Analyzing Prediction Performance between Wavelet Neural Network and Product-Unit Neural Network

Nur Suhailayani Suhaimi, Zalinda Othman, Mohd Ridzwan Yaakub

1Universiti Teknologi MARA Cawangan Melaka, Jasin, 77300, Melaka, Malaysia
2,3 Science and Information Technology Faculty, Universiti Kebangsaan Malaysia, Bangi, 91125, Selangor, Malaysia
suhailayani@uitm.edu.my

Abstract. Analyzing the performance of a particular approach in a field very dependent on the problem it’s aimed to solve. Artificial Neural Network (ANN) widely used for prediction in many areas including medical, environment, business intelligence and education. The uniqueness of ANN is the dynamic of hidden layer can be improvised mapped with the data problem and the structure of architecture can be enhanced such as Wavelet Artificial Neural Network (WANN) and Product Unit Neural Network (PUNN). This research aimed to analyzed the performance between WANN and PUNN towards water quality data of Chini Lake. Real world data comes with dynamic stream data and dynamic parameters based on its area of data collection method. Handling dynamic data would be misleading if the approach used very dependent towards data classes. The measurement to analyze the data based on performance accuracy, data sensitivity, data precision and specification of both method with regards of the regular ANN. The findings demonstrate the ability to obtain satisfactory prediction accuracy for both WANN and PUNN compared to regular ANN. The model accuracy for this case study by using WANN and PUNN were 75.34 % and 66.86 %, respectively. Therefore, WANN would be a competitive tool for prediction with conventional ANN.

Keywords. Artificial Neural Network; Wavelet Artificial Neural Network; Product Unit Neural Network; Water Quality Prediction.

1. Introduction

Predicting the flows of stream data by using machine learning, deep learning associates with Artificial Intelligence field seems promising. The usage of big data streams becoming current interest in data analysis and prediction nowadays. This research of data streams quality prediction focuses on water quality domain. Sample data were taken from Chini Lake by UKM research centre at Pusat Pengajian Tasik Chini (PPTC). Chini Lake water quality changes over time due to several factors such as climate exchanged, logging activities and dam development in nearby area. As a developing country, a large amount of contaminated are discharged into the lake. Artificial Neural Network widely used in classification and prediction field. The ability of ANN act as powerful predictor due to high tolerance towards noisy data. A Linear Artificial Neural Network usually constructs multiple layers of summation.
units in all the non-input layers and caused large number of summation units when approximating complex functions. Water quality prediction has been evolving from data-driven method until involving Artificial Intelligence touch to keep pace with the evolving data emergence. Various of enhanced ANN approaches has been applied in various domain. Diamantopoulou, Antonopoulos and Papamichail (2005) successfully predict water quality of river in Northern Greece by using Neural Network. Wavelet Artificial Neural Network is an approach based on time-dependent spectral analysis that decomposes time-series in time-frequency space to provide timescale relationship. Meanwhile, another approach called Product Unit Neural Network (PUNN) is a data driven model to do prediction. PUNN provides smaller architectures with increased information capacity. Engelbrecht and Ismail (1999) stated that PUNN provides more local minima, deep ravines and large valleys in the search space that trap local optimization algorithm. Various research in hybridizing and combining methods being done to manipulate the findings become more accurate and precise. Despite of venturing complicated methods for a very narrowed domain, enhancing a single method could be promising. These two ANN’s enhanced approaches aimed to produce better accuracy and high sensitivity than linear ANN. Therefore, this paper focusing on comparative analysis of predicting water quality among linear ANN, wavelet NN and Product-Unit NN.

2. Related works

In recent years, neural networks have been widely studied because of their outstanding capability of fitting nonlinear models. Li, Sha and Wang (2017) perform hybrid Neural Networks to predict water quality knowing that ANN alone yielded less accurate result. They injected Genetic Algorithm to improve the performance on ANN prediction ability. In 2018, Sonmez, Kale, Ozdemir and Kadak hybridize ANN with Fuzzy Inference to reduce time consumption in ANN learning and increase performance of the predictor. The hybrid method also used in water quality prediction where dynamic data occurs and too many uncertainty. Instead of combining two and more methods in a model to increase performance of predictor, this paper enhanced the structure of ANN itself in order to reduce time consumption and increasing performance of single approach. Two approach chose as comparison measure which are Wavelet Artificial Neural Network and Product-Unit Neural Network.

