate schedule = piecewise; learn rate drop period = 125; learn rate drop factor = 0.2; and Verbose = 0.

**III. RESULTS AND DISCUSSION**

The results of the measurement of TDS values can be seen in Figure 5. Measurement of TDS values is carried out for 720 hours. Data is taken every hour for 30 days. Hydroponic nutrition is carried out every noon.
The LSTM method is implemented to predict TDS values. There are several combinations of tests based on a comparison of the amount of training data and testing data. In each comparison, there is also a variation of the epoch value. The prediction result will be calculated the accuracy value using root mean square error (RMSE) according to equation (5). The value of \( y_j \) is predictive data while the value of \( \hat{y}_j \) is testing data.

\[
RMSE = \sqrt{\frac{\sum_{j=1}^{n} (y_j - \hat{y}_j)^2}{n}}
\]  

(5)

LSTM testing is done by varying the value of training data. There are 5 (five) combinations of training data and testing data, namely:

1. Testing with variations in data 50%:50%
2. Testing with variations in data 60%:40%
3. Testing with variations in data 70%:30%
4. Testing with variations in data 80%:20%
5. Testing with variations in data 90%:10%

Each combination will be tested with variations in the values of epoch 100, 200, 300, 400 and 500. The best test results for each combination can be seen in Figure 6.

Table I shows the RMSE values for each epoch test for each combination.

| Percentage Combination (%) | Epoch 100 | Epoch 200 | Epoch 300 | Epoch 400 | Epoch 500 | Avg |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|-----|
| 50%:50%                     | 82.12     | 57.41     | 89.85     | 92.22     | 99.34     | 84.19|
| 60%:40%                     | 55.74     | 73.55     | 64.97     | 63.59     | 50.90     | 61.75|
| 70%:30%                     | 74.32     | 62.55     | 74.77     | 77.57     | 57.81     | 69.40|
| 80%:20%                     | 93.10     | 67.60     | 69.69     | 83.53     | 79.81     | 78.74|
| 90%:10%                     | 46.32     | 32.18     | 28.33     | 43.94     | 26.62     | 35.48|

The smaller RMSE does not always produce a prediction graph that is closest to the actual result. In general, the smallest RMSE value produces the best prediction results close to the actual value, but in the combination of 60%: 40% the best graph is obtained at epoch 200 with an RMSE value of 73.55. Graphic representations of Table I can be seen in Figure 7. From Figure 7 can be seen that the increase in the number of epochs does not significantly reduce the value of RMSE. The higher the comparison of training data with test data, in general, is able to reduce the value of RMSE. The combination of 90%: 90% produces the smallest average RMSE value of 35.48. While the highest average RMSE occurs in a combination of 50%: 50% at 84.19.
IV. CONCLUSION

We have implemented LSTM as a forecasting algorithm in predicting TDS values using any combinations. From the test results obtained the best graphic results on each combination in sequence with the RMSE values of 57.41, 50.90, 57.81, 67.60 and 26.62. In general, the smallest RMSE value of each combination produces a graph that is close to except for a 70%: 30% combination. The greater use of training data compared to test data (90%: 10%) results in the smallest average RMSE value of 35.48.

ACKNOWLEDGMENT

This research was supported by LPPM Ahmad Dahlan University (No. PF-131/SP3/LPPM-UAD/IV/2019). We thank our colleagues from Ahmad Dahlan University who provide insight and expertise that greatly assisted the research.

REFERENCES

[1] K. Roberto, How To-Hydroponics, New York: The Futuregarden Press, 2003.
[2] M. H. Salmani and E. S. Jajaei, “Forecasting models for flow and total dissolved solids in Karoun river-Iran,” Journal of Hydrology, vol. 535, pp. 148-159, 2016.
[3] S. Heble, A. Kumar, K. V. V. D. Prasad, S. Samirana, P. Rajalakshmi and U. B. Desai, “A low power IoT network for smart agriculture,” 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 2018, pp. 609-614.
[4] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow and M. N. Hindia, "An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges," in IEEE Internet of Things Journal, vol. 5, no. 5, pp. 3758-3773, Oct. 2018.
[5] M. Mehra, S. Saxena, S. Sankaranarayanan, R. J. Tom and M. Veeramankandan, "IoT based hydroponics system using Deep Neural Networks," Computers and Electronics in Agriculture, vol. 155, pp. 473-486, 2018.
[6] M. H. Salmani dan E. S. Jajaei, “Forecasting models for flow and total dissolved solids in Karoun river-Iran,” Journal of Hydrology, vol. 535, pp. 148-159, 2016.
[7] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT press, 2016.