2.1. Wavelet Neural Network

Wavelet Artificial Neural Networks(WANN) are a class of neural networks consisting of wavelets. Wavelet was first to develop by Grossmann and Morlet in 1984. As wavelet has emerged as a new powerful tool for representing nonlinearity, a class of networks combining wavelets and neural networks has recently been investigated (Huang and Cui, 2005). Wavelet Neural Networks are a combination of neural networks and wavelets and have been mostly used in the area of time-series prediction and control. Recently, Evolutionary Wavelet Neural Networks have been employed by Khan, Mendes, and Chalup (2018) to develop cancer prediction models. WANN applied a time-dependent spectral analysis that decomposes time-series in time-frequency to provide a timescale relationship. Several advantages of WANN being exposed such as flexibility of choice in selecting a mother wavelet according to the properties of the time series (Conraria and Soares, 2011). WANN can incorporate different characteristic of data such as images and sequence data where it becomes another significance of implementing this approach into wider perspectives. Fujieda, Takayama, and Hachisuka (2018) successfully implemented Wavelet Convolutional Neural Network in their research of image processing which utilized spectral information in most image processing tasks.
2.2. Product-Unit Neural Network

Product Unit Neural Networks (PUNN) were introduced by Durbin and Rumelhart in 1989, then further enhanced by Janson and Frenzel (1991) and Leerink, Giles, Horne, and Jabri (1995). In 2014, De-Niet, Meijers, and Schep produced a significant prediction of the ecological quality ration with PUNN then successfully revealed the underlying relations between characteristics and ecological quality. PUNN’s structure adopts three-layer-feed-forward network as conventional ANN; with normalize activation function. In 2006, Valero, Hervas, Gimeno, and Zurera used PUNN as a predictor to increase the accuracy of prediction. Meanwhile, Martinez, Estudillo, and Ruz (2007) used a multi logistic regression model based on the combination of linear and product-unit models, where the product-unit nonlinear functions are constructed with the product of the inputs raised to arbitrary powers. Both pieces of research agreed on the advantage of PUNN in increasing the performance of predictor and reduce errors in calculating the probability of classification. In recent years, neural networks have been widely studied because of their outstanding capability of fitting nonlinear models. Li, Sha, and Wang (2017) perform hybrid Neural Networks to predict water quality knowing that ANN alone yielded a less accurate result. They injected a Genetic Algorithm to improve the performance of ANN prediction ability. In 2018, Sonmez, Kale, Ozdemir, and Kadak hybridize ANN with Fuzzy Inference to reduce time consumption in ANN learning and increase the performance of the predictor. The hybrid method also used in water quality prediction where dynamic data occurs and too many uncertainties. Instead of combining two and more methods in a model to increase the performance of predictor, this paper enhanced the structure of ANN itself to reduce time consumption and increasing performance of a single approach. Two approaches chose as a comparison measure which is Wavelet Artificial Neural Network and Product-Unit Neural Network.

3. Methodology

Data collected from seven (7) stations in Tasik Chini from the year 2011 until 2015. Six (6) parameters used as permanent input to determine the water quality. The parameters are pH, temperature, Optical Dissolved Oxygen (ODO), Total Dissolved Solid (TDS), turbidity and conductivity. Expected output would be five classes of water quality from Class I until Class V. Hidden layer neuron is set randomly in initial epoch and evolved accordingly. The data first run by using WEKA 3.0 tools for Linear ANN. For wavelet ANN and PUNN, the data run by R with flexible structure editing for dynamic architecture.

3.1. Perform Linear ANN

**Step 1:** Pre-process data and remove noise.
**Step 2:** Divide data into 70% for training and the remaining 30% for testing
**Step 3:** Classify data based on quality categories using Linear Artificial Neural Network with sigmoid activation function. The architecture of WANN as represented in Fig.1
**Step 4:** Test the finalized model architecture towards test data.
**Step 5:** Visualize findings for analysis.
Where:
- \( n \) = total data
- \( X_i \) = input neuron
- \( W_i \) = weight
- \( b \) = bias; and
- \( f(x) \) = sigmoid activation function

\[
 \sum_{i=1}^{n} x_i w_i + b \quad ; \quad f(x) = \frac{e^x}{e^x + 1}
\]  

(1)

3.2. Perform Wavelet Artificial Neural Network

**Step 1**: Pre-process data and remove noise.

**Step 2**: Divide data into 70% for training and the remaining 30% for testing.

**Step 3**: Classify data based on quality categories using Wavelet Neural Network formula inside the summation neuron. The architecture of WANN as represented in Fig.2

**Step 4**: Test the finalized model architecture towards test data.

**Step 5**: Visualize findings for analysis.
Where;

\[ Y_j(x) = \text{mother wavelet} \]

\[ X = \text{input neuron} \]

\[ m = \text{number of network input} \]

\[ w = \text{weight} \]

\[ \lambda = \text{number of hidden unit} \]

\[ \xi = \text{small variant} \]

### 3.3. Perform Product-Unit Neural Network

**Step 1:** Pre-process data and remove noise.

**Step 2:** Divide data into 70% for training and the remaining 30% for testing.

**Step 3:** Classify data based on quality categories using Wavelet Neural Network formula inside the summation neuron. The architecture of WANN as represented in Fig.3.

**Step 4:** Test the finalized model architecture towards test data.

**Step 5:** Visualize findings for analysis.

\[
\Psi_j(x) = \prod_{i=1}^{m} \Psi_{zij}
\]  

(2)

\[
\Psi_{zij} = (1 - Z_{ij}^2) e^{-\frac{1}{2} Z_{ij}^2}
\]  

(3)

\[
\Psi_{\lambda}(x) = \prod_{i=1}^{m} \Psi \left[ \frac{X - W^{[1]}_{\epsilon \nu \lambda}}{W^{[1]}_{\epsilon \nu \lambda}} \right]
\]  

(4)

**Figure 3:** Product-Unit Neural Networks
\[
\prod_{i=1}^{N} x_i^{w_i}
\]

(4)

\[
\theta_y = (\beta^y, w_1, \ldots, w_m)\ i(x_j^w_j) \quad ; \ y = 1, 2, \ldots, y
\]

\[
\beta^y = (\beta_0^y, \beta_1^y, \ldots, \beta_m^y)
\]

Where:
\(\Pi\) = product summation
\(X\) = input neuron
\(m\) = number of network input
\(\beta\) = Beta activation function
\(B\) = bias
\(\theta\) = theta activation function

4. Result Analysis

The data run for each technique with limit 100 epoch or halted when the accuracy reach maximum (highest accuracy set at 100 percent). Table 1 shows the overall result yielded by all three algorithms.

| Dataset (Year) | Algorithm | TT (s) | RMSE | TP (%) | FP (%) | Accuracy (%) |
|---------------|-----------|--------|------|--------|--------|--------------|
| 2011          | ANN       | 3.2    | 175.4| 68.3   | 31.7   | 68.3         |
|               | WANN      | 5.6    | 93.1 | 89.4   | 10.6   | 89.4         |
|               | PUNN      | 1.8    | 107.2| 71.8   | 28.2   | 71.8         |
| 2012          | ANN       | 0.2    | 67.3 | 49.8   | 50.2   | 49.8         |
|               | WANN      | 0.6    | 52.1 | 59.2   | 40.8   | 59.2         |
|               | PUNN      | 0.1    | 69.5 | 66.3   | 33.7   | 66.3         |
| 2013          | ANN       | 4.9    | 137.7| 63.7   | 36.3   | 63.7         |
|               | WANN      | 6.2    | 175.7| 65.9   | 34.1   | 65.9         |
|               | PUNN      | 4.8    | 126.5| 61.4   | 38.6   | 61.4         |
| 2014          | ANN       | 4.3    | 92.1 | 78.2   | 21.8   | 78.2         |
|               | WANN      | 4.9    | 77.2 | 84.9   | 15.1   | 84.9         |
|               | PUNN      | 3.7    | 103.6| 71.4   | 28.6   | 71.4         |
| 2015          | ANN       | 3.5    | 141.8| 65.8   | 34.2   | 65.8         |
|               | WANN      | 6.4    | 79.4 | 77.3   | 22.7   | 77.3         |
|               | PUNN      | 2.8    | 192.6| 63.4   | 36.6   | 63.4         |
Based on Table 1, dataset from 5 years being used to classify and predict water quality. TT represent time taken in seconds unit, RMSE represents root-means-squared-error, TP represent True Positive and FP represent False Positive. True positive is a value of water quality which correctly classified and predicted to its class while False positive is a value of water quality which incorrectly classified. Each algorithm run with maximum 10 epoch (iteration) for 5 years data from 2011 until 2015.

Fig. 4 indicates the performance result in term of time taken to predict the water quality class. Time takes was measured in seconds to determine the algorithm processing time in calculating the predicted class label according to each formula set for ANN, WANN and PUNN.

![Prediction Performance](image1)

**Figure 4:** Prediction performance (time taken)

Based on results in Fig.4, highest performance of time taken falls in year 2012 data where the data is mostly missing, thus the volume of data decreased and eventually increased the performance value. However, based on algorithm comparison, in average, PUNN produced better performance with shortest time taken in every run, followed by ANN and lastly by WANN.

![Root Mean Squared Error (RMSE)](image2)

**Figure 5:** Root-Mean-Squared-Error (RMSE)

Results shown in Fig. 5 focuses on RMSE yielded during training. Algorithm wise, ANN shows to have higher RMSE due to value bias added into the calculation. Drifting biases occurs where the classified water quality magnitudes increase gradually as the data become more dynamic. However, during 2013 data run, WANN shown to have higher RMSE due to the higher volume of data measured. Therefore, PUNN selected as the best performance in term of minimum errors compared to Linear ANN and Wavelet ANN except for data in 2013 where Wavelet ANN produced highest RMSE.
Figure 6: Prediction Accuracy

As shown in Fig. 6, prediction accuracy among ANN, WANN, and PUNN being compared. Basically, all three algorithms produced almost similar level of accuracy. To be precise, WANN may selected as better predictor since its proven to yield highest accuracy (75.34 in average) among the three. However, based on missing data in 2012, the accuracy of WANN affected and fell below PUNN. In average, PUNN (66.86 %) will be selected as second-best accurate predictor and leave ANN (65.15) as lowest predictor ability by only 1.71 % difference measure.

5. Conclusion

Enhancing an algorithm or modifying a method is considered as repairing mechanism towards its drawback. Most of the Intelligent tools designated for a very narrowed domain. To apply an algorithm to a various problem might not yield a very good accuracy as it’s trying to adjust the algorithm as possible it can be. This paper concludes the algorithm used to predict water class perform ANN with difference summation function (linear ANN, Wavelet ANN, and Product-Unit NN). Difference algorithm yielded different results in several measures. In term of performance analysis which used time taken as unit measured, Product-Unit NN selected as the most suitable algorithm due to its formula directly calculate the product summation and reduce calculation complexity. Due to a massive formula in determining mother wavelet function in Wavelet ANN summation, the time taken to classify the water quality become longer. However, Wavelet ANN produced higher accuracy prediction compared to linear ANN and Product-Unit NN. Based on root-mean-squared errors analysis, PUNN produced lesser errors compared to the other two algorithms. Thus, PUNN may yield a better performance algorithm with lesser time and lesser errors. Wavelet algorithm selected as the best predictor in classifying the water quality class level due to its ability in obtaining higher accuracy.

Acknowledgements

We are very grateful to Universiti Teknologi MARA for supporting this research paper and sincere thanks to Pusat Penyelidikan Tasik Chini (PPTC), Universiti Kebangsaan Malaysia for the data and expert’s knowledge. Huge appreciation to Universiti Kebangsaan Malaysia in providing space and computer lab as a venue to conduct research and paper writing.
References

[1] Diamantopoulou, M.J., Antonopoulos V.Z. & Papamichail D.M. (2005). The Use of a Neural Network Technique for the prediction of Water Quality Parameters. SpringerLink: Operational Research, Vol.5, Issue 1, 115-125.

[2] Engelbrecht A.P. & Ismail A. (1999). Training Product Unit Neural Networks. CiteseerX, doi=10.1.1.23.3872.

[3] Sonmez A.Y., Kale S., Ozdemir R.C. & Kadak A.E. (2018). An Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict of Cadmium (Cd) Concentrations in the Filyos River. Turkish Journal of Fisheries and Aquatic Sciences, doi : 10.4194/1303-2712-v18_12_01.

[4] Huang & Cui (2005)

[5] Khan MM, Mendes A, Chalup SK (2018) Evolutionary Wavelet Neural Network Ensembles For Breast Cancer And Parkinson’s Disease Prediction. PLOS ONE 13(2): e0192192. https://doi.org/10.1371/journal.pone.0192192.

[6] Conraria L.A. & Soares M.J. (2011). Business Cycle Synchronization and the Euro: A Wavelet Analysis. Journal of Macroeconomics, Vol. 33, Issues 3, 477-489.

[7] Fujieda S., Takayama K. & Hachisuka T. (2018). Wavelet Convolutional Neural Network. Cornell University, Computer Vision and Pattern Recognition.

[8] Janson D.J. & Frenzel J.F. (1991). Training Product-Unit Neural Network with Genetic Algorithms. IEEE, doi: 10.1109/64.236478.

[9] Leerink L., Giles C.L., Horne W.G. & Jabri M.A. (1995). Learning with Product Units. Advances in Neural Information Processing System, Vol. 7, 537-544.

[10] De-Niet A., Meijers E. & Schep S.A. (2014). Accurate Prediction of Ecological Quality Ratio with Product Unit Neural Networks. International Conference on Hydroinformatics.

[11] Valero A., Hervas C., Gimeno G. & Zurera G. (2006). Product Unit Neural Network models for predicting the Growth Limits of Listeria Monocytogenes. US National Library of Medicine, National Institute of Health.

[12] Martinez C.H., Estudillo F.J.M. & Ruz M.C. (2007). Evolutionary Combining of Basis Function Neural Network for Classification